Matematica Discreta per Informatica

Alberto Albano

Marco Burzio

Dipartimento di Matematica, Università di Torino, Via Carlo Alberto 10, 10123 Torino, ITALY

E-mail address: alberto.albano@unito.it

URL: http://www.dm.unito.it/personalpages/albano/index.htm

Dipartimento di Matematica, Università di Torino, Via Carlo Alberto 10, 10123 Torino, ITALY

E-mail address: marco.burzio@unito.it

URL: http://www2.dm.unito.it/paginepersonali/burzio/index.htm

Indice

Capitolo 1

Aritmetica modulare

1. L'algoritmo di Euclide

Da bambini abbiamo imparato che quando si divide 27 per 6 il *quoziente* è 4 ed il *resto* è 3, cioè che

$$27 = 6 \cdot 4 + 3.$$

La cosa importante è notare che il resto deve essere un intero maggiore o uguale a 0 e minore in senso stretto di 6. Anche se è vero che, ad esempio

$$27 = 6 \cdot 3 + 9,$$

si prende come resto 3 e non 9 perché 3 è il più piccolo intero positivo per cui vale l'identità.

La proprietà fondamentale di cui ci stiamo occupando si può enunciare in modo preciso con il seguente teorema.

Teorema 1.1 (della Divisione)**.** *Se $a \geq 0$ e $b > 0$ sono interi, esistono e sono univocamente determinati gli interi $q \geq 0$ ed r tali che*

$$a = bq + r \qquad e \qquad 0 \leq r < b.$$

q ed r sono detti rispettivamente il **quoziente** *e il* **resto** *della divisione.*

La determinazione del quoziente e del resto si può fare utilizzando l'algoritmo della divisione che abbiamo imparato nella scuola elementare. Ad esempio si provi ad applicare l'algoritmo per ottenere quoziente e resto della divisione di 2731 diviso per 43.

Definizione 1.2. Dati due interi a e b diciamo che b è un **divisore** di a, e scriviamo $b|a$, se, dividendo a per b si ottiene resto 0 cioè se $a = qb$.

Equivalentemente si dice che b **divide** a oppure che a è **divisibile** per b, oppure ancora che a è un **multiplo** di b.

Definizione 1.3. Dati due interi a e b, il numero intero positivo d tale che:

(1) $d|a$ e $d|b$,

(2) se $c|a$ e $c|b$ allora $c|d$

è detto il **massimo comun divisore** di a e b e denotato con

$$d = (a, b).$$

La condizione (1) dice che d è un divisore comune di a e b, e la condizione (2) dice che ogni divisore comune di a e b è anche un divisore di d. Per esempio, 6 è un divisore comune di 60 ed 84, ma non è il massimo comun divisore, dal momento che $12|60$ e $12|84$, ma $12 \nmid 6$. (Il simbolo $\nmid$ significa "non divide").

Definizione 1.4. Due interi a, b sono detti **coprimi** se $(a, b) = 1$.

Esiste un famoso metodo per calcolare il massimo comun divisore di due interi, basato sull'uso del quoziente e del resto.

Algoritmo di Euclide. *Il massimo comun divisore di due interi a, b si ottiene con il metodo delle divisioni successive seguenti:*

$$\begin{aligned} a &= bq + r_0 \\ b &= r_0 q_0 + r_1 \\ r_0 &= r_1 q_1 + r_2 \\ &\cdots \\ r_{n-2} &= r_{n-1} q_{n-1} + r_n \\ r_{n-1} &= r_n q_n + 0 \end{aligned}$$

Poiché i resti r_i $(0 \leq i \leq n)$ sono positivi e decrescenti, dopo un numero finito di passi la divisione dà resto 0. Allora l'ultimo resto non nullo r_n è il massimo comun divisore di a, b.

Dimostrazione. Proviamo inizialmente che se $a = bq + r_0$ allora $(a, b) = (b, r_0)$. Infatti se $c|a$ e $c|b$, da $r_0 = a - qb$ segue $c|r_0$. Quindi ogni divisore di a e b è anche un divisore di r_0. Viceversa, se $c'|b$ e $c'|r_0$ segue che $c'|a$. Allora a, b e b, r_0 hanno gli stessi divisori e quindi $(a, b) = (b, r_0)$. Ripetendo questo ragionamento si ottiene:

$$(a, b) = (b, r_0) = (r_0, r_1) = \cdots = (r_{n-1}, r_n) = r_n.$$

Si noti che l'ultima uguaglianza vale in quanto $r_n|r_{n-1}$. □

Esempio. Determinare il massimo comun divisore di 1776 e 1492.

Soluzione. Si ha:

$$\begin{aligned}
(1776, 1492) &= (1492, 284), && \text{essendo} && \mathbf{1776} = 1 \cdot \mathbf{1492} + \mathbf{284} \\
(1492, 284) &= (284, 72), && \text{essendo} && \mathbf{1492} = 5 \cdot \mathbf{284} + \mathbf{72} \\
(284, 72) &= (72, 68), && \text{essendo} && \mathbf{284} = 3 \cdot \mathbf{72} + \mathbf{68} \\
(72, 68) &= (68, 4), && \text{essendo} && \mathbf{72} = 1 \cdot \mathbf{68} + \mathbf{4} \\
(68, 4) &= 4, && \text{essendo} && \mathbf{68} = 17 \cdot \mathbf{4}.
\end{aligned}$$

Teorema 1.5 (Identità di Bezout)**.** *Se $d = (a, b)$ è il massimo comun divisore di a e b, allora esistono due interi m ed n tali che*

$$d = ma + nb.$$

$ma + nb$ è detta una **combinazione lineare** *di a e b con coefficienti m ed n.*

Dimostrazione. Ricaviamo i resti ottenuti nelle divisioni successive dell'algoritmo e sostituiamoli successivamente a partire dall'ultimo. Si ottiene:

$$\begin{aligned}
r_n &= r_{n-2} - r_{n-1}q_{n-1} \\
&= r_{n-2} - (r_{n-3} - r_{n-2}q_{n-2})q_{n-1} \\
&= (1 - q_{n-1}q_{n-2})r_{n-2} + q_{n-1}r_{n-3} \\
&= (\cdots)r_{n-3} + (\cdots)r_{n-4} \\
&= \cdots\cdots\cdots \\
&= (\cdots)a + (\cdots)b.
\end{aligned}$$

□

Esempio 1.6. Determinare $d = (1776, 1492)$ come combinazione lineare di 1776 e di 1492.

Soluzione. Ricavando i resti calcolati nell'esempio precedente, a partire dall'ultimo non nullo, si ha:

$$\begin{aligned}
\mathbf{4} &= \mathbf{72} - 1 \cdot \mathbf{68} \\
&= \mathbf{72} - 1 \cdot (\mathbf{284} - 3 \cdot \mathbf{72}) && = -1 \cdot \mathbf{284} + 4 \cdot \mathbf{72} \\
&= -1 \cdot \mathbf{284} + 4 \cdot (\mathbf{1492} - 5 \cdot \mathbf{284}) && = 4 \cdot \mathbf{1492} - 21 \cdot \mathbf{284} \\
&= 4 \cdot \mathbf{1492} - 21 \cdot (\mathbf{1776} - 1 \cdot \mathbf{1492}) && = -21 \cdot \mathbf{1776} + 25 \cdot \mathbf{1492}.
\end{aligned}$$

Occupiamoci ora dell'analisi dell'algoritmo di Euclide. L'efficienza di questo algoritmo dipende dal numero di passi o divisioni successive richieste. Cerchiamo quindi un estremo superiore di questo numero.

Vale al riguardo la seguente importante proprietà dell'algoritmo di Euclide.

Teorema 1.7. *Dati a e b interi positivi con $a \geq b$ e considerate le prime due divisioni successive $a = bq + r_0$, $b = r_0q_0 + r_1$ si ha*

$$r_1 < \frac{b}{2}$$

Dimostrazione. Supponiamo per assurdo che $r_1 \geq \frac{b}{2}$. Dall'identità $b = r_0q_0 + r_1$ segue allora $b \geq r_0q_0 + \frac{b}{2}$, da cui $r_0q_0 \leq \frac{b}{2}$.

