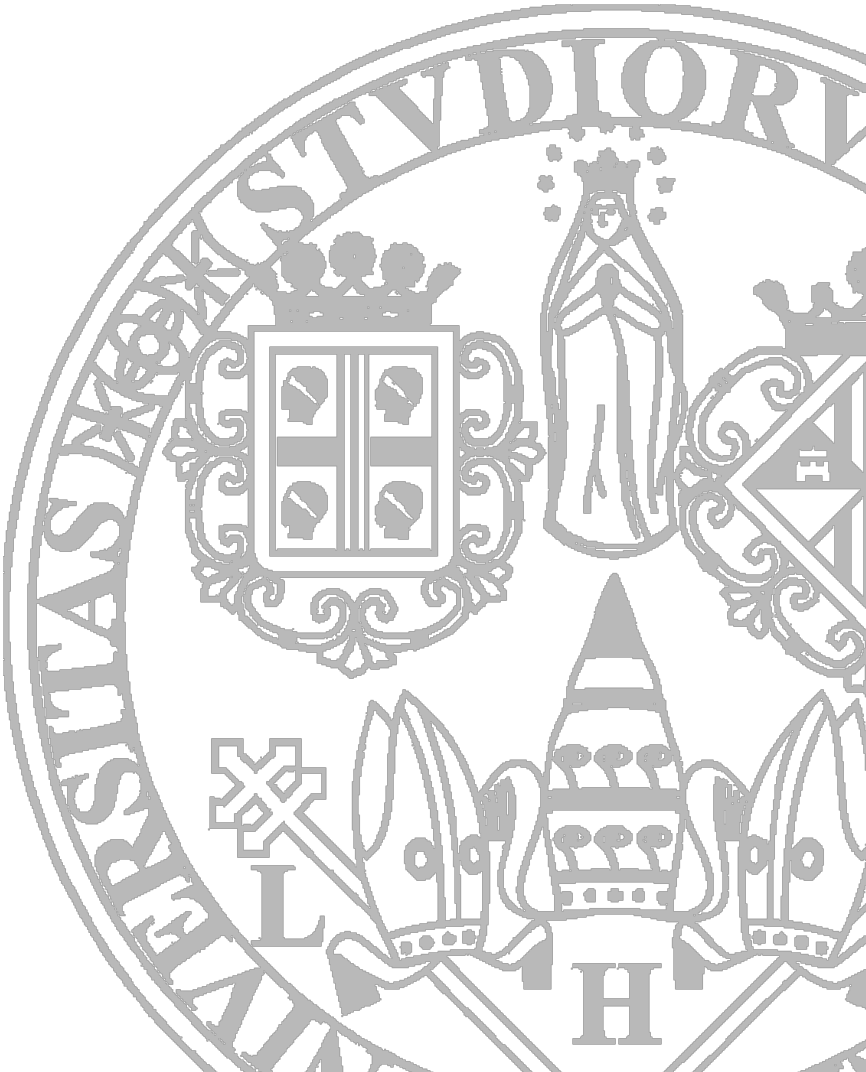
Appunti di Matematica Discreta

Università di Cagliari
Corso di Laurea in Informatica
Anno Accademico 2021/22
Secondo semestre



Aggiornato il
3 marzo 2022

Abstract

Nell'espressione "matematica discreta", l'aggettivo "discreto" va inteso in contrapposizione all'aggettivo "continuo". Senza pretesa di dare definizioni rigorose di queste due nozioni (che sono di pertinenza della branca della matematica detta topologia) possiamo dire che laddove la seconda descrive oggetti come una retta o una curva nel piano, la prima descrive insiemi di punti isolati.
Nel caso degli oggetti continui, i matematici hanno sviluppato metodi come il calcolo infinitesimale che, sfruttando il fatto che in un oggetto continuo ha senso considerare punti "sempre più vicini" a un punto dato, possono definire nozioni come quella di limite o derivata e studiare problemi quali la determinazione della tangente a una curva o la lunghezza della curva stessa. Tali metodi costituiscono quindi una sorta di matematica del continuo, e possiamo quindi pensare come una sorta di matematica discreta l'insieme delle branche della matematica che non usano tali metodi. In questo senso, è matematica discreta, ad esempio, lo studio dei numeri interi e delle loro proprietà e più in generale l'algebra nelle sue varie accezioni.
In questo corso studieremo alcune proprietà fondamentali dell'algebra dei numeri interi e a dispetto della sua apparente semplicità ne vedremo importanti applicazioni, ad esempio, in crittografia. Inoltre ci concentreremo sulla cosiddetta algebra vettoriale (le cui applicazioni hanno un ruolo centrale, per esempio, nella grafica e nella robotica) e sull'algebra di Boole, una sorta di formulazione algebrica della logica di cruciale importanza nell'informatica.
Alla base di tutte queste nozioni (ma in generale di ogni teoria matematica) sta il concetto di insieme. Per questo motivo nei primi due capitoli di questi appunti ci concentreremo sulla teoria degli insiemi.

Indice

Capitolo 1

Teoria degli insiemi e algebra Booleana

1.1 Insiemi, elementi, uguaglianza tra insiemi

Definizione 1.1. Un **insieme** è una collezione di oggetti (concreti o astratti). Questi sono detti gli **elementi** dell'insieme.

Osservazione 1.2. Quella che abbiamo appena dato è una definizione molto intuitiva, sufficiente per questo corso ma chiaramente insoddisfacente se si vuole rendere rigorosa la teoria. A questo scopo è stata elaborata la cosiddetta *teoria assiomatica degli insiemi* che evita di dare le definizioni di insieme e elemento assumendole come nozioni primitive. La teoria si basa su un insieme di affermazioni di partenza (gli assiomi, appunto) che devono essere soddisfatte dagli oggetti e che li definiscono implicitamente (analogamente a quanto succede in geometria euclidea con le nozioni di punto e retta). Non tratteremo qui la teoria assiomatica degli insiemi, che è oltre gli scopi di questo corso.

Per indicare che x è un elemento di un insieme A, scriveremo

$$x \in A$$

che si legge "x è un elemento di A", "x appartiene a A" o ancora "x in A".

Dare un insieme significa quindi specificare i suoi elementi. Questo può essere fatto elencandoli esplicitamente, convenzionalmente all'interno di parentesi graffe, ad esempio

$$A = \{0, 1, 2, 3\}.$$

Alternativamente gli elementi dell'insieme possono essere caratterizzati mediante una proprietà. Per esempio, lo stesso insieme A appena definito può essere descritto come l'insieme dei numeri naturali[1] minori di 4. In simboli

[1] Ricordiamo che si tratta dei numeri $0, 1, 2, 3, 4, 5...$ che si usano per contare: l'insieme di tali numeri

$$A = \{x \in \mathbb{N} \mid x < 4\}.$$

Useremo spesso la notazione

$$\{x \mid P(x)\}$$

per denotare l'insieme degli x dotati di una certa proprietà P.
Questo è chiaramente l'unico modo in cui è possibile definire un insieme infinito, nel qual caso non possiamo scrivere tutti i suoi elementi.

Osservazione 1.3. Accettare che per ogni proprietà P che possiamo enunciare esista l'insieme degli x con la proprietà P può però portare a contraddizioni. Un celebre esempio, dovuto a Russell, fu tra i motivi per cui emerse la necessità di una più rigorosa teoria assiomatica degli insiemi. Nello specifico, si scelga come proprietà P quella di "non appartenere a se stessi" e si definisca l'insieme degli x con tale proprietà, ovvero

$$A = \{x \mid x \notin x\}.$$

Otterremo una contraddizione chiedendoci se A appartiene a se stesso o no. Se infatti $A \in A$, allora non soddisfa la proprietà scelta e quindi $A \notin A$. Se invece $A \notin A$, allora A soddisfa la proprietà di appartenenza ad A e quindi $A \in A$. In conclusione qualunque ipotesi facciamo, arriviamo a una contraddizione.
Nella teoria assiomatica degli insiemi, si mostra che la contraddizione appena trovata può essere evitata se, ogni volta che si vuole definire un insieme mediante una certa proprietà, si precisa che gli elementi x che soddisfano tale proprietà devono appartenere a un insieme già esistente. Quindi solo una formula del tipo $\{x \in A \mid P(x)\}$, dove P è una qualunque proprietà e A è un insieme dato, definisce correttamente un insieme.
Tuttavia, continueremo a scrivere $\{x \mid P(x)\}$ sottointendendo che gli x considerati appartengano ad un insieme che contiene tutti gli oggetti e gli enti di cui abbiamo bisogno in questo corso. Chiameremo *insieme universo* questo ipotetico insieme e lo denoteremo con U.
Si badi tuttavia che l'insieme universo non è l'insieme di tutti gli insiemi possibili. Si può vedere che anche tale nozione porterebbe a una contraddizione.

Dal momento che un insieme è determinato dai suoi elementi, verificare che due insiemi dati sono uguali significa verificare che questi hanno esattamente gli stessi elementi. Per esempio, se come sopra consideriamo gli insiemi

$$A = \{0, 1, 2, 3\} \text{ e } B = \{x \in \mathbb{N} \mid x < 4\},$$

si denota $\mathbb{N}$. Si indicano invece con il simbolo $\mathbb{Z}$ i *numeri interi*, ovvero $0, \pm 1, \pm 2...$, con $\mathbb{Q}$ i *numeri razionali*, ovvero le frazioni (esprimibili anche come numeri decimali con numero finito di cifre dopo a virgola o periodici) e con $\mathbb{R}$ i *numeri reali*, ovvero tutti i numeri decimali (anche quelli con una sequenza infinita non periodica di cifre dopo la virgola, detti irrazionali, quali $\sqrt{2}$ o π).

per mostrare che $A = B$ dobbiamo verificare innanzitutto che ogni elemento di A appartiene anche a B. Per cui 0 è un elemento di A, e poiché $0 < 4$ si ha che 0 soddisfa la proprietà che definisce B, quindi $0 \in B$. Allo stesso modo si procede con $1, 2, 3$ e si dimostra che ogni elemento di A appartiene anche in B.
Questo non è però sufficiente: la stessa cosa sarebbe vera anche se avessimo definito $B = \{x \in \mathbb{N} \mid x < 5\}$. Tuttavia in tal caso A non sarebbe uguale a B in quanto $B = \{0, 1, 2, 3, 4\}$.
Questo perché per dimostrare che $A = B$ non basta dimostrare che ogni elemento di A sta anche in B, ma anche che ogni elemento di B appartiene ad A.
La verifica di uno solo di questi due fatti, per esempio che ogni elemento di A appartiene anche a B, ci dice solo che A *è contenuto in* B.
Un insieme A è un **sottoinsieme** di un insieme B se ogni elemento di A appartiene anche a B. In simboli scriveremo $A \subseteq B$.

1.2 Intersezione, unione e complementare

Siano A e B due insiemi. Possiamo allora considerare i seguenti insiemi.

(1) L'insieme, denotato con $A \cap B$ e detto **intersezione di** A **e** B, costituito dagli elementi comuni a A e B. In simboli

$$A \cap B = \{x \mid x \in A \text{ e } x \in B\}$$

dove il significato della congiunzione "e" è quello usuale. Cioè si intende che $x \in A$ e $x \in B$ è verificata solo quando sono verificate sia $x \in A$ che $x \in B$, ovvero solo per gli elementi x che stanno sia in A che in B.

Esempio 1.4. Se $A = \{0, 1, 2, 3\}$ e $B = \{2, 3, 4, 5\}$, si ha $A \cap B = \{2, 3\}$. Infatti $2 \in A$ e $2 \in B$ sono entrambe vere; e lo stesso per $3 \in A$, $3 \in B$. D'altra parte 0 non appartiene a $A \cap B$ perché $0 \in A$ è vera ma $0 \in B$ è falsa. Analogamente, 5 non appartiene a $A \cap B$ perché $5 \in B$ è vera ma $5 \in A$ è falsa. Lo stesso vale per 4.

(2) L'insieme, denotato da $A \cup B$ e detto **unione di** A **e** B, che si ottiene mettendo insieme gli elementi di A e gli elementi di B. In simboli

$$A \cup B = \{x \mid x \in A \text{ o } x \in B\}$$

dove la congiunzione "o" è usata nel senso che $x \in A$ o $x \in B$ è verificata quando è vera almeno una delle due tra $x \in A$ e $x \in B$. Eventualmente anche entrambe possono essere verificate: non va inteso come un "o" esclusivo, significato che tale congiunzione ha spesso nel linguaggio comune, nel quale deve valere l'una o l'altra ma non entrambe.

Esempio 1.5. Se, come sopra, $A = \{0,1,2,3\}$ e $B = \{2,3,4,5\}$, si ha $A \cup B = \{0,1,2,3,4,5\}$. Infatti, per $x = 0$ e $x = 1$ si ha che $x \in A$ o $x \in B$ è verificata in quanto è vera $x \in A$ (come abbiamo detto, basta che una sola delle due sia verificata). Per $x = 4$ e $x = 5$ invece $x \in A$ o $x \in B$ è verificata in quanto è vera $x \in B$ (pur non essendo vera $x \in A$). Infine, per $x = 2$ e $x = 3$ sono vere sia $x \in A$ che $x \in B$ e quindi, in virtù del fatto che la congiunzione "o" non va intesa in senso esclusivo, $x \in A$ o $x \in B$ è verificata.

Il significato non esclusivo della congiunzione "o" che abbiamo definito e usato sopra è esattamente quello che aveva in latino la congiunzione "vel" (contrapposta a "aut" che invece significava un "o" esclusivo: o l'uno o l'altro, ma non entrambi). Per questo motivo, tale congiunzione viene anche indicata con il simbolo $\vee$. In questa notazione, la congiunzione "e" viene indicata con il simbolo $\wedge$. Adotteremo questa notazione perché ha il vantaggio di usare simboli universali.

Esempio 1.6. Consideriamo ora gli insiemi $A = \{0,1,2,3\}$ e $B = \{4,5\}$. In tal caso, l'affermazione $x \in A \wedge x \in B$ non è mai verificata, in quanto non esiste nessun elemento x per cui siano vere sia $x \in A$ che $x \in B$. Quindi l'insieme $A \cap B = \{x \mid x \in A \wedge x \in B\}$ non ha elementi. Tale insieme è detto l'*insieme vuoto*.

Definizione 1.7. L'insieme che non contiene alcun elemento è detto **insieme vuoto** e si denota con $\varnothing$.

Quando $A \cap B = \varnothing$, si dice anche che gli insiemi A e B sono **disgiunti**.

Un'altra operazione su insiemi, in questo caso a partire da un solo insieme A, è quella del complementare. Il complementare di un insieme A è l'insieme, denotato da $\overline{A}$, degli elementi che *non* appartengono ad A. In simboli

$$\overline{A} = \{x \mid x \notin A\}. \tag{1.1}$$

In effetti, come abbiamo nell'Osservazione 1.3, la scrittura corretta per (1.1) è $\overline{A} = \{x \in U \mid x \notin A\}$. Generalizzando, si può definire anche il complementare di un insieme A dato *in un altro insieme* X, che non sia necessariamente l'insieme universo U.

Definizione 1.8. Il **complementare di un insieme** A **in un insieme** X è l'insieme, denotato da $\overline{A}$, degli elementi di X che *non* appartengono ad A. In simboli

$$\overline{A} = \{x \in X \mid x \notin A\}. \tag{1.2}$$

Esempio 1.9. Se A è l'insieme dei numeri naturali pari, ovvero $A = \{x \in \mathbb{N} \mid x \text{ è pari }\}$, il complementare di A in $\mathbb{N}$ è $\{x \in \mathbb{N} \mid x \text{ non è pari }\}$, ovvero l'insieme dei numeri dispari.

Così come le operazioni di intersezione e unione sono definite da simboli logici (rispettivamente $\wedge$ e $\vee$) anche il complementare è definito applicando un'operazione logica, ovvero la **negazione**, all'affermazione $x \in A$.
Il simbolo che si usa per indicare la negazione è $\neg$, oppure $\sim$.

L'intersezione e l'unione appena introdotte costituiscono a tutti gli effetti due operazioni che dati due insiemi ce ne danno un terzo, analogamente alla somma e alla moltiplicazione tra numeri, che ci danno un numero a partire da due numeri dati. Seguendo questa analogia, il complementare può essere pensato come un'operazione che associa un nuovo insieme ad uno dato, come succede ad esempio con l'opposto di un numero che manda a in $-a$ o l'inverso che manda a in a^{-1}.
Così come per fare calcoli e manipolare espressioni numeriche è necessario conoscere le proprietà di tali operazioni (es. la proprietà commutativa, per cui $a+b=b+a$, etc.) anche per l'intersezione, l'unione e il complementare è utile sapere quali proprietà son valide.
Ad esempio, possiamo chiederci

Domanda: *L'intersezione gode della proprietà distributiva rispetto all'unione?*

Ovvero ci chiediamo se, dati tre insiemi A, B, C, valga sempre l'uguaglianza

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)\,. \tag{1.3}$$

In base alla definizione di uguaglianza tra insiemi dobbiamo verificare se, dato $x \in A \cap (B \cup C)$, si ha $x \in (A \cap B) \cup (A \cap C)$ e viceversa.
Ora, da una parte $x \in A \cap (B \cup C)$ significa per definizione di intersezione "$x \in A \wedge x \in B \cup C$", che a sua volta per definizione di unione si legge

$$\text{“}x \in A \wedge (x \in B \vee x \in C)\text{”} \tag{1.4}$$

D'altra parte, $x \in (A \cap B) \cup (A \cap C)$ significa $x \in (A \cap B) \vee x \in (A \cap C)$" cioè

$$\text{“}(x \in A \wedge x \in B) \vee (x \in A \wedge x \in C)\text{”} \tag{1.5}$$

Quindi, per dimostrare la (1.3) bisogna dimostrare che se l'affermazione (1.4) è vera, allora lo è anche la (1.5), e viceversa.
È chiaro che questa verifica richieda un approfondimento delle proprietà delle congiunzioni $\wedge$ e $\vee$ che ci consenta di manipolare espressioni logiche come la (1.4) e la (1.5).
Inoltre possiamo interpretare le congiunzioni $\wedge$ e $\vee$ come operazioni che a due affermazioni o proposizioni P e Q (arbitrarie in generale ma riguardanti l'appartenenza nello specifico del caso precedente) associano una terza proposizione "$P \wedge Q$" oppure "$P \vee Q$".
A questo punto, la verifica delle proprietà di queste operazioni può essere pensata co-

me una sorta di "algebra della logica" o di "algebra delle proposizioni". Questa viene chiamata *algebra Booleana* e sarà l'oggetto della prossima sezione.

1.3 Algebra Booleana

L'algebra usuale studia espressioni nelle quali compaiono delle incognite (es. x, y, z etc.) che rappresentano numeri reali. Otteniamo quindi formule e identità che coinvolgono queste incognite e che sono vere qualunque siano i numeri reali che sostituiamo ad esse. Ad esempio, l'equazione

$$(x+y)^2 = x^2 + 2xy + y^2$$

ci sta dicendo che tale uguaglianza è verificata qualunque siano i valori numerici che sostituiamo al posto delle incognite x e y. In altre parole, scelti un valore numerico di x e un valore numerico di y e sostituendoli nel primo e nel secondo membro della formula otterremo lo stesso valore.

L'algebra Booleana si comporta in maniera analoga. Ossia le incognite P, Q, R etc. rappresentano proposizioni che possono essere o vere o false, cioè possono assumere solo due valori "V" o "F". Per ottenere un'analogia ancora più stretta con l'algebra usuale si rappresenta il valore di verità "V" (rispettivamente"F") con il valore numerico 1 (rispettivamente 0).
Le formule dell'algebra Booleana sono quindi formule che coinvolgono le incognite P, Q, R e le operazioni $\wedge$ e $\vee$. Un esempio è dato dall'espressione

$$P \wedge (Q \vee R) = (P \wedge Q) \vee (P \wedge R)\ . \tag{1.6}$$

La validità di una tale formula significa, esattamente come nell'algebra usuale, che qualunque valore numerico sostituiamo alle incognite (che può essere solo 0 o 1 a seconda che l'incognita rappresenti una proposizione falsa o vera rispettivamente) otteniamo lo stesso valore numerico (sempre 0 o 1) al primo e al secondo membro.

A questo scopo, iniziamo con il racchiudere in tabelle le proprietà base delle operazioni $\wedge$ e $\vee$ che abbiamo implicitamente usato nelle definizioni di intersezione e unione.
Abbiamo visto che un'affermazione come $P \wedge Q$ è vera solo se sono vere tutte e due le proposizioni P e Q, ed è falsa negli altri casi. Possiamo rappresentare questo in una tabella nella quale in corrispondenza dei vari possibili valori (0 o 1) di P e Q scriviamo il corrispondente valore di verità di $P \wedge Q$.

P	Q	P $\wedge$ Q
0	0	0
0	1	0
1	0	0
1	1	1

La tabella dice esattamente quanto abbiamo espresso sopra: $P \wedge Q$ è vera (cioè assume valore 1) solamente quando sia P che Q hanno valore 1, cioè solo quando sono entrambe vere. In tutti gli altri casi l'affermazione P $\wedge$ Q è falsa (cioè vale 0).
Una tabella come quella che abbiamo appena scritto prende il nome di **tavola di verità**.

Passiamo ora alla tavola di verità della congiunzione $\vee$. Abbiamo detto che affinchè un'affermazione del tipo $P \vee Q$ sia vera è sufficiente che sia vera almeno una tra P e Q. Quindi la tavola di verità è

P	Q	P $\vee$ Q
0	0	0
0	1	1
1	0	1
1	1	1

La tabella corrisponde esattamente quanto abbiamo visto sopra: $P \vee Q$ è vera (cioè assume valore 1) quando almeno una tra P e Q ha valore 1, cioè è vera.

A questo punto diamo la tavola di verità della negazione:

P	$\neg$ P
0	1
1	0

Questa rappresenta il fatto che che se P è vera, la sua negazione è falsa e viceversa.

Siamo ora pronti a verificare la validità della formula (1.6). A tal fine ci basterà considerare tutti i possibili valori di verità di P, Q ed R ed verificare che per ogni scelta di valori il primo ed il secondo membro assumano lo stesso valore. Prima di tutto, i possibili valori delle tre incognite coinvolte sono

P	Q	R
0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1

Visto che nel secondo membro della (1.6) compaiono $P \wedge Q$ e $P \wedge R$, aggiungiamo due colonne con i valori di verità di queste due affermazioni (in corrispondenza dei valori che vediamo nelle prime tre colonne):

P	Q	R	P $\wedge$ Q	P $\wedge$ R
0	0	0	0	0
0	0	1	0	0
0	1	0	0	0
0	1	1	0	0
1	0	0	0	0
1	0	1	0	1
1	1	0	1	0
1	1	1	1	1

Ora possiamo scrivere la colonna di valori di $(P \wedge Q) \vee (P \wedge R)$ (cioè il secondo membro della (1.6)). Usando la tavola di verità di $\vee$ vediamo che $(P \wedge Q) \vee (P \wedge R)$ vale 0 solo quando sia $P \wedge Q$ che $P \wedge R$ valgono 0:

P	Q	R	P $\wedge$ Q	$P \wedge R$	(P $\wedge$ Q) $\vee$ (P $\wedge$ R)
0	0	0	0	0	0
0	0	1	0	0	0
0	1	0	0	0	0
0	1	1	0	0	0
1	0	0	0	0	0
1	0	1	0	1	1
1	1	0	1	0	1
1	1	1	1	1	1

Fino a questo punto, la tabella che abbiamo scritto ci dice che valori assume il secondo membro della (1.6) in funzione di tutti i valori possibili di P, Q, R. Ora scriviamo la tavola di verità del primo membro della (1.6), cioè $P \wedge (Q \vee R)$ per vedere se otteniamo esattamente gli stessi valori. Iniziamo con lo scrivere la colonna di $Q \vee R$

P	Q	R	P $\wedge$ Q	$P \wedge R$	$(P \wedge Q) \vee (P \wedge R)$	Q $\vee$ R
0	0	0	0	0	0	0
0	0	1	0	0	0	1
0	1	0	0	0	0	1
0	1	1	0	0	0	1
1	0	0	0	0	0	0
1	0	1	0	1	1	1
1	1	0	1	0	1	1
1	1	1	1	1	1	1

Infine, guardando la prima e l'ultima colonna ed usando la tavola di verità di $\wedge$, scriviamo la colonna con i valori di $P \wedge (Q \vee R)$, cioè il secondo membro della (1.6):

P	Q	R	P $\wedge$ Q	$P \wedge R$	$(P \wedge Q) \vee (P \wedge R)$	Q $\vee$ R	P $\wedge$ (Q $\vee$ R)
0	0	0	0	0	0	0	0
0	0	1	0	0	0	1	0
0	1	0	0	0	0	1	0
0	1	1	0	0	0	1	0
1	0	0	0	0	0	0	0
1	0	1	0	1	1	1	1
1	1	0	1	0	1	1	1
1	1	1	1	1	1	1	1

Vediamo allora che la sesta e l'ottava colonna, corrispondenti rispettivamente al primo e al secondo membro della (1.6), sono uguali. In altre parole, le espressioni corrispondenti assumono lo stesso valore qualunque siano i valori che assegniamo a P, Q ed R. Questo dimostra la validità dell'uguaglianza (1.6).

La (1.6) ci dice che nell'algebra Booleana vale la proprietà distributiva di $\wedge$ rispetto a $\vee$.

Osservazione 1.10. Si noti che se usiamo la notazione (che si trova spesso in letteratura) della somma $+$ per indicare la congiunzione $\vee$ e del prodotto $\cdot$ per indicare $\wedge$, la (1.6) può essere riscritta come

$$P \cdot (Q + R) = P \cdot Q + P \cdot R$$

che ha il fascino di essere formalmente identica alla analoga proprietà distributiva valida nell'algebra usuale. Tuttavia, non tutte le proprietà dell'algebra Booleana risultano avere la forma di identità valide anche nell'algebra usuale. Ad esempio, con i metodi usati precedentemente, si può dimostrare che vale la distributività di $\vee$ rispetto a $\wedge$, ovvero che

$$P \vee (Q \wedge R) = (P \vee Q) \wedge (P \vee R)$$

(la dimostrazione di questa identità è proposta come esercizio alla fine del capitolo). Nella notazione con somma e prodotto usata sopra, questa formula diventa

$$P + Q \cdot R = (P + Q) \cdot (P + R)$$

che è chiaramente falsa nell'algebra usuale, ma è un'identità valida nell'algebra Booleana.

Ispirandoci alle proprietà delle operazioni somma e prodotto dell'algebra usuale e in particolare alle proprietà commutativa ($a + b = b + a$ e $ab = ba$) e associativa ($(a + b) + c = a + (b + c)$ e $(ab)c = a(bc)$), ci chiediamo se tali proprietà sono vere anche rispetto a $\vee$ e $\wedge$ in algebra Booleana. Ovvero ci chiediamo se valgono le identità

$$P \vee Q = Q \vee P, \tag{1.7}$$
$$P \wedge Q = Q \wedge P, \tag{1.8}$$
$$(P \vee Q) \vee R = P \vee (Q \vee R) \text{ e} \tag{1.9}$$
$$(P \wedge Q) \wedge R = P \wedge (Q \wedge R)\,. \tag{1.10}$$

La risposta è sì, e la validità di tali formule è facile da verificare. Per esempio, la (1.7) si dimostra ricordando che $\vee$ è falsa solo se P e Q sono entrambe false, e chiaramente questo non dipende dall'ordine in cui le scriviamo.
Si osservi che se per P, Q ed R prendiamo rispettivamente $x \in A$, $x \in B$ e $x \in C$, dove A, B, C sono tre insiemi dati, allora le uguaglianze di sopra si riscrivono rispettivamente

$$\begin{aligned}
x \in A \vee x \in B &= x \in B \vee x \in A,\\
x \in A \wedge x \in B &= x \in B \wedge x \in A,\\
(x \in A \vee x \in B) \vee x \in C &= x \in A \vee (x \in B \vee x \in C) \text{ e}\\
(x \in A \wedge x \in B) \wedge x \in C &= x \in A \wedge (x \in B \wedge x \in C)\,.
\end{aligned}$$

La validità di queste uguaglianze significa che se il loro primo membro vale 1 (cioè è vero) allora lo è anche il secondo membro, e viceversa. Allora, ricordando le definizioni

di unione e intersezione, esse dimostrano le uguaglianze

$$A \cup B = B \cup A, \tag{1.11}$$
$$A \cap B = B \cap A, \tag{1.12}$$
$$(A \cup B) \cup C = A \cup (B \cup C) \text{ e} \tag{1.13}$$
$$(A \cap B) \cap C = A \cap (B \cap C)\,. \tag{1.14}$$

Osservazione 1.11. Le uguaglianze (1.11) e (1.12) ci dicono che sia per l'unione che per l'intersezione vale la proprietà commutativa mentre la (1.13) e la (1.14) ci dicono che vale per entrambe le operazioni la proprietà associativa. In altri termini è indifferente dove scriviamo le parentesi tonde. Tale proprietà ci consente quindi di omettere le parentesi e di scrivere semplicemente $A \cup B \cup C$ e $A \cap B \cap C$, senza ambiguità[2].

In una espressione dell'algebra usuale possono comparire anche numeri reali (ad esempio $2x + 3y$), ovvero delle costanti, rispetto alle incognite che possono assumere qualunque valore reale. Allo stesso modo in algebra Booleana possono comparire valori numerici costanti, in questo caso solo 0 e 1. Si ha ad esempio la seguente identità algebrica base:

$$P \wedge 1 = P$$

dimostrata dalla tavola seguente, dove si vede che la colonna di $P \wedge 1$ è uguale alla colonna di P

P	1	P $\wedge$1
0	1	0
1	1	1

Analogamente, abbiamo le

$$P \wedge 0 = 0$$
$$P \vee 0 = P$$
$$P \vee 1 = 1$$

dimostrate rispettivamente dalle tabelle

[2]In un'operazione non associativa, invece, le parentesi sono necessarie. Per esempio l'espressione $2 : 2 : 2$ è ambigua. Infatti se la interpretiamo come $2 : (2 : 2)$ essa vale 2, mentre se la interpretiamo come $(2 : 2) : 2$ essa vale $\frac{1}{2}$.

P	0	P $\wedge$0
0	0	0
1	0	0

P	0	P $\vee$0
0	0	0
1	0	1

P	1	P $\vee$1
0	1	1
1	1	1

Si noti che nella notazione con somma $+$ per $\vee$ e prodotto $\cdot$ per $\wedge$, le identità di sopra diventano

$$P \cdot 1 = P$$

$$P \cdot 0 = 0$$

$$P + 0 = P$$

$$P + 1 = 1$$

Con questa notazione, vediamo che le prime tre sono vere anche nell'algebra usuale, mentre la quarta vale solo nell'algebra Booleana.

Vediamo ora anche alcune identità che coinvolgono la negazione $\neg$. La prima è

$$\neg(\neg P) = P$$

la cui validità è dimostrata dalla tabella

P	$\neg$ P	$\neg(\neg P)$
0	1	0
1	0	1

Come esempio di formule che valgono in algebra Booleana ma non in algebra usuale, diamo inoltre le seguenti:

$$P \wedge P = P$$

$$P \vee P = P$$

Si dice che le operazioni $\wedge$ e $\vee$ sono *idempotenti*. Nella notazione con somma e prodotto tali proprietà si scrivono rispettivamente $P \cdot P = P$ e $P + P = P$. La loro verifica mediante tavole di verità è immediata:

P	$P \wedge P$
0	0
1	1

P	$P \vee P$
0	0
1	1

Altre due identità molto importanti in algebra Booleana sono le seguenti:

$$\neg(P \wedge Q) = \neg P \vee \neg Q$$

$$\neg(P \vee Q) = \neg P \wedge \neg Q$$

note col nome di **leggi di De Morgan**.
La loro validità è dimostrata dal fatto che la quarta e l'ultima colonna in ognuna delle seguenti tabelle sono uguali:

P	Q	P $\wedge$ Q	$\neg(P \wedge Q)$	$\neg P$	$\neg Q$	$\neg P \vee \neg Q$
0	0	0	1	1	1	1
0	1	0	1	1	0	1
1	0	0	1	0	1	1
1	1	1	0	0	0	0

P	Q	P $\vee$ Q	$\neg(P \vee Q)$	$\neg P$	$\neg Q$	$\neg P \wedge \neg Q$
0	0	0	1	1	1	1
0	1	1	0	1	0	0
1	0	1	0	0	1	0
1	1	1	0	0	0	0

1.4 L'implicazione

Oltre a "e" e "o" c'è un terzo modo di ottenere una proposizione combinandone due date P e Q, ovvero l'implicazione "se P allora Q", denotata con il simbolo $P \Rightarrow Q$.
Ad esempio, se P è la proposizione "piove" e Q è la proposizione "prendo l'ombrello" allora $P \Rightarrow Q$ è la nuova proposizione "se piove allora prendo l'ombrello.
Nell'espressione $P \Rightarrow Q$ (che si legge anche "P implica Q") P si dice *antecedente* e Q *conseguente*.
Così come fatto per $\wedge$ e $\vee$, anche l'implicazione $\Rightarrow$ viene definita mediante una tavola di verità, che è la seguente:

P	Q	P $\Rightarrow$ Q
0	0	1
0	1	1
1	0	0
1	1	1

Tale tavola implica che affermazioni quali "se 1=2 allora Parigi è la capitale dell'Italia" e "se 1=2 allora Roma è la capitale dell'Italia" sono vere (la prima per la prima riga della tavola, in cui sia antecedente che conseguente sono falsi, e la seconda per la seconda riga, in cui l'antecedente è falso e il conseguente è vero). A parte l'impressione di spaesamento provocato dal fatto che in queste affermazioni antecedente e conseguente non hanno nessun legame di tipo causale o temporale tra loro (come invece succede quando si usa l'implicazione nel linguaggio comune), si potrebbe comprensibilmente avere qualche difficoltà nell'accettare che sia vera un'implicazione in cui l'antecedente è falso e il conseguente è vero oppure in cui sia antecedente che conseguente sono falsi. Tuttavia, ci sono dei motivi precisi per cui i logici (a partire da quelli dell'antichità per arrivare, attraverso i logici medievali, a quelli moderni) hanno assegnato all'implicazione la tavola di verità data sopra. Di seguito ne riportiamo due.

(1) La prima giustificazione alla tavola di verità dell'implicazione è che vogliamo che essa sia l'operazione logica che corrisponde all'inclusione tra insiemi. Più precisamente, dati due insiemi A e B vogliamo che $A \subseteq B$ esattamente quando l'implicazione

$$x \in A \Rightarrow x \in B \tag{1.15}$$

è vera per ogni x, cioè qualunque cosa sostituiamo a x.

Per esempio, supponiamo che $A = \{1,2,3\}$ e $B = \{1,2,3,4,5\}$, per i quali chiaramente vale $A \subseteq B$. Vogliamo che (1.15) sia vera qualunque cosa sostituiamo ad x.

In particolare, se poniamo $x = 4$ vogliamo che l'implicazione

$$4 \in A \Rightarrow 4 \in B$$

sia vera. Poiché qui l'antecedente è falso e il conseguente è vero, questa è esattamente la situazione descritta dalla seconda riga della tavola di verità di $\Rightarrow$.

Ponendo invece ad esempio $x = 8$, vogliamo che anche l'implicazione

$$8 \in A \Rightarrow 8 \in B$$

sia vera. Ora sia l'antecedente che il conseguente sono falsi, perciò questa è la situazione descritta dalla prima riga della tavola di verità di $\Rightarrow$.

Questo giustifica la prima e la seconda riga della tavola di verità data per $\Rightarrow$ che sono solitamente quelle di più difficile comprensione.

(2) Una seconda giustificazione della tavola di verità dell'implicazione è che $P \Rightarrow Q$ va letto come un sinonimo di

$$\neg(P \wedge \neg Q)\,.$$

Infatti, se il primo significa "P implica Q", il secondo si interpreta come "non è possibile che valgano contemporaneamente P e la negazione di Q", che è un'interpretazione ragionevole del fatto che P implichi Q. Se si accetta tale equivalenza, allora la tavola di verità di $P \Rightarrow Q$ è esattamente quella di $\neg(P \wedge \neg Q)$, che ricaviamo subito:

P	Q	$\neg Q$	$P \wedge \neg Q$	$\neg(P \wedge \neg Q)$
0	0	1	0	1
0	1	0	0	1
1	0	1	1	0
1	1	0	0	1

Come si vede questa coincide con la tavola di verità data sopra per $\Rightarrow$.

Osserviamo che per l'operazione di implicazione non vale la proprietà commutativa, ovvero l'uguaglianza $(P \Rightarrow Q) = (Q \Rightarrow P)$ non è vera. Basta verificare che le due espressioni non hanno la stessa tavola di verità (ad esempio, quando P vale 0 e Q vale 1, l'implicazione $P \Rightarrow Q$ vale 1 mentre l'implicazione opposta $Q \Rightarrow P$ vale 0). Si dimostra invece che vale l'uguaglianza

$$(P \Rightarrow Q) = (\neg Q \Rightarrow \neg P)$$

come si evince dalla tavola di verità di $\neg Q \Rightarrow \neg P$

P	Q	$\neg Q$	$\neg P$	$\neg Q \Rightarrow \neg P$
0	0	1	1	1
0	1	0	1	1
1	0	1	0	0
1	1	0	0	1

che risulta essere la stessa di $P \Rightarrow Q$.
La proposizione $\neg Q \Rightarrow \neg P$ è detta **contronominale** di $P \Rightarrow Q$. Il fatto che ogni implicazione sia logicamente equivalente alla sua contronominale è spesso usato per riformulare in maniera alternativa un teorema da dimostrare. Ad esempio, per dimostrare il teorema "se x è dispari allora x^2 è dispari" si può equivalentemente dimostrare l'affermazione "se x^2 è pari allora x è pari", ossia la sua contronominale.

Osservazione 1.12. Se vale l'implicazione $P \Rightarrow Q$, si dice che *P è condizione sufficiente per Q* mentre *Q è condizione necessaria per P*. Questo è in accordo con l'uso di tali espressioni nel linguaggio ordinario. Per esempio il fatto che l'implicazione "se x è veneto allora x è italiano" sia vera ci dice che essere veneti è condizione sufficiente per essere italiani (ma non necessaria: si può essere italiani anche essendo toscani, tra gli altri) e che essere italiani è condizione necessaria per essere veneti (ma non sufficiente: esistono italiani che non sono veneti, per esempio i toscani).

Un altro simbolo logico che si può definire a partire dall'implicazione è quello della **doppia implicazione** $\Leftrightarrow$. Scriveremo $P \Leftrightarrow Q$ per intendere

$$(P \Rightarrow Q) \wedge (Q \Rightarrow P)$$

Ricavando la tavola di verità di questa espressione come segue

P	Q	$P \Rightarrow Q$	$Q \Rightarrow P$	$(P \Rightarrow Q) \wedge (Q \Rightarrow P)$
0	0	1	1	1
0	1	1	0	0
1	0	0	1	0
1	1	1	1	1

otteniamo la tavola di verità della doppia implicazione è

P	Q	$P \Leftrightarrow Q$
0	0	1
0	1	0
1	0	0
1	1	1

Quindi una doppia implicazione è vera solamente quando le due proposizioni che la compongono hanno lo stesso valore di verità.

Osservazione 1.13. Dal momento che due insiemi A e B sono uguali esattamente quando per ogni x si ha $x \in A \Rightarrow x \in B$ e viceversa $x \in B \Rightarrow x \in A$, possiamo unire queste due condizioni usando la doppia implicazione e dire che due insiemi A e B sono uguali quando per ogni x si ha $x \in A \Leftrightarrow x \in B$.

1.5 Tautologie e contraddizioni

Consideriamo l'espressione $P \vee \neg P$ e ricaviamone la tavola di verità:

P	$\neg P$	$P \vee \neg P$
0	1	1
1	0	1

Notiamo che tale espressione assume sempre valore 1 (cioè è sempre vera) indipendentemente dal valore di P.

Definizione 1.14. Si dice **tautologia** una proposizione che assume sempre il valore di verità 1.

Il fatto che la $P \vee \neg P$ sia una tautologia risulta chiaro se la si interpreta come "o vale P o non vale P". Questo rende evidente il fatto che essa sia sicuramente e sempre vera (si tratta del cosiddetto *principio del terzo escluso*).

Consideriamo ora invece l'espressione $P \wedge \neg P$ e ricaviamone la tavola di verità:

P	$\neg P$	$P \wedge \neg P$
0	1	0
1	0	0

Notiamo stavolta che tale espressione assume sempre valore 0 (cioè è sempre falsa) indipendentemente dal valore di P.

Definizione 1.15. Una **contraddizione** è un'espressione che assume sempre valore di verità 0.

Il fatto che la $P \wedge \neg P$ sia una contraddizione risulta ovvio se la si interpreta come "si ha contemporaneamente P e non P", che rende più evidente il fatto che essa si contraddica ed è quindi sicuramente sempre falsa.

Osservazione 1.16. Grazie alla doppia implicazione possiamo riformulare tutte le identità logiche viste sopra come tautologie. Per esempio, la validità di

$$P \wedge (Q \vee R) = (P \wedge Q) \vee (P \wedge R)$$

equivale al fatto che

$$P \wedge (Q \vee R) \Leftrightarrow (P \wedge Q) \vee (P \wedge R)$$

è una tautologia. Infatti, nella prima forma l'uguaglianza significa che il primo e il secondo membro assumono lo stesso valore per qualunque assegnazione dei valori di verità di P, Q ed R. Dal momento che, come si vede nella sua tavola di verità, la doppia implicazione $P_1 \Leftrightarrow P_2$ è vera esattamente quando P_1 e P_2 hanno lo stesso valore di verità, la seconda forma risulta essere una tautologia.

Osservazione 1.17. L'algebra Booleana che abbiamo discusso prevede che le incognite (che rappresentano proposizioni) possano assumere solo i due valori 0 e 1 (ovvero essere o false o vere). Logici e matematici hanno però provato a sviluppare anche algebre nelle quali ci siano più valori di verità. Ad esempio 0, 1, $\frac{1}{2}$: una tale algebra vuole modellizzare una logica nella quale le proposizioni oltre a essere false o vere possono essere anche "possibili" (e assumono quindi un valore di verità intermedio, per l'appunto $\frac{1}{2}$).
Quando si aggiungono valori di verità, bisogna però stare attenti a definire le nuove tavole di verità (ovvero bisogna dire che valori assumono $P \wedge Q$, $P \vee Q$ e $\neg P$ in funzione dei vari valori di verità di P e Q, per i quali ora è ammesso anche $\frac{1}{2}$) in modo da mantenere alcune delle proprietà alle quali in nessuna logica si vuole rinunciare, ad esempio il fatto che $P \vee \neg P$ sia una tautologia o che $P \wedge \neg P$ sia una contraddizione.
Ad esempio, Lukasiewicz nel 1920 ha proposto di definire $P \wedge Q$ come il minimo dei valori di P e Q (quindi ad esempio se P vale $\frac{1}{2}$ e Q vale 1, allora $P \wedge Q$ vale $\frac{1}{2}$), $P \vee Q$ come il massimo dei valori di P e Q (quindi ad esempio se P vale $\frac{1}{2}$ e Q vale 1, allora $P \vee Q$ vale 1) e $\neg P$ come 0 non solo se P vale 1 ma anche se vale $\frac{1}{2}$.
Il problema di tali definizioni, che sembrano sensate (ad esempio,è ragionevole che la negazione di una proposizione possibile sia falsa), è che, mentre $P \wedge \neg P$ rimane una contraddizione (se P vale $\frac{1}{2}$ allora $\neg P$ in base alla scelta di Lukasiewicz vale 0, e quindi sempre in base alle sue scelte $P \wedge \neg P$ è comunque zero), si ha che $P \vee \neg P$ non è più una tautologia! Infatti, se P vale $\frac{1}{2}$, si ha che $\neg P$ vale 0 e quindi $P \vee \neg P$, che, per la definizione di Lukasiewicz è il massimo dei valori di verità di P e $\neg P$, vale $\frac{1}{2}$ e non 1.

Esistono anche le cosiddette *logiche fuzzy*, nelle quali le proposizioni possono assumere come valore di verità qualunque numero reale compreso tra 0 e 1 (l'idea è rendere infinite sfumature di verità). I problemi da affrontare in tali costruzioni sono gli stessi che abbiamo ilustrato nel caso della logica a 3 valori, ma omettiamo ulteriori dettagli.

1.6 I quantificatori universali

Concludiamo questo primo capitolo introducendo due importanti simboli della logica, detti *quantificatori universali*:

(1) il simbolo $\forall$, che si legge "per ogni". Per esempio, data una proprietà P, la formula $\forall x \in X\ P(x)$ significa "per ogni elemento x di X, x ha la proprietà P", ovvero "tutti gli elementi di X soddisfano la proprietà P.

(2) il simbolo $\exists$, che va letto "esiste". Ad esempio, data una proprietà P, la formula $\exists x \in X\ P(x)$ significa "esiste un elemento x di X tale che x ha la proprietà P", ovvero "almeno un elemento di X ha la proprietà P.

Osserviamo che la negazione di $\forall x \in X\ P(x)$ è $\exists x \in X\ \neg P(x)$. Infatti, se non è vero che per tutti gli elementi di X vale la proprietà P, allora ne esiste almeno uno che non la soddisfa (e non, come in un tipico errore, che non la soddisfi nessuno!!). Analogamente, la negazione di $\exists x \in X\ P(x)$ è $\forall x \in X\ \neg P(x)$. Infatti, se non è vero che esiste un elemento di X per cui vale la proprietà P, questo significa che nessun elemento di X soddisfa P (ovvero per ogni x, x non ha la proprietà P).

I quantificatori universali possono essere usati per estendere le definizioni di intersezione e unione di due insiemi ad un numero arbitrario di insiemi, anche infinito.
Più precisamente, iniziamo fissando una notazione opportuna per unioni e intersezioni che comprendano più di due insiemi. L'uso delle lettere dell'alfabeto A, B, C per indicare degli insiemi è estremamente limitante in quanto non ci consente, ad esempio, di scrivere l'unione o l'intersezione di più insiemi di quante siano le lettere dell'alfabeto. Tuttavia, tale problema si risolve semplicemente denotando tutti gli insiemi con la stessa lettera ma aggiungendo un indice che ci dica se si tratta del primo insieme, del secondo, del terzo e così via: A_1, A_2, A_3, etc.
In questo modo, possiamo scrivere l'unione e l'intersezione di un qualunque numero di insiemi (diciamo n insiemi, dove n rappresenta un qualunque numero naturale) con le notazioni seguenti:

$$A_1 \cap A_2 \cap \cdots \cap A_n, \quad A_1 \cup A_2 \cup \cdots \cup A_n\,.$$

Si noti che non abbiamo bisogno di parentesi quando uniamo o intersechiamo più di due insiemi grazie alla proprietà associativa di tali operazioni, cf. Osservazione 1.11.

Una notazione efficace nell'evitare di scrivere più volte il simbolo di intersezione (o unione) è la seguente

$$\cap_{j=1}^{n} A_j, \quad \cup_{j=1}^{n} A_j$$

che significano esattamente che si considera l'intersezione (o l'unione) degli insiemi A_j per j che varia da 1 a n.
In base alla definizione di intersezione, che usa la congiunzione "e", si ha quindi che $x \in \cap_{j=1}^{n} A_j$ se e solo se $x \in A_j$ è vera per tutti gli indici j da 1 a n. Analogamente, in base alla definizione di unione, che usa la congiunzione "o", si ha quindi che $x \in \cup_{j=1}^{n} A_j$ se e solo se $x \in A_j$ è vera per almeno un indice j da 1 a n, cioè se esiste almeno un j compreso tra 1 e n per cui si ha $x \in A_j$.
In questa riformulazione, l'intersezione e l'unione di insiemi si estende immediatamente a una qualunque famiglia di insiemi, anche infinita. Più precisamente, si consideri una famiglia A_i di insiemi, dove i è un indice che appartiene a un certo insieme di indici I (es. se $I = \mathbb{N}$, allora la famiglia è data da $A_0, A_1, A_2, A_3, \ldots$). L'intersezione di tutti gli insiemi di questa famiglia si denota con $\cap_{i \in I} A_i$ ed è definita come

$$\cap_{i \in I} A_i = \{x \mid \forall i \in I \; x \in A_i\},$$

ovvero l'insieme degli x che appartengono ad A_i per tutti gli indici $i \in I$.
Analogamente, l'unione di tutti gli insiemi di questa famiglia si denota con $\cup_{i \in I} A_i$ ed è definita come

$$\cup_{i \in I} A_i = \{x \mid \exists i \in I \; x \in A_i\},$$

ovvero l'insieme degli x che appartengono ad A_i per almeno un indice $i \in I$.

Capitolo 2

Relazioni, funzioni e calcolo combinatorio

2.1 Prodotto cartesiano

In questo capitolo introdurremo altre due importanti nozioni della teoria degli insiemi, quella di *relazione su un insieme* e quella di *funzione tra due insiemi*.
Per dare delle definizioni rigorose, abbiamo bisogno di introdurre prima una nuova operazione sugli insiemi, quella di *prodotto cartesiano*.

Definizione 2.1. Dati due insiemi A e B, si dice il **prodotto cartesiano di A per B**, denotato con $A \times B$, l'insieme di tutte le *coppie ordinate* (a,b) dove a è un elemento di A e b è un elemento di B. In simboli

$$A \times B = \{(a,b) \mid a \in A \wedge b \in B\} .$$

Dare una coppia ordinata (a,b) significa non solo dare due elementi a e b ma anche specificare che a è il primo e b è il secondo. In questo (a,b) si differenzia da $\{a,b\}$ (l'insieme degli elementi a e b) nel quale non importa l'ordine di a e b ma solo la loro appartenenza all'insieme dato[1]. Quindi, in generale, $(a,b) \neq (b,a)$ (a meno che non sia $a = b$).

Esempio 2.2. Siano $A = \{0,1,2,3\}$ e $B = \{3,4\}$. Ad esempio, la coppia $(0,3)$ appartiene a $A \times B$ in quanto la sua prima componente, 0, appartiene ad A e la seconda, 3, appartiene a B. Si ha in effetti

[1] Se volessimo dare una definizione insiemistica e rigorosa della coppia (a,b), questa potrebbe essere $(a,b) = \{\{a,b\},\{a\}\}$: l'insieme con due elementi ci dice quali sono gli elementi che compongono la coppia, l'insieme con un solo elemento ci dice chi è il primo della coppia. Tuttavia, non useremo mai tale notazione.

$$A \times B = \{(0,3),(0,4),(1,3),(1,4),(2,3),(2,4),(3,3),(3,4)\}$$

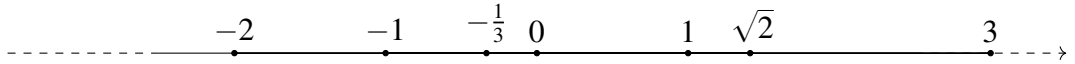
Si noti che il numero degli elementi di $A \times B$ è dato dal prodotto del numero degli elementi di A per il numero degli elementi di B. Infatti, per ogni elemento di A scelto come primo elemento della coppia, il secondo elemento può essere scelto tra uno qualunque degli elementi di B.

Come caso particolare possiamo considerare il prodotto di un insieme per se stesso. Per esempio, se $A = \{1,2,3\}$ si ha

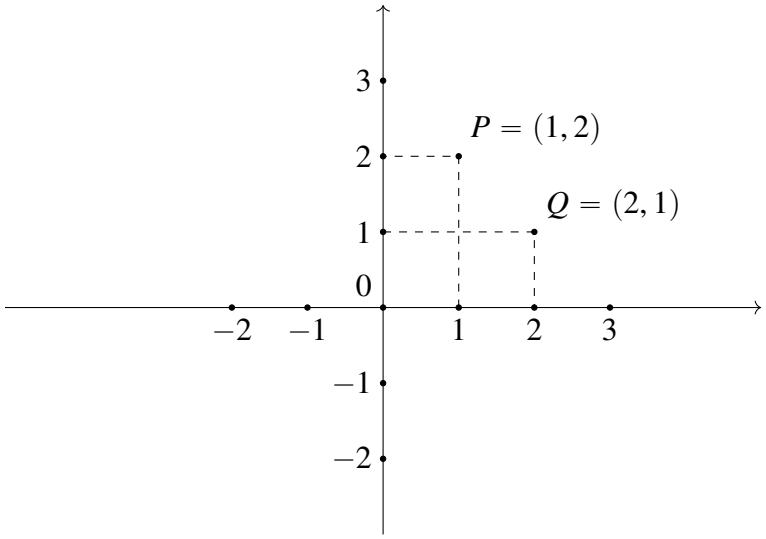
$$A \times A = \{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2),(3,3)\}$$

Il prodotto cartesiano $A \times A$ di un insieme con se stesso si denota anche con A^2, con una notazione presa in prestito da quella usata per il prodotto usuale tra numeri dove effettivamente si scrive $a \cdot a = a^2$.

Osservazione 2.3. Il prodotto cartesiano $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ ha particolare importanza in geometria in quanto, fissato un sistema di coordinate cartesiane, rappresenta l'insieme dei punti del piano. Più precisamente, sappiamo che i numeri reali si possono rappresentare come i punti di una retta scegliendo prima un punto per rappresentare lo 0. Poi a distanza fissata (da noi scelta come unità di misura) gli interi, positivi da un lato e negativi dall'altro (quindi la retta risulta orientata). Inserendo successivamente le frazioni tra i numeri interi e infine i restanti numeri reali che non possono essere rappresentati come frazioni (gli *irrazionali*, ovvero i numeri decimali con infinite cifre dopo la virgola, non periodici):



A questo punto, se nel piano scegliamo due rette orientate (dette *assi cartesiani*), perpendicolari tra loro e che si incontrano nel punto scelto per lo 0, a ogni punto P si può assegnare una coppia di numeri reali (le sue *coordinate*) che sono determinati proiettando P sui due assi.



Si noti che è importante che le coordinate siano rappresentate mediante una coppia ordinata, in quanto, ad esempio, il punto di coordinate $(1,2)$ è diverso dal punto di coordinate $(2,1)$.

Il prodotto cartesiano può anche essere definito per più di due insiemi. Ad esempio, per 3 insiemi A, B, C basta usare invece della nozione di coppia ordinata quella di *terna ordinata* (a,b,c), nella quale, analogamente a quanto fatto per la coppia, diamo tre elementi in un ordine assegnato. In questo caso definiamo

$$A \times B \times C = \{(a,b,c) \mid a \in A \wedge b \in B \wedge c \in C\}\,.$$

Possiamo generalizzare tale definizione a un numero arbitrario di insiemi $A_1, A_2, \ldots, A_n$: in tal caso, useremo la nozione di *n-upla ordinata* $(a_1, a_2, \ldots, a_n)$ (che coincide con una coppia per $n = 2$ e con una terna per $n = 3$) e definiremo

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \ldots, a_n) \mid a_1 \in A \wedge a_2 \in A_2 \wedge \cdots \wedge a_n \in A_n\}\,.$$

Analogamente a quanto già detto per il prodotto tra due insiemi, denoteremo con A^n il prodotto cartesiano $A \times A \times \cdots \times A$ di A con se stesso n volte. Vedremo in un capitolo successivo che il prodotto $\mathbb{R}^n$ ha una particolare importanza in algebra vettoriale.

2.2 Relazioni

Introduciamo la definizione rigorosa di relazione su un insieme tramite un esempio concreto. Supponiamo che X sia un insieme i cui elementi sono persone, e supponiamo che alcuni elementi dell'insieme siano figli di altri sempre appartenenti all'insieme dato. In

altre parole, alcuni elementi dell'insieme X sono legati dalla relazione padre-figlio ad altri elementi di X.
Per raccogliere l'informazione su quali sono gli elementi di X in tale relazione tra loro, si può fare ricorso proprio alla nozione di coppia ordinata: ogniqualvolta si ha che x è figlio di x' per due elementi $x, x' \in X$, formiamo la coppia ordinata (x, x'). L'insieme delle coppie così costruite è un sottoinsieme del prodotto cartesiano $X \times X$ che ci dice esattamente chi è figlio di chi all'interno dell'insieme.
Si noti che per avere l'informazione corretta è necessario usare la coppia *ordinata*, perchè distinguendo tra la prima e la seconda componente specifica quale dei due elementi è figlio dell'altro. Quest'informazione sarebbe persa se usassimo solo l'insieme $\{x, x'\}$, che ci direbbe solo che tra x e x' uno è figlio dell'altro senza specificare chi. Questo ci porta alla seguente definizione generale.

Definizione 2.4. Una **relazione** su un insieme X è un sottoinsieme del prodotto cartesiano $X \times X$.

Esempio 2.5. Consideriamo sull'insieme dei numeri naturali la relazione "essere minore di". Questa relazione è rappresentata dal sottoinsieme S (infinito) di tutte le coppie (n, m) di numeri naturali in cui $n < m$:

$$S = \{(0,1), (0,2), \ldots, (1,2), (1,3), \ldots\}.$$

Se invece la relazione fosse quella di uguaglianza, il sottoinsieme sarebbe

$$S = \{(0,0), (1,1), (2,2), (3,3) \ldots\}$$

in cui ogni elemento è in relazione solo con se stesso (in quanto uguale solo a se stesso).

Quando un elemento $x \in X$ è nella relazione data con un altro $x' \in X$, useremo la simbologia $x \sim x'$.

Come abbiamo osservato sopra, si usano coppie ordinate per formalizzare la definizione di relazione perchè in generale può essere $x \sim x'$ ma non $x' \sim x$. Ad esempio, per la relazione "essere figlio di", se x è figlio di x' sicuramente x' non è figlio di x.
Tuttavia, per alcune relazioni succede che

$$x \sim x' \Rightarrow x' \sim x. \tag{2.1}$$

In tal caso, si dice che la relazione è **simmetrica**.
Ad esempio, la relazione di uguaglianza è simmetrica in quanto se $x = x'$ sicuramente possiamo anche scrivere $x' = x$.
Un'altra proprietà soddisfatta dalla relazione di uguaglianza ma non dalle relazioni in generale, è che per ogni $x \in X$ si ha

$$x \sim x. \tag{2.2}$$

Ovvero ogni elemento è in relazione con se stesso (nel caso dell'uguaglianza, chiaramente questo vale in quanto $x = x$). Tra molte altre, la relazione "essere minore di" non soddisfa questa proprietà in quanto $x < x$ non è verificata. Quando vale questa proprietà, si dice che la relazione è **riflessiva**.

Infine, l'uguaglianza soddisfa una terza importante proprietà:

$$x \sim x' \text{ e } x' \sim x'' \Rightarrow x \sim x'' . \tag{2.3}$$

Ovvero se un elemento è in relazione con un secondo, e il secondo elemento è in relazione con un terzo, allora il primo è in relazione con il terzo. Quando vale ciò, la relazione si dice **transitiva**.

L'uguaglianza è chiaramente una relazione transitiva in quanto se $x = x'$ e $x' = x''$, allora $x = x''$. Anche la relazione "essere minore di" soddisfa tale proprietà in quanto $x < x'$ e $x' < x''$ implica $x < x''$. Diversamente la relazione "essere figlio di" invece non è transitiva, in quanto se x è figlio di x' e x' è figlio di x'', non è vero che x è figlio di x'' (sarà semmai il nipote di x'').

Definizione 2.6. Una relazione su un insieme X che sia riflessiva, simmetrica e transitiva si dice **relazione di equivalenza**.

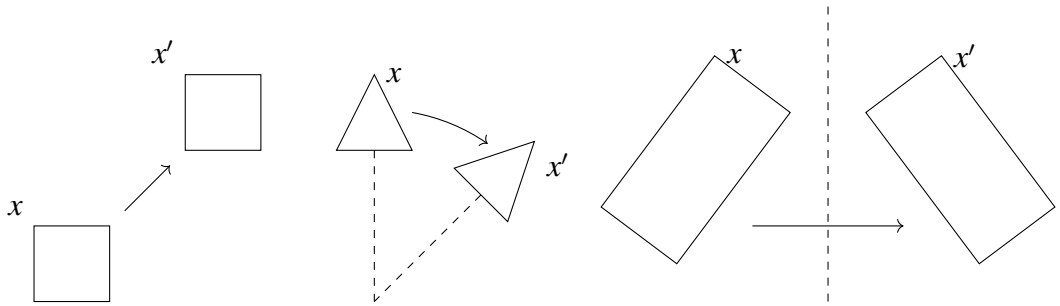
La definizione precedente è di grande importanza per la matematica e in questo corso in particolare. Per questo motivo sarà l'oggetto della prossima sezione.

2.3 Relazioni di equivalenza

Come visto nella sezione precedente, la relazione di uguaglianza è una relazione di equivalenza. In generale, una relazione di equivalenza è una sorta di "uguaglianza in senso largo", utile per identificare, ovvero considerare in un qualche senso uguali, oggetti che uguali non sono.

A sostegno di questa affermazione, consideriamo l'esempio seguente. Sia X l'insieme delle *figure* del piano, ovvero quadrati, triangoli, circonferenze e in generale qualunque insieme di punti.

Definiamo su tale insieme la seguente relazione: date due figure x e x', diciamo che $x \sim x'$ se x' può essere ottenuta da x attraverso una sequenza di traslazioni, rotazioni o riflessioni di x rispetto a qualche retta (ovvero tramite cosiddetti *movimenti rigidi*), come nel disegno seguente.



Mostriamo che si tratta di una relazione di equivalenza (che prende il nome di *congruenza*). Se $x \sim x'$, ovvero x' si ottiene da x mediante traslazioni, rotazioni o riflessioni, allora anche $x' \sim x$, in quanto x si può riottenere da x' applicando le trasformazioni inverse nell'ordine inverso. Questa relazione è quindi simmetrica.
Inoltre, se $x \sim x'$ e $x' \sim x''$, questo significa che x' si ottiene da x mediante certi movimenti rigidi, e x'' si ottiene da x' mediante altri movimenti di tale tipo. Ma allora anche x'' si ottiene da x mediante movimenti rigidi (e quindi $x \sim x''$): basta applicare in sequenza quelli che servono prima per ottenere x' e poi quelli che servono per ottenere x'' da x'. Quindi abbiamo anche la transitività.
Quanto alla riflessività, ovvero $x \sim x$, possiamo dire che vale in quanto x si ottiene da se stesso applicando per esempio una rotazione di angolo 0.
Ora, questo esempio chiarisce perché, come abbiamo detto sopra, una relazione di equivalenza può essere considerata come un'uguaglianza in senso largo. I due quadrati della figura precedente, così come i due triangoli o i due rettangoli, pur essendo figure diverse (non sono formate dagli stessi punti, quindi non possiamo dire che siano uguali), rappresentano tutte in un certo senso "la stessa figura".

Esempio 2.7. Tra le relazioni di parentela, un esempio di relazione di equivalenza è "essere fratello di", nel senso di avere gli stessi genitori. La riflessività è chiara: ogni x ha gli stessi genitori di se stesso. La simmetria è altrettanto palese, dal momento che se x è fratello di x', anche x' è fratello di x. Infine si ha anche la transitività, poiché se x è fratello di x' e x' è fratello di x'', allora è chiaro che x è fratello di x''.

Data una relazione di equivalenza, tutti gli elementi equivalenti tra loro possono essere raggruppati in in una cosiddetta classe di equivalenza.

Definizione 2.8. Sia X un insieme sul quale è definita una relazione d'equivalenza e sia $x \in X$. La **classe di equivalenza** di x è l'insieme

$$[x] = \{y \in X \mid y \sim x\} \tag{2.4}$$

cioè l'insieme di tutti gli elementi y che sono in relazione con (ovvero equivalenti a) x.

Elementi distinti x e x' possono dar luogo alla stessa classe di equivalenza, ovvero $[x] = [x']$. Questo accade esattamente quando $x \sim x'$.

Proposizione 2.9. *Si consideri una relazione di equivalenza $\sim$ su un insieme X. Per ogni $x, x' \in X$, si ha $[x] = [x'] \Leftrightarrow x \sim x'$.*

Dimostrazione. Trattandosi di una doppia implicazione, dobbiamo dimostrare sia $[x] = [x'] \Rightarrow x \sim x'$ che $x \sim x' \Rightarrow [x] = [x']$.
Per dimostrare la prima implicazione, ricordiamoci che, per definizione di uguaglianza tra insiemi, $[x] = [x']$ significa che ogni elemento della classe $[x]$ sta anche nella classe $[x']$ e viceversa ogni elemento di $[x']$ appartiene anche a $[x]$. Ma nella classe $[x]$ c'è sicuramente x stesso, perchè essendo la relazione di equivalenza riflessiva vale $x \sim x$, e quindi $x \in \{y \in X \mid y \sim x\} = [x]$. Quindi, essendo $x \in [x]$ e $[x] = [x']$, si ha anche $x \in [x']$, ovvero $x \sim x'$, come volevamo.
Per dimostrare l'implicazione inversa $x \sim x' \Rightarrow [x] = [x']$, dobbiamo far vedere che (sotto l'ipotesi $x \sim x'$), ogni elemento di $[x]$ appartiene anche a $[x']$ e viceversa.
Sia $y \in [x]$: questo significa che $y \sim x$; ma poiché abbiamo $x \sim x'$, possiamo applicare la transitività delle relazioni di equivalenza e concludere che $y \sim x'$, ovvero $y \in [x']$. Abbiamo allora dimostrato che ogni elemento di $[x]$ sta anche in $[x']$, ovvero che $[x] \subseteq [x']$. Per dimostrare che vale proprio l'uguaglianza, prendiamo viceversa un $y \in [x']$: questo significa che $y \sim x'$. Combinando con $x' \sim x$ (l'ipotesi sarebbe $x \sim x'$, ma è la stessa cosa visto che una relazione di equivalenza è anche simmetrica) e applicando la transitività si ottiene $y \sim x$, ovvero $y \in [x]$. Quindi è vero anche che ogni elemento di $[x']$ sta in $[x]$, cioè abbiamo finalmente $[x] = [x']$. □

Le classi di equivalenza hanno l'importante proprietà di ripartire l'insieme X su cui è definita la relazione in sottoinsiemi non vuoti e disgiunti, come affermato dal seguente risultato.

Proposizione 2.10. *Data una relazione di equivalenza $\sim$ su un insieme X, le classi di equivalenza hanno le seguenti tre proprietà:*

(i) *per ogni $x \in X$, si ha $[x] \neq \emptyset$,*

(ii) *l'unione $\bigcup_{x \in X} [x]$ di tutte le classi di equivalenza è uguale ad X e*

(iii) *due classi di equivalenza distinte sono disgiunte, ovvero $[x] \neq [x'] \Rightarrow [x] \cap [x'] = \emptyset$.*

Dimostrazione. La (i) si dimostra semplicemente ricordando (lo abbiamo visto nella dimostrazione della Proposizione 2.9) che x appartiene sempre alla sua classe di equivalenza $[x]$, quindi questa non è sicuramente vuota.

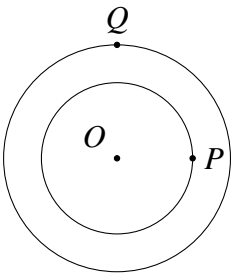
Per dimostrare la (ii), in base alla definizione di unione basta mostrare che ogni elemento di X sta in almeno una classe di equivalenza. Questo è vero per quanto appena detto, in quanto ogni elemento x sta almeno nella sua classe $[x]$.
Per dimostrare la (iii), dimostriamo la sua contronominale

$$[x] \cap [x'] \neq \varnothing \Rightarrow [x] = [x'] \tag{2.5}$$

che come sappiamo dal capitolo precedente è equivalente all'implicazione in (iii).
L'ipotesi $[x] \cap [x'] \neq \varnothing$ significa che esiste un z che appartiene sia a $[x]$ che a $[x']$. Allora, per definizione di classe , valgono sia $z \sim x$ (o, equivalentemente visto che la relazione è simmetrica, $x \sim z$) che $z \sim x'$. Mettendo insieme $x \sim z$ e $z \sim x'$, per la transitività abbiamo $x \sim x'$. Ma la Proposizione 2.9 ci dice che allora $[x] = [x']$, e la (2.5) è dimostrata. □

In generale, dato un insieme X, una famiglia di sottoinsiemi non vuoti, disgiunti e la cui unione sia tutto X si chiama una **partizione** di X. Quindi, la Proposizione 2.10 afferma che una relazione di equivalenza su X determina una partizione di X (in classi di equivalenza).

Esempio 2.11. Si consideri l'insieme X dei punti del piano, si fissi un punto O e per ogni $P, P' \in X$, si ponga $P \sim P'$ quando $|OP| = |OP'|$ (cioè quando i due punti sono alla stessa distanza da O).
È facile vedere che si tratta di una relazione di equivalenza. Infatti si ha $P \sim P$ in quanto ovviamente $|OP| = |OP|$. Inoltre $P \sim P'$ implica $P' \sim P$ in quanto $|OP| = |OP'|$ implica (per simmetria della relazione di uguaglianza) $|OP'| = |OP|$. Infine se $P \sim P'$ e $P' \sim P''$, ovvero $|OP| = |OP'|$ e $|OP'| = |OP''|$, visto che l'uguaglianza gode della proprietà transitiva si ha $|OP| = |OP''|$, ovvero $P \sim P''$.
Chi sono le classi di equivalenza di questa relazione? Dato un punto P, che dista $|OP| = R$ da O, la sua classe di equivalenza è data da tutti i punti che distano anch'essi R da O, ovvero è la circonferenza centrata in O su cui si trova P



Questo esempio illustra come le classi di equivalenza diano una partizione dell'insieme dato. In questo caso le circonferenze costituiscono una partizione del piano. Due cir-

conferenze di raggio diverso sono disgiunte e la loro unione, al crescere del raggio, ci dà tutto il piano[2].

Esempio 2.12. Se la relazione di equivalenza è l'uguaglianza, ogni classe di equivalenza contiene un solo elemento, cioè $[x] = \{x\}$: in questo caso la partizione è quella in singoli elementi.

Osservazione 2.13. Notiamo che se definissimo anche per una relazione $\sim$ che non sia di equivalenza le classi mediante la (2.4), ovvero dicendo che la classe $[x]$ di x è l'insieme di tutti gli elementi in relazione con x, non sarebbe più vero che tali classi determinano una partizione.
Ad esempio, consideriamo un insieme di quattro individui a_1, a_2, a_3, a_4 nel quale a_2 e a_3 sono figli di a_1 e a_4 è figlio di a_2. Rispetto alla relazione "essere figlio di", allora, la classe di a_1 (cioè l'insieme degli x tali che x è figlio di a_1) è $\{a_2, a_3\}$, la classe di a_2 è $\{a_4\}$, mentre le classi di a_3 e a_4 sono vuote, in quanto non esistono figli di a_3 e a_4 nell'insieme dato. Quindi non vale la (i) della Proposizione 2.10 (ci sono classi vuote) e neanche la (ii), in quanto unendo le classi $\{a_2, a_3\}$, $\{a_4\}$ e le due classi vuote si ottiene $\{a_2, a_3, a_4\}$, che non è tutto l'insieme.

Per vedere invece un esempio di relazione non di equivalenza in cui non vale la (iii) della Proposizione 2.10, si prenda ad esempio l'insieme $\mathbb{N}$ dei numeri naturali e la relazione $<$ (minore). Allora, la classe di 0 è vuota (non esiste nessun $x < 0$), la classe di 1 è $\{0\}$ (zero è l'unico numero naturale per cui $x < 1$ è verificata), la classe di 2 è $\{0, 1\}$, di 3 è $\{0, 1, 2\}$ e così via. Come si vede, ci sono classi diverse che non sono disgiunte, diversamente da quanto affermato dalla (iii) della Proposizione 2.10 nel caso di relazioni di equivalenza.

Nel prossimo capitolo studieremo un'altra importante relazione di equivalenza sull'insieme $\mathbb{Z}$ degli interi che ci porterà a comprendere alcuni moderni metodi crittografici.

Esempio 2.14. Per dare un'anticipazione, illustrata da un esempio concreto, prendiamo l'insieme degli interi $\mathbb{Z}$ e supponiamo che lo 0 rappresenti la mezzanotte di un dato giorno fissato, mentre ogni intero positivo (risp. negativo) rappresenti il numero di ore successive (risp. precedenti) alla mezzanotte.
Decidiamo di considerare equivalenti due interi x e y se essi corrispondono alla stessa ora del giorno (eventualmente di giorni diversi): definiamo cioè la relazione (nel capitolo successivo, dimostreremo che si tratta di una relazione di equivalenza[3]) $x \sim y$ se e solo se la differenza tra x e y è un multiplo di 24.
Ad esempio, 0, 24, 48, -24 sono equivalenti in quanto rappresentano la mezzanotte (rispettivamente del giorno in questione, del giorno successivo, di due giorni dopo e del

[2] Si noti che la classe di equivalenza di O contiene solo O, ovvero è una circonferenza di raggio 0.

[3] Anzi dimostreremo più in generale che, fissato un qualunque k, se per due interi n, m si definisce $n \sim m$ se e solo se la differenza $n - m$ è un multiplo di k questo dà una relazione di equivalenza.

giorno precedente). Gli interi $+1$, $+25$, $+49$, -23 sono equivalenti in quanto rappresentano l'una del mattino (rispettivamente del giorno in questione, del giorno successivo, di due giorni dopo e del giorno precedente).

Esempio 2.15. Si noti che se al posto di 24 avessimo scelto 2, ovvero avessimo definito due numeri sono equivalenti se e solo se la loro differenza è un multiplo di 2, avremmo avuto due sole classi di equivalenza: la classe di 0, data da $\{0, \pm 2, \pm 4, \pm 6 \dots\}$, e la classe di 1, data da $\{\pm 1, \pm 3, \pm 5, \pm 7 \dots\}$. In altre parole, la partizione di $\mathbb{Z}$ determinata da questa relazione è quella in numeri pari e numeri dispari.

2.4 Relazioni d'ordine e grafi

Vediamo ora un'altra importantissima classe di relazioni, le *relazioni d'ordine*.

Definizione 2.16. Una relazione $\sim$ su un insieme X si dice **relazione d'ordine** se soddisfa le seguenti tre proprietà:

(1) $x \sim x$ per ogni $x \in X$ (ovvero $\sim$ è riflessiva),

(2) se $x \sim y$ e $y \sim x$, allora $x = y$ (ovvero $\sim$ è antisimmetrica) e

(3) se $x \sim y$ e $y \sim z$, allora $x \sim z$ (ovvero $\sim$ è transitiva).

Esempio 2.17. La relazione di "minore o uguale" (denotata con $\leqslant$) sull'insieme dei numeri naturali $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ è un esempio di relazione d'ordine, come è facile verificare.

Esempio 2.18. Un altro importante esempio di relazione d'ordine è dato dall'inclusione di sottoinsiemi. Più precisamente, si consideri un insieme X e l'insieme $P(X)$ dei suoi sottoinsiemi (cioè il cosiddetto **insieme potenza** di X). La relazione d'inclusione $A \subseteq B$, dove $A, B \in P(X)$ (cioè A e B sono sottoinsiemi di X) è una relazione d'ordine. Infatti,

(1) Si ha sempre $A \subseteq A$, in quanto tale inclusione, per definizione stessa di inclusione tra insiemi, significa che per ogni x vale $x \in A \Rightarrow x \in A$, che è un'implicazione sempre vera. Quindi vale la proprietà riflessiva.

(2) se $A \subseteq B$ e $B \subseteq A$, allora la prima inclusione ci dice che ogni elemento di A sta anche in B, e la seconda che viceversa ogni elemento di B sta anche in A, quindi per definizione di uguaglianza tra insiemi si ha $A = B$. Vale quindi anche la proprietà antisimmetrica.

(3) Infine, se $A \subseteq B$ e $B \subseteq C$, discende che $A \subseteq C$ (e quindi vale la proprietà transitiva). Infatti se $x \in A$, allora dalla prima inclusione $A \subseteq B$ deduciamo che $x \in B$ e, usando la seconda inclusione, deduciamo che $x \in C$. Avendo quindi dimostrato che $x \in A \Rightarrow x \in C$, concludiamo che $A \subseteq C$, come richiesto.

Questi due semplici esempi di relazioni d'ordine presentano alcune interessanti caratteristiche e significative differenze tra loro.

Ad esempio, osserviamo che in $\mathbb{N}$ con la relazione $\leqslant$ di minore o uguale, per due numeri $x, y \in \mathbb{N}$ vale sempre $x \leqslant y$ o $y \leqslant x$. Due elementi che soddisfano questa proprietà rispetto ad una relazione d'ordine si dicono *confrontabili*.

D'altro canto questo non si verifica sempre in $P(X)$ dotato della relazione $\subseteq$ di inclusione. Ad esempio se $X = \{a, b, c\}$, l'insieme dei sottoinsiemi di X è dato da

$$P(X) = \{\varnothing, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}\,.$$

In questo caso esistono elementi *non confrontabili*. Ad esempio, per $\{a, b\}$ e $\{a, c\}$ non si ha né $\{a, b\} \subseteq \{a, c\}$ né $\{a, c\} \subseteq \{a, b\}$.

Nelle definizioni che daremo in seguito, scegliamo per ragioni di convenienza didattica di indicare la generica relazione d'ordine sempre con il simbolo "$\leqslant$", anche se non si tratterà necessariamente della relazione usuale di minore o uguale tra numeri (quando si tratterà invece di quest'ultima lo specificheremo esplicitamente).

Definizione 2.19. Una relazione d'ordine $\leqslant$ su insieme X si dice **ordine totale** se due elementi di X sono sempre confrontabili rispetto a $\leqslant$ (ovvero se, dati comunque $x_1, x_2 \in X$, si ha sempre o $x_1 \leqslant x_2$ o $x_2 \leqslant x_1$). In tal caso, l'insieme X dotato della relazione $\leqslant$ si dice **totalmente ordinato**.

Quindi possiamo dire che $\mathbb{N}$, dotato della relazione d'ordine usuale $\leqslant$ di minore o uguale tra numeri, è un insieme totalmente ordinato, mentre l'insieme $P(X)$ dei sottoinsiemi di un insieme dato, dotato della relazione d'ordine $\subseteq$, non è totalmente ordinato.

Un'altra importante differenza tra le relazioni $\leqslant$ su $\mathbb{N}$ e $\subseteq$ su $P(X)$ consiste nell'esistenza di *massimi e minimi*.

Definizione 2.20. Data una relazione d'ordine $\leqslant$ su insieme X, un elemento $x_0 \in X$ si dice **massimo** se per ogni $x \in X$ si ha $x \leqslant x_0$. Analogamente, x_0 si dice **minimo** se per ogni $x \in X$ si ha $x_0 \leqslant x$

Ad esempio, $\mathbb{N}$ dotato della relazione usuale $\leqslant$ di minore o uguale tra numeri ammette minimo, cioè lo zero 0, in quanto per ogni $x \in \mathbb{N}$ si ha $0 \leqslant x$. D'altra parte $\mathbb{N}$ non ammette massimo perché non esiste nessun numero naturale x_0 tale che $x \leqslant x_0$ per ogni $x \in \mathbb{N}$ (non esiste il numero naturale più grande).

Invece, qualunque sia X, l'insieme dei suoi sottoinsiemi $P(X)$ dotato della relazione di inclusione ha sempre sia minimo (dato dall'insieme vuoto $\varnothing$, in quanto per ogni altro sottoinsieme A si X si ha $\varnothing \subseteq A$) che massimo (dato da X stesso, in quanto per ogni altro sottoinsieme A si X si ha $A \subseteq X$)

Esempio 2.21. Consideriamo l'insieme

$$P' = \{\ \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}\ \}$$

ottenuto dall'insieme $P(X)$ dei sottoinsiemi di $X = \{a, b, c\}$ eliminando l'insieme vuoto e X stesso, e consideriamo su P' sempre la relazione $\subseteq$ d'inclusione. Allora, si vede che non c'è nessun S in P' tale che $Z \subseteq S$ per ogni Z di P', ovvero non c'è nessun massimo. Analogamente, si vede che non c'è nessun S in P' tale che $S \subseteq Z$ per ogni Z di P', ovvero non c'è nessun minimo.
Tuttavia, l'elemento $S = \{b, c\}$ ha la notevole proprietà che *non esiste nessun Z in P' diverso da S per cui $S \subseteq Z$*. Quindi, benché questo elemento S non possa essere considerato un massimo in base alla Definizione 2.20 data sopra, non esiste nessun elemento che sia "maggiore di lui" (rispetto alla relazione d'inclusione). Potremmo quindi considerare S una sorta di massimo in un senso più largo.

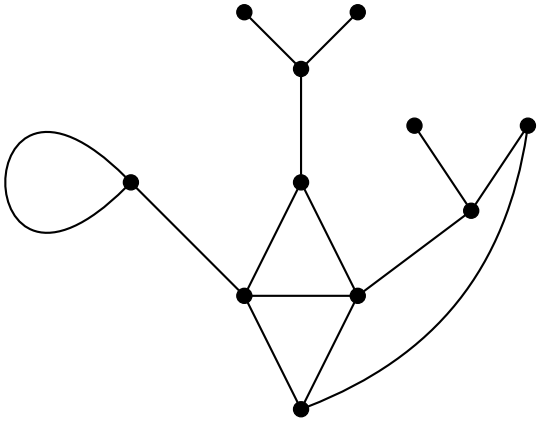
Definizione 2.22. Data una relazione d'ordine $\leqslant$ su insieme X, un elemento $x_0 \in X$ si dice **massimale** se non esiste nessun $y \in X$ diverso da x e tale che $x_0 \leqslant y$. Analogamente, un elemento $x_0 \in X$ si dice **minimale** se non esiste nessun $y \in X$ diverso da x e tale che $y \leqslant x_0$.

Equivalentemente, possiamo dire un elemento x_0 massimale se $x_0 \leqslant y$ implica $y = x_0$, o minimale se $y \leqslant x_0$ implica $y = x_0$.
Ad esempio, sempre nell'insieme $P' = \{ \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\} \}$ si ha che $\{a, b\}, \{a, c\}, \{b, c\}$ sono massimali, mentre $\{a\}, \{b\}, \{c\}$ sono minimali. Si noti quindi che in un insieme ordinato possiamo avere diversi elementi massimali e diversi elementi minimali (mentre, come si vede facilmente, se esiste un massimo o un minimo questo è unico).

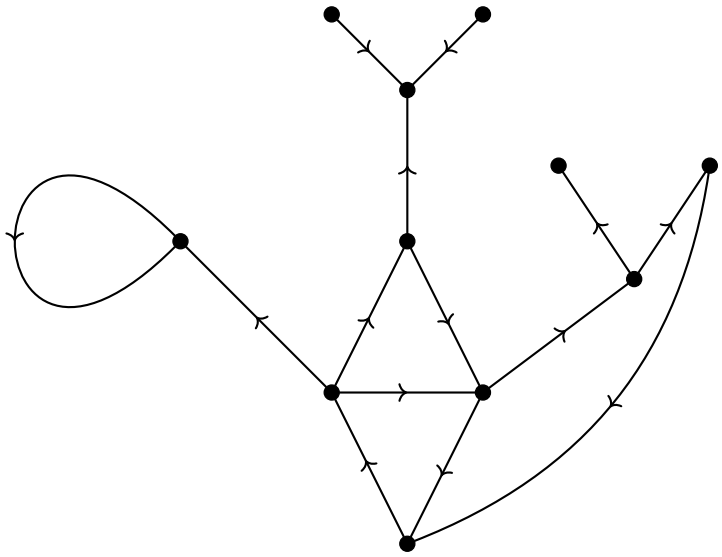
Osservazione 2.23. Si noti che la distinzione tra massimale e massimo (o tra minimale e minimo) ha senso se stiamo lavorando con un ordine non totale, cioè nel quale esistono elementi non confrontabili tra loro. Il motivo per cui $\{b, c\}$ è massimale ma non massimo in $P' = \{ \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\} \}$ rispetto alla relazione $\subseteq$ d'inclusione è che in P' ci sono elementi con i quali $\{b, c\}$ non è confrontabile, come ad esempio $\{a, c\}$. Quindi, pur non esistendo nessun elemento "maggiore" di $\{b, c\}$ non possiamo dire che $\{b, c\}$ sia maggiore di tutti gli altri, sicuramente non di quelli con i quali non è confrontabile.
Al contrario, in un ordine totale in cui dati due elementi x e y si ha sempre o $x \leqslant y$ o $y \leqslant x$, un massimale x_0 è automaticamente anche il massimo. Questo perché se per nessun y si ha $x_0 \leqslant y$ (come prevede la definizione di massimale), dovendo valere comunque una delle due disuguaglianze allora si ha necessariamente $y \leqslant x_0$ per ogni y, e quindi x_0 è in effetti un massimo.

Un modo efficace per visualizzare un insieme ordinato e l'eventuale presenza di massimali, minimali, massimi o minimi è quello di rappresentarlo mediante un *grafo*.

Definizione 2.24. Un **grafo** è un insieme di punti, detti *vertici*, e di segmenti di curva, detti *lati*, che collegano coppie di vertici.[4]



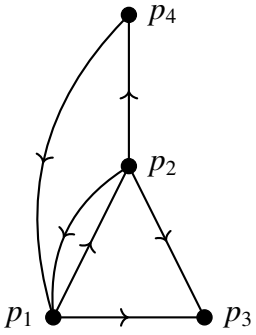
Si noti che un grafo può presentare **cappi**, ovvero segmenti che collegano un vertice con se stesso.
Se i lati del grafo sono dotati di un verso di percorrenza (simboleggiato da una freccia) si parla di **grafo orientato.**



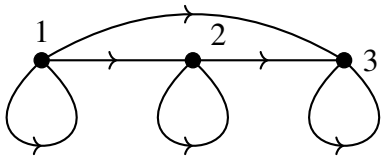
Una qualunque relazione $\sim$ su un insieme X può essere rappresentata mediante un grafo orientato nel modo seguente. Si rappresentano gli elementi di X mediante punti (i vertici del grafo) e si traccia un segmento orientato dal punto che rappresenta x al punto che rappresenta x' se e solamente se $x \sim x'$, ovvero x è in relazione con x'. Si osservi che

[4]Questa non è una definizione rigorosa di grafo ma è sufficiente ai nostri scopi. In generale un grafo è definito come una coppia di insiemi: quello dei vertici e quello dei lati.

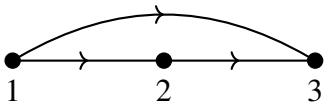
orientare il segmento è necessario per capire se $x \sim x'$ o $x' \sim x$ per cui una relazione d'ordine definisce sempre un grafo orientato. Ad esempio, supponiamo di avere un insieme $X = \{p_1, p_2, p_3, p_4\}$ di pagine internet. Vogliamo definire la relazione: $p \sim p'$ se e solo se esiste un link dalla pagina p alla pagina p'. Se ad esempio da p_1 abbiamo un link verso p_2 e uno verso p_3, da p_2 un link verso p_1, uno verso p_3 e uno verso p_4, da p_3 nessun link e da p_4 solo un link verso p_1, allora il grafo orientato che rappresenta la relazione corrispondente è



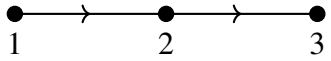
Parlando in particolare di relazioni d'ordine, se consideriamo ad esempio su $X = \{1, 2, 3\}$ l'usuale relazione d'ordine $\leqslant$ di minore o uguale, allora il grafo corrispondente è



Si noti che da ogni vertice parte un cappio, in quanto ogni elemento x è in relazione $x \leqslant x$ con se stesso (le relazione è riflessiva). Se, sapendo che la relazione è d'ordine e quindi riflessiva, sottointendiamo la presenza di questi cappi, potremmo semplificare il grafo nel modo seguente



Un'ulteriore semplificazione può essere fatta osservando che, essendo la relazione d'ordine transitiva, la presenza del segmento orientato da 1 a 3 può essere dedotta dalla presenza del segmenti orientati da 1 a 2 e da 2 a 3 (infatti, per la transitività $1 \leqslant 2$ e $2 \leqslant 3$ implica $1 \leqslant 3$). Sottointendendo quindi anche i segmenti che sono conseguenza della transitività, il grafico diventa

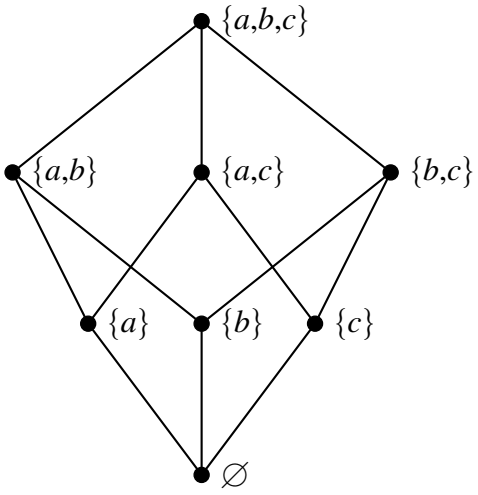


Infine, possiamo pensare di evitare di orientare i segmenti disponendo il grafico in modo che i versi degli stessi vadano sempre pensati dal basso verso l'alto:

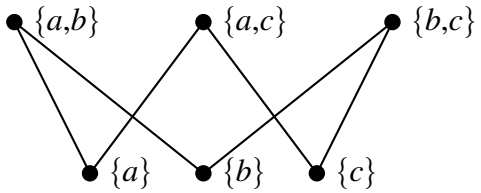
Il grafo di una relazione d'ordine così semplificato prende anche il nome di *diagramma di Hasse*.
Ad esempio, consideriamo di nuovo l'insieme

$$P(X) = \{\varnothing, \{a\}, \{b\}, \{c\}, \{a,b\}, \{a,c\}, \{b,c\}, \{a,b,c\}\}$$

dei sottoinsiemi di $X = \{a,b,c\}$ con la relazione d'ordine data dall'inclusione $\subseteq$. Allora, il suo diagramma di Hasse è



Come si vede, l'insieme ordinato possiede un massimo (il vertice più alto del diagramma) e un minimo (il vertice più basso del diagramma). Se ci limitassimo alla relazione d'inclusione sui sottoinsiemi $\{a\}, \{b\}, \{c\}, \{a,b\}, \{a,c\}, \{b,c\}$, il diagramma sarebbe invece



nel quale non esiste un solo vertice più alto o più basso, ma più vertici nel livello superiore e in quello inferiore (che rappresentano rispettivamente gli elementi massimali o minimali).