Se fosse $q_0 = 0$ si avrebbe $b = r_1$ ma, essendo i resti strettamente decrescenti, si avrebbe $r_1 < r_0 < b$, da cui $r_1 < b$ e quindi $q_0 \geq 1$. Dividendo allora per q_0 la disuguaglianza $r_0q_0 \leq \frac{b}{2}$ si ottiene $r_0 \leq \frac{b}{2}$. Poiché $r_1 < r_0$ si ottiene ancora $r_1 < r_0 \leq \frac{b}{2}$ contro l'ipotesi $r_1 \geq \frac{b}{2}$. Dunque l'ipotesi fatta è assurda e deve essere $r_1 < \frac{b}{2}$. □

Il ragionamento svolto nella dimostrazione precedente si può ripetere partendo, invece che dalle due prime divisioni dell'algoritmo, da due divisioni successive qualsiasi:

$$r_i = r_{i+1}q_{i+1} + r_{i+2}$$
$$r_{i+1} = r_{i+2}q_{i+2} + r_{i+3}$$

ottenendo in generale che

$$r_{i+3} < \frac{r_{i+1}}{2}$$

Questo significa che ogni due passi il resto almeno si dimezza. Si ha quindi

$$r_1 < \frac{b}{2}$$
$$r_3 < \frac{r_1}{2} < \frac{b}{2^2}$$
$$r_5 < \frac{r_3}{2} < \frac{b}{2^3}$$
$$\cdots\cdots$$
$$r_{2k-1} < \frac{b}{2^k}.$$

Se ora k è tale che $\frac{b}{2^k} < 1$, si ottiene che $r_{2k-1} < \frac{b}{2^k} < 1$, cioè $r_{2k-1} = 0$.

Quindi l'algoritmo termina sicuramente appena k è tale da aversi $\frac{b}{2^k} < 1$ cioè $b < 2^k$ cioè ancora $k < \log_2 b$. Poiché ad ogni variazione di k corrispondono 2 passi dell'algoritmo, l'algoritmo di Euclide termina in un numero intero di passi minore o uguale a $2\log_2 b$.

Ad esempio l'algoritmo per determinare $(1776, 1492)$ termina sicuramente in un numero intero di passi $\leq 2\log_2 1442$ cioè in al più $2 \cdot 10 = 20$ passi.

Esempio 1.8. Determinare il massimo comun divisore di 721 e 448 ed esprimerlo nella forma $721m + 448n$ con $m, n \in \mathbb{Z}$. Al massimo quante divisioni si dovranno fare per determinare $(721, 448)$?

Soluzione. Facendo i calcoli si ottiene $(721, 448) = 7$ e $7 = 23 \cdot \mathbf{721} - 37 \cdot \mathbf{448}$.

Siccome l'algoritmo termina in un numero intero di passi minore o uguale a $2\log_2 448$, essendo $\log_2 448 = 8.807\ldots$, l'algoritmo termina sicuramente in 17 passi.

2. Equazioni in $\mathbb{Z}$

Definizione 2.1. Una **equazione lineare in una incognita x a coefficienti in** $\mathbb{Z}$ è una espressione della forma

$$ax + b = 0, \qquad \text{con} \qquad a, b \in \mathbb{Z}.$$

Una **soluzione** in $\mathbb{Z}$ dell'equazione precedente è un intero $x_0 \in \mathbb{Z}$ tale che l'equazione sia soddisfatta, cioè si trasformi in una identità, sostituendo x_0 al posto della x.

Sull'esistenza di soluzioni si può notare immediatamente che l'equazione precedente, se $a \neq 0$, ha una ed una sola soluzione in $\mathbb{Z}$ quando $a|b$.

Esempio 2.2. L'equazione $3x - 9 = 0$ ha soluzione $x = 3$. L'equazione $3x - 8 = 0$ non ha soluzioni in $\mathbb{Z}$ in quanto $3 \nmid 8$.

Definizione 2.3. Una **equazione lineare in due incognite a coefficienti in** $\mathbb{Z}$ è una espressione della forma

$$ax + by = c, \qquad \text{con} \qquad a, b, c \in \mathbb{Z}.$$

Una **soluzione** dell'equazione precedente è una coppia (x_0, y_0) formata da due interi x_0, y_0, tale che l'equazione sia soddisfatta, cioè si trasformi in una identitità, sostituendo $x = x_0$, $y = y_0$.

L'insieme di tutte le soluzioni è detta la **soluzione generale** dell'equazione.

Esempio 2.4. $(18, -91)$ è una soluzione dell'equazione $365x + 72y = 18$ come si può verificare direttamente.

Sulle eventuali soluzioni dell'equazione $ax + by = c$ si ha il seguente risultato.

Teorema 2.5. *L'equazione $ax + by = c$ a coefficienti in $\mathbb{Z}$ ha soluzioni in $\mathbb{Z}$ se e solo se*

$$d|c, \qquad dove \qquad d = (a, b).$$

Inoltre, se (x_0, y_0) è una soluzione dell'equazione, allora anche (x_1, y_1) lo è, essendo

$$x_1 = x_0 + k\frac{b}{d}, \qquad y_1 = h_0 - k\frac{a}{d}, \qquad \forall k \in \mathbb{Z}.$$

Dimostrazione. Se (x_0, y_0) è una soluzione, $ax_0 + by_0 = c$ è una identità. Quindi, poiché $d|a$ e $d|b$, segue anche $d|c$.

Viceversa, sappiamo che d è una combinazione lineare di a e b, cioè $d = ma + nb$; inoltre per ipotesi $d|c$, cioè esiste $q \in \mathbb{Z}$ tale che $qd = c$. Quindi

$$qd = q(ma + nb) = a(qm) + b(qn) = c,$$

e cioè $x_0 = qm$ e $y_0 = qn$ trasformano l'equazione in una identità quando sono sostituiti al posto delle incognite. Ne segue che (x_0, y_0) è una soluzione dell'equazione.

La seconda parte dell'enunciato del teorema si verifica direttamente sostituendo (x_1, y_1) al posto delle incognite e ottenendo una identità in $\mathbb{Z}$. □

Esempio 2.6. Risolvere in $\mathbb{Z}$ l'equazione

$$365x + 72y = 18.$$

Soluzione. Si calcola con l'algoritmo di Euclide il massimo comun divisore ottenendo $(365, 72) = 1$. Si esprime poi $(365, 72)$ come combinazione lineare di 365 e 72 nella forma

$$1 = 29 \cdot 365 - 147 \cdot 72.$$

Segue

$$\begin{aligned} 18 = 1 \cdot 18 &= (29 \cdot 365 - 147 \cdot 72) \cdot 18 = (18 \cdot 29)365 - (18 \cdot 147)72 \\ &= 365 \cdot 522 + 72(-2646). \end{aligned}$$

Da cui la soluzione particolare $(522, -2646)$ e quindi la soluzione generale

$$x_1 = 522 + 72k, \qquad y_1 = -2646 - 365k \qquad \forall k \in \mathbb{Z}$$

che può anche essere scritta nella forma

$$x_1 = 18 + 72h, \qquad y_1 = -91 - 365h \qquad \forall h \in \mathbb{Z},$$

avendo notato nell'esempio precedente che anche $(18, -91)$ è una soluzione particolare dell'equazione.

Esempio 2.7. Risolvere in $\mathbb{Z}$ l'equazione

$$12x + 39y = 15.$$

Soluzione. Si ha $d = (12, 39) = 3$ che si può scrivere nella forma $3 = (-3)(12)+(1)(39)$, da cui, moltiplicando per 5, si ottiene l'identità $12(-15)+(39)(5) = 15$, che fornisce la soluzione particolare $(x_0, y_0) = (-15, 5)$. La soluzione generale è quindi (x_1, y_1) con $x_1 = -15+13k$, $y_1 = 5-4k$, $\forall k \in \mathbb{Z}$.

Esercizio 2.8. Date due clessidre, una da 6 minuti, l'altra da 11 minuti, misurare 13 minuti.

Esempio 2.9. Un contadino vuole comperare 100 animali per un totale di 2.000 euro. Se gli animali sono caprette, conigli e galline per un costo rispettivo di 50, 20 e 5 euro cadauno, quanti di ciascun tipo ne può comperare?

Soluzione. Facendo i conti si ottengono le possibili soluzioni seguenti, indicando con x (risp. y, risp. z) il numero delle caprette (risp. conigli, risp. galline):

$$x = k, \quad y = 100 - 3k, \quad z = 2k, \quad \forall k \in \mathbb{Z}, \quad 0 \leq k \leq 33.$$

3. Relazioni e funzioni

Definizione 3.1. Dati due insiemi A e B, l'insieme di tutte le coppie ordinate con primo elemento in A e secondo elemento in B è detto il **prodotto cartesiano** di A e B e denotato con $A \times B$.

Per esempio:

$$A \times B = \{1, 2\} \times \{a, b, c\} = \{(1, a), (1, b), (1, c), (2, a), (2, b), (2, c)\}.$$

Definizione 3.2. Dati due insiemi A e B, una **relazione da A in B** è un sottoinsieme del prodotto cartesiano $A \times B$. A è detto il **dominio** della relazione e B il **codominio** della relazione. Se $(a, b) \in A \times B$ si dice che a è **in relazione** con b, o anche che b è **immagine** di a nella relazione. Una **relazione su** un insieme C è una relazione con dominio e codominio coincidenti con C.