Osservazione 2.25. L'estrema importanza dei grafi in matematica e nelle sue applicazioni è giustificata dall fatto che un grafo può modellizzare numerose situazioni e permettere di visualizzare geometricamente una grande varietà di problemi. Una rete di trasporti (ad esempio la cartina delle strade di una città con i suoi incroci) può essere rappresentata da un grafo ordinato, un albero genealogico è un grafo, un circuito stampato può essere pensato come un grafo e molti altri.
Queste situazioni concrete hanno ispirato una serie di problemi e domande relativi ai grafi, quali: dato un grafo, è possibile percorrerlo tramite un cammino continuo che tocchi tutti i suoi vertici una sola volta ? (si parla di *cammino hamiltoniano*); o ancora, è possibile percorrerlo tramite un cammino continuo che tocchi tutti i suoi lati una sola volta ? (si parla di *cammino euleriano*). Questi e tanti altri problemi (alcuni risolti, altri ancora aperti) sono affrontati da una branca della matematica detta appunto teoria dei grafi.

2.5 Funzioni

Vedremo ora un'altra nozione fondamentale per tutta la matematica, quella di *funzione*.

Definizione 2.26. Dati due insiemi X e Y, una **funzione** f da X a Y (si scrive $f\colon X \longrightarrow Y$) è una legge che assegna ad ogni elemento x di X un unico elemento y di Y. L'elemento y è detto **immagine** di x e si denota con $f(x)$. Per indicare che f associa $f(x)$ ad x si scrive anche $x \mapsto f(x)$. Il tutto è espresso dalla scrittura compatta

$$\begin{aligned} f\colon X &\longrightarrow Y \\ x &\mapsto f(x) \end{aligned}$$

L'insieme X di partenza si dice **dominio** della funzione, quello Y di arrivo **codominio**.

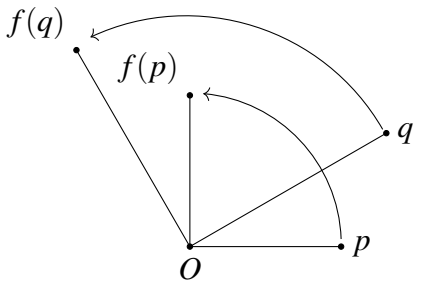
Esempi 2.27. Sia X l'insieme degli studenti presenti in un'aula e sia Y l'insieme $\mathbb{N}$ dei numeri naturali. Associando a ogni studente x il suo numero di matricola, si ottiene una funzione $X \longrightarrow \mathbb{N}$.

In questo esempio, la funzione rappresenta l'assegnazione di un dato a ogni entrata di una lista.

Le funzioni possono essere anche usate per descrivere l'andamento di un dato fenomeno. Si consideri ad esempio un corpo che a un certo istante di tempo parte da un punto O e inizia a muoversi lungo una traiettoria rettilinea. Identificando i punti della retta con i numeri reali con il punto O in corrispondenza dello zero, come descritto nell'Osservazione 2.3, possiamo assegnare a ogni $t \in \mathbb{R}$ la posizione $x(t)$ del corpo sulla retta lungo cui si muove dopo un tempo t dall'istante iniziale (se t è positivo, altrimenti sia il tempo precedente all'istante iniziale). Questo definisce una funzione $\mathbb{R} \longrightarrow \mathbb{R}$ che descrive il moto del corpo.

Ancora, le funzioni possono rappresentare delle operazioni. Ad esempio la funzione $f\colon \mathbb{Z} \longrightarrow \mathbb{Z}$ che assegna a ogni numero intero n il suo quadrato $f(n) = n^2$. Non otteniamo invece una funzione $\mathbb{Z} \longrightarrow \mathbb{Z}$ definendo $f(n) = \sqrt{n}$, perché alcuni elementi del dominio non hanno immagine nel codominio,. Infatti se $n = 2$ o $n = -1$, $f(n) = \sqrt{2}$ e $f(n) = \sqrt{-1}$ non sono numeri interi (il secondo non è neanche un numero reale).

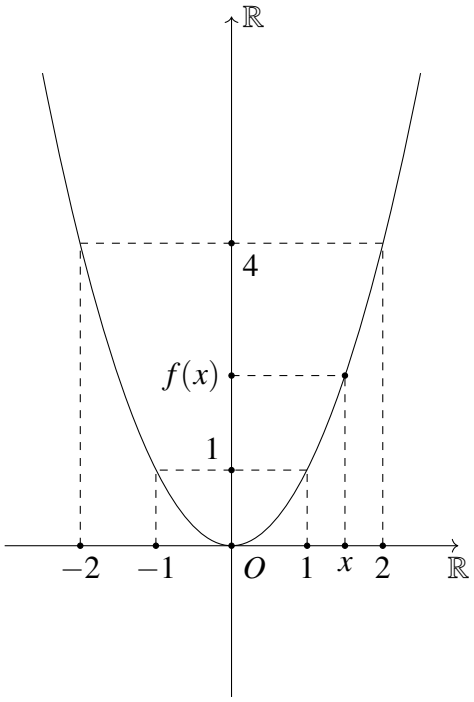
Infine, come ultimo esempio, le funzioni possono rappresentare trasformazioni di tipo geometrico. Se X è l'insieme dei punti del piano, scelto un punto O possiamo definire una funzione $X \longrightarrow X$ assegnando a ogni punto P il punto $f(P)$ che si ottiene ruotando P attorno a O di un angolo di 90 gradi in senso antiorario.



Osservazione 2.28. La definizione di funzione data sopra è sufficiente per gli scopi di questo corso ma non è la definizione rigorosa e formale di funzione, che afferma che una funzione da un insieme X a un insieme Y è un sottoinsieme S del prodotto cartesiano $X \times Y$ che per ogni $x \in X$ contiene una e una sola coppia (x, y) che ha x come prima componente (in simboli[5], $\forall x \in X\ \exists! y \in Y\ (x, y) \in S$)
In pratica, l'idea è di rappresentare l'informazione che a x stiamo assegnando un certo $y \in Y$ tramite la coppia (x, y). La richiesta che per ogni x esista un'unica coppia che ha x come primo elemento traduce il fatto che a ogni x dobbiamo assegnare un ben preciso y (uno e uno solo).

[5]Il quantificatore universale $\exists$, seguito dal punto esclamativo, significa "esiste un unico".

L'idea che una funzione $f\colon X \longrightarrow Y$ sia definita da un sottoinsieme del prodotto cartesiano $X \times Y$ non è tanto sorprendente se si pensa che ogni funzione $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ può essere rappresentata mediante il suo grafico, che può essere pensato proprio come un sottoinsieme del prodotto cartesiano $\mathbb{R} \times \mathbb{R}$. Più precisamente, se come già descritto nell'Osservazione 2.3 consideriamo un sistema di coordinate per cui a ogni punto del piano corrisponde una coppia (x,y) di numeri reali e viceversa (e quindi il piano si identifica con il prodotto cartesiano $\mathbb{R} \times \mathbb{R}$), possiamo rappresentare una funzione $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ disegnando l'insieme dei punti che hanno come prima coordinata x e come seconda coordinata l'immagine $f(x)$ (per ogni $x \in \mathbb{R}$).
Ad esempio, se la funzione fosse $f(x) = x^2$ si avrebbe



L'insieme di punti così ottenuto (nel disegno dato la parabola) è quindi un sottoinsieme del piano, che è come dire un sottoinsieme del prodotto cartesiano $\mathbb{R}\times\mathbb{R}$, che rappresenta la funzione.

Consideriamo ora alcune proprietà delle funzioni.
Iniziamo con l'osservare che la definizione di funzione non impedisce che due elementi diversi del dominio abbiano la stessa immagine. Ad esempio, la funzione quadrato $f\colon \mathbb{Z} \longrightarrow \mathbb{Z}$ considerata sopra che assegna a ogni $n \in \mathbb{Z}$ l'intero $f(n) = n^2$ è tale che $f(2) = f(-2) = 4$.

Questo non accade ad esempio per la funzione $X \longrightarrow \mathbb{N}$ che assegna a ogni studente di una certa aula il suo numero di matricola, in quanto non è possibile che studenti diversi abbiano lo stesso numero di matricola.
Alla luce di questi due esempi, diamo la seguente

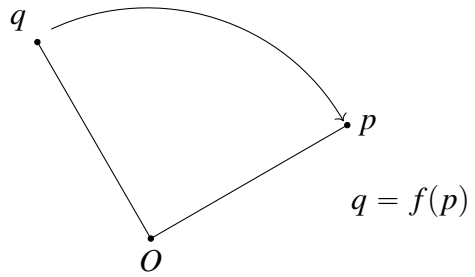
Definizione 2.29. Una funzione $f\colon X \longrightarrow Y$ si dice **iniettiva** se due diversi elementi del dominio hanno sempre immagini diverse. In simboli

$$x \neq x' \Rightarrow f(x) \neq f(x')\,.$$

Quindi, la funzione che assegna a ogni studente il suo numero di matricola è iniettiva, mentre la funzione quadrato da $\mathbb{Z}$ in $\mathbb{Z}$ non è iniettiva.

Osservazione 2.30. Usando il fatto che ogni implicazione è equivalente alla sua contronominale, possiamo anche dire che una funzione f è iniettiva se per ogni $x, x' \in X$ si ha $f(x) = f(x') \Rightarrow x = x'$ (ovvero se l'unica possibilità per cui x e x' abbiano la stessa immagine sia che $x = x'$).
Tale riformulazione è spesso più conveniente quando si vuole dimostrare che una funzione data è iniettiva. Se volessi verificare ad esempio che la funzione $f\colon \mathbb{N} \longrightarrow \mathbb{N}$ data da $f(n) = n + 1$ è iniettiva, dovrei dimostrare l'implicazione $f(n) = f(m) \Rightarrow n = m$, ovvero $n+1 = m+1 \Rightarrow n = m$. Ma questo è quasi immediato in quanto da $n+1 = m+1$, portando l'1 del primo membro a secondo si ottiene subito $n = m + 1 - 1 = m$.

Un'altra particolarità della funzione $X \longrightarrow \mathbb{N}$ che assegna a ogni studente il suo numero di matricola e della funzione quadrato $\mathbb{Z} \longrightarrow \mathbb{Z}$ è che non tutti gli elementi del codominio sono immagine di qualche elemento del dominio. Cioè non tutti i numeri naturali sono numeri di matricola di qualche studente, esattamente come non tutti gli interi sono quadrati di qualche intero di partenza (ad esempio, 5 non è quadrato, ovvero non è della forma n^2 per nessun $n \in \mathbb{Z}$).
Al contrario, questo non accade, ad esempio, per la rotazione che abbiamo definito nell'Esempio 2.27, cioè la funzione $f\colon X \longrightarrow X$ dal piano X in se stesso e che manda ogni punto P nel punto $f(P)$ che si ottiene ruotando P di 90 gradi in senso antiorario. Infatti qualunque punto Q del codominio X è immagine $f(P)$ di qualche punto P del dominio (basta prendere come P il punto che si ottiene ruotando Q in senso *orario*, come si vede nel disegno seguente)



Diamo allora la seguente

Definizione 2.31. Una funzione $f\colon X \longrightarrow Y$ si dice **suriettiva** se per ogni $y \in Y$ esiste un $x \in X$ tale che $y = f(x)$.

Quindi, sia la funzione $X \longrightarrow \mathbb{N}$ che assegna a ogni studente il suo numero di matricola che la funzione quadrato su $\mathbb{Z}$ sono non suriettive, mentre la rotazione del piano in sè è suriettiva.
In effetti, quest'ultima è anche iniettiva. Tali funzioni sono particolarmente importanti.

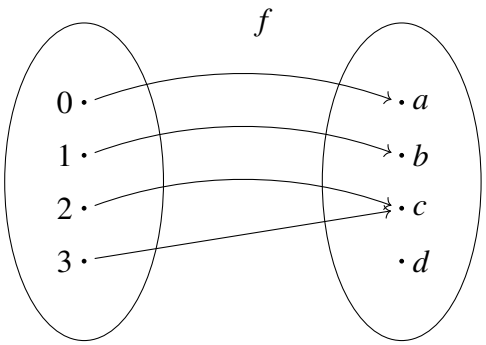
Definizione 2.32. Una funzione $f\colon X \longrightarrow Y$ si dice **biiettiva** se è sia iniettiva che suriettiva.

Una funzione biiettiva si chiama anche *biiezione* o *corrispondenza biunivoca*.

Un modo alternativo di definire l'iniettività è usare il concetto di *controimmagine*. Data una funzione $f\colon X \longrightarrow Y$ e un elemento $y \in Y$, si dice controimmagine di y (e si denota con $f^{-1}(y)$) l'insieme di tutti gli elementi di X a cui F associa y. In simboli

$$f^{-1}(y) = \{x \in X \mid f(x) = y\}\,.$$

Ad esempio, si consideri la funzione $f\colon \{0,1,2,3\} \longrightarrow \{a,b,c,d\}$ rappresentata nel seguente disegno



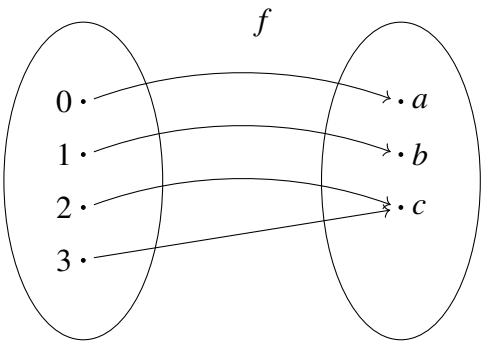
Si ha $f^{-1}(a) = \{0\}$, $f^{-1}(b) = \{1\}$, $f^{-1}(c) = \{2,3\}$, $f^{-1}(d) = \varnothing$.
Come in questo esempio la controimmagine di un elemento di Y potrebbe essere vuota. *In effetti una funzione è suriettiva se e solo se non esistono elementi $y \in Y$ tali che nessun $x \in X$ soddisfa $f(x) = y$.* Chiaramente, una funzione è iniettiva se non succede mai che $f^{-1}(y)$ abbia più di un elemento, perchè in tal caso vorrebbe dire che ci sono elementi diversi di X che vanno a finire in y. In altre parole, possiamo dire che *una funzione $f\colon X \longrightarrow Y$ è iniettiva se e solo se tutte le controimmagini degli elementi di Y hanno al massimo un elemento.*

Anche la suriettività ammette una definizione alternativa, basata sul concetto di *immagine della funzione*. L'immagine di una funzione $f\colon X \longrightarrow Y$ (denotata $Im(f)$) è semplicemente l'insieme di tutte le immagini degli elementi di X, in simboli

$$Im(f) = \{y \in Y \mid y = f(x) \text{ per qualche } x \in X\}$$

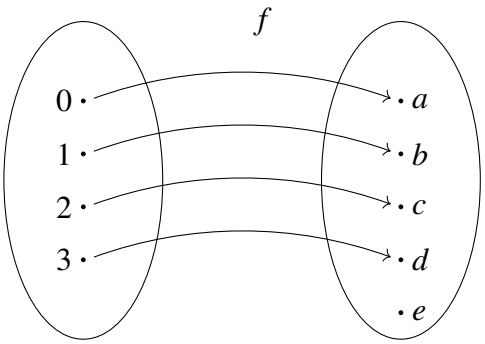
Ad esempio, nella funzione $f\colon \{0,1,2,3\} \longrightarrow \{a,b,c,d\}$ vista sopra si ha $Im(f) = \{a,b,c\}$ (solo d non si scrive come immagine di qualche elemento del dominio). Allora, è chiaro che una funzione è suriettiva solo se $Im(f) = Y$, in quanto questa uguaglianza si verifica esattamente quando ogni elemento $y \in Y$ si scrive come $y = f(x)$ per qualche $x \in X$, che è proprio la definizione di suriettività.

Osserviamo che iniettività e suriettività sono nozioni indipendenti: una funzione può essere iniettiva e non suriettiva, suriettiva e non iniettiva, nessuna delle due o entrambe. Ad esempio, la funzione $f\colon \{0,1,2,3\} \longrightarrow \{a,b,c\}$ rappresentata dal seguente disegno



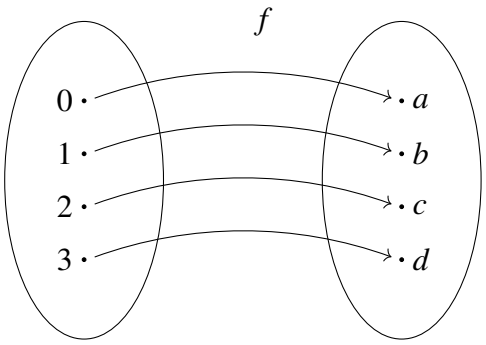
(ovvero $f(0) = a$, $f(1) = b$, $f(2) = c$, $f(3) = c$) è suriettiva (in quanto ogni elemento a,b,c del codominio è immagine di qualche elemento del dominio) ma non iniettiva in quanto esistono due elementi diversi del dominio (ossia 2 e 3) che hanno la stessa immagine.

Invece, la funzione $f\colon \{0,1,2,3\} \longrightarrow \{a,b,c,d,e\}$ rappresentata dal seguente disegno



è iniettiva ma non suriettiva (c'è un elemento del codomino, ovvero e, che non è immagine di nessun elemento del dominio).

Infine, la funzione $f\colon \{0,1,2,3\} \longrightarrow \{a,b,c,d\}$ rappresentata dal seguente disegno



è sia iniettiva che suriettiva, ovvero biiettiva.

Osserviamo che non c'è nessuna possibile scelta delle immagini di $0,1,2,3$ nella prima delle tre funzioni precedenti che la renda una funzione iniettiva. Infatti, essendoci più elementi nel dominio che nel codominio, almeno due elementi dovranno avere per forza la stessa immagine. In generale, dati due insiemi finiti X e Y, il primo con $|X|$ elementi e il secondo con $|Y|$ elementi, abbiamo che

$$|X| > |Y| \Rightarrow f\colon X \longrightarrow Y \text{ non è iniettiva} \tag{2.6}$$

o equivalentemente, usando la contronominale,

$$f\colon X \longrightarrow Y \text{ è iniettiva} \Rightarrow |X| \leqslant |Y|\,. \tag{2.7}$$

Osservazione 2.33. La (2.6) si chiama anche *principio della piccionaia* per il seguente motivo. Supponiamo che X sia un insieme di piccioni e Y un insieme di gabbie. Se dobbiamo mettere ogni piccione in una gabbia e ci sono più piccioni che gabbie almeno due piccioni dovranno essere messi nella stessa gabbia. A dispetto della sua elementarità può essere usato per dimostrare fatti a prima vista non banali.

Ad esempio, usando tale principio possiamo dimostrare che nel mondo esistono almeno due persone con lo stesso numero di capelli. Infatti, sia X l'insieme delle persone del mondo, che quindi diciamo contiene 8 miliardi di elementi, e consideriamo la funzione f che assegna a ogni persona il numero dei suoi capelli. Stimando che il numero massimo N di capelli che possono essere contenuti in una testa sia sicuramente minore di 8 miliardi, la funzione f sarà una funzione $X \longrightarrow Y = \{0,1,2,3,\ldots,N\}$ tale che $|X| > |Y|$. In base al principio enunciato f non può essere iniettiva, ovvero esistono almeno due elementi del dominio (due persone) con la stessa immagine (lo stesso numero di capelli).

Analogamente, per quanto riguarda la seconda delle tre funzioni rappresentate nei disegni sopra, ovvero la $f\colon \{0,1,2,3\} \longrightarrow \{a,b,c,d,e\}$, non c'è nessuna possibile scelta alternativa delle immagini che la renda una funzione suriettiva. Infatti, essendoci nel dominio meno elementi che nel codominio, non è possibile raggiungere tutti gli elementi del codominio facendo partire una freccia da ciascun elemento del dominio. Analogamente a sopra, possiamo dire che in generale dati due insiemi finiti X e Y, il primo con $|X|$ elementi e il secondo con $|Y|$ elementi, abbiamo che

$$|X| < |Y| \Rightarrow f\colon X \longrightarrow Y \text{ non è suriettiva} \tag{2.8}$$

o equivalentemente, usando la contronominale,

$$f\colon X \longrightarrow Y \text{ è suriettiva} \Rightarrow |X| \geqslant |Y|\,. \tag{2.9}$$

Dal momento che le funzioni biiettive sono sia iniettive che suriettive, combinando la (2.7) e la (2.9) otteniamo

$$f\colon X \longrightarrow Y \text{ è biiettiva} \Rightarrow |X| = |Y|\,. \tag{2.10}$$

Osservazione 2.34. La (2.10) ci dice che se esiste una funzione biiettiva tra due insiemi *finiti*, allora essi hanno lo stesso numero di elementi. Non è difficile verificare che vale il viceversa: se il numero di elementi di X e Y è lo stesso, allora non è difficile costruire una funzione biiettiva $X \longrightarrow Y$.

Nel caso in cui X e Y siano insiemi infiniti, parlare di numero di elementi sembra non avere più senso, o al più verrebbe spontaneo dire che due insiemi infiniti hanno lo stesso numero di elementi (infinito, appunto). Tuttavia, se decidiamo, procedendo per analogia con gli insiemi finiti, di dire anche per due insiemi infiniti X e Y che essi hanno lo stesso numero di elementi se esiste una funzione biiettiva $X \longrightarrow Y$, allora scopriremo che *non tutti gli insiemi infiniti hanno lo stesso numero di elementi* (detto anche *cardinalità*).

Per illustrare tale affermazione, usiamo come esempi gli insiemi numerici $\mathbb{N}$ (i numeri naturali), $\mathbb{Z}$ (i numeri interi), $\mathbb{Q}$ (i numeri razionali), $\mathbb{R}$ (i numeri reali)[6], che sono inclusi uno nel successivo:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}\,.$$

Iniziamo con il chiederci se esiste una funzione biiettiva $\mathbb{N} \longrightarrow \mathbb{Z}$. Questo significherebbe che è possibile associare a ogni naturale un intero in modo da non lasciare scoperto nessun intero e da non ripetere due volte lo stesso intero, o in altre parole mettere tutti gli interi in una successione infinita $a_0, a_1, a_2 \ldots$ senza ripetizioni.

[6]Pur non avendo definito formalmente questi insiemi numerici, possiamo portare avanti questa discussione ricordandoci che non ha la pretesa di essere rigorosa. L'insieme dei numeri naturali sarà definito rigorosamente nella Sezione 2.7.

La risposta è affermativa, ad esempio gli interi possono essere messi in sequenza come segue:

$$\begin{array}{cccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & \dots \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \dots \\ 0 & 1 & -1 & 2 & -2 & 3 & -3 & \dots \end{array}$$

Dunque esiste una corrispondenza biunivoca tra $\mathbb{N}$ e $\mathbb{Z}$, ovvero possiamo dire che i naturali e gli interi hanno "lo stesso numero di elementi" o, più precisamente, la stessa cardinalità.
Mentre da una parte questo può risultare banale (si potrebbe pensare che avendo entrambi infiniti elementi, necessariamente questi debbano poter essere messi in corrispondenza biunivoca), dall'altra il risultato ha un aspetto sorprendente. Specificatamente, dal momento che $\mathbb{N}$ si identifica con un sottoinsieme di $\mathbb{Z}$ (quello degli interi non negativi), stiamo dicendo che $\mathbb{Z}$ ha lo stesso numero di elementi di un suo sottoinsieme proprio.[7] Questo non è possibile per insiemi finiti: in effetti questa proprietà caratterizza gli insiemi infiniti ed è spesso assunto come definizione di insieme infinito. Ovvero, un insieme si dice infinito se può essere messo in corrispondenza biunivoca con un suo sottoinsieme proprio.
Inoltre possiamo mostrare la meno intuitiva esistenza di una biiezione tra i razionali $\mathbb{Q}$ e i naturali $\mathbb{N}$. Per mostrare come, consideriamo per facilità solo i razionali positivi (il ragionamento si modifica successivamente per accomodare tutti i razionali), che si scrivono come quoziente n/m di due naturali n, m (con $m \neq 0$).
Consideriamo la seguente tabella infinita di frazioni dove, nella riga n-esima, scriviamo le frazioni con numeratore uguale a n:

$$\begin{array}{c|ccccc|c} & 1 & 2 & 3 & 4 & 5 & \cdots \\ \hline 1 & 1/1 & 1/2 & 1/3 & 1/4 & 1/5 & \dots \\ 2 & 2/1 & 2/2 & 2/3 & 2/4 & 2/5 & \dots \\ 3 & 3/1 & 3/2 & 3/3 & 3/4 & 3/5 & \dots \\ 4 & 4/1 & 4/2 & 4/3 & 4/4 & 4/5 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Chiaramente in questa tabella alcune frazioni rappresentano lo stesso razionale (1/2 e 2/4, oppure 1/1 e 2/2)) ma se riusciamo a mostrare che è possibile realizzare una corrispondenza biunivoca tra le entrate della tabella e i naturali, a maggior ragione questo sarà vero per i razionali che corrispondono alle entrate della tabella senza ripetizioni.
La corrispondenza cercata con le frazioni della tabella si può per esempio realizzare scrivendo in sequenza tutte le diagonali della tabella a partire dall'angolo in alto a sini-

[7] Si dice sottoinsieme proprio di un insieme X un sottoinsieme che non sia X stesso (in base alla definizione di sottoinsieme, vale sempre l'inclusione $X \subseteq X$, ovvero ogni insieme è sottoinsieme di se stesso).

stra: prima $1/1$, poi la diagonale adiacente $1/2, 2/1$, poi continuando a scendere verso destra la terza diagonale $1/3, 2/2, 3/1$ e così via.

$$1/1, 1/2, 2/1, 1/3, 2/2, 3/1, 1/4, 2/3, 3/2, 4/1, \ldots$$

In questo modo, si riesce a mettere in una sequenza (ovvero in corrispondenza con i numeri naturali) tutti i razionali, e realizzare una funzione biiettiva $\mathbb{N} \longrightarrow \mathbb{Q}$. Possiamo concludere che anche $\mathbb{Q}$ ha la stessa cardinalità di $\mathbb{N}$.

Questi esempi potrebbero generare l'idea che qualunque insieme infinito possa essere messo in corrispondenza biunivoca con l'insieme dei naturali. In realtà questo è falso: si dimostra che per l'insieme $\mathbb{R}$ dei numeri reali questa corrispondenza non esiste.
Un'idea della dimostrazione è la seguente: supponiamo per assurdo di poter mettere in una sequenza in corrispondenza biunivoca con i numeri naturali tutti i numeri reali: $x_0, x_1, x_2, x_3, \ldots$.
Rappresentando ogni reale con la sua rappresentazione decimale, possiamo riscrivere tale sequenza come

$$\begin{aligned} x_0 &= a_0, b_0 c_0 d_0 \ldots \\ x_1 &= a_1, b_1 c_1 d_1 \ldots \\ x_2 &= a_2, b_2 c_2 d_2 \ldots \\ x_3 &= a_3, b_3 c_3 d_3 \ldots \\ &\ldots \end{aligned}$$

Ora, mostreremo che in realtà una tale sequenza non può mai contenere tutti i numeri reali, e lo faremo costruendo esplicitamente un numero reale che non è contenuto nella sequenza, il numero

$$x = 0, bcd \ldots$$

in cui $b, c, d, \ldots$ sono definiti come segue. Poniamo $b = 1$ se $b_0 = 0$ e $b = 0$ se $b_0 \neq 0$ (questo garantisce già che x non sia il primo numero della sequenza x_0). Successivamente definiamo $c = 1$ se $c_1 = 0$ e $c = 0$ se $c_1 \neq 0$ (questo garantisce che x non sia il secondo numero della sequenza x_1). Ancora poniamo $d = 1$ se $d_2 = 0$ e $d = 0$ se $d_2 \neq 0$ (questo garantisce che x non sia il terzo numero della sequenza x_2). Ripetendo questo ragionamento per ogni decimale otterremo il numero desiderato.

Dunque, $\mathbb{R}$ non ha lo stesso numero di elementi di $\mathbb{N}$ (pur essendo entrambi gli insiemi infiniti), e poichè $\mathbb{R}$ contiene $\mathbb{N}$ possiamo dire che $\mathbb{R}$ ha una cardinalità strettamente maggiore di quella di $\mathbb{N}$.
Non esiste dunque un unico "infinito", ma infiniti di tipo diverso. È lecito chiedersi se $\mathbb{R}$ sia il maggiore infinito possibile. La risposta è no. Infatti, dato un qualunque insieme

infinito, ne esiste sicuramente sempre uno di cardinalità più grande. Più precisamente, Cantor (al quale si deve la teoria degli insiemi infiniti che stiamo illustrando) mostrò che dato un insieme X, l'insieme potenza $P(X)$ di X (cf. Esempio 2.18) ha sempre cardinalità maggiore di X. Quindi, ad esempio, se vogliamo un insieme che abbia cardinalità maggiore di $\mathbb{R}$, basta prendere l'insieme potenza $P(\mathbb{R})$ di $\mathbb{R}$.

Concludiamo dicendo che si dimostra che la cardinalità di $\mathbb{R}$ coincide con quella dell'insieme potenza di $\mathbb{N}$.

L'affermazione che non esistono insiemi di cardinalità compresa tra quella di $\mathbb{N}$ e quella di $\mathbb{R}$ si chiama *ipotesi del continuo*. Si può dimostrare che tale affermazione non è né dimostrabile né confutabile nella teoria assiomatica degli insiemi.

Teorema 2.35. *Dato un qualunque insieme X, il suo insieme potenza $P(X)$ ha cardinalità strettamente maggiore di quella di X.*

Dimostrazione. Dobbiamo mostrare che esiste una funzione iniettiva $X \longrightarrow P(X)$ ma non esiste una funzione biiettiva $X \longrightarrow P(X)$. Per quello che riguarda la prima affermazione, un esempio di funzione iniettiva da X a $P(X)$ è dato dalla funzione che manda ogni $x \in X$ nel sottoinsieme $\{x\}$ che contiene solamente x. L'iniettività di tale funzione è immediata in quanto è chiaro che se $x \neq x'$ allora i due sottoinsiemi $\{x\}$ e $\{x'\}$ sono necessariamente diversi.

Dimostriamo ora la seconda affermazione per assurdo, ovvero supponiamo che esista una funzione $f\colon X \longrightarrow P(X)$ biiettiva e mostriamo che questo porta a una contraddizione. Tale funzione associa a ogni $x \in X$ un sottoinsieme $f(x)$ di X. Questo sottoinsieme $f(x)$ potrebbe contenere x o meno. Consideriamo allora il sottoinsieme

$$A = \{x \in X \mid x \notin f(x)\}$$

costituito da tutti gli elementi di X che non appartengono al sottoinsieme $f(x)$ che viene loro associato mediante la f.

Dal momento che A è un sottoinsieme di X, ovvero appartiene a $P(X)$, e la funzione $f\colon X \longrightarrow P(X)$ è biiettiva e in particolare suriettiva, deve esistere un elemento $a \in X$ che ha come immagine proprio A, ovvero $f(a) = A$.

Ma a questo punto otteniamo la contraddizione cercata chiedendoci se a appartiene a A. Infatti, se $a \in A$, allora essendo A l'insieme degli elementi x caratterizzati dalla proprietà $x \notin f(x)$, sarebbe $a \notin f(a)$, che essendo $f(a) = A$ significa $a \notin A$. Viceversa, se $a \notin A$ questo significa che $a \notin f(a)$ e quindi a gode della proprietà $x \notin f(x)$ che definisce gli elementi di A ovvero $a \in A$. Vediamo quindi che $a \in A$ se e solo se $a \notin A$.

Deduciamo quindi che l'ipotesi che $f\colon X \longrightarrow P(X)$ fosse biiettiva era assurda, ovvero non esiste nessuna funzione biiettiva da X a $P(X)$. □

2.6 Composizione di funzioni e funzioni invertibili

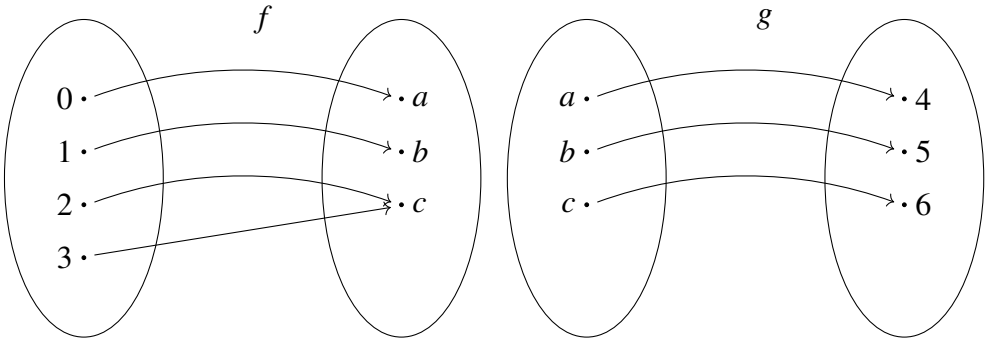
Cosí come negli insiemi numerici esistono operazioni (somma e prodotto) che dati due numeri ce ne danno un terzo, e nell'algebra di Boole esistono operazioni (le congiunzioni "e" e "o") che date due proposizioni ne danno una terza, anche per le funzioni, sotto ipotesi opportune, esiste un'operazione che produce una funzione a partire da due funzioni date: tale operazione si chiama **composizione**.

Più precisamente, supponiamo di avere due funzioni $f\colon X \longrightarrow Y$ e $g\colon Y \longrightarrow Z$, tali che *il codominio della prima coincida con il dominio della seconda*. Allora, per ogni elemento $x \in X$, possiamo applicare prima f ottenendo $f(x) \in Y$, e poi dal momento che Y è anche il dominio della g possiamo applicare la g a $f(x)$, ottenendo $g(f(x))$. In questo modo otteniamo una nuova funzione che associa a ogni elemento di X un elemento di Z:

$$\begin{aligned} g \circ f\colon X &\longrightarrow Y \\ x &\mapsto g(f(x)) \end{aligned}$$

A causa dell'ordine in cui appaiono le funzioni in $g(f(x))$, questa nuova funzione si denota con $g \circ f$ (che si legge "f composto g" a causa dell'ordine in cui vengono applicate).

Esempio 2.36. Consideriamo le funzioni $f\colon \{0,1,2,3\} \longrightarrow \{a,b,c\}$ e $g\colon \{a,b,c\} \longrightarrow \{4,5,6\}$ rappresentate nel seguente disegno



Essendo il codominio di f uguale al dominio di g, si può costruire la composizione $g \circ f$ applicando prima f e poi g ad ogni elemento di $\{0,1,2,3\}$. Otteniamo quindi $g \circ f\colon \{0,1,2,3\} \longrightarrow \{a,b,c\}$,

$$\begin{aligned} (g \circ f)(0) &= g(f(0)) = g(a) = 4 \\ (g \circ f)(1) &= g(f(1)) = g(b) = 5 \end{aligned}$$

$$(g \circ f)(2) = g(f(2)) = g(c) = 6$$
$$(g \circ f)(3) = g(f(3)) = g(c) = 6$$

Si noti che, in questo esempio, la composizione $f \circ g$ non è definita in quanto il codominio di g non coincide[8] con il dominio di f. Infatti non ha senso provare a calcolare, per esempio, $f(g(a)) = f(4)$.
In alcuni casi hanno senso invece tutte e due le composizioni, ad esempio se $f\colon X \longrightarrow Y$ e $g\colon Y \longrightarrow X$ allora hanno senso sia $g \circ f\colon X \longrightarrow X$ che $f \circ g\colon Y \longrightarrow Y$. Chiaramente, in generale, si tratta di due funzioni diverse, ovvero si avrà $g \circ f \neq f \circ g$. Per cui possiamo dire che la composizione di funzioni non gode sicuramente della proprietà commutativa.
Questo è vero anche nel caso specifico in cui componiamo due funzioni $f\colon X \longrightarrow X$ e $g\colon X \longrightarrow X$, per le quali si possono considerare sia $g \circ f$ che $f \circ g$ ed entrambe sono funzioni $X \longrightarrow X$.

Esempio 2.37. Si considerino le due funzioni

$$f\colon \mathbb{R} \longrightarrow \mathbb{R}, \quad f(x) = x^2$$
$$g\colon \mathbb{R} \longrightarrow \mathbb{R}, \quad g(x) = x + 1$$

Si ha allora

$$(g \circ f)(x) = g(f(x)) = g(x^2) = x^2 + 1$$

mentre

$$(f \circ g)(x) = f(g(x)) = f(x + 1) = (x + 1)^2 = x^2 + 2x + 1.$$

La composizione di funzioni gode invece sempre della proprietà associativa. Più precisamente, date tre funzioni $f\colon X \longrightarrow Y$, $g\colon Y \longrightarrow Z$, $h\colon Z \longrightarrow W$, si ha

$$(h \circ g) \circ f = h \circ (g \circ f)\,. \tag{2.11}$$

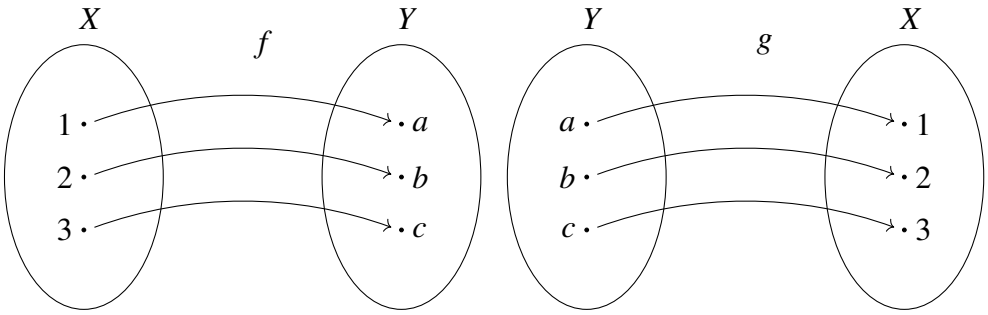
Infatti, si verifica subito che entrambi i membri di questa uguaglianza sono funzioni $X \longrightarrow W$; inoltre, per ogni $x \in X$, per definizione di composizione si ha

$$[(h \circ g) \circ f](x) = (h \circ g)(f(x)) = h(g(f(x))) \text{ e}$$
$$[h \circ (g \circ f)](x) = h((g \circ f)(x)) = h(g(f(x)))\,.$$

[8] In realtà per poter comporre due funzioni $f\colon X \longrightarrow Y$ e $g\colon Y' \longrightarrow Z$ basterebbe anche solo che il codominio Y della prima fosse *contenuto* nel dominio Y' della seconda. Infatti, se $Y \subseteq Y'$, allora ogni elemento di Y è anche un elemento di Y', quindi $f(x) \in Y$ sarebbe anche un elemento di Y' e potremmo applicargli la g ottenendo $g(f(x))$. Tuttavia, per semplicità supporremmo sempre che il codominio della prima funzione coincida col dominio della seconda.

Grazie a questa proprietà, quando dobbiamo comporre tre (o più) funzioni possiamo scrivere semplicemente $h \circ g \circ f$, omettendo le parentesi senza che l'espressione risulti ambigua.

Consideriamo ora $X = \{1, 2, 3\}$ e $Y = \{a, b, c\}$ e le funzioni $f\colon X \longrightarrow Y$ e $g\colon Y \longrightarrow X$ rappresentate nel disegno seguente.



Come si vede subito, si ha

$$(g \circ f)(1) = g(f(1)) = g(a) = 1,$$
$$(g \circ f)(2) = g(f(2)) = g(b) = 2,$$
$$(g \circ f)(3) = g(f(3)) = g(c) = 3,$$

ovvero $g \circ f\colon \{1, 2, 3\} \longrightarrow \{1, 2, 3\}$ è la funzione che manda ogni elemento dell'insieme $X = \{1, 2, 3\}$ in se stesso. Tale particolare funzione si chiama **funzione identica** di X e si denota con id_X. Possiamo quindi scrivere

$$g \circ f = id_X\,. \tag{2.12}$$

Analogamente, si ha

$$(f \circ g)(a) = f(g(a)) = f(1) = a,$$
$$(f \circ g)(b) = f(g(b)) = f(2) = b,$$
$$(f \circ g)(c) = f(g(c)) = f(3) = c,$$

ovvero $f \circ g\colon \{a, b, c\} \longrightarrow \{a, b, c\}$ è la funzione che manda ogni elemento dell'insieme $Y = \{a, b, c\}$ in se stesso, cioè

$$f \circ g = id_Y\,. \tag{2.13}$$

Le uguaglianze (2.12) e (2.13) sono analoghe alle uguaglianze

$$a \cdot b = 1 \text{ e } b \cdot a = 1$$

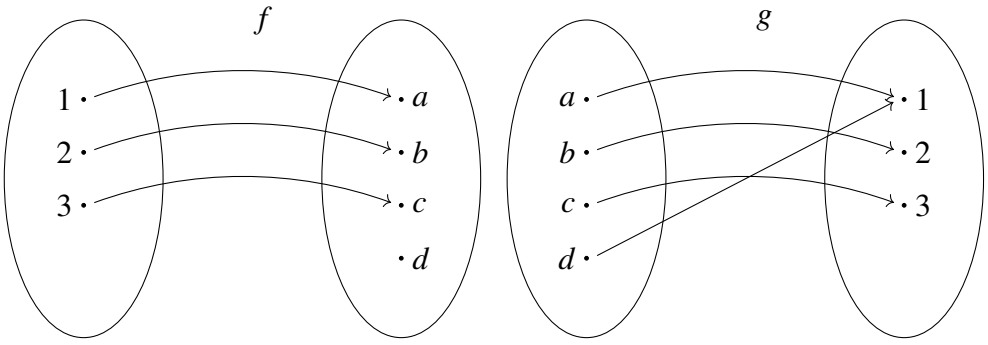
che sussistono quando a e b sono due numeri, uno l'inverso dell'altro (ad esempio, $a = 2$ e $b = \frac{1}{2}$). In effetti, per convincersi ulteriormente di quanto questa analogia sia appropriata, osserviamo che, così come il numero 1 ha la proprietà di fungere da cosiddetto "elemento neutro per la moltiplicazione" (cioè soddisfa $x \cdot 1 = 1 \cdot x = x$ per ogni numero reale x), anche le funzioni identiche fungono da elementi neutri per la composizione. Più precisamente, se abbiamo una funzione $f\colon X \longrightarrow Y$, allora $f \circ id_X = f$ e $id_Y \circ f = f$ in quanto, dal momento che la funzione identica manda ogni elemento in se stesso, si ha

$$(f \circ id_X)(x) = f(id_X(x)) = f(x),$$
$$(id_Y \circ f)(x) = id_Y(f(x)) = f(x)\,.$$

Questa analogia con quello che accade per l'inverso di un numero ci suggerisce la seguente, importante

Definizione 2.38. Una funzione $f\colon X \longrightarrow Y$ si dice **invertibile** se esiste una funzione $g\colon Y \longrightarrow X$ per cui $g \circ f = id_X$ e $f \circ g = id_Y$. In tal caso la funzione g si dice **inversa** di f e la si denota con f^{-1}.

Mentre nell'insieme dei numeri reali tutti i numeri sono invertibili tranne lo zero, per le funzioni e la composizione la situazione è un po' più complessa. Infatti esistono funzioni f per cui vale la $g \circ f = id_X$ ma $f \circ g \neq id_Y$, e in tal caso si dice che g è un'**inversa sinistra** di f. Al contempo esistono funzioni f per cui viceversa si ha $f \circ g = id_Y$ ma $g \circ f \neq id_X$, e in tal caso si dice che g è un'**inversa destra** di f.
Per un esempio di funzione del primo tipo, si prendano $f\colon \{1,2,3\} \longrightarrow \{a,b,c,d\}$ e $g\colon \{a,b,c,d\} \longrightarrow \{1,2,3\}$ definite come nel seguente disegno.

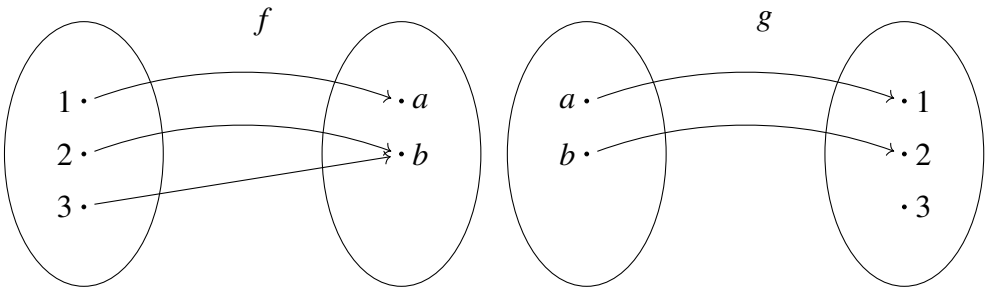


Si vede subito che $g \circ f$ è la funzione identica di $\{1,2,3\}$, ma $f \circ g$ non è la funzione identica di $\{a,b,c,d\}$. Infatti, pur essendo $f(g(a)) = f(1) = a$, $f(g(b)) = f(2) = b$ e $f(g(c)) = f(3) = c$, si ha $f(g(d)) = f(1) = a$, ovvero $f \circ g$ non manda d in se stesso.
Si osservi che non c'è alcun modo di modificare g in modo che sia anche un'inversa destra per f. Qualunque valore assegniamo a d non sarà mai $f(g(d)) = d$, perchè

$g(d)$ dovrebbe essere un elemento mandato da f in d, ma non esiste nessun elemento di $\{1,2,3\}$ che viene mandato da f in d.
In altre parole, il motivo per cui non esiste un'inversa destra di f è dovuto alla non suriettività di f. Questo esempio illustra la validità della seguente proposizione, di cui omettiamo la dimostrazione rigorosa.

Proposizione 2.39. *Una funzione $f\colon X \longrightarrow Y$ ammette un'inversa destra se e solo se è suriettiva*

Analogamente, per un esempio di funzione che ammette un'inversa destra ma non sinistra[9], si prendano $f\colon \{1,2,3\} \longrightarrow \{a,b\}$ e $g\colon \{a,b\} \longrightarrow \{1,2,3\}$ definite dal seguente disegno.



Si nota immediatamente che $f \circ g$ è la funzione identica di $\{a,b\}$, ma $g \circ f$ non è la funzione identica di $\{1,2,3\}$. Di fatto, pur essendo $g(f(1)) = g(a) = 1$ e $g(f(2)) = g(b) = 2$, si ha $g(f(3)) = g(b) = 2$, ovvero $g \circ f$ non manda 3 in se stesso.
Si osservi che anche qui non c'è alcun modo di modificare g in modo che sia anche un'inversa sinistra per f. Se avessimo posto $g(b) = 3$ avremmo sì ottenuto $g(f(3)) = g(b) = 3$ ma stavolta sarebbe stato $g(f(2)) = g(b) = 3$, ovvero $g \circ f$ non avrebbe mandato 2 in se stesso. Come si vede, il problema stavolta è che f manda entrambi gli elementi 2 e 3 in b, quindi necessariamente $g(f(2))$ e $g(f(3))$ coincideranno e non potranno mai essere il primo 2 e il secondo 3.
In altre parole, il motivo per cui non esiste un'inversa sinistra di f è dovuto alla non iniettività di f. Questo esempio illustra la validità della seguente proposizione, di cui omettiamo una dimostrazione rigorosa.

Proposizione 2.40. *Una funzione $f\colon X \longrightarrow Y$ ammette un'inversa sinistra se e solo se è iniettiva*

Poiché, per la Definizione 2.38, dire che f ha un'inversa g significa dire che g è sia inversa sinistra che inversa destra, possiamo combinare la Proposizione 2.39 e la Proposizione 2.40 ottenendo la

[9] In realtà basterebbe l'esempio già visto, in cui è g ad avere f come inversa destra (ma non sinistra) ma per una maggiore chiarezza facciamo un ulteriore esempio.

Proposizione 2.41. *Una funzione $f\colon X \longrightarrow Y$ ammette un'inversa se e solo se è sia iniettiva che suriettiva (ovvero biiettiva).*

2.7 Numeri naturali

In questa sezione useremo la nozione di funzione biiettiva e le sue proprietà per definire rigorosamente i concetti elementari di "numero" e "contare".

Definizione 2.42. Un insieme X si dice **equipotente** a un insieme Y se esiste una funzione biiettiva $f\colon X \longrightarrow Y$.

La Definizione 2.42 definisce una relazione tra insiemi, che denotiamo in questo capitolo con il simbolo $\sim$.

Proposizione 2.43. *La relazione di equipotenza definita nella Definizione 2.42 è una relazione di equivalenza.*

Prima di dimostrare la proposizione enunciamo un lemma indipendentemente per la sua importanza e per referenza futura.

Lemma 2.44. *Siano $f\colon X \longrightarrow Y$ e $g\colon Y \longrightarrow Z$ due funzioni. Allora*

(1) se f e g sono iniettive, anche $g \circ f$ è iniettiva

(2) se f e g sono suriettive, anche $g \circ f$ è suriettiva

In particolare, se f e g sono biiettive, anche $g \circ f$ è biiettiva.

Dimostrazione. Per dimostrare che $g \circ f\colon X \longrightarrow Z$ è iniettiva, dobbiamo mostrare che per ogni $x_1, x_2 \in X$, se $(g \circ f)(x_1) = (g \circ f)(x_2)$ allora $x_1 = x_2$. Ma, per definizione di composizione, $(g \circ f)(x_1) = (g \circ f)(x_2)$ significa $g(f(x_1)) = g(f(x_2))$. Da questa uguaglianza, siccome g è iniettiva segue che $f(x_1) = f(x_2)$, e siccome f è iniettiva a sua volta da quest'ultima segue $x_1 = x_2$, come volevamo.
Per dimostrare che $g \circ f\colon X \longrightarrow Z$ è suriettiva dobbiamo dimostrare che per ogni $z \in Z$ esiste un $x \in X$ tale che $(g \circ f)(x) = z$. Ora, dal momento che $g\colon Y \longrightarrow Z$ è suriettiva, per ogni $z \in Z$ esiste un $y \in Y$ tale che $z = g(y)$. A sua volta, dal momento che y appartiene al codominio di $f\colon X \longrightarrow Y$ che è anch'essa per ipotesi suriettiva, esiste un $x \in X$ tale che $y = f(x)$. Sostituendo allora $y = f(x)$ in $z = g(y)$ si ottiene $z = g(f(x))$, ovvero per definizione di composizione $z = (g \circ f)(x)$, come volevamo.
L'ultima parte del lemma è un corollario immediato delle prime due. □

Siamo ora pronti per la

Dimostrazione della Proposizione 2.43. Per dimostrare che $\sim$ è una relazione di equivalenza dobbiamo mostrare che è riflessiva, simmetrica e transitiva.

i) **Riflessività.** Dato un qualunque insieme X, la funzione identica $id_X : X \longrightarrow X$ che manda ogni elemento di X in se stesso è una funzione biiettiva (è chiaramente iniettiva in quanto due elementi x, x' diversi hanno come immagini x, x' stessi, quindi immagini diverse; è chiaramente suriettiva in quanto ogni elemento x del codominio è immagine di se stesso nel dominio).

ii) **Simmetria.** Supponiamo che sia $X \sim Y$, ossia che esista una funzione biiettiva $f\colon X \longrightarrow Y$. Per mostrare che $Y \sim X$ (e quindi la simmetria) ci basta trovare una funzione biiettiva da Y a X. A questo scopo, basta ricordare che le funzioni biiettive sono caratterizzate dal fatto di essere invertibili e prendere la funzione $f^{-1} : Y \longrightarrow X$ inversa di f. Essendo f^{-1} l'inversa di f, si ha $f \circ f^{-1} = id_Y$ e $f^{-1} \circ f = id_X$. Ma queste due uguaglianze ci dicono che f è a sua volta l'inversa (sia destra che sinistra) di f^{-1}. Quindi f^{-1} è invertibile e, per quanto appena ricordato, biiettiva. Abbiamo dunque trovato una funzione biiettiva da Y a X. Questo implica che $Y \sim X$, e possiamo concludere che la relazione è simmetrica.

iii) **Transitività.** Supponiamo che sia $X \sim Y$ e $Y \sim Z$, ovvero che esista una funzione $f\colon X \longrightarrow Y$ biiettiva da X a Y e una funzione $g\colon Y \longrightarrow Z$ biiettiva da Y a Z. Se riusciamo a mostrare che la composizione $g \circ f\colon X \longrightarrow Z$ è anch'essa biiettiva avremo dimostrato che esiste una funzione biiettiva da X a Z e quindi che $X \sim Z$, da cui la transitività. A questo scopo, dimostriamo separatamente che $g \circ f\colon X \longrightarrow Z$ è iniettiva e che è suriettiva.

   Per quanto riguarda l'iniettività, siano $x, x' \in X$, con $x \neq x'$. Vogliamo dimostrare che $(g \circ f)(x) \neq (g \circ f)(x')$. Iniziamo con l'osservare che, dal momento che f è biiettiva e quindi in particolare iniettiva, abbiamo $f(x) \neq f(x')$. Essendo $f(x)$ e $f(x')$ elementi diversi di Y, che è il codominio di f ma anche il dominio di $g\colon Y \longrightarrow Z$, possiamo applicare loro la g e, essendo anche quest'ultima biiettiva e in particolare iniettiva, si avrà $g(f(x)) \neq g(f(x'))$. Per definizione di composta questo significa proprio $(g \circ f)(x) \neq (g \circ f)(x')$, che è quello che volevamo.

   Per la suriettività di $g \circ f\colon X \longrightarrow Z$, dobbiamo dimostrare che per ogni $z \in Z$ esiste un $x \in X$ tale che $(g \circ f)(x) = z$. Iniziamo con l'osservare che, dal momento che $g\colon Y \longrightarrow Z$ è biiettiva e in particolare suriettiva, esiste (almeno) un $y \in Y$ tale che $g(y) = z$. Ma dal momento che $f\colon X \longrightarrow Y$ è biiettiva e in particolare suriettiva, a sua volta esiste un $x \in X$ tale che $f(x) = y$. Ma allora possiamo riscrivere la $g(y) = z$ come $g(f(x)) = z$. Per definizione di composizione questo significa proprio $(g \circ f)(x) = z$, che è quello che volevamo dimostrare.

□

Essendo l'equipotenza una relazione di equivalenza, possiamo considerare delle sue classi di equivalenza. Ebbene, i numeri naturali, che ci accingiamo a definire in modo

rigoroso, sono esattamente il modo in cui noi identifichiamo le classi di equivalenza degli insiemi finiti. Infatti, se possiamo dire per esempio che l'insieme $X = \{a, b, c, d\}$ ha 4 elementi, è perché abbiamo *contato* i suoi elementi. Cioè li abbiamo messi in corrispondenza con gli elementi dell'insieme $\{1, 2, 3, 4\}$ mediante una funzione biettiva (ad esempio, a 1 associamo a, a 2 associamo b, a 3 associamo c, a 4 associamo d). Chiaramente avremmo contato male se avessimo contato a due volte, associando sia a 1 che a 2 l'elemento a, e anche se dimenticassimo di contare qualche elemento di X, ad esempio d. Il primo errore equivarrebbe alla costruzione di una corrispondenza $\{1, 2, 3, 4\} \longrightarrow \{a, b, c, d\}$ non iniettiva (in cui gli elementi diversi 1 e 2 del dominio hanno la stessa immagine, a, nel codominio). Il secondo errore significherebbe aver costruito una corrispondenza non suriettiva (in cui l'elemento d, non essendo stato contato, non è immagine di nessuno degli elementi $1, 2, 3, 4$ del dominio).
In generale, *contare* gli elementi di un insieme X e concludere che questo ha n elementi significa costruire una funzione biiettiva tra X e l'insieme dei numeri naturali che vanno da 1 a n.
Per rendere rigoroso cosa significa "l'insieme dei numeri naturali da 1 a n" dobbiamo definire i numeri naturali e chiarire da dove nasce la successione naturale in cui li pensiamo quando contiamo. La definizione che daremo ora è assiomatica, ovvero definiremo l'insieme $\mathbb{N}$ dei numeri naturali mediante assiomi che, pur non dicendoci cosa siano i numeri naturali, ci dicono però le relazioni che intercorrono tra di essi e li definiscono così implicitamente.[10]
Enunciamo allora gli assiomi che implicitamente definiscono i numeri naturali, detti **assiomi di Peano** (dal nome del matematico italiano Giuseppe Peano, al quale sono dovuti).

(1) Lo zero 0 è un numero naturale.

    Questo assioma afferma che l'insieme dei numeri naturali è un insieme non vuoto, nel quale c'è almeno un elemento che chiamiamo "zero" (i successivi assiomi chiariranno la particolarità di questo elemento rispetto agli altri). Si noti che se ci fermassimo a questo assioma, l'insieme dei numeri naturali potrebbe anche contenere solo 0, ovvero essere $\mathbb{N} = \{0\}$.

(2) Per ogni numero naturale n, ne esiste un altro n' detto il suo *successore*.

    Questo assioma introduce in pratica una funzione $s : \mathbb{N} \longrightarrow \mathbb{N}$ dall'insieme $\mathbb{N}$ dei numeri naturali in se stesso, detta appunto funzione successore.

(3) Due numeri naturali diversi hanno successori diversi.

[10] Esattamente come nella costruzione mediante assiomi della geometria euclidea non diciamo cosa siano punto e retta ma specifichiamo mediante gli assiomi le relazioni che li legano - ad esempio "dati due punti distinti esiste ed è unica la retta che li contiene".

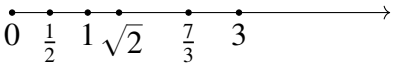
Questo assioma ci sta dicendo che la funzione successore è iniettiva.

(4) Lo zero 0 non è successore di nessun numero.

Prima di enunciare il quinto e ultimo assioma, fermiamoci a riflettere sulle prime conseguenze degli assiomi (1), (2), (3), (4). L'assioma (4) ci dice che il successore di 0 non può essere 0 stesso.[11] Chiamando "uno" il successore di zero e denotandolo 1, vediamo che il successore di 1 (che esiste per il secondo assioma) non può essere né zero (per il quarto assioma) nè nuovamente 1, altrimenti contrariamente a quello che afferma il terzo assioma la funzione successore non sarebbe iniettiva. Quindi abbiamo che il successore di 1 è un nuovo numero naturale, diverso da 0 e 1, che denotiamo con 2, e così via.

Questi primi quattro assiomi impediscono quindi ad esempio che i numeri naturali siano ad esempio l'insieme finito $\{0,1,2,3\}$ con successore di 0 dato da 1, successore di 1 dato da 2, successore di 2 dato da 3, e successore di 3 dato da 0 (o da un altro degli elementi $1,2,3$). Ovvero non è possibile che la funzione successore "permuti ciclicamente" gli elementi. Gli assiomi (1), (2), (3), (4) garantiscono quindi che l'insieme dei naturali sia un insieme infinito.[12]

Siamo ora pronti a enunciare il quinto e ultimo assioma di Peano. Allo scopo di comprendere meglio l'importanza di tale assioma, osserviamo che i primi 4 assiomi risultano soddisfatti se consideriamo ad esempio l'insieme $\mathbb{R}$ di tutti i numeri reali (cioè tutti i decimali, limitati o illimitati, periodici e non, con segno + o -) maggiori o uguali di zero, rappresentati lungo una semiretta come nel disegno.



Come funzione successore consideriamo la funzione che manda un qualsiasi numero x nel numero $x+1$. Questo ci fa capire che i primi quattro assiomi non sono sufficienti a definire in modo univoco i numeri naturali così come siamo abituati a pensarli. Come questo esempio illustra bene, i primi quattro assiomi non impediscono che oltre allo zero e ai numeri che ottengo applicando ripetutamente la funzione successore ci siano altri elementi. Per evitare questo e fare di $\mathbb{N}$ *il più piccolo insieme che soddisfi primi quattro assiomi*, serve il seguente, quinto e ultimo assioma.

[11] Si noti che questo non è impedito dagli altri assiomi: l'insieme $\{0\}$ costituito da un solo elemento denotato 0, con la funzione successore che associa a 0 se stesso, verifica i primi tre assiomi . Per l'iniettività, si osservi che l'implicazione che definisce una funzione iniettiva, cioè $x \neq x' \Rightarrow f(x) \neq f(x')$, è vera in quanto, non essendoci due elementi diversi, l'antecedente dell'implicazione è sempre falso, il che come sappiamo rende l'implicazione automaticamente vera.

[12] Daremo più avanti la definizione rigorosa di insieme infinito, e usiamo per il momento questa parola con il suo significato intuitivo nel linguaggio comune.

(5) Se un sottoinsieme U dell'insieme dei numeri naturali contiene 0, e per ogni numero n che sta in U appartiene a U anche il successore n' di n, allora U coincide con $\mathbb{N}$.

Questo assioma garantisce che si ottengano tutti i numeri naturali partendo dallo zero e applicando ripetutamente la funzione successore, e che non ci siano altri elementi in $\mathbb{N}$ oltre a quelli che si possono ottenere in questo modo.

Come abbiamo accennato sopra, gli assiomi di Peano caratterizzano completamente i numeri naturali così come li intendiamo comunemente, e a partire da essi è possibile costruire tutta l'aritmetica usuale. Ad esempio è possibile definire le operazioni di somma e prodotto tra naturali e verificarne le varie proprietà. In base a tali definizioni, che non diamo, il successore di un numero naturale n risulta $n + 1$, ed è così che lo denoteremo da ora in poi.
Tornando al punto da cui siamo partiti, gli assiomi chiariscono in che modo i naturali sono i numeri che "ci servono per contare". Quando stiamo contando gli elementi di un insieme, partiamo da uno e applichiamo ripetutamente la funzione successore, fino a che non avremo esaurito gli elementi dell'insieme che stiamo contando. Il numero a cui arriviamo ci dice "quanti elementi ha l'insieme dato" ovvero identifica la classe di equivalenza della relazione di equipotenza. Quando diciamo che due insiemi X e Y hanno entrambi 4 elementi stiamo dicendo che esiste una funzione biiettiva dall'insieme $\{1, 2, 3, 4\}$ in X e una funzione biiettiva dall'insieme $\{1, 2, 3, 4\}$ in Y, e quindi che, dato che la relazione di equipotenza è una relazione di equivalenza, esiste una funzione biiettiva da X a Y. Possiamo dire che X e Y stanno nella stessa classe di equivalenza, determinata dal numero naturale 4, e che "hanno lo stesso numero di elementi". Si noti che nel contare partiamo da 1: il numero zero è usato per denotare la classe di equivalenza dell'insieme vuoto rispetto alla relazione di equipotenza.

2.8 Applicazioni del quinto assioma (il principio d'induzione)

Il quinto assioma di Peano (detto anche *principio d'induzione*) è di fondamentale importanza quando si vuole dimostrare che una certa proprietà, un'uguaglianza o una formula vale per ogni numero naturale. Infatti, basterà dimostrare che tale proprietà/formula/uguaglianza etc. vale per 0 (questa prima parte si chiama spesso *passo base*), e poi che se essa vale per un numero naturale m allora vale anche per il suo successore $m + 1$ (questa seconda parte è spesso chiamata *passo induttivo*). In base al quinto assioma, concluderemo allora che l'insieme U dei naturali per i quali vale la nostra formula coincide con l'insieme di tutti i numeri naturali.

Esempi 2.45. Vediamo subito alcuni semplici ma significativi esempi.

(1) Supponiamo di voler dimostrare applicando il principio di induzione che la disuguaglianza

$$2^n \geqslant n+1 \tag{2.14}$$

vale per ogni numero naturale n.

A questo scopo, osserviamo innanzitutto che la disuguaglianza vale per $n = 0$, in quanto per tale valore primo e secondo membro sono uguali ($2^0 = 1$ e $0+1 = 1$). Il cosiddetto passo base è concluso.

Per applicare il cosiddetto passo induttivo, supponiamo ora che la disuguaglianza si vera per un certo numero naturale arbitrario m (questa ipotesi è detta *ipotesi induttiva*, cioè che $2^m \geqslant m+1$. Mostriamo che, *sotto tale ipotesi*, essa vale anche per il suo successore $m+1$, ovvero che $2^{m+1} \geqslant (m+1)+1$. Infatti, $2^{m+1} = 2 \cdot 2^m$: ma essendo per ipotesi induttiva $2^m \geqslant m+1$ sarà $2 \cdot 2^m \geqslant 2 \cdot (m+1)$ (una disuguaglianza rimane valida se moltiplichiamo primo e secondo membro per uno stesso numero positivo). Quindi

$$2^{m+1} \geqslant 2(m+1) = 2m+2$$

e per concludere la disuguaglianza che serva a noi, ovvero che $2^{m+1} \geqslant (m+1)+1$, ci basterà allora mostrare che $2m+2 \geqslant (m+1)+1$. Ma quest'ultima disuguaglianza significa semplicemente $2m+2 \geqslant m+2$, che semplificando il 2 da entrambi i membri e portando a primo membro l'm che compare a secondo ci dà $m \geqslant 0$, cioè essa è vera qualunque sia m.

Riassumendo, abbiamo dimostrato che la (2.14) è vera per $n = 0$ e che se essa è vera per m allora è vera per $m+1$. In base al principio d'induzione essa è vera per qualunque numero naturale.

(2) La somma dei numeri naturali da 0 a n vale $\frac{n(n+1)}{2}$, ovvero, usando il simbolo di sommatoria,

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2} \tag{2.15}$$

Per dimostrare per induzione che questa formula vale per ogni numero naturale $n \in \mathbb{N}$, verifichiamo prima che essa vale per $n = 0$: infatti, in tal caso la sommatoria contiene solo 0, mentre dall'altra si ha $\frac{0(0+1)}{2} = 0$, quindi l'uguaglianza è verificata.

Ora, supponiamo di sapere che la formula (2.15) valga per n, e dimostriamo che essa vale per $n+1$, ovvero dimostriamo che

$$\sum_{k=0}^{n+1} k = \frac{(n+1)[(n+1)+1]}{2} = \frac{(n+1)(n+2)}{2}\,. \tag{2.16}$$

Si ha chiaramente

$$\sum_{k=0}^{n+1} k = \sum_{k=0}^{n} k + (n+1)\,. \tag{2.17}$$

Ma poiché stiamo supponendo che la formula (2.15) valga per n, possiamo sostituire $\sum_{k=0}^{n} k$ con $\frac{n(n+1)}{2}$, e quindi la (2.17) si scrive

$$\sum_{k=0}^{n+1} k = \frac{n(n+1)}{2} + (n+1)\,.$$

Ma, svolgendo i conti, il secondo membro è uguale a

$$\frac{n(n+1)}{2} + (n+1) = \frac{n(n+1)+2(n+1)}{2} = \frac{(n+2)(n+1)}{2}$$

e quindi abbiamo dimostrato la (2.16), come volevamo.

Osservazione 2.46. Il metodo si applica anche quando si vuole dimostrare che una certa proprietà P vale non per ogni naturale ma a partire da un $n_0 \in \mathbb{N}$ dato. In tal caso, bisogna dimostrare che

(1) la proprietà P vale per n_0

(2) ogniqualvolta la P vale per un certo numero n, allora vale per il suo successivo $n+1$ (scriveremo $P(n) \Rightarrow P(n+1)$)

Esempio 2.47. Per ogni numero naturale $n \geqslant 1$, si ha

$$2^{n-1} \leqslant n! \tag{2.18}$$

Dimostriamo anche questa formula per induzione, nella versione vista nell'Osservazione 2.46. Poichè vogliamo mostrare che essa vale per ogni $n \geqslant 1$, iniziamo con il dimostrare che essa vale per $n = 1$. Sostituendo $n = 1$ in (2.18) si trova da un lato $2^{1-1} = 2^0 = 1$ e dall'altro $1! = 1$, ovvero $1 \leqslant 1$, che è vera.
Ora, come prevede il metodo di induzione, supponiamo che la (2.18) sia valida per un certo n e dimostriamo che essa vale per $n+1$, ovvero che

$$2^{(n+1)-1} \leqslant (n+1)!\,. \tag{2.19}$$

Si ha chiaramente $2^{(n+1)-1} = 2^n = 2 \cdot 2^{n-1}$, e allora poiché stiamo supponendo che valga la $2^{n-1} \leqslant n!$ possiamo scrivere

$$2^{(n+1)-1} = 2 \cdot 2^{n-1} \leqslant 2 \cdot n! \leqslant (n+1) \cdot n! = (n+1)!$$

dove nell'ultima disuguaglianza abbiamo sfruttato il fatto che per ogni $n \geqslant 1$ si ha $2 \leqslant n+1$. La (2.19) è dimostrata.

Osservazione 2.48. Quando si dimostra per induzione che una certa proprietà P vale per ogni numero naturale maggiore o uguale di un numero di partenza n_0, dopo aver dimostrato che $P(n_0)$ è vera, bisogna stare attenti che l'implicazione $P(n) \Rightarrow P(n+1)$ sia valida effettivamente per ogni $n \geqslant n_0$.
Ad esempio, consideriamo la disuguaglianza

$$2^n \geqslant n^2 \tag{2.20}$$

e supponiamo di voler dimostrare che essa è vera per tutti i numeri naturali.
Chiaramente, tale disuguaglianza è vera per $n = 0$ in quanto si riduce a $2^0 = 1 \geqslant 0^2 = 0$.
Ora, come prevede l'induzione, supponiamo che la formula sia vera per n (ipotesi induttiva) e dimostriamo che essa vale per $n+1$, ovvero che

$$2^{(n+1)} \geqslant (n+1)^2 . \tag{2.21}$$

Usando l'ipotesi induttiva $2^n \geqslant n^2$, si ha

$$2^{(n+1)} = 2 \cdot 2^n \geqslant 2 \cdot n^2$$

e quindi per dimostrare la (2.21) basterebbe mostrare che $2n^2 \geqslant (n+1)^2$, ovvero $2n^2 \geqslant n^2 + 2n + 1$, che portando tutto a primo membro equivale alla disuguaglianza

$$n^2 - 2n - 1 \geqslant 0 . \tag{2.22}$$

Ora, ricordiamo che una generica disuguaglianza di secondo grado $ax^2 + bx + c \geqslant 0$ è verificata per $x \geqslant x_1$ e $x \leqslant x_2$ se a è positivo, e per $x_1 \leqslant x \leqslant x_2$ se a è negativo, dove x_1 e x_2 sono le soluzioni di $ax^2 + bx + c = 0$.
Nel nostro caso, usando la nota formula risolutiva $x_{1,2} = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$, si vede che $n^2 - 2n - 1 = 0$ ha come soluzioni $n = 1 - \sqrt{2}$ e $n = 1 + \sqrt{2}$. Quindi la nostra disuguaglianza (2.22) è verificata per $n \geqslant 1 + \sqrt{2} \sim 2,4$ e $n \leqslant 1 - \sqrt{2} \sim -0,4$. Siccome stiamo lavorando nei naturali, il primo numero naturale per cui la disuguaglianza è verificata è 3. Ne consegue che l'implicazione $P(n) \Rightarrow P(n+1)$ che vogliamo dimostrare è vera solo per $n \geqslant 3$. In conclusione, pur essendo $P(0)$ vera non possiamo partire da 0 per innescare il meccanismo dell'induzione. Come minimo, dobbiamo partire da 3, tuttavia, $P(3)$ non è vera. Infatti, sostituendo $n = 3$ in (2.20) si ottiene $2^3 \geqslant 3^2$, cioè $8 \geqslant 9$.

Sostituendo $n = 4$ vediamo invece che $2^4 \geqslant 4^2$ (cioè $16 \geqslant 16$), ovvero $P(4)$ è vera. Quindi non solo possiamo usare 4 come punto di partenza dell'induzione, ma sapendo anche che per $n \geqslant 4$ la $P(n) \Rightarrow P(n+1)$ è vera, la formula risulta dimostrata per induzione per $n \geqslant 4$.

Il principio d'induzione viene a volta enunciato e usato anche nella seguente forma, detta **principio d'induzione forte**.

Se una proprietà

(i) vale per $n = 0$

(ii) vale per $n+1$ *sotto l'ipotesi che valga per ogni numero naturale* k *compreso tra* 0 *e* n *(cioè* $0 \leqslant k \leqslant n$*)*

allora essa vale per tutti i numeri naturali.

In questa forma, il passo base (i) è identico al passo base del principio d'induzione già enunciato sopra mentre il passo induttivo si differenzia invece dal passo induttivo descritto prima. Infatti non assumiamo più che la proprietà valga solo per n per concludere poi che vale per $n+1$ (cioè che vale per un numero ogniqualvolta essa vale per quello immediatamente precedente), ma assumiamo che valga per tutti i numeri tra 0 e n (cioè che vale per un numero ogniqualvolta essa vale per *tutti* quelli precedenti).
Benchè in questa forma l'ipotesi induttiva sembri più restrittiva (poiché stiamo chiedendo qualcosa in più), si può dimostrare che queste due formulazioni diverse sono in realtà del tutto equivalenti. In altre parole, se sostituissimo il principio d'induzione forte al posto del quinto assioma di Peano così come l'abbiamo enunciato sopra si otterrebbe sempre lo stesso insieme $\mathbb{N}$ dei numeri naturali.
Tuttavia a volte l'uso del principio d'induzione nella forma forte si rivela più efficace, come vediamo nel seguente

Esempio 2.49. Un numero naturale[13] $n > 1$ si dice *primo* se gli unici modi di decomporlo come prodotto di due numeri naturali sono $n = n \cdot 1$ e $n = 1 \cdot n$. Ad esempio, 5 è primo perché non ci sono altre decomposizioni a parte $5 = 5 \cdot 1$ e $5 = 1 \cdot 5$, mentre 6 non è primo perché oltre a $6 = 6 \cdot 1$ e $6 = 1 \cdot 6$ abbiamo ad esempio anche $6 = 2 \cdot 3$. Dimostriamo ora, usando il principio d'induzione forte, che ogni numero naturale $n \geqslant 2$ si può decomporre come prodotto (con uno, due o più fattori) di primi. Per $n = 2$ questo è banalmente vero in quanto essendo esso stesso primo la decomposizione in fattori primi è data da 2 stesso (decomposizione con un solo fattore). Ora, supponiamo che questo valga per ogni numero naturale $2 \leqslant k \leqslant n$, cioè che ogni numero naturale compreso tra 2 e n si possa esprimere come prodotto di primi (ipotesi induttiva) e mostriamo che sotto tale ipotesi la stessa cosa vale per $n+1$, cioè che $n+1$ può essere scritto come

[13] In realtà questa definizione vale sull'insieme dei numeri interi, dei quali ci occuperemo più diffusamente nel prossimo capitolo.

prodotto di primi. Si hanno due possibilità: o $n+1$ è già primo, e in tal caso, esattamente come abbiamo fatto per 2, la decomposizione cercata è data da $n+1$ stesso (un solo fattore), oppure $n+1$ non è primo. In questo secondo esisterà una decomposizione $n+1 = a \cdot b$ di $n+1$ come prodotto in cui i due fattori non sono 1 e quindi sono necessariamente compresi tra 2 e n. In base all'ipotesi induttiva fatta, per cui ogni numero compreso tra 2 e n può essere decomposto in fattori primi, sia a che b possono essere decomposti in fattori primi, ovvero diciamo $a = p_1 p_2 \cdots p_m$ e $a = p'_1 p'_2 \cdots p'_l$. Ma allora $n+1 = ab = p_1 p_2 \cdots p_m p'_1 p'_2 \cdots p'_l$, e questa è proprio una decomposizione di $n+1$ in fattori primi, della quale dovevamo dimostrare l'esistenza. Questo completa la verifica delle due condizioni espresse dal principio d'induzione forte e quindi in base a tale principio mostra che tale affermazione vale per tutti i numeri naturali, come voluto.

Sottolineiamo come in questa dimostrazione il principio d'induzione nella forma classica sarebbe stato del tutto inutile. Infatti, non possiamo dedurre che $n+1$ ha una decomposizione in primi dal fatto che n ce l'ha. Ad esempio non possiamo ottenere una decomposizione in fattori primi di $n+1 = 60$ usando la decomposizione in fattori primi del suo predecessore $n = 59$ (le due composizioni in generale non hanno nulla a che vedere l'una con l'altra). Però, notando che $60 = 6 \cdot 10$ e conoscendo le decomposizioni in fattori primi di tutti i numeri che precedono 60, quindi in particolare anche di $6 = 2 \cdot 3$ e di $10 = 2 \cdot 5$, ottengo subito la decomposizione $60 = 2 \cdot 3 \cdot 2 \cdot 5$ di 60 in fattori primi.

Un'altra importante applicazione del principio d'induzione (sia nella forma classica che in quella forte) sono le cosiddette *definizioni ricorsive*. Tali definizioni consentono ad esempio di definire una successione di numeri $a_0, a_1, a_2, \ldots$ senza dare l'espressione esplicita del termine generico a_n.
Ad esempio, la successione $a_n = 2^n$ è tale che il primo termine $a_0 = 1$ e i successivi termini sono $a_1 = 2, a_2 = 4, a_3 = 8$ e così via. Notiamo che ogni termine viene ottenuto moltiplicando per 2 il precedente. Possiamo sfruttare questa osservazione per scrivere la successione come

$$\begin{cases} a_0 = 1 \\ a_{n+1} = 2a_n \,. \end{cases}$$

Il fatto che queste due uguaglianze siano sufficienti a ottenere tutti i termini a_n e definire così completamente la successione è garantito proprio dal principio d'induzione (in questo caso nella forma classica). Difatti, detto U l'insieme dei numeri naturali per cui la successione è definita, la $a_0 = 1$ ci dice che la successione è definita per $n = 0$, ovvero $0 \in U$, La seconda condizione $a_{n+1} = 2a_n$ implica che se per un certo numero naturale n il termine a_n è definito (cioè $n \in U$), allora lo è anche a_{n+1} (che si ottiene infatti semplicemente moltiplicando per 2 il termine a_n), ovvero $n+1 \in U$. Quindi sono verificate le due condizioni del principio d'induzione che garantiscono che $U = \mathbb{N}$, cioè la successione è definita per qualunque $n \in \mathbb{N}$.

Un altro esempio di successione definita in questo modo é il fattoriale $a_n = n!$, che può essere definito ricorsivamente da

$$\begin{cases} a_0 = 1 \\ a_{n+1} = (n+1)a_n \end{cases}$$

come è facile verificare.
Una definizione ricorsiva può sfruttare anche il principio d'induzione forte: questo succede quando nella definizione del termine a_{n+1} della successione non si utilizza solo il termine immediatamente precedente ma anche gli altri (o qualcuno degli altri). Un esempio significativo è dato dalla successione

$$\begin{cases} a_0 = 1 \\ a_1 = 1 \\ a_{n+1} = a_n + a_{n-1} \end{cases} \tag{2.23}$$

dove il termine a_{n+1} risulta definito da a_n e a_{n-1}. I primi termini della successione sono

$$a_0 = 1,\ a_1 = 1,\ a_2 = 2,\ a_3 = 3,\ a_4 = 5,\ a_5 = 8,\ a_6 = 13,\ \ldots$$

Si noti che è necessario dare il valore non solo di a_0 ma anche di a_1 in quanto la $a_{n+1} = a_n + a_{n-1}$, per $n = 1$, ci dà $a_1 = a_0 + a_{-1}$, e non essendo a_{-1} definito (-1 non è un numero naturale), non è possibile ottenere a_1 solo a partire da a_0. A partire da $n = 1$ invece gli indici a secondo membro della $a_{n+1} = a_n + a_{n-1}$ sono sempre numeri naturali, questo problema non si presenta. Quindi solo a partire da $n = 1$ la $a_{n+1} = a_n + a_{n-1}$ ci consente sempre di definire a_{n+1} una volta che sono definiti i precedenti, come richiede l'induzione forte.[14]
La successione (2.23) è un significativo esempio di come definire una successione per ricorsione possa essere molto più semplice che dare l'espressione esplicita del termine a_n. Inoltre la (2.23) fornisce un'illustrazione di come possa essere tutt'altro che immediato capire quale sia tale espressione. Più precisamente, si può vedere che il generico termine a_n definito dalla (2.23) è

$$a_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{n+1}.$$

Osservazione 2.50. La successione (2.23) prende il nome di *successione di Fibonacci*, dal nome del matematico italiano del tredicesimo secolo che risolse tramite di essa

[14] In altre parole, il punto è che l'implicazione "se a_k è definito per ogni $0 \leqslant k \leqslant n$ allora a_{n+1} è definito" deve essere verificata per n generico, ma non lo è per $n = 0$. Di fatto per $n = 0$ essa diventa "se a_k è definito per ogni $0 \leqslant k \leqslant 0$ allora a_1 è definito", ovvero "se a_0 è definito allora a_1 è definito",. Questo non è vero proprio perché, come abbiamo osservato, dalla $a_{n+1} = a_n + a_{n-1}$ e conoscendo solo a_0 non riusciamo a ricavare a_1.

il seguente problema (proposto dall'imperatore Federico II). Supponiamo di partire al momento 0 da una coppia di conigli, che al mese 1 diventa fertile e al mese 2 genera una nuova coppia. Supponiamo inoltre che ogni nuova coppia si comporti nello stesso modo (cioè dal momento della nascita impiega un mese a diventare fertile e un altro mese per generare un'altra coppia). Quante coppie di conigli avremo all'n-esimo mese?
Il fatto che sia la successione (2.23) a dirci il numero di coppie di conigli all'n-esimo mese è il ragionamento seguente. Al mese 0 abbiamo la coppia di partenza ($a_0 = 1$). Al mese 1 questa coppia diventa fertile ma non ha ancora generato (quindi ancora $a_1 = 1$). Al mese 2 avremo tante coppie quante ne avevamo al mese 1 (stiamo ovviamente supponendo che le coppie non muoiano) più una per ogni coppia presente al mese 0 (le coppie dell'ultimo mese sono fertili ma non generano ancora). Al mese 3 avremo tante coppie quante ne avevamo al mese 2 più una nuova per ogni coppia presente al mese 1 (le coppie presenti al mese 2 ma non al mese 1 sono fertili ma non generano ancora). Iterando questo ragionamento vediamo che al mese $(n+1)$-esimo abbiamo tante coppie a_{n+1} quante sono quelle a_n del mese precedente più le nuove coppie generate da quelle a_{n-1} del mese precedente ancora (le coppie generate al mese n-esimo sono fertili ma non hanno ancora generato nel mese $n+1$).

2.9 Calcolo combinatorio

Ora che abbiamo visto cosa significa che due insiemi (finiti o infiniti) hanno lo stesso numero di elementi e abbiamo dato un preciso significato teorico al verbo "contare", vogliamo risolvere problemi pratici che consistono proprio nel dire quanti elementi ci sono in un insieme dato o nel contare quanti possibili configurazioni diverse può assumere un sistema dato sotto certe condizioni. Problemi di questo tipo sono l'oggetto di studio di quel ramo della matematica detta *combinatoria*.

Ci concentreremo inizialmente sulla seguente

Domanda: *In quanti modi è possibile permutare (ovvero cambiar d'ordine) n oggetti distinti?*

La risposta è data dal ragionamento seguente: dare una disposizione significa dire chi è il primo elemento, chi il secondo e così via. Ora, essendoci n elementi, il primo che scegliamo potrebbe essere uno qualunque di essi, quindi abbiamo n possibili scelte. Per ognuna di queste n scelte, il secondo elemento può essere scelto tra tutti gli elementi diversi da quello che abbiamo già scelto come primo, quindi abbiamo $n-1$ possibili scelte per ognuna delle n scelte precedenti. In tutto $n(n-1)$ scelte per i primi due elementi. Per ognuna di queste $n(n-1)$ scelte, il terzo elemento può essere scelto tra gli $n-2$ rimasti (tutti tranne i due già scelti come primo e secondo). Quindi abbiamo in

tutto $n(n-1)(n-2)$ scelte, etc., fino a che non esauriamo tutti gli elementi. Il numero di possibili scelte è quindi

$$n(n-1)(n-2)(n-3)\cdots 1 \tag{2.24}$$

ovvero n moltiplicato per tutti i numeri naturali precedenti escluso lo zero.
Il numero che compare nella (2.24) si dice *fattoriale di n* e si denota con il simbolo $n!$.
A scopo esemplificativo, scriviamo di seguito i primi fattoriali:

$$\begin{aligned}
1! &= 1\\
2! &= 2\cdot 1 = 2\\
3! &= 3\cdot 2\cdot 1 = 6\\
4! &= 4\cdot 3\cdot 2\cdot 1 = 24\\
5! &= 5\cdot 4\cdot 3\cdot 2\cdot 1 = 120\\
&\vdots
\end{aligned}$$

Per convenzione, si pone inoltre $0! = 1$.
Come si vede, il fattoriale di n cresce molto rapidamente all'aumentare di n.
Ad esempio, elenchiamo esplicitamente tutte le $3! = 6$ permutazioni di 3 elementi a, b, c:

$$abc,\ bac,\ acb,\ cba,\ bca, cab\,.$$

Il calcolo del numero di permutazioni di un insieme con n elementi fa parte del cosiddetto *calcolo combinatorio*, che si occupa di contare i modi per ordinare, raggruppare o scegliere gli elementi di insiemi finiti, sotto condizioni date.
Vediamo alcuni altri problemi tipici del calcolo combinatorio, nei quali il fattoriale ha un ruolo fondamentale.

(1) Supponiamo di voler contare quante sono le possili permutazioni di n elementi (possibilmente non distinti).

Un caso familiare di tale situazione si ha quando si vogliono considerare tutti i possibili anagrammi di una parola in cui alcune lettere si ripetono. Si consideri, per esempio, la parola "PAPPA": in tale parola compaiono due lettere di cui una, la A, si ripete due volte e l'altra, la P, si ripete tre volte, per un totale di 5 elementi non distinti tra loro. Quanti sono gli anagrammi di tale parola?

Iniziamo con l'osservare che, per quanto visto sopra, il numero totale di permutazioni possibili delle 5 lettere di tale parola è $5! = 120$. Tuttavia, questo non

significa che esistono 120 possibili anagrammi di tale parola. Infatti, in queste 120 permutazioni troveremo ogni anagramma di PAPPA ripetuto tante volte quante sono le permutazioni che scambiano solo le A tra loro o solo le P tra loro, perché tali permutazioni non modificano l'anagramma. Quante sono per ogni anagramma dato queste permutazioni che non lo modificano? Abbiamo $2! = 2$ permutazioni delle due A tra loro e, per ciascuna di queste, $3! = 6$ permutazioni delle tre P tra loro. In totale $2!3! = 12$ permutazioni che non cambiano un anagramma dato.

Quindi, il numero totale di permutazioni trovate, 120, va diviso per il numero di queste permutazioni che non cambiano l'anagramma corrispondente, cioè 12. Ne deriva che il numero di anagrammi distinti della parola PAPPA è $120/12 = 10$:

PAPPA, AAPPP, PPPAA, APAPP, PAAPP, PAPAP, APPAP, PPAAP, APPPA, PPAPA.

In generale, supponiamo di avere s elementi distinti in cui il primo si ripete k_1 volte, il secondo k_2 volte, e così via fino all'ultimo che si ripete k_s volte, per un totale di $k_1 + k_2 + \cdots + k_s = n$ elementi non distinti tra loro (nell'esempio precedente sarebbe $s = 2$, $k_1 = 2$, $k_2 = 3$ e $n = k_1 + k_2 = 5$). Esattamente come per il ragionamento fatto sopra, le possibili permutazioni di questi n elementi distinti sono $n!$, ma questo numero va diviso per il numero di permutazioni che non cambia veramente la configurazione trovata, che sono le permutazioni che scambiano i k_1 elementi che si ripetono tra loro, i k_2 che si ripetono tra loro, e così via fino all'ultimo gruppo di k_s elementi. Esattamente come abbiamo visto nel nostro esempio, tali permutazioni sono $k_1!k_2!\cdots k_s!$ (per ognuna delle $k_1!$ permutazioni del primo gruppo di elementi uguali, ne abbiamo $k_2!$ del secondo gruppo, $k_3!$ del terzo etc.).

In conclusione, il numero di permutazioni di n elementi in cui c'è un gruppo di k_1 elementi uguali, un altro di k_2 elementi uguali, e così via fino a k_s elementi uguali, è

$$\frac{n!}{k_1!k_2!\cdots k_s!} \tag{2.25}$$

dove $k_1 + k_2 + \cdots + k_s = n$. Queste permutazioni sono dette **permutazioni con ripetizioni** per distinguerle dalle permutazioni usuali in cui gli elementi sono tutti distinti, dette anche **permutazioni semplici**. Questo numero è anche detto **coefficiente polinomiale**.[15]

[15] Il nome deriva dal fatto che questo è il coefficiente del temine $x_1^{k_1}\cdots x_s^{k_s}$ nel polinomio $(x_1+\cdots+x_s)^n$.

Si noti che se gli elementi fossero tutti distinti senza ripetizioni, allora $k_1! = k_2! = \cdots = k_s! = 1$ e nella (2.25) rimarrebbe $n!$, che coincide col numero di permutazioni di n elementi distinti.

(2) Un'altra importante domanda a cui il calcolo combinatorio risponde è la seguente: dato un insieme X di n elementi e scelto un $k \leqslant n$, quanti sono i sottoinsiemi di X formati da k elementi?

Ad esempio, se $X = \{a, b, c\}$ (quindi $n = 3$) e $k = 2$, i sottoinsiemi di X con 2 elementi sono

$$\{a, b\},\ \{a, c\},\ \text{e}\ \{b, c\}\,.$$

Si noti che dal momento che un sottoinsieme è determinato dai suoi elementi indipendentemente dall'ordine in cui li mettiamo, il problema può essere riformulato nel modo seguente: in quanti modi diversi possiamo scegliere k elementi da un insieme di n (senza ripetizioni e senza tener conto dell'ordine)? Con una terminologia tipica del calcolo combinatorio, si dice anche che vogliamo sapere quante sono le **combinazioni semplici di n elementi di classe k.**

Rispondiamo ora alla domanda. Se X ha n elementi, e dobbiamo scegliere k elementi distinti tra questi n per formare un sottoinsieme S, per il primo elemento di S abbiamo n possibili scelte (uno qualunque degli elementi di X). Per ognuna di queste n scelte, il secondo elemento può essere scelto tra $n-1$ (tutti gli elementi di X tranne quello già scelto). Quindi per i primi due elementi abbiamo $n(n-1)$ possibili scelte. Per ognuna di queste, abbiamo $n-2$ possibilità per il terzo elemento (tutti gli elementi di X tranne i 2 già scelti), quindi $n(n-1)(n-2)$ possibilità per i primi 3 elementi. Iterando questo ragionamento risulta evidente che avremo $n(n-1)(n-2)(n-3)$ possibili scelte per i primi 4 elementi, e in generale saranno $n(n-1)(n-2)\cdots(n-k+1)$ scelte per k elementi.

Tuttavia, questo numero non è ancora la risposta corretta: ad esempio, per $n = 3$ e $k = 2$ come nell'esempio di sopra, otterremmo $n(n-1) = 3 \cdot 2 = 6$, mentre abbiamo visto che ci sono solo tre sottoinsiemi!

Questo perché, nel ragionamento appena fatto, ci sono scelte diverse del primo, del secondo etc. fino al k-esimo elemento che risultano nello stesso sottoinsieme. Ad esempio se in $\{a, b, c\}$ scegliamo a come primo e b come secondo, oppure b come primo a a come secondo, otteniamo chiaramente lo stesso sottoinsieme di due elementi $\{a, b\}$. In questo caso non ci sono altre scelte che danno questo stesso sottoinsieme, perché sono solo 2 le permutazioni possibili dei suoi due elementi.

In generale, lo stesso sottoinsieme di k elementi può essere ottenuto mediante esattamente $k!$ scelte diverse, ovvero tante quante sono le possibili permutazioni dei suoi elementi.

Concludiamo che, per avere il numero effettivo di sottoinsiemi di k elementi da un insieme di n, il numero $n(n-1)(n-2)\cdots(n-k+1)$ delle scelte che si possono fare va diviso per il numero $k!$ di scelte che in realtà danno gli stessi elementi disposti in ordine diverso. Quindi questo numero è

$$\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}.$$

Tale numero può essere riscritto in maniera compatta come segue. Moltiplicando numeratore e denominatore per $(n-k)(n-k-1)\cdots 2\cdot 1$, ovvero $(n-k)!$, si trova

$$\frac{n(n-1)(n-2)\cdots(n-k+1)(n-k)(n-k-1)\cdots 2\cdot 1}{k!(n-k)!}$$

Ma ora il numeratore è il prodotto di tutti i naturali da n fino ad arrivare a 1, ovvero $n!$. Quindi otteniamo

$$\frac{n!}{k!(n-k)!}. \tag{2.26}$$

L'espressione (2.26) appena scritta si denota con $\binom{n}{k}$, si chiama **coefficiente binomiale**[16] e si legge *n su k*.

Ad esempio, se $n=3$ e $k=2$, troviamo $\binom{n}{k}=\frac{3!}{2!1!}=\frac{6}{2}=3$, in accordo con l'esempio di sopra in cui abbiamo trovato esattamente 3 sottoinsiemi di due elementi in un insieme con 3 elementi.

Il coefficiente binomiale compare in molte importanti formule della matematica, quindi è importante conoscerne le proprietà. A tal fine dimostriamone due che, come vedremo tra poco, ci permettono di calcolare rapidamente i coefficienti binomiali al crescere di n senza usare la definizione.

Lemma 2.51. *I coefficienti binimiali soddisfano le seguenti proprietà:*

$$\binom{n}{k}=\binom{n}{n-k}, \tag{2.27}$$

$$\binom{n}{k}=\binom{n-1}{k}+\binom{n-1}{k-1}. \tag{2.28}$$

[16] In effetti, la (2.26) è un particolare caso del coefficiente polinomiale (2.25) quando $s=2$ (da cui binomiale). Infatti, in tal caso la (2.25) diventa $\frac{n!}{k_1!k_2!}$ con $k_1+k_2=n$, ovvero $k_2=n-k_1$. Quindi possiamo riscriverlo come $\frac{n!}{k_1!(n-k_1)!}$ che è, a parte il k_1 al posto di k, proprio la (2.26).

Dimostrazione. Per dimostrare la (2.27), basta scrivere il secondo membro in base alla definizione stessa di binomiale, cioè la (2.26) con $n-k$ al posto di k:

$$\binom{n}{n-k} = \frac{n!}{(n-k)![n-(n-k)]!} = \frac{n!}{(n-k)!(n-n+k)!} = \frac{n!}{(n-k)!k!} = \binom{n}{k}$$

Dimostriamo ora la (2.28). Per definizione di binomiale si ha

$$\begin{aligned}\binom{n-1}{k} + \binom{n-1}{k-1} &= \frac{(n-1)!}{k!(n-1-k)!} + \frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!} = \\ &= \frac{(n-1)!}{k!(n-1-k)!} + \frac{(n-1)!}{(k-1)!(n-k)!} \qquad (2.29)\end{aligned}$$

Per sommare le due frazioni, osserviamo che $k!(n-k)!$ è un multiplo comune dei denominatori. Infatti, tale numero si ottiene moltiplicando il primo denominatore per $n-k$ (che moltiplicato per $(n-1-k)!$ lo fa diventare $(n-k)!$), ma anche moltiplicando il secondo denominatore per k (che moltiplicato per $(k-1)!$ lo fa diventare $k!$). Usando allora $k!(n-k)!$ come multiplo comune, possiamo sommare le frazioni esattamente come nel caso della somma di due frazioni numeriche. Otteniamo che la (2.29) è uguale a

$$\frac{(n-k)\cdot(n-1)! + k\cdot(n-1)!}{k!(n-k)!},$$

ovvero, mettendo in evidenza $(n-1)!$ nei due addendi a numeratore,

$$\frac{[(n-k)+k]\cdot(n-1)!}{k!(n-k)!} = \frac{n\cdot(n-1)!}{k!(n-k)!} = \frac{n!}{k!(n-k)!} = \binom{n}{k}$$

proprio come volevamo dimostrare. □

Come abbiamo accennato, la (2.28) ci permette di calcolare velocemente i coefficienti binomiali al crescere di n. Incominciamo col disporre in riga, per $n = 0,1,2,\ldots,$ i coefficienti binomiali $\binom{n}{k}$ al variare di $k = 0,1,\ldots,n$. Ovvero scriviamo la tabella

	$(k=0)$	$(k=1)$	$(k=2)$	$(k=3)$	$(k=4)$	$(k=5)$
$(n=0)$	$\binom{0}{0}$					
$(n=1)$	$\binom{1}{0}$	$\binom{1}{1}$				
$(n=2)$	$\binom{2}{0}$	$\binom{2}{1}$	$\binom{2}{2}$			
$(n=3)$	$\binom{3}{0}$	$\binom{3}{1}$	$\binom{3}{2}$	$\binom{3}{3}$		
$(n=4)$	$\binom{4}{0}$	$\binom{4}{1}$	$\binom{4}{2}$	$\binom{4}{3}$	$\binom{4}{4}$	
$(n=5)$	$\binom{5}{0}$	$\binom{5}{1}$	$\binom{5}{2}$	$\binom{5}{3}$	$\binom{5}{4}$	$\binom{5}{5}$
			$\dots$			

La (2.28) ci dice che l'entrata $\binom{n}{k}$ della tabella (che si trova all'incrocio di riga n-esima e colonna k-esima) si ottiene sommando l'entrata sopra di lei e quella che si trova a sinistra di quest'ultima. Infatti l'entrata sopra si trova nella stessa colonna, la k-esima, ma una riga sopra, la $(n-1)$-esima (corrisponde quindi al coefficiente binomiale $\binom{n-1}{k}$). Allo stesso modo, l'entrata alla sua sinistra è sempre nella riga sopra, la $(n-1)$-esima, e nella colonna più a sinistra, la $(k-1)$-esima (cioè corrisponde al coefficiente $\binom{n-1}{k-1}$).

Con questa regola si ricostruiscono rapidamente tutte le entrate della tabella conoscendo le prime due righe (in quanto $\binom{0}{0} = \binom{1}{0} = \binom{1}{1} = 1$):

	$(k=0)$	$(k=1)$	$(k=2)$	$(k=3)$	$(k=4)$	$(k=5)$
$(n=0)$	1					
$(n=1)$	1	1				
$(n=2)$	1	2	1			
$(n=3)$	1	3	3	1		
$(n=4)$	1	**4**	**6**	4	1	
$(n=5)$	1	5	**10**	10	5	1
			$\dots$			

Abbiamo messo in evidenza in grassetto il fatto che il 10 dell'ultima riga è stato ottenuto sommando il 6 sopra di lui più il 4 a sinistra del 6.

La tabella appena scritta si dice anche *triangolo di Pascal-Tartaglia*.

I coefficienti ottenuti servono, tra le altre cose, per calcolare le potenze $(a+b)^n$

di un binomio. Infatti, vale la formula

$$(a+b)^n = \\ = \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \cdots + \binom{n}{n-1}ab^{n-1} + \binom{n}{n}b^n .$$

In altre parole vale il seguente risultato.

Proposizione 2.52. *Per ogni $n \in \mathbb{N}$ vale la seguente formula:*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k}b^k . \tag{2.30}$$

Qui il simbolo $\sum$, simbolo di sommatoria, indica che dobbiamo sommare tutti gli addendi del tipo $\binom{n}{k}a^{n-k}b^k$ con k che varia da 0 a n.

Ad esempio, applicando questa formula e guardando i coefficienti della tabella si trova

$$\begin{gathered}(a+b)^2 = a^2 + 2ab + b^2 \\ (a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3 \\ (a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4 \\ \vdots\end{gathered}$$

La (2.30) può essere dimostrata per induzione nel modo seguente.

Dimostrazione della Proposizione 2.52. Innanzitutto, tale formula è vera per $n = 0$, in quanto, da una parte $(a+b)^0 = 1$, e dall'altra, si ha $\sum_{k=0}^{0}\binom{0}{k}a^{0-k}b^k = \binom{0}{0}a^0b^0 = 1$.

Supponiamo ora che la formula valga per n (ipotesi induttiva), ovvero

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k}b^k \tag{2.31}$$

e dimostriamo che vale per $n+1$, ovvero

$$(a+b)^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k}b^k. \tag{2.32}$$

Poiché $(a+b)^{n+1} = (a+b)(a+b)^n$, possiamo usare l'ipotesi induttiva e scrivere

$$\begin{aligned}(a+b)^{n+1} =& (a+b)\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k = \\ =& a\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k + b\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k. \end{aligned} \tag{2.33}$$

Ora, possiamo portare l'a e il b che moltiplicano rispettivamente la prima e la seconda sommatoria all'interno delle sommatorie stesse, in quanto in generale per le sommatorie vale[17] la proprietà $c\sum_{j=1}^{N} x_j = \sum_{j=1}^{N} cx_j$. Si ottiene quindi

$$(a+b)^{n+1} = \sum_{k=0}^{n}\binom{n}{k}a^{n+1-k}b^k + \sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^{k+1}.$$

Per motivi che saranno chiari nei passaggi successivi, separiamo ora dalla prima sommatoria il termine corrispondente a $k=0$, cioè $\binom{n}{0}a^{n+1-0}b^0 = a^{n+1}$, e dalla seconda sommatoria il termine che corrisponde a $k=n$, cioè $\binom{n}{n}a^{n-n}b^{n+1} = b^{n+1}$:

$$(a+b)^{n+1} = a^{n+1} + \sum_{k=1}^{n}\binom{n}{k}a^{n+1-k}b^k + \sum_{k=0}^{n-1}\binom{n}{k}a^{n-k}b^{k+1} + b^{n+1}.$$

Ora utilizziamo un trucco spesso usato quando si ha a che fare con sommatorie: nella seconda sommatoria denotiamo $k+1$ con h (ovvero $k = h-1$) e usiamo h come nuovo indice di sommatoria. Quando $k=0$, si ha $h=1$ e quando $k=n-1$ si ha $h=n$. Cambiati in questo modo gli estremi della sommatoria e sostituito ovunque in tale sommatoria $k = h-1$, si ha

$$(a+b)^{n+1} = a^{n+1} + \sum_{k=1}^{n}\binom{n}{k}a^{n+1-k}b^k + \sum_{h=1}^{n}\binom{n}{h-1}a^{n-h+1}b^h + b^{n+1}.$$

A questo punto, dal momento che l'indice di una sommatoria può essere denotato in qualunque modo, ridenominiamo h come k:

$$(a+b)^{n+1} = a^{n+1} + \sum_{k=1}^{n}\binom{n}{k}a^{n+1-k}b^k + \sum_{k=1}^{n}\binom{n}{k-1}a^{n-k+1}b^k + b^{n+1}.$$

[17] Infatti, si ha $c\sum_{j=1}^{N} x_j = c(x_1 + x_2 + \cdots + x_N) = cx_1 + cx_2 + \cdots + cx_N = \sum_{j=1}^{N} cx_j$, dove abbiamo applicato la proprietà distributiva nella seconda uguaglianza.

Dopo queste trasformazioni, in entrambe le sommatorie l'indice k varia sempre da 1 a n e in entrambe compare $a^{n-k+1}b^k$, moltiplicato nella prima per $\binom{n}{k}$ e nella seconda per $\binom{n}{k-1}$. Quindi possiamo raggrupparle, usando la proprietà[18] generale delle sommatorie $\sum_{j=1}^{N} x_j + \sum_{j=1}^{N} y_j = \sum_{j=1}^{N}(x_j + y_j)$ e mettendo in evidenza $a^{n-k+1}b^k$:

$$(a+b)^{n+1} = a^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k} + \binom{n}{k-1} \right] a^{n+1-k}b^k + b^{n+1}.$$

A questo punto, all'interno della parentesi quadra possiamo usare la proprietà (2.28) del coefficiente binomiale, che ci dice che $\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$, ottenendo

$$(a+b)^{n+1} = a^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} a^{n+1-k}b^k + b^{n+1}.$$

Ora, possiamo inglobare di nuovo gli addendi a^{n+1} e b^{n+1}, che nei passaggi precedenti avevamo separato, all'interno della sommatoria centrale. Notiamo che $a^{n+1} = \binom{n+1}{0} a^{n+1-0}b^0$ e $b^{n+1} = \binom{n+1}{n+1} a^{n+1-(n+1)}b^{n+1}$. Quindi perchè nella sommatoria siano compresi anche a^{n+1} e b^{n+1} basta aggiungere $k = 0$ e $k = n+1$, ottenendo finalmente

$$(a+b)^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k}b^k.$$

Abbiamo quindi dimostrato la (2.32) partendo dall'ipotesi che fosse vera la (2.31), quindi come afferma il principio d'induzione la formula è vera per ogni $n \in \mathbb{N}$. □

(3) Un altro problema tipico del calcolo combinatorio è il seguente: dato un insieme X con n elementi, e scelto un $k \leqslant n$, in quanti modi possiamo scegliere k distinti elementi da X tenendo conto dell'ordine? In altre parole vogliamo contare le k-uple composte da elementi distinti (diversamente da quanto visto in (2) per i sottoinsiemi di cardinalità k in cui l'ordine degli elementi non conta).

Ad esempio, se $X = \{a, b, c\}$ (cioè $n = 3$) e scegliamo $k = 2$, abbiamo le seguenti possibilità:

$$ab, ba, ac, ca, bc, cb.$$

[18] Infatti, usando le proprietà associativa e commutativa della somma, che ci permettono di permutare e accorpare gli addendi come vogliamo, si nota che $\sum_{j=1}^{N} x_j + \sum_{j=1}^{N} y_j = (x_1 + x_2 + \cdots + x_N) + (y_1 + y_2 + \cdots + y_N)$ è uguale a $(x_1 + y_1) + (x_2 + y_2) + \cdots + (x_N + y_N) = \sum_{j=1}^{N}(x_j + y_j)$.

In calcolo combinatorio si dice che vogliamo determinare il numero delle **disposizioni semplici di n elementi di classe k.**

La risposta a questa domanda è facile se si tiene conto di quanto già detto per risolvere il problema (2). In quel caso abbiamo visto che il numero di sottoinsiemi di k elementi presi da X è $n(n-1)(n-2)\cdots(n-k+1)$ diviso per $k!$ (il numero di permutazioni dei k elementi del sottoinsieme) perché ordinamenti diversi di questi elementi danno lo stesso sottoinsieme.

Dal momento che qui vogliamo invece tener conto dell'ordine, non dobbiamo dividere per $k!$ e il numero cercato è semplicemente

$$n(n-1)(n-2)\cdots(n-k+1).$$

Allo scopo di scrivere questo numero in una forma più conveniente, moltiplichiamolo e dividiamolo per $(n-k)(n-k-1)\cdots 2\cdot 1$ (ovvero $(n-k)!$). In questo modo otteniamo

$$\begin{aligned} &n(n-1)(n-2)\cdots(n-k+1) = \\ =&\frac{n(n-1)(n-2)\cdots(n-k+1)(n-k)(n-k-1)\cdots 2\cdot 1}{(n-k)!} = \frac{n!}{(n-k)!}. \end{aligned}$$

(4) Se nel problema (3) ammettiamo anche le ripetizioni otteniamo le cosiddette **disposizioni con ripetizione di n elementi di classe k.**

Ovvero, dato un insieme X con n elementi, e scelto un $k \leqslant n$, ci chiediamo in quanti modi possiamo scegliere k elementi da X, eventualmente con ripetizioni e tenendo conto dell'ordine.

Ad esempio, se $X = \{a, b, c\}$ (cioè $n = 3$) e $k = 2$, abbiamo le seguenti possibilità:

$$aa, bb, cc, ab, ba, ac, ca, bc, cb.$$

In questo caso, la risposta è particolarmente semplice: il primo elemento può essere scelto in n modi (può essere un qualunque elemento di X), il secondo anche (non dobbiamo escludere il primo già scelto perché sono ammesse ripetizioni) e così via per tutti gli altri: quindi abbiamo $n \cdot n \cdots \cdot n = n^k$ possibilità, tenendo conto anche dell'ordine.

Si osservi che nello scegliere k elementi da X ammettendo ripetizioni e tenendo conto dell'ordine, stiamo equivalentemente considerando tutte le k-uple ordinate di elementi di X, ovvero l'insieme prodotto cartesiano X^k di X per se stesso k volte. Ad esempio per l'insieme $X = \{a, b, c\}$ considerato sopra, si ha

$$X^2 = \{(a,a),(b,b),(c,c),(a,b),(b,a),(a,c),(c,a),(b,c),(c,b)\}$$

ed è chiaro che dare le coppie ordinate o le disposizioni con ripetizione è sostanzialmente la stessa cosa.

(5) Infine, consideriamo le cosiddette **combinazioni con ripetizione di n elementi di classe k**: rispetto alle combinazioni semplici, ovvero i sottoinsiemi di k elementi, studiate in (2), ammettiamo anche le ripetizioni. Più precisamente, dato un insieme X con n elementi, e scelto un $k \leqslant n$, ci chiediamo in quanti modi possiamo scegliere k elementi da X, eventualmente con ripetizioni, ma sempre senza tener conto dell'ordine come nelle combinazioni semplici.

Ad esempio, se $X = \{a,b,c\}$ (cioè $n = 3$) e $k = 2$, abbiamo le seguenti possibilità:

$$aa, bb, cc, ab, ac, bc.$$

Notiamo che dare una di queste combinazioni significa semplicemente dire il numero di volte che si ripete a, il numero di volte che si ripete b e il numero di volte che si ripete c, con la condizione che la somma di questi tre numeri deve essere 2, perché in tutto dobbiamo avere solo 2 elementi. Questi numeri possono essere anche 0, ad esempio in aa, a si ripete 2 volte, b zero volte e c zero volte; in bc, a si ripete 0 volte, b una volta e c una volta.

Possiamo rappresentare allora ognuna di queste combinazioni come uno schema del tipo

$$* * \cdots * \circ * * \cdots * \circ * * \cdots *$$

dove il numero di asterischi della prima serie ci dice quante volte si ripete a, il numero di asterischi della seconda serie (dopo il primo cerchietto) ci dice quante volte si ripete b, e il numero di asterischi della terza serie (quelli dopo il secondo cerchietto) ci dice quante volte si ripete c.

Ad esempio, ab è rappresentato da $* \circ *\circ$ (il primo asterisco ci dice che a si ripete una volta, il secondo che b si ripete una volta, e l'assenza di asterischi dopo il secondo cerchietto significa che c si ripete 0 volte, ovvero non compare nella combinazione). O ancora, cc è rappresentato da $\circ\circ **$ (nessun asterisco prima del primo cerchietto significa che a si ripete 0 volte, nessun asterisco tra il primo e il secondo cerchietto significa che b si ripete 0 volte, 2 asterischi dopo il secondo cerchietto significano che c si ripete 2 volte).

In generale, se l'insieme X ha n elementi, avremo n serie di asterischi, separate da $n-1$ cerchietti, e il numero totale di asterischi deve essere k.

La nostra domanda equivale allora a chiedere: quanti possibili schemi di questo tipo con k asterischi e $n-1$ cerchietti possiamo formare?

Tali schemi saranno tanti quanti i modi diversi di permutare gli asterischi e i cerchietti tra loro a meno di permutazioni che lasciano lo schema invariato. Avendo in tutto $k+n-1$ elementi tra asterischi e cerchietti, le loro permutazioni sono $(k+n-1)!$. Dobbiamo però dividere per il numero di permutazioni che lasciano invariato uno schema dato, ovvero quelle che permutano i cerchietti tra loro e gli asterischi tra loro. Essendoci k asterischi, le permutazioni che li scambiano sono $k!$; mentre essendoci $n-1$ cerchietti, le permutazioni che li scambiano sono $(n-1)!$. Quindi dobbiamo dividere il numero totale di permutazioni per $k!(n-1)!$, ottenendo

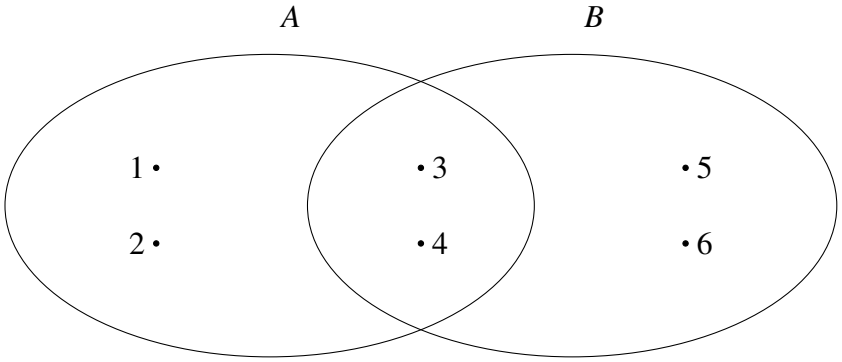
$$\frac{(k+n-1)!}{k!(n-1)!} = \binom{k+n-1}{k}.$$

Le formule appena viste possono essere riassunte nel seguente schema

	con ripetizioni	senza ripetizioni
ordine	n^k	$\frac{n!}{(n-k)!}$
no ordine	$\binom{n+k-1}{k}$	$\binom{n}{k}$

Concludiamo questa parte dedicata alla combinatoria con un altro problema di calcolo del numero di elementi di un certo insieme. Con le formule precedenti abbiamo imparato a contare il numero di elementi dell'insieme potenza $P(X)$ di un insieme finito dato e il numero di elementi del prodotto cartesiano X^n di un insieme finito dato per se stesso (n volte). Concludiamo ora il capitolo con una formula utile per il calcolo del numero di elementi di unioni e intersezioni di insiemi: il cosiddetto **principio di inclusione/esclusione**.

Per iniziare a illustrare tale principio, supponiamo di avere due insiemi finiti A e B e di voler calcolare il numero $|A \cup B|$ di elementi dell'unione $A \cup B$ in funzione del numero $|A|$ di elementi di A e del numero $|B|$ di elementi di B. Come si vede nel disegno seguente, se $A = \{1,2,3,4\}$ e $B = \{3,4,5,6\}$



non vale l'uguaglianza $|A \cup B| = |A| + |B|$, in quanto sommando $|A|$ e $|B|$ stiamo contando due volte gli elementi 3 e 4 dell'intersezione. Per ottenere un'uguaglianza dobbiamo quindi sottrarre il numero di elementi che stiamo contando più volte. La formula corretta diventa

$$|A \cup B| = |A| + |B| - |A \cap B|. \tag{2.34}$$

Usando questa semplice formula e uguaglianze elementari di teoria degli insiemi, possiamo vedere anche senza l'aiuto di disegni cosa succede se consideriamo l'unione $A \cup B \cup C$ di tre insiemi A, B, C. Infatti, per associatività dell'unione possiamo intendere l'unione $A \cup B \cup C$ come unione $(A \cup B) \cup C$ dei due insiemi $A \cup B$ e C. Per la formula (2.34) abbiamo allora

$$|A \cup B \cup C| = |A \cup B| + |C| - |(A \cup B) \cap C|. \tag{2.35}$$

Il primo addendo al secondo membro può essere riscritto usando di nuovo la formula (2.34), da cui

$$|A \cup B \cup C| = |A| + |B| - |A \cap B| + |C| - |(A \cup B) \cap C|. \tag{2.36}$$

Per quello che riguarda l'ultimo addendo, in base alla commutatività dell'intersezione a alla distributività dell'intersezione rispetto all'unione si ha $(A \cup B) \cap C = C \cap (A \cup B) = (C \cap A) \cup (C \cap B)$ e quindi possiamo applicare anche all'ultimo addendo della (2.36) la formula (2.35):

$$|(A \cup B) \cap C| = |(C \cap A) \cup (C \cap B)| = |C \cap A| + |C \cap B| - |(C \cap A) \cap (C \cap B)|.$$

Tenendo conto che $(C \cap A) \cap (C \cap B) = A \cap B \cap C$ scriviamo

$$|(A \cup B) \cap C| = |(C \cap A) \cup (C \cap B)| = |C \cap A| + |C \cap B| - |A \cap B \cap C|.$$

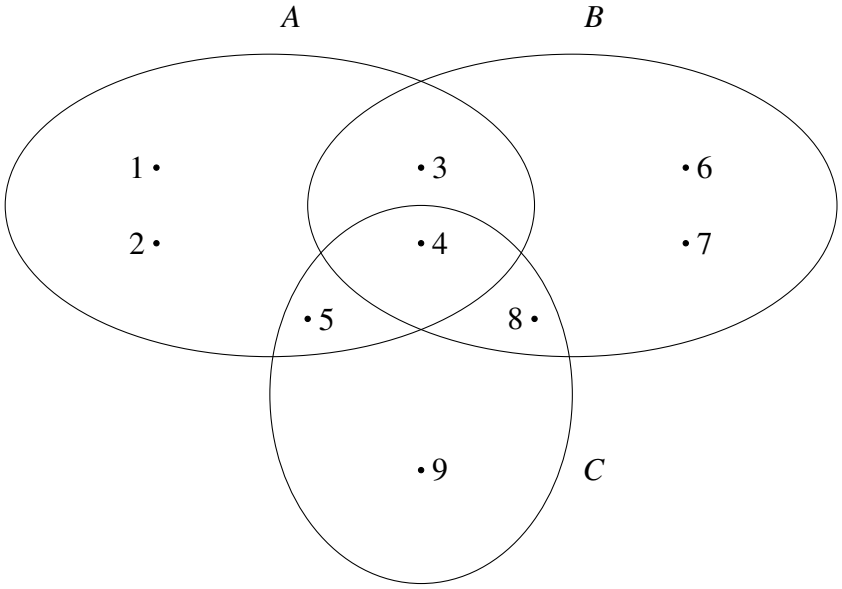
Sostituendo questa espressione nella (2.36) otteniamo allora

$$|A \cup B \cup C| = |A| + |B| - |A \cap B| + |C| - [|C \cap A| + |C \cap B| - |A \cap B \cap C|]$$

ovvero, riordinando gli addendi,

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|. \quad (2.37)$$

Questa formula mostra che per contare il numero di elementi dell'unione di tre insiemi dobbiamo sommare le cardinalità dei tre insiemi, escludere tutte le possibili intersezioni di questi insiemi a due a due (che sono state contate due volte nella somma $|A|+|B|+|C|$ e poi reincludere però l'intersezione di tutti e tre gli insiemi, che è stata esclusa una volta di troppo quando abbiamo sottratto $|A \cap B|$, $|A \cap C|$ e $|B \cap C|$. Il seguente disegno, in cui $A = \{1,2,3,4,5\}$, $B = \{3,4,6,7,8\}$, $C = \{4,5,8,9\}$, può essere utile a illustrare quello che sta succedendo.



A questo punto ci si potrebbe chiedere se è possibile ricavare una formula per la cardinalità dell'unione $A_1 \cup A_2 \cup \cdots \cup A_n$ di n insiemi, con n qualunque. La risposta è affermativa e la formula è la seguente:

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}| \,. \quad (2.38)$$

La formula esprime il fatto che per calcolare la cardinalità dell'unione bisogna sommare le cardinalità degli insiemi dati, sottrarre le cardinalità delle intersezioni di due

insiemi (tutte quelle possibili), sommare le cardinalità delle intersezioni di tre insiemi (tutte quelle possibili), sottrarre le cardinalità delle intersezioni di quattro insiemi (tutte quelle possibili), e così via, a segni alterni, includendo ed escludendo (da cui il nome di principio di inclusione/esclusione). La formula (2.38) può essere dimostrata per induzione.

Esempio 2.53. Vediamo un esempio di applicazione delle formule appena viste. Supponiamo che i calzini contenuti in un cassetto possano essere grossi, bianchi e lunghi. Sapendo che il cassetto contiene 7 calzini grossi, che i calzini bianchi e grossi sono 2, quelli lunghi e grossi 3, quelli contemporaneamente grossi, bianchi e lunghi 2, e che ci sono 5 calzini che sono bianchi o lunghi, quanti sono i calzini in totale?
Se denotiamo B l'insieme dei calzini bianchi, L l'insieme dei calzini lunghi e G l'insieme dei calzini grossi, il numero totale dei calzini è dato da $|G \cup B \cup L|$, che come sappiamo dal principio di inclusione/esclusione è

$$|G \cup B \cup L| = |G| + |B| + |L| - |B \cap L| - |B \cap G| - |G \cap L| + |B \cap G \cap L|.$$

In base ai dati del problema, abbiamo $|G| = 7$, $|B \cap G| = 2$, $|G \cap L| = 3$, $|G \cap B \cap L| = 2$ e $|B \cup L| = 5$, quindi

$$|G \cup B \cup L| = 7 + |B| + |L| - |B \cap L| - 2 - 3 + 2 = |B| + |L| - |B \cap L| + 4.$$

I dati del problema non ci danno le quantità $|B|$, $|L|$, $|B \cap L|$, ma ci danno $|B \cup L| = 5$, che in base al principio di inclusione/esclusione (caso di due insiemi) è uguale esattamente alla somma $|B| + |L| - |B \cap L|$ che compare nell'espressione di $|G \cup B \cup L|$. Sostituendo si ha allora

$$|G \cup B \cup L| = |B| + |L| - |B \cap L| + 4 = |B \cup L| + 4 = 5 + 4 = 9.$$

2.10 L'insieme delle permutazioni come gruppo

Definizione 2.54. Per ogni numero naturale n, le funzioni biiettive dall'insieme $\{1, 2, \ldots, n\}$ in se stesso si dicono **permutazioni di n elementi**.

Questo nome è dovuto al fatto che una tale funzione deve assegnare a ogni numero tra 1 e n un numero tra 1 e n evitando ripetizioni (altrimenti non sarebbe iniettiva) e in modo che tutti compaiano come immagini (altrimenti non sarebbe suriettiva): quindi, essa non fa altro che "cambiare l'ordine" degli elementi $1, 2, \ldots, n$. Ad esempio, se $n = 3$, un esempio di funzione biiettiva da $\{1, 2, 3\}$ in se stesso si ottiene ponendo

$$f(1) = 3$$
$$f(2) = 1$$
$$f(3) = 2$$

oppure

$$f(1) = 3$$
$$f(2) = 2$$
$$f(3) = 1$$

che scambia 3 e 1 tra loro lasciando fisso 2.

Si noti che anche la funzione identica su $\{1, 2, \ldots, n\}$, essendo chiaramente una funzione biiettiva, è una permutazione in base alla definizione data.

Definizione 2.55. In generale, dato un insieme $A = \{a_1, a_2, \ldots a_n\}$ con n elementi, chiameremo **permutazione di** A (o **permutazione di** $a_1, a_2, \ldots, a_n$) una funzione biiettiva $\{1, 2, \ldots, n\} \longrightarrow A$.

Ad esempio se $A = \{a, b, c\}$, la funzione $f\colon \{1, 2, 3\} \longrightarrow \{a, b, c\}$ definita da

$$f(1) = b$$
$$f(2) = a$$
$$f(3) = c$$

è una permutazione di a, b, c.

Quante sono le possibili permutazioni di n elementi $a_1, a_2, \ldots, a_n$? Ricordiamo che abbiamo già calcolato questo numero all'inizio della Sezione 2.9 rispondendo alla domanda a pagina 63. Aveveamo quindi trovato che le scelte possibili come immagini di $1, 2, \ldots, n$ che diano una funzione biiettiva sono $n!$.
In questa sezione vogliamo considerare l'insieme delle permutazioni di $\{1, 2, \ldots, n\}$, che da ora denoteremo con S_n, assieme all'operazione di composizione di funzioni, e studiarne le proprietà; non diversamente da quello che si fa quando si considera ad esempio l'insieme dei numeri naturali $\mathbb{N}$, assieme a un'operazione (ad esempio la somma o il prodotto) e se ne studiano le proprietà.

Prima di fare ciò, dobbiamo però essere sicuri che quando componiamo due permutazioni otteniamo ancora una permutazione. Non avrebbe infatti senso studiare le proprietà di questa operazione su S_n se succedesse che quando la applichiamo rischiamo di ottenere qualcosa che non sta più in S_n, allo stesso modo in cui non ha senso per esempio

studiare le proprietà della sottrazione su $\mathbb{N}$ quando non sempre questa operazione si può fare in questo insieme (pur essendo 2 e 5 numeri naturali, 2-5 non è un numero naturale).

Dal momento che una permutazione per definizione è una funzione biiettiva da $X = \{1, 2, \ldots, n\}$ in sè, basta ricordare che per il Lemma 2.44 la composizione di due funzioni biiettive è ancora biiettiva. Ora che siamo sicuri che la composizione di permutazioni sia ancora una permutazione, e quindi che l'operazione di composizione sia ben definita sull'insieme S_n delle permutazioni, possiamo studiarne le proprietà.
Abbiamo già visto che la composizione di funzioni è associativa, si veda la (2.11). Sappiamo inoltre che in generale la composizione non è commutativa, e non è difficile mostrare un esempio di permutazioni che non commutano, ad esempio sull'insieme $\{1, 2, 3\}$ se definiamo f e g come segue

$$\begin{array}{ccc} 1 & \mapsto & 2 \\ 2 & \mapsto & 1 \\ 3 & \mapsto & 3 \end{array}$$
$$f$$

$$\begin{array}{ccc} 1 & \mapsto & 3 \\ 2 & \mapsto & 2 \\ 3 & \mapsto & 1 \end{array}$$
$$g$$

allora si ha

$$\begin{array}{ccc} 1 & \mapsto & 2 \\ 2 & \mapsto & 3 \\ 3 & \mapsto & 1 \end{array}$$
$$g \circ f$$

mentre

$$\begin{array}{ccc} 1 & \mapsto & 3 \\ 2 & \mapsto & 1 \\ 3 & \mapsto & 2 \end{array}$$
$$f \circ g$$

Un'altra proprietà della composizione sull'insieme S_n è che in S_n esiste un elemento neutro per tale operazione. Come abbiamo visto a pagina 50, l'elemento neutro per la composizione è la funzione identica id_X, che appartiene all'insieme S_n delle permutazioni di $X = \{1, 2, \ldots, n\}$ in quanto biiettiva. Per semplicità di notazione, da questo momento in poi denoteremo la funzione identica semplicemente con id.

Infine, essendo ogni permutazione una funzione $f\colon X \longrightarrow X$ biiettiva, essa, come abbiamo visto nella Proposizione 2.41, è invertibile, ovvero esiste una funzione $g\colon X \longrightarrow X$

tale che $g \circ f = f \circ g = id$.
Chiaramente, anche la funzione g inversa di f è biiettiva, perché le uguaglianze $g \circ f = f \circ g = id$ ci dicono che anche lei è invertibile (ha f come sua inversa). Quindi anche g, l'inversa di f, che da questo momento denoteremo con f^{-1}, è una permutazione di X.
In altre parole S_n contiene l'inversa di ogni sua permutazione.

Le proprietà che stiamo mettendo in evidenza rientrano in quelle elencate nella seguente, importantissima

Definizione 2.56. Un **gruppo** è un insieme G dotato di un'operazione (che denotiamo con il simbolo $\cdot$) per cui valgono le tre seguenti proprietà.

1. (**Associatività.**) Per ogni $g_1, g_2, g_3 \in G$ si ha $(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$.
2. (**Elemento neutro.**) Esiste un elemento neutro, ovvero esiste un $u \in G$ tale che $g \cdot u = u \cdot g = g$ per ogni $g \in G$.
3. (**Inverso.**) Ogni $g \in G$ ha un inverso in G, ovvero per ogni $g \in G$ esiste $g^{-1} \in G$ tale che $g \cdot g^{-1} = g^{-1} \cdot g = u$

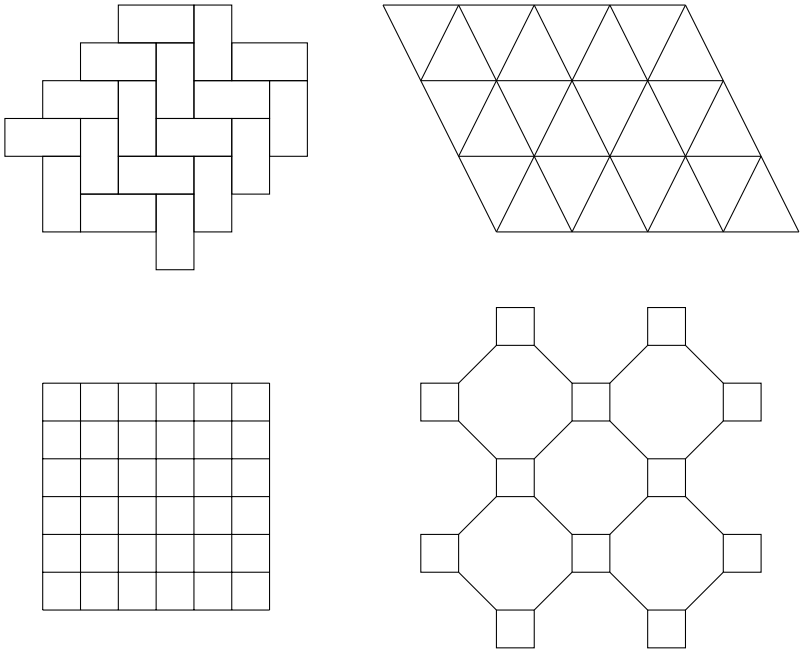
Si usa anche scrivere che $(G, \cdot)$ è un gruppo.

Definizione 2.57. Un gruppo $(G, \cdot)$ si dice **abeliano** o **commutativo** se $\cdot$ gode della proprietà commutativa, cioè se per due qualunque elementi $g_1, g_2 \in G$ vale $g_1 \cdot g_2 = g_2 \cdot g_1$.

Quindi $(S_n, \circ)$ (l'insieme delle permutazioni dotato dell'operazione di composizione) è un gruppo (non abeliano).

Esempio 2.58. Un altro esempio di gruppo, stavolta abeliano, è $(\mathbb{Z}, +)$, ovvero l'insieme degli interi dotato dell'operazione di somma: l'elemento neutro è il numero 0, e l'inverso di ogni numero intero n è semplicemente $-n$.
Invece, $(\mathbb{Z}, \cdot)$, cioè sempre l'insieme degli interi ma dotato della moltiplicazione, non è un gruppo. Pur essendo valide la proprietà associativa e l'esistenza dell'elemento neutro (in questo caso il numero 1), non è vero che ogni elemento di $\mathbb{Z}$ ha un inverso rispetto alla moltiplicazione (ad esempio, l'inverso di 2 è $\frac{1}{2}$, che è un razionale non intero).

Esempi di gruppi si trovano numerosi in ogni campo della matematica e nelle sue applicazioni. Ad esempio, in geometria, data una figura nel piano l'insieme delle trasformazioni che lasciano la figura invariata mandandola in se stessa (ad esempio, dato un quadrato la rotazione di 90 gradi attorno al centro del quadrato manda il quadrato in sè) forma un gruppo, detto *gruppo di simmetria della figura*. Questa definizione si estende anche a figure non limitate, ad esempio i motivi geometrici illlimitati, spesso usati nelle decorazioni o in architettura, che ricoprono in maniera regolare il piano per ripetizione di un motivo. Si consideri ad esempio il seguente disegno



Queste sono figure infinite con un loro gruppo di simmetria, e grazie alla teoria dei gruppi si può dimostrare che esistono solo 17 possibili gruppi di simmetria di tali motivi. Infine, la nozione di gruppo si rivela fondamentale in tutta la fisica contemporanea: grazie alla teoria dei gruppi si è riusciti a prevedere l'esistenza di particelle elementari che poi sono state effettivamente osservate.

Mettiamo ora in evidenza un'altra importante differenza tra il gruppo delle permutazioni $(S_n, \circ)$ e quello degli interi $(\mathbb{Z}, +)$. In $\mathbb{Z}$ ogni elemento può essere scritto usando solo il numero 1 o il suo inverso -1 e l'operazione $+$:

$$2 = 1 + 1, 3 = 1 + 1 + 1, 4 = 1 + 1 + 1 + 1, \ldots$$

$$-2 = (-1) + (-1), -3 = (-1) + (-1) + (-1), \ldots$$

$$0 = 1 + (-1)$$

In questo senso, quindi, è sufficiente un elemento a *generare* tutto $\mathbb{Z}$. Questo è un caso particolare della seguente

Definizione 2.59. Dato un gruppo $(G, \cdot)$, si dice che $g_1, g_2, \ldots, g_s$ **generano** G (o **sono generatori di** G) se ogni elemento di G può essere scritto come prodotto finito dei g_i e dei loro inversi.

Il gruppo $(\mathbb{Z}, +)$ ha quindi la caratteristica di essere generato da un solo elemento, 1. I gruppi che godono di questa proprietà son detti **ciclici**.

Si ha invece che il gruppo S_n non è ciclico, a meno che $n = 2$ (in questo caso, infatti, S_2 contiene solo l'identità id e la permutazione f che scambia 1 e 2, che genera id in quanto $f \circ f = id$).

Troveremo ora un insieme di generatori del gruppo S_n. Più precisamente, chiamiamo **trasposizioni** le permutazioni di S_n che scambiano tra loro due elementi e lasciano fissi tutti gli altri. Ad esempio, in S_4, la permutazione

$$1 \mapsto 1$$

$$2 \mapsto 3$$

$$3 \mapsto 2$$

$$4 \mapsto 4$$

è una trasposizione (scambia tra loro 2 e 3 lasciando fissi 1 e 4).

Mostreremo ora che qualunque permutazione può essere scritta come composizione di trasposizioni. Da questo si deduce che le trasposizioni generano il gruppo S_n.

Per far ciò, procederemo in modo costruttivo. Iniziamo col mostrare su un esempio come effettivamente, data una permutazione, si può trovare la sua decomposizione in trasposizioni. Consideriamo la seguente permutazione σ in S_6:

$$1 \mapsto 4$$

$$2 \mapsto 1$$

$$3 \mapsto 6$$

$$4 \mapsto 2$$

$$5 \mapsto 5$$

$$6 \mapsto 3$$

Procediamo come segue: la permutazione manda 1 in 4, 4 in 2, e 2 in 1. Questi tre elementi vengono quindi permutati tra loro in quello che si chiama un **ciclo** di lunghezza 3.

Usiamo la notazione $(1\ 4\ 2)$ per indicare tale ciclo (ogni numero che appare nel ciclo viene mandato nel successivo, e l'ultimo viene rimandato nel primo).

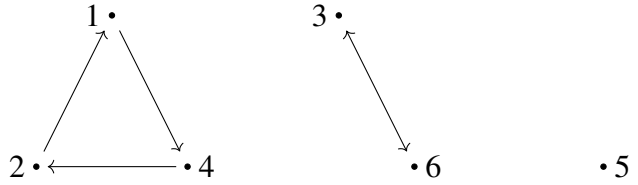
Concentriamoci ora sul primo elemento rimasto fuori da questo ciclo, cioè 3. Questo viene mandato in 6, e 6 viene mandato in 3. Quindi questi due elementi vengono permutati tra loro in un ciclo di lunghezza 2, che seguendo la notazione di sopra denotiamo con $(3\ 6)$ (si noti che un ciclo di lunghezza 2 è una trasposizione).

Infine, l'ultimo elemento che rimane fuori dai cicli trovati è 5, che viene fissato dalla

permutazione, quindi non appartiene a nessun ciclo[19].
La nostra permutazione è quindi scrivibile come *composizione di cicli disgiunti*:

$$\sigma = (3\ 6)(1\ 4\ 2) \tag{2.39}$$

(da questo momento per facilità di notazione ometteremo spesso il simbolo di composizione) e agisce sugli elementi $1, 2, 3, 4, 5, 6$ come nel disegno seguente.



Quanto visto sopra può essere generalizzato ad una qualunque permutazione per ottenere il seguente risultato.

Proposizione 2.60. *Ogni permutazione è la composizione di cicli disgiunti.*

Idea della dimostrazione. Sia $\sigma \in S_n$ una permutazione. Consideriamo l'immagine dell'elemento 1, cioè l'elemento $\sigma(1)$. Questo viene mandato da σ in un altro elemnto di $\{1, \dots, n\}$ potenzialmente diverso. Stiamo quindi considerando l'elemento $\sigma \circ \sigma(1)$. Iterando questo procedimento otterremo nuovamente l'elemento 1 dopo un certo numero k_1 di passaggi. In altre parole avremo che $\big(1,\ \sigma(1),\ \sigma(\sigma(1)),\ \dots, \sigma^{k_1}(1)\big)$ è un ciclo. Ora se $k_1 = n$ abbiamo dimostrato la tesi. In caso contrario, cioè se $k_1 < n$, consideriamo il primo elemento m in $\{1, \dots, n\}$ che non appartiene al ciclo appena trovato. Con lo stesso ragionamento otteniamo un ciclo $\big(m, \sigma(m), \dots, \sigma^{k_m}(m)\big)$. Dato che n è un numero finito, dopo al più n passaggi otterremo una decomposizione di σ in cicli. Per dimostrare che i cicli così ottenuti sono effettivamente disgiunti dobbiamo usare il fatto che una permutazione è una funzione biiettiva. □

Ora, il nostro obiettivo è decomporre ogni permutazione in trasposizioni. Se riusciamo a mostrare che ogni ciclo può essere decomposto in trasposizioni avremo raggiunto lo scopo. In effetti, dato un ciclo $(a_1\ a_2\ \cdots\ a_k)$ di lunghezza k, si può dimostrare che si ha sempre

$$(a_1\ a_2\ \cdots\ a_{k-1}\ a_k) = (a_1\ a_k)(a_1\ a_{k-1}) \cdots (a_1\ a_2). \tag{2.40}$$

Non diamo una dimostrazione generale di tale fatto: limitiamoci a illustrarlo con un esempio. Il ciclo $(1\ 4\ 2)$ in base alla (2.40) si dovrebbe decomporre come $(1\ 2)(1\ 4)$. Verifichiamolo esplicitamente: ricordando che le permutazioni si applicano da destra a sinistra, si ha che $(1\ 2)(1\ 4)$ agisce su 1,2,4 come segue:

[19] O, volendo, possiamo interpretarlo come un ciclo di lunghezza 1 e denotarlo (5), ma in generale lo si omette.

$$\begin{array}{ccccc} 1 & \mapsto & 4 & \mapsto & 4 \\ 2 & \mapsto & 2 & \mapsto & 1 \\ 4 & \mapsto & 1 & \mapsto & 2 \\ & (1\,4) & & (1\,2) & \end{array}$$

cioè esattamente come il ciclo $(1\ 4\ 2)$ che manda $1 \mapsto 4$, $4 \mapsto 2$, $2 \mapsto 1$.
In conclusione, la (2.39) può essere riscritta come

$$\sigma = (3\ 6)(1\ 2)(1\ 4)$$

ovvero come prodotto di trasposizioni.

Quello che abbiamo mostrato è un modo per determinare la decomposizione di una permutazione in trasposizioni, tuttavia si può vedere che tale decomposizione non è in generale unica. Ad esempio, è facile verificare che la permutazione

$$\begin{array}{ccc} 1 & \mapsto & 3 \\ 2 & \mapsto & 4 \\ 3 & \mapsto & 1 \\ 4 & \mapsto & 2 \end{array}$$

si decompone sia come $(1\ 3)(2\ 4)$ che come $(1\ 4)(1\ 2)(4\ 3)(1\ 4)$.
Tuttavia, vale il seguente risultato.

Teorema 2.61. *Il numero di trasposizioni in cui si decompone una permutazione data è o sempre pari o sempre dispari.*

Alla luce di ciò, possiamo dare la seguente

Definizione 2.62. Una permutazione si dice **pari** (risp. **dispari**) se si decompone in un numero pari (risp. dispari) di trasposizioni.

L'identità *id* è una permutazione pari in quanto può essere realizzata come prodotto di una qualunque trasposizione τ per se stessa. Infatti $\tau \circ \tau = id$ perché τ scambia due elementi tra loro lasciando gli altri fissi, e riapplicando τ gli unici due elementi invertiti vengono di nuovo scambiati tornando alla situazione iniziale.

Osserviamo che se componiamo due permutazioni σ e σ' entrambe pari, il risultato sarà ancora una permutazione pari. Questo perché se la prima si decompone in $2k$ trasposizioni

$$\sigma = \tau_1 \tau_2 \cdots \tau_{2k}$$

e la seconda in $2l$ trasposizioni

$$\sigma' = \tau'_1 \tau'_2 \cdots \tau'_{2l},$$

allora si ha chiaramente

$$\sigma\sigma' = \tau_1\tau_2\cdots\tau_{2k}\tau'_1\tau'_2\cdots\tau'_{2l}.$$

Cioè $\sigma\sigma'$ si decompone come prodotto di $2k + 2l = 2(k + l)$ trasposizioni, ed è quindi anche lei una permutazione pari.

Inoltre, l'inversa σ^{-1} di una permutazione pari σ, sarà una permutazione pari. Infatti, se $\sigma = \tau_1\tau_2\cdots\tau_{2k-1}\tau_{2k}$, è facile vedere che la sua inversa è

$$\sigma^{-1} = \tau_{2k}\tau_{2k-1}\cdots\tau_2\tau_1, \tag{2.41}$$

cioè il prodotto delle stesse trasposizioni ma nell'ordine inverso (quindi sempre un numero pari).

Questo è un caso particolare del seguente risultato, che ci dice come calcolare l'inversa di un prodotto in un gruppo qualunque.

Lemma 2.63. *Sia G un gruppo e siano $g_1, g_2, \ldots, g_k \in G$. Allora*

$$(g_1g_2\cdots g_k)^{-1} = g_k^{-1}\cdots g_2^{-1}g_1^{-1} \tag{2.42}$$

Prima di dimostrare il lemma, osserviamo che da esso segue subito la (2.41). Infatti, in base al lemma si ha che l'inversa di $\tau_1\tau_2\cdots\tau_{2k-1}\tau_{2k}$ è $\tau_{2k}^{-1}\tau_{2k-1}^{-1}\cdots\tau_2^{-1}\tau_1^{-1}$. Ma come abbiamo già osservato sopra, per qualunque trasposizione τ si ha $\tau \circ \tau = id$, che significa che τ ha come inversa se stessa, ovvero $\tau^{-1} = \tau$. Quindi $\tau_{2k}^{-1}\tau_{2k-1}^{-1}\cdots\tau_2^{-1}\tau_1^{-1} = \tau_{2k}\tau_{2k-1}\cdots\tau_2\tau_1$, che dimostra la (2.41).

Dimostrazione del Lemma 2.63. In base alla definizione di inversa basta vedere se moltiplicando $g_1g_2\cdots g_{k-1}g_k$ con $g_k^{-1}g_{k-1}^{-1}\cdots g_2^{-1}g_1^{-1}$ si ottiene l'elemento neutro:

$$(g_1g_2\cdots g_{k-1}g_k)(g_k^{-1}g_{k-1}^{-1}\cdots g_2^{-1}g_1^{-1}) =$$

(per associatività dell'operazione in un gruppo, possiamo disporre le parentesi come vogliamo)

$$= (g_1g_2\cdots g_{k-1})(g_kg_k^{-1})(g_{k-1}^{-1}\cdots g_2^{-1}g_1^{-1}) = \tag{2.43}$$

Ma la parentesi centrale $g_kg_k^{-1}$ è uguale all'elemento neutro u, quindi la (2.43) si riduce a

$$= (g_1g_2\cdots g_{k-1})u(g_{k-1}^{-1}\cdots g_2^{-1}g_1^{-1}) = \tag{2.44}$$

ovvero, tenendo conto della definizione di elemento neutro,

$$= (g_1g_2\cdots g_{k-1})(g_{k-1}^{-1}\cdots g_2^{-1}g_1^{-1}). \tag{2.45}$$

A questo punto possiamo nello stesso modo eliminare g_{k-1}: infatti, sempre per associatività la (2.45) si scrive

$$= (g_1 g_2 \cdots)(g_{k-1} g_{k-1}^{-1})(\cdots g_2^{-1} g_1^{-1}) = \tag{2.46}$$

e come prima la parentesi centrale si elimina in quanto $g_{k-1} g_{k-1}^{-1} = u$. Continuando così, si eliminano via via tutti i fattori del prodotto, fino a che non rimane solo $g_1 g_1^{-1} = u$. La dimostrazione è conclusa[20]. □

Quanto detto mostra in effetti che l'insieme delle permutazioni pari in S_n, che si denota A_n, può essere considerato un gruppo a sè, in quanto quando compongo due elementi di A_n rimango in A_n (cioè l'operazione di composizione è ben definita dentro A_n), l'elemento neutro sta dentro A_n e l'inversa di ogni elemento di A_n sta ancora dentro A_n.

Definizione 2.64. Un gruppo G' contenuto in un gruppo G si dice **sottogruppo di** G (si intende che devono essere gruppi rispetto alla stessa operazione).

Quindi A_n è un sottogruppo di S_n. Notiamo che invece il sottoinsieme di S_n costituito dalle permutazioni dispari non forma un sottogruppo: basterebbe già il fatto che l'elemento neutro *id* (che come abbiamo osservato sopra è una permutazione pari) non sta in tale sottoinsieme. Inoltre possiamo notare che il prodotto di due permutazioni dispari non è più dispari in quanto se $\sigma = \tau_1 \tau_2 \cdots \tau_{2k+1}$ e $\sigma' = \tau'_1 \tau'_2 \cdots \tau'_{2l+1}$, allora $\sigma\sigma' = \tau_1 \tau_2 \cdots \tau_{2k+1} \tau'_1 \tau'_2 \cdots \tau'_{2l+1}$ risulta essere prodotto di $(2k+1) + (2l+1) = 2(k+l+1)$ trasposizioni, cioè un numero pari.

Osservazione 2.65. Un'importante applicazione delle proprietà dei gruppi A_n e S_n in matematica è stata la dimostrazione del fatto che *non esiste una formula risolutiva generale per risolvere le equazioni di grado superiore al quarto*. In altre parole, non esiste una formula generale che permetta di trovare le radici di un polinomio di grado $d \geqslant 5$. La dimostrazione, che fa parte della cosiddetta teoria di Galois, fa uso del fatto che a ogni polinomio di grado n si può associare un gruppo, opportunamente definito, che permuta le sue radici, quindi si può pensare come un sottogruppo del gruppo delle permutazioni S_n. La risolubilità dell'equazione determinata da tale polinomio corrisponde a una particolare proprietà di questi gruppi di permutazioni. Si dimostra che S_n e A_n non soddisfano questa proprietà se $n \geqslant 5$.

Le decomposizione in cicli disgiunti di una permutazione σ è importante anche per ottenere un'altra informazione: l'*ordine di* σ.
Il fatto che S_n sia finito ha come conseguenza il fatto che per ogni permutazione $\sigma \in S_n$ esista un numero naturale k per cui $\sigma^k = id$, dove con σ^k intendiamo la composizione

[20] Dovremmo verificare anche che $(g_k^{-1} g_{k-1}^{-1} \cdots g_2^{-1} g_1^{-1})(g_1 g_2 \cdots g_{k-1} g_k) = u$, ma i calcoli sono analoghi.

di σ con se stessa k volte. Ad esempio, consideriamo in S_3 il ciclo $(1\ 2\ 3)$, ovvero la permutazione

$$\begin{array}{ccc} 1 & \mapsto & 2 \\ 2 & \mapsto & 3 \\ 3 & \mapsto & 1 \end{array}$$

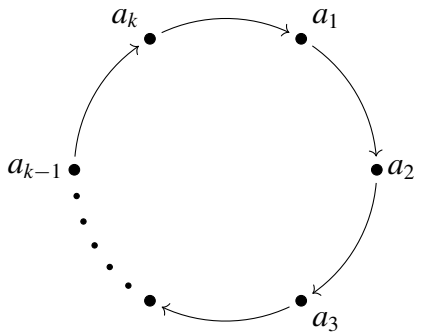
Si vede allora che $\sigma^3 = id$, ovvero che componendo σ tre volte con se stessa si ottiene l'identità:

$$\begin{array}{ccccccc} 1 & \mapsto & 2 & \mapsto & 3 & \mapsto & 1 \\ 2 & \mapsto & 3 & \mapsto & 1 & \mapsto & 2 \\ 3 & \mapsto & 1 & \mapsto & 2 & \mapsto & 3 \\ & \sigma & & \sigma & & \sigma & \end{array}$$

Chiaramente, se continuare ad applicare σ per la quarta volta sarebbe come comporre id con σ, ovvero $\sigma^4 = \sigma$, componendo ancora una volta otterremo $\sigma^5 = \sigma^2$ e componendo per la sesta volta avremmo $\sigma^6 = \sigma^3$, ovvero $\sigma^6 = id$. Iterando questo argomento, è facile vedere che in realtà vale $\sigma^k = id$ per tutti i multipli di 3, quindi ci sono in realtà infiniti numeri naturali per cui $\sigma^k = id$. Chiamiamo ordine il più piccolo numero naturale positivo per cui vale questa uguaglianza: nel nostro caso esso è chiaramente 3. In generale si può dare la seguente

Definizione 2.66. Sia G un gruppo con elemento neutro u e sia g un suo elemento. Si dice **ordine** di g il più piccolo intero positivo k tale che $g^k = u$.

Ora impareremo a calcolare l'ordine di qualunque permutazione σ senza dover applicare σ con se stessa fino a ottenere l'identità.
Per prima cosa, osserviamo che un ciclo $\sigma = (i_1\ i_2\ \ldots\ i_k)$ di lunghezza k ha ordine esattamente k.
Di ciò ci si può facilmente convincere guardando la rappresentazione grafica del ciclo:



Come si vede, ogni applicazione del ciclo manda un qualunque elemento nell'elemento successivo del ciclo (nel verso indicato dalla freccia), quindi applicando k volte il ciclo ogni elemento viene rimandato in se stesso.

A questo punto, per calcolare l'ordine di una generica permutazione σ, basta usare la sua decomposizione come prodotto di cicli disgiunti. Supponiamo che σ si decomponga in m cicli $C_1, C_2, \ldots, C_m$ rispettivamente di lunghezze $L_1, L_2, \ldots, L_m$. Essendo i cicli disgiunti, quando componiamo σ con se stessa ogni ciclo è composto con se stesso. Il primo ciclo C_1 ci darà l'identità se iteriamo la permutazione un numero di volte pari a un qualunque multiplo della sua lunghezza L_1; analogamente, il secondo ciclo C_2 ci darà l'identità se iteriamo la permutazione un numero di volte pari a un qualunque multiplo della sua lunghezza L_2, e così via per tutti i cicli. Quindi, avremo che tutti i cicli ci danno contemporaneamente l'identità se e solo se iteriamo la permutazione un numero di volte k che sia un multiplo comune di tutte le lunghezze $L_1, L_2, \ldots, L_m$. Dalla definizione otteniamo che *l'ordine della permutazione è il più piccolo di questi multipli comuni, ovvero il minimo comune multiplo di* $L_1, L_2, \ldots, L_m$.

Esempio 2.67. Consideriamo la permutazione σ di S_5 data da

$$\sigma(1) = 3,\ \sigma(2) = 5,\ \sigma(3) = 4,\ \sigma(4) = 1,\ \sigma(5) = 2.$$

Si vede subito che la permutazione si decompone in prodotto di cicli disgiunti come $\sigma = (1\ 3\ 4)(2\ 5)$. Poiché abbiamo due cicli di lunghezze due e tre, l'ordine della permutazione sarà il minimo comune multiplo di 2 e 3 ovvero 6. In altre parole, il più piccolo intero positivo k per cui $\sigma^k = id$ è $k = 6$. Per verificarlo direttamente, calcoliamo le potenze successive di σ.

		σ		σ^2		σ^3		σ^4		σ^5		σ^6
1	$\mapsto$	3	$\mapsto$	4	$\mapsto$	1	$\mapsto$	3	$\mapsto$	4	$\mapsto$	1
2	$\mapsto$	5	$\mapsto$	2	$\mapsto$	5	$\mapsto$	2	$\mapsto$	5	$\mapsto$	2
3	$\mapsto$	4	$\mapsto$	1	$\mapsto$	3	$\mapsto$	4	$\mapsto$	1	$\mapsto$	3
4	$\mapsto$	1	$\mapsto$	3	$\mapsto$	4	$\mapsto$	1	$\mapsto$	3	$\mapsto$	4
5	$\mapsto$	2	$\mapsto$	5	$\mapsto$	2	$\mapsto$	5	$\mapsto$	2	$\mapsto$	5

Come si vede, 1, 3 e 4 vengono rimandati in sè già da σ^3. Ma per quella potenza 2 e 5 non vengono mandati in se stessi perché questo accade solo per le potenze pari (cioè multiple di 2). Quindi vengono rimandati in sè contemporaneamente 1, 3, 4 e anche 2, 5 dopo il primo multiplo comune di 2 e 3, ovvero 6.

Come secondo esempio, consideriamo la seguente permutazione di S_{10}

$$\sigma(1) = 5,\ \sigma(2) = 7,\ \sigma(3) = 6,\ \sigma(4) = 10,\ \sigma(5) = 2,$$
$$\sigma(6) = 4,\ \sigma(7) = 1,\ \sigma(8) = 9,\ \sigma(9) = 3,\ \sigma(10) = 8$$

La sua decomposizione in prodotto di cicli disgiunti è $\sigma = (1\ 5\ 2\ 7)(3\ 6\ 4\ 10\ 8\ 9)$. Avendo due cicli di lunghezze rispettivamente 4 e 6, l'ordine della permutazione è il loro minimo comune multiplo, ovvero 12.

Capitolo 3

Interi e aritmetica modulare

In questo capitolo lavoreremo con i numeri interi. Nella prima parte richiameremo le proprietà fondamentali dell'aritmetica. Questo ci consentirà, nella seconda parte, di introdurre la cosiddetta aritmetica modulare, fondamentale in molte applicazioni pratiche.

3.1 Somma, prodotto e algoritmo della divisione

Iniziamo con la seguente

Definizione 3.1. Sia A un insieme dotato di due operazioni che denotiamo con $+$ e $\cdot$. Diremo che $(A, +, \cdot)$ è un **anello** se soddisfa le seguenti proprietà.

(1) *La somma è associativa*, ovvero per ogni $a, b, c \in A$ si ha

$$(a + b) + c = a + (b + c).$$

(2) *La somma è commutativa*, ovvero per ogni $a, b \in A$ si ha

$$a + b = b + a.$$

(3) *Esiste un elemento neutro per la somma* (denotato con 0), tale che per ogni $a \in A$ si ha

$$a + 0 = 0 + a = a.$$

(4) *Ogni elemento ammette un inverso rispetto alla somma* (che chiameremo il suo *inverso additivo* o *opposto*), ovvero per ogni $a \in A$ esiste un $b \in A$ tale che

$$a + b = b + a = 0$$

che verrà indicato con $b = -a$.

(5) *Il prodotto è associativo*, ovvero per ogni $a, b, c \in A$ si ha

$$(ab)c = a(bc).$$

(6) *Esiste un elemento neutro per il prodotto* (denotato con 1), tale che per ogni $a \in A$ si ha

$$a1 = 1a = a.$$

(7) *Vale la proprietà distributiva del prodotto rispetto alla somma*, ovvero per ogni $a, b, c \in A$ si ha

$$a(b + c) = ab + ac, \quad (a + b)c = ac + bc.$$

Se inoltre *Il prodotto è commutativo*, ovvero per ogni $a, b \in A$ si ha $ab = ba$, allora $(A, +, \cdot)$ è detto **anello commutativo**.

Chiaramente l'insieme $\mathbb{Z}$ dei numeri interi soddisfa tutte le proprietà nella Definizione 3.1 ed è quindi un anello commutativo. Si noti che negli interi non vale la proprietà analoga alla (4) per il prodotto. Cioè non è vero che ogni elemento $a \in \mathbb{Z}$ ha inverso rispetto al prodotto (non in $\mathbb{Z}$), che dovrebbe essere un $b \in \mathbb{Z}$ tale che $ab = ba = 1$. Ad esempio, non esiste nessun numero intero b tale che $2b = 1$.
Tale proprietà è invece verificata per tutti gli $a \neq 0$ nell'insieme $\mathbb{Q}$ dei numeri razionali.

Definizione 3.2. Un anello commutativo in cui ogni elemento diverso da 0 ha un inverso rispetto al prodotto si dice **campo**.

L'insieme dei razionali $\mathbb{Q}$ è quindi un esempio di campo (ad esempio, 2 ha come inverso moltiplicativo $\frac{1}{2}$ in $\mathbb{Q}$). Un altro esempio è dato dall'insieme $\mathbb{R}$ dei numeri reali e da quello $\mathbb{C}$ dei numeri complessi.
Nella matematica e nelle sue applicazioni si incontrano vari esempi di anelli diversi dall'insieme degli interi $\mathbb{Z}$ (ne vedremo altri nei prossimi paragrafi e capitoli). Lavorare con questa definizione astratta permette di dimostrare in piena generalità risultati e formule che poi saranno validi in ogni anello, senza doverli dimostrare caso per caso (esattamente come già visto con la nozione di gruppo). Ad esempio il Lemma 2.63, dimostrato per un gruppo qualsiasi, è valido per ogni gruppo che incontreremo.

Osservazione 3.3. Il fatto che l'anello $\mathbb{Z}$ si estenda al campo $\mathbb{Q}$, cioè il fatto che gli interi non zero, come razionali, ammettano (in $\mathbb{Q}$) un inverso moltiplicativo, ha un importante conseguenza. Questa è la cosiddetta *legge di annullamento del prodotto*:

$$ab = 0 \Rightarrow a = 0 \vee b = 0$$

cioè se un prodotto è nullo allora almeno uno dei due fattori deve essere nullo[1].
Possiamo dimostrare facilmente questa legge. Sia $ab = 0$ e supponiamo per assurdo che a e b siano entrambi diversi da zero. Considerando gli interi come particolari razionali, e quindi l'uguaglianza $ab = 0$ come un'uguaglianza tra razionali, possiamo moltiplicare entrambi i membri per $\frac{1}{a}$ (che esiste nei razionali sotto l'ipotesi $a \neq 0$). Così troviamo da una parte $\frac{1}{a}(ab) = \frac{1}{a}0 = 0$ e dall'altra

$$\frac{1}{a}(ab) = (\frac{1}{a}a)b = 1b = b\,.$$

Ne deduciamo $b = 0$, che contraddice l'ipotesi che a e b fossero entrambi nulli.

Sottolineiamo che tale proprietà non va data per scontata. Infatti vedremo, nei paragrafi e capitoli successivi, alcuni anelli in cui essa non vale. D'altra parte risulta chiaro dalla discussione precedente che questa proprietà è sempre soddisfatta in un campo.

Un operazione fondamentale definita nell'anello degli interi $\mathbb{Z}$, e che useremo continuamente in questo capitolo, è *la divisione con resto*. Più precisamente, vale il seguente risultato.

Teorema 3.4. *Siano $a, b \in \mathbb{Z}$, con $b \neq 0$. Allora esiste un'unica coppia (q, r) di interi tali che valgono le due seguenti condizioni:*

$$a = bq + r \tag{3.1}$$

$$0 \leqslant r < |b| \tag{3.2}$$

(dove $|b|$ indica il cosiddetto valore assoluto di b, uguale a b se $b \geqslant 0$ e a $-b$ se $b < 0$; in altre parole se necessario si cambia il segno di b per renderlo non negativo). L'intero q si dice **quoziente** *e l'intero r* **resto** *della divisione.*

Esempio 3.5. Siano $a = 7$, $b = 2$. Si ha

$$7 = 2 \cdot 3 + 1$$

quindi quoziente e resto della divisione sono rispettivamente $q = 3$ e $r = 1$. Si noti che ad esempio si ha anche

$$7 = 2 \cdot 2 + 3$$

oppure

$$7 = 2 \cdot 4 + (-1)$$

[1] Proprietà usata spesso quando si risolvono le equazioni. Ad esempio per risolvere $x^2 - x = 0$ osservo che $x^2 - x = x(x-1)$, e quindi la mia equazione si riscrive $x(x-1) = 0$. Per la legge di annullamento del prodotto questa è verificata se $x = 0$ o $x - 1 = 0$.

ma queste due uguaglianze, pur essendo del tipo (3.1), non verificano la (3.2). Questo perché nella prima il resto 3 non è minore di $b = 2$, nella seconda il resto -1, pur essendo minore di $|b| = 2$ non è maggiore o uguale a zero.
Questo esempio mostra che affinché quoziente e resto siano unici e determinati da a e b, la condizione (3.2) è necessaria.

Osservazione 3.6. Come enunciato nel Teorema 3.4, a e b possono essere anche negativi. Ad esempio se $a = -7$ e $b = 2$, si ha

$$-7 = 2 \cdot (-4) + 1 \,.$$

Oppure se $a = 7$ e $b = -2$ si ha

$$7 = (-2) \cdot (-3) + 1 \,.$$

Un'applicazione importante dell'algoritmo della divisione è la possibilità di *scrivere un qualunque numero intero positivo in qualsiasi base* $b \geqslant 2$.
Prima di dare la definizione precisa, osserviamo che quando scriviamo un qualunque numero in notazione decimale, ad esempio $n = 1375$, le cifre $1, 3, 7, 5$ indicano (a partire da destra a sinistra) quante sono le unità, quante le decine, quante le centinaia etc., ovvero

$$1375 = 1 \cdot 1000 + 3 \cdot 100 + 7 \cdot 10 + 5 \,.$$

Usando la notazione usuale per le potenze

$$1375 = 1 \cdot 10^3 + 3 \cdot 10^2 + 7 \cdot 10^1 + 5 \,.$$

Quindi qualunque numero non negativo può essere scritto come somma di potenze decrescenti di 10 (l'ultima, a moltiplicare 5, sarebbe $10^0 = 1$ e non la scriviamo) precedute da un coefficiente che può essere un qualunque numero tra 0 e 9 (ovvero un qualunque numero minore di 10 e maggiore o uguale a zero). Vedremo ora che, usando l'algoritmo della divisione, la stessa cosa può essere fatta con un qualunque altro numero naturale maggiore o uguale di 2 al posto di 10.

Teorema 3.7. *Sia $b \geqslant 2$ un intero fissato. Allora ogni intero positivo n può essere scritto in modo unico nella forma seguente*

$$n = r_N b^N + r^{N-1} b^{N-1} + \cdots + r_1 b + r_0 \tag{3.3}$$

con $0 \leqslant r_i < b$ per tutti gli $i = 0, 1, \ldots, N$ (e $r_N \neq 0$). Si scrive allora $n = (r_N\, r_{N-1} \cdots r_1\, r_0)_b$ e si dice che $(r_N\, r_{N-1} \cdots r_1\, r_0)_b$ è la scrittura di n in base b.

Dimostrazione. Come abbiamo anticipato, per dimostrare il teorema si usa l'algoritmo della divisione. Iniziamo col dividere n per b ottenendo

$$n = q_0 b + r_0 \,. \tag{3.4}$$

Se q_0 è minore di b, abbiamo finito: la (3.4) sarebbe già come nel teorema. Cioè una combinazione di potenze di b (in questo caso compaiono solo $b = b^1$ e, sottointesa, $b^0 = 1$ che moltiplica r_0) con coefficienti q_0 e r_0 minori di b (r_0 lo è sicuramente per definizione di resto), come prevede il Teorema 3.7.
Se invece $q_0 \geqslant b$, allora dividiamo q_0 per b:

$$q_0 = q_1 b + r_1 \tag{3.5}$$

e sostituiamo la (3.5) nella (3.4), ottenendo

$$n = (q_1 b + r_1) b + r_0 = q_1 b^2 + r_1 b + r_0 \,. \tag{3.6}$$

Di nuovo, se q_1 è minore di b, abbiamo finito: la (3.6) sarebbe proprio una combinazione di potenze di b (in questo caso compaiono b^2, $b^1 = b$ e il termine di grado zero) con coefficienti q_1, r_1, r_0 minori di b.
Se invece $q_1 \geqslant b$, iteriamo il procedimento. Dividiamo q_1 per b:

$$q_1 = q_2 b + r_2 \tag{3.7}$$

e sostituiamo la (3.7) nella (3.6), ottenendo

$$n = (q_2 b + r_2) b^2 + r_1 b + r_0 = q_2 b^3 + r_2 b^2 + r_1 b + r_0 \tag{3.8}$$

e se q_3 è minore di b, abbiamo finito.
Dal momento che i quozienti successivi sono sempre più piccoli, arriveremo sicuramente a un quoziente minore di b. A questo punto il procedimento si arresta e otteniamo la formula cercata.[2]. □

Osservazione 3.8. La condizione che il primo coefficiente r_N sia diverso da zero serve a evitare addendi inutili nella (3.3) e garantire l'unicità della scrittura in base b. Ad esempio, 1375 è anche uguale a $0 \cdot 10^4 + 1 \cdot 10^3 + 3 \cdot 10^2 + 7 \cdot 10^1 + 5$. Ovvero può esser scritto come 01375, ma chiaramente il primo 0 si può essere omesso.

Esempio 3.9. Scriviamo il numero $n = 19$ in base $b = 2$ con il procedimento descritto sopra: iniziamo dividendo 19 per 2:

$$19 = 9 \cdot 2 + 1 \,.$$

[2] Si noti che se avessimo ammesso come base $b = 1$ questo non sarebbe vero. Dividendo n per 1 si ottiene $n = 1 \cdot n + 0$ ovvero quoziente n, che può essere diviso per 1 rendendo sempre quoziente n, e il procedimento non si arresta mai.

Poichè il quoziente ottenuto $q_0 = 9$ è maggiore di 2, eseguiamo la seconda divisione:

$$9 = 4 \cdot 2 + 1$$

e sostituiamo

$$19 = (4 \cdot 2 + 1) \cdot 2 + 1 = 4 \cdot 2^2 + 1 \cdot 2 + 1 \,. \tag{3.9}$$

Questa non à ancora l'espressione di 19 in base 2 in quanto l'ultimo quoziente ottenuto, $q_1 = 4$, non è ancora minore di 2. Dividendo ulteriormente otteniamo

$$4 = 2 \cdot 2 + 0$$

e sostitutendo al posto di 4 nella (3.9) si ottiene

$$19 = (2 \cdot 2 + 0) \cdot 2^2 + 1 \cdot 2 + 1 = 2 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2 + 1 \,. \tag{3.10}$$

Non abbiamo ancora finito in quanto l'ultimo quoziente ottenuto, $q_2 = 2$, non è minore di 2. Eseguiamo allora un'ulteriore divisione per 2

$$2 = 1 \cdot 2 + 0 \,.$$

Ora l'ultimo quoziente ottenuto, $q_3 = 1$, è minore di 2. Sostituendo nella (3.10) si ottiene allora

$$19 = (1 \cdot 2 + 0) \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2 + 1 = 1 \cdot 2^4 + 0 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2 + 1 \tag{3.11}$$

che è l'uguaglianza cercata, che esprime 19 come combinazione di potenze di 2 con coefficienti minori di 2 (cioè 0 o 1). Possiamo anche scrivere che

$$19 = (1\,0\,0\,1\,1)_2$$

ovvero, più semplicemente, che 19 in base 2 si scrive come

$$10011 \,.$$

Osservazione 3.10. Si noti che, viceversa, dato un numero in una certa base, è facile verificare di quale numero si tratti in notazione decimale. Ad esempio, in base 2 il numero 1101 rappresenta, in base alla definizione data nella formula (3.3),

$$1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2 + 1 = 13 \,.$$

La numerazione in base $b = 2$ (o *numerazione binaria*) è di particolare importanza nell'informatica in quanto, esprimendo qualunque numero come successione di 0 e 1, consente a un computer di registrarlo o esprimerlo come successione di stati spento/acceso.

Storicamente, altri casi importanti sono $b = 16$ (numerazione *esadecimale*) o $b = 8$ (numerazione *ottale*). Tuttavia, come afferma il Teorema 3.7, si può usare qualunque $b \geqslant 2$. Ad esempio, se volessimo scrivere 23 in base 7, eseguiremmo la divisione

$$23 = 3 \cdot 7 + 2$$

e poiché il quoziente ottenuto $q_1 = 3$ è già minore della base $b = 7$, il procedimento si ferma qui. Il numero 23, in base 7, si scrive semplicemente come $(3\ 2)_7$.

3.2 Divisori e numeri primi

Se, dividendo un intero a per un intero b capita che il resto sia zero, ovvero che

$$a = b \cdot q$$

allora si dice che a **è un multiplo di** b, o equivalentemente che b **è un divisore di** a (o anche che b **divide** a). In formule, si scrive $b|a$.

Ad esempio, sia $a = 20$. È facile verificare[3] che i divisori di 20 sono

$$\pm 1, \pm 20, \pm 2, \pm 5, \pm 4, \pm 10\,.$$

Si noti che se b è un divisore di a, ovvero se $a = bq$ per qualche q, allora anche $-b$ è divisore di a in quanto vale chiaramente anche $a = (-b)(-q)$.
Si osservi inoltre che tra i divisori di $a \in \mathbb{Z}$ ci sono sicuramente ± 1 (in quanto $a = 1 \cdot a$ e $a = (-1) \cdot (-a)$) e $\pm a$ (in quanto $a = a \cdot 1$ e $a = (-a) \cdot (-1)$). I restanti divisori si chiamano **divisori propri**.
I numeri che non hanno divisori propri sono di fondamentale importanza nella matematica e in tutte le sue applicazioni.

Definizione 3.11. Un numero intero $p \neq \pm 1$ si dice **primo** se i suoi unici divisori sono ± 1 e $\pm p$.

Come vedremo, i numeri primi sono di grandissima importanza per gli scopi di questo corso, quindi è importante conoscerne le proprietà.
Ad esempio, una prima proprietà, che caratterizza i numeri primi, è la seguente:

$$p|ab \Rightarrow p|a \vee p|b,$$

ovvero se un numero primo divide un prodotto, esso divide necessariamente uno dei fattori.

[3] Vedremo dopo un metodo generale.

Per convincersi che tale proprietà non vale per numeri in generale, si noti ad esempio che 10 divide il prodotto $4 \cdot 15 = 60$ ma non è vero né che 10 divide 4 né che 10 divide 15. Il motivo è che essendo $10 = 2 \cdot 5$, esso divide $4 \cdot 15$ in quanto $2|4$ e $5|15$. Per un numero primo p che divida un prodotto ab una cosa del genere non è possibile in quanto non avendo divisori propri non può accadere che si scomponga in un prodotto in cui un fattore divide a e uno divida b.

La proprietà più importante dei numeri primi è però sicuramente quella espressa nel seguente risultato.

Teorema 3.12 (Teorema fondamentale dell'aritmetica)**.** *Dato un qualsiasi numero intero $a > 1$, esiste un unica decomposizione*

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s}$$

in primi positivi $p_1, p_2, \ldots, p_s$ tali che $p_1 > p_2 > \cdots > p_s$ e $\alpha_1, \alpha_2, \ldots, \alpha_s > 0$.

Il teorema afferma in un certo senso che i primi sono i "mattoni" di cui si compongono tutti i numeri interi.

Esempio 3.13. La decomposizione in primi di $a = 20$ è

$$20 = 5 \cdot 2^2 .$$

In questo caso si ha $s = 2$, $p_1 = 5$ e $p_2 = 2$ cioè 20 si decompone in un prodotto con fattori 2 e 5. Inoltre $\alpha_1 = 1$ e $\alpha_2 = 2$. Non esiste alcun altra decomposizione di 20 che soddisfi le condizioni del Teorema 3.12.

Osservazione 3.14. La decomposizione di un intero in fattori primi ci permette di trovare velocemente tutti i suoi divisori: essi saranno tutti quelli del tipo $b = \pm p_1^{\beta_1} p_2^{\beta_2} \cdots p_s^{\beta_s}$, con gli esponenti $0 \leqslant \beta_1 \leqslant \alpha_1$, $0 \leqslant \beta_2 \leqslant \alpha_2$, $\ldots$, $0 \leqslant \beta_s \leqslant \alpha_s$.
Ad esempio, per $20 = 5^1 \cdot 2^2$, i divisori positivi sono tutti e soli i numeri del tipo $5^{\beta_1} 2^{\beta_2}$ con $0 \leqslant \beta_1 \leqslant 1$ e $0 \leqslant \beta_2 \leqslant 2$, ovvero

$$5^0 2^0 = 1, \;\; 5^0 2^1 = 2, \;\; 5^0 2^2 = 4, \;\; 5^1 2^0 = 5, \;\; 5^1 2^1 = 10, \;\; 5^1 2^2 = 20 .$$

Dati due numeri interi a e b possiamo considerare i loro divisori in comune. Ad esempio, è facile verificare che $a = 30$ e $b = 45$ hanno come divisori comuni

$$\pm 1, \pm 3, \pm 5, \pm 15 .$$

Notiamo che tutti questi divisori comuni dividono 15 o, in altre parole, 15 è multiplo di tutti i divisori comuni. La seguente definizione estende questa osservazione al caso generale.

Definizione 3.15. Siano a e b due interi non entrambi nulli. Un intero positivo d si dice **massimo comune divisore di a e b** se

(1) d è un divisore comune di a e b (in simboli, $d|a \wedge d|b$) e

(2) ogni altro divisore comune di a e b divide d (in simboli, $c|a \wedge c|b \Rightarrow c|d$).

Osservazione 3.16. La condizione che a e b siano non entrambi nulli serve a garantire l'esistenza del massimo comune divisore: 0 ha come divisori tutti i numeri interi, in quanto $0 = n \cdot 0$ per ogni $n \in \mathbb{Z}$, quindi se $a = 0$ e $b = 0$ ogni intero sarebbe un divisore comune di a e b e non ce ne sarebbe uno massimo nel senso della Definizione 3.15[4].
Inoltre, la condizione che b sia positivo serve a garantire l'unicità. Per esempio, nell'esempio di sopra con $a = 30$ e $b = 45$ anche -15 è un divisore comune che è multiplo di ogni altro divisore comune.

Il massimo comune divisore di due interi a e b può essere calcolato semplicemente scrivendo prima tutti i divisori di a, poi tutti quelli di b (usando il metodo descritto nell'Osservazione 3.14) e poi guardando quale tra tutti i divisori comuni è multiplo degli altri.
Tuttavia, tale metodo non è molto efficace in quanto richiede la conoscenza della scomposizione in fattori primi di a e b, che non sempre è facile da trovare. Vedremo quindi un altro metodo, basato sull'algoritmo della divisione, che fornisce informazioni aggiuntive.

Il metodo funziona come segue. Si inizia dividendo a per b:

$$a = bq_1 + r_1 \,. \tag{3.12}$$

Si divide poi b per il resto r_1

$$b = r_1 q_2 + r_2 \tag{3.13}$$

ottenendo un secondo resto r_2. A questo punto si divide il primo resto per il secondo resto

$$r_1 = r_2 q_3 + r_3, \tag{3.14}$$

il secondo resto per il terzo resto

$$r_2 = r_3 q_4 + r_4 \tag{3.15}$$

e così via, fino a che non si ottiene resto zero:

$$\vdots$$

[4]Invece uno solo tra a e b può essere nullo, ad esempio se $a = 0$ e $b > 0$, i divisori comuni di a e b sono tutti i divisori di b, e il massimo sarebbe b stesso.

$$r_n = r_{n+1}q_{n+2} + r_{n+2} \tag{3.16}$$

$$r_{n+1} = r_{n+2}q_{n+3}\,. \tag{3.17}$$

Affermiamo che l'ultimo resto r_{n+2} non nullo di queste divisioni successive è esattamente il massimo comune divisore di a e b.
Prima di mostrare perché, vediamo alcuni esempi.

Esempio 3.17. Siano $a = 45$ e $b = 30$, come sopra. Dividendo a per b otteniamo

$$45 = 30 \cdot 1 + 15$$

e dividendo $b = 30$ per il resto $r_1 = 15$ otteniamo

$$30 = 15 \cdot 2$$

Quindi l'ultimo resto prima di ottenere resto zero è 15, che risulta essere come avevamo già detto sopra il massimo comune divisore.

Esempio 3.18. Consideriamo ora $a = 42$ e $b = 30$. Dividendo a per b otteniamo

$$42 = 30 \cdot 1 + 12\,.$$

Dividendo $b = 30$ per il resto $r_1 = 12$ si ha

$$30 = 12 \cdot 2 + 6$$

e dividendo $r_1 = 12$ per il secondo resto $r_2 = 6$ si ottiene

$$12 = 6 \cdot 2$$

cioè resto zero: l'ultimo resto non nullo, 6, è quindi il massimo comune divisore.

Esempio 3.19. Infine, siano $a = 120$ e $b = 23$. Procedendo come sopra, si ha

$$120 = 23 \cdot 5 + 5$$

$$23 = 5 \cdot 4 + 3$$

$$5 = 3 \cdot 1 + 2$$

$$3 = 2 \cdot 1 + 1$$

$$2 = 1 \cdot 2$$

L'ultimo resto non nullo è 1, che risulta essere quindi il massimo comune divisore.

L'ultimo esempio di sopra è un caso della seguente, importante

Definizione 3.20. Due interi a e b si dicono **primi tra loro** o **coprimi** se il loro massimo comune divisore è 1.

In altre parole, due interi sono coprimi se non hanno divisori comuni a parte il caso banale 1.