Esempio 3.3. La relazione $\{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)\}$ è una relazione familiare sull'insieme $\{1, 2, 3, 4\}$. Di quale relazione si tratta?

Soluzione. Notiamo che a è in relazione con b se e solo se $a < b$. Quindi la relazione assegnata è la relazione di *minore (in senso stretto)* tra i numeri 1, 2, 3 e 4.

Definizione 3.4. Dati due insiemi A e B, una **funzione** f **da** A **in** B, denotata con $f : A \to B$, è una relazione da A in B tale che ciascun a in A è in relazione con uno ed un solo b in B. A è detto il **dominio** della funzione e B il **codominio** della funzione.

Esempio 3.5. Se $f(x) = x + 2$ è usata come una formula per descrivere una funzione da $\{-1, 0, 1, 2\}$ in $\{1, 2, 3, 4\}$, quale relazione definisce f?

Soluzione. $f = \{(-1, 1), (0, 2), (1, 3), (2, 4)\}$.

Esempio 3.6. Mostrare che la relazione $\{(a, 2), (a, 3), (b, 4), (c, 5)\}$ non è una funzione da $A = \{a, b, c\}$ in $B = \{2, 3, 4, 5\}$.

Soluzione. L'elemento a del dominio è in relazione sia con 2 che con 3, quindi a non è in relazione con un solo elemento del codominio.

Esempio 3.7. Mostrare che la relazione $\{(a, 3), (c, 2)\}$ dal dominio $\{a, b, c\}$ al codominio $\{2, 3, 4, 5\}$ non è una funzione da $\{a, b, c\}$ in $\{2, 3, 4, 5\}$.

Soluzione. L'elemento b del dominio non è in relazione con alcun elemento del codominio. Non vale quindi che ciascun elemento del dominio è in relazione con un elemento del codominio; non si tratta quindi di una funzione.

Per riassumere quanto si è visto negli esempi precedenti, si controlla che una relazione è una funzione se sono verificate le due condizioni seguenti:

(1) Ogni elemento del dominio è in relazione con un elemento del codominio.

(2) Non c'è nessun elemento del dominio che è in relazione con più di un elemento del codominio.

Definizione 3.8. Una funzione è detta una **iniezione** o una **funzione iniettiva** se elementi distinti del dominio hanno immagini distinte, cioè se ciascun elemento del codominio è immagine di **al più** un elemento del dominio.

Definizione 3.9. Una funzione è detta una **suriezione** o una **funzione suriettiva** se ciascun elemento del codominio è immagine di **almeno** un elemento del dominio.

Definizione 3.10. Una funzione è detta una **biiezione** o una **funzione biiettiva** se è iniettiva e suriettiva, cioè se ogni elemento del codominio è immagine di **uno ed un solo** elemento del dominio.

Esercizio 3.11. Quali delle relazioni seguenti sono funzioni dal dominio A al codominio B? Per le funzioni dire se sono iniezioni, suriezioni o biiezioni.

(a) $A = \{a, b, c, d\}$, $B = \{1, 2, 3, 4\}$,
$f = \{(a, 1), (a, 2), (b, 1), (c, 2), (d, 3)\}$

(b) $A = \{-2, -1, 0, 1, 2\}$, $B = \{0, 1, 4\}$, $f(x) = x^2$

(c) $A = \{-2, -1, 0, 1, 2\}$, $B = \{0, 1, 2, 3, 4\}$, $f(x) = x^2$

(d) $A = \{0, 1, 2\}$, $B = \{0, 1, 4\}$, $f(x) = x^2$

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

(f) $A = \{a, b, c, d\}$, $B = \{1, 2, 3, 4, 5\}$, $f = \{(a, 1), (b, 5), (c, 4), (d, 3)\}$

(g) $A = \{a, b, c, d\}$, $B = \{1, 2, 3, 4\}$, $f = \{(a, 1), (b, 2), (c, 3)\}$

Definizione 3.12. Data una funzione $f : A \to B$, l'insieme degli elementi del codominio che sono immagini di qualche elemento del dominio è detto l'**immagine** della funzione e denotato con $f(A)$.

Esercizio 3.13. Qual è l'immagine di ciascuna delle funzioni dell'esempio precedente?

Molte funzioni che si utilizzano in Matematica Discreta hanno come dominio insiemi di numeri interi.

Definizione 3.14. Dato un insieme X, una funzione s da $\{1, 2, \dots\}$ in X è detta una **successione di elementi di** X. Si preferisce denotare con s_i l'immagine $s(i)$ dell'intero i, che è anche chiamato l'**i-esimo termine** della successione. Una **lista** o **stringa** o n**-pla di elementi di** X è una funzione da $\{1, 2, \dots, n\}$ in X.

Esempio 3.15. Se definiamo $s(n) = -2n$ otteniamo la successione $s : \{1, 2, \dots, n\} \to \mathbb{Z}$ seguente

$$-2, -4, -6, \dots.$$

Esempio 3.16. Scrivere il terzo termine della successione $s : \{1, 2, \dots, n\} \to \mathbb{N}$ definita da

$$s_i = i(i - 1) + 1.$$

Soluzione. $s_3 = 3(2) + 1 = 7$.

Esempio 3.17. Dato $X = \{a, b, c\}$, se definiamo $f : \{1, 2, 3, 4\} \to \{a, b, c\}$ con

$$f(1) = b, \quad f(2) = a, \quad f(3) = a, \quad f(4) = c,$$

otteniamo una stringa su X che può essere scritta nella forma *baac* o una 4-pla che può essere scritta nella forma (b, a, a, c).

Torniamo ad occuparci di relazioni su un insieme per dare le seguenti definizioni.

Definizione 3.18. Una relazione R su un insieme X è detta **riflessiva** se ogni elemento di X è in relazione con se stesso. In simboli, R è riflessiva se

$$\forall x \in X \quad (x, x) \in R.$$

Esercizio 3.19. Quali delle seguenti relazioni su $\{1, 2, 3\}$ è riflessiva?

(a) $\{(1,1), (1,2), (2,2), (1,3), (3,2), (3,3)\}$

(b) $\{(1,1), (2,2), (2,3), (3,2), (3,1), (1,3)\}$.

Definizione 3.20. Una relazione R su un insieme X è detta **simmetrica** se ogniqualvolta x è in relazione con y segue anche che y è in relazione con x.

In simboli, R è simmetrica se

$$\forall x, y \in R \quad (x, y) \in R \implies (y, x) \in R.$$

Esercizio 3.21. Quali delle seguenti relazioni su $\{1, 2, 3\}$ è simmetrica?

(a) $\{(1,1), (1,2), (2,2), (1,3), (3,2), (3,3)\}$

(b) $\{(1,1), (2,2), (2,3), (3,2), (3,1), (1,3)\}$.

Definizione 3.22. Una relazione R su un insieme X è detta **transitiva** se ogniqualvolta x è in relazione con y e y è in relazione con z segue anche che x è in relazione con z.

In simboli, R è transitiva se

$$\forall x, y, z \in R \quad (x, y) \in R \wedge (y, z) \in R \implies (x, z) \in R.$$

Esercizio 3.23. Scrivere la relazione *minore o uguale* sull'insieme $X = \{1, 2, 3, 4\}$ come sottoinsieme di $X \times X$ e verificare che è transitiva.

Definizione 3.24. Una relazione su un insieme che è riflessiva, simmetrica e transitiva è detta una **relazione di equivalenza**.

Esercizio 3.25. Determinare se ciascuna delle relazioni seguenti è di equivalenza sull'insieme $\{1, 2, 3, 4, 5, 6\}$.

(a) $\{(1,3), (3,5), (5,1), (3,1), (5,3), (1,5), (4,6), (6,4), (4,2), (2,4),$
$(2,6), (6,2), (1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$

(b) $(x, y) \in R$ se $x - y = \pm 1$

(c) $(x, y) \in R$ se $|x - y| \leq 1$

(d) $(x, y) \in R$ se $x^2 - 6x = y^2 - 6y$

Definizione 3.26. Data una relazione di equivalenza R su un insieme X, per ogni $x \in X$, l'insieme degli elementi y che sono in relazione con x, denotato con $\overline{x}$, è detta la **classe di equivalenza rappresentata da** x. L'elemento x è detto un **rappresentante** della classe $\overline{x}$. In simboli

$$\overline{x} = \{y \in X \mid (x, y) \in R\}.$$

Esercizio 3.27. Verificato che la relazione (a) dell'esempio precedente è di equivalenza, si determinino le classi di equivalenza.

Esempio 3.28. Data la relazione R sull'insieme $X = \{1, 2, 4, 5, 8, 12, 15, 16\}$ definita da

$$(a, b) \in R \quad \text{se } 7 \mid b - a,$$

verificare che R è di equivalenza e determinare le classi di equivalenza.

Soluzione.

R è riflessiva in quanto $7 \mid a - a = 0,\ \forall a \in X$.

R è simmetrica in quanto da $7 \mid b - a$ segue che $7 \mid a - b$.

R è transitiva in quanto se $7 \mid b - a$ e $7 \mid c - b$, allora $7 \mid (c - b) + (b - a)$, cioè $7 \mid c - a$.

Le classi di equivalenza sono

$$\overline{1} = \overline{8} = \overline{15} = \{1, 8, 15\}$$
$$\overline{2} = \overline{16} = \{2, 16\}$$
$$\overline{4} = \{4\}$$
$$\overline{5} = \overline{12} = \{5, 12\}.$$

4. Classi di equivalenza e classi di congruenza

Definizione 4.1. Una **partizione** di un insieme X è una famiglia di sottoinsiemi non vuoti di X tali che:

(a) X è l'unione di tutti questi sottoinsiemi;

(b) i sottoinsiemi sono a due a due **disgiunti**, cioè non hanno a due a due elementi in comune.

Esempio 4.2. Dato $X = \{1, 2, \ldots, 14\}$, i sottoinsiemi

$$X_1 = \{1, 7, 13\},\ X_2 = \{2, 3, 4, 8, 9\},\ X_3 = \{5\},\ X_4 = \{6, 10, 11, 12, 14\}$$

formano una partizione di X.

Esempio 4.3. Dato l'insieme $\mathbb{Z}$ degli interi relativi, i due sottoinsiemi $\mathbb{P}$ dei numeri pari e $\mathbb{D}$ dei numeri dispari formano una partizione di $\mathbb{Z}$.

Utilizzando le proprietà riflessiva, simmetrica e transitiva di una relazione di equivalenza R, si può dimostrare il seguente teorema

Teorema 4.4. *Se R è una relazione di equivalenza su un insieme X, si ha:*

(a) *se $(x, y) \in R$ allora $\overline{x} = \overline{y}$;*

(b) *la famiglia delle classi di equivalenza forma una partizione di X.*

Esempio 4.5. La partizione $\{\mathbb{P}, \mathbb{D}\}$ in pari e dispari di $\mathbb{Z}$ è individuata dalla relazione di equivalenza R data da:

$$(a, b) \in R \qquad \text{se e solo se} \quad 2 \mid b - a$$

con $\mathbb{P} = \overline{0}$ e $\mathbb{D} = \overline{1}$.

L'esempio precedente si può generalizzare scegliendo un qualunque intero n al posto di 2.

Definizione 4.6. Fissato un intero n, $(n \geq 2)$, due interi $a, b \in \mathbb{Z}$ sono detti **congruenti modulo** n se $n \mid b - a$. In questo caso si scrive

$$a \equiv b \mod n.$$

Esercizio 4.7. 87 è congruente a 3 modulo 12? 26 è congruente a -13 modulo 7?

Il significato di questa definizione è spiegato dai due seguenti teoremi.

Teorema 4.8. *La relazione di congruenza modulo n è una relazione di equivalenza su $\mathbb{Z}$. Le classi di equivalenza sono dette le* **classi di congruenza modulo** *n o anche* **classi di resto modulo** *n. L'insieme delle classi di congruenza modulo n è denotato con $\mathbb{Z}_n$.*

Dimostrazione. Si lascia per esercizio. (Suggerimento: imitare i passaggi dell'Esempio 3.28). □

Teorema 4.9. *Due interi sono congruenti modulo n se e solo se divisi per n danno lo stesso resto.*

Dimostrazione. Consideriamo gli interi a e b. Dividendoli per n, otteniamo

$$a = q_1 n + r_1, \quad b = q_2 n + r_2, \quad \text{con} \quad 0 \leq r_1, r_2 < n.$$

Se $r_1 = r_2$, allora $b - a = (q_2 - q_1)n$ cioè a e b sono congruenti modulo n.

Viceversa, se $a \equiv b \mod n$, allora $b - a = kn$, da cui

$$b = a + kn = q_1 n + r_1 + kn = (q_1 + k)n + r_1.$$

Dal Teorema della Divisione (Teorema 1.1) il resto è univocamente determinato e quindi, confrontando $b = q_2 n + r_2$ e $b = (q_1 + k)n + r_1$, si ha che $r_1 = r_2$. □

Il teorema precedente assicura che si può scegliere come rappresentante di ogni classe il resto della divisione per n. Possiamo allora usare la notazione seguente:

$$\mathbb{Z}_n = \{\overline{0}, \overline{1}, \overline{2}, \cdots, \overline{n-1}\},$$

essendo $0, 1, 2, \cdots, n-1$ i possibili resti della divisione di un numero per n.

Esempio 4.10. Dato $n = 3$, $\mathbb{Z}_3$ è formata dalle 3 classi di congruenza seguenti:

$$\overline{0} = \{\cdots, -6, -3, 0, 3, 6, 9, \cdots\},$$
$$\overline{1} = \{\cdots, -5, -2, 1, 4, 7, 10, \cdots\},$$
$$\overline{2} = \{\cdots, -4, -1, 2, 5, 8, 11, \cdots\}.$$

È possibile definire delle operazioni di somma e prodotto in $\mathbb{Z}_n$.

Teorema 4.11. *Dati gli interi a, b, a', b', se*

$$a \equiv b \mod n, \qquad a' \equiv b' \mod n$$

allora

$$a + a' \equiv b + b' \mod n, \qquad aa' \equiv bb' \mod n$$

cioè se sono congruenti gli argomenti, lo sono anche i risultati della composizione di somma e prodotto.

Dimostrazione. Da $a \equiv b \mod n$ segue che $b = a + kn$, e da $a' \equiv b' \mod n$ segue che $b' = a' + hn$. Quindi, sommando e moltiplicando, si ottiene

$$b + b' = a + a' + (k + h)n$$

e

$$bb' = aa' + (ka' + ha + khn)n$$

cioè la tesi. □

Il teorema precedente consente di introdurre una somma e un prodotto tra classi, utilizzando la somma e il prodotto tra gli interi che rappresentano le classi, nel seguente modo:

$$\overline{a} + \overline{b} = \overline{a + b}, \qquad \overline{a} \cdot \overline{b} = \overline{a \cdot b}.$$

Le due operazioni di somma e prodotto definite su $\mathbb{Z}_n$ soddisfano alle analoghe proprietà della somma e prodotto in $\mathbb{Z}$, cioè alla proprietà associativa di somma e prodotto, alla proprietà commutativa di somma e prodotto, alla proprietà distributiva del prodotto rispetto alla somma, all'esistenza degli elementi neutri rispetto alla somma e al prodotto, ...

Si costruisce in questo modo, per ogni intero n fissato, una struttura algebrica denotata con $(\mathbb{Z}_n, +, \cdot)$. L'**aritmetica modulare** è proprio lo studio delle proprietà di questa struttura.

Occupiamoci, come primo problema, della risoluzione delle equazioni lineari in una incognita in $\mathbb{Z}_n$.

Definizione 4.12. Una **equazione lineare in una variabile in** $\mathbb{Z}_n$ è una espressione della forma

$$\overline{a}x = \overline{b}, \qquad \text{con} \quad \overline{a}, \overline{b} \in \mathbb{Z}_n$$

Una **soluzione** è una classe $\overline{x}_0$ che soddisfa l'equazione.

Sull'esistenza di soluzioni si ha il seguente risultato.

Teorema 4.13. *L'equazione $\overline{a}x = \overline{b}$ ha soluzioni in $\mathbb{Z}_n$ se e solo se (a, n) divide b.*

Dimostrazione. L'equazione $\overline{a}x = \overline{b}$ ha soluzione $\overline{x_0}$ in $\mathbb{Z}_n$, cioè $\overline{a} \cdot \overline{x_0} = \overline{b}$,

$\iff ax_0 - b = hn$ (che si può riscrivere nella forma $ax_0 + n(-h) = b$),

$\iff (x_0, -h)$ è una soluzione dell'equazione in $\mathbb{Z} \quad ax + ny = b$,

$\iff (a, n)$ divide b.