Vediamo ora perchè il procedimento per divisioni successive descritto e illustrato fornisce effettivamente il massimo comune divisore. Dobbiamo dimostrare che l'ultimo resto non nullo r_{n+2} che compare nella (3.16) è un divisore comune di a e b e che ogni altro divisore comune di a e b divide r_{n+2}.
Per dimostrare la prima affermazione, osserviamo che l'ultima divisione svolta, la (3.17) (quella con resto nullo), mostra che r_{n+2} divide r_{n+1} o, equivalentemente, che r_{n+1} è un multiplo di r_{n+2}. Sostituendola nella (3.16), si trova

$$r_n = r_{n+1}q_{n+2} + r_{n+2} = r_{n+2}q_{n+3}q_{n+2} + r_{n+2} = r_{n+2}(q_{n+3}q_{n+2} + 1) \tag{3.18}$$

che ci dice che r_{n+2} divide anche r_n o, equivalentemente, che r_n è un multiplo di r_{n+2}. Ora per brevità, riscriviamo la (3.18) come $r_n = r_{n+2}k$.
La divisione immediatamente precedente alla (3.16) sarà

$$r_{n-1} = r_n q_{n+1} + r_{n+1}\,. \tag{3.19}$$

Sostituendo sia $r_{n+1} = r_{n+2}q_{n+3}$ che $r_n = r_{n+2}k$ nella (3.19) otteniamo

$$r_{n-1} = r_n q_{n+1} + r_{n+1} = r_{n+2}kq_{n+1} + r_{n+2}q_{n+3} = r_{n+2}(kq_{n+1} + q_{n+3}) \tag{3.20}$$

che dimostra che r_{n+2} divide anche r_{n-1}. In altre parole, stiamo mostrando che r_{n+2} divide tutti resti r_{n+1}, r_n, r_{n-1}... delle divisioni effettuate, dall'ultima alla prima. Quando saremo arrivati alla (3.15) otterremo che r_{n+2} divide r_2, la (3.14) ci dirà che r_{n+2} divide r_1. Infine la (3.13) ci dirà che r_{n+2} divide b e la (3.12) ci dirà che r_{n+2} divide a. Quindi r_{n+2} è un divisore comune di a e b.
Per dimostrare che si tratta effettivamente del massimo comune divisore, basta mostrare che r_{n+2} è multiplo di qualunque divisore comune di a e b. Sia allora c un tale divisore comune: possiamo scrivere $a = ca'$ e $b = cb'$. Sostituendo queste due uguaglianze nella (3.12), otteniamo

$$ca' = cb'q_1 + r_1\,.$$

Portando il primo addendo al primo membro e mettendo in evidenza c scriviamo

$$c(a' - b'q_1) = r_1\,. \tag{3.21}$$

Questa uguaglianza ci dice che anche il primo resto r_1 è un multiplo di c. Per brevità, riscriviamo la (3.21) come $r_1 = cr_1'$. Sostituendo questa uguaglianza e la $b = cb'$ nella

(3.13), si ottiene

$$cb' = cr'_1 q_2 + r_2 .$$

Ancora una volta, portando il primo addendo a primo membro e mettendo in evidenza c, troviamo

$$c(b' - r'_1 q_2) = r_2 . \tag{3.22}$$

Questa uguaglianza ci dice che anche il secondo resto r_2 è un multiplo di c. Continuando in questo modo per sostituzioni successive nelle divisioni, dalla prima all'ultima, vediamo che tutti i resti sono un multiplo di c, compreso r_{n+2}. Quindi r_{n+2} è proprio il massimo comune divisore, come volevamo.

Vediamo ora una importante proprietà del massimo comune divisore, che ci sarà molto utile nel paragrafo successivo.

Proposizione 3.21. *Sia d il massimo comune divisore di due interi a e b. Allora, esistono due interi $x, y \in \mathbb{Z}$ tali che*

$$d = ax + by . \tag{3.23}$$

Dimostrazione. Denotiamo il massimo comune divisore di a e b con r_{n+2} come nella (3.16). Iniziamo quindi riscrivendo la (3.16) come

$$r_{n+2} = r_n - r_{n+1} q_{n+2}$$

e cioè come

$$r_{n+2} = x_n r_n + y_{n+1} r_{n+1} \tag{3.24}$$

per due interi x_n, y_{n+1}. Ora notiamo che tutte le divisioni precedenti possono essere scritte nella forma

$$r_i = r_{i-2} - r_{i-1} q_i$$

per $i + 3, 4, \ldots, n + 1$. Quindi sostituendo questa scrittura di r_{n+1} in 3.24 otteniamo

$$\begin{aligned} r_{n+2} &= x_n r_n + y_{n+1} \left(r_{n-1} - r_n q_{n+1} \right) \\ &= \left(x_n - y_{n+1} q_{n+1} \right) r_n + y_{n+1} r_{n-1} \end{aligned}$$

e cioè come

$$r_{n+2} = x_{n-1} r_{n-1} + y_n r_n \tag{3.25}$$

per due interi x_{n-1}, y_n. Continuando con le sostituzioni dei resti precedenti sino a r_3 otteniamo

$$r_{n+2} = x_1 r_1 + y_2 r_2 .$$

Sostituendo in quest'ultima la (3.13) nella forma $r_2 = b - r_1 q_2$ si arriva a

$$\begin{aligned} r_{n+2} &= x_1 r_1 + y_2 (b - r_1 q_2) \\ &= y_2 b + (x_1 - y_2 q_2) r_1 . \end{aligned} \tag{3.26}$$

Ora dalla (3.12) otteniamo $r_1 = a - bq_1$ che sostituita nella (3.26) rende

$$\begin{aligned} r_{n+2} &= y_2 b + (x_1 - y_2 q_2)(a - bq_1) \\ &= (x_1 - y_2 q_2) a + (y_2 + y_2 q_2 q_1 - x_1 q_1) b . \end{aligned}$$

Abbiamo quindi scritto il massimo comune divisore d di a e b come $d = ax + by$ dove $x = x_1 - y_2 q_2$ e $y = y_2 + y_2 q_2 q_1 - x_1 q_1$. □

Consideriamo il caso dell'Esempio 3.19 in cui, tramite la successione di divisioni

$$120 = 23 \cdot 5 + 5$$

$$23 = 5 \cdot 4 + 3$$

$$5 = 3 \cdot 1 + 2$$

$$3 = 2 \cdot 1 + 1$$

$$2 = 1 \cdot 2$$

abbiamo mostrato che il massimo comune divisore di 120 e 23 è $d = 1$ (l'ultimo resto non nullo delle divisioni). Allo scopo di trovare i due interi x e y tali che $120x + 23y = 1$, la cui esistenza è prevista nella Proposizione 3.21, iniziamo con l'osservare che se nella penultima divisione $3 = 2 \cdot 1 + 1$ portiamo il primo addendo al primo membro, ovvero

$$1 = 3 - 2 \cdot 1 = 1 \cdot 3 + (-1) \cdot 2 , \tag{3.27}$$

otteniamo una scrittura del massimo comune divisore come combinazione degli ultimi due resti 2 e 3 ottenuti prima del massimo comune divisore stesso. Ora, combinando la (3.27) con la divisione precedente $5 = 3 \cdot 1 + 2$ possiamo scrivere 1 come combinazione dei penultimi due resti 5 e 3. Più precisamente, portando al primo membro nella $5 = 3 \cdot 1 + 2$ otteniamo $2 = 5 - 3 \cdot 1$ che sostituito al posto di 2 nella (3.27) ci dà

$$1 = 1 \cdot 3 + (-1) \cdot (5 - 3 \cdot 1) = 3 - 5 + 3 = 3 \cdot 2 + (-1) \cdot 5. \tag{3.28}$$

Iteriamo questo procedimento: usando la divisione $23 = 5 \cdot 4 + 3$, riscritta come $3 =$

$23 - 5 \cdot 4$, e sostituendo al 3 che compare nella (3.28) si ottiene

$$\begin{aligned} 1 &= 3 \cdot 2 + (-1) \cdot 5 = (23 - 5 \cdot 4) \cdot 2 + (-1) \cdot 5 = \\ &= 23 \cdot 2 - 5 \cdot 4 \cdot 2 + (-1) \cdot 5 = 23 \cdot 2 + (-9) \cdot 5 \end{aligned} \tag{3.29}$$

cioè siamo arrivati a scrivere il massimo comune divisore come combinazione di 23 e 5. A questo punto, possiamo usare la prima divisione $120 = 23 \cdot 5 + 5$, riscritta $5 = 120 - 23 \cdot 5$, per esprimere nella combinazione (3.29) il 5 in funzione di 120 e 23. In formule, sostituendo,

$$\begin{aligned} 1 &= 23 \cdot 2 + (-9) \cdot 5 = 23 \cdot 2 + (-9) \cdot (120 - 23 \cdot 5) = \\ &= 23 \cdot 2 + (-9) \cdot 120 + (-9) \cdot (-5) \cdot 23 = \\ &= 47 \cdot 23 + (-9) \cdot 120 \end{aligned} \tag{3.30}$$

e siamo riusciti ad esprimere 1 nella forma $23x + 120y$, con $x = 47$ e $y = -9$.

Vediamo un ulteriore esempio. Sia $a = 42$ e $b = 30$, per i quali abbiamo già mostrato che il massimo comune divisore è 6 nell'Esempio 3.18 mediante la successione di divisioni

$$\begin{aligned} 42 &= 30 \cdot 1 + 12 \\ 30 &= 12 \cdot 2 + 6 \\ 12 &= 6 \cdot 2\,. \end{aligned}$$

Come sopra, usiamo la penultima divisione $30 = 12 \cdot 2 + 6$ per esprimere il massimo comune divisore $6 = 30 + (-2) \cdot 12$ come combinazione di 30 e 12. In tale uguaglianza sostituiamo $12 = 42 - 30$, ricavata dalla prima divisione, ottenendo

$$6 = 30 + (-2) \cdot (42 - 30).$$

Procedendo come sopra, otteniamo

$$6 = 30 + (-2) \cdot 42 + 2 \cdot 30 = 3 \cdot 30 + (-2) \cdot 42\,,$$

cioè siamo riusciti, come volevamo, ad esprimere 6 nella forma $30x + 42y$, con $x = 3$ e $y = -2$.

Osservazione 3.22. Se a e b sono interi uno multiplo dell'altro, diciamo a è multiplo di b, allora il procedimento descritto sopra non si applica. Ad esempio se $a = 140$ e $b = 14$, la prima divisione $140 = 14 \cdot 10 + 0$ dà già resto zero, e quindi non possiamo usare il fatto che il massimo comune divisore è l'ultimo resto non nullo delle divisioni successive. Tuttavia, in tal caso il massimo comune divisore è semplicemente uguale a

b, che in tale situazione è chiaramente divisore di a e anche di b, ed è chiaramente il massimo. Anche per ottenere la forma $d = ax + by$ non è necessario applicare nessun procedimento di divisioni, in quanto essendo $d = b$ basta scrivere $d = a \cdot 0 + b \cdot 1$. Lo stesso ragionamento si applica al caso in cui $a = 0$ o $b = 0$.

Concludiamo questo paragrafo con la seguente domanda: dato un intero positivo N, quanti sono gli interi positivi $k < N$ coprimi con N (ovvero quelli che hanno come unico divisore in comune con N l'intero 1)?
Il numero di tali interi si denota con $\phi(N)$, e definisce quindi una funzione ϕ detta **funzione di Eulero**, che è di fondamentale importanza nelle applicazioni che vedremo nell'ultima sezione di questo capitolo.
Vediamo allora come calcolare $\phi(N)$ per ogni N. L'idea è sfruttare il fatto che N, in base al teorema fondamentale dell'aritmetica, si può scrivere come $N = P_1^{\alpha_1} P_2^{\alpha_2} \cdots P_s^{\alpha_s}$ con $P_1, P_2, \ldots, P_s$ primi.
Innanzitutto, se N è un numero primo P, si ha

$$\phi(P) = P - 1\,. \tag{3.31}$$

Infatti, dal momento che P è primo, per definizione esso non ha altri divisori (positivi) oltre 1 e P stesso. Quindi, un numero k che ammette divisori diversi da 1 in comune con P è necessariamente un multiplo di P. Questo chiaramente non può succedere se $k < P$: tutti gli interi $k < P$, ovvero $k = 1, 2, \ldots, P - 1$ sono coprimi con P, da cui la (3.31).
Generalizziamo ora al caso in cui $N = P^\alpha$ è potenza di un primo:

$$\phi(P^\alpha) = P^\alpha - P^{\alpha-1}\,. \tag{3.32}$$

Infatti, $\phi(P^\alpha)$ è dato dai numeri positivi minori di P^α che sono coprimi con P^α. Possiamo contarli come tutti i numeri da 1 a P^α, che sono proprio P^α, meno quelli che hanno divisori in comune (oltre a 1) con P^α. Se dimostriamo allora che gli interi minori o uguali a P^α che hanno divisori in comune con P^α sono $P^{\alpha-1}$, avremo finito.
Dal momento che P^α è potenza di P, un intero k può avere un divisore in comune con P^α solo se contiene P tra i suoi fattori, ovvero solo se k è un multiplo di P. Tali multipli sono chiaramente

$$P,\ 2P,\ 3P, \ldots, P^{\alpha-1}P$$

(ci siamo fermati a $P^{\alpha-1}P = P^\alpha$ perché dobbiamo considerare solo interi minori o uguali a P^α). Quindi il numero di interi in questo intervallo non coprimi con P^α è $P^{\alpha-1}$, come volevamo.

Infine, per calcolare $\phi(N)$ per ogni N useremo il seguente risultato, che non dimostriamo.

Lemma 3.23. *Se N_1 e N_2 sono due interi primi tra loro, allora*

$$\phi(N_1N_2) = \phi(N_1)\phi(N_2)\,. \tag{3.33}$$

Possiamo combinare questo lemma con la formula (3.32) per calcolare la funzione di Eulero per ogni N.
Supponiamo che sia, in base al teorema fondamentale dell'aritmetica, $N = P_1^{\alpha_1}P_2^{\alpha_2}\cdots P_s^{\alpha_s}$. Chiaramente i fattori $P_1^{\alpha_1}$, $P_2^{\alpha_2}, \ldots, P_s^{\alpha_s}$ sono tutti primi tra loro in quanto potenze di primi diversi. Concludiamo che, in base al Lemma 3.23,

$$\phi(N) = \phi(P_1^{\alpha_1}P_2^{\alpha_2}\cdots P_s^{\alpha_s}) = \phi(P_1^{\alpha_1})\phi(P_2^{\alpha_2})\cdots\phi(P_s^{\alpha_s}) =$$

(per la (3.32))

$$= (P_1^{\alpha_1} - P_1^{\alpha_1-1})(P_2^{\alpha_2} - P_2^{\alpha_2-1})\cdots(P_s^{\alpha_s} - P_s^{\alpha_s-1})\,. \tag{3.34}$$

Se nelle parentesi tonde che compaiono nella (3.34) mettiamo in evidenza rispettivamente $P_1^{\alpha_1}, P_2^{\alpha_2}, \ldots, P_s^{\alpha_s}$, notiamo che

$$\phi(N) = \left[P_1^{\alpha_1}\left(1 - \frac{1}{P_1}\right)\right]\left[P_2^{\alpha_2}\left(1 - \frac{1}{P_2}\right)\right]\cdots\left[P_s^{\alpha_s}\left(1 - \frac{1}{P_s}\right)\right]\,. \tag{3.35}$$

Ovvero, permutando i fattori,

$$\phi(N) = P_1^{\alpha_1}P_2^{\alpha_2}\cdots P_s^{\alpha_s}\left(1 - \frac{1}{P_1}\right)\left(1 - \frac{1}{P_2}\right)\cdots\left(1 - \frac{1}{P_s}\right) \tag{3.36}$$

che, ricordando che $P_1^{\alpha_1}P_2^{\alpha_2}\cdots P_s^{\alpha_s}$ è proprio N, significa

$$\phi(N) = N\left(1 - \frac{1}{P_1}\right)\left(1 - \frac{1}{P_2}\right)\cdots\left(1 - \frac{1}{P_s}\right)\,. \tag{3.37}$$

Questa è la formula che si trova solitamente nei libri per la funzione di Eulero ϕ.
Ad esempio, se $N = 100$, abbiamo $N = 5^2 \cdot 2^2$, quindi

$$\phi(100) = 100\left(1 - \frac{1}{5}\right)\left(1 - \frac{1}{2}\right) = 100 \cdot \frac{4}{5} \cdot \frac{1}{2} = 40.$$

Osserviamo che il calcolo di ϕ mediante la formula (3.37) richiede di conoscere la decomposizione di N in fattori primi, che non è sempre facile da trovare.

Osservazione 3.24. Lo studio dei numeri primi, e in particolare della loro distribuzione nell'insieme dei numeri interi, costituisce un capitolo fondamentale nella matematica. Intanto, sappiamo che i numeri primi sono infiniti. La dimostrazione di questo fatto è dovuta ad Euclide ed è apprezzata ancora oggi per la sua semplicità ed eleganza. Se per assurdo i primi fossero in numero finito, diciamo $p_1, p_2, \ldots, p_k$, il numero

$N = p_1 p_2 \cdots p_k + 1$ sarebbe un numero non divisibile per nessun primo. Infatti il resto della divisione di N per uno qualunque dei p_i sarebbe sempre 1, ma questo per il teorema fondamentale dell'aritmetica non è possibile a meno che non sia primo con $p_1, p_2, \ldots, p_k$. Quindi deve esistere un altro primo che divide N (possibilmente anche N stesso), contraddicendo l'ipotesi che i primi fossero solo $p_1, p_2, \ldots, p_k$.
Come è distribuito tale insieme infinito all'interno degli interi? Quanto è raro trovare primi quando si considerano numeri sempre più grandi?
Da una parte, si congettura che esistano infinite coppie di numeri primi a distanza 2 l'uno dall'altro (ad esempio 3 e 5, oppure 17 e 19), detti *numeri primi gemelli*. Dall'altra esistono sequenze lunghe quanto vogliamo di numeri che non contengono nessun primo. Dato n, la sequenza

$$n! + 2, n! + 3, \ldots, n! + n$$

è una sequenza di $n - 1$ numeri consecutivi (quindi possiamo renderla lunga a piacere scegliendo n) in cui nessuno è primo. Infatti, $n! + 2$ è divisibile per 2 in quanto sia $n!$ che 2 lo sono (ricordiamo che $n!$ è il prodotto di tutti i numeri naturali da 1 a n), $n! + 3$ è divisibile per 3 in quanto sia $n!$ che 3 lo sono, e così via.
Un importante teorema afferma che se $\pi(n)$ indica il numero dei primi minori di n, allora la quantità $\frac{\pi(n)}{n}$ (che descrive la percentuale di primi minori di n rispetto al totale dei numeri da 1 a n) si avvicina sempre di più a $\frac{1}{\log n}$ (ovvero il rapporto tra $\frac{\pi(n)}{n}$ e $\frac{1}{\log n}$ tende a 1 al crescere di n).

3.3 Congruenze e aritmetica modulare

Basandoci sui risultati esposti nella prima parte di questo capitolo, introdurremo e studieremo ora un'importante relazione di equivalenza sull'insieme dei numeri interi, la cosiddetta *congruenza modulo n*.

Definizione 3.25. Sia n un intero positivo fissato. Dati $a, b \in \mathbb{Z}$, diremo che a è **congruo a** b **modulo** n se a e b divisi per n danno lo stesso resto. Equivalentemente, a è congruo a b modulo n se la differenza $a - b$ è un multiplo di n.

Che le due formulazioni della definizione appena data siano equivalenti si vede facilmente. Infatti, da una parte, se a e b divisi per n danno lo stesso resto, ovvero $a = nq_1 + r$ e $b = nq_2 + r$, allora

$$a - b = (nq_1 + r) - (nq_2 + r) = nq_1 + r - nq_2 - r = nq_1 - nq_2 = n(q_1 - q_2),$$

ovvero la differenza $a - b$ è un multiplo di n.
Viceversa, supponiamo che $a - b = kn$ per qualche k, ovvero $a = b + kn$. Se il resto della divisione di b per n è r, ovvero $b = nq + r$, allora sostituendo in $a = b + kn$ si

trova

$$a = b + kn = nq + r + kn = n(q + k) + r$$

da cui si vede che anche a, divisa per n, ha resto r (con quoziente $q + k$).

Per indicare che a è congruo a b modulo n, scriveremo

$$a \equiv b \pmod{n}$$

o anche

$$a \equiv_n b\,.$$

Mostriamo ora che la congruenza è una relazione di equivalenza, ovvero che gode delle proprietà riflessiva, simmetrica e transitiva (Definizione 2.6). Innanzitutto, si ha

$$a - a = 0 = 0 \cdot n$$

ovvero la differenza tra a e se stesso è un multiplo di n. Per la definizione data, questo significa che $a \equiv_n a$, cioè la congruenza è riflessiva.
Inoltre, se $a \equiv_n b$, ovvero $a - b = kn$ per qualche $k \in \mathbb{Z}$, si ha, semplicemente cambiando di segno,

$$b - a = -kn = (-k)n\,.$$

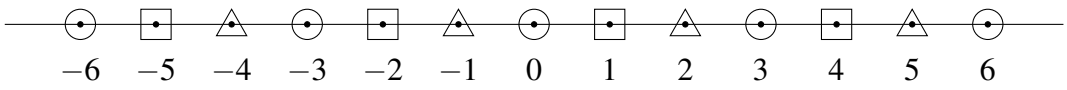
Perciò anche $b - a$ è un multiplo di n, cioè $b \equiv_n a$: la relazione di congruenza è simmetrica.
Infine, se $a \equiv_n b$ (ovvero $a - b = kn$ per qualche $k \in \mathbb{Z}$) e $b \equiv_n c$ (ovvero $b - c = ln$ per qualche $l \in \mathbb{Z}$) allora si ha

$$a - c = a - b + b - c = kn + ln = (k + l)n\,.$$

Quindi anche $a - c$ è un multiplo di n, cioè $a \equiv_n c$: la relazione di congruenza è transitiva.

In base alla teoria vista nel capitolo precedente, tale relazione ripartisce allora $\mathbb{Z}$ in classi di equivalenza non vuote e disgiunte. Una classe contiene tutti i numeri interi che hanno lo stesso resto quando li si divide per n.
Dal momento che il resto r della divisione per n deve soddisfare $0 \leqslant r < n$, i possibili resti sono $0, 1, 2, \ldots, n - 1$, quindi avremo esattamente n classi di equivalenza: $[0], [1], \ldots, [n - 1]$. L'insieme $\{[0], [1], \ldots, [n - 1]\}$ delle classi di equivalenza su $\mathbb{Z}$ della congruenza modulo n si denota con $\mathbb{Z}_n$.
Ad esempio, per $n = 3$ i possibili resti della divisione sono $0, 1, 2$. La classe $[0]$ di equivalenza di 0 contiene tutti i numeri a tali che $a = a - 0 = 3k$, ovvero i multipli di 3. La classe $[1]$ di equivalenza di 1 contiene tutti gli interi a tali che $a - 1 = 3k$, ovvero

$a = 3k+1$ (ad esempio 4, 7, 10, -2, -5 etc.). La classe $[2]$ di equivalenza di 2 contiene tutti gli interi a tali che $a - 2 = 3k$, ovvero $a = 3k + 2$ (ad esempio 5, 8, 11, -1, -4 etc.).



Nel disegno stiamo etichettando con lo stesso simbolo interi che appartengono alla stessa classe.

Se $n = 2$, abbiamo solo due resti possibili, 0 e 1 e quindi due classi, $[0]$ e $[1]$. La prima contiene tutti gli interi a del tipo $2k$ e la seconda tutti gli interi del tipo $2k + 1$. In altre parole, la partizione indotta su $\mathbb{Z}$ è quella in numeri pari e numeri dispari.

Come tutte le relazioni di equivalenza, la relazione di congruenza è una sorta di "uguaglianza in senso largo": consideriamo due numeri "uguali" se lo sono "a meno di mutipli di n".
La relazione di equivalenza per eccellenza, l'uguaglianza, verifica l'ovvia proprietà che se $a = b$ e $a' = b'$ allora $a + a' = b + b'$ e $aa' = bb'$. Lo stesso vale per la congruenza modulo n.

Lemma 3.26. *La congruenza modulo n verifica le due seguenti proprietà.*

(1) Se $a \equiv_n a'$ e $b \equiv_n b'$, allora $a + b \equiv_n a' + b'$.

(2) Se $a \equiv_n a'$ e $b \equiv_n b'$, allora $ab \equiv_n a'b'$.

Dimostrazione. (1): se $a \equiv_n a'$ e $b \equiv_n b'$ allora $a - a' = kn$ e $b - b' = ln$ per qualche $k, l \in \mathbb{Z}$. Dobbiamo dimostrare che anche la differenza tra $a + b$ e $a' + b'$ è un multiplo di n. Infatti

$$a + b - (a' + b') = a - a' + b - b' = kn + ln = (k + l)n$$

come volevamo.
(2): sotto la stessa ipotesi dobbiamo dimostrare che la differenza tra ab e $a'b'$ è un multiplo di n o, equivalentemente, che $ab = a'b' + qn$ per qualche $q \in \mathbb{Z}$. Usando $a - a' = kn$ e $b - b' = ln$ (riscrivibili come $a = a' + kn$ e $b = b' + ln$) abbiamo

$$ab = (a' + kn)(b' + ln) = a'b' + a'ln + knb' + kln^2 = a'b' + n(a'l + kb' + kln)\,,$$

il che dimostra l'asserto (con $q = a'l + kb' + kln$). □

Le due proprietà appena dimostrate, come vedremo, hanno molte applicazioni.
La prima è la data da alcuni criteri di divisibilità. È noto che un numero n è divisibile per 3 se e solo se la somma delle sue cifre è divisibile per 3 (lo stesso enunciato vale per 9).
Ad esempio, il numero $n = 1917$ è divisibile per 3 in quanto lo è la somma $1+9+1+7 = 18$ delle sue cifre; il numero 313 non lo è perché $3 + 1 + 3 = 10$ non lo è.
La spiegazione di tale criterio si basa proprio sul Lemma 3.26. Infatti, dire che le cifre di n (nella notazione decimale) sono $a_k, a_{k-1}, \ldots, a_1, a_0$ significa dire che

$$n = a_k 10^k + a_{k-1} 10^{k-1} + \cdots + a_1 10 + a_0 \,. \tag{3.38}$$

Ora, poiché $10 = 3 \cdot 3 + 1$, si ha $10 \equiv_3 1$; applicando la (2) del Lemma 3.26 al particolare caso $a = a' = 10$, $b = b' = 1$, si ha che $10 \cdot 10 \equiv_3 1 \cdot 1$, ovvero $10^2 \equiv_3 1$; sempre applicando la stessa proprietà, si ha $10^2 \cdot 10 \equiv_3 1 \cdot 1$, ovvero $10^3 \equiv_3 1$ e così via. Si vede quindi che tutte le potenze positive 10^i di 10 sono congrue a 1 modulo 3.
Ora, questo implica che per ogni addendo della rappresentazione (3.38), che è del tipo $a_i 10^i$, valga $a_i 10^i \equiv_3 a_i$. Infatti, applicando di nuovo la (2) del Lemma 3.26, si ha che $a_i \equiv_3 a_i$ (la congruenza gode della proprietà riflessiva) e $10^i \equiv_3 1$ implicano proprio $a_i \cdot 10^i \equiv_3 a_i \cdot 1$.
A questo punto, poichè $a_k 10^k \equiv_3 a_k$, $a_{k-1} 10^{k-1} \equiv_3 a_{k-1}$ e così via fino a $a_1 10 \equiv_3 a_1$ (l'ultima, $a_0 \equiv_3 a_0$ è chiara per la riflessività della relazione di congruenza), possiamo sommare membro a membro usando la (1) del Lemma 3.26 e ottenere

$$a_k 10^k + a_{k-1} 10^{k-1} + \cdots + a_1 10 + a_0 \equiv_3 a_k + a_{k-1} + \cdots + a_1 + a_0 \,. \tag{3.39}$$

In altre parole, un qualunque n è congruente modulo 3 alla somma delle sue cifre.
Ma allora, poiché n divisibile per 3 significa che n è congruente a 0 modulo 3, si ha che n è divisibile per 3 se e solamente se la somma delle sue cifre è congruente a zero modulo 3, ovvero anche lei divisibile per 3. Il criterio è dimostrato.

Lo stesso criterio vale per la divisibilità per 9. Il motivo è che si ha anche $10 \equiv_9 1$, quindi si la stessa dimostrazione che abbiamo usato per 3 è valida in questo caso.

Il Lemma 3.26 ci consente quindi di definire una somma e una moltiplicazione sull'insieme $\mathbb{Z}_n$ delle classi di congruenza modulo n. Questo ci permetterà di fare calcoli su questo insieme, e sviluppare quella che si chiama **aritmetica modulare**.
Piú precisamente, date due classi $[i]$ e $[j]$ della congruenza modulo n, definiamo la loro somma come

$$[i] + [j] := [i + j] \tag{3.40}$$

e il loro prodotto come

$$[i] \cdot [j] := [ij]. \tag{3.41}$$

Il problema di tali definizioni è che definire la somma o il prodotto tra due classi mediante la somma o il prodotto dei rappresentanti i e j potrebbe essere ambiguo. Infatti, una stessa classe può essere rappresentata da interi diversi (ad esempio, modulo 3 si ha $[1] = [4]$). Quindi dobbiamo assicurarci che anche scegliendo rappresentanti diversi i' e j' per le classi $[i]$ e $[j]$ rispettivamente, la somma e il prodotto definiti da (3.40) e (3.41) diano sempre lo stesso risultato (ovvero la stessa classe).
Ma questo è garantito proprio dal Lemma 3.26. Infatti, supponiamo che sia $[i'] = [i]$ e $[j'] = [j]$: questo significa che $i \equiv_n i'$ e $j \equiv_n j'$. Ma allora, applicando la (1) del Lemma 3.26 si ha $i + j \equiv_n i' + j'$, ovvero $[i + j] = [i' + j']$. Quindi la somma che abbiamo definito non dipende dai rappresentanti scelti per le classi e non c'è ambiguità. Analogamente, sempre supponendo che sia $[i'] = [i]$ e $[j'] = [j]$, ovvero $i \equiv_n i'$ e $j \equiv_n j'$, applicando la (2) del Lemma 3.26 si ha $ij \equiv_n i'j'$, ovvero $[ij] = [i'j']$. Anche il prodotto per come lo abbiamo definito non dipende dai rappresentanti scelti per le classi e non c'è ambiguità.

A questo punto, come abbiamo fatto ogni volta che ci siamo trovati in presenza di un insieme con una o più operazioni, verifichiamo quali proprietà sono soddisfatte da queste operazioni.

(1) La somma è associativa:

$$([i] + [j]) + [k] = [i + j] + [k] = [(i + j) + k] =$$

(sfruttando l'associatività della somma usuale tra gli interi i, j, k)

$$[i + (j + k)] = [i] + [j + k] = [i] + ([j] + [k]).$$

(2) La somma è commutativa:

$$[i] + [j] = [i + j] = [j + i] = [j] + [i]$$

dove nell'uguaglianza centrale abbiamo sfruttato il fatto che la somma usuale tra interi è commutativa.

(3) Esiste un elemento neutro per la somma: la classe $[0]$ di 0. Infatti si ha

$$[i] + [0] = [i + 0] = [i] = [0 + i] = [0] + [i].$$

Si noti che, dal momento che $[0] = [kn]$ per ogni multiplo kn di n, possiamo anche scrivere $[i] + [kn] = [i]$ o, equivalentemente, $i + kn \equiv_n i$.

(4) Per ogni classe $[i]$ esiste un elemento inverso per la somma (ovvero un opposto), dato dalla classe $[-i]$:

$$[i] + [-i] = [i + (-i)] = [0]\,.$$

A volte si preferisce vedere ogni classe rappresentata da un numero compreso tra 0 e $n-1$. In questo caso si può anche rappresentare l'opposto di $[i]$ come $[n-i]$ in quanto chiaramente $-i \equiv_n n - i$. Ad esempio, in $\mathbb{Z}_5$ l'opposta della classe $[2]$ è certamente $[-2]$, che può essere anche espressa come $[5-2] = [3]$ (e infatti $[2] + [3] = [5] = [0]$).

(5) Il prodotto è associativo:

$$([i][j])[k] = [ij][k] = [(ij)k] =$$

(sfruttando l'associatività del prodotto usuale tra gli interi i, j, k)

$$[i(jk)] = [i][jk] = [i]([j][k])\,.$$

(6) Il prodotto è commutativo:

$$[i][j] = [ij] = [ji] = [j][i]$$

dove nell'uguaglianza centrale abbiamo sfruttato il fatto che il prodotto usuale tra interi è commutativo.

(7) Esiste un elemento neutro per il prodotto, che è la classe $[1]$ di 1. Infatti si ha

$$[i][1] = [i \cdot 1] = [i] = [1 \cdot i] = [1][i]\,.$$

(8) Il prodotto è distributivo rispetto alla somma:

$$[i]([j] + [k]) = [i][j + k] = [i(j + k)] = [ij + ik] = [ij] + [ik] = [i][j] + [i][k]\,.$$

Analogamente si mostra che $([i] + [j])[k] = [i][j] + [i][k]$.

Le proprietà appena elencate di dicono che $\mathbb{Z}_n$, dotato delle operazioni di somma e prodotto definite dalle (3.40) e (3.41), è un anello commutativo (Definizione 3.1).
Tuttavia, rispetto all'anello $\mathbb{Z}$ degli interi, $\mathbb{Z}_n$ presenta alcune importanti differenze. La prima, più evidente, è che si tratta di un anello finito, diversamente da $\mathbb{Z}$ che è infinito. La seconda importante differenza tra $\mathbb{Z}$ e $\mathbb{Z}_n$ è che in quest'ultimo non vale la legge di annullamento del prodotto (che afferma che un prodotto è zero solo se almeno uno dei due fattori è nullo). Ad esempio, in $\mathbb{Z}_6$ consideriamo le classi $[2]$ e $[3]$: nessuna di

queste due è la classe nulla (ovvero $[2] \neq [0]$ e $[3] \neq [0]$) in quanto né 2 né 3 sono congruenti a zero modulo 6. Tuttavia il loro prodotto è dato da

$$[2][3] = [2 \cdot 3] = [6] = [0]\,.$$

Un'altra domanda fondamentale nell'aritmetica di $\mathbb{Z}_n$ riguarda l'esistenza dell'inverso moltiplicativo di una classe $[a]$ data. Nell'anello degli interi $\mathbb{Z}$ sappiamo che gli unici interi x ad avere un inverso moltiplicativo (ovvero un elemento y per cui $xy = yx = 1$) sono $+1$ e -1. Vediamo ora che in $\mathbb{Z}_n$ la questione è più complessa.
Sia $[a]$ una classe in $\mathbb{Z}_n$. l'esistenza di un inverso moltiplicativo di $[a]$ equivale all'esistenza di una classe $[x]$ tale che $[a][x] = [1]$. Ma quest'ultima uguaglianza , per definizione di prodotto tra classi, significa $[ax] = [1]$ ovvero che ax e 1 sono congrui modulo n. In altre parole la loro differenza è un multiplo di n:

$$ax - 1 = yn$$

per un certo intero $y \in \mathbb{Z}$.
Portando yn a primo membro e -1 a secondo, concludiamo che la classe di a ha inverso modulo n se e solo esistono due interi x, y tali che

$$ax - yn = 1 \tag{3.42}$$

e in tal caso l'inverso di $[a]$ è dato proprio dalla classe $[x]$.
Ora, la relazione (3.42), come sappiamo dalla Proposizione 3.21, è verificata sicuramente se il massimo comune divisore di a e n è 1. Viceversa, se vale tale relazione allora necessariamente il massimo comune divisore di a e n è 1. Questo perché se d è un divisore comune di a e n (ovvero $a = da'$ e $n = dn'$) allora sostituendo nella (3.42) otteniamo

$$ax - yn = da'x - ydn' = d(a'x - yn') = 1\,.$$

Quindi necessariamente $d = \pm 1$ (perché non esiste nessun intero che moltiplicato per d dà 1).

In altre parole, abbiamo appena dimostrato che *una classe $[a]$ in $\mathbb{Z}_n$ è invertibile se e solo se il massimo comune divisore di a e n è 1, ovvero se e solo se a e n sono primi tra loro.*

Quanto detto sopra per ottenere questo risultato ci fornisce anche un modo pratico per il calcolo dell'inverso, quando questo esiste. Ad esempio, determiniamo l'inverso di $[5]$ in $\mathbb{Z}_{14}$ (o, detto in altre parole, l'inverso di 5 modulo 14).
Innanzitutto, tale inverso esiste perché 5 e 14 sono primi tra loro. Per trovare tale inverso, ovvero l'intero x per cui è soddisfatta la (3.42), eseguiamo il procedimento di divisioni successive visto nella sezione precedente:

$$14 = 5 \cdot 2 + 4$$
$$5 = 4 \cdot 1 + 1$$
$$4 = 1 \cdot 4\,.$$

Dalla seconda divisione troviamo $1 = 5 - 4$, e sostituendo 4 tramite $4 = 14 - 5 \cdot 2$ (ricavata dalla prima divisione) otteniamo

$$1 = 5 - 4 = 5 - (14 - 5 \cdot 2) = 5 - 14 + 5 \cdot 2 = 5 \cdot 3 + (-1) \cdot 14\,.$$

Quindi, in base a quanto visto sopra, 3 è l'inverso di 5 modulo 14, ovvero $[3]$ è l'inverso di $[5]$ in $\mathbb{Z}_{14}$ (infatti, $[5][3] = [15] = [1]$ in $\mathbb{Z}_{14}$).

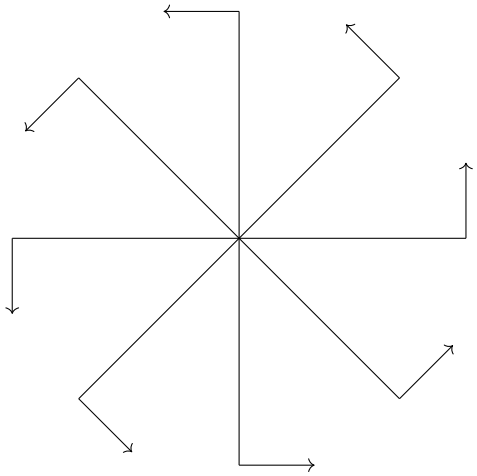
Come corollario di quanto appena detto, notiamo che se n è un numero primo, allora per ogni $i = 1, 2, \ldots, n-1$ si ha che i è primo con n. Quindi in $\mathbb{Z}_n$ le classi $[1], [2], \ldots, [n-1]$ (cioè tutte le classi tranne la classe nulla $[0]$) ammettono un inverso. Questo, in base alla Definizione 3.2, significa che *se n è primo allora $\mathbb{Z}_n$ è un campo* (esattamente come $\mathbb{Q}$ e $\mathbb{R}$).

In particolare, in base all'Osservazione 3.3, la legge dell'annullamento del prodotto è soddisfatta in $\mathbb{Z}_n$ con n primo, come in tutti i campi.

Osservazione 3.27. Le proprietà (1), (2), (3) e (4) dell'operazione di somma tra classi che abbiamo visto nella Definizione 3.1 ci dicono che $\mathbb{Z}_n$, rispetto a tale operazione, è un gruppo commutativo (Definizione 2.56).

Una interessante "realizzazione geometrica" di questo gruppo può essere data come segue. Consideriamo il gruppo delle simmetrie della seguente figura del piano[5], una sorta di girandola con tutti i bracci della stessa lunghezza, ottenuta dividendo l'angolo giro di $360^\circ = 2\pi$ in 8 parti uguali.

[5]Ricordiamo che le simmetrie di una figura sono i movimenti rigidi (traslazioni, rotazioni, riflessioni) che la lasciano invariata, rispetto all'operazione di composizione.



Sicuramente la rotazione di angolo $\frac{2\pi}{8} = 45^\circ$ in senso antiorario lascia invariata la figura (comprese le freccette che spuntano da ogni braccio), così come la rotazione di $\frac{2\pi}{8} \cdot 2 = 90^\circ$ (ottenuta componendo la rotazione di 45° con se stessa), la rotazione di $\frac{2\pi}{8} \cdot 3 = 135^\circ$ (ottenuta componendo la rotazione di 45° con se stessa 3 volte) e così via tutte le rotazioni di angolo $\frac{2\pi}{8} \cdot k$, ottenute componendo k volte la rotazione di 45°.
Tuttavia, è chiaro che se considero la rotazione di angolo $\frac{2\pi}{8} \cdot 8$, ottengo esattamente la funzione identica che manda ogni punto in se stesso (8 rotazioni di 45° una dietro l'altra danno un giro completo). La stessa cosa sarà vera se ruoto di $\frac{2\pi}{8} \cdot 16$, $\frac{2\pi}{8} \cdot 24$ e in generale per qualunque multiplo di 8. Analogamente, una rotazione di $\frac{2\pi}{8} = 45^\circ$ muove i punti nello stesso modo di una rotazione di $\frac{2\pi}{8} \cdot 9$ (un giro completo più altri 45°), o di $\frac{2\pi}{8} \cdot 17$ (due giri completi più altri 45°) e così via per tutte le rotazioni di $\frac{2\pi}{8} \cdot k$ con k che differisce da 1 per un multiplo di 8.
In generale, avremo che un angolo di $\frac{2\pi}{8} \cdot k$ e un angolo di $\frac{2\pi}{8} \cdot l$ rappresentano la stessa rotazione se e solo se k e l differiscono per un multiplo di 8. Questa è esattamente la relazione di congruenza modulo 8 che definisce le classi di $\mathbb{Z}_8$.
Quindi c'è una corrispondenza biunivoca tra le classi $\{[0],[1],[2],[3],[4],[5],[6],[7]\}$ e le rotazioni che lasciano invariata la figura: alla classe $[k]$ corrisponde la rotazione di angolo $\frac{2\pi}{8} \cdot k$. Chiaramente $[k] = [k']$ se e solo se la rotazione di $\frac{2\pi}{8} \cdot k$ coincide con la rotazione di $\frac{2\pi}{8} \cdot k'$ (visto che k e k' differiscono per un multiplo di 8).
Questa corrispondenza identifica il gruppo $\mathbb{Z}_n$ con il gruppo delle simmetrie della figura non solo come insiemi, ma come gruppi, cioè anche rispetto alle operazioni in ciascuno dei due. Infatti se compongo due rotazioni di angoli $\frac{2\pi}{8} \cdot k$ e $\frac{2\pi}{8} \cdot l$ (che corrispondono alle classi $[k]$ e $[l]$ in $\mathbb{Z}_n$), il risultato sarà la rotazione di $\frac{2\pi}{8} \cdot k + \frac{2\pi}{8} \cdot l = \frac{2\pi}{8} \cdot (k+l)$, che corrisponde alla classe $[k+l]$. Quindi la composizione in un gruppo corrisponde esattamente alla somma nell'altro, e così vengono preservate anche tutte le altre relazioni. Ad esempio, il fatto che in $\mathbb{Z}_8$ si abbia $[5] + [3] = [0]$ corrisponde nel gruppo delle simmetrie della figura al fatto che se compongo una rotazione di $\frac{2\pi}{8} \cdot 5$ con una di $\frac{2\pi}{8} \cdot 3$

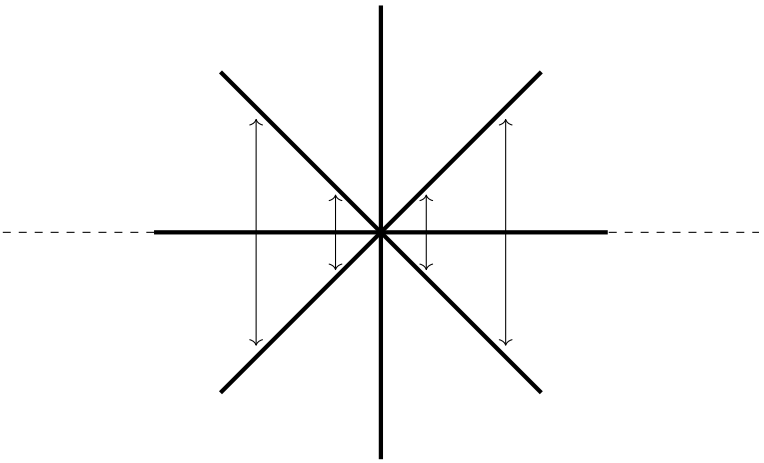
ottengo l'identità (l'elemento neutro per la composizione), in quanto avrò fatto un giro completo.

Definizione 3.28. Un **isomorfismo** f tra due gruppi $(G,\cdot)$ e $(H,\cdot')$ è una funzione biiettiva $f\colon G \longrightarrow H$ tale che

$$f(g_1 \cdot g_2) = f(g_1) \cdot' f(g_2)$$

per qualsiasi coppia di elementi $g_1, g_2 \in G$. Se esiste un isomorfismo tra due gruppi, questi si dicono **isomorfi**.

In pratica, se due gruppi sono isomorfi possiamo considerarli come se fossero lo stesso gruppo, anche se sono di natura molto diversa tra loro come nel nostro esempio (il primo è costituito di classi di numeri interi e l'operazione è una somma, il secondo da trasformazioni geometriche e l'operazione è la composizione).
Per concludere, si noti che le freccette nella figura servono a far sì che le rotazioni siano le sue *uniche* simmetrie, così da poterle identificare con gli elementi di $\mathbb{Z}_8$. Se non avessimo messo le freccette, ci sarebbero state altre simmetrie (ad esempio le riflessioni rispetto alle rette che contengono i bracci) e tale identificazione non sarebbe stata possibile:



Abbiamo visto che trovare l'inverso di un intero a modulo n (cioè un x tale che $[a][x] = [1]$, ovvero $[ax] = [1]$) equivale a risolvere l'equazione

$$ax \equiv_n 1.$$

Affrontiamo ora il problema più generale di determinare le soluzioni dell'equazione

$$ax \equiv_n b \tag{3.43}$$

dove b è un qualunque numero intero. Le domande a cui vogliamo rispondere sono: questa equazione è risolubile? se sì, quali e quante sono le sue soluzioni?
La risposta alla prima domanda ci è fornita dal seguente

Teorema 3.29. *L'equazione $ax \equiv_n b$ ammette soluzioni se e solo se il massimo comune divisore d di a e n divide b.*

Dimostrazione. Mostriamo prima che se il massimo comune divisore $d = (a,n)$ divide b, allora l'equazione ha soluzioni. Come sappiamo dalla Proposizione 3.21, se d è il massimo comune divisore di a e n, esistono due interi t e s tali che

$$d = at + ns\,. \tag{3.44}$$

Ora, l'ipotesi che d divida b significa che $b = db'$ per qualche intero b'. Moltiplicando allora la (3.44) per b' si ottiene

$$b = db' = atb' + nsb'\,. \tag{3.45}$$

Ma quest'ultima uguaglianza ci dice che atb' è congruo a b modulo n (perché differisce da b per nsb', che è un multiplo di n), ovvero $atb' \equiv_n b$: allora $x = tb'$ è una soluzione di $ax \equiv_n b$. Quindi, come volevamo, abbiamo mostrato che l'equazione ammette soluzioni sotto l'ipotesi che il massimo comune divisore di a e n divida b.
Viceversa, supponiamo che l'equazione $ax \equiv_n b$ ammetta soluzioni. Questo significa che esiste un intero x_0 tale che b e ax_0 differiscono solo per un multiplo di n, ovvero

$$b = ax_0 + ny \tag{3.46}$$

per qualche $y \in \mathbb{Z}$. Ma da questa uguaglianza è facile mostrare che il massimo comune divisore d di a e n divide b. Infatti, essendo d un divisore comune di a e n si ha che $a = da'$ e $n = dn'$ per certi interi a', n'. Ora sostituendo queste due uguaglianze nella (3.46) si trova

$$b = da'x_0 + dn'y = d(a'x_0 + n + y)$$

che ci dice proprio che d divide b. □

La dimostrazione precedente, oltre a darci il criterio necessario e sufficiente perché l'equazione $ax \equiv_n b$ sia risolubile, ci suggerisce anche come trovarne almeno una soluzione.
Esempio 3.30. Supponiamo di voler risolvere l'equazione

$$12x \equiv_{39} 15.$$

Poiché il massimo comune divisore tra $a = 12$ e $n = 39$ è $d = 3$, e 3 divide $b = 15$, l'esistenza di una soluzione è garantita dal Teorema 3.29. Per trovarla, seguiamo passo passo la dimostrazione. Esprimiamo prima $d = 3$ come combinazione di $n = 39$ e $a = 12$ come fatto nella (3.44), tramite il metodo delle divisioni successive.

$$39 = 12 \cdot 3 + 3$$

$$12 = 3 \cdot 4$$

Le divisioni confermano che il massimo comune divisore di 12 e 39 è 3, in quanto è l'ultimo resto non nullo. Dalla prima divisione otteniamo $3 = 39 - 12 \cdot 3$, ovvero

$$3 = 12 \cdot (-3) + 39 \cdot 1 \tag{3.47}$$

che sarebbe la (3.44) con $t = -3$ e $s = 1$. Ora, esattamente come fatto nella dimostrazione, osserviamo che $b = 15 = 3 \cdot 5$ (cioè, sempre in riferimento alla dimostrazione, sarebbe $b' = 5$) e moltiplichiamo la (3.47) proprio per 5, ottenendo

$$15 = 3 \cdot 5 = 12 \cdot (-3) \cdot 5 + 39 \cdot 5 = 12 \cdot (-15) + 39 \cdot 5\,. \tag{3.48}$$

Ma allora $x_0 = -15$ è una soluzione dell'equazione. Infatti, la (3.48) ci dice proprio che $12 \cdot (-15)$ è uguale a 15 modulo 39 (essi differiscono per $39 \cdot 5$).

Quella che abbiamo trovato è solo una soluzione dell'equazione: ce ne sono altre? La risposta è data dal seguente risultato, di cui omettiamo la dimostrazione.

Teorema 3.31. *Se l'equazione $ax \equiv_n b$ ha soluzioni, esse si ottengono tutte sommando a una qualunque di esse x_0 i multipli di $\frac{n}{d}$ (ovvero $x = x_0 + \frac{n}{d}k$, per ogni $k \in \mathbb{Z}$), dove d è il massimo comune divisore di a e n.*

Per illustrare il teorema, applichiamolo all'equazione $12x \equiv_{39} 15$, di cui sopra abbiamo trovato una soluzione $x_0 = -15$.
Poiché in quel caso $n = 39$ e $d = 3$, abbiamo che le soluzioni sono tutti e soli gli interi del tipo

$$x = x_0 + k\frac{n}{d} = -15 + \frac{39}{3}k = -15 + 13k$$

Queste, al variare di $k \in \mathbb{Z}$, sono tutte le soluzioni dell'equazione $12x \equiv_{39} 15$, che quindi risulta avere infinite soluzioni in $\mathbb{Z}$.
Per comprendere meglio la struttura dell'insieme delle soluzioni, distinguiamo però nell'espressione generale della soluzione $x = -15 + 13k$ tre casi:

(1) k è un multiplo di 3, ovvero $k = 3k'$ per qualche $k' \in \mathbb{Z}$ (ovvero $[k] = [0] \in \mathbb{Z}_3$).

In tal caso le soluzioni sono date da

$$x = -15 + 13k = -15 + 13 \cdot 3k' = -15 + 39k'$$

Questo ci dice che le soluzioni ottenute quando k è multiplo di 3, pur essendo interi diversi, sono tutte equivalenti modulo 39, ovvero appartengono tutte alla classe $[-15]$ di -15.

(2) k è un intero del tipo $k = 3k' + 1$ per qualche $k' \in \mathbb{Z}$ (ovvero $[k] = [1] \in \mathbb{Z}_3$).

In tal caso le soluzioni sono date da

$$x = -15 + 13k = -15 + 13 \cdot (3k' + 1) = -15 + 39k' + 13 = -2 + 39k'$$

Di nuovo otteniamo una famiglia di soluzioni tutte equivalenti modulo 39, e più precisamente tutte congruenti a -2. Esse formano la classe $[-2]$ di -2 modulo 39. Si noti che questa classe è diversa in $\mathbb{Z}_{39}$ dalla classe ottenuta nel primo caso (infatti, $[-15] \neq [-2]$ in quanto la differenza $-15 - (-2) = -13$ non è un multiplo di 39).

(3) k è un intero del tipo $k = 3k' + 2$ per qualche $k' \in \mathbb{Z}$ (ovvero $[k] = [2] \in \mathbb{Z}_3$).

In tal caso le soluzioni sono date da

$$x = -15 + 13k = -15 + 13 \cdot (3k' + 2) = -15 + 39k' + 26 = 11 + 39k' .$$

Ancora una volta, otteniamo una famiglia di soluzioni tutte equivalenti modulo 39, e più precisamente tutte congruenti a 11. Esse formano la classe $[11]$ di 11 modulo 39. Si noti che questa classe è diversa in $\mathbb{Z}_{39}$ da entrambe le classi ottenute nei casi precedenti. Infatti, $[11] \neq [-2]$ e $[11] \neq [-15]$ in quanto nessuna delle due differenze $11 - (-2) = 13$ e $11 - (-15) = 26$ è un multiplo di 39.

Si noti che non abbiamo altri casi: le uniche possibilità quando dividiamo k per 3 è che il resto sia 0, 1 o 2, che corrispondono rispettivamente ai tre casi considerati $k = 3k'$, $k = 3k' + 1$ o $k = 3k' + 2$.

Possiamo riassumere quanto appena visto come segue. L'equazione $12x \equiv_{39} 15$ ha infinite soluzioni se la consideriamo come un'equazione da risolvere in $\mathbb{Z}$. Ma possiamo pensarla come l'equazione $[12][x] = [15]$ in $\mathbb{Z}_{39}$. Infatti, l'equazione $12x \equiv_{39} 15$ significa $[12x] = [15]$, ovvero, per definizione di prodotto tra classi, proprio $[12][x] = [15]$. In questo caso essa ha esattamente tre soluzioni distinte: $[-15], [-2], [11] \in \mathbb{Z}_{39}$.

Aiutati da questo esempio, possiamo allora dire che, in generale, *se l'equazione* $ax \equiv_n b$ *ha soluzioni se e solo se il massimo comune divisore d di a e n divide b, e in tal caso l'equazione ha esattamente d soluzioni modulo n*. In particolare, *se a e n sono primi tra loro, ovvero il loro massimo comune divisore d è* 1*, allora la soluzione esiste per ogni b ed è unica (modulo n)*. Questo è in accordo con il fatto che se a e n sono primi tra loro, allora a ammette un inverso moltiplicativo modulo n, che possiamo denotare a^{-1}. Quindi l'equazione $ax \equiv_n b$ può essere risolta (cioè, x può essere determinato in modo unico modulo n) semplicemente moltiplicando entrambi i membri per a^{-1}:

$$ax \equiv_n b \Rightarrow a^{-1}(ax) \equiv_n a^{-1}b \Rightarrow (a^{-1}a)x \equiv_n a^{-1}b \Rightarrow 1x \equiv_n a^{-1}b \Rightarrow x \equiv_n a^{-1}b\,.$$

In altre parole, se esiste l'inverso di a la soluzione può essere trovata esattamente come si risolve normalmente l'equazione $ax = b$ nei reali o nei razionali.

3.4 Applicazioni: cenni alla crittografia e ai numeri casuali

Vedremo adesso alcune importanti applicazioni pratiche dell'aritmetica modulare e dei teoremi visti nel capitolo precedente.
La prima di tali applicazioni riguarda la crittografia, ovvero lo studio delle tecniche con le quali un mittente può spedire un messaggio rendendolo illeggibile a tutti tranne al destinatari, l'unico a conoscere il modo di decriptarlo.

Per capire in che modo l'aritmetica modulare possa entrare in questo discorso, supponiamo di voler criptare un messaggio scritto con le 21 lettere dell'alfabeto italiano

$$A, B, C, D, \ldots, T, U, V, Z.$$

Per farlo potremmo far corrispondere a ogni lettera dell'alfabeto un'altra lettera e poi sostituire le lettere che compaiono nel messaggio dato in base alla corrispondenza scelta. Se, per esempio, la corrispondenza che scegliamo manda A in P, D in V, T in L , la parola $DATA$ verrà trasformata nella parola $VPLP$.
Il destinatario, per decriptare tale messaggio, deve essere a conoscenza della corrispondenza scelta e deve applicarla a ritroso per ottenere il messaggio originale.
Chiaramente, quella che stiamo chiamando corrispondenza tra le lettere dell'alfabeto deve essere più precisamente una corrispondenza biunivoca di tali lettere, ovvero una funzione biiettiva tra l'alfabeto e se stesso (in altre parole, una permutazione delle lettere dell'alfabeto). Infatti se la corrispondenza non fosse iniettiva, cioè se per esempio mandassimo A in P e anche O in P, il destinatario leggendo $VPLP$ non saprebbe se

sostituire una A o una O al posto della P. In questo modo non saprebbe se il messaggio originale era per esempio $DATA$, $DATO$, $DOTO$ o $DOTA$.
Quindi, una qualunque permutazione realizza una cifratura del messaggio, e la decifratura si realizza applicando la permutazione inversa (che esiste ben definita perché una permutazione è una funzione biiettiva, ovvero invertibile).

Ora, tra tutte le permutazioni del nostro insieme di 21 elementi può essere conveniente (per comodità di comunicazione con il destinatario) scegliere quelle che possono essere descritte tramite una regola generale, senza dover per forza dare il corrispondente di ogni lettera.
Ad esempio, è noto che Cesare per cifrare i suoi messaggi usava la permutazione che si ottiene associando ad ogni lettera quella che si trova 3 posizioni più avanti nell'alfabeto. Più precisamente, alla A associa la D, alla B la E, e così via, fino alla T, a cui verrà associata la Z, che è l'ultima lettera dell'alfabeto. A quel punto per le ultime tre lettere si riparte da capo: alla U si associa la A, alla V la B e alla Z la C.
Nel linguaggio del gruppo delle permutazioni imparato nel capitolo precedente, la permutazione usata da Cesare non è nient'altro che il prodotto di tre cicli disgiunti, ciascuno di lunghezza 7.

$$(A\ \ D\ \ G\ \ L\ \ O\ \ R\ \ U)(B\ \ E\ \ H\ \ M\ \ P\ \ S\ \ V)(C\ \ F\ \ I\ \ N\ \ Q\ \ T\ \ Z)$$

Benché fino a questo punto l'oggetto matematico che più si presta allo studio dei cifrari sembrino essere le permutazioni, vediamo ora che il modo più semplice di scrivere la corrispondenza di Cesare si ottiene usando il linguaggio delle congruenze.
A questo scopo, iniziamo con il tradurre numericamente l'alfabeto, assegnando a ogni lettera il numero della sua posizione nell'alfabeto: la A sarà quindi rappresentata dal numero 1, la B dal numero 2 e così via fino alla Z che sarà rappresentata dal numero 21.
Allora, la regola stabilita da Cesare per il suo cifrario dovrebbe essere semplicemente

$$x \mapsto x + 3\,,$$

cioè alla lettera corrispondente al numero x associamo la lettera che corrisponde alla posizione $x + 3$. Per esempio ad 1, cioè A, corrisponde $1 + 3 = 4$, cioè D, e così via.
Tuttavia, la formula sembra funzionare solo fino a $x = 18$ (ovvero T) a cui viene associata come abbiamo detto sopra $x + 3 = 21$ (ovvero la Z), dal momento che per $x = 19$ (cioè la U) abbiamo $x + 3 = 22$, che non corrisponde a nessuna lettera dell'alfabeto.
In realtà il problema si risolve se lavoriamo nell'aritmetica modulo 21 invece che nell'aritmetica usuale degli interi. Infatti, in $\mathbb{Z}_{21}$, 22 corrisponde a 1 (ovvero alla A) ed il cifrario di Cesare prevede di associare alla U la A. Analogamente, a $x = 20$ (cioè la V) corrisponde $x + 3 = 23$ che è congruo a 2 (ovvero la B) modulo 21 e a $x = 21$ (cioè la Z) corrisponde $x + 3 = 24$ che è congruo a 3 (ovvero la C) modulo 21.
Quindi la formula che descrive correttamente il cifrario di Cesare è

$$x \mapsto x + 3 \pmod{21}. \tag{3.49}$$

Il fatto che il cifrario di Cesare sia un procedimento crittografico corretto, cioè una funzione biiettiva sull'insieme delle lettere dell'alfabeto, equivale al fatto che la (3.49) sia una funzione biiettiva sull'insieme $\mathbb{Z}_{21} = \{1, 2, \ldots, 21\}$ (per comodità di notazione stiamo omettendo le parentesi quadre nell'indicare le classi).
Infatti, la biiettività della (3.49) discende dal fatto che se $x + 3 \equiv y \pmod{21}$, allora x è determinata da y semplicemente risolvendo l'equazione. In questo caso si fa semplicemente portando il 3 a secondo membro: $x \equiv y - 3 \pmod{21}$. Equivalentemente, essendo -3 congruo a $21-3 = 18$ modulo 21, potremmo scrivere $x = y+18 \pmod{21}$: ad esempio, se $y = 4$, che corrisponde alla D, decriptando abbiamo $x = y + 18 = 22$ che è congruo a 1 modulo 21, ovvero la A.

Ispirati da questo, possiamo pensare di realizzare nuovi cifrari semplicemente sostituendo una qualunque funzione $f(x)$ al posto di $x + 3$ purché questa sia invertibile in $\mathbb{Z}_{21}$. Come primo esempio, possiamo considerare, per ogni b, la generalizzazione del cifrario di Cesare

$$x \mapsto x + b \pmod{21} \tag{3.50}$$

che corrisponde ad associare a ogni lettera che si trova nella posizione x dell'alfabeto la lettera che si trova b posizioni più avanti (se $b > 0$) o più indietro (se $b < 0$).
L'invertibilità di tale funzione si mostra esattamente come fatto per il cifrario di Cesare, ovvero basta osservare che, per ogni y, l'equazione $x + b \equiv y \pmod{21}$ ha unica soluzione data da $x \equiv y - b \pmod{21}$.
Un ulteriore passo avanti nella generalizzazione del cifrario di Cesare si ottiene come segue:

$$x \mapsto ax + b \pmod{21}. \tag{3.51}$$

Ad esempio, se $a = 5$ e $b = 3$, allora la lettera D, che corrisponde a $x = 4$, viene mandata in $5 \cdot 4 + 3 = 23 \pmod{21}$ ovvero $2 \pmod{21}$, che corrisponde alla lettera B.
Stavolta, però, dobbiamo fare più attenzione, in quanto non ogni scelta di a dà un procedimento crittografico corretto. Infatti, come abbiamo detto sopra, la funzione (3.51) deve essere biiettiva, e questo corrisponde a dire che per ogni y l'equazione $ax + b \equiv y \pmod{21}$ deve avere soluzione e questa deve essere unica.
Ora, portando b a secondo membro questo equivale a dire che deve essere unica la soluzione di $ax \equiv y - b \pmod{21}$.
Ma alla fine della sezione precedente abbiamo visto qual è la condizione per l'unicità di tale soluzione: a deve essere primo con 21. Questo infatti garantisce che a abbia un inverso a^{-1} modulo 21. Come abbiamo visto nel paragrafo precedente, la soluzione si ottiene moltiplicando entrambi i membri per a^{-1}:

$$x \equiv a^{-1}(y-b) \pmod{21}$$

che è la formula per decriptare il messaggio.
Un cifrario dato dalla (3.51) è detto *cifrario affine*.

Esempio 3.32. Supponiamo, sbagliando, di voler criptare un messaggio con la regola $x \mapsto 6x \pmod{21}$ (l'errore consiste nel fatto che si tratta di una formula del tipo (3.51) con $b = 0$ e $a = 6$ non primo con $n = 21$).
Allora, vediamo ad esempio che 1 (cioè la A) viene mandato in 6 (cioè F); ma anche H, che corrisponde a 8, viene mandato in $6x = 48 \equiv 6 \pmod{21}$, ovvero nuovamente la F.
Tale metodo di cifratura non è quindi corretto in quanto si hanno lettere diverse mandate nella stessa lettera, e il messaggio criptato sarebbe ambiguo.

Quando si trasmette un messaggio criptato si vuole che questo sia il più difficile possibile da decriptare per chi, non essendo il destinatario, lo intercettasse.
Ora, per decriptare un messaggio bisogna innanzitutto conoscere la funzione $x \mapsto f(x)$ con cui si cripta il messaggio, e poi saper calcolare l'inversa di tale funzione.
Per quello che riguarda la conoscenza della funzione f, supponendo che chi intercetta il messaggio sappia di che tipo di cifrario si tratta, ad esempio un cifrario affine (3.51), questi dovrebbe conoscere la coppia (a,b) per avere esplicitamente la funzione e invertirla. Tale coppia, e più in generale i parametri che servono per costruire la funzione con cui criptiamo il messaggio, si chiama la *chiave del cifrario*.
Chiaramente, più sono le chiavi possibili di un tipo di cifrario, più sarà difficile per chi lo intercetta di determinarle. Ad esempio, il cifrario di Cesare generalizzato $x \mapsto x + b \pmod{21}$ ha come chiave la b e questa può essere scelta esattamente in 21 modi ($b = 1, 2, \ldots, 21$). Escludendo $b = 21$ con il quale assoceremmo a ogni lettera se stessa, otteniamo 20 cifrari diversi.
Nel caso del cifrario affine (3.51), b può variare tra 1 e 21 mentre a può essere un qualunque intero positivo minore di 21 e primo con esso. Il numero di interi positivi minori di 21 e primi con esso è dato, come abbiamo visto a pagina 105, da $\phi(21)$, dove ϕ è la funzione di Eulero. Essendo $21 = 3 \cdot 7$, in base alla formula (3.37) abbiamo $\phi(21) = 21(1-1/3)(1-1/7) = 12$.
Quindi per un cifrario affine, tenendo conto che per ognuna delle 12 scelte di a abbiamo 20 scelte possibili di b, abbiamo $12 \cdot 21 = 252$ possibili chiavi. Da queste escludiamo però il caso $a = 1, b = 21$, che, essendo 21 congruente a 0 modulo 21, manda ogni numero in stesso. Quindi, chiunque sia interessato a decriptare il messaggio, sospettando che si tratti di un cifrario affine, potrebbe doverle provare tutte (il che chiaramente non è troppo complicato con i computer a disposizione oggi).

Il secondo aspetto della sicurezza di un codice, come abbiamo detto, riguarda la difficoltà di calcolare la funzione inversa di quella usata per la cifratura. A prima vista,

sembrerebbe che una volta nota la chiave, ovvero la funzione f, calcolare l'inversa di f sia una cosa relativamente semplice. In realtà stiamo per mostrare un codice, detto RSA[6], nel quale il calcolo della funzione inversa si rivela talmente complesso che, conoscendo solo in che modo il messaggio è stato criptato e la chiave, la decifratura risulta praticamente impossibile.
Vediamo i dettagli: supponiamo che il messaggio che vogliamo trasmettere sia un numero intero x (che può rappresentare una lettera dell'alfabeto ma anche una qualunque altra informazione). Il destinatario del messaggio comunica (pubblicamente) al mittente in che modo deve criptare il messaggio, che poi solo lui saprà decriptare invertendo la funzione usata. Più precisamente, il destinatario procede come segue.

(1) Sceglie due numeri primi p_1 e p_2 il più grandi possibile e li moltiplica ottenendo $N = p_1 p_2$.

(2) Calcola la funzione di Eulero $\phi(N)$ (in questo caso si ha $\phi(N) = (p_1-1)(p_2-1)$).

(3) Sceglie un intero positivo $e < \phi(N)$ e che sia primo con $\phi(N)$.

(4) Comunica al mittente, che vuole spedire l'intero x, di calcolare il valore di $x^e \pmod N$ e spedirglielo.

In altre parole, la funzione di cifratura è $x \mapsto y \equiv x^e \pmod N$.
In realtà, per la correttezza della cifratura, x deve essere minore di N, ma tale condizione non è difficile da realizzare visto che il destinatario ha ottenuto N moltiplicando due primi molto grandi.
Ora vediamo che, sotto le condizioni date, tale funzione è invertibile ma il calcolo della sua inversa richiede un'informazione aggiuntiva, oltre alla chiave (N, e) comunicata pubblicamente, che solo il destinatario possiede e che risulta estremamente difficile ricavare dalla chiave, anche con l'ausilio di un computer.
Più precisamente, il destinatario che ha ricevuto $y \equiv x^e \pmod N$ e lo decripta come segue. Dal momento che e è primo con $\phi(N)$, esso è invertibile modulo $\phi(N)$ e ne calcola il suo inverso, che chiamiamo d. A questo punto, affermiamo che la x può essere ricostruita dalla y tramite la formula

$$x = y^d \pmod N . \tag{3.52}$$

Prima di spiegare il perché, osserviamo che il calcolo di d, che è l'inverso di e modulo $\phi(N)$, richiede la conoscenza del valore $\phi(N)$, e qui sta esattamente la difficoltà di violare il codice. Infatti, tutti conoscono N, che fa parte della chiave pubblica, ma il calcolo di $\phi(N) = (p_1 - 1)(p_2 - 1)$ richiede la conoscenza della decomposizione $N = p_1 p_2$

[6] Si tratta di un esempio di cosiddetta *crittografia a chiave pubblica* che deve il nome RSA alle iniziali dei nomi Rivest, Shamir e Adleman degli studiosi che l'hanno introdotta nel 1977.

di N come prodotto di primi, e questo è un calcolo estremamente difficile, per numeri grandi, anche per computer molto potenti.

Ora spieghiamo perché la (3.52) fornisca l'inversa che serve per decifrare il messaggio cifrato. Dal momento che $y \equiv x^e \pmod N$, elevando entrambi i membri della congruenza alla potenza d si ha

$$y^d \equiv (x^e)^d \equiv x^{ed} \pmod N . \tag{3.53}$$

Quindi, per concludere che $y^d \equiv x \pmod N$ ci basta mostrare che

$$x^{ed} \equiv x \pmod N . \tag{3.54}$$

Per dimostrare la (3.54) bisogna distinguere due casi: quello in cui x è primo con N e quello in cui x non è primo con N. Per semplicità, mostreremo la (3.54) solo nel primo caso.
In tal caso, si sfrutta il seguente cosiddetto *teorema di Eulero*, che non dimostriamo:

Teorema 3.33 (Teorema di Eulero)**.** *Se x e N sono interi primi tra loro, allora*

$$x^{\phi(N)} \equiv 1 \pmod N \tag{3.55}$$

In che modo il teorema di Eulero ci serve per dimostrare la (3.54)?
Per definizione, d è l'inverso di e modulo $\phi(N)$, ovvero $ed \equiv_{\phi(N)} 1$. In altre parole questo significa che esiste un intero k per cui $ed = 1 + k\phi(N)$. Ma allora si ha

$$x^{ed} \equiv x^{1+k\phi(N)} \pmod N . \tag{3.56}$$

Ora, in base al teorema di Eulero, $x^{\phi(N)} \equiv 1 \pmod N$, ovvero, elevando entrambi i membri di tale congruenza alla potenza k si ha

$$x^{k\phi(N)} \equiv (x^{\phi(N)})^k \equiv 1^k \equiv 1 \pmod N \tag{3.57}$$

ovvero anche $x^{k\phi(N)} \equiv 1 \pmod N$. Ma allora, moltiplicando entrambi i membri di questa uguaglianza per x si ottiene

$$x \cdot x^{k\phi(N)} \equiv x \cdot 1 \pmod N$$

ovvero

$$x^{1+k\phi(N)} \equiv x \pmod N . \tag{3.58}$$

Combinando la (3.56) e la (3.58) si ha proprio la (3.54), e abbiamo concluso.

Esempio 3.34. Supponiamo di prendere $p_1 = 7$, $p_2 = 5$: allora $N = 35$ e $\phi(N) = (p_1 - 1)(p_2 - 1) = 6 \cdot 4 = 24$. Scegliamo allora $e = 5$ che come previsto è primo con $\phi(N) = 24$.
Supponiamo che il mittente voglia mandare al destinatario il messaggio $x = 3$. Allora, il destinatario gli chiede prima di criptarlo secondo la formula $x \mapsto x^e \pmod N$, dopo avergli comunicato pubblicamente la chiave $(N, e) = (35, 5)$. Si ha

$$x^e = 3^5 = 243 \equiv 33 \pmod{35}$$

Il mittente comunica quindi al destinatario il messaggio criptato $y = 33$.
A questo punto, per decriptarlo, il destinatario calcola prima l'inverso di $e = 5$ modulo $\phi(N) = 24$. Essendo

$$24 = 5 \cdot 4 + 4$$
$$5 = 4 \cdot 1 + 1$$
$$4 = 4 \cdot 1$$

dalla seconda divisione troviamo $1 = 5 - 4$, e sostituendo $4 = 24 - 5 \cdot 4$ (ricavato dalla prima) si trova

$$1 = 5 - (24 - 5 \cdot 4) = 5 - 24 + 5 \cdot 4 = -24 + 5 \cdot 5$$

che ci dice che 5 è l'inverso di 5 modulo 24, ovvero $d = 5$.
Allora, il destinatario decripta il messaggio calcolando

$$33^5 = 39.135.393 \equiv 3 \pmod{35}$$

(in quanto $39.135.393 = 35 \cdot 1.118.154 + 3$) e ritrova quindi il messaggio originale $x = 3$.

Osservazione 3.35. Se N è un numero primo P, allora sapendo che $\phi(P) = P - 1$ si trova che il teorema di Eulero in questo caso si scrive

$$x^{P-1} \equiv 1 \pmod P \tag{3.59}$$

(sempre valida sotto l'ipotesi che x sia primo con P).
Questo risultato è noto sotto il nome di *piccolo teorema di Fermat*.
Moltiplicando entrambi i membri della (3.59) per x si ottiene

$$x^P \equiv x \pmod P \tag{3.60}$$

che ha il vantaggio di essere vera anche senza ipotesi su x. Infatti, se x non è primo con P, essendo P primo, deve essere un multiplo di P. Questo implica che sia x^P che x sono

congruenti a zero modulo P, e quindi congruenti tra loro. A volte quando si parla di piccolo teorema di Fermat si intende la (3.60).

Concludiamo il capitolo accennando ad un'altra importante applicazione delle congruenze. Spesso, nella pratica, serve avere a disposizione sequenze di numeri casuali o meglio *pseudocasuali*. Per capire di cosa si tratta, facciamo un esempio. I sistemi di identificazione online di certe banche sono basati sull'uso di una chiavetta che premendo un tasto fornisce una sequenza di numeri che l'utente deve inserire nella pagina di accesso, assieme ad altre credenziali. Succede è che dall'altra parte un meccanismo genera, in base allo stesso algoritmo usato nella chiavetta, la stessa sequenza di numeri, riconoscendo così l'utente.
Affinché il sistema sia sicuro, la sequenza di numeri generata deve essere "apparentemente casuale". Chiaramente non lo è davvero visto che sia l'utente che il sito della banca usano un algoritmo per generare la stessa sequenza e riconoscersi, ma deve essere tale da avere caratteristiche simili a quelle di una sequenza di numeri casuali (si parla allora appunto di numeri pseudocasuali), in modo che individuare l'algoritmo usato e entrare anche senza la chiavetta sia il più difficile possibile per chi volesse violare il sistema. La definizione precisa di cosa significhi "caratteristiche simili a quelle di una sequenza casuale" spetta alla statistica, ma non approfondiamo questo aspetto: basti sapere che esiste una definizione precisa e un modo di testare se una sequenza ha queste caratteristiche.
Solitamente un algoritmo che generi una sequenza funziona nel seguente modo: parte da un numero iniziale x_0 (detto il *seme*) e genera la sequenza applicando ripetutamente una funzione f scelta. In altre parole la sequenza è data da

$$x_0, x_1 = f(x_0), x_2 = f(x_1), x_3 = f(x_2) \dots$$

ovvero è data dalla regola[7] che $x_{i+1} = f(x_i)$.
Ad esempio, se f fosse la funzione $f(x) = x + 1$, partendo da $x = 2$ avremmo

$$x_0 = 2, x_1 = f(2) = 3, x_2 = f(3) = 4, x_3 = f(x_2) = 5 \dots$$

Chiaramente se si usasse questa funzione come generatrice dei numeri sulla chiavetta si otterrebbero sempre sequenze di numeri consecutivi, e chi riuscisse a vedere i numeri usati capirebbe subito il meccanismo e non avrebbe difficoltà a violare il sistema.
Nuovamente l'aritmetica modulare ci viene in aiuto: è tramite le congruenze che si riescono a generare sequenze di numeri pseudocasuali soddisfacenti. Ad esempio, si fissa un certo n e si lavora modulo n, definendo una sequenza di numeri tramite la formula

[7] Si dice che abbiamo definito la sequenza *induttivamente*: infatti, la stiamo definendo per $i = 0$ dando il seme x_0 e poi tramite la $x_{i+1} = f(x_i)$ stiamo dicendo che, una volta che è definita per i, risulta definita anche per $i + 1$. In base al principio di induzione, la sequenza è allora definita per ogni numero naturale i.

$$x_{i+1} = ax_i + b \pmod{n}. \tag{3.61}$$

Sotto opportune ipotesi su a e b (omettiamo i dettagli per semplicità) si dimostra che la (3.61) genera, a partire da un seme x_0, una sequenza di numeri pseudocasuali. La (3.61) si chiama *generatore lineare* di numeri pseudocasuali

La combinazione dell'uso di generatori di numeri casuali e di crittografia RSA è alla base di molte importanti procedure di identificazione on line (ad esempio, la firma digitale).

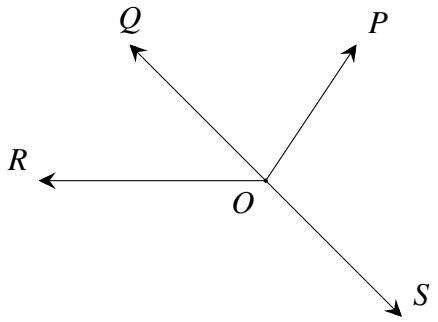
Capitolo 4

Vettori, coordinate e geometria

In questo capitolo vedremo come, passando per i cosiddetti vettori geometrici, si possano introdurre nel piano e nello spazio sistemi di coordinate che permettono la traduzione di oggetti e problemi geometrici in oggetti numerici e problemi algebrici. Tale costruzione, oltre ad essere di fondamentale importanza per le applicazioni alla grafica, ci permetterà di introdurre alcune nozioni basilari della cosiddetta *algebra lineare* che saranno approfondite nei capitoli successivi e che presentano a loro volta numerose applicazioni in tutte le scienze applicate e in informatica in particolare.

4.1 Vettori geometrici

Un **vettore geometrico applicato nel piano** è un segmento orientato che va da un punto fissato O verso un secondo punto P del piano, come nel disegno seguente[1].

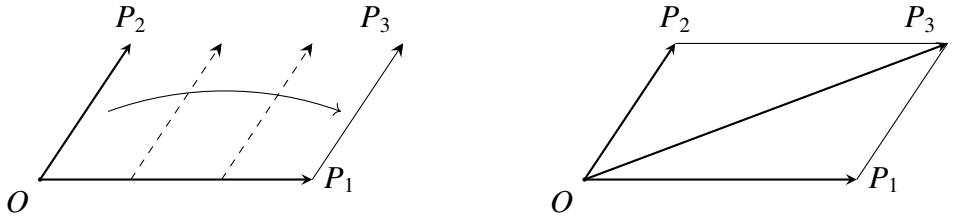


Analogamente, parliamo di **vettori geometrici applicati nello spazio** se il punto P (e quindi il segmento) è libero di variare in tutto lo spazio tridimensionale. In entrambi i casi il vettore sarà denotato con $\vec{OP}$. Si noti che il punto finale P può essere uguale a O, ovvero il vettore può essere "schiacciato" sul punto O.

[1] L'orientazione viene messa in evidenza con una freccia

I vettori orientati sono importanti in fisica, dove vengono usati per rappresentare forze applicate sul punto O. Ad esempio si può immaginare che in O si trovi un oggetto sul quale viene esercitata una forza che lo "trascina" nella direzione e nel verso dati dalla freccia, mentre l'intensità della forza esercitata è rappresentata dalla lunghezza del segmento.

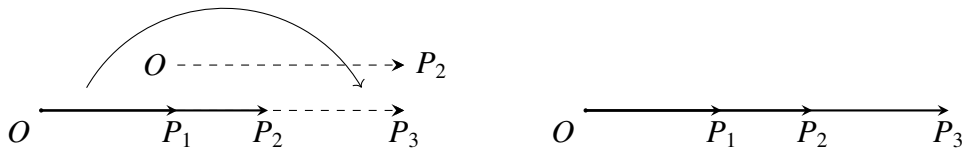
Con questa interpretazione, è naturale chiedersi cosa succede quando su O si esercitano contemporaneamente due forze rappresentate da due vettori geometrici $\vec{OP}_1$ e $\vec{OP}_2$. In tal caso si verifica sperimentalmente che la forza totale "percepita" da O è quella rappresentata dal terzo vettore $\vec{OP}_3$ ottenuto nel modo seguente. Si trasla parallelamente uno dei due vettori, diciamo $\vec{OP}_2$, in modo da spostare il suo punto di applicazione da O nel punto finale P_1 dell'altro vettore e si individua così il punto P_3. Si noti che si otterrebbe lo stesso punto se traslassimo invece $\vec{OP}_1$ spostando il suo punto di applicazione in P_2. Dal momento che $\vec{OP}_3$ rappresenta la forza totale esercitata su O quando si applicano contemporaneamente $\vec{OP}_1$ e $\vec{OP}_2$, è naturale pensare $\vec{OP}_3$ come la somma di $\vec{OP}_1$ e $\vec{OP}_2$. Scriveremo cioè



$$\vec{OP}_3 = \vec{OP}_1 + \vec{OP}_2$$

definendo in tal modo un'operazione di somma sull'insieme dei vettori geometrici (del piano o dello spazio).

Per vettori che non hanno la stessa direzione, si osserva che OP_3 è la diagonale del parallelogramma che ha OP_1 e OP_2 come lati (si parla infatti anche di **regola del parallelogramma**). Il metodo descritto funziona comunque anche per sommare due vettori che hanno la stessa direzione:



$$\vec{OP}_3 = \vec{OP}_1 + \vec{OP}_2$$

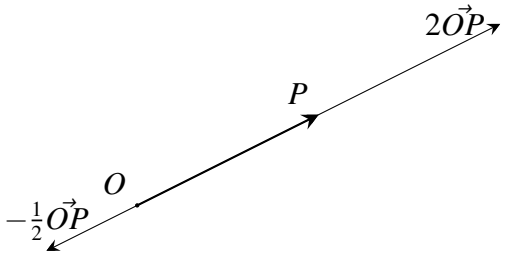
Un'altra operazione che si può introdurre nell'insieme dei vettori geometrici è il prodotto del vettore per un numero reale: nel contesto delle forze, l'idea è quella di rappresentare una variazione dell'intensità e eventualmente del verso della forza rappresentata dal vettore.

Più precisamente, dati un vettore geometrico $\vec{OP}$ e un numero reale $c \in \mathbb{R}$, definiamo $c \cdot \vec{OP}$ come il vettore che sta sulla stessa retta a cui appartiene $\vec{OP}$, ma avente

(1) stesso verso e lunghezza c volte la lunghezza di $\vec{OP}$, se c è positivo;

(2) verso opposto e lunghezza $-c$ volte quella di $\vec{OP}$, se c è negativo;

(3) lunghezza nulla se $c = 0$, cioè $0\vec{OP} = \vec{OO}$.



In questo contesto, i numeri reali si chiamano anche **scalari**.

Ora, come sappiamo e abbiamo visto in vari esempi nei capitoli precedenti, ogni volta che si introducono delle operazioni su un insieme è bene conoscere le proprietà di cui godono tali operazioni. Per quelle che abbiamo appena introdotto, si può verificare (con dimostrazioni di geometria euclidea, che omettiamo) che valgono le seguenti:

(1) La somma è *associativa*:

$$(\vec{OP}_1 + \vec{OP}_2) + \vec{OP}_3 = \vec{OP}_1 + (\vec{OP}_2 + \vec{OP}_3)\,. \tag{4.1}$$

(2) La somma è *commutativa*:

$$\vec{OP}_1 + \vec{OP}_2 = \vec{OP}_2 + \vec{OP}_1\,. \tag{4.2}$$

(3) Il vettore $\vec{OO}$ è l'elemento neutro della somma:

$$\vec{OP} + \vec{OO} = \vec{OO} + \vec{OP} = \vec{OP}\,. \tag{4.3}$$

(4) Per ogni vettore $\vec{OP}$, il vettore $(-1)\vec{OP}$ (ovvero il vettore che si ottiene da $\vec{OP}$ semplicemente invertendo il verso, senza modificare direzione e lunghezza) è il suo inverso additivo (o opposto):

$$\vec{OP} + (-1)\vec{OP} = (-1)\vec{OP} + \vec{OP} = \vec{OO} \tag{4.4}$$

(5) Dati due numeri reali c_1, c_2 e un vettore $\vec{OP}$, si ha

$$c_1(c_2\vec{OP}) = (c_1c_2)\vec{OP} \tag{4.5}$$

(una sorta di proprietà associativa del prodotto).

(6) Per ogni vettore $\vec{OP}$, si ha

$$1\vec{OP} = \vec{OP} \tag{4.6}$$

ovvero il numero 1 funge da elemento neutro rispetto al prodotto per scalari.

(7) Dati due numeri reali c_1, c_2 e un vettore $\vec{OP}$, si ha

$$(c_1 + c_2)\vec{OP} = c_1\vec{OP} + c_2\vec{OP}\,. \tag{4.7}$$

(8) Dati un numero reale c e due vettori $\vec{OP}_1$, $\vec{OP}_2$ si ha

$$c(\vec{OP}_1 + \vec{OP}_2) = c\vec{OP}_1 + c\vec{OP}_2. \tag{4.8}$$

(le ultime due ci dicono che vale la proprietà distributiva rispetto alla somma di numeri reali o rispetto alla somma di vettori).

Osservazione 4.1. Ispirati dall'esempio dei vettori geometrici, si definisce **spazio vettoriale** un qualunque insieme i cui elementi possano essere sommati tra loro e moltiplicati per numeri reali (o più in generale per gli elementi di un campo), in modo che valgano proprietà analoghe alle (4.1)-(4.8). Tuttavia in questo corso non approfondiamo tale approccio astratto.

Come esempio di applicazione delle proprietà appena elencate, mostriamo che in una uguaglianza tra vettori, esattamente come si fa in un'uguaglianza tra numeri, si possono"spostare i vettori" da un membro all'altro cambiandoli di segno:

$$\vec{OP}_1 + \vec{OP}_2 = \vec{OP}_3 \;\rightarrow\; \vec{OP}_1 = \vec{OP}_3 - \vec{OP}_2$$

dove, come si fa anche per i numeri, stiamo scrivendo $\vec{OP}_3 - \vec{OP}_2$ come forma semplificata di $\vec{OP}_3 + (-1)\vec{OP}_2$.
Per vederlo, sommiamo il vettore $(-1)\vec{OP}_2$ ad entrambi i membri:

$$(\vec{OP}_1 + \vec{OP}_2) + (-1)\vec{OP}_2 = \vec{OP}_3 + (-1)\vec{OP}_2$$

e applichiamo la proprietà associativa (4.1) al primo membro:

$$\vec{OP}_1 + (\vec{OP}_2 + (-1)\vec{OP}_2) = \vec{OP}_3 + (-1)\vec{OP}_2\,.$$

Ora applichiamo la proprietà (4.4) che afferma che $(-1)\vec{OP}_2$ è l'opposto di $\vec{OP}_2$ per ottenere

$$\vec{OP}_1 + \vec{OO} = \vec{OP}_3 + (-1)\vec{OP}_2$$

e infine applichiamo la (4.3) che ci dice che il vettore nullo funge da elemento neutro:

$$\vec{OP}_1 = \vec{OP}_3 + (-1)\vec{OP}_2$$

come volevamo.