Da questo segue che, se (x_1, y_1) è la soluzione generale di $ax + ny = b$, cioè se

$$x_1 = x_0 + \frac{n}{d}k, \qquad y_1 = -h - \frac{na}{d}k, \qquad \forall k \in \mathbb{Z}, \quad \text{con} \quad d = (a, n),$$

gli interi $x_1 = x_0 + \dfrac{n}{d}k, \quad \forall k \in \mathbb{Z}$ sono i rappresentanti delle soluzioni di $\overline{a}x = \overline{b}$ in $\mathbb{Z}_n$.

Si ottengono quindi le d soluzioni date da:

$$\overline{x_0}, \qquad \overline{x_1} = \overline{x_0 + \frac{n}{d}}, \qquad \overline{x_2} = \overline{x_0 + \frac{2n}{d}}, \ldots, \qquad \overline{x_{d-1}} = \overline{x_0 + \frac{(d-1)n}{d}}.$$

□

Esempio 4.14. L'equazione $\overline{6}x = \overline{5}$ non ha soluzioni in $\mathbb{Z}_4$ in quanto $(6, 4) = 2$ non divide 5.

Esempio 4.15. L'equazione $\overline{12}x = \overline{15}$ in $\mathbb{Z}_{39}$ ha soluzioni perché $3 = (12, 39)$ divide 15. Per determinarle si risolve l'equazione $12x + 39y = 15$ in $\mathbb{Z}$ che dà soluzione generale $(-15 + 13k, 5 - 4k), \quad \forall k \in \mathbb{Z}$. Allora $\overline{x_1} = \overline{-15 + 13k}, \quad \forall k \in \mathbb{Z}$, è la soluzione generale dell'equazione $\overline{12}x = \overline{15}$. Si hanno quindi le $3 = (12, 39)$ soluzioni:

$$\overline{-15 + 0} = \overline{24}, \qquad \overline{-15 + 13} = \overline{37}, \qquad \overline{-15 + 26} = \overline{11}.$$

5. Il teorema di Eulero-Fermat e il metodo di crittografia a chiave pubblica RSA

In questa sezione vedremo una sorprendente applicazione della teoria delle congruenze. Nel 1736, Eulero diede una dimostrazione di una affermazione di Fermat (1601-1665) che riguardava la relazione fra le potenze di un numero e le sue classi di congruenza modulo un numero primo. Eulero riuscì anche a dimostrare una generalizzazione di questa affermazione, valida per ogni numero intero n, e non solo per i numeri primi. Ed è proprio usando questo teorema (o meglio, una sua variante) che recentemente, Rivest, Shamir e Adleman hanno inventato un metodo di crittografia estremamente efficiente e sicuro. Il metodo RSA è uno dei più usati oggi (per esempio, è implementato nel programma di pubblico dominio PGP (Pretty Good Privacy)) ed è usato da molte autorità di certificazione riconosciute in rete.

Il Teorema di Eulero-Fermat. Enunciamo subito quello che è comunemente noto come il *Piccolo Teorema di Fermat*.

Teorema 5.1 ((Piccolo Teorema di Fermat))**.** *Sia p un numero primo, e sia a un intero non divisibile per p. Allora*

$$a^{p-1} \equiv 1 \mod p$$

Dimostrazione. Consideriamo i multipli di a

$$m_1 = a, \quad m_2 = 2a, \quad m_3 = 3a, \quad \ldots, \quad m_{p-1} = (p-1)a.$$

Nessuna coppia di questi numeri interi può essere congrua modulo p: infatti, in tal caso, p sarebbe un divisore di $m_r - m_s = (r-s)a$, dove r ed s sono numeri interi $1 \leq r < s \leq p-1$, e poiché $r-s$ è minore di p ed inoltre p non divide a per ipotesi, questo non è possibile. Dunque i numeri $m_1, m_2, \ldots, m_{p-1}$ sono congrui ai numeri 1, 2, ..., $p-1$, considerati in un ordine opportuno. Ma allora moltiplicando si ha

$$m_1 m_2 \cdots m_{p-1} = 1 \cdot 2 \cdot 3 \cdots (p-1)a^{p-1} \equiv 1 \cdot 2 \cdot 3 \cdots (p-1) \mod p,$$

e se per brevità poniamo $K = 1 \cdot 2 \cdot 3 \cdots (p-1)$ possiamo scrivere

$$K(a^{p-1} - 1) \equiv 0 \mod p.$$

Ma K non è divisibile per p, perché nessuno dei suoi fattori lo è, e dunque deve essere $(a^{p-1} - 1)$ divisibile per p, cioè

$$a^{p-1} \equiv 1 \mod p.$$

□

Notiamo che se n non è un numero primo, allora non si può dire che $a^{n-1} \equiv 1 \mod n$, in generale. Per esempio, se $n = 4$ e $a = 3$, si ha $3^3 = 27 \equiv 3 \mod 4$. Ci si può dunque chiedere se ci sia una potenza di

a che sia congrua ad 1 modulo n. La risposta è sì, e l'esponente corretto è stato trovato da Eulero.

Definizione 5.2. Per un numero intero positivo n, indichiamo con $\varphi(n)$ il numero degli interi compresi fra 1 ed n coprimi con n.

Per esempio,

$\varphi(1) = 1$	perché 1 è coprimo con 1,
$\varphi(2) = 1$	1 è coprimo con 2,
$\varphi(3) = 2$	1 e 2 sono coprimi con 3,
$\varphi(4) = 2$	1 e 3 sono coprimi con 4,
$\varphi(5) = 4$	1, 2, 3, 4 sono coprimi con 5,
$\varphi(6) = 2$	1 e 5 sono coprimi con 6,
$\varphi(7) = 6$	1, 2, 3, 4, 5, 6 sono coprimi con 7,
$\varphi(8) = 4$	1, 3, 5, 7 sono coprimi con 8,
$\varphi(9) = 6$	1, 2, 4, 5, 7, 8 sono coprimi con 9,
$\varphi(10) = 4$	1, 3, 7, 9 sono coprimi con 10,
$\dots$	$\dots$

Una proprietà della funzione φ è evidente:

$$\text{se } p \text{ è un numero primo, allora } \varphi(p) = p - 1,$$

in quanto *tutti* i numeri 1, 2, ..., $p - 1$ sono coprimi con p. Altrettanto semplice è calcolare $\varphi(n)$ per un numero che sia potenza di un primo. La formula è la seguente:

$$\text{se } n = p^a, \text{ con } p \text{ numero primo, allora } \varphi(n) = p^a - p^{a-1}.$$

Infatti, i numeri non coprimi con p^a sono tutti i multipli di p, e ce ne sono p^{a-1} minori di p^a. Notiamo che se $a = 1$ (cioè $n = p$ è primo) ritroviamo la stessa formula di prima. Per esempio, $\varphi(125) = \varphi(5^3) = 5^3 - 5^2 = 100$. L'ultima proprietà della funzione φ è quella che consente di calcolare $\varphi(n)$ per un numero qualsiasi. La formula è

$$\text{se } m \text{ e } n \text{ sono coprimi, cioè } (m, n) = 1, \text{ allora } \varphi(m \cdot n) = \varphi(m) \cdot \varphi(n).$$

Per esempio, usando la tabella precedente, si ha $\varphi(21) = \varphi(3 \cdot 7) = 12$, $\varphi(30) = \varphi(5 \cdot 6) = 8$, e così via. Come vediamo, è molto semplice calcolare $\varphi(n)$ se si conosce la fattorizzazione di n. Questa è la proprietà che verrà sfruttata nell'algoritmo RSA.

Vediamo ora il teorema di Eulero-Fermat:

Teorema 5.3. *Sia n un intero positivo e a un intero coprimo con n. Allora*

$$a^{\varphi(n)} \equiv 1 \mod n.$$

Osserviamo che, se n è primo, allora $\varphi(n) = n - 1$ e quindi ritroviamo il Teorema di Fermat.

Ci si può chiedere cosa capiti se $(a, n) \neq 1$. Per esempio, se $n = 6$, $a = 2$, si ha

$$2^2 \equiv 4 \mod 6$$
$$2^3 \equiv 2 \mod 6$$
$$2^4 \equiv 4 \mod 6$$

e così via, e vediamo che le potenze di 2 modulo 6 alternano fra 2 e 4 e non si ottiene mai 1. Osserviamo però che il teorema di Eulero-Fermat si può enunciare in modo equivalente come

Teorema 5.4 (di Eulero-Fermat)**.** *Sia n un intero positivo ed a un intero coprimo con n. Allora*

$$a^{\varphi(n)+1} \equiv a \mod n.$$

e in questa forma la conclusione del teorema vale anche per $n = 6$, $a = 2$. Non è però possibile eliminare del tutto l'ipotesi che a ed n siano coprimi come mostra l'esempio $n = 12$, $a = 2$: si ha

$$2^2 \equiv 4 \mod 12$$
$$2^3 \equiv 8 \mod 12$$
$$2^4 \equiv 4 \mod 12$$
$$2^5 \equiv 8 \mod 12$$

e tutte le potenze successive alternano fra 4 e 8, e dunque non si riottiene più 2. Se però facciamo un'ipotesi su n, è possibile eliminare la condizione che a ed n siano coprimi.