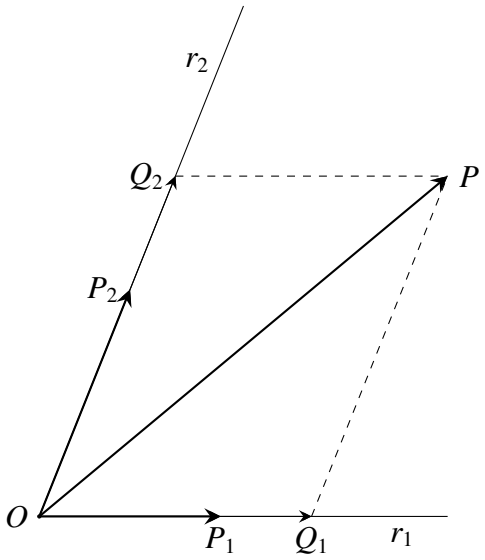
4.2 Coordinate

Consideriamo ora due vettori geometrici $\vec{OP}_1$ e $\vec{OP}_2$ nel piano, e supponiamo che $\vec{OP}_1$ e $\vec{OP}_2$ non abbiano la stessa direzione.
Affermiamo che ogni altro vettore $\vec{OP}$ può essere ottenuto sommando multipli opportuni di $\vec{OP}_1$ e $\vec{OP}_2$, ovvero

$$\vec{OP} = c_1\vec{OP}_1 + c_2\vec{OP}_2$$

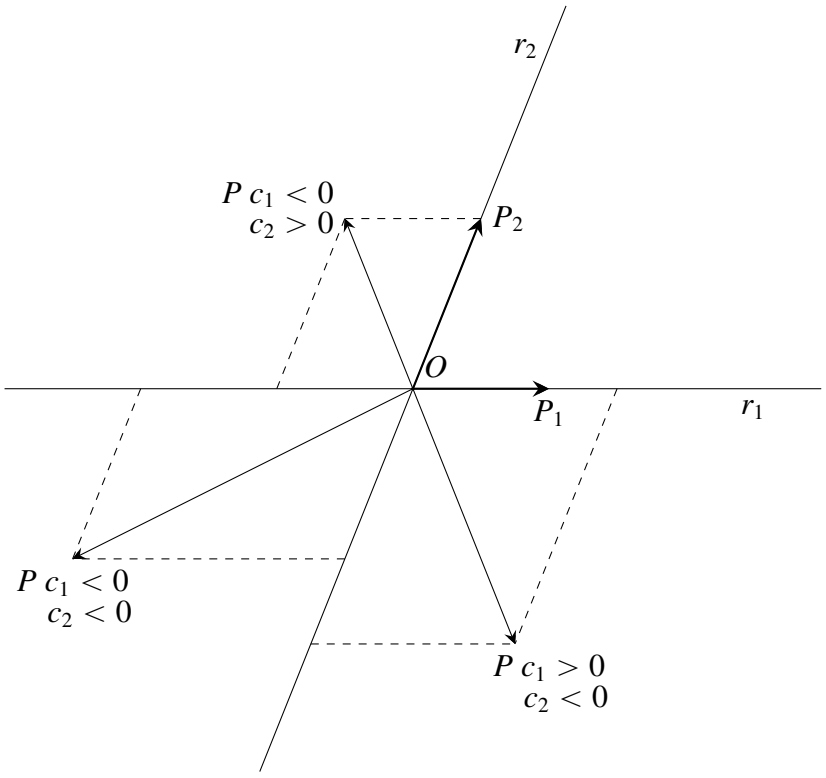
dove c_1, c_2 sono opportuni numeri reali.

Questo può essere visto graficamente. Come nel disegno seguente, prolunghiamo i vettori $\vec{OP}_1$ e $\vec{OP}_2$ disegnando le due rette r_1 e r_2. Proiettiamo quindi il punto P su r_1 seguendo la direzione parallela a $\vec{OP}_2$, e chiamiamo il punto proiettato Q_1. Analogamente, proiettiamo il punto P su r_2 seguendo la direzione parallela a $\vec{OP}_1$, e chiamiamo il punto proiettato Q_2.



Avendo fatto le due proiezioni parallelamente a $\vec{OP}_1$ e $\vec{OP}_2$, risulta determinato un parallelogramma che ha $\vec{OQ}_1$ e $\vec{OQ}_2$ come lati e $\vec{OP}$ come diagonale. Quindi per definizione di somma tra vettori geometrici si ha $\vec{OP} = \vec{OQ}_1 + \vec{OQ}_2$.
Ma dal momento che $\vec{OQ}_1$ si trova sulla stessa retta di $\vec{OP}_1$ e per definizione di prodotto dei vettori per i numeri reali, esisterà un numero reale c_1 tale che $\vec{OQ}_1 = c_1\vec{OP}_1$. Qui c_1 dipende solamente dal rapporto tra la lunghezza di $\vec{OQ}_1$ e quella di $\vec{OP}_1$ e dal loro verso.
Analogamente esisterà un numero reale c_2 tale che $\vec{OQ}_2 = c_2\vec{OP}_2$ (dove c_2 dipende dal rapporto tra la lunghezza di $\vec{OQ}_2$ e quella di $\vec{OP}_2$ e dal loro verso).
Abbiamo quindi concluso, come volevamo, che $\vec{OP} = c_1\vec{OP}_1 + c_2\vec{OP}_2$.

Si noti che, nella situazione considerata nel disegno, $c_1, c_2 > 0$ in quanto $\vec{OQ}_1$ e $\vec{OQ}_2$ hanno lo stesso verso di $\vec{OP}_1$ e $\vec{OP}_2$ rispettivamente. In generale, la stessa costruzione può essere effettuata per qualunque vettore $\vec{OP}$ del piano e i coefficienti c_1 e c_2 potranno anche essere negativi[2] a seconda del quadrante nel quale si trova $\vec{OP}$, ovvero a seconda che la proiezione di P sulle rette r_1, r_2 cada dalla stessa parte o dalla parte opposta dei punti P_1 e P_2.

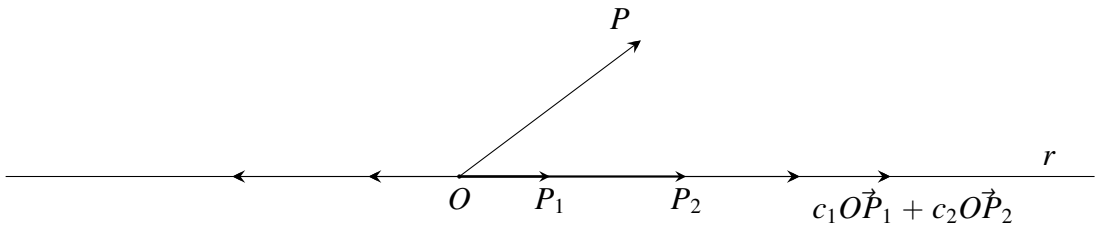


In ogni caso si avrà sempre $\vec{OP} = c_1\vec{OP}_1 + c_2\vec{OP}_2$.

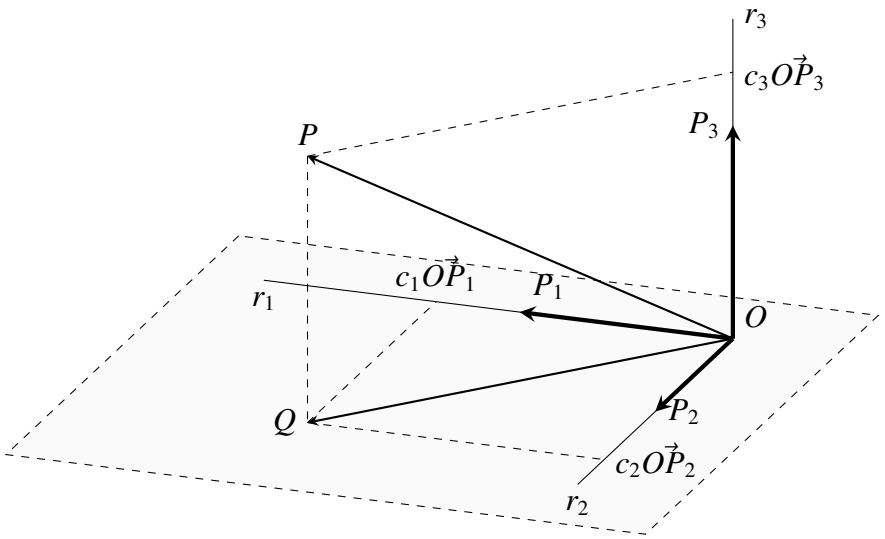
[2]Può essere anche $c_1 = 0$ o $c_2 = 0$. Nel primo caso, si ha $\vec{OP} = c_2\vec{OP}_2$, nel secondo $\vec{OP} = c_1\vec{OP}_1$. Cioè $\vec{OP}$ non sta all'interno di uno dei quadranti in cui le rette r_1, r_2 dividono il piano, ma sta sulla retta r_2 (se $\vec{OP} = c_2\vec{OP}_2$) o sulla retta r_1 (se $\vec{OP} = c_1\vec{OP}_1$).

Definizione 4.2. Una coppia $\vec{OP}_1$ e $\vec{OP}_2$ di vettori non allineati nel piano si dice **base**. La coppia (c_1, c_2) di numeri reali tale che $\vec{OP} = c_1\vec{OP}_1 + c_2\vec{OP}_2$ si dice la **coppia delle coordinate** del vettore $\vec{OP}$ rispetto a tale base (sottointendendo la base fissata, scriveremo anche $\vec{OP} \equiv (c_1, c_2)$).

Osservazione 4.3. Si noti che se i due vettori $\vec{OP}_1$ e $\vec{OP}_2$ *stanno sulla stessa retta r*, allora non è più vero che ogni vettore $\vec{OP}$ si può scrivere come loro combinazione. Infatti, qualunque multiplo $c_1\vec{OP}_1$ di $\vec{OP}_1$ sta ancora sulla retta r. Analogamente qualunque multiplo $c_2\vec{OP}_2$ di $\vec{OP}_2$ sta su r. Quindi starà su r anche la somma $c_1\vec{OP}_1 + c_2\vec{OP}_2$, e non riusciremo mai a ottenere i vettori $\vec{OP}$ che stanno fuori dalla retta come combinazione di $\vec{OP}_1$ e $\vec{OP}_2$.



Un risultato analogo a quello appena visto per i vettori nel piano può essere ottenuto anche nel caso dei vettori geometrici nello spazio tridimensionale. In questo caso non dobbiamo però partire da una coppia di vettori non allineati ma da una terna di vettori $\vec{OP}_1$, $\vec{OP}_2$ e $\vec{OP}_3$ *che non stiano tutti e tre sullo stesso piano*. In questo caso è facile vedere graficamente, utilizzando proiezioni come abbiamo fatto nel caso di due vettori nel piano, che ogni vettore $\vec{OP}$ dello spazio può essere scritto come combinazione $c_1\vec{OP}_1 + c_2\vec{OP}_2 + c_3\vec{OP}_3$.



Come rappresentato nel disegno, si proietta il punto P sul piano su cui stanno $\vec{OP}_1$ e $\vec{OP}_2$ seguendo la direzione di $\vec{OP}_3$ e si individua così un punto Q. Proiettando poi P sulla retta r_3 parallelamente al vettore $\vec{OQ}$, risulta individuato un parallelogramma, che ci dice che $\vec{OP}$ si scrive come somma $\vec{OP} = \vec{OQ} + c_3\vec{OP}_3$ di $\vec{OQ}$ e di un opportuno multiplo $c_3\vec{OP}_3$ di $\vec{OP}_3$. A questo punto si osserva che $\vec{OQ}$, stando sul piano di $\vec{OP}_1$ e $\vec{OP}_2$ si scriverà come loro combinazione lineare $\vec{OQ} = c_1\vec{OP}_1 + c_2\vec{OP}_2$, e sostituendo nella $\vec{OP} = \vec{OQ} + c_3\vec{OP}_3$ si conclude che $\vec{OQ} = c_1\vec{OP}_1 + c_2\vec{OP}_2 + c_3\vec{OP}_3$.

In modo analogo a quanto già fatto per vettori geometrici del piano, possiamo allora dare la seguente

Definizione 4.4. Una terna $\vec{OP}_1$, $\vec{OP}_2$, $\vec{OP}_3$ di vettori non complanari nello spazio si dice **base**. La terna (c_1, c_2, c_3) di numeri reali tale che $\vec{OP} = c_1\vec{OP}_1 + c_2\vec{OP}_2 + c_3\vec{OP}_3$ si dice la **terna delle coordinate** del vettore $\vec{OP}$ rispetto a tale base (sottointendendo la base fissata, scriveremo anche $\vec{OP} \equiv (c_1, c_2, c_3)$).

L'importanza delle coordinate consiste nel fatto che esse, rappresentando i vettori mediante coppie o terne di numeri, ci permettono di tradurre in calcoli numerici i calcoli tra vettori. Innanzitutto questa è un'importante semplificazione per noi, in quanto è più semplice lavorare con le coordinate che con costruzioni o dimostrazioni di geometria euclidea che sarebbero altrimenti necessarie per lavorare con i vettori, che sono oggetti geometrici. Inoltre, tale traduzione è indispensabile se si vogliono far trattare tali problemi a un computer.
Un esempio di tale semplificazione è dato dal seguente risultato.

Proposizione 4.5. *Si fissi una base B nell'insieme dei vettori del piano. Allora*

(1) Se $\vec{OP}$ e $\vec{OP}'$ hanno coordinate rispettivamente (x_1, x_2) e (x'_1, x'_2) rispetto a B, le coordinate di $\vec{OP} + \vec{OP}'$ rispetto a B sono date dalla coppia $(x_1 + x'_1, x_2 + x'_2)$ ottenuta sommando componente per componente le coppie delle coordinate dei due vettori.

(2) Se $\vec{OP}$ ha coordinate (x_1, x_2) rispetto a B e $c \in \mathbb{R}$ è un numero reale, allora le coordinate di $c\vec{OP}$ sono date dalla coppia (cx_1, cx_2) ottenuta moltiplicando per c le componenti delle coordinate di $\vec{OP}$.

Dimostrazione. Siano $\vec{OP}_1$ e $\vec{OP}_2$ i vettori che costituiscono la base B fissata. Per definizione il fatto che $\vec{OP}$ abbia coordinate (x_1, x_2) rispetto a B significa che $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2$, e, analogamente, il fatto che $\vec{OP}'$ abbia coordinate (x'_1, x'_2) rispetto a B significa che $\vec{OP}' = x'_1\vec{OP}_1 + x'_2\vec{OP}_2$. Ma allora

$$\vec{OP} + \vec{OP}' = (x_1\vec{OP}_1 + x_2\vec{OP}_2) + (x'_1\vec{OP}_1 + x'_2\vec{OP}_2) =$$

(riordiniamo gli addendi e raccogliamoli diversamente sfruttando le proprietà associativa e commutativa della somma tra vettori)

$$= (x_1\vec{OP}_1 + x'_1\vec{OP}_1) + (x_2\vec{OP}_2 + x'_2\vec{OP}_2) =$$

(sfruttiamo la proprietà (4.7) sia nella prima parentesi che nella seconda)

$$= (x_1 + x'_1)\vec{OP}_1 + (x_2 + x'_2)\vec{OP}_2 \,.$$

Ma questo, per definizione di coordinate, significa proprio che le coordinate di $\vec{OP}+\vec{OP'}$ sono date dalla coppia $(x_1 + x'_1, x_2 + x'_2)$, come affermato nella (1).

Per dimostrare la (2), ricordiamo ancora le coordinate di $\vec{OP}$ sono date da (x_1, x_2) rispetto a B se abbiamo $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2$. Allora

$$c\vec{OP} = c(x_1\vec{OP}_1 + x_2\vec{OP}_2) =$$

(applichiamo la proprietà (4.8))

$$= c(x_1\vec{OP}_1) + c(x_2\vec{OP}_2) =$$

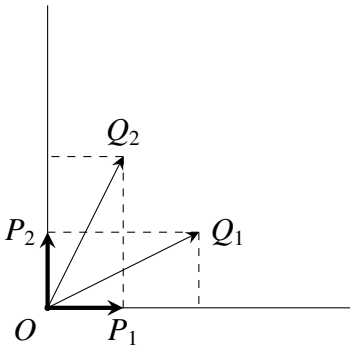
(applicando la (4.5) a entrambi gli addendi)

$$= (cx_1)\vec{OP}_1 + (cx_2)\vec{OP}_2$$

Ma questo, per definizione di coordinate, ci dice proprio che le coordinate di $c\vec{OP}$ sono date dalla coppia (cx_1, cx_2), come affermato nella (2).

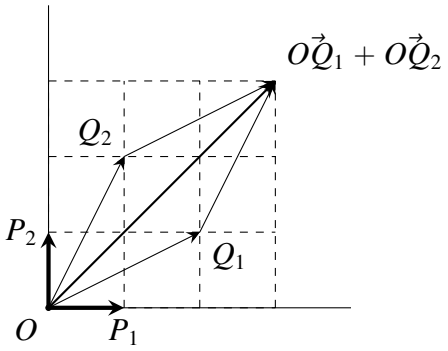
□

Esempio 4.6. Si considerino come nel disegno seguente la base $B = \{\vec{OP}_1, \vec{OP}_2\}$ e i due vettori $\vec{OQ}_1$ e $\vec{OQ}_2$.



Come si vede graficamente, si ha $\vec{OQ}_1 = 2\vec{OP}_1 + \vec{OP}_2$ e $\vec{OQ}_2 = \vec{OP}_1 + 2\vec{OP}_2$. Ovvero le coordinate di $\vec{OQ}_1$ rispetto alla base B sono date dalla coppia $(2, 1)$, mentre le coordinate di $\vec{OQ}_2$ rispetto alla base B sono date dalla coppia $(1, 2)$.

Allora, in base alla (1) della Proposizione 4.5, la somma $O\vec{Q}_1 + O\vec{Q}_2$ ha coordinate (sempre rispetto alla base B) date da $(2+1, 1+2) = (3,3)$, ovvero si ha $O\vec{Q}_1 + O\vec{Q}_2 = 3O\vec{P}_1 + 3O\vec{P}_2$.
In effetti, questo può essere verificato graficamente costruendo con la regola del parallelogramma la somma $O\vec{Q}_1 + O\vec{Q}_2$, come nel disegno seguente.



La cosa notevole è che siamo stati in grado di individuare il vettore $O\vec{Q}_1 + O\vec{Q}_2$ (in coordinate) con un semplicissimo conto aritmetico, ancora prima di disegnarlo con la costruzione geometrica del parallelogramma.

Osservazione 4.7. Affermazioni del tutto analoghe a quelle della Proposizione 4.5 valgono anche nel caso dei vettori nello spazio tridimensionale, e la dimostrazione è perfettamente analoga a quella appena vista per il piano. Più precisamente, si ha che fissata una base B nell'insieme dei vettori dello spazio tridimensionale, allora

(1) Se $\vec{OP}$ e $\vec{OP'}$ hanno coordinate rispettivamente (x_1, x_2, x_3) e (x'_1, x'_2, x'_3) rispetto a B, le coordinate di $\vec{OP}+\vec{OP'}$ rispetto a B sono date dalla terna $(x_1+x'_1, x_2+x'_2, x_3+x'_3)$ ottenuta sommando componente per componente le terne delle coordinate dei due vettori.

(2) Se $\vec{OP}$ ha coordinate (x_1, x_2, x_3) rispetto a B e $c \in \mathbb{R}$ è un numero reale, allora le coordinate di $c\vec{OP}$ sono date dalla terna (cx_1, cx_2, cx_3) ottenuta moltiplicando per c le componenti dellle coordinate di $\vec{OP}$.

Osservazione 4.8. Il fatto che, fissata una base B, ogni vettore del piano (risp. dello spazio tridimensionale) sia determinato in modo univoco da una coppia (risp. una terna) di numeri reali, ci dice che l'**insieme dei vettori del piano (risp. dello spazio tridimensionale) si può identificare con l'insieme** $\mathbb{R}^2$ **(risp.** $\mathbb{R}^3$**)**. La Proposizione 4.5 ci dice inoltre che lavorare con l'insieme dei vettori geometrici dotato della somma e del prodotto per scalari definito all'inizio del capitolo è la stessa cosa che lavorare con $\mathbb{R}^2$ o $\mathbb{R}^3$ dotati della somma e del prodotto per numeri reali definiti nel modo seguente:

$$(x_1, x_2) + (x'_1, x'_2) \coloneqq (x_1 + x'_1, x_2 + x'_2), \quad c(x_1, x_2) \coloneqq (cx_1, cx_2)$$

$$(x_1, x_2, x_3) + (x_1', x_2', x_3') \coloneqq (x_1 + x_1', x_2 + x_2', x_3 + x_3'), \quad c(x_1, x_2, x_3) \coloneqq (cx_1, cx_2, cx_3).$$

In generale, in uno spazio n-dimensionale avremmo dovuto usare lo spazio $\mathbb{R}^n$ dotato della somma e del prodotto definiti da

$$\begin{aligned}(x_1, x_2, \ldots, x_n) + (x_1', x_2', \ldots, x_n') &\coloneqq (x_1 + x_1', x_2 + x_2', \ldots, x_n + x_n') \\ c(x_1, x_2, \ldots, x_n) &\coloneqq (cx_1, cx_2, \ldots, cx_n)\,.\end{aligned}$$

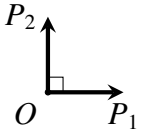
Notiamo che lo spazio $\mathbb{R}^n$ può rappresentare le coordinate di un ipotetico spazio n-dimensionale (che per alcune teorie fisiche è una possibilità concreta). Inoltre può rappresentare un qualunque insieme o sistema in cui servano più di 3 parametri indipendenti (che non necessariamente rappresentano una posizione spaziale). Ad esempio, $\mathbb{R}^4$ potrebbe rappresentare l'insieme dei punti dello spazio tridimensionale con un'informazione aggiuntiva (la quarta componente) che può, ad esempio, essere la temperatura in quel punto (o la frequenza di un colore da associare a quel punto, se stiamo traducendo numericamente un disegno).

4.3 Lunghezze e angoli

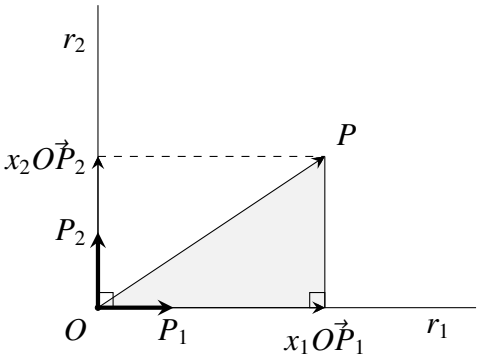
Lavorare in coordinate rispetto a una base ci permette di tradurre numericamente costruzioni geometriche con i vettori e risolvere numericamente problemi relativi ai vettori. Questo è vero qualunque sia la base scelta, tuttavia a seconda del problema che dobbiamo risolvere, alcune basi possono essere più convenienti di altre. Questo accade in particolare quando si vuole rispondere, lavorando in coordinate, alle domande seguenti: qual è la lunghezza di un vettore dato? qual è l'angolo tra due vettori dati?
In tal caso, le basi più convenienti da usare, come stiamo per vedere, sono quelle dela tipologia seguente.

Definizione 4.9. Una base di vettori del piano (o dello spazio) si dice **ortogonale** se i suoi vettori sono tra loro ortogonali (ovvero formano tra loro angoli di 90 gradi). Se, inoltre, i vettori della base hanno lunghezza 1 rispetto ad un'unità di misura scelta, allora la base si dice **ortonormale**.

Consideriamo una base ortonormale nel piano



Ora, consideriamo un vettore $\vec{OP}$di coordinate (x_1, x_2) rispetto a tale base (ovvero, per definizione di coordinate, $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2$). È possibile calcolare la lunghezza del vettore $\vec{OP}$ a partire dalle coordinate? Per rispondere a tale domanda, consideriamo il seguente disegno, nel quale è rappresentata la decomposizione $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2$.

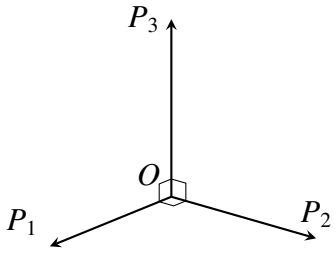


Come abbiamo visto nella sezione precedente proiettiamo P sulla retta r_1 che contiene $\vec{OP}_1$ seguendo la direzione di $\vec{OP}_2$. Dal momento che abbiamo scelto i vettori di base perpendicolari tale proiezione incontra r_1 con un angolo di 90 gradi. Si viene quindi a formare un triangolo rettangolo (evidenziato nel disegno) avente come ipotenusa proprio $\vec{OP}$ e al quale possiamo quindi applicare il teorema di Pitagora per calcolare la lunghezza di $\vec{OP}$, che denoteremo con $|\vec{OP}|$.
A questo scopo, notiamo che il cateto orizzontale di tale triangolo è dato dal vettore $x_1\vec{OP}_1$, e quindi la sua lunghezza è data dal prodotto di x_1 per la lunghezza di $\vec{OP}_1$. Per cui, avendo scelto i vettori di base di lunghezza unitaria, la lunghezza di tale cateto è semplicemente x_1. Per quanto riguarda il cateto verticale, esso per costruzione ha la stessa lunghezza del vettore $x_2\vec{OP}_2$, ovvero x_2 (in quanto $\vec{OP}_2$ ha lunghezza 1). Quindi per il teorema di Pitagora

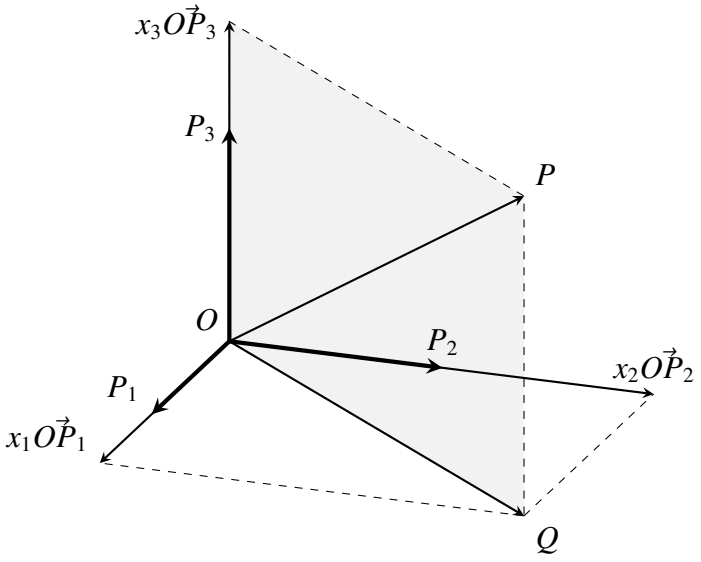
$$|\vec{OP}| = \sqrt{x_1^2 + x_2^2} \tag{4.9}$$

che rappresenta la formula cercata, overo fornisce la lunghezza di $\vec{OP}$ in funzione delle sue coordinate.
Si noti che nei ragionamenti che abbiamo fatto è stato fondamentale aver scelto una base fatta di vettori ortogonali (in modo da costruire un triangolo rettangolo a cui applicare il teorema di Pitagora) e di lunghezza 1 (in modo da esprimere le lunghezze dei cateti in funzione delle sole coordinate).

Vediamo adesso come si ottiene una formula analoga nello spazio tridimensionale. Sia $\vec{OP}_1, \vec{OP}_2, \vec{OP}_3$ una base ortonormale B per i vettori applicati nello spazio tridimensionale: B è composta di vettori ortogonali tra loro e di lunghezza 1.



Dato un vettore $\vec{OP}$ calcoliamone la lunghezza, che denotiamo ancora con $|\vec{OP}|$, in funzione delle sue coordinate x_1, x_2, x_3 rispetto alla base B scelta. Per definizione di coordinate, $\vec{OP}$ si decompone come somma $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2 + x_3\vec{OP}_3$, come nel disegno seguente.



La decomposizione è stata ottenuta graficamente come segue. Si proietta P sul piano su cui stanno P_1 e P_2 parallelamente a $\vec{OP}_3$ ottenendo il punto Q. Si noti che l'angolo in Q è retto perché $\vec{OP}_3$ è ortogonale sia a $\vec{OP}_1$ che a $\vec{OP}_2$. Si ottiene così un rettangolo, ombreggiato nel disegno, da cui si evince che $\vec{OP} = \vec{OQ} + x_3\vec{OP}_3$. Dal momento che $\vec{OQ}$ giace sul piano di $\vec{OP}_1$ e $\vec{OP}_2$ lo si può decomporre come $\vec{OQ} = x_1\vec{OP}_1 + x_2\vec{OP}_2$ (sempre ottenendo angoli retti in quanto $\vec{OP}_1$ e $\vec{OP}_2$ sono perpendicolari). Quindi $\vec{OP} = \vec{OQ} + x_3\vec{OP}_3 = x_1\vec{OP}_1 + x_2\vec{OP}_2 + x_3\vec{OP}_3$ come volevamo.

Ora, essendo OP l'ipotenusa del triangolo OPQ rettangolo in Q, per il teorema di Pitagora avremo

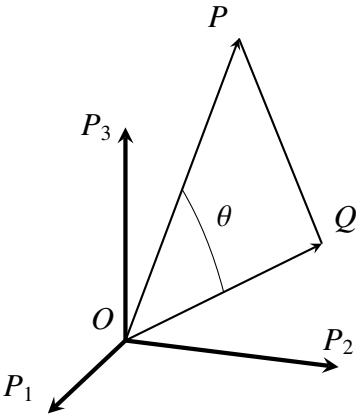
$$|OP|^2 = |OQ|^2 + |PQ|^2 \tag{4.10}$$

Ora da una parte, il segmento PQ, essendo un lato del rettangolo ombreggiato in figura, è lungo esattamente quanto il vettore $x_3\vec{OP}_3$, ovvero x_3 (in quanto $\vec{OP}_3$ ha lunghezza 1). D'altra parte, OQ è la diagonale del rettangolo che ha come lati i vettori $x_1\vec{OP}_1$ e $x_2\vec{OP}_2$ di lunghezze rispettivamente x_1 e x_2 (in quanto $\vec{OP}_1$ e $\vec{OP}_2$ hanno lunghezza 1). Quindi, sempre per il teorema di Pitagora, si ha $|OQ|^2 = x_1^2 + x_2^2$. Combinando queste osservazioni con la (4.10), si ha $|OP|^2 = x_1^2 + x_2^2 + x_3^2$, ovvero

$$|\vec{OP}| = \sqrt{x_1^2 + x_2^2 + x_3^2}\,. \tag{4.11}$$

Questa formula, analoga alla (4.9), fornisce la lunghezza di un vettore geometrico $\vec{OP}$ dello spazio in funzione delle sue coordinate rispetto alla base ortonormale scelta.

Ora, poniamoci il problema di calcolare l'angolo tra due vettori non nulli $\vec{OP}$ e $\vec{OQ}$ una volta note le loro coordinate rispetto ad una stessa base ortonormale. Supponiamo che tali coordinate siano rispettivamente (x_1, x_2, x_3) e (y_1, y_2, y_3).
Consideriamo il triangolo OPQ del disegno seguente.

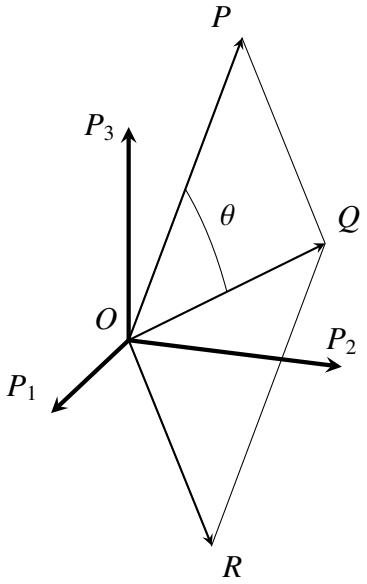


Per un risultato di trigonometria, l'angolo θ tra $\vec{OP}$ e $\vec{OQ}$ è collegato alle lunghezze dei segmenti OP, OQ e PQ dalla formula[3]

$$|PQ|^2 = |OP|^2 + |OQ|^2 - 2\cos\theta|OP| \cdot |OQ|\,. \tag{4.12}$$

Ora, per la (4.11), si ha $|OP| = \sqrt{x_1^2 + x_2^2 + x_3^2}$ e $|OQ| = \sqrt{y_1^2 + y_2^2 + y_3^2}$. Ci resta da calcolare la lunghezza $|PQ|$. Dal momento che la (4.11) ci consente di calcolare lunghezze solo dei vettori applicati in O, tracciamo come nel disegno seguente

[3]Si tratta di una sorta di "teorema di Pitagora per triangoli qualunque". Infatti, se il triangolo è rettangolo in O, ovvero $\theta = \frac{\pi}{2}$, allora $\cos\theta = 0$ e la formula si riduce a $|PQ|^2 = |OP|^2 + |OQ|^2$, il classico teorema di Pitagora.



il vettore $\vec{OR}$ parallelo al segmento PQ e avente la sua stessa lunghezza, ovvero $|PQ| = |\vec{OR}|$.
Ora, essendo $\vec{OR}$ parallelo a PQ e della stessa lunghezza, il quadrilatero di vertici O, R, Q, P è un parallelogramma che ha $\vec{OR}$ e $\vec{OP}$ come lati e $\vec{OQ}$ come diagonale. Quindi, dalla definizione di somma tra vettori applicati, si ha $\vec{OQ} = \vec{OR} + \vec{OP}$, ovvero $\vec{OR} = \vec{OQ} - \vec{OP}$.
Per l'Osservazione 4.7, le coordinate di $\vec{OR} = \vec{OQ} - \vec{OP}$ sono date dalle coordinate di $\vec{OQ}$ meno le coordinate di $\vec{OP}$, ovvero $(y_1 - x_1, y_2 - x_2, y_3 - x_3)$. Infine dalla (4.11) otteniamo

$$|PQ| = |\vec{OR}| = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2}\,. \tag{4.13}$$

La formula (4.12) diventa allora

$$\begin{aligned} |PQ|^2 &= (y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2 = \\ &= x_1^2 + x_2^2 + x_3^2 + y_1^2 + y_2^2 + y_3^2 - 2\cos\theta \sqrt{x_1^2 + x_2^2 + x_3^2}\sqrt{y_1^2 + y_2^2 + y_3^2}\,. \end{aligned} \tag{4.14}$$

Poichè il primo membro, per la formula del quadrato di binomio, è uguale a

$$x_1^2 + y_1^2 - 2x_1y_1 + x_2^2 + y_2^2 - 2x_2y_2 + x_3^2 + y_3^2 - 2x_3y_3,$$

semplificando con i quadrati al secondo membro si ottiene

$$-2x_1y_1 - 2x_2y_2 - 2x_3y_3 = -2\cos\theta \sqrt{x_1^2 + x_2^2 + x_3^2}\sqrt{y_1^2 + y_2^2 + y_3^2}\,. \tag{4.15}$$

Da quest'ultima si ricava

$$\cos\theta = \frac{x_1y_1 + x_2y_2 + x_3y_3}{\sqrt{x_1^2 + x_2^2 + x_3^2}\sqrt{y_1^2 + y_2^2 + y_3^2}} \tag{4.16}$$

che è la formula che esprime l'angolo tra due vettori in funzione delle loro coordinate rispetto alla base data.

Lo stesso calcolo si traduce mutatis mutandis nel caso bidimensionale per ottenere la formula analoga

$$\cos\theta = \frac{x_1y_1 + x_2y_2}{\sqrt{x_1^2 + x_2^2}\sqrt{y_1^2 + y_2^2}}\,. \tag{4.17}$$

Esempio 4.10. Consideriamo i vettori $\vec{OP}$ e $\vec{OQ}$ aventi coordinate $(1,0,1)$ e, rispettivamente, $(1,-1,0)$ rispetto a una base ortonormale $\{\vec{OP}_1, \vec{OP}_2, \vec{OP}_3\}$. In base alla definizione di coordinate, sono quindi $\vec{OP} = 1\vec{OP}_1 + 0\vec{OP}_2 + 1\vec{OP}_3 = \vec{OP}_1 + \vec{OP}_3$ e $\vec{OQ} = 1\vec{OP}_1 + (-1)\vec{OP}_2 + 0\vec{OP}_3 = \vec{OP}_1 - \vec{OP}_2$.
Allora l'angolo tra $\vec{OP}$ e $\vec{OQ}$, in base alla (4.16), è dato da

$$\cos\theta = \frac{1\cdot 1 + 0\cdot(-1) + 1\cdot 0}{\sqrt{1^2+0^2+1^2}\sqrt{1^2+(-1)^2+0^2}} = \frac{1}{\sqrt{2}\sqrt{2}} = \frac{1}{2}$$

ovvero, dalla trigonometria, $\theta = \frac{\pi}{3}$ (in gradi, 60°)
Le formule (4.16) e (4.17) ci forniscono anche un criterio per verificare in coordinate se due vettori sono perpendicolari: l'angolo θ è $\frac{\pi}{2}$ (ovvero 90 gradi) se e solo se $\cos\theta = 0$. Questo si verifica solo se i numeratori della (4.16) e della (4.17) sono nulli.
Nello spazio, abbiamo che due vettori $\vec{OP} \equiv (x_1,x_2,x_3)$ e $\vec{OQ} \equiv (y_1,y_2,y_3)$ sono perpendicolari se e solo se si verifica

$$x_1y_1 + x_2y_2 + x_3y_3 = 0\,. \tag{4.18}$$

Ad esempio, i due vettori di coordinate $(1,2,1)$ e $(3,1,-5)$ sono perpendicolari in quanto

$$1\cdot 3 + 2\cdot 1 + 1\cdot(-5) = 3 + 2 - 5 = 0\,.$$

Osservazione 4.11. In base al criterio (4.18), il vettore nullo $\vec{OO}$ risulta essere perpendicolare a qualunque altro vettore $\vec{OP}$. Questo perché le sue coordinate sono $(0,0,0)$ e, qualunque siano le coordinate (x_1,x_2,x_3) di $\vec{OP}$ si ottiene $x_1\cdot 0 + x_2\cdot 0 + x_3\cdot 0 = 0$.
Tuttavia, si noti che le formule (4.16) e (4.17) sono applicabili per calcolare un angolo solo se nessuno dei due vettori è nullo in quanto le lunghezze dei vettori appaiono al denominatore.

Il numeratore che compare nella (4.16), o nella (4.17) nel caso del piano, può essere interpretato come una nuova operazione che date due terne (due coppie nel caso del piano) di numeri reali, ci dà come risultato un numero reale. Se denotiamo con $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$ le due terne (nel caso del piano, $x = (x_1, x_2)$) e $y = (y_1, y_2)$), possiamo definire

$$x \cdot y \coloneqq x_1y_1 + x_2y_2 + x_3y_3 \tag{4.19}$$

(nel caso del piano, $x \cdot y \coloneqq x_1y_1 + x_2y_2$).
Per quello che riguarda i denominatori della (4.16) e della (4.17), dal momento che $\sqrt{x_1^2 + x_2^2 + x_3^2}$ rappresenta la lunghezza di un vettore $\vec{OP}$, che abbiamo denotato con $|\vec{OP}|$, possiamo definire

$$|x| \coloneqq \sqrt{x_1^2 + x_2^2 + x_3^2} \tag{4.20}$$

(nel caso del piano, $|x| = \sqrt{x_1^2 + x_2^2}$).
Con queste notazioni, sia la formula (4.16) che la (4.17) si riscrivono come

$$\cos\theta = \frac{x \cdot y}{|x||y|}. \tag{4.21}$$

Definizione 4.12. L'operazione $x \cdot y$ tra coppie o terne di vettori (cioè su su $\mathbb{R}^2$ e su $\mathbb{R}^3$) definita da (4.19) si chiama **prodotto scalare**. Il nome è dovuto al fatto che il risultato è un numero reale, ovvero uno scalare.

Si noti che non solo il prodotto scalare compare a numeratore dell'espressione dell'angolo tra due vettori in coordinate, ma anche la lunghezza $|x|$ di un vettore può essere espressa in funzione di esso, in quanto

$$|x| = \sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{x_1 \cdot x_1 + x_2 \cdot x_2 + x_3 \cdot x_3} = \sqrt{x \cdot x}$$

Osservazione 4.13. Il prodotto scalare $x \cdot y$ che abbiamo definito sopra è solo un caso particolare di una nozione più generale.
Più precisamente, noi abbiamo definito $x \cdot y$ solo per $x, y \in \mathbb{R}^2$ o $x, y \in \mathbb{R}^3$, perché stiamo lavorando con vettori nel piano o nello spazio che, fissata una base, vengono descritti rispettivamente da coppie o terne di numeri reali, e vogliamo tramite questo prodotto calcolare lunghezze e angoli. Come abbiamo visto nell'Osservazione 4.8, può capitare di avere a che fare con oggetti e situazioni che richiedono più di tre parametri indipendenti per essere descritti. In questi casi si deve ricorrere a $\mathbb{R}^n$ con $n > 3$.
Anche in tali situazioni può emergere l'esigenza di introdurre una "misura" sulle n-uple $(x_1, x_2, \ldots, x_n)$ che rappresentano gli oggetti o gli stati del nostro sistema, magari allo

scopo di controllare, o massimizzare o minimizzare qualche quantità che tenga conto di tutti i parametri usati.
Ispirandoci al prodotto scalare definito sopra per misurare le lunghezze e gli angoli tra vettori, potremmo porre anche in tale situazione generale $x \cdot y = x_1y_1 + x_2y_2 + \cdots + x_ny_n$, da cui $|x| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$. Tuttavia a seconda del significato delle componenti della n-upla potremmo voler assegnare loro pesi diversi. Ad esempio considerando 4-uple, invece di prendere $x \cdot y = x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4$ potremmo porre $x \cdot y = x_1y_1 + x_2y_2 + x_3y_3 + 3x_4y_4$ per far pesare di più il contributo della quarta componente. Potremmo anche voler far interagire i parametri tra loro in modo più complesso e usare espressioni più complicate, per esempio $x \cdot y = x_1y_1 + x_2y_2 + 2x_3y_3 + x_3y_4 + x_4y_3 + x_4y_4$.
In generale, possiamo definire $x \cdot y$ in base alle nostre esigenze, ma con un avvertimento: $x \cdot y$ deve essere tale da poter poi definire la lunghezza tramite la $|x| = \sqrt{x \cdot x}$ e, qualunque significato vogliamo dargli nel nostro contesto, l'angolo tramite $\cos\theta = \frac{x \cdot y}{|x||y|}$. Quindi, ad esempio, deve essere $x \cdot x \geqslant 0$ (perché questa quantità possa stare sotto radice) e $-1 \leqslant \frac{x \cdot y}{|x||y|} \leqslant 1$ (affinché questa quantità possa essere un coseno, che è appunto sempre compreso tra -1 e 1).
Ebbene, si dimostra che queste e altre importanti proprietà sono garantite se assumiamo che il prodotto $x \cdot y$ che stiamo scegliendo goda delle seguenti proprietà:

- Il prodotto è commutativo: $x \cdot y = y \cdot x$.
- Il prodotto è distributivo rispetto alla somma: $x \cdot (y + y') = x \cdot y + x \cdot y'$.
- Il prodotto è compatibile con la moltiplicazione per scalari: $x \cdot (cy) = (cx) \cdot y = c(x \cdot y)$ per ogni $c \in \mathbb{R}$.
- Il prodotto soddisfa $x \cdot x \geqslant 0$ e $x \cdot x = 0$ se e solo se $x = (0, 0, \ldots, 0)$.

In generale un **prodotto scalare** è una qualunque operazione che abbia esattamente queste proprietà, ma non approfondiamo ulteriormente questo argomento.

Vediamo ora che in $\mathbb{R}^3$ è possibile introdurre anche un'altra operazione molto utile nella grafica ma anche in altre applicazioni (soprattutto in fisica). Questa è il **prodotto vettoriale**, che date due terne di numeri reali dà come risultato non uno scalare (come nel caso del prodotto scalare) ma una nuova terna (cioè un nuovo vettore, come suggerisce il nome). La definizione è la seguente: se $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$ allora si pone

$$x \wedge y \coloneqq (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)\,. \tag{4.22}$$

Ad esempio, se $x = (1, 2, 3)$ e $y = (2, 5, -1)$, si ha

$$x \wedge y \coloneqq (2 \cdot (-1) - 3 \cdot 5, 3 \cdot 2 - 1 \cdot (-1), 1 \cdot 5 - 2 \cdot 2) = (-17, 7, 1)$$

Il motivo di questa particolare definizione è che si vuole che la terna $x \wedge y$ rappresenti (in coordinate) un vettore che è perpendicolare sia al vettore rappresentato da x che a quello rappresentato da y.
Per verificarlo, basta usare il criterio di perpendicolarità visto nella (4.19), cioè moltiplicare le rispettive componenti di x e $x \wedge y$ (la prima con la prima, la seconda con la seconda, la terza con la terza) e sommare:

$$\begin{aligned} &x_1(x_2y_3 - x_3y_2) + x_2(x_3y_1 - x_1y_3) + x_3(x_1y_2 - x_2y_1) = \\ &x_1x_2y_3 - x_1x_3y_2 + x_2x_3y_1 - x_2x_1y_3 + x_3x_1y_2 - x_3x_2y_1 = 0 \end{aligned}$$

in quanto come si vede facilmente tutti i termini si semplificano.
Analogamente, per verificare che anche il vettore di coordinate y è perpendicolare al vettore rappresentato dal prodotto vettoriale $x \wedge y$, svolgiamo il prodotto scalare tra y e $x \wedge y$:

$$\begin{aligned} &y_1(x_2y_3 - x_3y_2) + y_2(x_3y_1 - x_1y_3) + y_3(x_1y_2 - x_2y_1) = \\ &y_1x_2y_3 - y_1x_3y_2 + y_2x_3y_1 - y_2x_1y_3 + y_3x_1y_2 - y_3x_2y_1 = 0 \end{aligned}$$

come volevamo.
È facile verificare che il prodotto vettoriale *non* è commutativo bensì *anticommutativo*, cioè si ha

$$x \wedge y = -y \wedge x\,.$$

In altre parole, quando cambiamo l'ordine dei fattori il risultato finale cambia di segno. Questo segue immediatamente dalla seguente equazione.

$$\begin{aligned} x \wedge y &= (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1) = \\ &-(-x_2y_3 + x_3y_2, -x_3y_1 + x_1y_3, -x_1y_2 + x_2y_1) = -y \wedge x\,. \end{aligned}$$

Ad esempio, per i due vettori $x = (1,2,3)$ e $y = (2,5,-1)$ per cui sopra abbiamo già calcolato $x \wedge y$, si ha

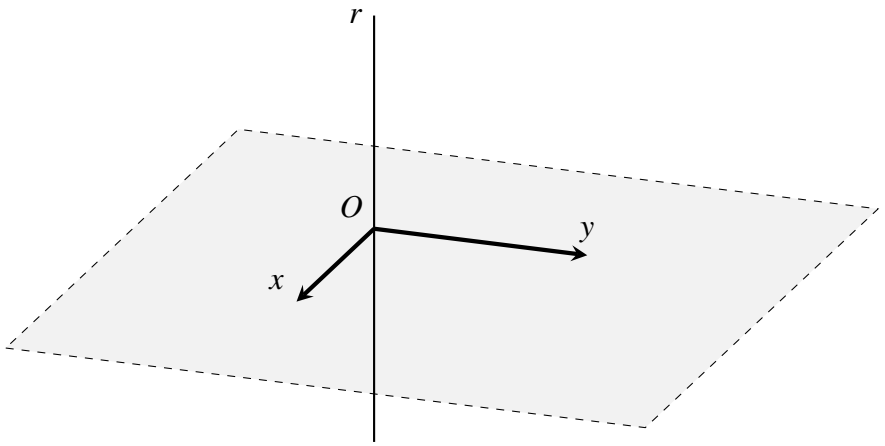
$$y \wedge x = (5 \cdot 3 + (-1) \cdot 2, -1 \cdot 1 - 2 \cdot 3, 2 \cdot 2 - 5 \cdot 1) = (17, -7, -1)$$

cioè proprio l'opposto della terna $(17,7,-1)$ ottenuta sopra.
Ancora, nei calcoli è necessario fare attenzione al fatto che il prodotto vettoriale *non* è neanche associativo, cioè in generale si ha

$$x \wedge (y \wedge z) \neq (x \wedge y) \wedge z\,.$$

Ad esempio, se prendiamo $x = (1,0,0)$ e $y = z = (0,1,0)$, si vede facilmente che $x \wedge y = (0,0,1)$ e $(x \wedge y) \wedge z = (-1,0,0)$, mentre dall'altra si ha $y \wedge z = (0,0,0)$ e $x \wedge (y \wedge z) = (0,0,0)$.

Abbiamo detto che il prodotto vettoriale $x \wedge y$ di due terne $x, y \in \mathbb{R}^3$ ci dà le coordinate di un vettore perpendicolare a entrambi i vettori rappresentati da x e da y, e che quindi si trova sulla retta rappresentata nel disegno seguente.



Conoscendo quindi la direzione di tale vettore, per determinarlo completamente dobbiamo trovarne lunghezza e verso.
Per quanto riguarda la lunghezza, sappiamo come calcolarla mediante la formula (4.20). In base a tale formula e alla (4.22), si ha

$$|x \wedge y|^2 = (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 + (x_1y_2 - x_2y_1)^2 .$$

Svolgendo i conti (omettiamo i passaggi), non è difficile vedere che tale espressione è uguale a

$$(x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1y_1 + x_2y_2 + x_3y_3)^2$$

ovvero, ricordando le notazioni introdotte nella (4.19) e nella (4.20),

$$|x|^2|y|^2 - (x \cdot y)^2 .$$

Riscrivendo questa espressione come

$$|x|^2|y|^2 \left(1 - \frac{(x \cdot y)^2}{|x|^2|y|^2}\right)$$

e ricordando la (4.21), concludiamo che

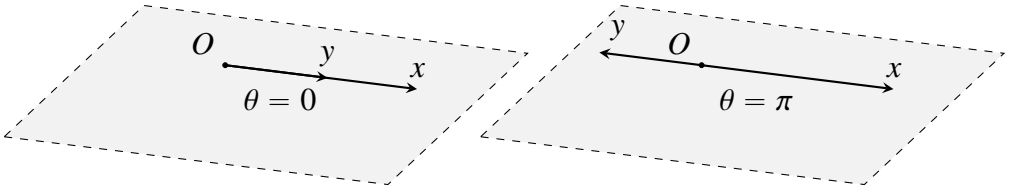
$$|x \wedge y|^2 = |x|^2|y|^2(1 - \cos^2\theta) = |x|^2|y|^2 \sin^2\theta$$

(dove abbiamo usato l'identità trigonometrica $\cos^2\theta + \sin^2\theta = 1$). A questo punto, estraendo la radice a entrambi i membri, si ottiene

$$|x \wedge y| = |x||y| \sin\theta \tag{4.23}$$

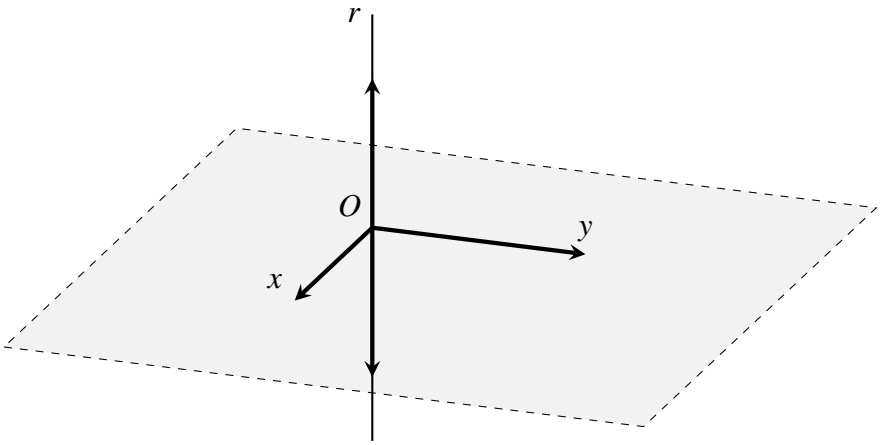
che è una formula semplice per la lunghezza del vettore rappresentato in coordinate da $x \wedge y$, in funzione della lunghezza $|x|$ del vettore rappresentato da x, della lunghezza $|y|$ del vettore rappresentato da y e dell'angolo θ formato da questi due vettori.

Tale formula ci dice ad esempio che $|x \wedge y| = 0$ (ovvero $x \wedge y = (0,0,0)$ rappresenta il vettore nullo $\vec{OO}$) esattamente quando $\sin\theta = 0$. Questo, come ci dice la trigonometria, avviene quando $\theta = 0$ oppure $\theta = \pi$ (180 gradi). Come si vede nel disegno seguente



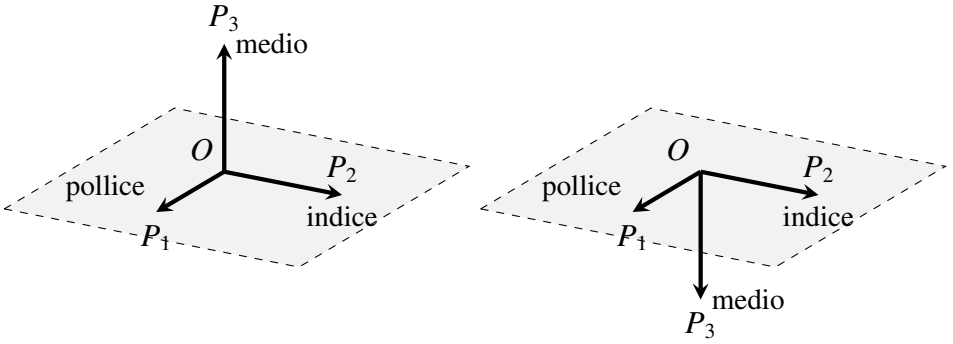
questo equivale a dire che i vettori sono allineati.

Conoscendo direzione e lunghezza del vettore rappresentato da $x \wedge y$, per il verso abbiamo solo due possibilità:



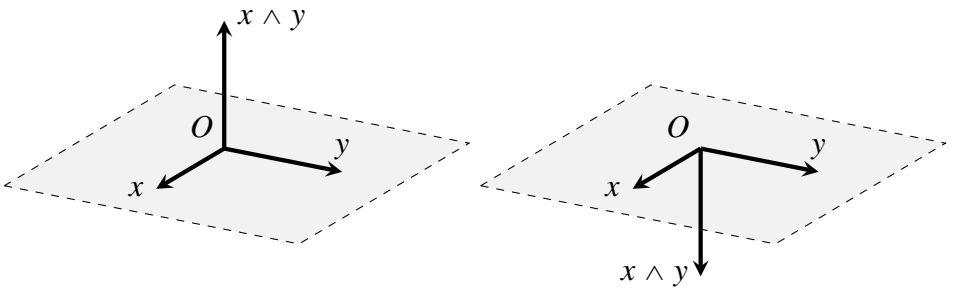
Come vedremo, il verso del vettore rappresentato da $x \wedge y$ non è determinabile in modo univoco, ma dipende da quale base ortonormale abbiamo scelto per tradurre i vettori in coordinate.

Più precisamente, esistono due tipi di basi ortonormali. Le prime sono quelle **destrorse**, cosiddette perché i tre vettori $\vec{OP_1}, \vec{OP_2}, \vec{OP_3}$ che le costituiscono sono orientati rispettivamente come pollice, indice e medio della mano destra (disposti a formare angoli ortogonali tra loro). Le seconde sono quelle **sinistrorse**, cosiddette perché i tre vettori $\vec{OP_1}, \vec{OP_2}, \vec{OP_3}$ che le costituiscono sono orientati rispettivamente come pollice, indice e medio della mano sinistra (sempre disposti a formare angoli ortogonali tra loro).



Base destrorsa Base sinistrorsa

È facile convincersi che si tratta effettivamente di tipi diversi: non è possibile sovrapporre perfettamente tutte e tre le dita pollice, indice e medio delle mani destra e sinistra, disposte a formare angoli ortogonali tra loro, comunque ruotiamo le mani. Riusciamo sempre a sovrapporre due dita, ma il terzo risulta di verso opposto tra una mano e l'altra. Si può dimostrare che, se stiamo usando una base destrorsa, allora il vettore rappresentato da $x \wedge y$ ha verso tale che i tre vettori rappresentati da x, y e $x \wedge y$ sono, nell'ordine, ancora orientati "in modo destrorso" (ovvero come pollice, indice e medio della mano destra). Se invece usiamo una base sinistrorsa, allora i tre vettori rappresentati da x, y e $x \wedge y$ sono, nell'ordine, ancora orientati "in modo sinistrorso" (ovvero come pollice, indice e medio della mano sinistra).

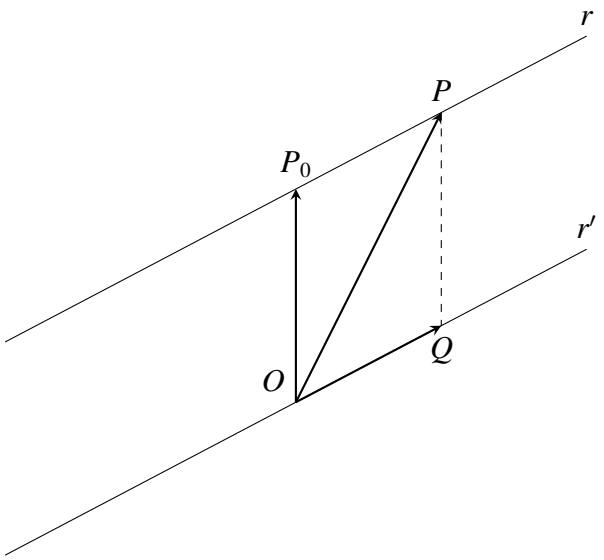


Rispetto a una base destrorsa Rispetto a una base sinistrorsa

Nella maggior parte dei testi si sceglie di usare basi ortonormali destrorse. Si dice allora, soprattutto nei testi di fisica, che il prodotto vettoriale è definito usando la *regola della mano destra*.

4.4 Sistemi di riferimento nello spazio e equazioni di rette e piani

Quanto visto nella sezione precedente ci consente di tradurre problemi geometrici in problemi numerici e risolverli mediante equazioni algebriche.
Per assegnare coordinate a un punto P (del piano o dello spazio) bisogna fissare un punto O, e una base B per i vettori applicati in O (sarà $B = \{\vec{OP}_1, \vec{OP}_2, \vec{OP}_3\}$ se siamo nello spazio tridimensionale e $B = \{\vec{OP}_1, \vec{OP}_2\}$ se siamo nel piano). Queste due scelte definiscono un **sistema di riferimento**, e il punto O si dice **origine del sistema**.
A questo punto, diciamo che le **coordinate del punto** P rispetto al sistema di riferimento fissato sono semplicemente le coordinate del vettore $\vec{OP}$ rispetto alla base scelta.
Scriveremo $P \equiv (x, y, z)$ (nel piano, $P \equiv (x, y)$) per dire che x, y, z (rispettivamente, x, y) sono le coordinate di P rispetto al sistema di riferimento fissato. In altre parole, in base alla definizione appena data, $\vec{OP} = x\vec{OP}_1 + y\vec{OP}_2 + z\vec{OP}_3$ (rispettivamente, $\vec{OP} = x\vec{OP}_1 + y\vec{OP}_2$).
Chiaramente, le coordinate di P dipendono sia dalla scelta iniziale dell'origine O (cambiando O, cambia il vettore $\vec{OP}$ di cui dobbiamo calcolare le coordinate) che dalla scelta della base. Solitamente, il sistema di riferimento è scelto in modo da semplificare il più possibile il problema che si vuole risolvere.
Ora che, fissato un sistema di riferimento, ogni punto dello spazio (o del piano) ha una sua terna (rispettivamente, coppia) di coordinate, ci chiediamo se sia possibile ricavare una formula che ci dia le coordinate dei punti di una retta r data. A questo scopo, fissiamo un punto P_0 della retta r. Osserviamo che, dato un qualunque altro punto P su r, il vettore $\vec{OP}$ si decompone come somma di $\vec{OP}_0$ e un vettore $\vec{OQ}$ che giace sulla retta r' parallela a r e passante per O.

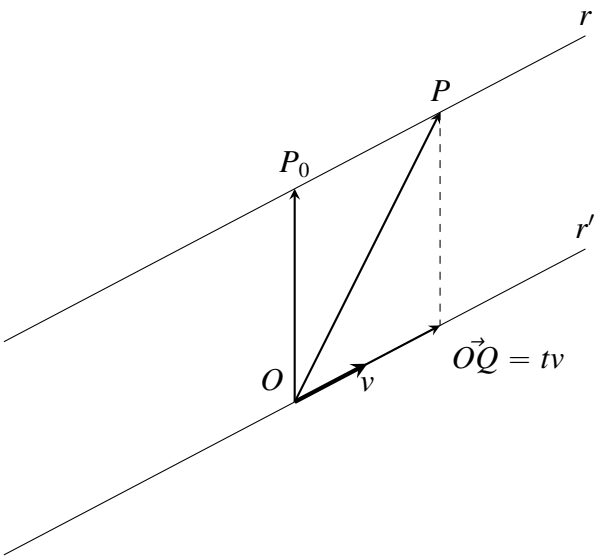


Fissato P_0, quindi, esiste una corrispondenza tra i punti P della retta e i vettori $\vec{OQ}$ che giacciono su r', data proprio da

$$\vec{OP} = \vec{OP_0} + \vec{OQ}\,. \tag{4.24}$$

Il vantaggio di passare dai punti della retta r ai vettori $\vec{OQ}$ che giacciono su r' consiste nel fatto che tali vettori $\vec{OQ}$ sono tutti multipli tra loro. Basta quindi fissarne uno (non nullo), diciamo v, e si avrà sempre $\vec{OQ} = tv$ per qualche $t \in \mathbb{R}$.
Quindi, la (4.24) può essere riscritta come

$$\vec{OP} = \vec{OP_0} + tv\,. \tag{4.25}$$



Vogliamo ora esprimere questa uguaglianza in coordinate rispetto a un sistema di riferimento utilizzando le proprietà delle coordinate dei vettori viste nell'Osservazione 4.7. Supponiamo di essere nello spazio e diciamo che le coordinate di $\vec{OP_0}$ e v sono date rispettivamente da (x_0, y_0, z_0) e (l, m, n). Quindi le coordinate di tv sono $t(l, m, n)$. Ora da $\vec{OP} = \vec{OP_0} + tv$ abbiamo che le coordinate (x, y, z) di $\vec{OP}$ sono uguali alle coordinate di $\vec{OP_0}$ più le coordinate di tv, ovvero

$$(x, y, z) = (x_0, y_0, z_0) + t(l, m, n) = (x_0 + lt, y_0 + mt, z_0 + nt)\,. \tag{4.26}$$

La (4.26) ci dice quindi che i punti della retta sono dati esattamente dalle terne (x, y, z) per le quali si ha $x = x_0 + lt$, $y = y_0 + mt$, $z = z_0 + nt$ per qualche $t \in \mathbb{R}$.
È convenzione comune riunire queste tre condizioni nella forma

$$\begin{cases} x = x_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt \end{cases} \tag{4.27}$$

Dal momento che le (4.27) ci danno esplicitamente le coordinate di tutti i punti della retta r al variare del parametro $t \in \mathbb{R}$, queste prendono il nome di **equazioni parametriche** della retta r.

Se invece stiamo lavorando nel piano, si procede nello stesso modo per arrivare ad un risultato simile ma con una coordinata in meno. Denotiamo con (x_0, y_0) le coordinate di $\vec{OP}_0$ e con (l, m) quelle di v. Da $\vec{OP} = \vec{OP}_0 + tv$ abbiamo che le coordinate (x, y) di $\vec{OP}$ sono date dalla somma delle coordinate di $\vec{OP}_0$ con quelle di tv, ovvero

$$(x, y) = (x_0, y_0) + t(l, m) = (x_0 + lt, y_0 + mt)\,. \tag{4.28}$$

La (4.28) ci dice quindi che i punti della retta sono dati esattamente dalle coppie (x, y) per le quali si ha $x = x_0 + lt$ e $y = y_0 + mt$ per qualche $t \in \mathbb{R}$.

Come nel caso dello spazio, è convenzione comune riunire queste condizioni nella forma

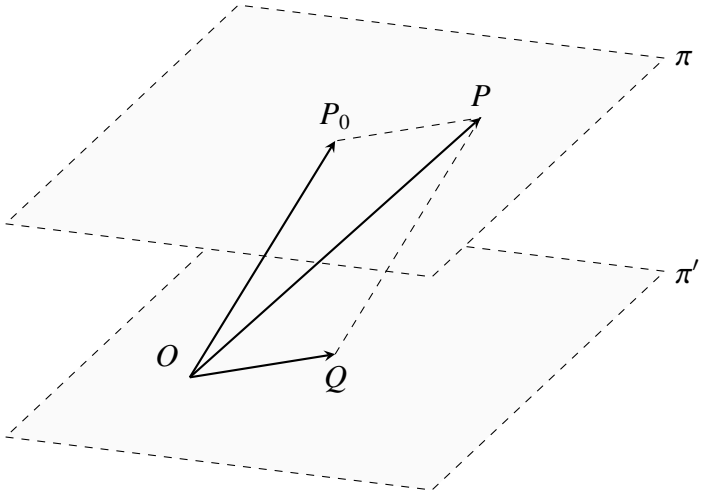
$$\begin{cases} x = x_0 + lt \\ y = y_0 + mt \end{cases} \tag{4.29}$$

e si parla sempre di **equazioni parametriche** di r.

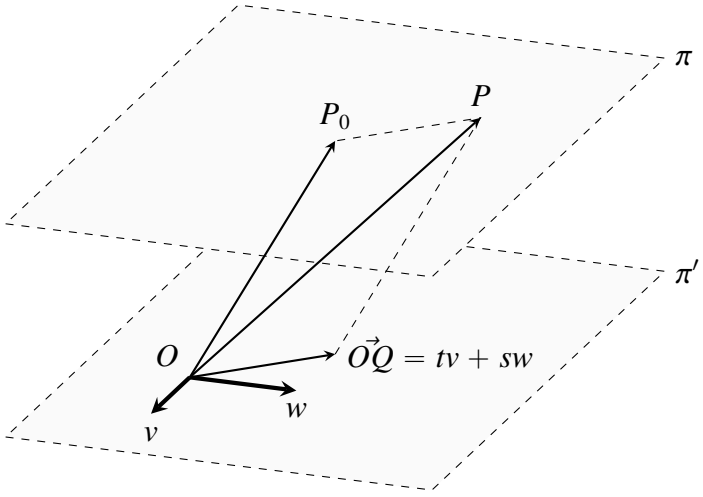
Osservazione 4.14. Si osservi che è ragionevole, per indivuidare la retta r, dover scegliere un punto P_0 e un vettore v sulla retta a lei parallela e passante per l'origine. Infatti, fissato un punto P_0 esistono infinite rette passanti per P_0, e il vettore v ha il ruolo di fissare la direzione della retta. Per questo motivo, il vettore v deve essere non nullo, ovvero non devono essere nulle le sue coordinate l, m, n. Questa è l'unica condizione che deve essere soddisfatta affinché una terna di uguaglianze come le (4.27) rappresenti effettivamente una retta r.

Vediamo ora come determinare i punti di un piano π nello spazio con un procedimento analogo.

Come nel disegno seguente, fissiamo un punto P_0 del piano π. Osserviamo che per ogni $P \in \pi$, il vettore $\vec{OP}$ si decompone come somma $\vec{OP} = \vec{OP}_0 + \vec{OQ}$ di $\vec{OP}_0$ e di un vettore $\vec{OQ}$ che giace sul piano π' parallelo a π e passante per O.



Esattamente come nel caso della retta, il vantaggio di ricondurre la descrizione dei punti P del piano π ai vettori $\vec{OQ}$ che giacciono su π' consiste nel fatto che tali vettori possono essere ottenuti come combinazione di due vettori non allineati v e w fissati. Ovvero $\vec{OQ} = tv + sw$, per opportuni $t, s \in \mathbb{R}$.



Riassumendo, per tutti i punti P sul piano π vale la relazione

$$\vec{OP} = \vec{OP}_0 + tv + sw\,. \tag{4.30}$$

Ora, come fatto sopra per la retta, traduciamo tale uguaglianza in coordinate rispetto al sistema di riferimento fissato. Usando ancora le proprietà delle coordinate viste nell'Osservazione 4.7, da $\vec{OP} = \vec{OP}_0 + tv + sw$ abbiamo che le coordinate (x, y, z) di $\vec{OP}$ sono uguali alle coordinate, diciamo (x_0, y_0, z_0), di $\vec{OP}_0$ più le coordinate di tv, (che sono uguali a t per le coordinate, diciamo (l, m, n), di v), più le coordinate di sw (che sono uguali a s per le coordinate, diciamo (l', m', n'), di w). In formule

$$\begin{aligned}(x,y,z) &= (x_0,y_0,z_0) + t(l,m,n) + s(l',m',n') \\ &= (x_0 + lt + l's, y_0 + mt + m's, z_0 + nt + n's)\,. \end{aligned} \tag{4.31}$$

La (4.31) ci dice quindi che i punti del piano sono dati esattamente dalle terne (x,y,z) per le quali si ha $x = x_0 + lt + l's$, $y = y_0 + mt + m's$, $z = z_0 + nt + n's$ per certi $t, s \in \mathbb{R}$.

È convenzione comune riunire queste tre condizioni nella forma

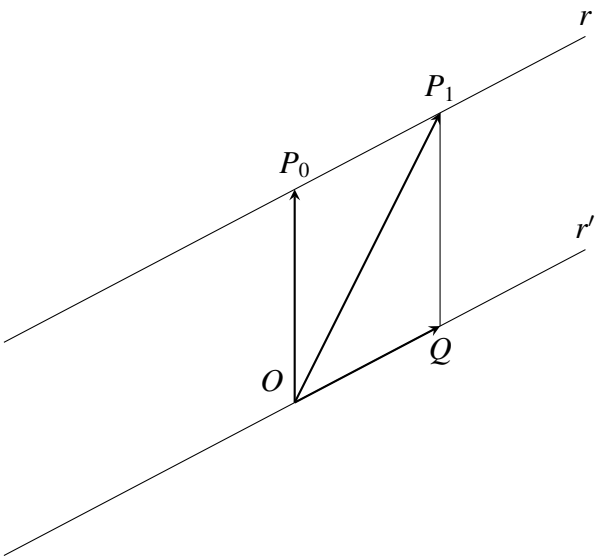
$$\begin{cases} x = x_0 + lt + l's \\ y = y_0 + mt + m's \\ z = z_0 + nt + n's \end{cases} \tag{4.32}$$

Come nel caso della retta, dal momento che le (4.32) ci danno esplicitamente le coordinate di tutti i punti del piano π al variare dei parametri $t, s \in \mathbb{R}$, si dice che tali uguaglianze sono le **equazioni parametriche** di π.

Osservazione 4.15. Una retta è determinata da un suo punto e da un vettore che ne determina la direzione. Allo stesso modo un piano π è determinato da un punto e da due vettori v e w che, individuando il piano π' parallelo al piano dato e passante per l'origine, determinano la cosiddetta **giacitura**. In altre parole, come il piano giace, cioè come è orientato o disposto, nello spazio: non avrebbe senso parlare di direzione del piano perché questo si sviluppa in due dimensioni e non in una sola. Poiché i vettori v e w devono generare il piano per O, essi non devono essere allineati. Questo equivale a dire che le terne (l,m,n) e (l',m',n') che rappresentano i vettori non devono essere proporzionali. Questa è l'unica condizione che deve essere soddisfatta affinché una terna di uguaglianze come le (4.32) rappresenti effettivamente un piano.

La geometria euclidea classica ci dice che dati due punti distinti P_0 e P_1 nel piano o nello spazio, esiste un'unica retta r che li contiene. Ci proponiamo ora di determinare questa retta a partire dalle coordinate dei due punti, che supponiamo note.

Supponiamo quindi che sia $P_0 \equiv (x_0,y_0,z_0)$ e $P_1 \equiv (x_1,y_1,z_1)$. Per determinare le equazioni parametriche della retta, come sappiamo, ci servono le coordinate (l,m,n) di un vettore non nullo che giaccia sulla retta r' parallela a r e passante per l'origine O del sistema di riferimento. A questo scopo, tracciamo la retta r' e, come nel disegno seguente, tracciamo partendo da P_1 il segmento P_1Q parallelo a OP_0.



Chiaramente, possiamo usare il vettore $\vec{OQ}$ come generatore della giacitura. Essendo il quadrilatero OP_0P_1Q un parallelogramma, per definizione di somma tra vettori geometrici si ha $\vec{OQ} + \vec{OP}_0 = \vec{OP}_1$, da cui otteniamo $\vec{OQ} = \vec{OP}_1 - \vec{OP}_0$.
Le coordinate (l, m, n) di $\vec{OQ}$ saranno quindi, per la Proposizione 4.5, uguali alle coordinate (x_1, y_1, z_1) di $\vec{OP}_1$ meno le coordinate (x_0, y_0, z_0) di $\vec{OP}_0$. Cioè

$$(l, m, n) = (x_1, y_1, z_1) - (x_0, y_0, z_0) = (x_1 - x_0, y_1 - y_0, z_1 - z_0).$$

Sostituendo nell'espressione generale (4.27) delle equazioni parametriche di una retta si trova allora

$$\begin{cases} x = x_0 + (x_1 - x_0)t \\ y = y_0 + (y_1 - y_0)t \\ z = z_0 + (z_1 - z_0)t \end{cases} \tag{4.33}$$

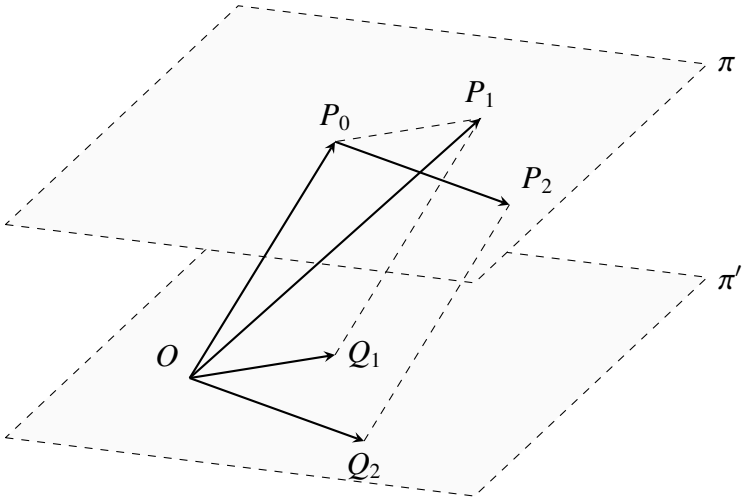
che sono quindi le equazioni della retta nello spazio che passa per P_0 e P_1.
Nel piano, si ha lo stesso risultato ma con una coordinata in meno. Ovvero le uguaglianze

$$\begin{cases} x = x_0 + (x_1 - x_0)t \\ y = y_0 + (y_1 - y_0)t \end{cases} \tag{4.34}$$

determinano la retta nel piano passante per i punti $P_0 \equiv (x_0, y_0)$ e $P_1 \equiv (x_1, y_1)$.

Si noti che se i punti non fossero distinti il vettore $(x_1 - x_0, y_1 - y_0, z_1 - z_0)$, che si ottiene sottraendo le loro coordinate, sarebbe nullo e non potrebbe servire da vettore direzione della retta.

Analogamente, la geometria euclidea classica ci dice che dati tre punti P_0, P_1, P_2 non allineati (ovvero che non appartengono alla stessa retta), allora esiste un unico piano π che li contiene. Troviamo ora le equazioni parametriche di questo piano a partire dalle coordinate dei tre punti.
Supponiamo che sia $P_0 \equiv (x_0, y_0, z_0)$, $P_1 \equiv (x_1, y_1, z_1)$ e $P_2 \equiv (x_2, y_2, z_2)$.
Per determinare le equazioni parametriche del piano, come sappiamo, ci servono le coordinate (l, m, n), (l', m', n') di due vettori non proporzionali che giacciano sul piano π' parallelo a π e passante per l'origine O del sistema di riferimento. A questo scopo, disegniamo il piano π' e, analogamente a quanto fatto nel caso della retta, consideriamo i parallelogrammi $OP_0P_1Q_1$ e $OP_0P_2Q_2$.



I due vettori $\vec{OQ_1}$ e $\vec{OQ_2}$, rispettivamente paralleli ai segmenti P_0P_1 e P_0P_2, giacciono sul piano π' e non sono proporzionali, poichè i tre punti P_0, P_1, P_2 sono non allineati. Quindi possiamo usarli come generatori della giacitura.
Per determinare le coordinate di tali vettori, osserviamo che per definizione di somma tra vettori geometrici si ha $\vec{OQ_1} + \vec{OP_0} = \vec{OP_1}$ e $\vec{OQ_2} + \vec{OP_0} = \vec{OP_2}$, da cui $\vec{OQ_1} = \vec{OP_1} - \vec{OP_0}$ e $\vec{OQ_2} = \vec{OP_2} - \vec{OP_0}$.
Allora, le coordinate (l, m, n) e (l', m', n') di $\vec{OQ_1}$ e $\vec{OQ_2}$ rispettivamente saranno

$$(l, m, n) = (x_1, y_1, z_1) - (x_0, y_0, z_0) = (x_1 - x_0, y_1 - y_0, z_1 - z_0),$$
$$(l', m', n') = (x_2, y_2, z_2) - (x_0, y_0, z_0) = (x_2 - x_0, y_2 - y_0, z_2 - z_0).$$

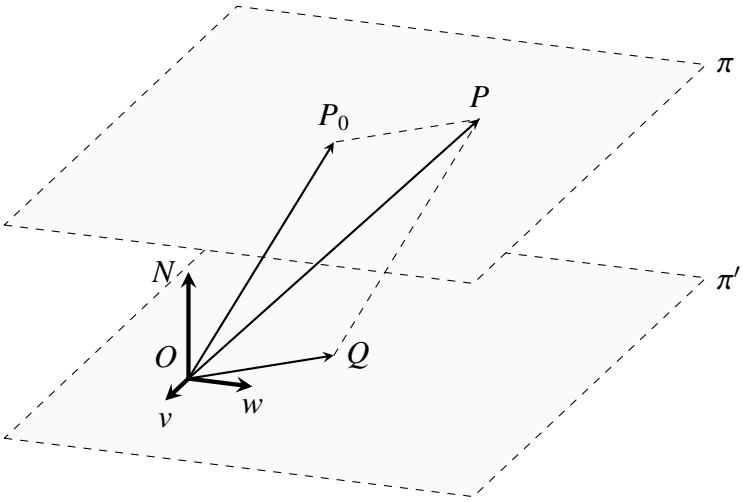
Sostituendo nell'espressione generale (4.32) delle equazioni parametriche di un piano si trova allora

$$\begin{cases} x = x_0 + (x_1 - x_0)t + (x_2 - x_0)s \\ y = y_0 + (y_1 - y_0)t + (y_2 - y_0)s \\ z = z_0 + (z_1 - z_0)t + (z_2 - z_0)s \end{cases} \tag{4.35}$$

Queste sono quindi le equazioni del piano che passa per P_0, P_1 e P_2.

Vediamo ora un modo alternativo per descrivere rette e piani, quello dato dalle cosiddette **equazioni cartesiane**, note anche come **equazioni implicite**.

Iniziamo dai piani. Come visto nel disegno a pagina 154, fissato un punto P_0 del piano, per ogni altro suo punto P il vettore $\vec{OP}$ si decompone come somma $\vec{OP} = \vec{OP}_0 + \vec{OQ}$, dove $\vec{OQ}$ è un vettore che sta sul piano parallelo al piano dato e passante per O.
Avevamo poi osservato che ogni tale $\vec{OQ}$ si scrive come somma $tv + sw$ di due vettori v e w non allineati, ottenendo così le parametriche. L'osservazione alternativa che facciamo ora invece, è che tutti i vettori $\vec{OQ}$ sono perpendicolari a un qualunque vettore N perpendicolare al piano.



Quindi i punti P del piano sono caratterizzati dalla condizione che $\vec{OP} - \vec{OP}_0 = \vec{OQ}$ è perpendicolare a N. Ora, esprimiamo tale condizione in coordinate. Se denotiamo con (x, y, z) le coordinate di P (ovvero le coordinate del vettore $\vec{OP}$) e con (x_0, y_0, z_0) le coordinate di P_0 (ovvero le coordinate del vettore $\vec{OP}_0$), le coordinate del vettore $\vec{OP} - \vec{OP}_0$ sono date, come sappiamo dalle proprietà delle coordinate, dalla terna $(x - x_0, y - y_0, z - z_0)$. Diciamo (a, b, c) le coordinate del vettore N e supponiamo che *la base fissata per il nostro sistema di riferimento sia ortonormale*. Allora, per la (4.18), la condizione che $\vec{OP} - \vec{OP}_0$ sia perpendicolare a N equivale a

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0\,.$$

Questa quindi è la condizione che devono soddisfare le coordinate (x, y, z) di un punto P perché il punto appartenga al piano.
Svolgendo i conti otteniamo

$$ax - ax_0 + by - by_0 + cz - cz_0 = 0$$

ovvero, portando a secondo membro tutti i termini che non contengono x, y, z,

$$ax + by + cz = ax_0 + by_0 + cz_0 \,.$$

Denotando $ax_0 + by_0 + cz_0 = d$, otteniamo infine

$$ax + by + cz = d \,. \tag{4.36}$$

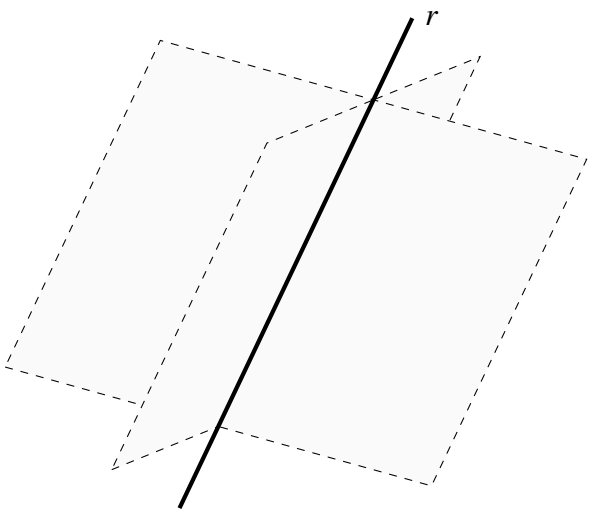
In conclusione, i punti di un piano dato sono caratterizzati dal fatto che le loro coordinate (x, y, z) soddisfano un'equazione di primo grado in tre incognite. Se la base del nostro sistema di riferimento è ortonormale, i coefficienti di x, y, z nell'equazione, ovvero a, b, c, rappresentano le coordinate di un vettore perpendicolare al piano (detto anche una **normale al piano**). Inoltre il termine noto d dell'equazione contiene l'informazione relativa a un punto per cui passa il piano, come abbiamo visto ricavando l'equazione.

Poichè piani con la stessa normale risultano paralleli, deduciamo che fissando a, b, c e facendo variare solo d "muoviamo" il piano parallelamente a se stesso.

Esempio 4.16. L'equazione $x + 2y + 3z = 6$ rappresenta un piano nello spazio, la cui giacitura è normale al vettore di coordinate $(1, 2, 3)$. Un punto che appartiene al piano è, ad esempio, $P \equiv (1, 1, 1)$, in quanto sostituendo tali coordinate al posto di x, y, z nell'equazione si ottiene $1 + 2 + 3 = 6$ che è verificata. Al contrario il punto $P \equiv (1, 2, 0)$ non appartiene al piano in quanto $1 + 2 \cdot 2 + 3 \cdot 0 = 6$ non è verificata.

L'equazione cartesiana del piano rappresenta un modo *implicito* di dare i punti del piano, diversamente dalle equazioni parametriche che ci danno i punti *esplicitamente*. Infatti, mentre nel caso delle parametriche i punti del piano sono dati semplicemente scegliendo valori reali dei parametri t e s, nel caso nelle cartesiane per ottenere i punti del piano bisogna trovare le soluzioni dell'equazione. Tale equazione rappresenta un vincolo che imponiamo ai punti dello spazio per appartenere al piano.

Anche le rette possono essere rappresentate in maniera implicita come segue.



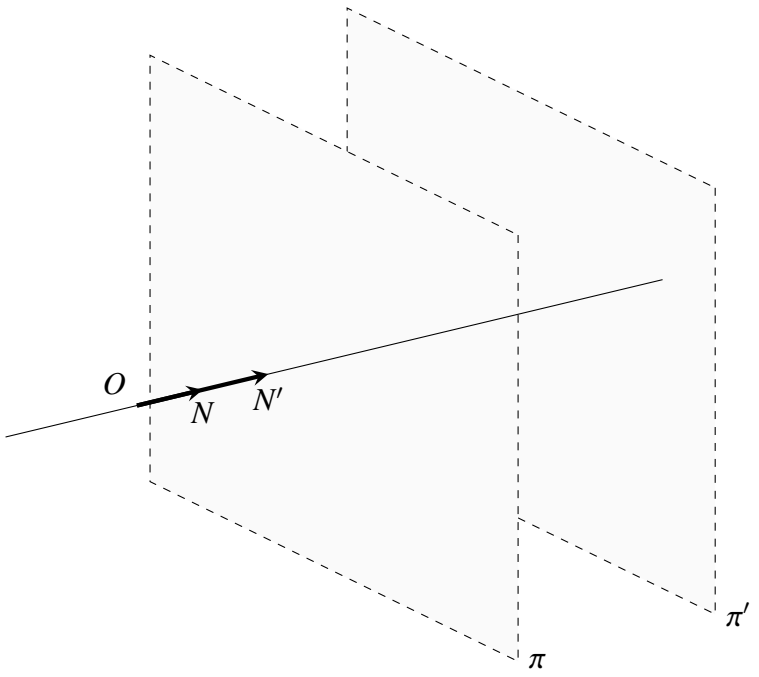
Come si vede nel disegno qualunque retta può essere pensata come intersezione di due piani (esistono infiniti piani che contengono una retta data, ma basta prenderne due per ottenere la retta come intersezione).

Supponiamo che il primo piano sia descritto dall'equazione cartesiana $ax + by + cz = d$ (ovvero i punti del primo piano sono, in coordinate, le soluzioni di questa equazione) e il secondo piano dall'equazione $a'x + b'y + c'z = d'$ (ovvero i punti del secondo piano sono, in coordinate, le soluzioni di quest'altra equazione). Allora i punti della retta, essendo i punti che appartengono ad entrambi i piani, sono dati dalle soluzioni comuni a queste due equazioni, ovvero dalle soluzioni del sistema

$$\begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \end{cases}$$

Abbiamo quindi dimostrato che, oltre che esplicitamente attraverso equazioni parametriche, una retta nello spazio può essere descritta implicitamente mediante un sistema di due equazioni di primo grado in tre incognite.

Osservazione 4.17. Per rappresentare una retta come intersezione di due piani, è ovviamente necessario che i due piani non siano paralleli (l'intersezione sarebbe l'insieme vuoto) o addirittura coincidenti. Questo accade solo se le normali dei due piani sono uguali o al più una è mutiplo dell'altra (cambierebbe lunghezza o verso ma la direzione rimarrebbe la stessa).



Esempio 4.18. l seguente sistema di equazioni

$$\begin{cases} x - y + 2z = 3 \\ 2x - 2y + 4z = 2 \end{cases}$$

non rappresenta una retta pur essendo un sistema di due equazioni di primo grado in tre incognite. Si noti infatti che la normale al piano rappresentato dalla prima equazione è $N \equiv (1,-1,2)$ mentre la normale al piano rappresentato dalla seconda equazione è $N' \equiv (2,-2,4)$. Essendo le coordinate di N' date dalle coordinate di N moltiplicate per 2, deduciamo che $N' = 2N$ e quindi i due piani, avendo normali con la stessa direzione, sono paralleli e non si incontrano lungo una retta.
Invece il sistema

$$\begin{cases} x + y + z = 3 \\ 2x - y + z = 2 \end{cases}$$

rappresenta una retta nello spazio. Iil punto $(1,1,1)$ appartiene alla retta in quanto le sue coordinate soddisfano entrambe le equazioni. D'altro canto il punto $(2,0,1)$ non appartiene alla retta in quanto benché le sue coordinate soddisfino la prima equazione non soddisfano la seconda.

Quanto visto sino ad ora suggerisce che per risolvere problemi geometrici riguardanti rette e piani è necessario saper risolvere i sistemi di equazioni di primo grado. Infatti, nelle espressioni delle parametriche i parametri t ed s compaiono con grado 1, e le

cartesiane sono sistemi di equazioni di primo grado. Inoltre se volessimo capire, ad esempio, se un piano e una retta dati rispettivamente da equazioni cartesiane $ax + by + cz = d$ e

$$\begin{cases} a'x + b'y + c'z = d' \\ a''x + b''y + c''z = d'' \end{cases}$$

si intersecano, e trovare gli eventuali punti di intersezione, dovremmo risolvere il sistema

$$\begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \\ a''x + b''y + c''z = d'' \end{cases}$$

che si ottiene mettendo insieme tutte le equazioni (in quanto i punti comuni al piano e alla retta sono i punti che soddisfano sia l'equazione del piano che le due equazioni della retta).

Anche in parametriche, il problema di individuare i punti comuni si traduce in un sistema di equazioni di primo grado. Ad esempio, per trovare i punti in comune a due piani, definiti dalle parametriche

$$\begin{cases} x = 1 + t + s \\ y = 2 + t - s \\ z = t + 2s \end{cases} \quad \text{e} \quad \begin{cases} x = -2 + t + 3s \\ y = 1 - t - s \\ z = 1 + t + s \end{cases}$$

si deve capire se esistono valori di t ed s nelle prime e valori t' ed s' nelle seconde (eventualmente diversi dai t e s scelti nelle prime) che danno gli stessi valori di x, y, z. In altre parole bisognerebbe impostare le uguaglianze

$$\begin{cases} 1 + t + s = -2 + t' + 3s' \\ 2 + t - s = 1 - t' - s' \\ t + 2s = 1 + t' + s' \end{cases}$$

ovvero

$$\begin{cases} t + s - t' - 3s' = -3 \\ t - s + t' + s' = -1 \\ t + 2s - t' - s' = 1 \end{cases}$$

che è un sistema di tre equazioni di primo grado in quattro incognite.