Definizione 5.5. Un intero n si dice **libero da quadrati** se è il prodotto di numeri primi distinti, cioè se non è divisibile per nessun quadrato > 1.

Per esempio, $6 = 2 \cdot 3$ è libero da quadrati, mentre $12 = 2 \cdot 2 \cdot 3$ non lo è. Con questa definizione si può enunciare il teorema seguente:

Teorema 5.6. *Se $n = p_1 \cdot p_2 \cdots p_k$ è libero di quadrati allora*

$$a^{\varphi(n)+1} \equiv a \mod n$$

per **ogni** *intero a.*

Osserviamo che la conclusione è la stessa del Teorema di Eulero-Fermat, ma l'ipotesi $(a, n) = 1$ è stata sostituita dall'ipotesi n libero da quadrati. Usando il Teorema 5.6 si può dimostrare facilmente per induzione che vale il seguente risultato più generale:

Corollario 5.7. *Se $n = p_1 \cdot p_2 \cdots p_k$ è libero di quadrati allora*

$$a^{h\varphi(n)+1} \equiv a \mod n$$

per **ogni** *intero a e per ogni intero $h \geq 1$.*

Crittografia a chiave pubblica. Il problema della crittografia è semplice da enunciare: vi sono due persone, il Mittente e il Ricevente, che vogliono comunicare fra loro senza che nessun altro possa leggere i loro messaggi. Dato un messaggio M, che possiamo immaginare essere una sequenza di simboli o di numeri, il Mittente usa una funzione f (un algoritmo) che trasforma M in un'altra sequenza C e invia questa sequenza. Il Ricevente usa un'altra funzione g che trasforma C di nuovo nel messaggio originale M. Naturalmente, se qualcuno intercetta C non deve essere in grado di usare la funzione g per decodificare il messaggio.

Per ottenere ciò, di solito si usa una **chiave** (per esempio, una password). Dunque la funzione che codifica f non ha come argomento solo il messaggio M ma anche un altro parametro, la chiave k. In simboli

$$f(M, k) = C$$

e la funzione g è tale che

$$g(C, k) = M$$

cosicché, conoscendo la chiave si può decrittare il messaggio. Questo schema, che usa la stessa chiave k sia per la codifica che per la decodifica si chiama **simmetrico**.

In crittografia si suppone che tutti siano a conoscenza dell'algoritmo che usano f e g, perché è sempre possibile avere questa informazione mediante attività di spionaggio o simili. Basare la segretezza delle comunicazioni sulla segretezza dell'algoritmo non è una buona idea. Invece la segretezza è basata sul fatto che solo il Mittente e il Ricevente conoscono la chiave. Infatti, l'algoritmo viene usato sempre, mentre la chiave può essere cambiata ogni volta e così, anche se qualcuno "ruba" una chiave, la può usare solo poche volte.

Questa situazione (cambiare spesso la chiave) però richiede un modo di comunicare sicuro fra il Mittente e il Ricevente. Un altro problema è il seguente: se vi è un Ricevente, ma molti Mittenti, tutti i Mittenti conoscono la chiave, e possono leggere i messaggi l'uno dell'altro, cosa che un Mittente certamente non vuole. Questo è il problema che il metodo a chiave pubblica vuole risolvere. Lo schema è il seguente:

> *Il Ricevente comunica pubblicamente una chiave e. Il Mittente codifica il messaggio usando la funzione $f(M, e) = C$ e*

> *trasmette il messaggio. Il Ricevente possiede un'altra chiave d privata. Il Ricevente decodifica il messaggio usando $g(C, d) = M$.*

Questo elimina il problema della distribuzione delle chiavi, e anche la chiave che serve per codificare non è quella che decodifica, e quindi solo il Ricevente può leggere i messaggi. Questo schema è stato proposto per la prima volta da Whitfield Diffie e Martin Hellman nel 1976. Uno schema differente di crittografia a chiave pubblica è stato proposto nello stesso anno da Ralph Merkle.

Per fare un'analogia, un metodo simmetrico è come una cassaforte, e la chiave è la combinazione per aprirla. Una persona che ha la chiave può aprire la cassaforte e depositare un documento (codificare un messaggio). Un'altra persona con la chiave può aprire la cassaforte e prendere il documento (decodificare il messaggio). Chi non ha la chiave (l'avversario, la spia, ecc.) deve scassinare la cassaforte. Un metodo a chiave pubblica invece è come una buca delle lettere. Chiunque può infilare una lettera nella buca (codificare il messaggio con la chiave pubblica), ma solo il proprietario della buca ha la chiave per aprirla facilmente (decodificare il messaggio). Tutti gli altri devono scassinare la buca delle lettere, e questo è possibile, e anche più facile che scassinare una cassaforte, ma richiede tempo. Se il padrone della buca ritira spesso la sua posta oppure se cambia spesso la serratura, la sua posta è al sicuro.

Naturalmente, la domanda è: funzioni f e g con le proprietà richieste da questo schema esistono? Si può pensare che se si conosce la chiave e usata per codificare il messaggio, e la funzione f che effettua la codifica, deve essere possibile in qualche modo decodificare il messaggio, e cioè un simile schema non può funzionare. L'intuizione di Rivest, Shamir e Adleman (1978) è stata che è possibile costruire funzioni f e g tali che:

(1) mediante la chiave privata d la decodifica è semplice e rapida, e

(2) usando la conoscenza della funzione f e della chiave pubblica e, la decodifica è possibile ma richiede troppo tempo per essere utile.

Per esempio, troppo tempo può voler dire che eseguire tutti i calcoli necessari, anche usando i computer più veloci a disposizione, richiede alcuni anni. È chiaro che leggere un messaggio segreto alcuni anni dopo che è stato scritto non è di grande utilità.

Descriviamo ora l'algoritmo RSA che, fra tutti gli algoritmi proposti per la crittografia a chiave pubblica che si sono rivelati sicuri, è il più semplice sia da capire che da implementare. Il Ricevente sceglie due numeri primi p e q, molto grandi, li moltiplica e forma il numero $n = pq$, che è libero da

quadrati. Sappiamo che $\varphi(n) = (p-1)(q-1)$. Ora sceglie un numero e che sia coprimo con $\varphi(n)$. La chiave pubblica per codificare i messaggi è la coppia (n, e). La funzione f per la codifica è la seguente:

- Si trasforma il messaggio in un numero M, minore di n. Se il messaggio è molto lungo, lo si spezza in blocchi più piccoli $M_1, \ldots, M_k$.
- Si calcola $C_i = M_i^e \mod n$ per ogni i.
- Si trasmette $C_1, \ldots C_k$.

Per decodificare, si procede nel modo seguente: consideriamo l'equazione

$$ex \equiv 1 \mod \varphi(n).$$

Poiché e è coprimo con $\varphi(n)$, questa equazione ha una soluzione d, per cui vale $ed = 1 + h\varphi(n)$. Allora la funzione g di decodifica è:

- Si calcola $M = C^d \mod n$.

Infatti

$$C^d = (M^e)^d = M^{ed} = M^{1+h\varphi(n)} \equiv M \mod n$$

dove la congruenza è vera in virtù del Corollario 5.7.

Facciamo un esempio semplice, lasciando i calcoli per esercizio. Se prendiamo $p = 47$ e $q = 71$, allora

$$n = 47 \cdot 71 = 3337$$

La chiave pubblica e deve essere coprima con

$$(p-1)(q-1) = 46 \cdot 70 = 3220$$

Scegliamo (a caso) $e = 79$. In questo caso, calcolando con l'algoritmo di Euclide si ottiene

$$\begin{aligned} \mathbf{3220} &= 40 \cdot \mathbf{79} + \mathbf{60} \\ \mathbf{79} &= 1 \cdot \mathbf{60} + \mathbf{19} \\ \mathbf{60} &= 3 \cdot \mathbf{19} + \mathbf{3} \\ \mathbf{19} &= 6 \cdot \mathbf{3} + \mathbf{1} \end{aligned}$$

Calcolando con i resti si ottiene

$$(79, 3220) = 1 = 1019 \cdot 79 - 25 \cdot 3220$$

e dunque abbiamo la chiave privata $d = 1019$. Rendiamo pubblici n ed e, teniamo segreto d ed eliminiamo p e q.

Per codificare il messaggio

$$M = 6882326879666683$$

per prima cosa lo spezziamo in blocchi più piccoli. Poiché $n = 3337$ ha quattro cifre, blocchi di tre cifre vanno bene. Dunque il messaggio si spezza in

$$M_1 = 688$$
$$M_2 = 232$$
$$M_3 = 687$$
$$M_4 = 966$$
$$M_5 = 668$$
$$M_6 = 003$$

Il primo blocco è codificato da

$$688^{79} \mod 3337 = 1570 = C_1$$

Notiamo che non bisogna elevare 688 alla 79-esima potenza (numero molto grande) e poi ridurre modulo 3337. Invece, si calcola 688^2 e si riduce modulo 337, poi si moltiplica di nuovo per 688 e si riduce nuovamente modulo 3337 e così via, e in ogni passo i numeri sono piccoli. Ci sono metodi per calcolare queste potenze in modo ancora più veloce di quello descritto, rendendo così la codifica del messaggio veloce.

Ripetendo gli stessi passi su tutti gli altri blocchi otteniamo il messaggio codificato

$$C = 1570\ 2756\ 2091\ 2276\ 2423\ 158$$

Decodificare il messaggio richiede le stesse operazioni usando l'esponente $d = 1019$ sui blocchi di C. Si ha

$$1570^{1019} \mod 3337 = 688 = M_1$$

e in modo simile per i blocchi successivi.

Perché questo metodo è sicuro? In fondo, per ottenere la chiave privata d basta conoscere $\varphi(n)$ e tutti conoscono n. Il modo più rapido di calcolare $\varphi(n)$ è avere la decomposizione $n = pq$ e cioè dobbiamo fattorizzare n. Ora, gli algoritmi più veloci noti al giorno d'oggi per fattorizzare un numero n hanno bisogno di un numero di passi proporzionale a $\sqrt{n}$. Se p e q hanno 100 cifre l'uno, n ha 200 cifre, cioè è dell'ordine di grandezza di 10^{200} e la sua radice quadrata è dell'ordine di 10^{100}.

In un anno vi sono circa $3 \cdot 10^7$ secondi, e l'età dell'Universo è attualmente stimata in circa 15 miliardi di anni e cioè in $(1.5) \cdot 10^{10}$ anni, che equivale a circa $(4.5) \cdot 10^{17}$ secondi. Un processore con velocità 1 GHz (1 gigahertz) esegue 1 miliardo, cioè 10^9, di cicli al secondo. Se supponiamo di avere un miliardo di processori, ognuno che lavora alla frequenza di 1 GHz,

e che eseguono uno dei passi dell'algoritmo in ognuno dei cicli, abbiamo la possibilità di eseguire $10^9 \cdot 10^9 = 10^{18}$ passi al secondo. Se il nostro calcolatore ha la possibilità di funzionare per un tempo pari alla vita dell'Universo, farà $(4.5) \cdot 10^{17} \cdot 10^{18} = (4.5) \cdot 10^{35}$ passi. È evidente che non vi è speranza di fattorizzare il numero n.

Si potrebbe osservare: non devo trovare i fattori p e q, ma solo calcolare $\varphi(n)$. Nel nostro caso però, questo è equivalente. Se conosciamo p e q sappiamo calcolare $\varphi(n)$ semplicemente, bastano due sottrazioni e una moltiplicazione:

$$\varphi(n) = (p-1)(q-1).$$

Viceversa, conoscendo n e $\varphi(n)$ è facile trovare p e q. Infatti sappiamo che $n = pq$ e che $\varphi(n) = (p-1)(q-1) = pq - p - q + 1$ e dunque si ha

$$p + q = n - \varphi(n) + 1$$

Dunque conosciamo la somma e il prodotto di p e q, e basta risolvere una equazione di secondo grado per trovare p e q. Ponendo $B = pq$ e $A = p + q$ si ha che l'equazione

$$x^2 - Ax + B = 0$$

ha come soluzioni esattamente p e q. Quindi, conoscendo $\varphi(n)$ si ha un algoritmo che in pochi passi, una somma e la risoluzione di un'equazione di secondo grado, consente di fattorizzare il numero n. Dunque la complessità di calcolo per determinare $\varphi(n)$ è equivalente alla complessità di fattorizzare il numero n che, come abbiamo visto, è piuttosto grande.

Questo spiega perché il metodo è sicuro. È però importante osservare che nessuno ha mai dimostrato che fattorizzare il numero n è veramente così difficile, ma solo che i migliori algoritmi **conosciuti** sono troppo lenti. Potrebbe capitare che un giorno qualcuno scopra un metodo più efficiente per fattorizzare i numeri, ed allora il metodo RSA non sarebbe più sicuro.

La crittografia è un argomento affascinante, ed è un campo in cui l'avvento dei computer ha portato grandi novità e possibilità di utilizzo. È bene osservare che non sempre il Mittente ed il Ricevente sono esseri umani, ma potrebbero essere entrambi computer che devono comunicare fra loro, per esempio un telefono cellulare che comunica con la rete di telefonia (non vogliamo che un telefono non autorizzato possa mettersi in contatto ed usare le risorse della rete oppure che qualcuno possa intercettare la trasmissione e ascoltare le nostre telefonate), oppure un computer che manda una richiesta ad un server, per esempio per fare una stampa (non vogliamo che tutti possano usare le nostre stampanti) e così via. Quindi è importante avere dei metodi semplici per trasmettere in sicurezza.

La sicurezza di un sistema però non è solo nella matematica che si usa, ma anche nel protocollo. Possiamo avere il sistema più sicuro del mondo,

ma se riveliamo la chiave ad un estraneo la nostra sicurezza è compromessa. Per saperne di più sui problemi legati alla crittografia si può consultare il libro:

- Bruce Schneier, "Applied Cryptography Second Edition: protocols, algorithms, and source code in C", John Wiley & Sons, 1996.

È un libro abbastanza recente, molto completo e orientato agli aspetti concreti dei sistemi crittografici. Contiene una bibliografia di articoli e libri di crittografia con 1653 referenze.

Esercizi

(1) Trovare il massimo comun divisore di 721 e 448 ed esprimerlo nella forma

$$721m + 448n, \qquad \text{con} \quad m, n \in \mathbb{Z}.$$

(2) Trovare le soluzioni intere delle seguenti equazioni:
   (a) $5x + 73y = 1$
   (b) $112x + 6y = 2$
   (c) $112x + 6y = 4$
   (d) $966x + 686y = 70$.

(3) Trovare gli interi positivi x e y tali che $7x + 19y = 1921$ e $x + y$ sia minimo.

(4) Stabilire quando ciascuna delle seguenti congruenze ha soluzioni e, in caso positivo, trovare la più piccola soluzione non negativa:
   (a) $12x \equiv 7 \mod 21$
   (b) $12x \equiv 7 \mod 73$
   (c) $12x \equiv 7 \mod 35$
   (d) $12x \equiv 7 \mod 84$
   (e) $12x \equiv 7 \mod 46$.

(5) Risolvere l'equazione $\overline{14}x = \overline{21}$ in $\mathbb{Z}_{77}$.

(6) In $\mathbb{Z}_n$ una classe $\overline{a}$ è **invertibile** se esiste una classe $\overline{b} \in Z_n$, detta **inversa di** $\overline{a}$, tale che $\overline{a}\overline{b} = \overline{1}$, cioè se l'equazione $\overline{a}x = \overline{1}$ ha soluzione.
   (a) La classe $\overline{4}$ ha inversa in $\mathbb{Z}_{15}$? In caso positivo calcolarla.
   (b) La classe $\overline{7}$ ha inversa in $\mathbb{Z}_{21}$? In caso positivo calcolarla.
   (c) Quali sono le classi invertibili in $\mathbb{Z}_{15}$?
   (d) Quante sono le classi invertibili in $\mathbb{Z}_n$?

(7) Verificare che in $\mathbb{Z}_{26}$ si ha $\overline{3^{12}} = \overline{1}$.

(8) Provare che, per ogni intero n, $n^9 + 2n^7 + 3n^3 + 4n$ è divisibile per 5.
Suggerimento: Considerare i due casi

(a) $5 \mid n$ che si risolve direttamente;
(b) $5 \nmid n$ che si risolve utilizzando il teorema di Eulero-Fermat.

(9) Dividendo 3^{47} per 23 quale resto si ottiene?
Dividendo 3^{47} per 22 quale resto si ottiene?
Dividendo 3^{47} per 21 quale resto si ottiene?
Specificare nei tre casi quale risultato teorico si utilizza per semplificare i calcoli.

Capitolo 2

Calcolo combinatorio

1. Tecniche elementari di enumerazione

Dato un insieme S, diciamo che S ha **cardinalità** n, e scriviamo $|S| = n$, se esiste una biiezione f tra $\{1, 2, ..., n\}$ ed S.