Emerge quindi l'esigenza di saper risolvere sistemi di primo grado con un qualsiasi numero di equazioni in qualunque numero di incognite. Sarà questo lo scopo del prossimo capitolo.

4.5 Appendice: il campo dei numeri complessi

Successivamente alla Definizione 3.2 abbiamo citato vari esempi di campo. Oltre all'insieme dei razionali $\mathbb{Q}$ e dei reali $\mathbb{R}$ abbiamo nominato il campo dei numeri complessi $\mathbb{C}$. In questa sezione presentiamo in dettaglio questo campo.

Il campo dei **numeri complessi**, denotato con $\mathbb{C}$, può essere definito come l'insieme delle espressioni del tipo

$$x + iy$$

dove x e y sono numeri reali e i (detta **unità immaginaria**) è un numero che soddisfa l'identità $i^2 = -1$. Si noti che questa identità non è soddisfatta da nessun numero reale: il quadrato di un qualunque numero reale è un numero non negativo. Ad esempio, sono numeri complessi

$$2 + i3\,, \quad \sqrt{2} - i\,, \quad \frac{1}{5} - i\pi\,, \ \ldots$$

In ogni numero complesso $x + iy$, il numero reale x si chiama **parte reale** e il numero reale y **parte immaginaria**.

Ovviamente, x e y possono essere nulli. In particolare, quando y è nullo si ha $x+i0 = x \in \mathbb{R}$. Quindi i numeri reali sono particolari numeri complessi, quelli con parte immaginaria nulla (in altre parole, vale l'inclusione $\mathbb{R} \subseteq \mathbb{C}$).

Per poter dotare $\mathbb{C}$ della struttura di campo, dobbiamo la somma e il prodotto tra numeri complessi e poi verificare che per tali operazioni valgono le proprietà della Definizione 3.2.

Definiamo la somma di due numeri complessi come il numero complesso ottenuto sommando le loro parti reali e le loro parti immaginarie. In altre parole dati due numeri complessi $x + iy$ e $x' + iy'$ la loro somma è

$$(x + iy) + (x' + iy') = (x + x') + i(y + y').$$

Ad esempio

$$(2 + i3) + (4 + i5) = (2 + 4) + i(3 + 5) = 6 + i8.$$

Il prodotto di due numeri complessi $x + iy$ e $x' + iy'$ è definito come segue

$$(x + iy) \cdot (x' + iy') = (xx' - yy') + i(yx' + xy').$$

In altri termini per moltiplicare due numeri complessi, basta prima eseguire il prodotto come se si trattasse di un'espressione algebrica nell'incognita i:

$$(x + iy) \cdot (x' + iy') = xx' + ixy' + iyx' + i^2yy'$$

e poi sostituire $i^2 = -1$ e raccogliere la parte reale e quella immaginaria

$$xx' + ixy' + iyx' + i^2yy' = xx' + ixy' + iyx' - yy' = (xx' - yy') + i(yx' + xy').$$

Per esempio, dati i due numeri complessi $(2 + 3i)$ e $(4 + 5i)$ si ha

$$(2 + 3i) \cdot (4 + 5i) = 2 \cdot 4 + 2 \cdot 5i + 3i \cdot 4 + 3i \cdot 5i = 8 + 10i + 12i + 15i^2$$

e poi ricordando che $i^2 = -1$:

$$8 + 10i + 12i + 15i^2 = 8 + 22i - 15 = -7 + 22i.$$

In questo modo abbiamo definito una somma e un prodotto che, come si verifica facilmente, soddisfano tutte le proprietà della definizione di campo.

Ad esempio, c'è un elemento neutro per la somma: lo 0 (come numero complesso, sarebbe $0 + 0i$).

Per quello che riguarda tutte le altre proprietà, verifichiamo solo l'esistenza dell'opposto (cioè dell'inverso rispetto alla somma) e dell'inverso rispetto al prodotto (per ogni numero complesso diverso da zero). Chiaramente, l'opposto di un numero complesso $z = x + iy$ è il numero complesso $-x - iy$ che si ottiene da z cambiando di segno la sua parte reale e la sua parte immaginaria. Per come abbiamo definito la somma, è immediato vedere che la somma di questi due numeri è uguale a zero, come prevede la definizione di opposto:

$$(x + iy) + (-x - iy) = (x - x) + i(y - y) = 0 + i0 = 0.$$

Ad esempio

$$(2 + i3) + (-2 - i3) = (2 + (-2)) + i(3 + (-3)) = 0 + i0 = 0.$$

Per quanto riguarda l'inverso di un numero complesso , affermiamo che se $x + iy \neq 0$ (cioè x e y non sono entrambi nulli) allora il suo inverso è il numero complesso

$$\frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}\,. \tag{4.37}$$

Possiamo quindi scrivere

$$\frac{1}{x + iy} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}.$$

Per verificarlo, scriviamo, ponendo tutto a un unico denominatore, la (4.37) nella forma

seguente

$$\frac{x - iy}{x^2 + y^2}\,. \tag{4.38}$$

Allora, è facile verificare che moltiplicando $x + iy$ per tale numero si ottiene 1 (come prevede la definizione di inverso):

$$(x + iy)\frac{x - iy}{x^2 + y^2} = \frac{(x + iy)(x - iy)}{x^2 + y^2} =$$

(sfruttando la nota formula $(a + b)(a - b) = a^2 - b^2$ nel caso $a = x$ e $b = iy$)

$$= \frac{x^2 - i^2y^2}{x^2 + y^2} = \frac{x^2 + y^2}{x^2 + y^2} = 1$$

dove nella penultima uguaglianza abbiamo usato il fatto che $i^2 = -1$.

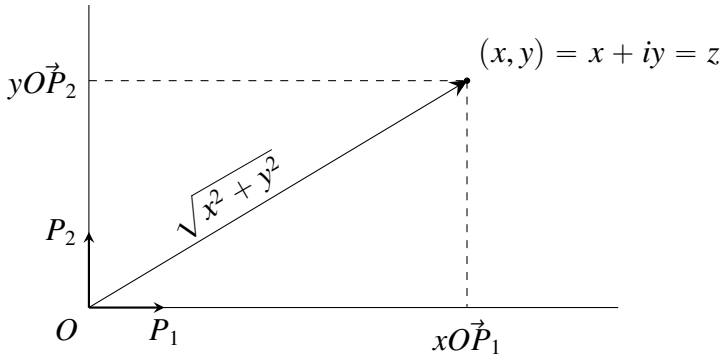
Ad esempio se $x+iy = 2+i3$, ovvero $x = 2, y = 3$, si ha $x^2+y^2 = 2^2+3^2 = 4+9 = 13$ e quindi in base alla formula

$$\frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2} = \frac{2}{13} - i\frac{3}{13}$$

è l'inverso di $2 + i3$.

L'espressione $x - iy$ che compare a numeratore della (4.38) si dice il **complesso coniugato** del numero complesso $x + iy$. Il coniugato di un numero $z \in \mathbb{C}$ si ottiene quindi semplicemente cambiando di segno la parte immaginaria di z. Il complesso coniugato di z si denota con $\bar{z}$.

Si chiama invece **norma** di un numero complesso $z = x+iy$ il numero reale $\sqrt{x^2 + y^2}$ e si denota con $|z|$ (analogamente alla notazione usata per il valore assoluto di un numero reale). La norma può essere visualizzata geometricamente come segue. Nel linguaggio delle sezioni precedenti, si fissi una base ortonormale del piano. Allora ogni numero complesso $z = x + iy$ è determinato da due numeri reali, la sua parte reale x e la sua parte immaginaria y. Per cui possiamo identificarlo con la coppia ordinata (x, y) e rappresentarlo quindi sul piano dove (x, y) è la coppia delle coordinate di un punto P ripetto ad una certa base ortonormale fissata $\{\vec{OP}_1, \vec{OP}_2\}$. Allora, come si vede nel disegno seguente



la norma $|z| = \sqrt{x^2+y^2}$ è, in base a quanto visto precedentemente, la lunghezza del segmento che congiunge l'origine O del piano cartesiano con il punto che identifica z. Essa rappresenta quindi la "lunghezza" del numero complesso.

Si noti che

$$z \cdot \bar{z} = |z|^2.$$

come il seguente conto dimostra

$$z \cdot \bar{z} = (x+iy)\cdot(x-iy) = x^2 - (iy)^2 = x^2 - i^2y^2 = x^2 + y^2 = (\sqrt{x^2+y^2})^2 = |z|^2.$$

Non è difficile verificare che l'operazione di coniugato gode delle seguenti proprietà rispetto alla somma e al prodotto di due numeri complessi $z_1 = x_1 + iy_1$ e $z_2 = x_2 + iy_2$.

(1) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$.

Dimostrazione.

$$\begin{aligned}\overline{z_1 + z_2} &= \overline{(x_1+x_2) + i(y_1+y_2)} \\ &= (x_1+x_2) - i(y_1+y_2) \\ &= (x_1 - iy_1) + (x_2 - iy_2) \\ &= \bar{z}_1 + \bar{z}_2.\end{aligned}$$

□

(2) $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$.

Dimostrazione.

$$\begin{aligned}
\overline{z_1 z_2} &= \overline{(x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1)} \\
&= (x_1 x_2 - y_1 y_2) - i(x_1 y_2 + x_2 y_1) \\
&= (x_1 x_2 - (-y_1)(-y_2)) + i(x_1(-y_2) + x_2(-y_1)) \\
&= (x_1 - i y_1)(x_2 - i y_2) \\
&= \bar{z}_1 + \bar{z}_2.
\end{aligned}$$

□

Capitolo 5

Sistemi di equazioni lineari e matrici

Per definire rigorosamente cosa intendiamo per equazioni lineari e scrivere una generica equazione lineare dobbiamo prima stabilire una notazione conveniente per denotare tali equazioni.
Infatti, per non avere limitazioni sul numero delle incognite, non possiamo continuare a indicarle con le lettere dell'alfabeto x, y, z etc., che sono in numero limitato. Per cui useremo sempre la stessa lettera, tradizionalmente la x, con degli indici numerici che ci dicono di quale incognita si tratta: x_1 indicherà la prima incognita, x_2 la seconda, e così via. In generale x_n indicherà la n-esima incognita, dove n è un numero naturale. Inoltre per ogni incognita x_i, denotiamo il suo coefficiente con una lettera, a, con lo stesso indice dell'incognita.

Definizione 5.1. Si definisce **equazione lineare in n incognite** $x_1, x_2, \ldots, x_n$ a coefficienti in un campo $\mathbb{K}$ un'equazione del tipo

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{5.1}$$

dove $b, a_1, a_2, \ldots, a_n$ sono elementi del campo $\mathbb{K}$ che vengono detti rispettivamente **termine noto** e **coefficienti delle incognite**.

Per noi i coefficienti $a_1, a_2, \ldots, a_n$ ed il termine noto b saranno spesso elementi del campo $\mathbb{R}$ ovvero saranno numeri reali.
Dare una soluzione dell'equazione (5.1) significa trovare degli elementi del campo, ovvero dei numeri, che sostituiti alle incognite rendano l'uguaglianza vera.
Ad esempio, nell'equazione lineare in due incognite $x_1 - x_2 = 1$ a coefficienti nel campo dei reali $\mathbb{R}$, ponendo $x_1 = 2$ e $x_2 = 1$ si ottiene l'uguaglianza vera $2 - 1 = 1$, mentre ad esempio ponendo $x_1 = 1$ e $x_2 = 2$ si ottiene $1 - 2 = 1$ che è falsa.
Da questo semplice esempio si vede come dare una soluzione dell'equazione $x_1 - x_2 = 1$ significa non solo dare *due* valori numerici, da sostituire alle due incognite dell'equazione, ma anche precisare quale valore vada sostituito alla prima incognita e quale alla seconda, ovvero specificare in quale ordine stiamo prendendo questi due elementi.

Risulta quindi chiaro che la soluzione data è la *coppia ordinata* di numeri $(2,1)$. Al contrario la coppia $(1,2)$ non è una soluzione dell'equazione $x_1 - x_2 = 1$.
Analogamente una soluzione di un'equazione lineare in 3 incognite sarà data da una terna ordinata. Ad esempio, la terna ordinata $(3,2,1)$ è una soluzione dell'equazione è $x_1 - x_2 + x_3 = 2$ in quanto sostituendo $x_1 = 3, x_2 = 2, x_3 = 1$ si ottiene l'uguaglianza vera $3 - 2 + 1 = 2$. La terna $(2,1,3)$ invece, non è una sua soluzione.
In generale, per equazioni con n incognite dovremo usare n-uple ordinate $(v_1, v_2, \ldots, v_n)$: possiamo allora dare la seguente:

Definizione 5.2. Data un'equazione lineare $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ in n incognite a coefficienti in un campo $\mathbb{K}$, si dice **soluzione** dell'equazione una n-upla ordinata $(v_1, v_2, \ldots, v_n) \in \mathbb{K}^n$ tale che l'uguaglianza $a_1v_1 + a_2v_2 + \cdots + a_nv_n = b$ risulti vera.

In molte occasioni ci si trova a dover considerare più equazioni lineari contemporaneamente. Per queste occorrenze abbiamo bisogno di definire i *sistemi di equazioni lineari*.
Per scrivere un sistema generico dobbiamo risolvere un problema di notazione simile a quello affrontato quando abbiamo scritto la generica equazione lineare. In questo caso abbiamo bisogno di una notazione efficace per indicare i diversi coefficienti delle incognite nelle diverse equazioni del sistema.
A questo scopo, nell'espressione della generica equazione lineare $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ doteremo sia i coefficienti che il termine noto di un ulteriore pedice che serve ad indicare di quale equazione del sistema si tratta. La prima equazione del sistema sarà cioè denotata con $a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$, la seconda con $a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$ e così via.

Definizione 5.3. Un **sistema di m equazioni lineari in n incognite** è un insieme di m equazioni lineari nelle stesse n incognite e viene denotato con

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ . \\ . \\ . \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{5.2}$$

Definizione 5.4. Una soluzione del sistema (5.2) è una n-upla $(v_1, v_2, \ldots, v_n) \in \mathbb{K}^n$ che è soluzione di tutte le equazioni del sistema.

Ovviamente, per conoscere un sistema abbiamo bisogno solo di sapere quali sono i coefficienti che moltiplicano ogni singola incognita e i termini noti di tutte le equazioni.

Quindi, se, dato un sistema, scriviamo una tabella di numeri disposti in righe e in colonne in modo che in ogni riga ci siano i coefficienti delle incognite di una certa equazione (ordinati secondo le incognite) e il termine noto, tale tabella conterrà tutte le informazioni che ci servono sul sistema. Ad esempio, il sistema

$$\begin{cases} x_1 + 3x_2 = 5 \\ 2x_1 - x_2 = 4 \end{cases} \tag{5.3}$$

può essere rappresentato dalla tabella

$$\begin{pmatrix} 1 & 3 & 5 \\ 2 & -1 & 4 \end{pmatrix} \tag{5.4}$$

In generale, tutte le informazioni relative al sistema (5.2) sono contenute nella tabella

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix} \tag{5.5}$$

che chiameremo **matrice completa del sistema**.
Come vedremo, non solo la matrice completa costituisce una "fotografia" fedele di un sistema e contiene tutte le informazioni necessarie a determinarlo, ma sarà anche l'oggetto sul quale conviene lavorare per risolverlo.
Se ci limitiamo invece ai coefficienti delle incognite otterremo la cosiddetta **matrice dei coefficienti del sistema**. Ad esempio, la matrice dei coefficienti del sistema (5.3) è

$$\begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}$$

mentre quella del sistema (5.2) è

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Come vedremo, il concetto di matrice è di fondamentale importanza e comparirà in molti contesti in questo corso. Nei capitoli successivi ne faremo una trattazione indipendente e più approfondita. Per il momento, limitiamoci a definire una matrice come una tabella rettangolare di elementi di $\mathbb{K}$, detti le sue **entrate**, disposti in righe e in colonne. Analogamente alla notazione che abbiamo introdotto per identificare i coefficienti delle incognite di un sistema, per denotare la generica entrata di una matrice

useremo due indici. Il primo ci dice in quale riga della matrice si trova, il secondo in quale colonna. Una generica matrice sarà quindi

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{5.6}$$

Come si vede, tale matrice ha m righe e n colonne. La sua generica entrata è del tipo a_{ij}, dove il primo indice, detto *indice di riga*, va da 1 a m, mentre il secondo, detto *indice di colonna*, va da 1 a n. Si dice anche che a_{ij} è *l'entrata di posto i j*.
Ad esempio, nella matrice

$$\begin{pmatrix} 5 & 3 \\ 2 & -1 \\ 0 & 7 \end{pmatrix}$$

che ha tre righe e due colonne, 5 è la prima entrata della prima riga, e quindi $a_{11} = 5$. In corrispondenza della terza riga e seconda colonna troviamo il numero 7, quindi $a_{32} = 7$. E così via per le altre entrate.

5.1 Equazioni superflue e equazioni incompatibili

Nel prossimo paragrafo vedremo come determinare tutte le soluzioni di un sistema lavorando sulla sua matrice completa. In particolare, scopriremo che possono verificarsi solamente tre possibilità[1]:

- il sistema non ha nessuna soluzione
- il sistema ha una sola soluzione
- il sistema ha infinite soluzioni

Prima di entrare nei dettagli, vediamo un esempio di ciascuna di queste possibilità, con l'obiettivo di capire le ragioni per cui esse possono verificarsi.

Non è difficile esibire un esempio di sistema con infinite soluzioni. Ad esempio, consideriamo il seguente sistema formato da una sola equazione in due incognite

$$\begin{cases} x_1 + x_2 = 0. \end{cases}$$

[1] Questo è un fatto caratteristico delle equazioni lineari: per una generica equazione possono verificarsi anche altri casi. Ad esempio l'equazione $x^2 = 9$ ha due soluzioni, $x = 3$ e $x = -3$.

Una soluzione del sistema è una coppia di numeri reali la cui somma è zero. Questo significa che i numeri devono essere uno l'opposto dell'altro. Scelto un qualunque $t \in \mathbb{R}$, la coppia $(t, -t)$ è una soluzione: le soluzioni sono quindi infinite, tante quanti i numeri reali.
Aggiungiamo ora una seconda equazione, ottenendo quindi un sistema di due equazioni in due incognite:

$$\begin{cases} x_1 + x_2 = 0 \\ x_1 - x_2 = 0 \end{cases} \tag{5.7}$$

Le soluzioni del sistema sono quindi le coppie che soddisfano non solo la prima equazione, ma anche la seconda. Quest'ultima afferma semplicemente che $x_1 = x_2$, cioè i due elementi della coppia devono essere non solo opposti ma anche uguali tra loro. L'unico numero reale uguale al suo opposto è lo zero, e quindi il sistema ha come unica soluzione la coppia $(0, 0)$.
Questo esempio suggerisce che in generale più equazioni ci sono in un sistema, maggiori sono i vincoli che imponiamo sulle incognite e quindi esisteranno meno n-uple che soddisfino tutte le equazioni, ovvero meno soluzioni:. L'esempio del sistema (5.7) sembra suggerire che con due incognite, due condizioni siano sufficienti a ottenere una sola soluzione.
Tuttavia, è facile fornire un altro esempio che sembra contraddire questa prima impressione. Consideriamo il sistema

$$\begin{cases} x_1 + x_2 = 0 \\ 2x_1 + 2x_2 = 0 \end{cases} \tag{5.8}$$

Ora, è immediato vedere che le soluzioni $(t, -t)$ della prima equazione soddisfano anche la seconda, quindi il sistema continua ad avere le infinite soluzioni della forma $(t, -t)$. Questo accade perché la seconda equazione è in realtà del tutto equivalente alla prima. Infatti mettendo in evidenza il 2, si può riscrivere $2x_1 + 2x_2 = 0$ come $2(x_1 + x_2) = 0$, ovvero, dividendo per 2, proprio la prima equazione. Per cui la seconda equazione non aggiunge nessun nuovo vincolo sulle incognite.

Definizione 5.5. Un'equazione di un sistema si dice **superflua** se ogni soluzione comune delle altre equazioni è anche una sua soluzione.

Le equazioni superflue presenti in un sistema possono essere tuttavia molto meno evidenti che nel caso appena visto. Ad esempio, consideriamo il sistema di due equazioni in tre incognite

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 + x_2 + 3x_3 = 2 \end{cases} \tag{5.9}$$

Una qualunque terna (x_1, x_2, x_3) che verifica le due equazioni soddisfa necessariamente anche l'uguaglianza che si ottiene sommandole membro a membro. In altre parole soddisfa l'equazione

$$(x_1 + x_2 + x_3) + (2x_1 + x_2 + 3x_3) = 1 + 2$$

cioè, svolgendo i conti,

$$3x_1 + 2x_2 + 4x_3 = 3\,.$$

Essendo tale equazione una conseguenza delle prime due, aggiungerla al sistema non modifica l'insieme delle soluzioni. Quindi il sistema

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 + x_2 + 3x_3 = 2 \\ 3x_1 + 2x_2 + 4x_3 = 3 \end{cases} \tag{5.10}$$

contiene un'equazione superflua, cioè dipendente dalle altre, certamente meno evidente che nel caso del sistema (5.8).

Si noti che, nella matrice completa del sistema (5.10)

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 3 & 2 \\ 3 & 2 & 4 & 3 \end{pmatrix} \tag{5.11}$$

il fatto che la terza equazione sia stata ottenuta sommando le prime due membro a membro si traduce nel fatto che la terza riga della matrice è somma delle prime due. Più precisamente ogni entrata della terza riga è la somma delle corrispondenti entrate delle altre due righe (la prima con la prima: 1+2 = 3; la seconda con la seconda: 1+1=2 etc.). Denotando con R_1, R_2, R_3 le tre righe, possiamo esprimere questa relazione usando la notazione $R_3 = R_1 + R_2$.

Naturalmente, equazioni superflue possono essere ottenute anche con combinazioni più complicate della somma delle prime due equazioni. Ad esempio, sempre in riferimento al sistema (5.9), una terna che soddisfi le due equazioni necessariamente soddisfa anche l'uguaglianza

$$5(x_1 + x_2 + x_3) + (-3)(2x_1 + x_2 + 3x_3) = 5 \cdot 1 + (-3) \cdot 2$$

cioè, svolgendo i conti,

$$-x_1 + 2x_2 - 4x_3 = -1.$$

Quindi anche nel sistema

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 + x_2 + 3x_3 = 2 \\ -x_1 + 2x_2 - 4x_3 = -1 \end{cases} \tag{5.12}$$

la terza equazione è superflua, in un modo forse ancora meno evidente.
Anche qui, nella matrice completa del sistema (5.12)

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 3 & 2 \\ -1 & 2 & -4 & -1 \end{pmatrix} \tag{5.13}$$

la relazione di dipendenza tra le equazioni si traduce nella corrispondente relazione di dipendenza tra le righe, che stavolta possiamo scrivere come $R_3 = 5R_1 + (-3)R_2$. In altri termini ogni entrata della terza riga si ottiene moltiplicando la corrispondente entrata della prima riga per 5 e sommando la corrispondente entrata della seconda riga moltiplicata per -3.

Definizione 5.6. Le righe di una matrice con m righe $R_1, R_2, \ldots, R_m$, si dicono **dipendenti tra loro** se una di queste si scrive come combinazione delle altre, ovvero se

$$R_i = c_1R_1 + c_2R_2 + \cdots + c_{i-1}R_{i-1} + c_{i+1}R_{i+1} + \cdots + c_mR_m$$

per qualche $i \in \{1, 2, \ldots, m\}$.

Osservazione 5.7. Se la matrice è la matrice completa di un sistema, il fatto che le sue righe siano dipendenti equivale al fatto che nel sistema ci sono equazioni superflue, nel senso della Definizione 5.5.

Per quello che riguarda i sistemi senza soluzioni, è altrettanto semplice esibirne uno. Ad esempio, il sistema di due equazioni in due incognite

$$\begin{cases} x_1 + x_2 = 0 \\ x_1 + x_2 = 1 \end{cases}$$

è evidentemente privo di soluzioni, in quanto se la somma di due numeri è uguale a 0 non può allo stesso tempo essere uguale a 1.
In altre parole, le due equazioni del sistema sono tra loro incompatibili, ovvero esprimono condizioni contraddittorie.

Definizione 5.8. Due equazioni si dicono **incompatibili** se gli insiemi delle loro soluzioni sono disgiunti, cioè se nessuna soluzione della prima è soluzione della seconda (o viceversa).

Un sistema si dice **incompatibile** se contiene due equazioni incompatibili (ed in tal caso non ha soluzioni). Per contro, un sistema si dice **compatibile** se ammette almeno una soluzione.

Analogamente a quanto fatto sopra per le equazioni superflue, si possono costruire esempi di sistemi in cui l'incompatibilità di una equazione con le altre non è così evidente come nel sistema precedente.
Ad esempio, prendiamo sempre come punto di partenza il sistema (5.9). Come abbiamo visto sopra, una terna che soddisfi le due equazioni soddisfa anche l'uguaglianza $3x_1 + 2x_2 + 4x_3 = 3$ che si ottiene sommando le due equazioni membro a membro.
Ma allora, se modifichiamo solo il termine noto di quest'ultima uguaglianza, ne otteniamo una che è incompatibile con le altre due. Per esempio, il sistema

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 + x_2 + 3x_3 = 2 \\ 3x_1 + 2x_2 + 4x_3 = 5 \end{cases} \tag{5.14}$$

non ha soluzioni, perchè per una qualunque terna (x_1, x_2, x_3) che soddisfi le prime due equazioni si deve avere che $3x_1 + 2x_2 + 4x_3$ è uguale a 3, e non a 5.
Confrontando la matrice dei coefficienti e la matrice completa del sistema (5.14)

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 3 \\ 3 & 2 & 4 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 3 & 2 \\ 3 & 2 & 4 & 5 \end{pmatrix}$$

notiamo che l'incompatibilità delle equazioni si traduce nel fatto che la terza riga è somma delle prime due nella matrice dei coefficienti ma non nella matrice completa (l'ultima entrata non soddisfa $5 = 1 + 2$). Riassumendo, la matrice dei coefficienti presenta una relazione di dipendenza tra le sue righe che nella matrice completa non vale (in quanto l'incompatibilità è stata ottenuta sommando i primi membri delle due equazioni, che contengono i coefficienti delle incognite, ma non i termini noti).

Definizione 5.9. Un sistema in cui i termini noti sono tutti nulli si dice **omogeneo**.

Osservazione 5.10. Un sistema omogeneo ha sempre almeno la soluzione $(0, 0, \ldots, 0)$, in quanto ponendo tutte le incognite uguali a zero si ottengono uguaglianze vere. Quindi i sistemi omogenei sono sempre compatibili.

5.2 La risoluzione di un sistema lineare

Il metodo di risoluzione di un qualunque sistema con m equazioni lineari in n incognite che trattiamo in questa sezione può essere considerato come una generalizzazione dei

metodi tradizionalmente usati per la risoluzione dei sistemi di due equazioni in due incognite. Per ricordare quali sono questi metodi, prendiamo ad esempio il sistema

$$\begin{cases} x_1 + x_2 = 0 \\ -x_1 + x_2 = 1 \end{cases} \tag{5.15}$$

Solitamente, per risolvere tale sistema si ricava una delle incognite in funzione dell'altra usando una delle due equazioni. Per esempio dalla prima equazione si trova $x_1 = -x_2$, e si sostituisce l'espressione così ottenuta nell'altra equazione:

$$-(-x_2) + x_2 = 1$$

ovvero

$$2x_2 = 1.$$

In questo modo, abbiamo *eliminato* la prima incognita dalla seconda equazione che è diventata una semplice equazione di primo grado con una sola incognita, che ha come soluzione $x_2 = \frac{1}{2}$. A questo punto, per ricavare x_1 basta sostituire il valore di x_2 nella prima equazione:

$$x_1 + \frac{1}{2} = 0 \;\rightarrow\; x_1 = -\frac{1}{2}.$$

Quello che ci ha permesso di risolvere il sistema è stato quindi aver ridotto il numero di incognite presenti in una delle equazioni.
Allo stesso risultato si può arrivare, ad esempio, sommando membro a membro le due equazioni. Se $x_1 + x_2 = 0$ e $-x_1 + x_2 = 1$ allora

$$(x_1 + x_2) + (-x_1 + x_2) = 0 + 1$$

e quindi, svolgendo i conti, si ottiene come sopra $2x_2 = 1$.
Questo secondo metodo, apparentemente più artificioso, si rivela in realtà più semplice se si lavora sulla matrice completa del sistema invece che sulle equazioni. Infatti, la matrice completa del sistema (5.15) è

$$\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}$$

Poiché le due righe della matrice ci danno coefficienti e termini noti delle equazioni, sommare membro a membro le due equazioni equivale a sommare tra loro le due righe. Sostituendo poi tale somma alla seconda riga originale si ottiene, senza dover maneggiare le incognite e dover fare sostituzioni o semplificazioni, la matrice

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$$

Questa corrisponde proprio al sistema ridotto

$$\begin{cases} x_1 + x_2 = 0 \\ 2x_2 = 1 \end{cases}$$

le cui soluzioni si ottengono, come abbiamo visto sopra, risolvendo prima l'equazione con una sola incognita.

Questo stesso procedimento di eliminazione di incognite, realizzato lavorando sulle righe della matrice completa, ci permette di risolvere qualunque sistema lineare, indipendentemente dal numero di equazioni e di incognite.
Più precisamente, ci porremo l'obiettivo di trasformare le equazioni in modo che a partire dalla prima in esse compaiano sempre meno incognite. Se, per darci un criterio, scegliamo di eliminarle seguendo l'ordine $x_1, x_2, \ldots, x_n$, questo significa che vogliamo che le righe della matrice completa inizino con un numero sempre maggiore di zeri. Ad esempio, la seguente matrice

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 2 \\ 0 & 0 & 4 & 5 \end{pmatrix}$$

nella quale le righe iniziano con un numero sempre maggiore di zeri, ha come sistema corrispondente

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_2 + 3x_3 = 2 \\ 4x_3 = 5 \end{cases}$$

nel quale le equazioni presentano un numero decrescente di incognite.

Definizione 5.11. Una matrice si dice **a gradini** se, andando dalla prima all'ultima, ogni riga inizia con un numero sempre maggiore di zeri.
Il primo elemento non nullo in ogni riga di una matrice a gradini si chiama **pivot**.

In altre parole, una matrice è a gradini se in ogni riga il pivot ha un indice di colonna strettamente maggiore del pivot della riga precedente. Ad esempio, delle matrici seguenti

$$\begin{pmatrix} 7 & 1 & 1 & 3 \\ 0 & 4 & 3 & 5 \\ 0 & 0 & 0 & 6 \end{pmatrix}, \quad \begin{pmatrix} 7 & 1 & 1 & 3 \\ 0 & 1 & 3 & 5 \\ 0 & 2 & 0 & 6 \end{pmatrix}, \quad \begin{pmatrix} 7 & 1 & 1 & 3 \\ 0 & 0 & 3 & 5 \\ 0 & 4 & 0 & 6 \end{pmatrix}$$

la prima è a gradini perché i suoi pivot (7 nella prima riga, 4 nella seconda e 6 nella terza) si trovano, nell'ordine, sulla prima, seconda e quarta colonna (indice di colonna sempre più grande), mentre le altre due no (nella seconda, il primo elemento non nullo della terza riga sta nella stessa colonna del primo elemento non nullo della seconda riga; nella terza matrice, il primo elemento non nullo della terza riga sta in una colonna di indice più piccolo del primo elemento non nullo della seconda riga).
Un sistema si dice a gradini se la sua matrice completa è una matrice a gradini.

Vogliamo ora presentare un metodo per ridurre una matrice qualunque in una matrice a gradini in modo da applicarlo alle matrice completa di un sistema di equazioni per risolverlo. Questo procedimento si chiama appunto **metodo di riduzione a gradini** o, dal momento che consiste nell'eliminare incognite, **metodo di eliminazione di Gauss-Jordan**.
Come stiamo per vedere, il procedimento di riduzione a gradini, oltre a semplificare il sistema, fa emergere anche le eventuali incompatibilità e le eventuali equazioni superflue.

Per trasformare un sistema in un sistema a gradini, trasformeremo la sua matrice completa in una matrice a gradini effettuando le seguenti operazioni sulle sue righe, dette *operazioni elementari di primo, secondo e terzo tipo*:

(I) Scambiare tra loro due righe della matrice (in simboli, $R_i \leftrightarrow R_j$)

(II) Moltiplicare una riga della matrice per un coefficiente non nullo (in simboli, $R_i \to cR_i$, con $c \neq 0$)

(III) Sommare a una riga della matrice un'altra riga moltiplicata per un numero qualunque (in simboli, $R_i \to R_i + dR_j$)

Il fatto importante è che tali operazioni, che modificano le righe, corrispondono a modificare le equazioni del sistema *in modo però da non cambiare l'insieme delle soluzioni*.

Proposizione 5.12. *La matrice ottenuta dalla matrice completa di un sistema col metodo di Gauss-Jordan è la matrice completa di un sistema* **equivalente** *a quello iniziale (ovvero avente le stesse soluzioni del sistema iniziale).*

Dimostrazione. Per quanto riguarda le prime due operazioni questo è abbastanza chiaro. In particolare scambiare tra loro le righe della matrice equivale a cambiare l'ordine delle equazioni corrispondenti, e questo chiaramente non modifica le soluzioni del sistema che sono soluzioni comuni alle equazioni indipendentemente dall'ordine in cui le mettiamo. Inoltre moltiplicare una riga della matrice per $c \neq 0$ equivale a moltiplicare entrambi i membri dell'equazione corrispondente per c, e anche questo produce

un'equazione equivalente.[2] Denotiamo ora le equazioni di un sistema con le lettere $R_1, \ldots, R_m$ con un abuso di notazione. Ora per le operazioni del terzo tipo supponiamo di voler dimostrare che i sistemi

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{i1}x_1 + \cdots + a_{in}x_n = b_i \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \qquad \text{e}$$

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ (a_{i1}x_1 + \cdots + a_{in}x_n) + d(a_{j1}x_1 + \cdots + a_{jn}x_n) = b_i + db_j \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

sono equivalenti. Vogliamo cioè dimostrare che tutte le soluzioni del primo sono soluzioni del secondo e viceversa. Sia dunque $(v_1, v_2, \ldots, v_n)$ una soluzione del primo sistema. Allora dobbiamo solo mostrare che questa soddisfa anche l'equazione $(a_{i1}x_1 + \cdots + a_{in}x_n) + d(a_{j1}x_1 + \cdots + a_{jn}x_n) = b_i + db_j$. Ma questo è chiaro poichè $(v_1, v_2, \ldots, v_n)$ è in particolare soluzione della j-esima e i-esima equazione del primo sistema. Se, viceversa, $(v_1, v_2, \ldots, v_n)$ è soluzione del secondo sistema, allora in particolare è soluzione della sua j-esima equazione. Ma allora

$$\begin{aligned} (a_{i1}v_1 + \cdots + a_{in}v_n) + d(a_{j1}v_1 + \cdots + a_{jn}v_n) &= b_i + db_j \\ (a_{i1}v_1 + \cdots + a_{in}v_n) + d(b_j) &= b_i + db_j \\ a_{i1}v_1 + \cdots + a_{in}v_n &= b_i \end{aligned}$$

e quindi $(v_1, v_2, \ldots, v_n)$ è soluzione di tutte le equazioni del primo sistema. □

Ora, mostriamo tramite alcuni esempi come, usando le operazioni elementari, si possa trasformare un qualunque sistema in un sistema a gradini (che, in base alla Proposizione 5.12, sarà equivalente al sistema originale) e come poi risolvere tale sistema.
Si consideri il sistema

[2] Qui si capisce anche la condizione che c non debba essere zero: moltiplicando entrambi i membri dell'equazione per zero si ottiene $0 = 0$, che è sicuramente un'uguaglianza vera ma non può essere sostituita all'equazione originale, perché questo equivarrebbe a cancellarla dal sistema.

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ -x_1 + x_2 - 3x_3 = 0 \\ -x_1 + x_2 + x_3 = -3 \end{cases} \tag{5.16}$$

la cui matrice completa[3] è

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ -1 & 1 & -3 & 0 \\ -1 & 1 & 1 & -3 \end{array}\right). \tag{5.17}$$

Ora, vogliamo trasformare tale matrice in una matrice a gradini usando le operazioni elementari, in modo da ottenere un sistema a gradini equivalente al sistema (5.16).
Osserviamo che il primo elemento a_{11} della prima riga è diverso da zero. In base alla definizione di matrice a gradini i pivot della seconda e della terza riga non possono stare nella prima colonna. In altre parole, dobbiamo trasformare la matrice in modo che a_{21} e a_{31} siano uguali a zero.
Otteniamo sicuramente questo scopo se applichiamo le operazioni elementari del terzo tipo $R_2 \to R_2 + R_1$ e $R_3 \to R_3 + R_1$:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ -1 & 1 & -3 & 0 \\ -1 & 1 & 1 & -3 \end{array}\right) \underset{\substack{R_2 \to R_2 + R_1 \\ R_3 \to R_3 + R_1}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 2 & -2 & 1 \\ 0 & 2 & 2 & -2 \end{array}\right)$$

La matrice trasformata non è ancora una matrice a gradini in quanto il pivot della terza riga si trova in corrispondenza della stessa colonna (la seconda) del pivot della seconda riga. Dobbiamo far sì che $a_{32} = 0$. A questo scopo, basta applicare l'operazione elementare $R_3 \to R_3 - R_2$:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 2 & -2 & 1 \\ 0 & 2 & 2 & -2 \end{array}\right) \underset{R_3 \to R_3 - R_2}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 2 & -2 & 1 \\ 0 & 0 & 4 & -3 \end{array}\right)$$

Abbiamo ottenuto quindi una matrice a gradini. Il sistema

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_2 - 2x_3 = 1 \\ 4x_3 = -3 \end{cases} \tag{5.18}$$

corrispondente alla matrice trasformata è, come sappiamo, equivalente al sistema originale (5.16), quindi trovando le sue soluzioni avremo risolto il sistema (5.16).

[3] D'ora in poi, nella matrice completa tracceremo a volte una linea per separare la matrice dei coefficienti dalla colonna dei termini noti.

Il principale vantaggio di un sistema a gradini consiste nel fatto che nelle equazioni compaiono sempre meno incognite (leggendole dalla prima all'ultima). Per risolverlo, basta quindi iniziare a risolvere le equazioni da quella che contiene meno incognite (l'ultima) e risalire mediante sostituzioni fino alla prima. Più precisamente, dall'ultima equazione $4x_3 = -3$ ricaviamo subito $x_3 = -\frac{3}{4}$; sostituendo il valore così trovato nella seconda equazione troviamo

$$2x_2 - 2x_3 = 1 \longrightarrow 2x_2 = 1 + 2x_3 = 1 + 2(-\frac{3}{4}) = 1 - \frac{3}{2} = -\frac{1}{2} \longrightarrow x_2 = -\frac{1}{4}$$

e analogamente, sostituendo i valori di x_2 e x_3 così ottenuti nella prima equazione troviamo

$$x_1 + x_2 + x_3 = 1 \longrightarrow x_1 = 1 - x_2 - x_3 = 1 - (-\frac{1}{4}) - (-\frac{3}{4}) = 2$$

Riassumendo, la terna $\left(2, -\frac{1}{4}, -\frac{3}{4}\right)$ è l'unica soluzione del sistema (5.18), ovvero del sistema iniziale (5.16).

Ora vediamo altri due esempi significativi di risoluzione di un sistema lineare, che metteranno in evidenza ulteriori vantaggi della riduzione a gradini.
Consideriamo il sistema

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 - x_2 - x_3 = 0 \\ x_1 + 3x_2 + 3x_3 = 1 \end{cases} \tag{5.19}$$

che ha come matrice completa

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 0 \\ 1 & 3 & 3 & 1 \end{array}\right). \tag{5.20}$$

Come fatto per il sistema precedente, trasformiamo tale matrice in una matrice a gradini mediante operazioni elementari.

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 0 \\ 1 & 3 & 3 & 1 \end{array}\right) \underset{\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & -1 \\ 0 & 2 & 2 & 0 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & -1 \\ 0 & 2 & 2 & 0 \end{array}\right) \underset{R_3 \to R_3 + R_2}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & -1 \\ 0 & 0 & 0 & -1 \end{array}\right)$$

Notiamo che la terza riga della matrice trasformata corrisponde all'equazione $0x_1 + 0x_2 + 0x_3 = -1$, ovvero $0 = -1$. Poichè questa uguaglianza è falsa, non esiste nessuna terna che soddisfi le tre condizioni del sistema ridotto corrispondente, ovvero tale sistema non ha soluzioni. Questo, in virtù dell'equivalenza tra il sistema originale e quello ridotto, ci dice che il sistema di partenza non ha soluzioni, ovvero è incompatibile.
Necessariamente tra le equazioni del sistema di partenza vi era una incompatibilità non evidente che il procedimento di riduzione a gradini ha fatto emergere. In effetti, se moltiplichiamo membro a membro la prima equazione per 2 e le sottraiamo la seconda equazione otteniamo

$$2(x_1 + x_2 + x_3) - (x_1 - x_2 - x_3) = 2 \cdot 1 - 0$$

e quindi, svolgendo i calcoli,

$$x_1 + 3x_2 + 3x_3 = 2.$$

Questa condizione, che è conseguenza delle prime due equazioni ed è quindi soddisfatta da qualunque terna le soddisfi, è chiaramente incompatibile con la terza equazione. Il procedimento di riduzione a gradini ha messo in luce questa incompatibilità trasformandola nell'incompatibilità evidente $0 = -1$.

Consideriamo ora come ultimo esempio il sistema

$$\begin{cases} x_1 + x_2 + 3x_3 = 1 \\ x_1 - 2x_2 + x_3 = 0 \\ x_1 - 5x_2 - x_3 = -1 \end{cases} \tag{5.21}$$

che ha come matrice completa

$$\left(\begin{array}{ccc|c} 1 & 1 & 3 & 1 \\ 1 & -2 & 1 & 0 \\ 1 & -5 & -1 & -1 \end{array}\right). \tag{5.22}$$

Applicando operazioni elementari per ridurre a gradini,

$$\left(\begin{array}{ccc|c} 1 & 1 & 3 & 1 \\ 1 & -2 & 1 & 0 \\ 1 & -5 & -1 & -1 \end{array}\right) \underset{\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 3 & 1 \\ 0 & -3 & -2 & -1 \\ 0 & -6 & -4 & -2 \end{array}\right) \tag{5.23}$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 3 & 1 \\ 0 & -3 & -2 & -1 \\ 0 & -6 & -4 & -2 \end{array}\right) \underset{R_3 \to R_3 - 2R_2}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & 1 & 3 & 1 \\ 0 & -3 & -2 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right) \tag{5.24}$$

notiamo che la terza riga della matrice trasformata corrisponde all'equazione $0x_1 + 0x_2 + 0x_3 = 0$, ovvero $0 = 0$.

Quest'ultima condizione è un'identità vera indipendentemente dal valore che diamo alle incognite, quindi essa può essere cancellata dal sistema senza influire sulle sue soluzioni. In altre parole, il sistema iniziale di tre equazioni si è trasformato nel sistema di due equazioni equivalente

$$\begin{cases} x_1 + x_2 + 3x_3 = 1 \\ -3x_2 - 2x_3 = -1 \end{cases} \tag{5.25}$$

Benché non sia rimasta un'equazione con una sola incognita come nel primo sistema che abbiamo risolto, possiamo comunque procedere come segue. Ricaviamo x_2 dalla seconda equazione.

$$-3x_2 - 2x_3 = -1 \longrightarrow -3x_2 = 2x_3 - 1 \longrightarrow x_2 = -\frac{2}{3}x_3 + \frac{1}{3} \tag{5.26}$$

Poi sostituiamo l'espressione ottenuta nella prima equazione per ricavare x_1:

$$x_1 + x_2 + 3x_3 = 1 \longrightarrow x_1 = 1 - x_2 - 3x_3 = 1 - \left(-\frac{2}{3}x_3 + \frac{1}{3}\right) - 3x_3 = \frac{2}{3} - \frac{7}{3}x_3. \tag{5.27}$$

Ora, per qualunque valore $t \in \mathbb{R}$ assegniato a x_3, la (5.26) e la (5.27) implicano che ponendo $x_2 = -\frac{2}{3}t + \frac{1}{3}$ e $x_1 = \frac{2}{3} - \frac{7}{3}t$, le equazioni del sistema sono soddisfatte, ovvero si ottiene una soluzione del sistema. In altre parole, le soluzioni del sistema sono esattamente tutte le terne del tipo $(\frac{2}{3} - \frac{7}{3}t, -\frac{2}{3}t + \frac{1}{3}, t)$ al variare di $t \in \mathbb{R}$: il sistema ha infinite soluzioni.

Più precisamente, dal momento che le infinite soluzioni del sistema dipendono da un solo parametro libero t, si dice che il sistema ha "infinito alla uno" (si scrive ∞^1) soluzioni.

In generale, possiamo dare la seguente

Definizione 5.13. Un sistema di equazioni lineari ha ∞^k **soluzioni** se l'espressione generale della sua soluzione dipende da k parametri liberi.

Osservazione 5.14. Quando risolviamo un sistema ridotto a gradini, procediamo dall'ultima equazione alla prima ricavando per ogni equazione del sistema una incognita (eventualmente in funzione di altre incognite). Questo significa che, ammesso che il sistema sia compatibile, il numero di parametri liberi nell'espressione della soluzione sarà uguale alla differenza tra il numero delle incognite e il numero di equazioni non nulle rimaste dopo la riduzione.

In particolare, avremo un'unica soluzione (cioè nessun parametro libero) solo nel caso in cui il numero di equazioni non nulle rimaste *dopo la riduzione* sia uguale al numero di incognite[4].

Dal momento che il procedimento di riduzione a gradini produce una riga nulla quando tale riga era superflua (ovvero dipendente dalle altre), il numero di righe non nulle dopo la riduzione a gradini ci dà l'importante informazione di quante righe indipendenti aveva la matrice iniziale. Questo numero, che si chiama *rango della matrice*, ha un significato che va ben oltre la risoluzione di sistemi lineari.

Definizione 5.15. Data una matrice A, si dice **rango** di A il numero di righe non nulle nella matrice ottenuta riducendo A a gradini.

Osservazione 5.16. Il rango di una matrice, come lo abbiamo appena definito, dovrebbe essere chiamato più precisamente **rango per righe**, in quanto esso ci dice quante *righe* indipendenti ci sono in una matrice data. A priori, si potrebbe definire l'analoga nozione per le colonne, il **rango per colonne**, che ci dica quante *colonne* indipendenti ci sono in una matrice data. Tuttavia, questa distinzione è inutile in quanto un importante risultato, che non dimostriamo, afferma che *il rango per righe è sempre uguale al rango per colonne*, ovvero il numero di righe indipendenti di una matrice è uguale al numero delle sue colonne indipendenti.
Ad esempio, nella matrice

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 4 & 5 \end{pmatrix}$$

la seconda riga R_2 è evidentemente dipendente dalle altre, in quanto $R_2 = 2R_1$ (se volessimo far apparire anche la terza riga in questa relazione di dipendenza, potremmo equivalentemente scrivere $R_2 = 2R_1 + 0R_3$).
Per il risultato appena citato, allora anche una delle colonne della matrice deve essere dipendente dalle altre: in effetti, si ha

$$\begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

che era molto meno evidente della relazione di dipendenza esistente tra le righe.

Allora, quanto detto nell'Osservazione 5.14 si può riassumere nella seguente

Proposizione 5.17. *Se un sistema di equazioni lineari è compatibile, allora esso ha ∞^{n-r} soluzioni, dove n denota il numero di incognite e r è il rango della matrice. In particolare, se $r = n$ allora il sistema ha una sola soluzione.*

[4]Osserviamo che questa affermazione è falsa se non ci limitiamo a sistemi ridotti a gradini, come dimostra il semplice esempio del sistema (5.8) che ha due equazioni in due incognite ma infinite soluzioni.

5.3 Qualche applicazione geometrica

Vediamo ora alcuni problemi geometrici che possono essere risolti grazie ai concetti appena visti.
Ad esempio, consideriamo le due rette r e r' di equazioni cartesiane

$$r\colon \begin{cases} x+y+z=2 \\ 2x-y+z=5 \end{cases} \quad , \quad r'\colon \begin{cases} x-y-2z=-2 \\ x+3y+2z=2 \end{cases} \tag{5.28}$$

e supponiamo di voler determinare se esse hanno punti in comune.
Dal momento che i punti di una retta espressa in equazioni cartesiane sono proprio le soluzioni del sistema formato dalle due equazioni, i punti comuni alle due rette sono dati dalle soluzioni comuni a tutte e quattro le equazioni delle due rette, ovvero le soluzioni del sistema

$$\begin{cases} x+y+z=2 \\ 2x-y+z=5 \\ x-y-2z=-2 \\ x+3y+2z=2 \end{cases} \tag{5.29}$$

Riducendo la matrice completa otteniamo

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 2 & -1 & 1 & 5 \\ 1 & -1 & -2 & -2 \\ 1 & 3 & 2 & 2 \end{array}\right) \xrightarrow[\substack{R_3\to R_3-R_1 \\ R_4\to R_4-R_1}]{R_2\to R_2-2R_1} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -3 & -1 & 1 \\ 0 & -2 & -3 & -4 \\ 0 & 2 & 1 & 0 \end{array}\right)$$

$$\xrightarrow[R_4\to 3R_4+2R_2]{R_3\to 3R_3-2R_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -3 & -1 & 1 \\ 0 & 0 & -7 & -14 \\ 0 & 0 & 1 & 2 \end{array}\right) \xrightarrow{R_4\to 7R_4+R_3} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -3 & -1 & 1 \\ 0 & 0 & -7 & -14 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Vediamo quindi che il sistema è compatibile e, essendosi annullata una riga, la matrice ha rango 3. Quindi avendo 3 incognite, in base alla Proposizione 5.17, il sistema ha $\infty^{3-3}=1$ soluzioni, cioè una sola soluzione. Questa può essere determinata risolvendo il sistema ridotto corrispondente

$$\begin{cases} x+y+z=2 \\ -3y-z=1 \\ -7z=-14 \end{cases}$$

Dall'ultima equazione si ottiene $z=2$, che sostituito nella seconda dà

$$-3y = z + 1 = 2 + 1 = 3$$

ovvero $y = -1$. Sostituendo nella prima equazione per ottenere x si ha

$$x = 2 - y - z = 2 - (-1) - 2 = 1$$

Quindi l'unica equazione del sistema è data dalla terna $(1, -1, 2)$, che sono le coordinate del punto in cui si incontrano le due rette.

Supponiamo invece che le rette siano

$$r\colon \begin{cases} x + y + z = 0 \\ 2x + y - z = 1 \end{cases} \quad , \quad r'\colon \begin{cases} 2x - y = 3 \\ x + y - z = 1 \end{cases} \tag{5.30}$$

e supponiamo ancora di voler determinare se esse hanno punti in comune.

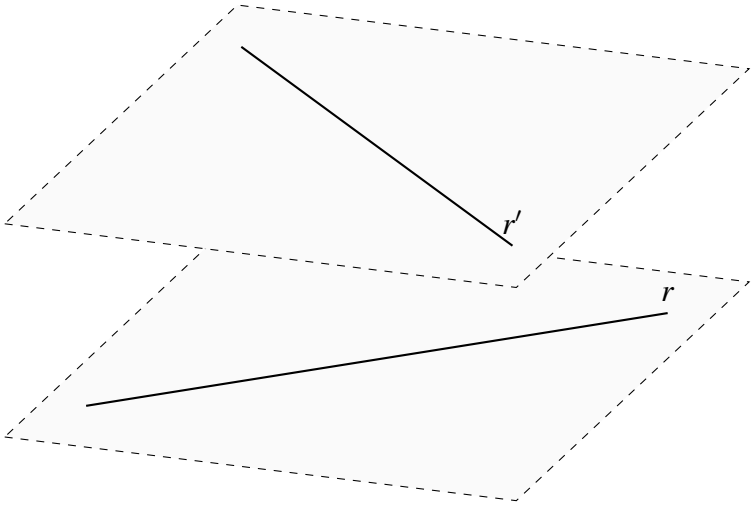
Come sopra, mettiamo insieme le quattro equazioni e riduciamo la matrice completa del sistema così ottenuto:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 1 & -1 & 1 \\ 2 & -1 & 0 & 3 \\ 1 & 1 & -1 & 1 \end{array}\right) \xrightarrow[\substack{R_3 \to R_3 - 2R_1 \\ R_4 \to R_4 - R_1}]{R_2 \to R_2 - 2R_1} \tag{5.31}$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -1 & -3 & 1 \\ 0 & -3 & -2 & 3 \\ 0 & 0 & -2 & 1 \end{array}\right) \xrightarrow{R_3 \to R_3 - 3R_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -1 & -3 & 1 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & -2 & 1 \end{array}\right) \xrightarrow{R_4 \to 7R_4 + 2R_3} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -1 & -3 & 1 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 7 \end{array}\right)$$

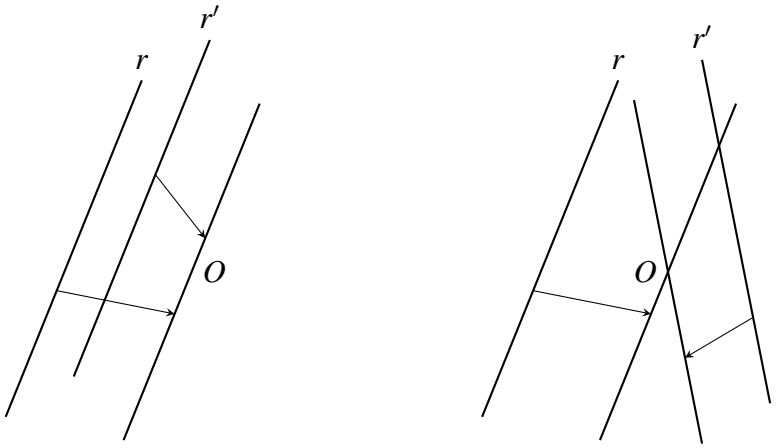
Essendo il sistema incompatibile (l'ultima riga corrisponde all'uguaglianza falsa $0 = 7$) deduciamo che le due rette non hanno punti in comune.

Mentre nel piano due rette che non hanno punti in comune sono necessariamente parallele, nello spazio tridimensionale questo non è più vero. Come si vede nel seguente disegno, due rette nello spazio, grazie alla dimensione in più presente rispetto al piano, possono trovarsi su piani paralleli e quindi non intersecarsi pur non avendo la stessa direzione.



Definizione 5.18. Due rette senza punti in comune si dicono **sghembe** se non sono parallele.

Vediamo ora come sia possibile determinare se le rette date sono sghembe o parallele senza ulteriori calcoli, ma sfruttando la riduzione già fatta.
Due rette che non hanno punti in comune sono parallele se e solo se coincidono quando traslate parallelamente a se stesse sull'origine (ovvero hanno infiniti punti in comune). Sono invece sghembe se e solo se hanno un solo punto in comune (l'origine) quando traslate parallelamente a se stesse sull'origine.



Possiamo così tradurre il problema di capire se le due rette hanno la stessa direzione nel problema di determinare un'intersezione (tra le rette traslate).
Per traslare una retta espressa in equazioni cartesiane, parallelamente a se stessa, basta modificare i termini noti delle equazioni lasciando invariati i primi membri poichè, come visto in precedenza questi traslano i piani parallelamente. In particolare, otteniamo la

traslazione sull'origine se poniamo i termini noti uguali a zero. Infatti in tal caso le equazioni risultano soddisfatte dalla terna $x = 0$, $y = 0$, $z = 0$, che sono le coordinate dell'origine, il che significa che la retta traslata è proprio quella che passa per l'origine. In altre parole, la retta parallela passante per l'origine è data in equazioni cartesiane dal sistema omogeneo associato.
Nel caso delle due rette r e r' date da (5.30), le parallele per l'origine sono rappresentate rispettivamente dalle equazioni

$$\begin{cases} x + y + z = 0 \\ 2x + y - z = 0 \end{cases}, \quad \begin{cases} 2x - y = 0 \\ x + y - z = 0 \end{cases} \tag{5.32}$$

Per determinare se tali rette traslate hanno infiniti punti in comune o uno solo dobbiamo risolvere il sistema che si ottiene mettendo insieme le 4 equazioni, che ha come matrice associata

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 1 & -1 & 0 \\ 2 & -1 & 0 & 0 \\ 1 & 1 & -1 & 0 \end{array}\right)$$

Per ridurre a gradini questa matrice dovremmo applicare esattamente le stesse operazioni usate per ridurre la (5.31), che differisce da essa solo per il fatto di avere tutti i termini noti uguali a zero. L'unica differenza è che i termini noti rimarrangono sempre nulli qualunque operazione elementare si applichi. Quindi, senza dover rifare i conti, sappiamo che arriveremmo alla stessa matrice ridotta ma con l'ultima colonna (quella dei termini noti) tutta nulla:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -1 & -3 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Guardando questa matrice, che rappresenta ora un sistema compatibile con 3 incognite e rango 3, concludiamo che traslando le rette sull'origine avremmo una sola soluzione (l'origine stessa) e quindi le rette di partenza non erano parallele.
Per vedere invece cosa succederebbe se le rette fossero parallele, consideriamo l'esempio seguente:

$$r\colon \begin{cases} x + y + z = 1 \\ x - y + 2z = 0 \end{cases}, \quad r'\colon \begin{cases} 2x + 3z = 3 \\ x + 3y = -3 \end{cases} \tag{5.33}$$

Allo scopo di controllare se le rette hanno punti in comune, mettiamo insieme le quattro equazioni e riduciamo la matrice completa del sistema così ottenuto:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 0 \\ 2 & 0 & 3 & 3 \\ 1 & 3 & 0 & -3 \end{array}\right) \begin{array}{c} \longrightarrow \\ {\scriptstyle R_2 \to R_2 - R_1} \\ {\scriptstyle R_3 \to R_3 - 2R_1} \\ {\scriptstyle R_4 \to R_4 - R_1} \end{array} \tag{5.34}$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -2 & 1 & -1 \\ 0 & -2 & 1 & 1 \\ 0 & 2 & -1 & -4 \end{array}\right) \begin{array}{c} \longrightarrow \\ {\scriptstyle R_3 \to R_3 - R_2} \\ {\scriptstyle R_4 \to R_4 + R_2} \end{array} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -2 & 1 & -1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & -5 \end{array}\right) \begin{array}{c} \longrightarrow \\ {\scriptstyle R_4 \to 2R_4 + 5R_3} \end{array} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -2 & 1 & -1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Come si vede, le rette non hanno punti in comune in quanto la terza riga corrisponde all'uguaglianza falsa $0 = 2$. D'altra parte il sistema omogeneo associato (formato dalle due rette traslate sull'origine) avrebbe come matrice ridotta la stessa matrice ottenuta ora ma con ultima colonna di zeri:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

che rappresenta la matrice di un sistema ridotto compatibile con 3 incognite e rango 2. Questo sistema ha quindi infinite soluzioni: questo significa che le due rette, traslate sull'origine, hanno infiniti punti in comune, ovvero coincidono. Concludiamo che le due rette di partenza erano parallele.

Osservazione 5.19. Per rette date in equazioni parametriche verificare se esse sono parallele o meno è immediato, in quanto come sappiamo dal capitolo precedente nelle equazioni parametriche

$$\begin{cases} x = x_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt \end{cases}$$

un vettore che rappresenta la direzione della retta è dato dalla terna (l, m, n) dei coefficienti di t. Basta quindi confontare i due vettori così ottenuti per ognuna delle due rette: queste avranno la stessa direzione se tali vettori sono proporzionali. Nel caso in cui le rette siano date in equazioni cartesiane, si può passare alle parametriche semplicemente risolvendo i sistemi dati dalle cartesiane stesse. Ad esempio, per le rette viste sopra in (5.33), riducendo la matrice completa delle cartesiane di r otteniamo

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 0 \end{array}\right) \begin{array}{c} \longrightarrow \\ {\scriptstyle R_2 \to R_2 - R_1} \end{array} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & 1 & -1 \end{array}\right)$$

ovvero il sistema ridotto

$$\begin{cases} x + y + z = 1 \\ -2y + z = -1 \end{cases}$$

Ponendo $z = t$ si ricava $-2y = -1 - t$ (ovvero $y = \frac{1}{2} + \frac{1}{2}t$) e $x = 1 - y - z = 1 - (\frac{1}{2} + \frac{1}{2}t) - t = \frac{1}{2} - \frac{3}{2}t$. Quindi r ha equazioni parametriche

$$\begin{cases} x = \frac{1}{2} - \frac{3}{2}t \\ y = \frac{1}{2} + \frac{1}{2}t \\ z = t \end{cases} \tag{5.35}$$

Analogamente, riducendo la matrice completa delle cartesiane di r' otteniamo

$$\left(\begin{array}{ccc|c} 2 & 0 & 3 & 3 \\ 1 & 3 & 0 & -3 \end{array}\right) \xrightarrow[R_2 \to 2R_2 - R_1]{} \left(\begin{array}{ccc|c} 2 & 0 & 3 & 1 \\ 0 & 6 & -3 & -9 \end{array}\right)$$

che corrisponde al sistema ridotto

$$\begin{cases} 2x + 3z = 3 \\ 6y - 3z = -9 \end{cases}$$

Ponendo $z = t$ si ricava $6y = 9 + 3t$ (ovvero $y = \frac{3}{2} + \frac{1}{2}t$) e $2x = 3 - 3z = 3 - 3t$, ovvero $x = \frac{3}{2} - \frac{3}{2}t$. Quindi r ha equazioni parametriche

$$\begin{cases} x = \frac{3}{2} - \frac{3}{2}t \\ y = \frac{3}{2} + \frac{1}{2}t \\ z = t \end{cases} \tag{5.36}$$

Confrontando i coefficienti di t nelle parametriche (5.35) e (5.36) vediamo che le rette r ed r' hanno entrambe direzione rappresentata dal vettore di coordinate $(-\frac{3}{2}, \frac{1}{2}, 1)$, e quindi sono parallele[5].

Un altro problema geometrico che può essere risolto con l'aiuto delle tecniche viste in questo capitolo è il seguente. Supponiamo di avere una retta data in equazioni cartesiane

$$r\colon \begin{cases} Ax + By + Cz = D \\ A'x + B'y + C'z = D' \end{cases} \tag{5.37}$$

Come abbiamo visto nel capitolo precedente, la (5.37) significa che la retta r è intersezione del piano dato dall'equazione cartesiana $Ax + By + Cz = D$ e dal piano di

[5] Si noti che sappiamo che le rette hanno la stessa direzione, che non esclude il caso in cui esse siano parallele coincidenti, ovvero che le equazioni cartesiane date rappresentassero in realtà la stessa retta.

equazione cartesiana $A'x + B'y + C'z = D'$. Le due equazioni che compongono le cartesiane sono quindi le equazioni di due particolari piani che contengono la retta.
Ora, vogliamo determinare *tutti* i piani che contengono la retta.

Proposizione 5.20. *La generica equazione cartesiana del piano che contiene la retta* (5.37) *è data da*

$$\alpha(Ax + By + Cz - D) + \beta(A'x + B'y + C'z - D') = 0 \tag{5.38}$$

al variare di $\alpha, \beta \in \mathbb{R}$ *(non entrambi nulli).*

Dimostrazione. Iniziamo con l'osservare che se un piano ha equazione della forma (5.38), allora esso contiene la retta. Infatti, dire che una retta è contenuta in un piano significa che se un punto appartiene alla retta allora esso appartiene anche al piano. Ma se un punto appartiene alla retta, allora le sue coordinate (x, y, z) soddisfano entrambe le equazioni $Ax + By + Cz = D$ e $A'x + B'y + C'z = D'$ della retta, e quindi

$$\alpha(Ax+By+Cz-D)+\beta(A'x+B'y+C'z-D') = \alpha(D-D)+\beta(D'-D') = \alpha\cdot 0+\beta\cdot 0 = 0.$$

Quindi (x, y, z) soddisfa anche l'equazione (5.38): il punto appartiene al piano rappresentato da tale equazione. Questo dimostra che, per ogni $\alpha, \beta \in \mathbb{R}$ (non entrambi nulli), l'equazione (5.38) rappresenta un piano che contiene la retta.

Viceversa, vogliamo mostrare che qualunque piano che contenga la retta può essere rappresentato nella forma (5.38). Per vederlo, osserviamo che un generico piano di equazione $A''x + B''y + C''z = D''$ contiene tutta la retta di equazioni (5.37) se e solo se ogni terna che soddisfa le equazioni $Ax + By + Cz = D$ e $A'x + B'y + C'z = D'$ soddisfa automaticamente anche l'equazione $A''x + B''y + C''z = D''$ del piano.
In altre parole, nel sistema

$$\begin{cases} Ax + By + Cz = D \\ A'x + B'y + C'z = D' \\ A''x + B''y + C''z = D'' \end{cases}$$

che si ottiene mettendo insieme tutte le cartesiane, la terza equazione è superflua ovvero dipendente dalle altre. A livello della matrice completa

$$\begin{pmatrix} A & B & C & D \\ A' & B' & C' & D' \\ A'' & B'' & C'' & D'' \end{pmatrix}$$

questo si traduce nel fatto che la terza riga deve essere combinazione lineare delle altre due, ovvero devono esistere $\alpha, \beta \in \mathbb{R}$ tali che

$$\begin{pmatrix} A'' & B'' & C'' & D'' \end{pmatrix} = \alpha \begin{pmatrix} A & B & C & D \end{pmatrix} + \beta \begin{pmatrix} A' & B' & C' & D' \end{pmatrix}$$

In altri termini devono essere soddisfatte le quattro condizioni

$$A'' = \alpha A + \beta A', \quad B'' = \alpha B + \beta B', \quad C'' = \alpha C + \beta C', \quad D'' = \alpha D + \beta D'.$$

Quindi l'equazione $A''x + B''y + C''z = D''$ si riscrive

$$(\alpha A + \beta A')x + (\alpha B + \beta B')y + (\alpha C + \beta C')z = \alpha D + \beta D'$$

che equivale proprio alla (5.38).

□

Esempio 5.21. Data la retta

$$r\colon \begin{cases} x + y + z = 1 \\ x - y + 2z = 0 \end{cases}$$

vogliamo trovare l'equazione cartesiana del piano π che contiene r e passa per il punto P_0 di coordinate $(1,1,1)$.
Determiniamo prima tutti i piani che contengono r che, secondo la Proposizione 5.20, sono dati al variare di $\alpha, \beta \in \mathbb{R}$ dall'equazione

$$\alpha(x + y + z - 1) + \beta(x - y + 2z) = 0. \tag{5.39}$$

Poiché tra tutti questi piani cerchiamo quello che contiene il punto $(1,1,1)$, vogliamo che tale equazione sia soddisfatta quando poniamo $x = 1, y = 1, z = 1$. Sostituendo otteniamo

$$\alpha(1 + 1 + 1 - 1) + \beta(1 - 1 + 2) = 2\alpha + 2\beta = 0$$

da cui $\alpha = -\beta$. Sostituendo questa relazione nella (5.39), si ottiene

$$-\beta(x + y + z - 1) + \beta(x - y + 2z) = 0$$

ovvero, svolgendo i calcoli,

$$-2\beta y + \beta z + \beta = 0.$$

Al variare del parametro β, queste equazioni rappresentano tutte lo stesso piano (il piano π cercato) in quanto si tratta di equazioni proporzionali, tutte equivalenti. Dividendo per il parametro β (o, equivalentemente, scegliendo $\beta = 1$), possiamo allora scrivere che π ha equazione cartesiana

$$-2y + z + 1 = 0.$$

Se, invece del passaggio per il punto, volessimo imporre che il piano, oltre a contenere r, fosse parallelo a un altro piano, ad esempio quello di equazione cartesiana $x+2y+3z = -1$, dovremmo procedere come segue.

Come abbiamo visto nel capitolo precedente, due piani $Ax + By + Cz = D$ e $A'x + B'y + C'z = D'$ sono paralleli se e solo se le terne (A, B, C) e (A', B', C') sono proporzionali, in quanto rappresentano le coordinate di vettori normali (perpendicolari) ai piani. In realtà, poiché abbiamo la libertà di moltiplicare l'equazione di un piano per qualunque coefficiente senza che il piano venga modificato, possiamo sempre far sì che sia $(A, B, C) = (A', B', C')$. Allora, poiché svolgendo i calcoli nella (5.39), vediamo che il generico piano che contiene r è della forma

$$(\alpha + \beta)x + (\alpha - \beta)y + (\alpha + 2\beta)z - \alpha = 0, \tag{5.40}$$

la condizione di parallelismo tra questo piano e il piano di equazione $x + 2y + 3z = -1$ è

$$(\alpha + \beta, \alpha - \beta, \alpha + 2\beta) = (1, 2, 3)$$

ovvero

$$\begin{cases} \alpha + \beta = 1 \\ \alpha - \beta = 2 \\ \alpha + 2\beta = 3 \end{cases}$$

Per trovare il piano dato, basta quindi risolvere tale sistema e sostituire i valori di α e β trovati nella (5.40). In questo caso, si vede riducendo a gradini la sua matrice completa

$$\left(\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 1 & 2 & 3 \end{array}\right) \underset{\substack{R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1}}{\longrightarrow} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 1 & 2 \end{array}\right) \underset{R_3 \to 2R_3 + R_2}{\longrightarrow} \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & 5 \end{array}\right)$$

il sistema è incompatibile e quindi la condizione di parallelismo non può essere soddisfatta. Possiamo concludere che tra i piani che contengono la retta r, non ne esiste uno parallelo al piano dato.

Esempio 5.22. Date le due rette

$$r_1 \colon \begin{cases} x + y + z = 3 \\ x - 2y + z = 0 \end{cases} \qquad r_2 \colon \begin{cases} 2x + y - z = 2 \\ x - y - z = -1 \end{cases}$$

determiniamo, se esiste, il piano che le contiene.

Come abbiamo visto sopra, il generico piano che contiene r_1 ha equazione

$$\alpha_1(x + y + z - 3) + \beta_1(x - 2y + z) = 0$$

ovvero

$$(\alpha_1+\beta_1)x+(\alpha_1-2\beta_1)y+(\alpha_1+\beta_1)z-3\alpha_1=0. \tag{5.41}$$

Il generico piano che contiene r_2 ha invece equazione

$$\alpha_2(2x+y-z-2)+\beta_2(x-y-z+1)=0$$

ovvero

$$(2\alpha_2+\beta_2)x+(\alpha_2-\beta_2)y+(-\alpha_2-\beta_2)z+(-2\alpha_2+\beta_2)=0. \tag{5.42}$$

Per trovare, se esiste, il piano che contiene entrambe le rette basta controllare se esistono valori di $\alpha_1,\beta_1,\alpha_2,\beta_2$ tali che la (5.41) e la (5.42) coincidano. Uguagliando i coefficienti di x, y, z e il termine noto in tali equazioni si ottiene

$$\begin{cases}\alpha_1+\beta_1=2\alpha_2+\beta_2\\ \alpha_1-2\beta_1=\alpha_2-\beta_2\\ \alpha_1+\beta_1=-\alpha_2-\beta_2\\ -3\alpha_1=-2\alpha_2+\beta_2\end{cases}$$

ovvero il sistema omogeneo di 4 equazioni in 4 incognite

$$\begin{cases}\alpha_1+\beta_1-2\alpha_2-\beta_2=0\\ \alpha_1-2\beta_1-\alpha_2+\beta_2=0\\ \alpha_1+\beta_1+\alpha_2+\beta_2=0\\ -3\alpha_1+2\alpha_2-\beta_2=0\end{cases}$$

Tale sistema ha sicuramente sempre la soluzione $\alpha_1=\beta_1=\alpha_2=\beta_2=0$, ma se sostituissimo tali valori nelle (5.41) e la (5.42) otterremmo $0=0$, che non è l'equazione di un piano. Quindi per l'esistenza del piano che contiene entrambe le rette deve esistere una soluzione non nulla di tale sistema.
Riducendo a gradini la sua matrice dei coefficienti troviamo

$$\begin{pmatrix}1&1&-2&-1\\1&-2&-1&1\\1&1&1&1\\-3&0&2&-1\end{pmatrix}\xrightarrow[\substack{R_3\to R_3-R_1\\R_4\to R_4+3R_1}]{R_2\to R_2-R_1}\begin{pmatrix}1&1&-2&-1\\0&-3&1&2\\0&0&3&2\\0&3&-4&-4\end{pmatrix}$$

$$\xrightarrow[R_4\to R_4+R_2]{}\begin{pmatrix}1&1&-2&-1\\0&-3&1&2\\0&0&3&2\\0&0&-3&-2\end{pmatrix}\xrightarrow[R_4\to R_4+R_3]{}\begin{pmatrix}1&1&-2&-1\\0&-3&1&2\\0&0&3&2\\0&0&0&0\end{pmatrix}$$

Poichè la matrice ha rango 3 il sistema ha sicuramente altre soluzioni oltre alla 4-upla nulla $(0,0,0,0)$, quindi esiste il piano che contiene le rette.

Per trovare tale piano, basta risolvere il sistema, che abbiamo ridotto alla forma equivalente

$$\begin{cases} \alpha_1 + \beta_1 - 2\alpha_2 - \beta_2 = 0 \\ -3\beta_1 + \alpha_2 + 2\beta_2 = 0 \\ 3\alpha_2 + 2\beta_2 = 0 \end{cases}$$

Vediamo però che non è necessario determinare completamente la soluzione del sistema. Infatti l'ultima equazione non nulla ci dà il valore di α_2 (in funzione di β_2), mentre le prime due i valori di α_1 e β_1 (sempre in funzione di β_2). Quindi se sostituiamo i valori di α_1 e β_1 così trovati nella (5.41) o quello di α_2 nella (5.42) otteniamo lo stesso piano (il sistema esprime proprio la condizione che i due piani siano uguali). Dunque per trovare il piano basta determinare α_2 dall'ultima equazione non nulla $3\alpha_2 + 2\beta_2 = 0$ senza dover risolvere le altre due.

Questa equazione è equivalente a $\alpha_2 = -\frac{2}{3}\beta_2$ che sostituita nella (5.42) dà

$$(-\frac{4}{3}\beta_2 + \beta_2)x + (-\frac{2}{3}\beta_2 - \beta_2)y + (\frac{2}{3}\beta_2 - \beta_2)z + (\frac{4}{3}\beta_2 + \beta_2) = 0$$

ovvero

$$-\frac{1}{3}\beta_2 x - \frac{5}{3}\beta_2 y - \frac{1}{3}\beta_2 z + \frac{7}{3}\beta_2 = 0.$$

Dividendo per β_2 e moltiplicando per -3 (solamente per ottenere coefficienti interi) si ottiene infine

$$x + 5y + z - 7 = 0$$

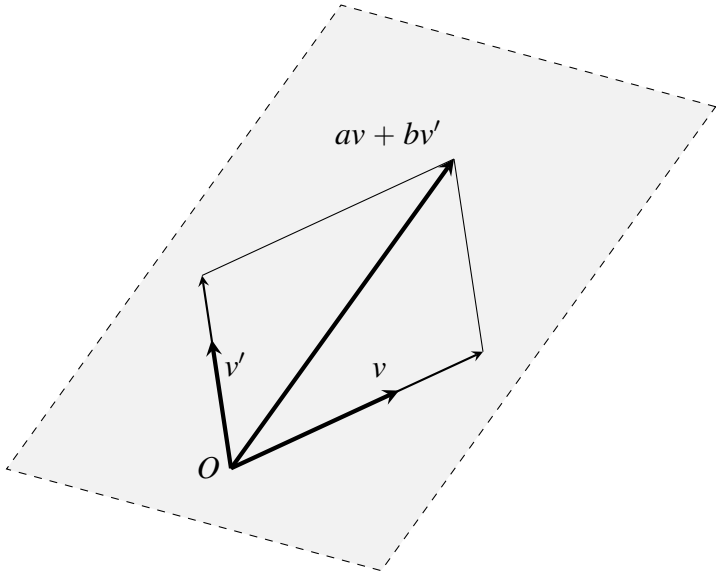
che è l'equazione del piano cercato.

Concludiamo questo capitolo con un ultimo esempio di applicazione geometrica delle tecniche apprese per risolvere i sistemi.

È possibile determinare se tre vettori applicati nello spazio tridimensionale v, v', v'' sono complanari?

La risposta è affermativa, e per vederlo si sfrutta proprio il procedimento di riduzione a gradini. Fissiamo una base e supponiamo che rispetto a questa i vettori abbiano coordinate $v \equiv (-1, 2, 1)$, $v' \equiv (-2, -1, 1)$ e $v'' \equiv (3, 4, -1)$.

Ora, due qualunque di questi vettori, ad esempio v e v', appartengono sicuramente a un piano, e un qualunque vettore che stia anch'esso su tale piano si scrive come combinazione $av + bv'$.



Quindi, il terzo vettore v'' appartiene allo stesso piano su cui stanno v e v' se e solo se esistono $a, b \in \mathbb{R}$ tali che $v'' = av + bv'$.
In coordinate, questo equivale a dire che

$$(3, 4, -1) = a(-1, 2, 1) + b(-2, -1, 1).$$

Questo significa che la terna $(3, 4, -1)$ si scrive come combinazione delle altre due e quindi è dipendente da esse. Ma allora, per capire se questo accade o no, basta scrivere la matrice che ha tali terne come righe

$$\begin{pmatrix} -1 & 2 & 1 \\ -2 & -1 & 1 \\ 3 & 4 & -1 \end{pmatrix}$$

ed effettuare una riduzione a gradini. Infatti, come sappiamo, una riga è dipendente dalle altre solo se essa si annulla in seguito alla riduzione a gradini. Calcoliamo quindi

$$\begin{pmatrix} -1 & 2 & 1 \\ -2 & -1 & 1 \\ 3 & 4 & -1 \end{pmatrix} \underset{\substack{R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 + 3R_1}}{\longrightarrow} \begin{pmatrix} -1 & 2 & 1 \\ 0 & -5 & -1 \\ 0 & 10 & 2 \end{pmatrix} \underset{R_3 \to R_3 + 2R_2}{\longrightarrow} \begin{pmatrix} -1 & 2 & 1 \\ 0 & -5 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Essendosi annullata la terza riga, questa era combinazione delle prime due: quindi i tre vettori rappresentati in coordinate dalle righe della matrice sono complanari.

Capitolo 6

Algebra matriciale

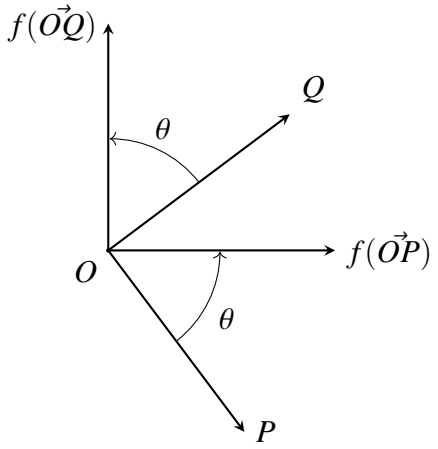
In questo capitolo vedremo che le matrici, oltre a rappresentare sistemi di equazioni lineari, formano una particolare classe di funzioni che comprende molte importanti trasformazioni geometriche. Questo ci consentirà di definire una operazione tra matrici, che chiameremo prodotto, e della quale studieremo le proprietà algebriche.

6.1 Matrici e trasformazioni

Mostriamo in questa sezione come alcune importanti trasformazioni geometriche possano essere rappresentate mediante una matrice.

Più precisamente, supponiamo di aver fissato nel piano una base ortonormale $\{\vec{OP}_1, \vec{OP}_2\}$. Come abbiamo visto nel Capitolo 4, a ogni punto P del piano risulta associata allora una coppia (x_1, x_2) di numeri reali (le sue coordinate rispetto al sistema di riferimento fissato) definiti dalla condizione che $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2$.

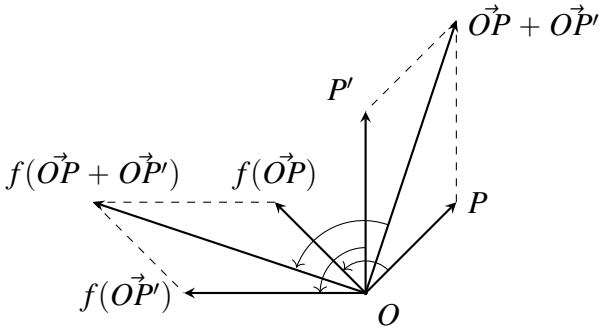
Supponiamo ora di voler applicare ai punti del piano la trasformazione f che consiste in una rotazione di un certo angolo fissato θ attorno all'origine O. Dal momento che ruotare il punto P attorno a O equivale a ruotare il vettore $\vec{OP}$ dello stesso angolo, possiamo considerare questa rotazione come una funzione sui vettori anziché sui punti. Possiamo quindi descrivere la funzione f che va dall'insieme dei vettori applicati in se stesso e che associa a ogni vettore $\vec{OP}$ il vettore $f(\vec{OP})$ che si ottiene ruotando $\vec{OP}$ dell'angolo θ.



Ora ci poniamo l'obiettivo di rappresentare tale trasformazione in coordinate: se (x_1, x_2) sono le coordinate di $\vec{OP}$, vogliamo determinare le coordinate di $f(\vec{OP})$. A questo scopo, osserviamo preliminarmente che la rotazione soddisfa le seguenti due proprietà.

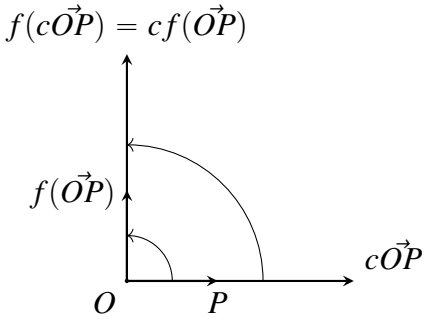
(1) Siano $\vec{OP}$ e $\vec{OP'}$ due vettori qualunque, Allora sommarli e successivamente ruotare il vettore risultante è equivalente a ruotarli prima e successivamente sommare i vettori così ottenuti. In formule

$$f(\vec{OP} + \vec{OP'}) = f(\vec{OP}) + f(\vec{OP'}) \tag{6.1}$$



(2) Sia $\vec{OP}$ un qualunque vettore e $c \in \mathbb{R}$. Allora moltiplicare il vettore $\vec{OP}$ per c e successivamente ruotarlo equivale a ruotarlo prima ed in seguito moltiplicarlo per c. In formule

$$f(c\vec{OP}) = cf(\vec{OP}) \tag{6.2}$$



Le due proprietà appena osservate ci permettono di rappresentare la rotazione in coordinate. Infatti, se P ha coordinate (x_1, x_2), per definizione si ha $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2$. Ma allora per il vettore ruotato $f(\vec{OP})$ si ha

$$f(\vec{OP}) = f(x_1\vec{OP}_1 + x_2\vec{OP}_2) =$$

(applicando la proprietà (6.1))

$$= f(x_1\vec{OP}_1) + f(x_2\vec{OP}_2) =$$

(applicando la proprietà (6.2) a entrambi gli addendi)

$$f(\vec{OP}) = x_1 f(\vec{OP}_1) + x_2 f(\vec{OP}_2). \tag{6.3}$$

Ora, essendo $f(\vec{OP}_1)$ e $f(\vec{OP}_2)$ vettori del piano, essi potranno essere scritti come combinazione dei vettori $\vec{OP}_1$ e $\vec{OP}_2$ della base fissata. Cioè si hanno le uguaglianze

$$f(\vec{OP}_1) = a_{11}\vec{OP}_1 + a_{21}\vec{OP}_2 \tag{6.4}$$

$$f(\vec{OP}_2) = a_{12}\vec{OP}_1 + a_{22}\vec{OP}_2 \tag{6.5}$$

per certi numeri reali $a_{11}, a_{21}, a_{12}, a_{22}$. Ma allora, sostituendo (6.4) e (6.5) nella (6.3), si ottiene

$$f(\vec{OP}) = x_1(a_{11}\vec{OP}_1 + a_{21}\vec{OP}_2) + x_2(a_{12}\vec{OP}_1 + a_{22}\vec{OP}_2)$$

e, raccogliendo i coefficienti di $\vec{OP}_1$ e $\vec{OP}_2$ secondo le proprietà (4.1)-(4.8), si ottiene

$$f(\vec{OP}) = (a_{11}x_1 + a_{12}x_2)\vec{OP}_1 + (a_{21}x_1 + a_{22}x_2)\vec{OP}_2. \tag{6.6}$$

Da quest'ultima uguaglianza, vediamo che le coordinate (sempre rispetto alla base fissata $\vec{OP}_1, \vec{OP}_2$) del vettore $f(\vec{OP})$ ottenuto ruotando il vettore $\vec{OP}$ di coordinate (x_1, x_2) sono date dalla coppia $(a_{11}x_1 + a_{12}x_2, a_{21}x_1 + a_{22}x_2)$.
Quindi conoscere la rotazione in coordinate è equivalente a conoscere i coefficienti

$a_{11}, a_{21}, a_{12}, a_{22}$. Disponendo tali numeri nella matrice

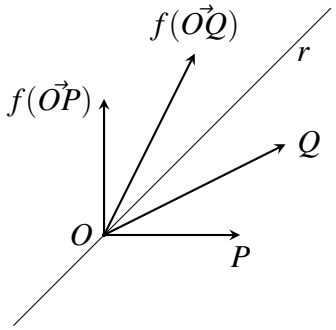
$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

possiamo dire che questa rappresenta la rotazione, nel senso che la funzione

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} \tag{6.7}$$

che associa a ogni coppia[1] (x_1, x_2) la coppia $(a_{11}x_1 + a_{12}x_2, a_{21}x_1 + a_{22}x_2)$ fornisce le coordinate di un vettore ruotato a partire dalle coordinate del vettore di partenza.

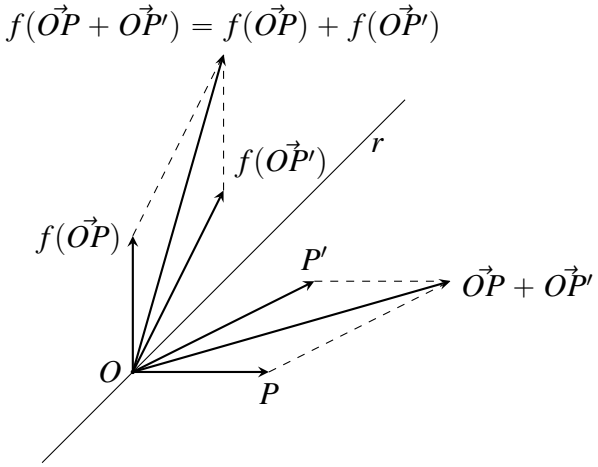
Lo stesso ragionamento fatto per arrivare alla matrice che rappresenta una rotazione nel piano può essere ripetuto anche per altre importanti trasformazioni geometriche che soddisfano le proprietà (6.1) e (6.2). Ad esempio la riflessione rispetto a una retta per O, che manda ogni vettore $\vec{OP}$ nel vettore simmetrico rispetto alla retta, come nel seguente disegno



ha esattamente le stesse proprietà (6.1) e (6.2) viste nel caso della rotazione. Infatti

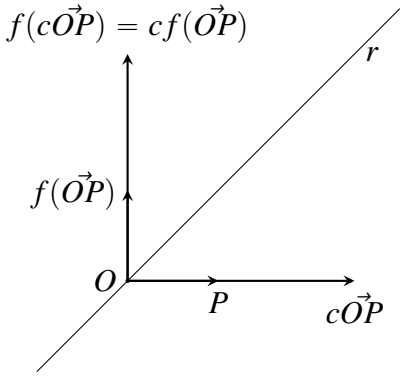
(1) dati due vettori $\vec{OP}$ e $\vec{OP'}$, sommarli e poi riflettere il vettore risultante oppure prima rifletterli e poi sommare i vettori riflessi è equivalente

[1] Qui e nel seguito indicheremo spesso le n-uple in colonna anziché in riga. Questo ci è comodo per questioni di spazio ma, sopratutto, è giustificato dall'utilizzo del prodotto di matrici che vedremo nel seguito del capitolo.



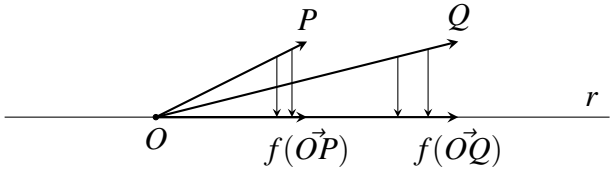
(cioè $f(\vec{OP} + \vec{OP'}) = f(\vec{OP}) + f(\vec{OP'})$).

(2) dato un vettore $\vec{OP}$ e un numero reale c, moltiplicare il vettore per c e poi rifletterlo oppure prima rifletterlo e poi moltiplicarlo per c è equivalente

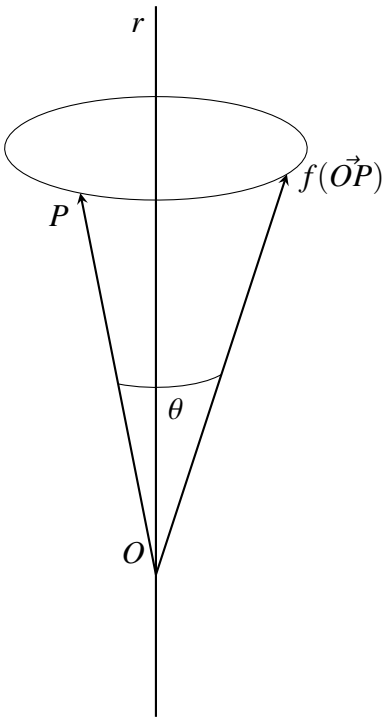


(cioè $f(c\vec{OP}) = cf(\vec{OP})$).

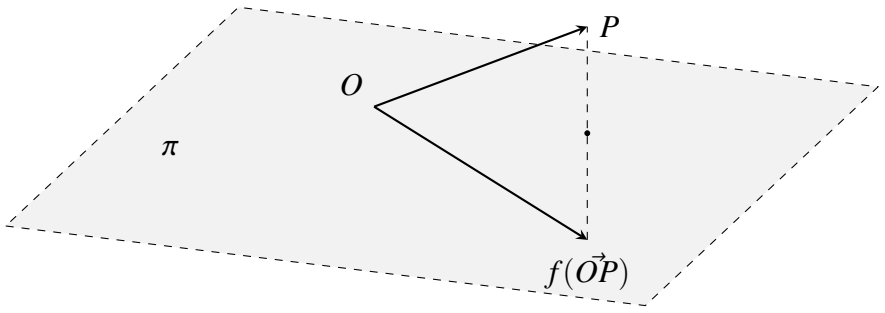
Quindi possiamo ripetere esattamente gli stessi passaggi visti sopra per la rotazione e rappresentare anche la riflessione mediante una matrice con due righe e due colonne.
Come terzo esempio di importante trasformazione geometrica che soddisfa le proprietà (6.1) e (6.2), e quindi può essere rappresentata da una matrice, citiamo la proiezione su una retta fissata passante per O.



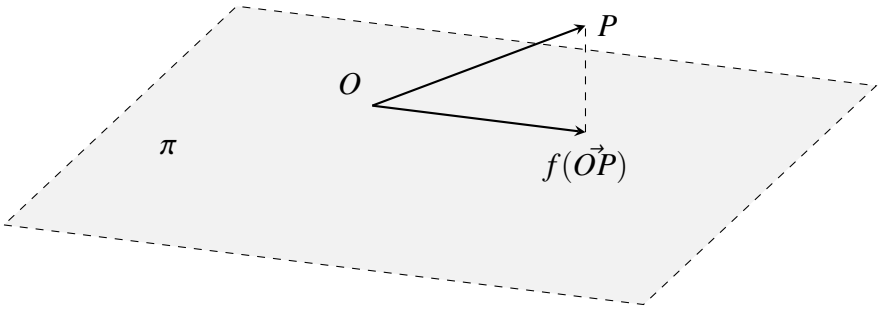
Quanto visto fino ad ora sul piano risulta vero anche per alcune importanti trasformazioni geometriche dello spazio tridimensionale. Ad esempio, non è difficile vedere che le proprietà (6.1) e (6.2) sono soddisfatte anche dalla rotazione di un angolo fissato attorno a una retta fissata passante per O (detta *asse della rotazione*)



dalla riflessione rispetto a un piano passante per O



e dalla proiezione ortogonale su un piano passante per l'origine O.



In tutti e tre questi esempi, essendo soddisfatte le proprietà $f(\vec{OP} + \vec{OP'}) = f(\vec{OP}) + f(\vec{OP'})$ e $f(c\vec{OP}) = cf(\vec{OP})$, possiamo ripetere gli stessi passaggi visti nel caso della rotazione nel piano, con la differenza che stavolta avremo una base $\vec{OP}_1, \vec{OP}_2, \vec{OP}_3$ formata da tre vettori. Per ogni vettore $\vec{OP}$ avremo $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2 + x_3\vec{OP}_3$ e, con ragionamenti simili a quelli visti sopra per ottenere la (6.6), si otterrà

$$\begin{aligned} f(\vec{OP}) = {} & (a_{11}x_1 + a_{12}x_2 + a_{13}x_3)\vec{OP}_1 \\ & + (a_{21}x_1 + a_{22}x_2 + a_{23}x_3)\vec{OP}_2 \\ & + (a_{31}x_1 + a_{32}x_2 + a_{33}x_3)\vec{OP}_3. \end{aligned}$$

Ovvero le coordinate (rispetto alla base fissata $\vec{OP}_1, \vec{OP}_2, \vec{OP}_3$) del vettore $f(\vec{OP})$ ottenuto trasformando il vettore $\vec{OP}$ di coordinate (x_1, x_2, x_3) sono date dalla terna

$$(a_{11}x_1 + a_{12}x_2 + a_{13}x_3, a_{21}x_1 + a_{22}x_2 + a_{23}x_3, a_{31}x_1 + a_{32}x_2 + a_{33}x_3). \tag{6.8}$$

Quindi la trasformazione data si traduce in coordinate in una funzione del tipo

$$f : \mathbb{R}^3 \to \mathbb{R}^3, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{pmatrix} \tag{6.9}$$

In questo caso, conoscere la trasformazione considerata in coordinate equivale a conoscere la matrice

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

con tre righe e tre colonne che contiene i coefficienti che appaiono nella (6.8).

Definizione 6.1. Dati due spazi vettoriali V e W (si veda l'Osservazione 4.1), una funzione $f : V \to W$ si dice **funzione lineare** (o **applicazione lineare**) se soddisfa le proprietà

(1) $f(v + v') = f(v) + f(v')$ per qualunque $v, v' \in V$

(2) $f(cv) = cf(v)$ per ogni $v \in V$ e ogni scalare c

Le formule (6.7) e (6.9), che ci danno la rappresentazione in coordinate delle trasformazioni geometriche viste sopra, sono casi particolari di funzioni del tipo

$$f : \mathbb{K}^n \to \mathbb{K}^m, \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} \tag{6.10}$$

(dove $\mathbb{K}$ è un campo, per esempio $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$) che risultano più in generale determinate dalla matrice con m righe e n colonne

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Chiameremo l'applicazione (6.10) associata a tale matrice l'**applicazione lineare determinata da** A e la denoteremo con L_A.
Nel paragrafo successivo concentreremo la nostra attenzione su tali funzioni. Usando la nozione di composizione di funzioni vista nel Capitolo 2 definiremo un'operazione di prodotto tra matrici. Potremo quindi considerare l'insieme delle matrici, analogamente a quanto fatto per gli insiemi numerici, come un insieme sul quale è definita un'operazione della quale studiare le proprietà. In particolare vedremo che certi insiemi di matrici sono anelli.
Introduciamo alcune notazioni preliminari.

Definizione 6.2. Una matrice con m righe e n colonne è detta **matrice** m **per** n. L'insieme delle matrici m per n le cui entrate appartengono a un campo dato $\mathbb{K}$ (ad esempio il campo $\mathbb{R}$ dei numeri reali, o quello $\mathbb{C}$ dei numeri complessi) si denota $M_{m,n}(\mathbb{K})$.
Nel caso in cui $m = n$, ovvero il numero di righe sia uguale al numero di colonne, si usa scrivere $M_n(\mathbb{K})$ al posto di $M_{n,n}(\mathbb{K})$.
Un elemento di $M_n(\mathbb{K})$ si dice **matrice quadrata di ordine** n.

6.2 Composizione e prodotto di matrici

In questo paragrafo vogliamo calcolare la matrice che rappresenta la composizione di due funzioni del tipo (6.10) visto alla fine del paragrafo precedente. Più precisamente,

consideriamo le applicazioni lineari $f = L_A : \mathbb{K}^n \to \mathbb{K}^m$ definita da

$$f\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} \tag{6.11}$$

e $g = L_B : \mathbb{K}^p \to \mathbb{K}^n$ data da

$$g\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} = \begin{pmatrix} b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p \\ b_{21}y_1 + b_{22}y_2 + \cdots + b_{2p}y_p \\ \vdots \\ b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p \end{pmatrix} \tag{6.12}$$

In base a quanto visto nella Sezione 2.6, la composizione $f \circ g$ può essere calcolata in quanto il codominio di g, ovvero $\mathbb{K}^n$, è anche il dominio di f. Si ha

$$(f \circ g)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} = f\left(g\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}\right) = f\begin{pmatrix} b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p \\ b_{21}y_1 + b_{22}y_2 + \cdots + b_{2p}y_p \\ \vdots \\ b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11}(b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p) + \cdots + a_{1n}(b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p) \\ a_{21}(b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p) + \cdots + a_{2n}(b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p) \\ \vdots \\ a_{m1}(b_{11}y_1 + b_{12}y_2 + \cdots + b_{1p}y_p) + \cdots + a_{mn}(b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{np}y_p) \end{pmatrix} =$$

(raccogliendo $y_1, y_2, \cdots, y_p$)

$$= \begin{pmatrix} (a_{11}b_{11} + \cdots + a_{1n}b_{n1})y_1 + \cdots + (a_{11}b_{1p} + \cdots + a_{1n}b_{np})y_p \\ (a_{21}b_{11} + \cdots + a_{2n}b_{n1})y_1 + \cdots + (a_{21}b_{1p} + \cdots + a_{2n}b_{np})y_p \\ \vdots \\ (a_{m1}b_{11} + \cdots + a_{mn}b_{n1})y_1 + \cdots + (a_{m1}b_{1n} + \cdots + a_{mn}b_{np})y_p \end{pmatrix}$$

Da quest'ultima espressione vediamo che la composizione $f \circ g$ è ancora una funzione del tipo (6.10), cioè è determinata da una matrice C, e più precisamente

$$C = \begin{pmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{n1} & a_{11}b_{12} + \cdots + a_{1n}b_{n2} & \cdots & a_{11}b_{1p} + \cdots + a_{1n}b_{np} \\ a_{21}b_{11} + \cdots + a_{2n}b_{n1} & a_{21}b_{12} + \cdots + a_{2n}b_{n2} & \cdots & a_{21}b_{1p} + \cdots + a_{2n}b_{np} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{11} + \cdots + a_{mn}b_{n1} & a_{m1}b_{12} + \cdots + a_{mn}b_{n2} & \cdots & a_{m1}b_{1n} + \cdots + a_{mn}b_{np} \end{pmatrix}$$

Ora, notiamo che le entrate c_{ij} di tale matrice sono tutte espressioni del tipo

$$c_{ij} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj} \tag{6.13}$$

in cui il primo indice dell'entrata di A è fisso (sempre uguale a i) e il secondo indice dell'entrata di B è fisso (sempre uguale a j). Invece gli indici "interni" (il secondo dell'entrata di A e il primo dell'entrata di B) variano da 1 a n. Usando la notazione di sommatoria, potremmo quindi scrivere $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$. Possiamo quindi definire il prodotto di due matrici A e B come la matrice $C = AB$ che rappresenta la composizione $L_A \circ L_B$ delle funzioni associate ad A e B.

Definizione 6.3. La matrice C le cui entrate sono date dalla (6.13) è detta **prodotto di** A **per** B, e si scrive $C = AB$.

Ricordiamo che la composizione $L_A \circ L_B$ delle due funzioni L_A e L_B può essere fatta solo sotto opportune condizioni (il codominio di $L_B : \mathbb{K}^p \to \mathbb{K}^n$ deve essere uguale al dominio di $L_A : \mathbb{K}^n \to \mathbb{K}^m$). Di conseguenza anche il prodotto di due matrici può essere fatto solo sotto opportune condizioni. Più precisamente, dal momento che la matrice A di L_A ha m righe e n colonne, mentre la matrice B di L_B ha n righe e p colonne, si possono moltiplicare tra loro due matrici A e B (in quest'ordine) se e solo se *il numero di colonne di A è uguale al numero di righe di B*. In tal caso, il risultato AB è una matrice con m righe e p colonne:

$$A \in M_{m,n}(\mathbb{K}),\ B \in M_{n,p}(\mathbb{K}) \Rightarrow AB \in M_{m,p}(\mathbb{K}).$$

Il prodotto di matrici dato dalla Definizione 6.3 si chiama anche **prodotto righe per colonne**. Il motivo è che nell'espressione (6.13) della generica entrata di posto $i\ j$ appaiono tutte e sole le entrate con primo indice i di A e tutte e sole le entrate con secondo indice j di B. Dal momento che il primo è l'indice di riga e il secondo quello di colonna, l'entrata di posto $i\ j$ di AB è la somma dei prodotti entrata per entrata della

i-esima riga di A per la j-esima colonna di B.

$$\begin{pmatrix} & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ & & & \end{pmatrix}\begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix} = \begin{pmatrix} & \cdots & & \\ \cdots & a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} & \cdots \\ & \cdots & & \end{pmatrix}$$

Esempio 6.4. Consideriamo le due matrici

$$A = \begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 3 & 6 \end{pmatrix} \in M_{3,2}(\mathbb{R}), \quad B = \begin{pmatrix} 5 & 2 & 0 & 1 \\ 3 & 1 & -1 & 2 \end{pmatrix} \in M_{2,4}(\mathbb{R}).$$

In base a quello che abbiamo detto, il prodotto di A per B è ben definito e il risultato AB sarà una matrice 3 per 4.
Per trovare l'entrata di AB che sta *nella prima riga* e *prima colonna* si considerano *la prima riga di* A $\begin{pmatrix} 1 & 2 \end{pmatrix}$, la *prima colonna* di B $\begin{pmatrix} 5 \\ 3 \end{pmatrix}$, se ne moltiplicano gli elementi corrispondenti (il primo con il primo, il secondo con il secondo) e si sommano i risultati: $1 \cdot 5 + 2 \cdot 3 = 5 + 6 = 11$. Quindi l'entrata 1 1 di AB sarà 11.
Analogamente, per trovare l'entrata di AB che sta *nella prima riga* e *seconda colonna* si considerano *la prima riga di* A $\begin{pmatrix} 1 & 2 \end{pmatrix}$, la *seconda colonna* di B $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, se ne moltiplicano gli elementi corrispondenti e si sommano i risultati: $1 \cdot 2 + 2 \cdot 1 = 2 + 2 = 4$. Quindi l'entrata 1 2 di AB sarà 4.
Svolgendo questo calcolo per tutte le entrate, si vede che

$$AB = \begin{pmatrix} 11 & 4 & -2 & 5 \\ 35 & 13 & -5 & 14 \\ 33 & 12 & -6 & 15 \end{pmatrix}.$$

Vediamo ora le proprietà fondamentali del prodotto di matrici.
In primo luogo, è importante osservare che tale prodotto *non gode della proprietà commutativa*. Date due matrici A e B e ammesso che si possano eseguire entrambi i prodotti AB e BA, in generale si avrà $AB \neq BA$.[2]

Esempio 6.5. Siano $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ e $B = \begin{pmatrix} -1 & 0 \\ 5 & 2 \end{pmatrix}$. Allora, per definizione di prodotto righe per colonne vediamo che

$$AB = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 5 & 2 \end{pmatrix} = \begin{pmatrix} 9 & 4 \\ 17 & 8 \end{pmatrix}$$

[2]Questo fatto può essere spiegato ricordando che il prodotto tra matrici rappresenta la composizione delle funzioni da loro determinate, e come sappiamo in generale la composizione non gode della proprietà commutativa (si veda l'Esempio 2.37).

$$BA = \begin{pmatrix} -1 & 0 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} -1 & -2 \\ 11 & 18 \end{pmatrix}$$

ovvero $AB \neq BA$.

Si può dimostrare invece che il prodotto di matrici soddisfa la proprietà associativa, ovvero $(AB)C = A(BC)$. Questa è una conseguenza della stessa proprietà per la composizione di funzioni.

Esempio 6.6. Siano $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$ e $C = \begin{pmatrix} 1 & 5 \\ 2 & 0 \end{pmatrix}$. Allora

$$(AB)C = \left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \right) \begin{pmatrix} 1 & 5 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 3 \\ 2 & 7 \end{pmatrix} \begin{pmatrix} 1 & 5 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 16 & 10 \end{pmatrix}$$

$$A(BC) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \left(\begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 5 \\ 2 & 0 \end{pmatrix} \right) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 4 & 10 \\ 1 & -5 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 16 & 10 \end{pmatrix}$$

Ci chiediamo ora se il prodotto di matrici ammetta un elemento neutro come nel caso del prodotto tra numeri reali, in cui esiste il numero 1 per cui $a \cdot 1 = 1 \cdot a = a$ per ogni a. La risposta è affermativa. Più precisamente, per ogni n consideriamo la seguente matrice con n righe e n colonne

$$Id_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{6.14}$$

ovvero la matrice che le cui entrate con lo stesso indice di riga e di colonna (a_{11}, a_{22} etc.) sono 1 e tutte le altre entrate sono 0.

Tale matrice si chiama **matrice identica** (o **matrice identità**) **di ordine** n e si denota con Id_n. Ad esempio,

$$Id_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad Id_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Osservazione 6.7. Data una matrice quadrata A di ordine n, le sue entrate $a_{11}, a_{22}, \cdots, a_{nn}$ che hanno stesso indice di riga e di colonna formano la cosiddetta **diagonale della matrice**. La matrice identica Id_n può essere quindi descritta come la matrice che ha 1 sulla diagonale e 0 nelle altre entrate. Le entrate di Id_n si denotano solitamente con il simbolo

δ_{ij}, detto **delta di Kronecker**, definito da

$$\delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}.$$

Ora, si può verificare che, per ogni $A \in M_{m,n}(\mathbb{K})$ si ha

$$AId_n = A, \quad Id_m A = A$$

(l'ordine della matrice identica cambia perché deve essere tale che si possa svolgere il prodotto). Quindi la matrice identica è l'elemento neutro per il prodotto righe per colonne.

Esempio 6.8.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

Osservazione 6.9. Si noti che la matrice identica Id_n rappresenta esattamente la funzione identica $id_{\mathbb{K}^n} : \mathbb{K}^n \to \mathbb{K}^n$ che manda ogni elemento in se stesso. Infatti

$$id_{\mathbb{K}^n} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1x_1 + 0x_2 + \cdots + 0x_n \\ 0x_1 + 1x_2 + \cdots + 0x_n \\ \vdots \\ 0x_1 + 0x_2 + \cdots + 1x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{6.15}$$

Questo spiega perché tale matrice sia l'elemento neutro per il prodotto. Infatti il prodotto tra matrici rappresenta la composizione delle applicazioni corrispondenti e la funzione identica è l'elemento neutro per la composizione, come visto a pagina 50.

Ora, ispirandoci a quello che succede in un campo numerico, dove ogni elemento a diverso da zero ha un inverso rispetto al prodotto (ovvero un elemento b tale che $ab = ba = 1$) e dalla composizione di funzioni diamo la seguente

Definizione 6.10. Una matrice A è detta **invertibile** se esiste una matrice B tale che $AB = Id$ e $BA = Id$, dove Id è una matrice identica di ordine opportuno, dipendente da A. In tal caso la matrice B sarà detta **inversa di** A e verrà denotata con A^{-1}.

Si ha il seguente risultato che dimostreremo in seguito.

Teorema 6.11. *Una matrice $A \in M_{m,n}(\mathbb{K})$ è invertibile se e solo se $m = n$ e il rango di A è uguale a n.*

In altre parole, sono invertibili solo le matrici quadrate $A \in M_n(\mathbb{K})$ e di rango massimo, ovvero in cui nessuna riga è dipendente dalle altre. Sotto tale condizione, affermiamo che esiste una matrice $B \in M_n(\mathbb{K})$ (cioè quadrata dello stesso ordine) per cui $AB = Id_n$ e $BA = Id_n$.

Ora, per capire invece quale sia il ruolo giocato dal rango, facciamo alcune considerazioni che ci suggeriranno anche un modo per calcolare l'inversa quando essa esiste.

Concentriamoci sulla condizione $AB = Id_n$, che equivale a

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ & & \cdots & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Tenuto conto della definizione di prodotto righe per colonne, vediamo che moltiplicando le righe di A per la prima colonna di B deve valere

$$\begin{cases} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} = 1 \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} = 0 \\ \cdots \\ a_{n1}b_{11} + a_{n2}b_{21} + \cdots + a_{nn}b_{n1} = 0 \end{cases}$$

ovvero la prima colonna di B soddisfa il sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{cases} \tag{6.16}$$

con matrice dei coefficienti uguale a A e termini noti uguali alla prima colonna della matrice identica.

Analogamente, moltiplicando le righe di A per la seconda colonna di B si vede che devono essere soddisfatte le seguenti uguaglianze

$$\begin{cases} a_{11}b_{12} + a_{12}b_{22} + \cdots + a_{1n}b_{n2} = 0 \\ a_{21}b_{12} + a_{22}b_{22} + \cdots + a_{2n}b_{n2} = 1 \\ \cdots \\ a_{n1}b_{12} + a_{n2}b_{22} + \cdots + a_{nn}b_{n2} = 0 \end{cases}$$

ovvero la seconda colonna di B deve soddisfare il sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 1 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0 \end{cases} \tag{6.17}$$

la cui matrice dei coefficienti è ancora A ma con termini noti uguali alla seconda colonna della matrice identica. Possiamo ragionare allo stesso modo fino all'ultima colonna di B che dovrà soddisfare

$$\begin{cases} a_{11}b_{1n} + a_{12}b_{2n} + \cdots + a_{1n}b_{nn} = 0 \\ a_{21}b_{1n} + a_{22}b_{2n} + \cdots + a_{2n}b_{nn} = 0 \\ \cdots \\ a_{n1}b_{1n} + a_{n2}b_{2n} + \cdots + a_{nn}b_{nn} = 1 \end{cases}$$

ovvero essere soluzione del sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 1 \end{cases} \tag{6.18}$$

Riassumendo, si ha $AB = Id_n$ se e solo se le colonne

$$\begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{pmatrix}, \quad \begin{pmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{pmatrix}, \quad \cdots, \quad \begin{pmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{nn} \end{pmatrix}$$

sono soluzione rispettivamente degli n sistemi (6.16), (6.17), ... , (6.18).
Ora, come sappiamo, per verificare se un sistema ha soluzioni basta scrivere la sua matrice completa e ridurla a gradini col metodo di Gauss-Jordan.
Poiché i sistemi (6.16), (6.17), ... , (6.18) hanno tutti la stessa matrice dei coefficienti, cioè A, e differiscono solo per i termini noti, le operazioni elementari che eseguiremo per ridurli sono le stesse, quindi conviene svolgerle una sola volta. A questo scopo, basta scrivere la matrice

$$(A|Id_n) = \left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ & & \cdots & & & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{array}\right) \tag{6.19}$$

ottenuta affiancando tutti i termini noti dei sistemi (6.16), (6.17), ... , (6.18) e risolverli contemporaneamente con una sola riduzione.

Se la matrice A avesse rango minore di n, in seguito alla riduzione si annullerebbe almeno una sua riga. In tal caso, affinché gli n sistemi rappresentati dalla (6.19) siano tutti compatibili, dovrebbe annullarsi almeno una riga del blocco Id_n sulla destra che rappresenta i termini noti (altrimenti in uno dei sistemi corrispondenti ci sarebbe un'uguaglianza falsa $0 = b \neq 0$).

Tuttavia, questo non può mai accadere. Se in seguito alle stesse operazioni elementari applicate per ridurre A si annullasse anche una riga del blocco di destra, allora le righe di Id_n sarebbero linearmente dipendenti. Ma questo è falso come si evince facilmente dal fatto che Id_n è una matrice a gradini senza righe nulle (ovvero di rango massimo).

Quindi, se la matrice di A avesse rango minore di n in seguito alla riduzione in contemporanea degli n sistemi (6.16), (6.17), ... , (6.18), comparirebbe necessariamente un'uguaglianza falsa $0 = b \neq 0$. Di conseguenza almeno uno di tali sistemi sarebbe incompatibile e l'uguaglianza $AB = Id_n$ non potrebbe essere verificata.

Questo spiega perché se A non ha rango massimo allora essa sicuramente non è invertibile.

Viceversa, se A ha rango massimo allora riducendo la matrice (6.19) nella parte dei coefficienti non si annulla nessuna riga e tutti i sistemi (6.16), (6.17), ... , (6.18) hanno soluzione. Esiste quindi una matrice B tale che $AB = Id_n$ le cui colonne sono date dalle soluzioni dei sistemi. Si può a questo punto mostrare (ma omettiamo i dettagli) che la matrice B così trovata soddisfa anche l'uguaglianza inversa $BA = Id_n$, ovvero A è invertibile con inversa uguale a B.

Ispirandoci a quanto appena detto, verifichiamo che la matrice $A = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 2 \end{smallmatrix}\right)$ è invertibile e calcoliamone l'inversa. Come fatto sopra, affianchiamo a tale matrice la matrice identica dello stesso ordine

$$(A|Id_2) = \left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right)$$

che rappresenta i due sistemi le cui soluzioni sono le colonne della matrice inversa. Applichiamo il procedimento di riduzione a gradini: in questo caso basta il singolo passaggio

$$(A|Id_n) = \left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right) \underset{R_2 \to R_2 - R_1}{\longrightarrow} \left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 3 & -1 & 1 \end{array}\right) \tag{6.20}$$

Notiamo che, dopo la riduzione a gradini, A non ha righe nulle e quindi, come abbiamo detto sopra, A è invertibile. La prima colonna dell'inversa B di A è data dalla soluzione del sistema ridotto

$$\begin{cases} x_1 - x_2 = 1 \\ 3x_2 = -1 \end{cases} \tag{6.21}$$

Risolvendo dal basso otteniamo la coppia $\left(\frac{2}{3}, -\frac{1}{3}\right)$, che è quindi la prima colonna della matrice inversa. Analogamente, la seconda colonna dell'inversa B di A è data dalla soluzione del sistema ridotto

$$\begin{cases} x_1 - x_2 = 0 \\ 3x_2 = 1 \end{cases} \tag{6.22}$$

Questa, come si vede risolvendo dal basso, è la coppia $\left(\frac{1}{3}, \frac{1}{3}\right)$, che è quindi la seconda colonna della matrice inversa.

In conclusione, l'inversa della matrice A è

$$A^{-1} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

Infatti, si verifica subito con un calcolo che

$$\begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2/3 & 1/3 \\ -1/3 & 1/3 \end{pmatrix} = \begin{pmatrix} 2/3 & 1/3 \\ -1/3 & 1/3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

concordemente con la definizione di inversa.

Per evitare di scrivere e risolvere separatamente i sistemi (6.21) e (6.22) e trovare invece in modo più diretto la matrice inversa, si può procedere come segue. Dopo aver effettuato la riduzione a gradini in (6.20), si applicano ulteriori operazioni elementari fino a trasformare la matrice A del blocco di sinistra nella matrice identica. A questo punto nel blocco di destra si leggerà direttamente l'inversa. Per vederlo, riprendiamo da (6.20) e facciamo comparire prima uno zero in posizione 1 2 eseguendo

$$\left(\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 3 & -1 & 1 \end{array}\right) \xrightarrow[R_1 \to 3R_1 + R_2]{} \left(\begin{array}{cc|cc} 3 & \mathbf{0} & 2 & 1 \\ 0 & 3 & -1 & 1 \end{array}\right)$$

Applichiamo poi a ogni riga l'operazione elementare del secondo tipo che consiste nel dividerla per l'elemento che si trova sulla diagonale:

$$\left(\begin{array}{cc|cc} 3 & 0 & 2 & 1 \\ 0 & 3 & -1 & 1 \end{array}\right) \xrightarrow[\substack{R_1 \to (1/3)R_1 \\ R_2 \to (1/3)R_2}]{} \left(\begin{array}{cc|cc} 1 & 0 & 2/3 & 1/3 \\ 0 & 1 & -1/3 & 1/3 \end{array}\right). \tag{6.23}$$

Come si vede la matrice identica che avevamo affiancato ad A si è trasformata nella matrice inversa di A già trovata sopra.

Per capire perché, ricordiamo che le operazioni elementari che stiamo eseguendo servono a risolvere contemporaneamente i sistemi che poi ci danno come soluzione le colonne della matrice inversa. Ma allora, riducendo la matrice A (cioè la matrice dei coefficienti di tali sistemi) alla matrice identica come in (6.23) non stiamo facendo altro che ridurre

i due sistemi alla forma

$$\begin{cases} x_1 = \frac{2}{3} \\ x_2 = -\frac{1}{3} \end{cases}, \quad \begin{cases} x_1 = \frac{1}{3} \\ x_2 = \frac{1}{3} \end{cases}$$

Abbiamo quindi fatto comparire direttamente nel blocco di destra le soluzioni cercate (che sono proprio le colonne della matrice inversa).

Esempio 6.12. Vogliamo trovare ora l'inversa della matrice

$$A = \begin{pmatrix} 1 & 1 & 2 \\ -1 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix}.$$

Iniziamo con il trasformare la matrice $(A|Id_n)$ in una matrice a gradini

$$(A|Id_n) = \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \underset{\substack{R_2 \to R_2 + R_1 \\ R_3 \to R_3 - 2R_1}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 1 & 1 & 0 \\ 0 & -1 & -3 & -2 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 1 & 1 & 0 \\ 0 & -1 & -3 & -2 & 0 & 1 \end{array}\right) \underset{R_3 \to 2R_3 + R_2}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 1 & 1 & 0 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array}\right)$$

Il fatto che non si sia annullata nessuna riga nel blocco di sinistra ci dice che la matrice A è invertibile.

Ora, come spiegato sopra, facciamo comparire zeri sopra la diagonale, effettuando una sorta di riduzione a gradini "inversa", dal basso verso l'altro e da destra verso sinistra.

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 1 & 1 & 0 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array}\right) \underset{\substack{R_2 \to 2R_2 + R_3 \\ R_1 \to 2R_1 + R_3}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 2 & 2 & \mathbf{0} & -1 & 1 & 2 \\ 0 & 4 & \mathbf{0} & -1 & 3 & 2 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 2 & 2 & 0 & -1 & 1 & 2 \\ 0 & 4 & 0 & -1 & 3 & 2 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array}\right) \underset{R_1 \to 2R_1 - R_2}{\longrightarrow} \left(\begin{array}{ccc|ccc} 4 & \mathbf{0} & 0 & -1 & -1 & 2 \\ 0 & 4 & 0 & -1 & 3 & 2 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array}\right)$$

Infine, dividiamo ogni riga per l'elemento sulla diagonale applicando operazioni elementari del secondo tipo

$$\left(\begin{array}{ccc|ccc} 4 & 0 & 0 & -1 & -1 & 2 \\ 0 & 4 & 0 & -1 & 3 & 2 \\ 0 & 0 & -4 & -3 & 1 & 2 \end{array}\right) \underset{\substack{R_1 \to (1/4)R_1 \\ R_1 \to (1/4)R_2 \\ R_3 \to -(1/4)R_3}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/4 & -1/4 & 1/2 \\ 0 & 1 & 0 & -1/4 & 3/4 & 1/2 \\ 0 & 0 & 1 & 3/4 & -1/4 & -1/2 \end{array}\right)$$

Abbiamo quindi ottenuto

$$A^{-1} = \begin{pmatrix} -1/4 & -1/4 & 1/2 \\ -1/4 & 3/4 & 1/2 \\ 3/4 & -1/4 & -1/2 \end{pmatrix}.$$

Come anticipato sopra, diamo una spiegazione del perché le matrici $A \in M_{m,n}(\mathbb{K})$ non quadrate (ovvero con numero di righe m diverso dal numero di colonne n) non possano essere invertibili. Più precisamente, mostriamo che se è verificata la condizione $AB = I$ allora necessariamente $m \leqslant n$ e, analogamente, se è verificata la condizione $BA = I$ allora necessariamente $m \geqslant n$.

Proposizione 6.13. *Sia A una matrice con m righe e n colonne. Allora L_A è iniettiva se e solo se il rango di A è uguale a n. In particolare, se L_A è iniettiva si deve avere $n \leqslant m$.*

Dimostrazione. Consideriamo la funzione

$$L_A : \mathbb{R}^n \to \mathbb{R}^m, \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

Per definizione, L_A è iniettiva se ogni m-upla $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$ nell'immagine di L_A è l'immagine di un'*unica* n-upla $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, ovvero se e solo se il sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2 \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = y_m \end{cases}$$

ammette un'unica soluzione per una certa m-upla di arrivo fissata. Tuttavia per la Proposizione 5.17 un sistema compatibile ha soluzione unica se e solo se il rango della matrice dei coefficienti è uguale al numero di incognite. In altre parole il numero di righe non nulle della matrice A dopo la sua riduzione a gradini dovrebbe essere n. Questo equivale a dire che il rango di A è esattamente n.

Per quanto riguarda l'ultima affermazione, si ricordi che il rango per righe è uguale al rango della matrice, si veda l'Osservazione 5.16. Per cui non si può avere $m < n$ in quanto, in tal caso, il rango di A sarebbe necessariamente minore o uguale al numero di righe m e conseguentemente ad n. □

Analogamente possiamo dimostrare la seguente

Proposizione 6.14. *Sia A una matrice con m righe e n colonne: allora L_A è suriettiva se e solo se il rango di A è uguale a m. In particolare, se L_A è suriettiva si deve avere $m \leqslant n$.*

Dimostrazione. Per definizione, L_A è suriettiva se e solo se per ogni m-upla $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{K}^m$ esiste una n-upla $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$ tale che $L_A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$, ovvero se e solo se il sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2 \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = y_m \end{cases}$$

è compatibile per qualunque scelta dei termini noti $y_1, y_2, \cdots, y_m$. Questo accade se e solo se nessuna riga di A si annulla quando viene ridotta a gradini (ovvero se e solo se il rango di A è uguale al numero di righe m). Infatti se nessuna riga si annulla, allora il sistema ammette soluzione qualunque siano i termini noti. Viceversa, se una riga è nulla allora possiamo scegliere i termini noti in modo che il corrispondente termine non si annulli ed ottenere così un sistema incompatibile.
Per quanto riguarda l'ultima affermazione, si ricordi che il rango per colonne è uguale al rango della matrice, si veda l'Osservazione 5.16. Per cui non si può avere $n < m$ in quanto, in tal caso, il rango di A sarebbe necessariamente minore o uguale al numero di colonne n e conseguentemente ad m. □

Siamo quindi pronti a dimostrare il Teorema 6.11.

Dimostrazione del Teorema 6.11. Una matrice $A \in M_{m,n}(\mathbb{K})$ ammette inversa se e solo se L_A è invertibile. Ovvero se e solo se L_A è suriettiva ed iniettiva. Combinando la Proposizione 6.14 e la Proposizione 6.13, otteniamo che questo accade se e solo se il rango di A è uguale sia ad m che ad n. Questo conclude la dimostrazione.

□

Concludiamo questo paragrafo osservando che l'insieme

$$\{A \in M_n(\mathbb{K}) \mid A \text{ è invertibile}\}$$

delle matrici invertibili di ordine n forma un gruppo (Definizione 2.56) rispetto all'operazione di prodotto di matrici.
Infatti osserviamo che il prodotto AB di due matrici A, B invertibili è ancora invertibile perché ha come inversa $B^{-1}A^{-1}$:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AId_nA^{-1} = AA^{-1} = Id_n.$$

Analogamente notiamo che $(B^{-1}A^{-1})(AB) = Id_n$. Questo mostra che il prodotto tra due matrici invertibili è ancora invertibile.
Essendo il prodotto di matrici associativo, la proprietà 1 della Definizione 2.56 è verificata. Inoltre,l'elemento neutro per il prodotto (la matrice identica Id_n) è invertibile (essendo $Id_n \cdot Id_n = Id_n$, essa ha come inversa se stessa). Quindi anche la proprietà 2 della Definizione 2.56 è verificata. Per quello che riguarda la proprietà 3, ci basta mostrare che l'inversa A^{-1} di una matrice A è anch'essa invertibile. Ma questo è chiaro in quanto A stessa è l'inversa di A^{-1}.

Definizione 6.15. Il gruppo $GL_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}) \mid A \text{ è invertibile}\}$ la cui operazione è la moltiplicazione tra matrici è detto **gruppo generale lineare** di ordine n.

Il fatto che $GL_n(\mathbb{K})$ sia un gruppo ci consente di usare i risultati che valgono in generale per qualunque gruppo senza doverli ridimostrare in questo caso specifico. Ad esempio, usando il Lemma 2.63, possiamo affermare che l'inversa del prodotto $A_1A_2 \cdots A_k$ di k matrici invertibili $A_1, A_2, \cdots, A_k$ è data da $A_k^{-1} \cdots A_2^{-1}A_1^{-1}$, senza bisogno di dimostrarlo.

6.3 Invertibilità e determinante di una matrice

In questa sezione introdurremo uno strumento di calcolo alternativo alla riduzione a gradini per determinare se una matrice quadrata A è invertibile e, in tal caso, determinarne l'inversa. Questo è il *determinante*, definito per matrici quadrate.
Come vedremo, il determinante è una funzione che associa un elemento di $\mathbb{K}$, denotato con $\det(A)$, ad ogni matrice A quadrata di ordine n con entrate appartenenti un campo $\mathbb{K}$. Il determinante rileva l'invertibilità di una matrice nel senso che A è invertibile se e solo se $\det(A) \neq 0$. Dal momento che, come visto nel paragrafo precedente, una matrice è invertibile esattamente quando le sue righe sono indipendenti, questo equivale a dire che $\det(A) = 0$ se e solo se le righe della matrice A sono dipendenti.

Prima di dare la definizione generale, mostriamo che nel caso di matrici quadrate di ordine 2 e di ordine 3 non è difficile trovare una funzione delle entrate che sia nulla se e solo se le righe della matrice sono dipendenti.

Infatti, data una generica matrice

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

di ordine 2, essendoci solo due righe queste saranno dipendenti se e solo se sono proporzionali. Questo, supponendo per il momento che a_{21} e a_{22} siano diversi da zero, significa che

$$\frac{a_{11}}{a_{21}} = \frac{a_{12}}{a_{22}} \tag{6.24}$$

ovvero, moltiplicando l'equazione per a_{21} (così da semplificarlo al primo membro) e per a_{22} (così da semplificarlo al secondo),

$$a_{11}a_{22} = a_{12}a_{21}.$$

Portando tutto al primo membro, si ottiene

$$a_{11}a_{22} - a_{12}a_{21} = 0 \tag{6.25}$$

che ci dà, come volevamo, una funzione delle entrate della matrice che si annulla esattamente quando le righe sono dipendenti.

Si noti che la (6.25) rileva la dipendenza delle righe anche senza l'ipotesi, che era necessaria per poter scrivere la (6.24), che a_{21} e a_{22} fossero diversi da zero. Infatti se, ad esempio, $a_{21} = 0$ allora la (6.25) diventa $a_{11}a_{22} = 0$, che equivale a dire che o $a_{11} = 0$ o $a_{22} = 0$. Nel primo caso la matrice è della forma $A = \left(\begin{smallmatrix} 0 & a_{12} \\ 0 & a_{22} \end{smallmatrix}\right)$, nel secondo $A = \left(\begin{smallmatrix} a_{11} & a_{12} \\ 0 & 0 \end{smallmatrix}\right)$, e in entrambi i casi le righe sono proporzionali.

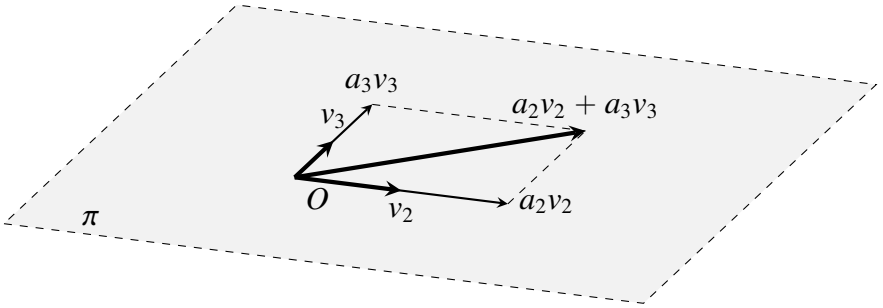
Anche per matrici di ordine 3 non è difficile ottenere, aiutandosi con un'interpretazione geometrica, una funzione delle entrate che si annulli se e solo se le righe sono dipendenti. Più precisamente, consideriamo

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

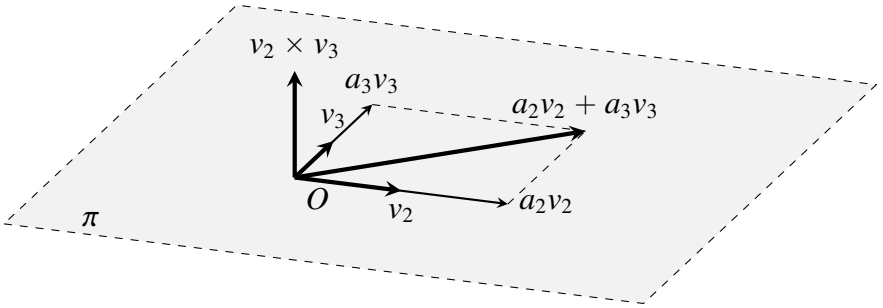
Possiamo pensare che le righe R_1, R_2, R_3 della matrice rappresentino le terne delle coordinate, rispetto a un sistema di riferimento ortonormale, di tre vettori geometrici v_1, v_2, v_3 applicati nello spazio.

Per cui esiste una relazione di dipendenza tra R_1, R_2, R_3 se e solo se la stessa relazione di dipendenza vale tra i vettori corrispondenti (ad esempio, $R_3 = R_1 + R_2$ se e solo se

$v_3 = v_1 + v_2$, oppure $R_1 = 2R_2 - R_3$ se e solo se $v_1 = 2v_2 - v_3$ e così via).
Ora, se esistesse una relazione di dipendenza tra i vettori questo vorrebbe dire che essi stanno sullo stesso piano. Ad esempio, se avessimo $v_1 = a_2v_2 + a_3v_3$ allora v_1 sarebbe sullo stesso piano su cui stanno v_2 e v_3 in quanto una qualunque combinazione di v_2 e v_3 appartiene a tale piano



Ma appartenere al piano su cui stanno v_2 e v_3 significa essere perpendicolare al prodotto vettoriale $v_2 \wedge v_3$ che è infatti perpendicolare a v_2 e v_3 e quindi a qualunque altro vettore si trovi su quel piano.



Traduciamo allora tale condizione in coordinate. Assumendo, come abbiamo detto sopra, che le coordinate di v_1, v_2, v_3 siano date rispettivamente dalla prima, seconda e terza riga della matrice A, ovvero $v_1 \equiv (a_{11}, a_{12}, a_{13})$, $v_2 \equiv (a_{21}, a_{22}, a_{23})$, $v_3 \equiv (a_{31}, a_{32}, a_{33})$, abbiamo

$$v_2 \wedge v_3 \equiv (a_{22}a_{33} - a_{23}a_{32}, a_{23}a_{31} - a_{21}a_{33}, a_{21}a_{32} - a_{22}a_{31}).$$

Ricordando che due vettori sono perpendicolari se e solo se il loro prodotto vettoriale è nullo, si ha che v_1 è perpendicolare a $v_2 \wedge v_3$ (ovvero sta sullo stesso piano su cui stanno v_2 e v_3) se e solo se

$$a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(a_{23}a_{31} - a_{21}a_{33}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) = 0 \qquad (6.26)$$

ovvero, svolgendo i calcoli,

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} = 0. \qquad (6.27)$$

Come volevamo, abbiamo ottenuto anche nel caso di matrici di ordine 3 una funzione delle entrate che si annulla se e solo se le sue righe sono dipendenti.
Ora, notiamo che i primi membri sia dell'espressione (6.25) che della (6.27) sono somma (con segno $+$ o $-$) di addendi del tipo

$$a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} \tag{6.28}$$

dove $\sigma(1),\sigma(2),\cdots,\sigma(n)$ è una permutazione degli indici di colonna. Ad esempio, nella (6.27) abbiamo che

l'addendo $a_{11}a_{22}a_{33}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ con σ data dalla permutazione $\sigma(1)=1,\ \sigma(2)=2,\ \sigma(3)=3$

l'addendo $a_{11}a_{23}a_{32}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ con σ data dalla permutazione $\sigma(1)=1,\ \sigma(2)=3,\ \sigma(3)=2$

l'addendo $a_{12}a_{23}a_{31}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ con σ data dalla permutazione $\sigma(1)=2,\ \sigma(2)=3,\ \sigma(3)=1$

l'addendo $a_{12}a_{21}a_{33}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ con σ data dalla permutazione $\sigma(1)=2,\ \sigma(2)=1,\ \sigma(3)=3$

l'addendo $a_{13}a_{21}a_{32}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ con σ data dalla permutazione $\sigma(1)=3,\ \sigma(2)=1,\ \sigma(3)=2$

l'addendo $a_{13}a_{22}a_{31}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ con σ data dalla permutazione $\sigma(1)=3,\ \sigma(2)=2,\ \sigma(3)=1$

Analogamente, nella (6.25) abbiamo che

l'addendo $a_{11}a_{22}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}$ con σ data dalla permutazione $\sigma(1)=1,\ \sigma(2)=2$

l'addendo $a_{12}a_{21}$ è del tipo $a_{1\sigma(1)}a_{2\sigma(2)}$ con σ data dalla permutazione $\sigma(1)=2,\ \sigma(2)=1$

In entrambi i casi, si vede che abbiamo un addendo per ognuna delle possibili permutazioni degli indici di colonna (nel caso di ordine 2 abbiamo due sole permutazioni dell'insieme $\{1,2\}$, nel caso di ordine 3 abbiamo 6 permutazioni dell'insieme $\{1,2,3\}$). Per quanto riguarda il segno di tali addendi, possiamo notare che sia nella (6.25) che nella (6.27) gli addendi che corrispondono a permutazioni pari (si veda da Definizione 2.62) hanno segno positivo, mentre quelli dati da permutazioni dispari hanno segno negativo. Ad esempio, l'addendo $a_{11}a_{22}$ ha segno positivo perché corrisponde alla permutazione identica, che è pari; mentre l'addendo $a_{12}a_{21}$ ha segno - perché corrisponde

alla trasposizione che scambia 1 e 2, che è una permutazione dispari. Analogamente, nella (6.27) l'addendo $a_{11}a_{22}a_{33}$ ha segno positivo perché corrisponde alla permutazione identica, l'addendo $a_{11}a_{23}a_{32}$ ha segno negativo perché corrisponde alla trasposizione che scambia 2 e 3, l'addendo $a_{12}a_{23}a_{31}$ ha segno positivo perché corrisponde al ciclo di lunghezza tre $(1\ 2\ 3)$, che si può ottenere tramite la composizione $(1\ 3) \circ (1\ 2)$ di due trasposizioni, e quindi è pari, e così via.

Definizione 6.16. Il **segno** $s(\sigma)$ di una permutazione $\sigma \in S_n$ è definito da

$$s(\sigma) = \begin{cases} 1 & \text{se } \sigma \text{ è pari} \\ -1 & \text{se } \sigma \text{ è dispari} \end{cases}$$

Con questa definizione sia la (6.25) che la (6.27) possono essere scritte come

$$\sum_{\sigma \in S_n} s(\sigma)\ a_{1\sigma(1)}a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

rispettivamente per $n = 2$ e $n = 3$.
Quello che si può dimostrare è che, anche per matrici di ordine $n > 3$ si ha che tale espressione si annulla se e solo se le righe della matrice sono dipendenti. Possiamo allora dare la seguente

Definizione 6.17. Data una matrice quadrata $A \in M_n(\mathbb{K})$, il suo determinante $\det(A)$ è l'elemento di $\mathbb{K}$ dato da

$$\sum_{\sigma \in S_n} s(\sigma)\ a_{1\sigma(1)}a_{2\sigma(2)} \cdots a_{n\sigma(n)} \tag{6.29}$$

Notiamo che la (6.29) ha tanti addendi quante sono le possibili permutazioni degli n indici di colonna che, come sappiamo dalla Sezione 2.9, sono $n!$. Per le matrici di ordine 4, si ha $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$, per quelle di ordine 5 si ha $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$, quindi vediamo che il numero di addendi nell'espressione del determinante cresce molto rapidamente, il che rende l'uso di tale definizione non molto pratico per il calcolo.
Per questo motivo, introduciamo ora una formula che ci consente di calcolare il determinante senza dover utilizzare la (6.29).

Definizione 6.18. Data una matrice quadrata $A \in M_n(\mathbb{K})$ di ordine $n \geqslant 2$ si dice **cofattore** (o **complemento algebrico**) di una sua entrata a_{ij}, l'elemento $C_{ij} \in \mathbb{K}$ dato dal determinante della matrice che si ottiene da A cancellando la i-esima riga e la j-esima colonna, moltiplicato per $(-1)^{i+j}$.

Esempio 6.19. Calcoliamo i cofattori delle entrate della matrice $A \in M_3(\mathbb{R})$ data da

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 2 \end{pmatrix}.$$

I cofattori sono

$$C_{11} = (-1)^{1+1} \det \begin{pmatrix} -1 & 1 \\ 0 & 2 \end{pmatrix} = -2, \quad C_{12} = (-1)^{1+2} \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = -1$$

$$C_{13} = (-1)^{1+3} \det \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} = 1, \quad C_{21} = (-1)^{2+1} \det \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} = -4$$

$$C_{22} = (-1)^{2+2} \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = 1, \quad C_{23} = (-1)^{2+3} \det \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} = 2$$

$$C_{31} = (-1)^{3+1} \det \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} = 3, \quad C_{32} = (-1)^{3+2} \det \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 0$$

$$C_{33} = (-1)^{3+3} \det \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} = -3.$$

Definizione 6.20. Sia A una matrice quadrata di ordine $n \geqslant 2$ e siano $\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix}$ la sua i-esima riga e $\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$ la sua j-esima colonna.
Si dice **sviluppo di Laplace rispetto alla i-esima riga** l'elemento di $\mathbb{K}$ definito da

$$a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}. \tag{6.30}$$

Analogamente si dice **sviluppo di Laplace rispetto alla j-esima colonna** l'elemento di $\mathbb{K}$ definito da

$$a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}. \tag{6.31}$$

Di seguito enunciamo un fondamentale risultato senza dimostrarlo.

Proposizione 6.21. *Sia A una matrice quadrata di ordine $n \geqslant 2$. Allora lo sviluppo di Laplace rispetto ad una qualunque sua riga o colonna è uguale al determinante di A. In formule*

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

indipendentemente dalla riga o dalla colonna scelta.

Ad esempio, nel caso di una matrice di ordine 3, scegliendo la prima riga di

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

si ha

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} =$$

$$= a_{11}(-1)^{1+1} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} +$$

$$+a_{12}(-1)^{1+2} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} +$$

$$+a_{13}(-1)^{1+3} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} =$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

che coincide proprio con la (6.27).

Osserviamo che lo sviluppo di Laplace è una formula *ricorsiva*. Infatti per calcolare il determinante di una matrice A di ordine n prevede che si conoscano i cofattori, che sono determinanti di matrici di ordine $n-1$ (quelle che si ottengono cancellando una riga e una colonna da A). A loro volta, ciascuno di questi determinanti va calcolato con la stessa formula, riducendoli al calcolo di determinanti di matrici di ordine $n-2$. Si continua così fino ad arrivare a determinanti di matrici di ordine 2, per i quali si ha la semplice formula $\det(A) = a_{11}a_{22} - a_{12}a_{21}$.

Esempio 6.22. Si consideri la matrice

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 \end{pmatrix}.$$

Scegliamo di sviluppare il determinante rispetto alla terza riga, usando cioè la formula (6.30) nel caso $i = 3$:

$$\det(A) = a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} + a_{34}C_{34}.$$

Tenendo conto che $a_{32} = a_{33} = 0$ (quindi senza calcolare i corrispondenti cofattori) si

ha

$$\det(A) = 1 \cdot (-1)^{3+1} \det \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 1 \\ 2 & 1 & 0 \end{pmatrix} + 1 \cdot (-1)^{3+4} \det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 2 & 1 \end{pmatrix}. \tag{6.32}$$

Ora, ciascuno dei due determinanti di ordine 3 che compaiono in questa uguaglianza va di nuovo calcolato usando la formula data. Ad esempio, se per il primo determinante scegliamo la terza colonna (e usiamo quindi la formula $\det(A) = a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33}$) tenendo conto che $a_{13} = a_{33} = 0$, si ha

$$\det \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 1 \\ 2 & 1 & 0 \end{pmatrix} = 1 \cdot (-1)^{2+3} \det \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = -1$$

dove il determinante di ordine 2 è stato calcolato usando la formula $\det(A) = a_{11}a_{22} - a_{12}a_{21}$. Analogamente, se per l'altro determinante scegliamo la prima riga (quindi usiamo $\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$) tenendo conto che $a_{13} = 0$, si ha

$$\det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 2 & 1 \end{pmatrix} = 1 \cdot (-1)^{1+1} \det \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} + 1 \cdot (-1)^{1+2} \det \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} = 3 - 1 = 2.$$

Sostituendo questi risultati nella (6.32), si trova allora

$$\det(A) = 1 \cdot (-1)^{3+1}(-1) + 1 \cdot (-1)^{3+4}\, 2 = -3.$$

Ora, le nozioni di determinante e cofattore appena introdotte, oltre a dirci se una matrice è invertibile, ci danno un modo alternativo al procedimento descritto nel paragrafo precedente per calcolarne l'inversa. Infatti, si ha il seguente risultato, che non dimostriamo.

Proposizione 6.23. *Sia $A \in M_n(\mathbb{K})$ una matrice invertibile (ovvero con $\det(A) \neq 0$). Allora la sua inversa A^{-1} è data da*

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{pmatrix} \tag{6.33}$$

dove C_{ij} indica, come sopra, il cofattore di a_{ij}, e $\frac{1}{\det(A)}$ davanti alla matrice dei cofattori significa che ogni entrata di tale matrice deve essere moltiplicata per $\frac{1}{\det(A)}$.

Osservazione 6.24. A proposito della disposizione dei cofattori nella (6.33), si noti che i cofattori delle entrate della prima *riga* di A sono nella prima *colonna* della (6.33), i cofattori delle entrate della seconda *riga* di A sono nella seconda *colonna* della (6.33), e così via. In altri termi le righe sono scambiate con le colonne e vicerversa.

Esempio 6.25. Calcoliamo l'inversa della matrice

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 2 \end{pmatrix}$$

di cui abbiamo calcolato i cofattori nell'Esempio 6.19. Da quel calcolo risulta

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{2n} & C_{33} \end{pmatrix} = -\frac{1}{3} \begin{pmatrix} -2 & -4 & 3 \\ -1 & 1 & 0 \\ 1 & 2 & -3 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{4}{3} & -1 \\ \frac{1}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & -\frac{2}{3} & 1 \end{pmatrix}.$$

Il determinate di A è stato calcolato sviluppandolo secondo Laplace rispetto alla terza riga, usando i cofattori già calcolati:

$$\det(A) = a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} = 1 \cdot 3 + 0 \cdot 0 + 2 \cdot (-3) = -3.$$

Osservazione 6.26. Si noti che la (6.33), nel caso $n = 2$, diventa la semplice formula

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \tag{6.34}$$

in quanto i cofattori sono semplicemente $C_{11} = (-1)^{1+1}a_{22}$, $C_{12} = (-1)^{1+2}a_{21}$, $C_{21} = (-1)^{2+1}a_{12}$, $C_{22} = (-1)^{2+2}a_{11}$. In pratica, a parte dividere per il determinante, la matrice inversa si ottiene scambiando tra loro i due elementi sulla diagonale e cambiando di segno le restanti entrate.

Essendo il determinante uno strumento fondamentale per la determinazione dell'invertibilità di una matrice, il calcolo della sua inversa e per molte altre apllicazioni, è fondamentale conoscerne le proprietà, anche allo scopo di agevolarne il calcolo.
Iniziamo con il chiederci come si comporta il determinante di una matrice rispetto alle operazioni elementari.

(1) Se A' si ottiene da una matrice $A \in M_n(\mathbb{K})$ moltiplicando una riga di A per $c \in \mathbb{K}$, allora $\det(A') = c\det(A)$.

Dimostrazione. In base alla Definizione 6.17 il determinante della matrice A' è

$$\det(A') = \sum s(\sigma)a_{1\sigma(1)} \cdots (ca_{i\sigma(i)}) \cdots a_{n\sigma(n)}.$$

Il fattore c è comune a tutti gli addendi della sommatoria, possiamo quindi metterlo in evidenza davanti alla somma e scrivere

$$\det(A') = c \sum s(\sigma) a_{1\sigma(1)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)},$$

ovvero $\det(A') = c \det(A)$, come volevamo. □

Esempio 6.27. Sia $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$ e sia $A' = \left(\begin{smallmatrix} 5 & 10 \\ 3 & 4 \end{smallmatrix}\right)$ è la matrice ottenuta da A moltiplicando la prima riga per $c = 5$. Allora si ha

$$\begin{aligned} \det(A) &= 1 \cdot 4 - 3 \cdot 2 = -2, \\ \det(A') &= 5 \cdot 4 - 10 \cdot 3 = -10. \end{aligned}$$

(2) Se una riga R_i della matrice $A \in M_n(\mathbb{K})$ si scrive come somma $R_i = R'_i + R''_i$, si ha

$$\det \begin{pmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_n \end{pmatrix} = \det \begin{pmatrix} R'_1 \\ \vdots \\ R_i \\ \vdots \\ R_n \end{pmatrix} + \det \begin{pmatrix} R_1 \\ \vdots \\ R''_i \\ \vdots \\ R_n \end{pmatrix}. \tag{6.35}$$

Dimostrazione. Siano

$$\begin{aligned} R'_i &= \begin{pmatrix} a'_{i1} & a'_{i2} & \cdots & a'_{in} \end{pmatrix} \\ R''_i &= \begin{pmatrix} a''_{i1} & a''_{i2} & \cdots & a''_{in} \end{pmatrix} \end{aligned}$$

e di conseguenza

$$R_i = R'_i + R''_i = \begin{pmatrix} a'_{i1} + a''_{i1} & a'_{i2} + a''_{i2} & \cdots & a'_{in} + a''_{in} \end{pmatrix}.$$

Allora, in base alla Definizione 6.17, si ha

$$\det \begin{pmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_n \end{pmatrix} = \sum s(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{i-1\sigma(i-1)} \left(a'_{i\sigma(i)} + a''_{i\sigma(i)} \right) a_{i+1\sigma(i+1)} \cdots a_{n\sigma(n)} =$$

$$= \sum s(\sigma) a_{1\sigma(1)} \cdots a'_{i\sigma(i)} \cdots a_{n\sigma(n)} + \sum s(\sigma) a_{1\sigma(1)} \cdots a''_{i\sigma(i)} \cdots a_{n\sigma(n)} =$$

$$= \det \begin{pmatrix} R_1 \\ \vdots \\ R'_i \\ \vdots \\ R_n \end{pmatrix} + \det \begin{pmatrix} R_1 \\ \vdots \\ R''_i \\ \vdots \\ R_n \end{pmatrix}$$

dove abbiamo usato la proprietà distributiva in ogni termine della sommatoria. □

Esempio 6.28. Se $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$, la seconda riga si scrive come somma $\begin{pmatrix} 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 3 \end{pmatrix}$ e, in accordo con la (6.35), si vede che

$$\det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \det \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} + \det \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$$

come può essere verificato col calcolo esplicito (si ottiene $-2 = -3 + 1$).

(3) Se la matrice A' si ottiene da $A \in M_n(\mathbb{K})$ scambiando tra loro due righe, o due colonne, allora $\det(A') = -\det(A)$.

Dimostrazione. Sia A' la matrice ottenuta da A scambiando la riga R_i con la riga R_j: per ogni k si ha $a'_{ik} = a_{jk}$ e $a'_{jk} = a_{ik}$, mentre tutte le altre entrate sono uguali a quelle di A.

In base alla Definizione 6.17, si ha

$$\begin{aligned} \det(A') &= \sum_{\sigma \in S_n} s(\sigma)\, a'_{1\sigma(1)} \cdots a'_{i\sigma(i)} \cdots a'_{j\sigma(j)} \cdots a'_{n\sigma(n)} = \\ &= \sum_{\sigma \in S_n} s(\sigma)\, a_{1\sigma(1)} \cdots a_{j\sigma(i)} \cdots a_{i\sigma(j)} \cdots a_{n\sigma(n)}. \end{aligned} \tag{6.36}$$

Confrontando la (6.36) con l'espressione

$$\det(A) = \sum_{\sigma \in S_n} s(\sigma)\, a_{1\sigma(1)} \cdots a_{i\sigma(i)} \cdots a_{j\sigma(j)} \cdots a_{n\sigma(n)} \tag{6.37}$$

vediamo che in ogni addendo le entrate delle colonne $\sigma(i)$ e $\sigma(j)$ hanno gli indici di riga i e j scambiati tra loro.

Ora, allo scopo di dimostrare il risultato, riscriviamo la (6.36) come segue. Consideriamo la trasposizione $\tau \in S_n$ che scambia i con j e lascia invariati tutti gli altri elementi, e definiamo la funzione $f : S_n \to S_n$ che associa a ogni permutazione σ la composizione $\sigma \circ \tau$.

Tale funzione è invertibile e ha come inversa se stessa: $f(f(\sigma)) = f(\sigma \circ \tau) = \sigma \circ \tau \circ \tau = \sigma$ (dove abbiamo usato il fatto che, per ogni trasposizione, $\tau \circ \tau = id$).

Quindi f è biiettiva; il che significa che, al variare di $\sigma \in S_n$, $\sigma \circ \tau$ ci dà tutte le permutazioni di S_n.

Allora possiamo sostituire il σ che compare in ognuno degli addendi di (6.36) con $\sigma \circ \tau$ e ottenere comunque, al variare di $\sigma \in S_n$, tutte le permutazioni di S_n in modo che la somma sia invariata. Si ha quindi

$$\det(A') = \sum_{\sigma \in S_n} s(\sigma \circ \tau)\, a_{1(\sigma\circ\tau)(1)} \cdots a_{j(\sigma\circ\tau)(i)} \cdots a_{i(\sigma\circ\tau)(j)} \cdots a_{n(\sigma\circ\tau)(n)}.$$

Dal momento che τ scambia i e j lasciando invariati tutti gli altri elementi di $\{1, 2, \cdots, n\}$, si ha

$$(\sigma \circ \tau)(k) = \begin{cases} k & \text{se } k \neq i, j \\ j & \text{se } k = i \\ i & \text{se } k = j \end{cases}$$

e quindi otteniamo

$$\det(A') = \sum_{\sigma \in S_n} s(\sigma \circ \tau)\, a_{1\sigma(1)} \cdots a_{j\sigma(j)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)}. \tag{6.38}$$

Notiamo che la (6.38) è uguale a quella (6.37) del determinante di A a parte che in essa compare $s(\sigma \circ \tau)$ invece che $s(\sigma)$. Ora, in base alla definizione di segno, si ha $s(\sigma \circ \tau) = -s(\sigma)$. Infatti, se $s(\sigma) = +1$ allora questo significa che σ si scrive come prodotto di un numero pari di trasposizioni, e quindi $\sigma \circ \tau$, contenendo in più la trasposizione τ, risulterà prodotto di un numero dispari di trasposizioni, da cui $s(\sigma \circ \tau) = -1$. Analogamente, se $s(\sigma) = -1$, allora σ è una permutazione dispari e, conseguentemente, $\sigma \circ \tau$ è una permutazione pari, da cui $s(\sigma \circ \tau) = +1$. Quindi la (6.38) diventa

$$\begin{aligned} \det(A') &= \sum_{\sigma \in S_n} -s(\sigma)\, a_{1\sigma(1)} \cdots a_{j\sigma(j)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} = \\ &= -\sum_{\sigma \in S_n} s(\sigma)\, a_{1\sigma(1)} \cdots a_{j\sigma(j)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} = -\det(A) \end{aligned}$$

come volevamo.

□

Esempio 6.29. Sia

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

e sia

$$A' = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}$$

la matrice ottenuta da A scambiando tra loro la prima e la seconda riga. Allora si ha

$$\det(A) = 1 \cdot 4 - 2 \cdot 3 = -2,$$
$$\det(A') = 3 \cdot 2 - 4 \cdot 1 = +2.$$

Le proprietà (1) e (2) appena viste ci consentono di capire come si comporta il determinante quando eseguiamo operazioni elementari del terzo tipo. Più precisamente, supponiamo che la matrice A' sia ottenuta da $A \in M_n(\mathbb{K})$ sommando alla sua i-esima riga R_i la j-esima riga R_j moltiplicata per c, ovvero

$$A' = \begin{pmatrix} R_1 \\ \vdots \\ R_i + cR_j \\ \vdots \\ R_j \\ \vdots \\ R_n \end{pmatrix}.$$

Ma allora, per la proprietà (2) abbiamo

$$\det(A') = \det \begin{pmatrix} R_1 \\ \vdots \\ R_i + cR_j \\ \vdots \\ R_j \\ \vdots \\ R_n \end{pmatrix} = \det \begin{pmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_j \\ \vdots \\ R_n \end{pmatrix} + \det \begin{pmatrix} R_1 \\ \vdots \\ cR_j \\ \vdots \\ R_j \\ \vdots \\ R_n \end{pmatrix}. \tag{6.39}$$

Mentre il primo addendo è il determinante della matrice A, il secondo addendo è nullo, in quanto determinante di una matrice con le righe dipendenti (compaiono R_j e cR_j che sono proporzionali).[3]

[3] Se non si vuole utilizzare questo risultato che non abbiamo dimostratato basta usare prima la proprietà (1) ed applicare poi la proprietà (3) alle righe R_j ed R_j che si trovano in i-esima e j-esima posizione.

Quindi la (6.39) si legge

$$\det(A') = \det(A) + 0 = \det(A).$$

Concludiamo che *quando eseguiamo un'operazione elementare del terzo tipo sulle righe di una matrice, il determinante non cambia*.
Questo risultato può essere usato per facilitare il calcolo del determinante di una matrice, in quanto mediante operazioni elementari del terzo tipo si può modificare la matrice trasformando alcune entrate in 0 senza cambiare il determinante. Si può poi calcolare lo sviluppo di Laplace rispetto ad una riga o colonna che presenta il numero maggiore di questi zeri così comparsi (ricordiamo che se un'entrata è nulla non è necessario calcolare il corrispondente cofattore).
Mediante tali operazioni potremmo anche trasformare la matrice in una matrice a gradini

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1n-1} & a_{n-1n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{pmatrix} \tag{6.40}$$

Il determinante di tale matrice è dato semplicemente dal prodotto $a_{11}a_{22}\cdots a_{nn}$ degli elementi della sua diagonale. Infatti sviluppando secondo Laplace rispetto alla prima colonna, essendo a_{11} l'unico elemento non nullo, si avrà

$$\det(A) = a_{11}(-1)^{1+1}\det\begin{pmatrix} a_{22} & \cdots & a_{2n-1} & a_{2n} \\ 0 & \ddots & \vdots & \vdots \\ 0 & \cdots & a_{n-1n-1} & a_{n-1n} \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

Se ora per lo sviluppo di Laplace del cofattore scegliamo nuovamente la prima colonna, che ha come unico elemento non nullo a_{22}, otteniamo

$$\det(A) = a_{11}a_{22}\det\begin{pmatrix} a_{33} & \cdots & a_{3n-1} & a_{3n} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & a_{n-1n-1} & a_{n-1n} \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

e continuando a sviluppare sempre rispetto alla prima colonna otteniamo proprio $\det(A) = a_{11}a_{22}\cdots a_{nn}$.

Osservazione 6.30. Tutte le proprietà del determinante che abbiamo visto valide per operazioni sulle righe sono vere anche rispetto alle colonne. Per cui valgono le seguenti

proprietà che si dimostrano in modo del tutto analogo.

(1) Se la matrice A' si ottiene da $A \in M_n(\mathbb{K})$ moltiplicando una colonna di A per $c \in \mathbb{K}$, allora $\det(A') = c\det(A)$.

Ad esempio, se $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$ e $A' = \left(\begin{smallmatrix} 5 & 2 \\ 15 & 4 \end{smallmatrix}\right)$ è la matrice ottenuta da A moltiplicando la prima colonna per $c = 5$, allora si ha

$$\det(A) = -2 \text{ e } \det(A') = 5 \cdot 4 - 10 \cdot 3 = -10.$$

(2) Se una colonna C_i della matrice $A \in M_n(\mathbb{K})$ si scrive come somma $C_i = C_i' + C_i''$, si ha

$$\begin{aligned} &\det\begin{pmatrix} C_1 & \cdots & C_i' + C_i'' & \cdots & C_n \end{pmatrix} = \\ &= \det\begin{pmatrix} C_1 & \cdots & C_i' & \cdots & C_n \end{pmatrix} + \det\begin{pmatrix} C_1 & \cdots & C_i'' & \cdots & C_n \end{pmatrix}. \end{aligned}$$

Ad esempio, se $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$, la seconda colonna si scrive come somma

$$\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \end{pmatrix}$$

e si ha

$$\det\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \det\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} + \det\begin{pmatrix} 1 & 0 \\ 3 & 3 \end{pmatrix}$$

come può essere verificato calcolando i determinanti (si ottiene $-2 = -5 + 3$).

(3) se la matrice A' si ottiene da $A \in M_n(\mathbb{K})$ scambiando tra loro due colonne, allora $\det(A') = -\det(A)$.

Ad esempio, se $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$ e $A' = \left(\begin{smallmatrix} 2 & 1 \\ 4 & 3 \end{smallmatrix}\right)$ è la matrice ottenuta da A scambiando tra loro la prima e la seconda colonna, si ha

$$\det(A) = 1 \cdot 4 - 2 \cdot 3 = -2 \text{ e } \det(A') = 2 \cdot 3 - 1 \cdot 4 = +2.$$

Come ulteriore proprietà del determinante, citiamo senza dimostrarlo il seguente

Teorema 6.31 (Teorema di Binet)**.** *Siano $A, B \in M_n(\mathbb{K})$ due matrici quadrate. Allora*

$$\det(AB) = \det(A)\det(B).$$

Ad esempio, sia $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ e $B = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$. Allora si ha $AB = \begin{pmatrix} 4 & 7 \\ 10 & 15 \end{pmatrix}$, e si vede che $\det(A) = -2$, $\det(B) = 5$ e $\det(AB) = -10 = (-2) \cdot 5$, come afferma il teorema di Binet.

Dal teorema di Binet discende immediatamente il seguente

Corollario 6.32. *Se $A \in M_n(\mathbb{K})$ è una matrice invertibile, allora*

$$\det(A^{-1}) = \frac{1}{\det(A)}. \tag{6.41}$$

Dimostrazione. Basta applicare il teorema di Binet al caso $B = A^{-1}$. In questo caso infatti abbiamo

$$\det(AA^1) = \det(A)\det(A^{-1})$$

ovvero, tenuto conto che $AA^{-1} = Id_n$,

$$\det(Id_n) = \det(A)\det(A^{-1}).$$

Ma la matrice identica Id_n è una matrice del tipo (6.40), con $a_{11}, a_{22}, \cdots, a_{nn}$ tutti uguali a 1, quindi il suo determinante è 1. In definitiva, si ottiene

$$1 = \det(A)\det(A^{-1})$$

che implica subito la (6.41). □

Vediamo ora un'applicazione del determinante e più precisamente della formula (6.33) per l'inversa che ci fornisce un modo alternativo alla riduzione a gradini per risolvere i sistemi di equazioni lineari sotto certe ipotesi.

A tale scopo, osserviamo prima che un generico sistema lineare di m equazioni in n incognite

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{6.42}$$

può essere riscritto in forma molto più concisa grazie al prodotto di matrici. Più precisamente, denotiamo con

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

la matrice dei coefficienti del sistema, con

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

la matrice (costituita da una sola colonna) che ha come entrate le incognite, e con

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

la matrice (sempre costituita da una sola colonna) che ha come entrate i termini noti del sistema. Allora svolgendo il prodotto righe per colonne, notiamo che

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

Quindi il sistema (6.42) può essere riscritto come prodotto di matrici nella forma

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

o nella forma più compatta

$$Ax = b. \tag{6.43}$$

Ora, supponiamo che la matrice A dei coefficienti del sistema sia quadrata di ordine n (quindi il sistema ha n equazioni e n incognite) e che abbia determinante diverso da zero. Allora A è invertibile e possiamo moltiplicare a sinistra entrambi i membri dell'uguaglianza (6.43) per l'inversa A^{-1} ottenendo

$$A^{-1}(Ax) = A^{-1}b.$$

Tenendo conto che il prodotto di matrici gode della proprietà associativa, questo può essere riscritto come

$$(A^{-1}A)x = A^{-1}b$$

ovvero, visto che per definizione di inversa $A^{-1}A = Id_n$,

$$Id_n x = A^{-1}b.$$

Per cui, essendo Id_n elemento neutro per il prodotto, otteniamo

$$x = A^{-1}b. \tag{6.44}$$

Quindi la formula (6.44) fornisce la soluzione x di un sistema con lo stesso numero di equazioni e di incognite e matrice dei coefficienti con determinante diverso da zero.

Osservazione 6.33. Si noti che non abbiamo fatto altro che applicare gli stessi passaggi, normalmente sottointesi, che si applicano quando si vuole risolvere una semplice equazione di primo grado in una sola incognita $ax = b$. Se, ad esempio, vogliamo risolvere l'equazione $2x = 3$, dividiamo entrambi i membri per 2 o, equivalentemente, moltiplichiamo per l'inverso $\frac{1}{2}$ di 2 ottenendo $\frac{1}{2}(2x) = \frac{1}{2}3 = \frac{3}{2}$. Per la proprietà associativa al primo membro si ha $(\frac{1}{2}2)x = \frac{3}{2}$ ovvero, essendo $\frac{1}{2}2 = 1$, si ha $1x = \frac{3}{2}$. Per cui, essendo 1 elemento neutro per la moltiplicazione tra reali, $x = \frac{3}{2}$, che è la soluzione dell'equazione.
L'unica differenza con il caso dei sistemi $Ax = b$ è che non sempre A è invertibile, mentre, a meno che a non sia zero, nell'equazione $ax = b$ tale ipotesi è sempre garantita.

Se combiniamo la (6.44) con la formula per l'inversa (6.33), vediamo che la soluzione di un sistema con n equazioni e n incognite e matrice dei coefficienti A invertibile è data da

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det(A)} \begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}. \tag{6.45}$$

Quindi le componenti x_i della soluzione del sistema si ottengono moltiplicando le righe della matrice dei cofattori (divisa per il determinante di A) per la colonna dei termini noti:

$$x_1 = \frac{b_1C_{11} + b_2C_{21} + \cdots + b_nC_{n1}}{\det(A)}$$
$$x_2 = \frac{b_1C_{12} + b_2C_{22} + \cdots + b_nC_{n2}}{\det(A)}$$
$$\vdots$$

In generale

$$x_i = \frac{b_1C_{1i} + b_2C_{2i} + \cdots + b_nC_{ni}}{\det(A)} \tag{6.46}$$

Confrontando il numeratore della (6.46) con l'espressione del determinante di A calcolato con lo sviluppo di Laplace rispetto alla i-esima colonna

$$\det(A) = a_{1i}C_{1i} + a_{2i}C_{2i} + \cdots + a_{ni}C_{ni}$$

si nota che i termini noti $b_1, b_2, \cdots, b_n$ compaiono al posto degli elementi $a_{1i}, a_{2i}, \cdots, a_{ni}$ della colonna i-esima di A. In altre parole, tale numeratore coincide con lo sviluppo di Laplace del determinante della matrice che si ottiene da A ponendo i termini noti al posto della i-esima colonna. Riassumendo, abbiamo dimostrato il

Teorema 6.34 (Teorema di Cramer)**.** *Sia $Ax = b$ un sistema di n equazioni lineari in n incognite, con* $\det(A) \neq 0$*. Allora tale sistema ha un'unica soluzione $(x_1, x_2, \cdots, x_n)$ le cui componenti sono date da*

$$x_i = \frac{\det(B_i)}{\det(A)}, \quad i = 1, \cdots, n$$

dove B_i è la matrice che si ottiene da A ponendo la matrice b dei termini noti al posto della i-esima colonna di A.

Esempio 6.35. Consideriamo il sistema

$$\begin{cases} 2x_1 + x_2 = 5 \\ x_1 - x_2 = 3 \end{cases}$$

La matrice dei coefficienti del sistema

$$A = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$$

ha determinante $\det(A) = -3$, quindi è invertibile e possiamo applicare il metodo di Cramer. Le componenti dell'unica soluzione sono date da

$$x_1 = \frac{\det(B_1)}{\det(A)} = \frac{\det\begin{pmatrix} 5 & 1 \\ 3 & -1 \end{pmatrix}}{\det(A)} = \frac{-8}{-3} = \frac{8}{3}$$

$$x_2 = \frac{\det(B_2)}{\det(A)} = \frac{\det\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}}{\det(A)} = \frac{1}{-3} = -\frac{1}{3}$$

Osservazione 6.36. Quanto visto finora ci permette di tracciare un parallelo tra il comportamento delle soluzioni di un sistema $Ax = b$ di n equazioni lineari in n incognite e

il comportamento delle soluzioni di un'equazione congruenziale $ax \equiv_n b$, descritto dal Teorema 3.29 e dal Teorema 3.31.
Infatti, in tali risultati abbiamo visto che $ax \equiv_n b$ ha soluzioni se e solo se il massimo comune divisore d di a e n divide b. In tal caso le soluzioni *distinte modulo n* sono esattamente d. Da ciò si evince che nel caso in cui il massimo comune divisore d di a e n sia 1, la soluzione esiste ed è unica qualunque sia b (1 divide qualunque b). Ma $d = 1$ significa che a e n sono primi tra loro, il che a sua volta equivale a dire, come visto a pagina 113, che a è invertibile modulo n. Possiamo allora riformulare quanto appena detto come segue: se a è invertibile modulo n, l'equazione $ax \equiv_n b$ ha una sola soluzione qualunque sia b, mentre se a non è invertibile (ovvero il massimo comune divisore d di a e n è maggiore di 1) allora la soluzione non esiste per certi valori di b (quelli per cui d non divide b), e per i valori per cui esiste sicuramente non è unica (come abbiamo detto, ci sono d soluzioni).
Un sistema lineare $Ax = b$ di n equazioni in n incognite si comporta esattamente allo stesso modo. Se la matrice A è invertibile, allora la soluzione esiste qualunque sia b ed è unica, data come abbiamo visto sopra da $x = A^{-1}b$. Alternativamente, se A è invertibile il suo rango r è uguale a n e quindi dopo la riduzione a gradini non si annulla nessuna riga; quindi non non ci sono incompatibilità del tipo $0 = c$ con $c \neq 0$ e la soluzione esiste qualunque sia la colonna dei termini noti b. Inoltre questa è unica, come afferma la Proposizione 5.17, in quanto il rango r è uguale al numero n di incognite.
Invece, se A non è invertibile, il suo rango r è minore di n e quindi dopo la sua riduzione a gradini si annulla almeno una riga. Questo comporta che per certe scelte della colonna dei termini noti b si otterrà in corrispondenza della riga annullata un'incompatibilità del tipo $0 = c$ con $c \neq 0$, ovvero per certi b non avremo soluzione. D'altra parte, per i b per cui abbiamo soluzione, questa non sarà certamente unica, in quanto come sappiamo le soluzioni di un sistema compatibile sono ∞^{n-r} (sempre per la Proposizione 5.17), ed essendo $r < n$ per ipotesi si avrà $n - r > 0$ e quindi le soluzioni saranno infinite.

Oltre al prodotto, tra matrici può essere definita anche un'operazione di somma.
L'introduzione di tale operazione si rivela necessaria se vogliamo manipolare espressioni e uguaglianze tra matrici esattamente come manipoliamo espressioni e uguaglianze tra numeri. Ad esempio, come abbiamo già messo in evidenza in capitoli precedenti, per risolvere l'equazione tra numeri $ab = ac$ portiamo ac al primo membro cambiandolo di segno e ottenendo $ab - ac = 0$ e poi mettiamo in evidenza a ottenendo $a(b - c) = 0$. In questo caso stiamo sfruttando il fatto che tra numeri è definita un'operazione di somma per la quale valgono la proprietà associativa, l'esistenza di un elemento neutro, l'esistenza di un inverso e la proprietà distributiva. Infatti, per spostare ac al primo membro stiamo in realtà sommando entrambi i membri di $ab = ac$ con l'inverso additivo (ovvero l'opposto) di ac, cioè $-ac$, ottenendo $ab + (-ac) = ac + (-ac) = 0$, ovvero $ab - ac = 0$. Inoltre per mettere in evidenza a sfruttiamo la proprietà distributiva della somma rispetto al prodotto.

Allo scopo di poter eseguire manipolazioni analoghe anche nelle espressioni tra matrici diamo la seguente

Definizione 6.37. La **somma di due matrici** $A, B \in M_{m,n}(\mathbb{K})$, denotata con $A + B$, è la matrice di $M_{m,n}(\mathbb{K})$ la cui entrata di posto ij è la somma delle entrate di posto ij di A e B, ovvero

$$(A + B)_{ij} = A_{ij} + B_{ij}.$$

Esempio 6.38. Consideriamo le matrici

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 & -2 \\ 3 & 1 & 1 \end{pmatrix}.$$

Allora

$$A + B = \begin{pmatrix} 1+1 & 2+0 & 3+(-2) \\ 4+3 & 5+1 & 6+1 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 1 \\ 7 & 6 & 7 \end{pmatrix}.$$

La somma di matrici così definita ha, tra le altre, proprio le proprietà desiderate citate sopra. Date matrici $A, B, C \in M_{m,n}(\mathbb{K})$ e $D, E \in M_{n,p}(\mathbb{K})$ si ha

(1) Proprietà commutativa: $A + B = B + A$

(2) Proprietà associativa: $(A + B) + C = A + (B + C)$

(3) Esistenza dell'elemento neutro: la matrice 0 le cui entrate sono nulle soddisfa

$$A + 0 = 0 + A = A.$$

In ogni insieme $M_{m,n}(\mathbb{K})$, chiameremo 0 la *matrice nulla*.

(4) Esistenza dell'inverso additivo: ogni matrice A ha un'inversa additiva o opposta, data dalla matrice $-A$ che ha come entrata di posto ij l'opposto dell'entrata di posto ij di A:

$$(-A)_{ij} = -A_{ij}.$$

(5) Proprietà distributiva rispetto al prodotto: si ha $A(D + E) = AD + AE$ e, analogamente, $(A + B)D = AD + BD$.

Quindi, data un'uguaglianza tra matrici $AB = AC$, possiamo eseguire gli stessi passaggi che si eseguono tra numeri: $AB - AC = 0$ e $A(B - C) = 0$.

Osservazione 6.39. Si noti che da un'uguaglianza come l'ultima appena scritta, $A(B - C) = 0$, non possiamo però dedurre $A = 0$ o $B - C = 0$ come faremmo per un'uguaglianza numerica. Questo perché il prodotto di matrici *non* gode della cosiddetta

"proprietà di annullamento del prodotto", ovvero può capitare che il prodotto di due matrici non nulle sia la matrice nulla. Ad esempio

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Ricordiamo che abbiamo già visto altri contesti in cui tale proprietà non vale (ad esempio per il prodotto tra classi nell'aritmetica modulare).

Avendo visto le operazioni di prodotto e somma tra matrici, che sono operazioni *binarie* in quanto ci danno una matrice come risultato di un'operazione tra due matrici, concludiamo vedendo un'operazione *unaria* che ci dà una nuova matrice a partire da una sola data[4], l'operazione di *trasposta di una matrice*.

Definizione 6.40. Data una matrice $A \in M_{m,n}(\mathbb{K})$, definiamo la sua **trasposta** come la matrice $A^T \in M_{n,m}(\mathbb{K})$ che si ottiene da A scambiando le righe con le colonne:

$$A^T_{ij} = A_{ji}.$$

In pratica, in seguito allo scambio gli elementi di A che si trovavano nella riga i e colonna j si vengono a trovare sulla riga j e colonna i.

Esempio 6.41.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \Longrightarrow A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

L'importanza della trasposta consiste nel fatto che il suo utilizzo consente una più agevole manipolazione di alcune espressioni. Ad esempio, il prodotto scalare $x \cdot y = x_1y_1 + x_2y_2 + x_3y_3$ tra due terne (x_1, x_2, x_3) e (y_1, y_2, y_3) può essere riscritto usando il prodotto di matrici e la trasposta. Più precisamente, se consideriamo x e y come matrici con una sola colonna

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

il prodotto scalare $x \cdot y$ coincide con il prodotto righe per colonne tra le matrici

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

[4] Ad esempio, in algebra booleana le operazioni di $\wedge$ e $\vee$ sono operazioni binarie in quanto ci danno una nuova proposizione (rispettivamente $P \wedge Q$ e $P \vee Q$) a partire da due proposizioni P e Q date, mentre $\neg$ è un'operazione unaria che ci dà una nuova proposizione $\neg P$ a partire da una sola proposizione P data.

Questo non è nient'altro che il prodotto righe per colonne $x^T y$ tra la trasposta x^T di x e y. Come vedremo nel prossimo capitolo, scrivere il prodotto scalare in questa forma ha vari vantaggi. Ad esempio, si possono così determinare quali matrici rappresentano trasformazioni rigide dello spazio (ovvero trasformazioni che non modificano lunghezze e angoli). Avremo allora bisogno di conoscere le proprietà della trasposta, tra cui le più importanti sono le seguenti (dimostriamo quelle meno immediate).

(1) $(A+B)^T = A^T + B^T$

(2) $(AB)^T = B^T A^T$

Dimostrazione. Siano $A \in M_{m,n}(\mathbb{K})$ e $B \in M_{n,p}(\mathbb{K})$. L'identità discende dalla definizione di prodotto righe per colonne:

$$(AB)^T_{ij} = (AB)_{ji} = \sum_{k=1}^{n} A_{jk} B_{ki} = \sum_{k=1}^{n} B^T_{ik} A^T_{kj} = (B^T A^T)_{ij}.$$

□

(3) $(A^T)^T = A$

(4) $\det(A^T) = \det(A)$

Dimostrazione. Sia $A \in M_n(\mathbb{K})$ di entrate a_{ij} e denotiamo con a^T_{ij} le entrate di A^T. Allora, direttamente dalla Definizione 6.17, si ha

$$\det(A) = \sum_{\sigma \in S_n} s(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

Notiamo che ogni prodotto può essere riscritto come $a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n}$ semplicemente riordinando rispetto agli indici di colonna. Per cui abbiamo

$$\det(A) = \sum_{\sigma \in S_n} s(\sigma^{-1}) a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n}.$$

Si noti che $s(\sigma) = s(\sigma^{-1})$. Poiché inoltre associare l'inversa, vista come funzione da S_n in S_n, è biiettiva possiamo concludere che

$$\det(A) = \sum_{\sigma \in S_n} s(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} = \sum_{\sigma \in S_n} s(\sigma) a^T_{1\sigma(1)} \cdots a^T_{n\sigma(n)} = \det(A^T).$$

□

Illustriamo con un esempio la (2) allo scopo di mettere in evidenza l'ordine invertito delle trasposte a secondo membro

Esempio 6.42. Siano

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$$

allora si ha

$$AB = \begin{pmatrix} -1 & 1 \\ -1 & 3 \end{pmatrix} \Longrightarrow (AB)^T = \begin{pmatrix} -1 & -1 \\ 1 & 3 \end{pmatrix}.$$

Si ha inoltre

$$A^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \quad B^T = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$$

e quindi, come dimostrato nella (2),

$$B^T A^T = \begin{pmatrix} -1 & -1 \\ 1 & 3 \end{pmatrix}.$$

D'altra parte si vede subito che

$$A^T B^T = \begin{pmatrix} 4 & -1 \\ 6 & -2 \end{pmatrix} \neq (AB)^T.$$

Un'ultima operazione riguardante le matrici, spesso usata nel fare i conti, è il prodotto di un elemento del campo numerico $\mathbb{K}$ per una matrice. Più precisamente, si ha la seguente

Definizione 6.43. Dato $c \in \mathbb{K}$ e una matrice $A \in M_{m,n}(\mathbb{K})$, si definisce il **prodotto di** A **per lo scalare** c come la matrice $cA \in M_{m,n}(\mathbb{K})$ che si ottiene moltiplicando ogni entrata di A per c:

$$(cA)_{ij} = c(A_{ij}).$$

Osservazione 6.44. Abbiamo usato implicitamente tale operazione nella formula (6.33) per l'inversa quando abbiamo moltiplicato la matrice dei cofattori per il numero $\frac{1}{\det(A)}$.

Segnaliamo le seguenti proprietà del prodotto di una matrice per uno scalare le cui dimostrazioni sono una conseguenza immediata delle definizioni.

$$c(A + B) = cA + cB \tag{6.47}$$

$$A(cB) = c(AB) \tag{6.48}$$

Capitolo 7

Autovalori e autovettori

7.1 Definizione, esempi e applicazioni

Definizione 7.1. Sia $A \in M_n(\mathbb{K})$ una matrice quadrata di ordine n a entrate in un campo $\mathbb{K}$. Un elemento $\lambda \in \mathbb{K}$ si dice **autovalore di** A se esiste una n-upla non nulla $0 \neq (x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$ tale che

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \tag{7.1}$$

In tal caso il vettore $(x_1, x_2, \ldots, x_n)$ è detto **autovettore di** A **relativo all'autovalore** λ.

Se x denota la n-upla scritta in colonna, scriveremo la (7.1) in modo più conciso come prodotto di matrici

$$Ax = \lambda x.$$

Un autovettore è quindi una n-upla che viene mandata in un suo multiplo quando le viene applicata la matrice A. Ad esempio, se $A = \left(\begin{smallmatrix} 1 & 3 \\ 3 & 1 \end{smallmatrix}\right)$ e $x = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$, si ha

$$Ax = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 4x$$

e quindi x è un autovettore di A associato all'autovalore $\lambda = 4$.
Per contro il vettore $x = \left(\begin{smallmatrix} 2 \\ 1 \end{smallmatrix}\right)$ non è un autovettore di A in quanto

$$Ax = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \end{pmatrix}$$

e $(5,7)$ non è multiplo di $(2,1)$.
Prima di vedere come si calcolano autovalori e autovettori di una matrice data, citiamo alcune applicazioni in fisica, matematica ed informatica.

(1) Una massa collegata a una molla, se spostata dalla posizione di equilibrio, inizia a oscillare con una certa frequenza, dipendente dalla massa stessa e dalle caratteristiche fisiche della molla. Se invece abbiamo più masse collegate tra loro tramite delle molle, l'oscillazione di ognuna di esse dipende anche dalla posizione e dal movimento delle altre masse del sistema e bisogna quindi tener conto di tutte le interazioni reciproche. Si mostra che tali interazioni possono essere descritte mediante una matrice quadrata, e gli autovalori di questa matrice ci danno le cosiddette *frequenze naturali di oscillazione del sistema*. Il modo in cui il sistema oscilla è dato da una combinazione di queste frequenze tramite una formula che coinvolge anche gli autovettori corrispondenti.

(2) Abbiamo visto nel Capitolo 4 come, fissato un sistema di riferimento nello spazio tridimensionale, una generica equazione $Ax + By + Cz = D$ di primo grado in tre incognite rappresenti un piano, nel senso che i punti appartenenti al piano sono tutti e soli quelli le cui coordinate (x,y,z) verificano l'equazione.

Ci chiediamo ora cosa sia rappresentato da una generica equazione di secondo grado nelle tre incognite date, ovvero

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + L = 0. \tag{7.2}$$

Si può dimostrare che le equazioni di questo tipo rappresentano superfici dette *quadriche* (tra le quali rientrano sfere, ellissoidi, iperboloidi, coni, cilindri) e che, per capire quale tra queste superfici rappresenti l'equazione (7.2), bisogna calcolare gli autovalori di una matrice di ordine 4 costruita opportunamente dai coefficienti $A, B, C, \ldots, L$:

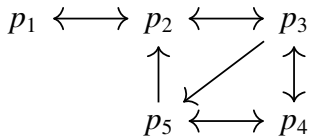
$$\begin{pmatrix} A & D & E & G \\ D & B & F & H \\ E & F & C & I \\ G & H & I & L \end{pmatrix}$$

Inoltre, gli autovettori forniscono informazioni sulla posizione di tale superficie nello spazio, dicendoci ad esempio quali sono i suoi assi di simmetria.

(3) Come ultima, importante applicazione della teoria degli autovalori e autovettori, vediamo che tali nozioni sono alla base dell'algoritmo con cui Google ordina le pagine internet per importanza. Visto l'interesse di tale esempio per l'informatica, diamo qualche dettaglio in più rispetto agli esempi precedenti.

La rete internet consiste di un numero (enorme ma finito) di pagine $p_1, p_2, \ldots, p_N$ ciascuna delle quali può avere uno o più link verso le altre pagine. Ad esempio,

supponiamo che internet sia formato solo da 5 pagine p_1, p_2, p_3, p_4, p_5. Lo schema seguente



rappresenta una rete nella quale la notazione $p_i \to p_j$ significa che c'è un link dalla pagina p_i alla pagina p_j e una doppia freccia $p_i \leftrightarrow p_j$ significa che esiste sia un link da p_i verso p_j che viceversa.

L'idea, abbastanza naturale, dell'algoritmo di Google, è che una pagina p è tanto più importante quanto maggiore è il numero di pagine che hanno un link verso p. Tuttavia i link non hanno tutti lo stesso valore: questo dipenderà a sua volta dall'importanza della pagina da cui proviene.

Per formalizzare quest'idea e dare delle formule rigorose, iniziamo con il denotare con $|p|$ il numero di link che vanno da una pagina p verso altre pagine. Ad esempio, facendo riferimento alla rete di 5 pagine rappresentata nel disegno precedente, si ha

$$|p_1| = 1,\ |p_2| = 2\ |p_3| = 3,\ |p_4| = 2,\ |p_5| = 2.$$

Ora, per ogni pagina p definiamo l'importanza o il *rank* di p come il numero positivo $r(p)$ che soddisfa l'uguaglianza

$$r(p) = \sum_{q \to p} \frac{r(q)}{|q|} \tag{7.3}$$

dove la somma è presa su tutte le pagine q che hanno un link verso p. Tale formula formalizza matematicamente l'idea di importanza di una pagina che abbiamo descritto intuitivamente sopra. Infatti, maggiore è il numero di link che riceve p maggiore sarà il risultato della sommatoria (avendo più addendi). Inoltre ognuno degli addendi, $\frac{r(q)}{|q|}$, è dato dall'importanza $r(q)$ della pagina q (quanto più q è importante, tanto più sarà importante p, verso la quale q ha un link), ma diviso per il numero $|q|$ totale dei link che q ha verso le altre pagine della rete (se q ha link verso più pagine, ciascuna di queste pagine "eredita" solo una frazione dell'importanza di q nel calcolo della propria importanza).

Il punto è che la formula (7.3) definisce l'importanza $r(p)$ di p in funzione delle importanza $r(q)$ di altre pagine per calcolare la quale dovremmo conoscere anche $r(p)$. Questo apparente circolo vizioso si risolve riformulando la definizione (7.3) proprio in termini di autovalori e autovettori.

Più precisamente, costruiamo una matrice A quadrata di ordine N, dove N è il numero delle pagine della rete, e che rappresenti la rete nel modo seguente. Definiamo le entrate a_{ij} di A come

$$a_{ij} = \begin{cases} 0 & \text{se non esiste nessun link da } p_j \text{ a } p_i, \\ \frac{1}{|p_j|} & \text{se esiste un link da } p_j \text{ a } p_i. \end{cases}$$

Ad esempio, si vede che per la rete di 5 pagine rappresentata nel disegno sopra si ha

$$A = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 1 & 0 & \frac{1}{3} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{3} & \frac{1}{2} & 0 \end{pmatrix} \tag{7.4}$$

In pratica, nella j-esima colonna di tale matrice leggiamo verso quali altre pagine ci sono link. Ad esempio, la prima colonna di A ha l'unica entrata non nulla in seconda posizione, il che corrisponde al fatto che p_1 ha un link solo verso p_2; la seconda colonna ha due entrate non nulle, la prima e la terza, che corrisponde al fatto che da p_2 partono due link, uno verso p_1 e un altro verso p_3, e il valore assegnato è $\frac{1}{2}$, e così via.

Ora mostriamo che trovare l'importanza delle pagine equivale a risolvere un problema di autovettori per A. Più precisamente, mostriamo che la (7.3) corrisponde all'uguaglianza

$$Ar = r \tag{7.5}$$

dove

$$r = \begin{pmatrix} r(p_1) \\ r(p_2) \\ \vdots \\ r(p_N) \end{pmatrix}$$

è la n-upla le cui componenti sono i rank delle pagine, che dobbiamo determinare. In altre parole, in base alla (7.5), tale n-upla è un autovettore di A relativo all'autovalore $\lambda = 1$.

Infatti, la (7.5) ci dice che la j-esima componente $(Ar)_j$ di Ar deve essere uguale alla j-esima componente di r, cioè $r(p_j)$. Ma la j-esima componente $(Ar)_j$ del prodotto di A per r si ottiene dal prodotto righe per colonne della j-esima riga di A per r, cioè

$$(Ar)_j = a_{j1}r(p_1) + a_{j2}r(p_2) + \cdots + a_{jN}r(p_N) = \sum_{k=1}^{N} a_{jk}r(p_k). \tag{7.6}$$

Tuttavia, in base alla definizione della matrice A, in tale sommatoria un'entrata a_{jk} è diversa da zero solo se esiste un link da p_k verso p_j. Quindi, tenendo conto solo degli addendi non nulli, possiamo riscrivere la (7.6) come

$$(Ar)_j = \sum_{p_k \to p_j} a_{jk} r(p_k).$$

Ma allora, ricordando che, se esiste un link da p_k verso p_j allora $a_{jk} = \frac{1}{|p_k|}$, otteniamo

$$(Ar)_j = \sum_{p_k \to p_j} \frac{1}{|p_k|} r(p_k).$$

Quindi per la formula (7.3), si vede che il secondo membro di quest'ultima uguaglianza è esattamente $r(p_j)$, il che mostra che $(Ar)_j = r(p_j)$, come volevamo.

Ad esempio, per la nostra rete di 5 pagine, rappresentata dalla matrice A data in (7.4), si può dimostrare che

$$r = (r(p_1), r(p_2), r(p_3), r(p_4), r(p_5)) = (\frac{2}{13}, \frac{4}{13}, \frac{3}{13}, \frac{2}{13}, \frac{2}{13})$$

il che ci dice in particolare che p_2 è la pagina più importante (il che è ragionevole in quanto è la pagina che riceve più link), seguita da p_3, che riceve 2 link esattamente come p_4 e p_5, ma uno dei suoi link viene da p_2 che è la pagina più importante.

Come vedremo nel prossimo paragrafo, non è detto che una matrice ammetta l'autovalore $\lambda = 1$, quindi potrebbe non esistere un vettore r che soddisfi $Ar = r$ le cui componenti siano date dalle importanza delle pagine così come le abbiamo definite tramite la (7.3). In tal caso, la matrice A che descrive la rete va leggermente modificata in modo che tale ipotesi sia soddisfatta. Un ulteriore problema è quello relativo all'ordine della matrice, uguale al numero delle pagine della rete, che nella realtà è un numero enorme, il che rende praticamente impossibile il calcolo esatto di autovalori e autovettori. Si ricorre quindi a soluzioni approssimate.[1]

Osservazione 7.2. Gli esempi precedenti ci danno un'ulteriore conferma delle numerose applicazioni che le matrici hanno nei più svariati contesti: non solo esse possono rappresentare sistemi di equazioni o trasformazioni geometriche, come abbiamo visto nei capitoli precedenti, ma, come si vede sia dall'esempio dei sistemi vibranti che da quello di Google, sono lo strumento matematico ideale per rappresentare sistemi costituiti da un numero finito di elementi che interagiscono tra loro in qualche modo.

[1]Per ulteriori dettagli rimandiamo alla lettura dell'articolo originale di Larry Page e Sergei Brin, fondatori di Google: *The Anatomy of a Large-Scale Hypertextual Web Search Engine* (1998) Seventh International World-Wide Web Conference (WWW 1998), April 14-18, 1998, Brisbane, Australia.

7.2 Calcolo di autovalori e autovettori

Come abbiamo visto nella Definizione 7.1, trovare gli autovettori di una matrice A significa trovare gli x diversi dalla n-upla nulla tali che $Ax = \lambda x$ per qualche $\lambda \in \mathbb{K}$. Vediamo ora come questa uguaglianza può essere riscritta come un sistema di equazioni lineari al quale applicare la teoria vista nei capitoli precedenti.

Come abbiamo visto alla fine del capitolo precedente, sfruttando le proprietà delle matrici, l'uguaglianza $Ax = \lambda x$ può essere riscritta come

$$Ax - \lambda x = 0 \tag{7.7}$$

dove stiamo indicando con 0 la n-upla nulla, scritta in colonna. Ora, per poter mettere in evidenza x nella (7.7), sfruttando la proprietà distributiva $AB + AC = A(B + C)$ del prodotto di matrici, usiamo prima $x = \mathrm{Id}_n x$ ottenendo

$$Ax - \lambda \mathrm{Id}_n x = 0$$

da cui possiamo ora mettere in evidenza

$$(A - \lambda \mathrm{Id}_n)x = 0. \tag{7.8}$$

Questa uguaglianza rappresenta un sistema omogeneo di n equazioni lineari in n incognite (dove n è l'ordine della matrice A), e trovare gli autovettori x di A equivale a trovare le soluzioni $x \neq 0$ di tale sistema.

Ricordiamo che un sistema omogeneo è sempre compatibile e ammette almeno la soluzione nulla. Il sistema (7.8) con n incognite ammette solo la soluzione nulla se e solo se il rango della sua matrice dei coefficienti è uguale a n.

Dal momento che siamo interessati all'esistenza di altre soluzioni oltre a quella nulla, la condizione necessaria e sufficiente per l'esistenza di tali soluzioni è quindi che il rango della matrice dei coefficienti $A - \lambda \mathrm{Id}_n$ del sistema (7.8) sia minore di n.

Tale sistema dipende però da λ e esso ammetterà soluzioni non nulle solo per certi valori di λ. Quello che faremo sarà quindi prima determinare i $\lambda \in \mathbb{K}$ per cui tali soluzioni esistono, e poi risolvere il sistema solo in corrispondenza di quei valori. Ma, come abbiamo visto nel capitolo precedente, una matrice quadrata di ordine n ha rango minore di n se e solo se il suo determinante si annulla. In conclusione, gli autovalori di A sono i $\lambda \in \mathbb{K}$ che soddisfano

$$\det(A - \lambda \mathrm{Id}_n) = 0 \tag{7.9}$$

e gli autovettori di A si trovano come soluzioni del sistema (7.8) per i rispettivi autovalori.

Definizione 7.3. Come vedremo negli esempi, il determinante della matrice $A - \lambda \mathrm{Id}_n$ è un polinomio di grado n (uguale all'ordine della matrice) in λ che si chiama **polinomio caratteristico di** A, e l'equazione (7.9) si dice **equazione caratteristica**.

Osservazione 7.4. Si noti che

$$A - \lambda \mathrm{Id}_n = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} - \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix} = \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix}$$

ovvero $A - \lambda \mathrm{Id}_n$ è semplicemente la matrice A in cui abbiamo sottratto λ alle entrate della diagonale.

Esempio 7.5. Determiniamo autovalori e autovettori della seguente matrice a entrate reali

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}.$$

Iniziamo col calcolare il polinomio caratteristico di A. Si ha

$$\det(A - \lambda \mathrm{Id}_2) = \det \begin{pmatrix} 1 - \lambda & 3 \\ 3 & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2 - 9 = \lambda^2 - 2\lambda - 8.$$

L'equazione caratteristica, le cui soluzioni reali sono gli autovalori di A, è quindi

$$\lambda^2 - 2\lambda - 8 = 0.$$

Applicando la formula risolutiva

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

per la generica equazione di secondo grado $ax^2 + bx + c = 0$, si trova subito $\lambda = 4$ e $\lambda = -2$, che sono quindi gli autovalori di A.
Per trovare gli autovettori relativi a $\lambda = 4$, basta risolvere il sistema omogeneo che ha $A - 4\mathrm{Id}_2$ come matrice dei coefficienti. Analogamente, per trovare gli autovettori relativi a $\lambda = -2$, basta risolvere il sistema omogeneo che ha $A + 2\mathrm{Id}_2$ come matrice dei coefficienti.

Per $\lambda = 4$ si ha $A - \lambda \mathrm{Id}_2 = \left(\begin{smallmatrix} -3 & 3 \\ 3 & -3 \end{smallmatrix}\right)$. Come si vede, le righe rappresentano le equazioni equivalenti (una opposta dell'altra) $-3x_1 + 3x_2 = 0$ e $3x_1 - 3x_2 = 0$. Dividendo per 3, si ha quindi l'unica equazione $-x_1 + x_2 = 0$. Posto $x_2 = t$, si trova $x_1 = x_2 = t$. Quindi le soluzioni del sistema $(A - 4\mathrm{Id}_2)x = 0$, ovvero gli autovettori di A relativi all'autovalore 4, sono esattamente tutti i vettori del tipo (t, t), al variare di $t \in \mathbb{R}$. Ad esempio, per $t = 1$ si ritrova proprio l'autovettore visto subito dopo la Definizione 7.1.

Analogamente, per $\lambda = -2$ si ha $A - \lambda \mathrm{Id}_2 = \left(\begin{smallmatrix} 3 & 3 \\ 3 & 3 \end{smallmatrix}\right)$, quindi le righe rappresentano la stessa equazione $3x_1 + 3x_2 = 0$. Dividendo per 3 e posto $x_2 = t$, si trova $x_1 = -x_2 = -t$. Quindi le soluzioni del sistema $(A + 2\mathrm{Id}_2)x = 0$, ovvero gli autovettori di A relativi all'autovalore -2, sono esattamente tutti i vettori del tipo $(-t, t)$, al variare di $t \in \mathbb{R}$.

Esempio 7.6. Sia

$$A = \begin{pmatrix} 5 & 1 & -1 \\ 4 & 5 & -2 \\ 2 & 1 & 2 \end{pmatrix}.$$

Per calcolare il polinomio caratteristico svolgiamo

$$\det(A - \lambda \mathrm{Id}_3) = \det \begin{pmatrix} 5-\lambda & 1 & -1 \\ 4 & 5-\lambda & -2 \\ 2 & 1 & 2-\lambda \end{pmatrix} =$$

(sviluppando secondo Laplace rispetto alla prima riga)

$$= (5-\lambda) \det \begin{pmatrix} 5-\lambda & -2 \\ 1 & 2-\lambda \end{pmatrix} - \det \begin{pmatrix} 4 & -2 \\ 2 & 2-\lambda \end{pmatrix} - \det \begin{pmatrix} 4 & 5-\lambda \\ 2 & 1 \end{pmatrix} =$$

$$(5-\lambda)[(5-\lambda)(2-\lambda)+2] - [4(2-\lambda)+4] - [4-2(5-\lambda)]$$

ovvero, svolgendo i calcoli,

$$-\lambda^3 + 12\lambda^2 - 45\lambda + 54.$$

Si noti che, come abbiamo anticipato sopra, il polinomio caratteristico è un polinomio di grado uguale all'ordine della matrice, ovvero 3.
Per trovare le *radici* del polinomio caratteristico, ovvero le soluzioni dell'equazione caratteristica

$$-\lambda^3 + 12\lambda^2 - 45\lambda + 54 = 0 \tag{7.10}$$

possiamo usare il seguente risultato algebrico.

Lemma 7.7. *Se un polinomio a coefficienti interi ammette una radice razionale, questa è necessariamente della forma* $\frac{a}{b}$ *dove* a *è un divisore del termine noto e* b *è un divisore del coefficiente del termine di grado massimo.*

Nel caso della nostra equazione (7.10), essendo -1 il coefficiente del termine di grado massimo (e non avendo quindi altri divisori oltre a ± 1) le eventuali soluzioni razionali, se esistono, sono tra i divisori del termine noto 54, ovvero ± 1, ± 2, ± 3, ± 6, ± 9, ± 18, ± 27, ± 54.
Ad esempio, sostituendo $\lambda = 1$ nella (7.10) si trova

$$-1^3 + 12 \cdot 1^2 - 45 \cdot 1 + 54 = 20 \neq 0$$

e quindi 1 non è soluzione dell'equazione. D'altra parte sostituendo $\lambda = 3$ si ottiene

$$-3^3 + 12 \cdot 3^2 - 45 \cdot 3 + 54 = -27 + 108 - 135 + 54 = 0.$$

Quindi $\lambda = 3$ è soluzione dell'equazione caratteristica (ed è quindi un autovalore di A). Una volta trovata una soluzione di un'equazione polinomiale, le altre possono essere trovate grazie al seguente risultato, che ci permette di ridurre l'equazione a una di grado più basso.

Teorema 7.8 (Teorema di Ruffini). *Se un polinomio P di grado n in λ ammette una radice λ_0, allora esso è divisibile per $\lambda - \lambda_0$. Ovvero P si decompone come prodotto $P = (\lambda - \lambda_0)P'$ dove P' è un polinomio di grado $n-1$ in λ.*

A questo punto, l'equazione $P = 0$ equivale a $(\lambda - \lambda_0)P' = 0$. Questa, per la legge di annullamento del prodotto, può essere verificata solo in due casi: $\lambda - \lambda_0 = 0$, cioè $\lambda = \lambda_0$ (che ci dà la soluzione che già conoscevamo) oppure $P' = 0$. Ci siamo quindi ridotti a risolvere un'equazione di grado più basso rispetto a quella iniziale.
Per applicare nella pratica questo metodo, bisogna conoscere un modo per trovare il fattore P' nella decomposizione $P = (\lambda - \lambda_0)P'$. A questo scopo, si usa un algoritmo che descriviamo tramite l'esempio della nostra equazione (7.10), per cui $P = -\lambda^3 + 12\lambda^2 - 45\lambda + 54$ e $\lambda_0 = 3$. Si inizia riportando, come nello schema seguente, i coefficienti che moltiplicano i monomi che compongono il polinomio (da quello di grado più alto a quello di grado più basso) e in basso a sinistra la radice già trovata, ovvero $\lambda = 3$:

	-1	12	-45	54
3				

Ora "abbassiamo", riportandolo nella riga in basso, il primo coefficiente del polinomio

	-1	12	-45	54
3				
	-1			

e moltiplichiamo il -1 così abbassato per la radice 3 e riportiamo il risultato sotto il secondo coefficiente del polinomio

	-1	12	-45	54
3		-3		
	-1			

Ora, sommiamo 12 e -3 e riportiamo il risultato nella riga in basso

	-1	12	-45	54
3		-3		
	-1	9		

A questo punto iteriamo il procedimento. Così come prima abbiamo moltiplicato -1 per la radice $\lambda = 3$, ora moltiplichiamo il 9 appena aggiunto nella riga in basso per $\lambda = 3$, scriviamo il risultato nella colonna successiva sotto il -45, sommiamo e riportiamo nella riga in basso:

	-1	12	-45	54
3		-3	27	
	-1	9	-18	

Come prevede l'algoritmo iteriamo il procedimento: moltiplichiamo il -18 appena aggiunto nella riga in basso per $\lambda = 3$, scriviamo il risultato nella colonna successiva sotto il 54, sommiamo e riportiamo nella riga in basso.

	-1	12	-45	54
3		-3	27	- 54
	-1	9	-18	0

L'ultimo 0 riportato è la conferma del fatto che abbiamo svolto i calcoli correttamente, e i tre coefficienti $-1, 9, -18$ ottenuti prima dello zero sono proprio i coefficienti del polinomio P' (di grado 2) tale che $P = (\lambda - 3)P'$, ordinati dal termine di grado più alto fino al termine noto. Possiamo quindi scrivere la decomposizione in fattori di P come

$$-\lambda^3 + 12\lambda^2 - 45\lambda + 54 = (\lambda - 3)(-\lambda^2 + 9\lambda - 18).$$

A questo punto, come abbiamo detto sopra, l'equazione caratteristica ha come soluzioni, oltre a $\lambda = 3$, anche i λ che annullano l'altro fattore, ovvero tali che $-\lambda^2 + 9\lambda - 18 = 0$. Risolvendo con la formula risolutiva per le equazioni di secondo grado si trova che

questa equazione ha come soluzioni $\lambda = 3$ e $\lambda = 6$.
Quindi gli autovalori della matrice A sono $\lambda = 3$ e $\lambda = 6$, con $\lambda = 3$ che si ripete due volte come soluzione dell'equazione.

Possiamo ora determinare gli autovettori corrispondenti. Come abbiamo visto nella (7.8), dobbiamo risolvere il sistema omogeneo che ha $A - \lambda \mathrm{Id}_3$ come matrice dei coefficienti, prima con $\lambda = 3$ e poi con $\lambda = 6$. Per $\lambda = 3$ si ha

$$A - \lambda \mathrm{Id}_3 = \begin{pmatrix} 2 & 1 & -1 \\ 4 & 2 & -2 \\ 2 & 1 & -1 \end{pmatrix}.$$

Riducendo tale matrice a gradini[2] si trova subito

$$\begin{pmatrix} 2 & 1 & -1 \\ 4 & 2 & -2 \\ 2 & 1 & -1 \end{pmatrix} \underset{\substack{R_2 \to R_2 - 2R_1 \\ R_3 \to R_3 - R_1}}{\longrightarrow} \begin{pmatrix} 2 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

e quindi il sistema $(A - \lambda \mathrm{Id}_3)x = 0$ si riduce all'unica equazione $2x_1 + x_2 - x_3 = 0$. Posti allora $x_2 = t$ e $x_3 = s$, si trova $2x_1 = -t + s$, ovvero $x_1 = \frac{-t+s}{2}$. Le soluzioni del sistema, ovvero gli autovettori di A relativi all'autovalore $\lambda = 3$, sono quindi tutte le terne del tipo

$$\left(\frac{-t+s}{2}, t, s\right)$$

al variare di $t, s \in \mathbb{R}$.
Analogamente, per $\lambda = 6$ si ha

$$A - \lambda \mathrm{Id}_3 = \begin{pmatrix} -1 & 1 & -1 \\ 4 & -1 & -2 \\ 2 & 1 & -4 \end{pmatrix}.$$

Riducendo tale matrice a gradini si trova

$$\begin{pmatrix} -1 & 1 & -1 \\ 4 & -1 & -2 \\ 2 & 1 & -4 \end{pmatrix} \underset{\substack{R_2 \to R_2 + 4R_1 \\ R_3 \to R_3 + 2R_1}}{\longrightarrow} \begin{pmatrix} -1 & 1 & -1 \\ 0 & 3 & -6 \\ 0 & 3 & -6 \end{pmatrix} \underset{R_3 \to R_3 - R_2}{\longrightarrow} \begin{pmatrix} -1 & 1 & -1 \\ 0 & 3 & -6 \\ 0 & 0 & 0 \end{pmatrix}$$

Dividendo inoltre la seconda riga per 3, vediamo che il sistema $(A - \lambda \mathrm{Id}_3)x = 0$ si riduce a

$$\begin{cases} -x_1 + x_2 - x_3 = 0 \\ x_2 - 2x_3 = 0 \end{cases}$$

[2] Ma basta anche osservare che la seconda e la terza riga sono dipendenti dalla prima.

Posto $x_3 = t$, dalla seconda equazione si trova $x_2 = 2t$, e sostituendo nella prima $-x_1 + 2t - t = 0$, ovvero $x_1 = t$. Le soluzioni del sistema, ovvero gli autovettori di A relativi all'autovalore $\lambda = 6$, sono quindi tutte le terne del tipo

$$(t, 2t, t)$$

al variare di $t \in \mathbb{R}$.

Questo esempio mostra che un autovalore λ_0 può presentarsi più volte come soluzione dell'equazione caratteristica e che gli autovettori relativi ad esso possono dipendere da più parametri.

Definizione 7.9. Sia A una matrice quadrata di ordine n e λ_0 un suo autovalore. La **molteplicità algebrica di** λ_0 è il numero di volte che $(\lambda - \lambda_0)$ divide il polinomio caratteristico $P = \det(A - \lambda \mathrm{Id}_n)$. Si chiama invece **molteplicità geometrica di** λ_0 il numero di parametri indipendenti che descrivono l'insieme dei suoi autovettori.

Nell'Esempio 7.6 $\lambda = 3$ ha molteplicità sia algebrica che geometrica pari a 2, mentre l'autovalore $\lambda = 6$ ha molteplicità algebrica e geometrica entrambe uguali a 1.
In generale, non è detto che molteplicità algebrica e geometrica di un autovalore siano uguali. Si può anzi dimostrare il seguente risultato.

Teorema 7.10. *La molteplicità geometrica di un autovalore è sempre minore o uguale alla sua molteplicità algebrica.*

Osservazione 7.11. La molteplicità geometrica di un autovalore λ_0 è in effetti uguale a $n - r$, dove n è l'ordine della matrice e r è il rango della matrice $A - \lambda_0 \mathrm{Id}_n$. Infatti, come sappiamo dalla Proposizione 5.17 il sistema $(A - \lambda_0 \mathrm{Id}_n)x = 0$ ha proprio ∞^{n-r} soluzioni, ovvero le sue soluzioni (che ci danno gli autovettori di A relativi all'autovalore λ_0) si scrivono in funzione di $n - r$ parametri. La molteplicità geometrica di un autovalore è quindi sempre maggiore o uguale a 1. In particolare se un autovalore ha molteplicità algebrica 1, allora, per il Teorema 7.10, anche la sua molteplicità geometrica sarà 1.

Esempio 7.12. Sia

$$A = \begin{pmatrix} 3 & \frac{1}{2} & 2 \\ -1 & 1 & -1 \\ -1 & -\frac{1}{2} & 0 \end{pmatrix}.$$

Il polinomio caratteristico è dato da

$$\det(A - \lambda \mathrm{Id}_3) = \det \begin{pmatrix} 3-\lambda & \frac{1}{2} & 2 \\ -1 & 1-\lambda & -1 \\ -1 & -\frac{1}{2} & -\lambda \end{pmatrix} =$$

sviluppando secondo Laplace rispetto alla prima riga

$$= (3-\lambda)\det\begin{pmatrix} 1-\lambda & -1 \\ -\frac{1}{2} & -\lambda \end{pmatrix} - \frac{1}{2}\det\begin{pmatrix} -1 & -1 \\ -1 & -\lambda \end{pmatrix} + 2\det\begin{pmatrix} -1 & 1-\lambda \\ -1 & -\frac{1}{2} \end{pmatrix} =$$

e svolgendo i calcoli si ottiene

$$-\lambda^3 + 4\lambda^2 - 5\lambda + 2$$

Per il Lemma 7.7 le eventuali radici razionali vanno cercate tra i divisori del termine noto fratto i divisori del termine di grado massimo, quindi $\pm 1, \pm 2$.

Sostituendo $\lambda = 1$ si trova

$$-1 + 4 + 2 - 5 = 0$$

ovvero $\lambda = 1$ è soluzione dell'equazione caratteristica (ed è quindi un autovalore di A). Per determinare le altre soluzioni, applichiamo l'algoritmo descritto nell'Esempio 7.6. Disponiamo nello schema già visto i coefficienti che moltiplicano i monomi che compongono il polinomio, da quello di grado più alto a quello di grado più basso, e in basso a sinistra la radice già trovata, ovvero $\lambda = 1$:

	-1	4	-5	2
1				

Riportiamo nella riga in basso il primo coefficiente del polinomio

	-1	4	-5	2
1				
	-1			

Moltiplichiamo il -1 così abbassato per la radice 1 e riportiamo il risultato sotto il secondo coefficiente del polinomio

	-1	4	-5	2
1		-1		
	-1			

Ora, sommiamo 4 e -1 e riportiamo il risultato nella riga in basso

	-1	4	-5	2
1		-1		
	-1	3		

Continuiamo ad applicare l'algoritmo come descritto nell'esempio precedente:

	-1	4	-5	2
1		-1		
	-1	3		

$\longrightarrow$

	-1	4	-5	2
1		-1	3	
	-1	3		

$\longrightarrow$

$\longrightarrow$

	-1	4	-5	2
1		-1	3	-2
	-1	3	-2	0

Come abbiamo detto, l'ultimo 0 comparso è la conferma del fatto che abbiamo svolto i calcoli correttamente. I tre coefficienti $-1, 3, -2$ ottenuti prima dello zero sono proprio i coefficienti del polinomio P' di grado 2 tale che $P = (\lambda - 1)P'$, ordinati dal termine di grado più alto fino al termine noto. Si ha quindi

$$-\lambda^3 + 4\lambda^2 - 5\lambda + 2 = (\lambda - 1)(-\lambda^2 + 3\lambda - 2).$$

A questo punto per trovare le altre radici del polinomio dobbiamo risolvere $-\lambda^2 + 3\lambda - 2 = 0$. Utilizzando la formula risolutiva per le equazioni di secondo grado si trovano le soluzioni $\lambda = 1$ e $\lambda = 2$.

Concludiamo che gli autovalori della matrice A sono $\lambda = 2$ e $\lambda = 1$, con quest'ultimo che si ripete quindi due volte come soluzione dell'equazione (in base alla Definizione 7.9 esso ha quindi molteplicità algebrica 2).

Possiamo ora determinare gli autovettori corrispondenti. Come abbiamo visto nella (7.8), dobbiamo risolvere il sistema omogeneo che ha $A - \lambda \mathrm{Id}_3$ come matrice dei coefficienti, prima con $\lambda = 1$ e poi con $\lambda = 2$.

Per $\lambda = 1$ si ha

$$A - \mathrm{Id}_3 = \begin{pmatrix} 2 & \frac{1}{2} & 2 \\ -1 & 0 & -1 \\ -1 & -\frac{1}{2} & -1 \end{pmatrix}.$$

Riducendo tale matrice a gradini si trova subito

$$\begin{pmatrix} 2 & \frac{1}{2} & 2 \\ -1 & 0 & -1 \\ -1 & -\frac{1}{2} & -1 \end{pmatrix} \underset{\substack{R_2 \to 2R_2+R_1 \\ R_3 \to 2R_3+R_1}}{\longrightarrow} \begin{pmatrix} 2 & \frac{1}{2} & 2 \\ 0 & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 \end{pmatrix} \underset{R_3 \to R_3+R_2}{\longrightarrow} \begin{pmatrix} 2 & \frac{1}{2} & 2 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Quindi il sistema $(A - \mathrm{Id}_3)x = 0$ si riduce a

$$\begin{cases} 2x_1 + \frac{1}{2}x_2 + 2x_3 = 0 \\ \frac{1}{2}x_2 = 0 \end{cases}$$

La seconda equazione ci dice che $x_2 = 0$; posto $x_3 = t$, dalla prima equazione si trova $2x_1 + 2t = 0$, ovvero $x_1 = -t$. Le soluzioni del sistema, ovvero gli autovettori di A relativi all'autovalore $\lambda = 1$, sono quindi tutte le terne del tipo

$$(-t, 0, t)$$

al variare di $t \in \mathbb{R}$. Si noti quindi che benché la molteplicità algebrica di $\lambda = 1$ sia uguale a 2, la sua molteplicità geometrica è uguale a 1 (in ogni caso, minore, come prevede il Teorema 7.10).

Completiamo l'esempio ricavando i restanti autovettori: per $\lambda = 2$ si ha invece

$$A - 2\mathrm{Id}_3 = \begin{pmatrix} 1 & \frac{1}{2} & 2 \\ -1 & -1 & -1 \\ -1 & -\frac{1}{2} & -2 \end{pmatrix}.$$

Riduciamo a gradini:

$$\begin{pmatrix} 1 & \frac{1}{2} & 2 \\ -1 & -1 & -1 \\ -1 & -\frac{1}{2} & -2 \end{pmatrix} \underset{\substack{R_2 \to R_2+R_1 \\ R_3 \to R_3+R_1}}{\longrightarrow} \begin{pmatrix} 1 & \frac{1}{2} & 2 \\ 0 & -\frac{1}{2} & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Quindi il sistema $(A - 2\mathrm{Id}_3)x = 0$ si riduce a

$$\begin{cases} x_1 + \frac{1}{2}x_2 + 2x_3 = 0 \\ -\frac{1}{2}x_2 + x_3 = 0 \end{cases}.$$

Posto $x_3 = t$, dalla seconda equazione si trova $-\frac{1}{2}x_2 + t = 0$, ovvero $x_2 = 2t$; sostituendo nella prima, si ha $x_1 + t + 2t = 0$, da cui $x_1 = -3t$. Le soluzioni del sistema, ovvero gli autovettori di A relativi all'autovalore $\lambda = 2$, sono quindi tutte le terne del tipo

$$(-3t, 2t, t)$$

al variare di $t \in \mathbb{R}$.

Vediamo ora un ultimo esempio di una matrice di ordine 3 in cui tutte le molteplicità algebriche sono uguali a 1 (e quindi anche tutte le molteplicità geometriche, si veda l'Osservazione 7.11).

Esempio 7.13. Sia

$$A = \begin{pmatrix} 4 & 3 & -3 \\ -3 & -4 & 3 \\ -1 & -3 & 2 \end{pmatrix}.$$

Calcoliamo il polinomio caratteristico. Dopo un calcolo diretto si ha

$$\det(A - \lambda \mathrm{Id}_3) = \det \begin{pmatrix} 4-\lambda & 3 & -3 \\ -3 & -4-\lambda & 3 \\ -1 & -3 & 2-\lambda \end{pmatrix} = -\lambda^3 + 2\lambda^2 + \lambda - 2$$

Come si vede subito con una sostituzione, $\lambda = 1$ è radice del polinomio.

Applichiamo l'algoritmo visto negli esempi precedenti per ottenere le altre radici:

	-1	2	1	-2
1				

$\longrightarrow$

	-1	2	1	-2
1				
	-1			

$\longrightarrow$

	-1	2	1	-2
1		-1		
	-1	1		

$\longrightarrow$

	-1	2	1	-2
1		-1	1	
	-1	1	2	

$\longrightarrow$

	-1	2	1	-2
1		-1	1	2
	-1	1	2	0

Quindi

$$-\lambda^3 + 2\lambda^2 + \lambda - 2 = (\lambda - 1)(-\lambda^2 + \lambda + 2)$$

Poiché, come si verifica subito mediante la formula risolutiva, le soluzioni di $-\lambda^2 + 3\lambda - 2 = 0$ sono $\lambda = 2$ e $\lambda = -1$, concludiamo che gli autovalori della matrice A sono $\lambda = 1$, $\lambda = 2$ e $\lambda = -1$, tutti e tre con molteplicità algebrica 1.

Concludiamo allora l'esercizio calcolando gli autovettori corrispondenti. Per $\lambda = 1$ si ha

$$A - \mathrm{Id}_3 = \begin{pmatrix} 3 & 3 & -3 \\ -3 & -5 & 3 \\ -1 & -3 & 1 \end{pmatrix} \underset{\substack{R_2 \to R_2 + R_1 \\ R_3 \to 3R_3 + R_1}}{\longrightarrow} \begin{pmatrix} 3 & 3 & -3 \\ 0 & -2 & 0 \\ 0 & -6 & 0 \end{pmatrix} \underset{R_3 \to R_3 - 3R_2}{\longrightarrow} \begin{pmatrix} 3 & 3 & -3 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Dividendo anche la prima riga per 3, si vede che il sistema $(A - \mathrm{Id}_3)x = 0$ si riduce a

$$\begin{cases} x_1 + x_2 - x_3 = 0 \\ -2x_2 = 0 \end{cases}.$$

La seconda equazione ci dice che $x_2 = 0$; posto $x_3 = t$, dalla prima equazione si trova $x_1 - t = 0$, ovvero $x_1 = t$. Le soluzioni del sistema, ovvero gli autovettori di A relativi all'autovalore $\lambda = 1$, sono quindi tutte le terne del tipo

$$(t, 0, t), \quad t \in \mathbb{R}.$$

Per $\lambda = 2$ si ha

$$A - 2\mathrm{Id}_3 = \begin{pmatrix} 2 & 3 & -3 \\ -3 & -6 & 3 \\ -1 & -3 & 0 \end{pmatrix} \underset{\substack{R_2 \to 2R_2 + 3R_1 \\ R_3 \to 2R_3 + R_1}}{\longrightarrow} \begin{pmatrix} 2 & 3 & -3 \\ 0 & -3 & -3 \\ 0 & -3 & -3 \end{pmatrix} \underset{R_3 \to R_3 - R_2}{\longrightarrow} \begin{pmatrix} 2 & 3 & -3 \\ 0 & -3 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

Dividendo anche la seconda riga per -3, si vede che il sistema $(A - 2\mathrm{Id}_3)x = 0$ si riduce a

$$\begin{cases} 2x_1 + 3x_2 - 3x_3 = 0 \\ x_2 + x_3 = 0 \end{cases}$$

posto $x_3 = t$, la seconda equazione dice che $x_2 = -t$, e sostituendo nella prima si trova $2x_1 - 3t - 3t = 0$, da cui $x_1 = 3t$. Le soluzioni del sistema, ovvero gli autovettori di A relativi all'autovalore $\lambda = 2$, sono quindi tutte le terne del tipo

$$(3t, -t, t), \quad t \in \mathbb{R}.$$

Infine, per $\lambda = -1$ si ha

$$A + \mathrm{Id}_3 = \begin{pmatrix} 5 & 3 & -3 \\ -3 & -3 & 3 \\ -1 & -3 & 3 \end{pmatrix} \underset{\substack{R_2 \to 5R_2 + 3R_1 \\ R_3 \to 5R_3 + R_1}}{\longrightarrow} \begin{pmatrix} 5 & 3 & -3 \\ 0 & -6 & 6 \\ 0 & -12 & 12 \end{pmatrix} \underset{R_3 \to R_3 - 2R_2}{\longrightarrow} \begin{pmatrix} 5 & 3 & -3 \\ 0 & -6 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

Dividendo anche la seconda riga per 6, si vede che il sistema $(A + \mathrm{Id}_3)x = 0$ si riduce a

$$\begin{cases} 5x_1 + 3x_2 - 3x_3 = 0 \\ -x_2 + x_3 = 0 \end{cases}$$

posto $x_3 = t$, la seconda equazione dice che $x_2 = t$, e sostituendo nella prima si trova $5x_1 + 3t - 3t = 0$, da cui $x_1 = 0$. Le soluzioni del sistema, ovvero gli autovettori di A

relativi all'autovalore $\lambda = -1$, sono quindi tutte le terne del tipo

$$(0, t, t)\,, \quad t \in \mathbb{R}.$$

Osservazione 7.14. Data una matrice le cui entrate appartengono a un certo campo, può capitare che le soluzioni dell'equazione caratteristica siano fuori da quel campo. Ad esempio, la matrice a entrate reali

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

ha come equazione caratteristica

$$\det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1 = 0,$$

che non ha soluzioni nel campo reale $\mathbb{R}$, mentre nel campo complesso $\mathbb{C}$ ha le due soluzioni $+i$ e $-i$.

Quindi, *vista come matrice reale* A non ha autovalori e quindi neanche autovettori, mentre vista come matrice complessa (i numeri reali sono particolari numeri complessi) A ha i due autovalori i e $-i$, e possiamo determinarne i corrispondenti autovettori, che saranno elementi di $\mathbb{C}^2$.

Più in dettaglio, per $\lambda = i$ si ha

$$A - i\mathrm{Id}_2 = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \underset{R_2 \to iR_2 - R_1}{\longrightarrow} \begin{pmatrix} -i & 1 \\ 0 & 0 \end{pmatrix}$$

quindi gli autovettori sono le soluzioni dell'unica equazione $-ix_1 + x_2 = 0$. Ponendo $x_2 = t$ (dove t stavolta varia tra tutti i numeri complessi) si ha $x_1 = \frac{1}{i}t$ e quindi gli autovettori relativi all'autovalore i sono dati da tutte le coppie del tipo $(\frac{1}{i}t, t)$, al variare di $t \in \mathbb{C}$.

Analogamente, per $\lambda = -i$ si ha

$$A + i\mathrm{Id}_2 = \begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} \underset{R_2 \to iR_2 + R_1}{\longrightarrow} \begin{pmatrix} i & 1 \\ 0 & 0 \end{pmatrix}$$

quindi gli autovettori sono le soluzioni dell'unica equazione $ix_1 + x_2 = 0$. Ponendo $x_2 = t$ (dove t stavolta varia tra tutti i numeri complessi) si ha $x_1 = -\frac{1}{i}t$ e quindi gli autovettori relativi all'autovalore $-i$ sono dati da tutte le coppie del tipo $(-\frac{1}{i}t, t)$, al variare di $t \in \mathbb{C}$.

7.3 Matrici diagonalizzabili

La molteplicità geometrica, definita sopra come il numero di parametri che servono per descrivere l'insieme degli autovettori relativi a un dato autovalore, ci dà in effetti il massimo numero di autovettori indipendenti che ogni autovalore fornisce.
Ad esempio, nell'Esempio 7.6, abbiamo visto che gli autovettori di A relativi all'autovalore $\lambda = 3$ sono tutti e soli i vettori della forma $\left(\frac{-t+s}{2}, t, s\right)$. Decomponendo tale vettore come somma della parte che contiene t e di quella che contiene s

$$\left(\frac{-t+s}{2}, t, s\right) = \left(\frac{-t}{2}, t, 0\right) + \left(\frac{s}{2}, 0, s\right) = t\left(\frac{-1}{2}, 1, 0\right) + s\left(\frac{1}{2}, 0, 1\right)$$

vediamo che tutti gli autovettori sono combinazione dei due vettori indipendenti $\left(-\frac{1}{2}, 1, 0\right)$ e $\left(\frac{1}{2}, 0, 1\right)$, quindi non troviamo più di due autovettori indipendenti relativi a $\lambda = 3$.
Analogamente, sempre nell'Esempio 7.6, gli autovettori di A relativi all'autovalore $\lambda = 6$ sono tutte e sole le terne del tipo $(t, 2t, t)$. Poiché

$$(t, 2t, t) = t(1, 2, 1)$$

vediamo che tale autovalore fornisce sostanzialmente un solo autovettore assieme a tutti i suoi multipli (quindi non esistono due autovettori indipendenti relativi a tale autovalore).
In effetti, quanto appena visto vale in generale:

Proposizione 7.15. *Sia A una matrice quadrata di ordine n. Una combinazione lineare*

$$v = \sum_{i=1}^{k} c_i v_i$$

di autovettori di A è un autovalore di A relativo all'autovalore λ se e solo se tutti gli autovettori v_i sono relativi allo stesso autovalore λ. In particolare, autovettori relativi ad autovalori distinti sono indipendenti.

Dimostrazione. Dimostriamo prima la sufficienza combinando le due seguenti proprietà.
Siano $v_1, \ldots, v_k$ autovettori di A relativi all'autovalore λ e sia

$$v = \sum_{i=1}^{k} c_i v_i$$

per certi coefficienti $c_i \in \mathbb{K}$. Allora vogliamo dimostrare che anche v è un autovettore relativo a λ. In effetti si ha

$$A(v) = A\left(\sum_{i=1}^{k} c_i v_i\right) = \sum_{i=1}^{k} c_i A(v_i) = \lambda \sum_{i=1}^{k} c_i v_i = \lambda v \tag{7.11}$$

dove abbiamo usato le proprietà del prodotto di matrici e del prodotto di uno scalare per una matrice viste nel capitolo precedente. Quindi anche v è autovettore di A relativo all'autovalore λ. Per quanto riguarda la necessità, possiamo considerare la somma senza coefficienti e gli autovalori tutti distinti per quanto appena dimostrato. Dobbiamo quindi dimostrare che un autovettore v relativo a λ è una somma

$$v = \sum_{i=1}^{k} v_i$$

di autovettori v_i relativi ad autovalori λ_i solo se $\lambda = \lambda_i$ per ogni i. Procediamo per induzione su k. Quando $k = 1$ si ha

$$v = v_1 \Longrightarrow A(v) = A(v_1) \Longrightarrow \lambda v = \lambda_1 v_1 \Longrightarrow \lambda = \lambda_1.$$

Ora supponiamo che questo valga per la somma di $k-1$ autovettori e dimostriamo che deve valere anche per la somma di k autovettori. Sia dunque per assurdo

$$v = \sum_{i=1}^{k} v_i$$

dove $v, v_1, \ldots, v_k$ sono autovettori relativi agli autovalori $\lambda, \lambda_1, \ldots, \lambda_k$. Allora moltiplicando per A si ottiene

$$\lambda v = \sum_{i=1}^{k} \lambda_i v_i$$

e sostituendo la scrittura di v come somma si ha

$$0 = \sum_{i=1}^{k} (\lambda_i - \lambda) v_i.$$

Poiché per ipotesi induttiva i v_i sono linearmente indipendenti, si deve avere $\lambda_i - \lambda = 0$, e quindi $\lambda_i = \lambda$, per ogni $i = 1, \ldots, k$. □

Quindi, nell'Esempio 7.6, mettendo insieme gli autovettori indipendenti forniti da ogni autovalore otteniamo in tutto tre autovettori indipendenti $\left(\frac{-1}{2}, 1, 0\right)$, $\left(\frac{1}{2}, 0, 1\right)$, $(1, 2, 1)$, tanti quanti l'ordine della matrice.

Definizione 7.16. Le matrici $A \in M_n(\mathbb{K})$ per cui esistono n autovettori indipendenti si dicono **matrici diagonalizzabili**.

Quindi, la matrice dell'Esempio 7.6 è diagonalizzabile. Di contro la matrice dell'Esempio 7.12 non lo è, in quanto come autovettori abbiamo trovato $(-t, 0, t)$ (relativi all'autovalore $\lambda = 1$) e $(-3t, 2t, t)$ (relativi all'autovalore $\lambda = 2$). Poiché $(-t, 0, t) = t(-1, 0, 1)$ e $(-3t, 2t, t) = t(-3, 2, 1)$, ogni autovalore fornisce un solo autovettore indipendente, quindi in tutto troviamo due autovettori indipendenti, mentre la matrice ha ordine tre.

Osservazione 7.17. Si noti che per la Proposizione 7.15 e per la definizione di molteplicità geometrica, una matrice è diagonalizzabile se e solo se la somma delle molteplicità geometriche dei sui autovalori è pari all'ordine della matrice.

Il motivo per cui una matrice con la proprietà detta si dice diagonalizzabile è il seguente. Supponiamo che A abbia ordine n e ammetta n autovettori indipendenti

$$v_1 = (v_{11}, v_{21}, \ldots, v_{n1}),\ v_2 = (v_{12}, v_{22}, \ldots, v_{n2}), \ldots, v_n = (v_{1n}, v_{2n}, \ldots, v_{nn})$$

Disponendo tali autovettori in colonna, si ottiene una matrice

$$V = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix}$$

quadrata di ordine n le cui colonne, che sono gli autovettori trovati, sono per ipotesi indipendenti.
Quindi il suo rango (ricordiamo che il rango si può calcolare indifferentemente per righe o per colonne) è n e per il Teorema 6.11 V è invertibile.
Affermiamo che $V^{-1}AV$ è una matrice diagonale e più precisamente vale l'uguaglianza

$$\begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix}^{-1} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{7.12}$$

cioè

$$V^{-1}AV = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

dove λ_1 è l'autovalore di v_1, λ_2 è l'autovalore di v_2, e così via.
Infatti, moltiplicando entrambi i membri di tale uguaglianza a sinistra per V, vediamo

che essa equivale a

$$AV = V \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{7.13}$$

Per mostrare l'uguaglianza scritta in questa forma, osserviamo che, per definizione di prodotto righe per colonne, la prima colonna del prodotto di matrici al primo membro della (7.13) si ottiene moltiplicando le righe di A per la prima colonna di V, ovvero

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix}$$

Ma, essendo la prima colonna di V l'autovettore v_1 di A, si ha

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix} = \lambda_1 \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{pmatrix} = \begin{pmatrix} \lambda_1 v_{11} \\ \lambda_1 v_{21} \\ \vdots \\ \lambda_1 v_{n1} \end{pmatrix}$$

Facendo lo stesso ragionamento sulle altre colonne del prodotto di matrici al primo membro della (7.13), si ottiene allora

$$AV = \begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} & \cdots & \lambda_n v_{1n} \\ \lambda_1 v_{21} & \lambda_2 v_{22} & \cdots & \lambda_n v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 v_{n1} & \lambda_2 v_{n2} & \cdots & \lambda_n v_{nn} \end{pmatrix} \tag{7.14}$$

Questa, come si verifica subito sempre svolgendo un prodotto righe per colonne, è uguale al secondo membro della (7.13).

Esempio 7.18. Consideriamo la matrice

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

di cui abbiamo calcolato autovalori e autovettori nell'Esempio 7.5, trovando (t,t) come autovettori relativi all'autovalore $\lambda = 4$ e $(-t,t)$ come autovettori relativi all'autovalore $\lambda = -2$. Poiché $(t,t) = t(1,1)$ e $(-t,t) = t(-1,1)$, troviamo due autovettori indipendenti $(1,1)$ e $(-1,1)$ e la matrice è diagonalizzabile. Per verificare la validità della

(7.12), disponiamo tali autovettori in colonna, ottenendo

$$V = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Come si verifica subito, l'inversa di tale matrice è

$$V^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

e si ha

$$V^{-1}AV = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 4 & 2 \\ 4 & -2 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$$

cioè, come previsto dalla (7.12), otteniamo la matrice diagonale che ha sulla diagonale gli autovalori di A.

La diagonalizzazione è fondamentale quando si vuole cambiare sistema di riferimento in modo da passare a un nuovo sistema in cui il problema in esame assume una forma più semplice. Ad esempio, se A rappresenta una superficie quadrica come nel secondo esempio visto nel primo paragrafo, la diagonalizzazione permette di ottenere un nuovo sistema di riferimento in cui l'equazione della superficie è la più semplice possibile.

7.4 Applicazioni alla grafica: rotazioni nello spazio e stitching di immagini

In quest'ultimo paragrafo, accenneremo ad alcune applicazioni di natura grafica della teoria di autovalori e autovettori.

La prima applicazione che vedremo consiste nella determinazione dell'asse e dell'angolo di una rotazione data nello spazio tridimensionale. Prima di entrare nei dettagli, vediamo alcune generalità su tali trasformazioni.

Come abbiamo visto nella Sezione 6.1, fissato nello spazio un sistema di riferimento con origine O, le applicazioni lineari sono descritte mediante una matrice A di ordine 3. Più precisamente se x rappresenta la terna delle coordinate di un punto P (o del corrispondente vettore $\vec{OP}$), le coordinate del punto (o del vettore) trasformato sono date dal prodotto Ax.

Tra le applicazioni lineari che abbiamo visto, rotazioni e riflessioni hanno un'altra caratteristica in comune (che ad esempio le proiezioni non hanno). Queste sono infatti trasformazioni rigide, ovvero non modificano né le lunghezze dei vettori né gli angoli tra i vettori.

Ricordiamo che nel Capitolo 4 abbiamo visto come, se il sistema di riferimento fissato è ortonormale, allora la lunghezza di un vettore $\vec{OP}$ di coordinate $x = (x_1, x_2, x_3)$ è dato da $|x| = \sqrt{x_1^2 + x_2^2 + x_3^2}$, cf. (4.20). Inoltre l'angolo θ tra due vettori $\vec{OP}$ e $\vec{OQ}$ di coordinate rispettivamente $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$ è dato dalla formula

$$\cos(\theta) = \frac{x \cdot y}{|x||y|},$$

dove a numeratore abbiamo il prodotto scalare $x \cdot y = x_1y_1 + x_2y_2 + x_3y_3$, cf. (4.21). Osservando che anche la lunghezza $|x|$ di un vettore si scrive in termini del prodotto scalare, in quanto $\sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{x \cdot x}$, deduciamo che una trasformazione è rigida, ovvero non modifica né lunghezze né angoli, se e solo se non modifica il prodotto scalare.

In formula, dal momento che le coordinate del vettore x dopo la trasformazione sono date da Ax, la trasformazione rappresentata dalla matrice A sarà una trasformazione rigida se e solo se

$$(Ax) \cdot (Ay) = x \cdot y \tag{7.15}$$

qualunque siano le terne x e y.

Ora, come abbiamo visto a pagina 241, il prodotto scalare $x \cdot y$ può essere anche scritto come prodotto $x^T y$ tra la trasposta della colonna formata dalle componenti di x e la colonna delle componenti di y. Questo ci permette di riscrivere la condizione (7.15) come

$$(Ax)^T (Ay) = x^T y. \tag{7.16}$$

Il vantaggio di questa riformulazione è che possiamo applicare la proprietà (2) della trasposta, per cui $(Ax)^T = x^T A^T$. Quindi la (7.16) si riscrive

$$x^T A^T A y = x^T y. \tag{7.17}$$

Ora, questa uguaglianza può essere verificata per ogni x e ogni y solo se il prodotto $A^T A$ che compare al primo membro è uguale alla matrice identica Id_3.

In altre parole, una trasformazione lineare è rigida se e solo se la matrice che la rappresenta rispetto a un sistema di riferimento ortonormale fissato è tale che $A^T A = \mathrm{Id}_3$.

Definizione 7.19. Una matrice quadrata di ordine n a entrate reali $A \in M_n(\mathbb{R})$ si dice **ortogonale** se $A^T A = \mathrm{Id}_n$ (equivalentemente, se $A^{-1} = A^T$).

Osservazione 7.20. In generale, il determinante di una matrice ortogonale può essere

uguale solo a $+1$ o -1. Infatti, dalla condizione $A^TA = \mathrm{Id}_n$ abbiamo che

$$\det(A^TA) = \det(\mathrm{Id}_n) = 1$$

ovvero, applicando il teorema di Binet al primo membro

$$\det(A^T)\det(A) = 1.$$

Poiché, come abbiamo visto nella proprietà (4), il determinante di una matrice è uguale a quello della sua trasposta, l'ultima uguaglianza equivale a

$$\det(A)^2 = 1$$

da cui deduciamo che $\det(A) = \pm 1$, come affermato.

Alla luce di questa definizione, possiamo riassumere quanto detto finora affermando che una trasformazione lineare è rigida se e solo se la matrice che la rappresenta rispetto a un sistema di riferimento ortonormale fissato è una matrice ortogonale.
Quindi la matrice che rappresenta una rotazione è sicuramente ortogonale, ma ovviamente non vale il viceversa. Ad esempio, anche le riflessioni sono rappresentate da matrici ortogonali in quanto sono trasformazioni rigide. Tuttavia, per riconoscere la rotazioni tra le matrici ortogonali basta calcolare un determinante, come afferma il seguente

Teorema 7.21. *Fissato un sistema di riferimento ortonormale nello spazio, una matrice A quadrata di ordine 3 rappresenta una rotazione se e solo se A è ortogonale e* $\det(A) = +1$.

Definizione 7.22. L'insieme delle matrici ortogonali di ordine n si denota con $\mathrm{O}(n)$. Il sottoinsieme di $\mathrm{O}(n)$ costituito dalle matrici ortogonali di ordine n e determinante $+1$ (dette anche **matrici ortogonali speciali**) si denota con $\mathrm{SO}(n)$.

Proposizione 7.23. *Gli insiemi $O(n)$ e $SO(n)$ sono sottogruppi di $GL_n(\mathbb{R})$ detti rispettivamente* **gruppo ortonormale** *e* **gruppo ortonormale speciale** *di ordine n.*

Dimostrazione. Incominciamo col dimostrare che $\mathrm{O}(n)$ è un sottogruppo di $GL_n(\mathbb{R})$ rispetto al prodotto di matrici. Questo discende dalle seguenti proprietà:

(i) Chiaramente $\mathrm{Id} \in \mathrm{O}(n)$ poiché $\mathrm{Id}^T = \mathrm{Id}$ e $\mathrm{Id}\,\mathrm{Id} = \mathrm{Id}$.

(ii) Se $A, B \in \mathrm{O}(n)$ allora anche $AB \in \mathrm{O}(n)$. Questo deriva da

$$(AB)^T(AB) = B^TA^TAB = B^T\mathrm{Id}B = B^TB = \mathrm{Id}$$

dove abbiamo utilizzato la proprietà (2) ed il fatto che $A, B \in \mathrm{O}(n)$.

(iii) Se $A \in \mathrm{O}(n)$ allora $A^{-1} \in \mathrm{O}(n)$. Qua usiamo il fatto che $(A^T)^T = A$ per cui

$$(A^T)^T A^T = AA^T = AA^{-1} = \mathrm{Id}$$

dove la penultima uguaglianza deriva da $A^T = A^{-1}$.

Ora dobbiamo mostrare le stesse proprietà per $\mathrm{SO}(n)$, ossia aggiungere la condizione $\det(A) = \det(B) = 1$ e richiedere $\det(\mathrm{Id}) = \det(AB) = \det(A^{-1}) = 1$.

(i) Chiaramente $\mathrm{Id} \in \mathrm{SO}(n)$ poiché $\det(\mathrm{Id}) = 1$ e $\mathrm{Id} \in \mathrm{O}(n)$.

(ii) Se $A, B \in \mathrm{SO}(n)$ allora anche $AB \in \mathrm{SO}(n)$. Questo deriva da $AB \in \mathrm{O}(n)$ e da

$$\det(AB) = \det(A)\det(B) = 1$$

dove abbiamo utilizzato che il teorema di Binet ed il fatto che $A, B \in \mathrm{SO}(n)$.

(iii) Se $A \in \mathrm{SO}(n)$ allora $A^{-1} \in \mathrm{SO}(n)$. Qua usiamo il fatto che $(A^T) = A^{-1}$ per cui

$$\det(A^{-1}) = \det(A^T) = \det(A) = 1.$$

□

Ora vediamo alcuni esempi concreti di matrici di rotazione.
Iniziamo dal caso del piano, che è il più semplice. In tal caso, fissato un sistema di riferimento ortonormale con origine del sistema data dal punto O, la matriche che rappresenta la rotazione in senso antiorario di angolo θ attorno a O è determinata da come agisce sui vettori della base $O\vec{P}_1$ e $O\vec{P}_2$. In altre parole le entrate della matrice

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

che rappresenta la rotazione sono determinate dal prodotto per i vettori colonna $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ e $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$. Si può facilmente verificare graficamente che

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

da cui ricaviamo $a_{11} = \cos\theta$ e $a_{21} = \sin\theta$. Allo stesso modo da

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

si ottiene $a_{22} = \cos\theta$ e $a_{12} = -\sin\theta$.

Per cui un elemento generico di $\mathrm{SO}(2)$, ovvero una rotazione di un angolo θ in senso antiorario attorno all'origine, è dato da

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{7.18}$$

Ovvero la rotazione è rappresentata in coordinate dalla funzione

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos\theta x_1 - \sin\theta x_2 \\ \sin\theta x_1 + \cos\theta x_2 \end{pmatrix} \tag{7.19}$$

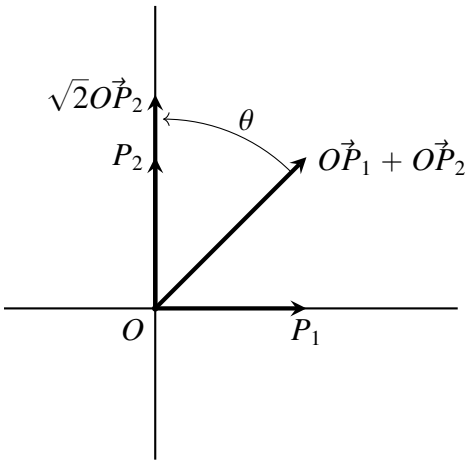
Esempio 7.24. Posto $\theta = \frac{\pi}{4}$, si ha $\cos\theta = \sin\theta = \frac{\sqrt{2}}{2}$. Quindi la rotazione di angolo $\frac{\pi}{4}$ è rappresentata dalla matrice

$$\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

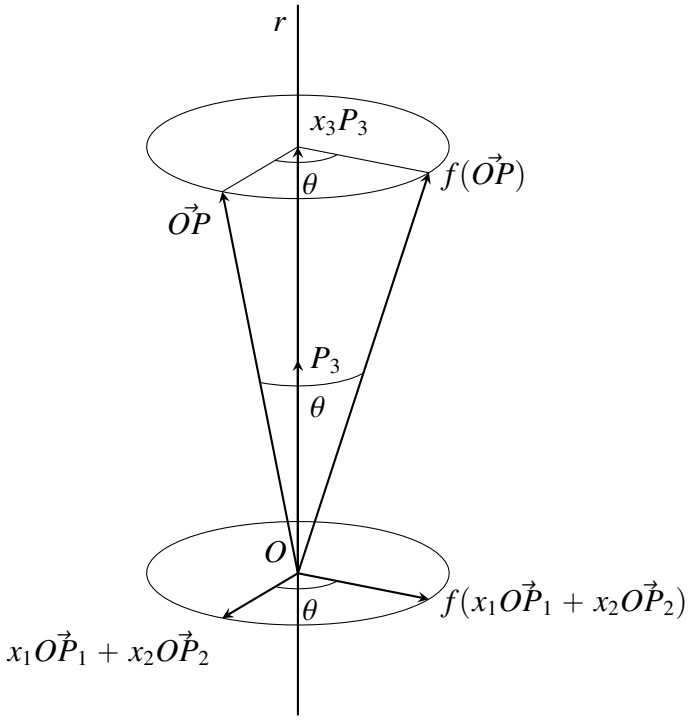
Ad esempio, il vettore $\vec{OP}$ del piano di coordinate $(1,1)$ rispetto al sistema di riferimento fissato viene mandato dalla rotazione nel vettore di coordinate

$$\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \sqrt{2} \end{pmatrix}$$

Questo può essere verificato graficamente: se $\vec{OP}_1, \vec{OP}_2$ sono i due vettori che formano la base del sistema di riferimento fissato, il fatto che $\vec{OP}$ abbia coordinate $(1,1)$ significa che $\vec{OP} = \vec{OP}_1 + \vec{OP}_2$, mentre il fatto che il suo trasformato abbia coordinate $(0, \sqrt{2})$ significa che esso è $0\vec{OP}_1 + \sqrt{2}\vec{OP}_2 = \sqrt{2}\vec{OP}_2$. Infatti, come si vede nel seguente disegno, $\sqrt{2}\vec{OP}_2$ è esattamente il vettore che si ottiene ruotando di 45 gradi in senso antiorario il vettore $\vec{OP} = \vec{OP}_1 + \vec{OP}_2$:



Nello spazio tridimensionale, la situazione è un po' più complessa in quanto le rotazioni dello spazio sono determinate non solo dall'angolo di rotazione ma anche dal loro asse, ovvero dalla retta attorno alla quale effettuiamo la rotazione. Per questo motivo, non esiste una classificazione esplicita completa come quella data dalla (7.19) nel caso del piano.
Assumono però forma particolarmente semplice le rotazioni attorno agli *assi coordinati*, ovvero alle rette che contengono i vettori $\vec{OP}_1, \vec{OP}_2, \vec{OP}_3$ che formano la base del sistema di riferimento scelto: le chiameremo rispettivamente *asse x*, *asse y* e *asse z*.
Iniziamo dalle rotazioni attorno all'asse z. Come si vede nel seguente disegno, una rotazione di angolo θ di un vettore $\vec{OP} = x_1\vec{OP}_1 + x_2\vec{OP}_2 + x_3\vec{OP}_3$ attorno a tale asse mantiene fissa la componente $x_3\vec{OP}_3$ che sta sull'asse di rotazione, mentre ruota di un angolo θ la componente $x_1\vec{OP}_1 + x_2\vec{OP}_2$ che sta nel piano determinato da $\vec{OP}_1$ e $\vec{OP}_2$.



Quindi, applicando la matrice che rappresenta la rotazione alla terna (x_1, x_2, x_3) delle coordinate del vettore $\vec{OP}$, la terza componente x_3 non deve variare mentre le prime due componenti x_1 e x_2 si trasformeranno esattamente come per le rotazioni nel piano, ovvero secondo la (7.19). Deduciamo che la matrice che rappresenta una rotazione di angolo θ attorno all'asse z è

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{7.20}$$

Infatti, come si vede applicando tale matrice a una terna (x_1, x_2, x_3),

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \cos\theta x_1 - \sin\theta x_2 \\ \sin\theta x_1 + \cos\theta x_2 \\ x_3 \end{pmatrix}$$

come previsto.

Ragionamenti analoghi permettono di dedurre che la matrice di una rotazione di angolo θ attorno all'asse y è

$$\begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \tag{7.21}$$

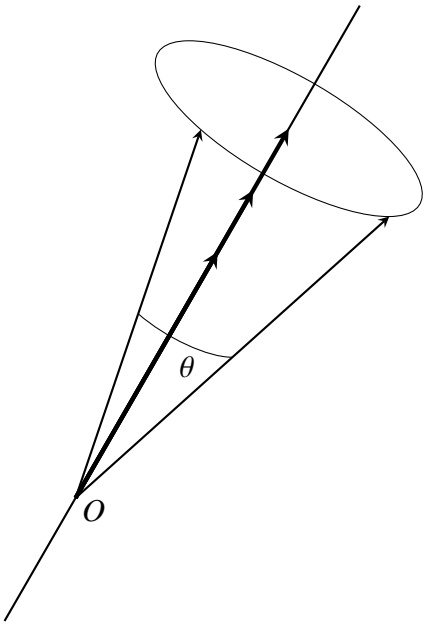
Infatti applicando tale matrice a una terna di coordinate (x_1, x_2, x_3) non si modifica x_2 che rappresenta la componente corrispondente all'asse di rotazione, mentre x_1 e x_3 si trasformano come una rotazione nel piano determinato da $\vec{OP}_1$ e $\vec{OP}_3$. Analogamente la matrice di una rotazione di angolo θ attorno all'asse x è

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{7.22}$$

Applicando tale matrice a una terna di coordinate (x_1, x_2, x_3) non si modifica x_1 che rappresenta la componente corrispondente all'asse x di rotazione, mentre x_2 e x_3 si trasformano come una rotazione nel piano su cui giacciono $\vec{OP}_2$ e $\vec{OP}_3$.

Data una qualunque matrice $A \in \mathrm{SO}(3)$ possiamo trovare l'asse di rotazione e l'angolo di rotazione come segue.

Per quello che riguarda l'asse, osserviamo che i vettori che appartengono all'asse di una rotazione rimangono invariati quando applichiamo la rotazione.



Poiché applicando la rotazione A il vettore di coordinate x viene mandato nel vettore di coordinate Ax, le coordinate x dei vettori che stanno sull'asse devono soddisfare l'uguaglianza $Ax = x$. Ma questa è esattamente l'uguaglianza che soddisfano gli autovettori di A relativi all'autovalore 1, e quindi per determinare l'asse basta determinare tali autovettori.

Quanto all'angolo di rotazione, utilizzando l'identificazione tra R^2 e $\mathbb{C}$ ed il fatto che un'equazione caratteristica ammette sempre 3 soluzioni (in $\mathbb{C}$), si può dimostrare che esso è determinato dalle restanti soluzioni dell'equazione caratteristica. Più precisamente, si mostra che gli autovalori complessi di una matrice in $\mathrm{SO}(2)$ sono numeri complessi del tipo

$$\cos\theta \pm i \sin\theta \tag{7.23}$$

dove θ è proprio l'angolo di rotazione. Lo stesso vale per le ulteriori radici complesse del polinomio caratteristico di A.

Ora, a partire dalle rotazioni date attorno agli assi, possiamo ricavare altre rotazioni per composizione: infatti, il Teorema 7.21 e la Proposizione 7.23 ci permettono di dedurre che la composizione di due rotazioni nello spazio è ancora una rotazione. Più precisamente, se abbiamo due rotazioni rappresentate da due matrici (ortogonali speciali) $A_1, A_2 \in \mathrm{SO}(n)$, la loro composizione come sappiamo è rappresentata dal prodotto A_1A_2. Per la Proposizione 7.23 questa è ancora un elemento di $\mathrm{SO}(n)$ e per il Teorema 7.21 questa è una rotazione.

Illustriamo ora con un esempio come determinare, usando il calcolo di autovalori e autovettori, asse e angolo di una rotazione.

Esempio 7.25. Dalla (7.22) abbiamo che la matrice che rappresenta una rotazione attorno all'asse x di angolo $\theta = \frac{\pi}{2}$ è

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix},$$

mentre la (7.21) ci dice che la matrice che rappresenta una rotazione attorno all'asse y di angolo $\theta = \frac{\pi}{2}$ è

$$\begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

La composizione di tali rotazioni è quindi data dalla matrice A che si ottiene eseguendo il prodotto

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{7.24}$$

Ora, vogliamo determinare asse e angolo di tale rotazione. Il polinomio caratteristico della matrice data in (7.24) è

$$\det \begin{pmatrix} -\lambda & 0 & -1 \\ -1 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{pmatrix} = -\lambda \det \begin{pmatrix} -\lambda & 0 \\ 1 & -\lambda \end{pmatrix} - \det \begin{pmatrix} -1 & -\lambda \\ 0 & 1 \end{pmatrix} = -\lambda^3 + 1$$

Come previsto, una radice del polinomio caratteristico è $\lambda = 1$, mentre le altre possono essere determinate con il metodo descritto nel paragrafo precedente.

	-1	0	0	1
1				

$\longrightarrow$

	-1	0	0	1
1				
	-1			

$\longrightarrow$

	-1	0	0	1
1		-1		
	-1	-1		

$\longrightarrow$

	-1	0	0	1
1		-1	-1	
	-1	-1	-1	

$\longrightarrow$

	-1	0	0	1
1		-1	-1	-1
	-1	-1	-1	0

Quindi le altre soluzioni dell'equazione caratteristica sono le soluzioni di $-\lambda^2 - \lambda - 1 = 0$ o equivalentemente cambiando di segno per comodità $\lambda^2 + \lambda + 1 = 0$. Applicando la formula risolutiva si trova

$$\lambda = \frac{-1 \pm \sqrt{1-4}}{2} = \frac{-1 \pm \sqrt{-3}}{2} = \frac{-1 \pm \sqrt{3}\sqrt{-1}}{2} = -\frac{1}{2} \pm i\frac{\sqrt{3}}{2}.$$

Come si vede confrontando con la (7.23), l'angolo di rotazione θ soddisfa $\cos\theta = -\frac{1}{2}$, $\sin\theta = \frac{\sqrt{3}}{2}$, e concludiamo $\theta = \frac{2}{3}\pi$
Per trovare l'asse di rotazione, calcoliamo gli autovettori relativi all'autovalore $\lambda = +1$ risolvendo il sistema omogeneo $(A - \mathrm{Id}_3)x = 0$. Si ha

$$A - \mathrm{Id}_3 = \begin{pmatrix} -1 & 0 & -1 \\ -1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \underset{R_2 \to R_2 - R_1}{\longrightarrow} \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \underset{R_3 \to R_3 + R_2}{\longrightarrow} \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Dunque gli autovettori relativi a $\lambda = 1$sono le soluzioni del sistema ridotto

$$\begin{cases} -x_1 - x_3 = 0 \\ -x_2 + x_3 = 0 \end{cases}$$

Posto $x_3 = t$, si trova $x_2 = x_3 = t$ e $x_1 = -x_3 = -t$; quindi le soluzioni sono le triple $(-t, t, t) = t(-1, 1, 1)$, al variare di $t \in \mathbb{R}$. Queste rappresentano in coordinate i vettori $-t\vec{OP}_1 + t\vec{OP}_2 + t\vec{OP}_3 = t(-\vec{OP}_1 + \vec{OP}_2 + \vec{OP}_3)$ che formano l'asse di rotazione.

Concludiamo la discussione sulle matrici ortogonali enunciando una importante caratterizzazione di queste.

Proposizione 7.26. *Una matrice $A \in M_n(\mathbb{R})$ è ortogonale se e solo se le sue colonne (o equivalentemente le sue righe) rappresentano n vettori ortogonali a due a due.*

Chiudiamo il capitolo citando un'importante applicazione della teoria degli autovettori e autovettori all'informatica e in particolare alla cosiddetta computer vision.
Una categoria importante di matrici che sooddisfano svariate proprietà è data dalle cosiddette *matrici simmetriche*

Definizione 7.27. Una matrice $A \in M_n(\mathbb{R})$ si dice **simmetrica** se $A^T = A$, ovvero se le sue entrate soddisfano $A_{ij} = a_{ji}$ per ogni coppia di indici.

Ad esempio, la matrice $A = \left(\begin{smallmatrix} 3 & 2 \\ 2 & 5 \end{smallmatrix}\right)$ è, come si verifica immediatamente, simmetrica.
enunciamo ora alcune proprietà delle matrici simmetriche senza dimostrarle tutte:

(i) Le matrici simmetriche di ordine n non formano un sottogruppo di $M_n(\mathbb{K})$. Infatti il prodotto di due matrici simmetriche non è necessariamente una matrice simmetrica. In effetti possiamo vedere che per due matrici simmetriche A e B vale la seguente condizione

$$AB = BA \iff AB \text{ è simmetrica.}$$

Infatti le matrici al primo membro hanno entrate

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} \text{ e } (BA)_{ij} = \sum_{k=1}^{n} b_{ik}a_{kj}.$$

Scambiando gli indici delle entrate di A e B vediamo queste due espressioni si equivalgono se e solo se AB è una matrice simmetrica.

(ii) Una matrice simmetrica ammette sempre n autovalori reali (possibilmente con ripetizioni) ed n autovettori linearmente indipendenti.

(iii) Gli autovettori di una matrice simmetrica A relativi ad autovalori distinti sono ortogonali. Infatti siano v_1 e v_2 autovettori relativi rispettivamente a λ_1 e λ_2 con $\lambda_1 \neq \lambda_2$. Allora vogliamo mostrare che il prodotto scalare $v_1 \cdot v_2$ è nullo. Questo discende dal calcolo

$$\lambda_1 v_1 \cdot v_2 = \lambda_1 v_1^T v_2 = (Av_1)^T v_2 = v_1^T A^T v_2 = v_1^T A v_2 = \lambda_2 v_1^T v_2 = \lambda_2 v_1 \cdot v_2$$

dove abbiamo usato che A è una matrice simmetrica. Ora l'uguaglianza può valere solo se $v_1 \cdot v_2 = 0$ poiché $\lambda_1 \neq \lambda_2$ per assunto.

Abbiamo detto sopra che se una matrice A è diagonalizzabile (cioè ammette n autovettori indipendenti), esiste sempre una matrice invertibile V, le cui colonne sono gli autovettori di A, tale che $V^{-1}AV$ sia una matrice diagonale (avente sulla diagonale gli autovalori di A). Per quanto appena visto, se A è simmetrica questo è sempre possibile e, per la Proposizione 7.26, V è una matrice ortogonale. A meno di scambiare due colonne possiamo supporre che $V \in \mathrm{SO}(n)$ per cui abbiamo il seguente

Teorema 7.28. *Sia $A \in M_n(\mathbb{R})$ una matrice simmetrica con autovalori $\lambda_1, \ldots, \lambda_n$. Allora esiste una matrice ortogonale speciale $V \in SO(n)$ tale che*

$$V^{-1}AV = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Denotando con D la matrice diagonale $V^{-1}AV$, l'uguaglianza $V^{-1}AV = D$ si può anche scrivere, moltiplicando entrambi i membri a destra per V^{-1} e a sinistra per V come $A = VDV^{-1}$.
Se siamo nel caso di matrici di ordine 3, tale uguaglianza ci dice che ogni trasformazione dello spazio che sia rappresentata da una matrice simmetrica può essere decomposta come prodotto (quindi composizione delle trasformazioni corrispondenti) di una rotazione, poi di una matrice diagonale e poi della rotazione inversa alla prima.

Tale risultato trova applicazione ad esempio in alcuni algoritmi di “stitching di immagini”, ovvero il procedimento mediante il quale si combinano più fotografie, sovrapponendole, per ottenere un’unica fotografia panoramica o in un video stabilizzato in caso si tratti di frame di un video. Il riconoscimento delle parti corrispondenti nelle varie immagini che devono essere incollate usa il risultato di decomposizione delle matrici simmetriche che abbiamo appena citato, ma non entriamo nei dettagli.

Appendice A

Complementi

In questa appendice presentiamo una serie di approfondimenti e applicazioni di quanto visto in questo corso. Questo materiale è da considerarsi facoltativo.

A.1 Teorema cinese del resto

A.2 Matrici simmetriche ed ortogonali

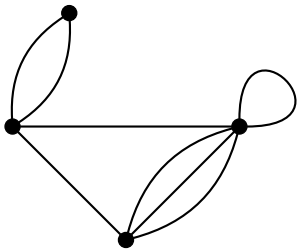
A.3 Il campo dei quaternioni e le rotazioni dello spazio

A.4 Matrici di adiacenza di un grafo

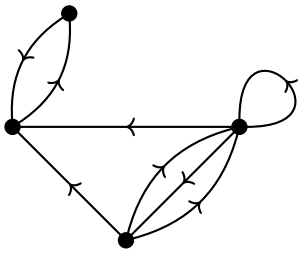
Come abbiamo visto, le matrici sono un utilissimo strumento di calcolo in vari contesti. Ad esempio, nella risoluzione dei sistemi di equazioni lineari, dove la compatibilità del sistema e le sue soluzioni possono essere determinate facilmente lavorando sulla sua matrice dei coefficienti e sulla matrice completa. O ancora, le applicazioni lineari tra vettori nel piano e nello spazio (quali rotazioni, riflessioni e proiezioni) possono essere rappresentate in coordinate mediante matrici, lavorando sulle quali possiamo comporre tali applicazioni tra loro, determinarne l'applicazione inversa e ottenere altre importanti informazioni geometriche e non.

Vedremo ora un'altra interessante applicazione del calcolo matriciale nel contesto della teoria dei grafi.

Ricordiamo che un grafo è un insieme di punti, detti *vertici*, e di segmenti di curva, detti *lati*, che collegano coppie di vertici, cf. Definizione 2.24.

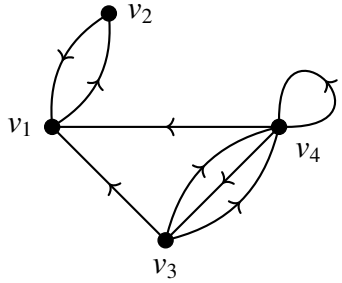


Inoltre i lati del grafo possono anche essere dotati di un'orientazione



e in tal caso si parla di grafo orientato.

Ad un grafo orientato può essere associata una matrice nel seguente modo: si numerano i suoi vertici



e si costruisce quindi una matrice A quadrata di ordine uguale al numero di vertici del grafo e le cui entrate sono definite nel modo seguente:

$$A_{ij} = \text{numero di lati che vanno dal vertice } v_i \text{ al vertice } v_j$$

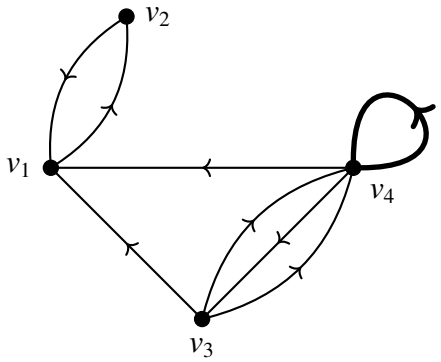
La matrice così costruita si dice *matrice di adiacenza del grafo*. Ad esempio, la matrice di adiacenza del grafo dato dalla figura precedente è la matrice quadrata di ordine 4

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 2 \\ 1 & 0 & 1 & 1 \end{pmatrix}$$

(si noti che un elemento A_{ii} della diagonale è diverso da zero se e solo se esiste un lato che va da v_i a v_i, ovvero un cappio centrato in v_i).
Se un grafo è non orientato, il lato da un vertice v_i a un vertice v_j va inteso come "percorribile" in entrambi i sensi, quindi si deve intendere che il numero di lati da v_i a v_j è uguale al numero di lati da v_j a v_i, ovvero $A_{ij} = A_{ji}$ (la matrice è quindi *simmetrica*).
Ad esempio, la matrice di adiacenza del grafo non orientato di sopra è

$$A = \begin{pmatrix} 0 & 2 & 1 & 1 \\ 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 3 \\ 1 & 0 & 3 & 1 \end{pmatrix}$$

La matrice di adiacenza di un grafo è estremamente utile per ottenere molte informazioni sul grafo stesso: noi ne esamineremo in dettaglio una in particolare, relativa all'esistenza e al numero di cammini che vanno da un vertice a un altro all'interno del grafo.
Un cammino in un grafo da un vertice v a un vertice v' è una sequenza di lati consecutivi (ovvero tali che il secondo vertice di un lato coincide con il primo vertice del successivo), percorsi secondo l'orientazione nel caso in cui il grafo sia orientato, che portano da v a v'. Diciamo *lunghezza di un cammino* il numero di lati che lo compongono , comprese eventuali ripetizioni.
Ad esempio, nel grafo precedente



(in cui per agevolare la descrizione abbiamo numerato anche alcuni lati) percorrendo nell'ordine i lati L_1, poi il cappio L_2 due volte, poi L_3 e infine L_4, si ottiene un cammino di lunghezza 5 da v_3 a v_2.

Ebbene, si ha il seguente

Teorema A.1. *Sia dato un grafo con vertici $v_1, v_2, \ldots, v_n$ e sia A la sua matrice di adiacenza. Allora, per ogni intero positivo k, l'entrata di posto ij della matrice A^k (che si ottiene moltiplicando A per se stessa k volte) è uguale al numero di cammini di lunghezza k che vanno da v_i a v_j.*

Prima di dare un'idea della dimostrazione del teorema e spiegare cosa abbia a che fare il prodotto righe per colonne tra matrici con i cammini nel grafo, facciamo subito alcuni esempi per illustrare e verificare il teorema.
Si consideri ad esempio il semplice grafo orientato con due vertici del disegno seguente
La matrice di adiacenza del grafo è $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$
Si ha allora ad esempio

$$A^2 = AA = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

In base al Teorema A.1, il fatto ad esempio che $(A^2)_{12} = 2$ ci dice che devono esistere esattamente due cammini di lunghezza 2 da v_1 a v_2: infatti, il primo si ottiene percorrendo il lato da v_1 a v_2 e poi il cappio centrato in v_2; il secondo cammino si ottiene percorrendo prima il cappio centrato in v_1 e poi il lato da v_1 a v_2. Ancora, l'entrata $(A^2)_{11} = 1$ ci dice che esiste un solo cammino di lunghezza 2 da v_1 a se stesso, che è infatti quello che si ottiene percorrendo due volte il cappio centrato in v_1.
Come secondo esempio si consideri il grafo
che ha come matrice di adiacenza $A = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$. Si ha

$$A^2 = \begin{pmatrix} 2 & 0 & 2 \\ 1 & 2 & 0 \\ 0 & 2 & 0 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 2 & 4 & 0 \\ 2 & 2 & 2 \\ 2 & 0 & 2 \end{pmatrix}$$

Ad esempio, $(A^2)_{11} = 2$ ci dice che ci sono esattamente due cammini di lunghezza 2 da v_1 a se stesso (si noti di lunghezza 1 non ce n'era neanche uno, non essendoci cappi su v_1), che si ottengono percorrendo prima uno dei due lati che va da v_1 a v_2 e poi l'unico lato da v_2 a v_1; ancora, mentre non ci sono cammini di lunghezza 2 da v_3 a v_1, ce ne sono due di lunghezza 3, essendo $(A^3)_{31} = 2$: si tratta dei cammini che si ottengono percorrendo prima l'unico lato da v_3 a v_1, poi scegliendo da v_1 uno dei due lati che vanno verso v_2 e infine percorrendo l'unico lato da v_2 a v_1.

Diamo ora una bozza di dimostrazione del Teorema A.1.
Se $k = 1$, il fatto che $(A^k)_{ij} = A_{ij}$ ci dia il numero di cammini di lunghezza 1 da v_i a v_j è dovuto alla definizione stessa di matrice di adiacenza: infatti, un cammino di lunghezza 1 è esattamente un lato, e come sappiamo la matrice di adiacenza ha come entrata A_{ij} proprio il numero di lati da v_i a v_j.
Nel caso $k = 2$, si tratta di contare i cammini di lunghezza 2 da v_i a v_j e dimostrare che il loro numero è dato da $(A^2)_{ij}$.
Ora, un tale cammino, essendo di lunghezza 2, sarà composto da due lati: il primo che va da v_i a un vertice intermedio v_k, il secondo che va da v_k a v_j. Ma il numero di lati che

vanno da v_i a v_k è dato per definizione di matrice di adiacenza da A_{ik}, mentre il numero di lati che vanno da v_k a v_j è dato da A_{kj}. Poiché per comporre un cammino di lunghezza 2 possiamo scegliere per ognuno degli A_{ik} lati da v_i a v_k uno degli A_{kj} lati da v_k a v_j, in tutto abbiamo esattamente $A_{ik}A_{kj}$ possibili cammini di lunghezza 2 da v_i a v_j. Ma poichè il vertice intermedio v_k può essere uno qualunque dei vertici $v_1, v_2, \ldots, v_n$, per ottenere il numero totale di cammini dobbiamo sommare i numeri $A_{ik}A_{kj}$ per k che va da 1 a n, ovvero

$$A_{i1}A_{1j} + A_{i2}A_{2j} + A_{i3}A_{3j} + \cdots + A_{in}A_{nj}$$

che per definizione di prodotto righe per colonne è proprio $(A^2)_{ij}$, come volevamo.
Per potenze successive della matrice di adiacenza, si ragiona nello stesso modo: ad esempio, un cammino di lunghezza 3 da v_i a v_j si può pensare come un cammino di lunghezza 2 che va da v_i a un vertice intermedio v_k e poi un lato finale da v_k a v_j. Ma il numero di cammini di lunghezza 2 da v_i a v_k per quanto appena dimostrato è uguale a $(A^2)_{ik}$: ognuno di questi può essere completato con uno degli A_{kj} lati da v_k a v_j, ottenendo in tutto $(A^2)_{ik}A_{kj}$ cammini. Facendo variare come prima k da 1 a n e sommando, si ottiene esattamente $(A^3)_{ij}$.
Una dimostrazione rigorosa del teorema, valida per ogni intero positivo k, può essere fatta per induzione, ma omettiamo i dettagli.