Esempio 1.1. $S = \{x, y, z\}$ ha cardinalità $n = |S| = 3$ perché esiste una biiezione f tra l'insieme $\{1, 2, 3\}$ ed S, per esempio quella data da

$$f = \{(1, x), (2, y), (3, z)\}.$$

Principio della somma. *La cardinalità dell'unione di due insiemi disgiunti è la somma delle cardinalità dei due insiemi, cioè se* $|S| = m$, $|T| = n$ *ed* $S \cap T = \emptyset$ *allora* $|S \cup T| = m + n = |S| + |T|$.

Dimostrazione. Poiché $|S| = m$, $|T| = n$ esistono due biiezioni

$$f : \{1, 2, \ldots, m\} \to S, \quad g : \{1, 2, \ldots, n\} \to T.$$

È allora possibile definire una biiezione $h : \{1, 2, \ldots, m + n\} \to S \cup T$ nel seguente modo:

$$h(i) = f(i), \quad \forall i = 1, 2, \ldots, m, \qquad h(i + m) = g(i), \quad \forall i = 1, 2, \ldots, n.$$

Segue che $|S \cup T| = m + n = |S| + |T|$. □

Esempio 1.2. Quanti sono i numeri interi x, con $3 \leq x \leq 20$, che sono o numeri pari o numeri primi?

Soluzione. Si ha

$$S = \{x \mid 3 \leq x \leq 20 \text{ e } x \text{ pari}\} = \{4, 6, 8, 10, 12, 14, 16, 18, 20\}$$
$$T = \{x \mid 3 \leq x \leq 20 \text{ e } x \text{ primo}\} = \{3, 5, 7, 11, 13, 17, 19\}$$

Poiché $|S| = 9$, $|T| = 7$ e $S \cap T = \emptyset$, segue che $|S \cup T| = 9 + 7 = 16$.

Esempio 1.3.

(a) Quanti sono i modi differenti di pescare un asso o una regina da un mazzo di 52 carte?

(b) Quanti sono i modi di pescare un asso o una carta rossa da un mazzo di 52 carte?

Soluzione.

(a) Se indichiamo con A l'insieme dei modi di pescare un asso, si ha $|A| = 4$. Se indichiamo con B l'insieme dei modi di pescare una regina, si ha $|B| = 4$. Inoltre $A \cap B = \emptyset$. Segue $|A \cup B| = 4 + 4 = 8$.

(b) Se indichiamo con A l'insieme dei modi di pescare un asso, si ha $|A| = 4$. Se indichiamo con B l'insieme dei modi di pescare una carta rossa, si ha $|B| = 26$. Ma non possiamo concludere che la soluzione è $4 + 26 = 30$, in quanto si può pescare una carta rossa che è contemporaneamente un asso, cioè $A \cap B \neq \emptyset$. In questo caso il modo corretto di ragionare è, ad esempio, quello di considerare gli insiemi disgiunti $A' = \{\text{modi di pescare un asso nero}\}$ e $B' = \{\text{modi di pescare una carta rossa}\}$. Si ottiene $|A'| = 2$, $|B'| = 26$ e $A' \cap B' = \emptyset$. Inoltre in B' sono già conteggiate le due scelte per gli assi rossi. Quindi $|A' \cup B'| = 2 + 26 = 28$.

Una conseguenza immediata del principio della somma è:

Teorema 1.4. *Se S_1, S_2, ..., S_n sono insiemi* **a due a due disgiunti**, *allora la cardinalità dell'unione è la somma delle cardinalità, e cioè*

$$|S_1 \cup S_2 \cup \cdots \cup S_n| = |S_1| + |S_2| + \cdots + |S_n|$$

Dimostrazione. Per induzione su n. □

Osserviamo che dire a due a due disgiunti significa che per ogni $i \neq j$ si ha $S_i \cap S_j = \emptyset$, e questo è più restrittivo di $S_1 \cap S_2 \cap \cdots \cap S_n = \emptyset$.

Esempio 1.5. Se $S_1 = \{1, 2, 3, 4\}$, $S_2 = \{3, 4, 5\}$ e $S_3 = \{5, 6\}$, si ha $S_1 \cap S_2 \cap S_3 = \emptyset$, ma gli insiemi non sono a due a due disgiunti. Dunque in questo caso non si può applicare il Teorema 1.4 e in effetti l'unione $S_1 \cup S_2 \cup S_3 = \{1, 2, 3, 4, 5, 6\}$ ha cardinalità 6, mentre la somma delle cardinalità è 9.

Esempio 1.6. In quanti modi 6 persone possono sceglierne due per farne un presidente ed un segretario, se non si vuole che la stessa persona ricopra due ruoli?

Soluzione. Fare una scelta significa determinare una coppia (presidente, segretario). Ci sono 6 modi di scegliere il presidente ed una volta che il presidente è scelto ci sono 5 modi di scegliere il segretario e dunque l'insieme di tutte le possibili coppie è l'unione di 6 insiemi di cardinalità 5 che sono

a due a due disgiunti in quanto differiscono per il primo elemento. Quindi per il Teorema 1.4, l'insieme cercato ha cardinalità $5 + \cdots + 5 = 6 \cdot 5 = 30$.

Principio del prodotto. *Siano S e T due insiemi con $|S| = m$ e $|T| = n$ e sia $S \times T$ il prodotto cartesiano dei due insiemi, allora*

$$|S \times T| = mn.$$

Dimostrazione. Poniamo $S = \{s_1, s_2, \ldots, s_m\}$ e, per ogni indice i fra 1 e m, sia A_i il sottoinsieme del prodotto cartesiano formato delle coppie con primo elemento s_i. Ogni A_i ha cardinalità n e gli A_i sono a due a due disgiunti e allora, per il Teorema 1.4 si ha

$$|S \times T| = |A_1 \cup A_2 \cup ... \cup A_m| = mn.$$

□

Esempio 1.7. Se si hanno 3 tipi di panini e 5 tipi di affettati, in quanti modi possiamo fare un sandwich usando un solo tipo di panino e uno di affettato?

Soluzione. I possibili sandwiches sono coppie ordinate (panino, affettato) e sono quindi in numero di $3 \cdot 5 = 15$.

Il principio del prodotto consente di contare in quanti modi si possono scegliere due elementi (ad esempio presidente e segretario) da un insieme dato. Se vogliamo scegliere più di due elementi dobbiamo utilizzare m-ple al posto di coppie e per contarle usiamo il seguente principio.

Principio generalizzato del prodotto. *Siano S_1, S_2, ..., S_m insiemi con cardinalità $|S_1| = n_1$, $|S_2| = n_2$, ... $|S_m| = n_m$. Sia poi*

$$S_1 \times S_2 \times \cdots \times S_m = \{(s_1, s_2, \ldots, s_m), \quad \forall s_i \in S_i, \quad \forall i = 1, 2, \ldots, m\}$$

l'insieme di tutte le m-ple con elementi scelti rispettivamente in S_1, S_2, ..., S_m, detto il **prodotto cartesiano di S_1 per S_2 ... per S_m**. *Allora*

$$|S_1 \times S_2 \times \cdots \times S_m| = n_1 \cdot n_2 \cdot \cdots \cdot n_m.$$

Cioè ci sono in tutto $\prod_{i=1}^{m} n_i = n_1 \cdot n_2 \cdot \cdots \cdot n_m$ m-ple.

Esempio 1.8. Se si hanno 3 tipi di panini, 5 tipi di affettati e 4 tipi di salse, in quanti modi possiamo fare un sandwich usando un solo tipo di panino, un solo tipo di affettato e un solo tipo di salsa ?

Soluzione. In $3 \cdot 5 \cdot 4 = 60$ modi diversi.

Permutazioni su k elementi. Nell'esempio 1.6 abbiamo scelto 2 elementi (presidente e segretario) in un insieme di 6 persone e abbiamo contato i modi differenti di fare questa scelta. Fare una scelta di questo tipo significa, in altre parole, determinare una funzione iniettiva f da $\{1, 2\}$ nell'insieme delle 6 persone, dove $(f(1), f(2))$ rappresenta la coppia (presidente, segretario) scelta. Se vogliamo generalizzare questo ragionamento, diamo allora la seguente definizione.

Definizione 1.9. Dato un insieme S, una **permutazione su k elementi di S** è una funzione iniettiva di $K = \{1, 2, \ldots, k\}$ in S. Alla funzione iniettiva $f : K \to Y$ si associa la k-pla delle immagini distinte

$$(f(1), f(2), \ldots, f(k)) \in S \times S \times \cdots \times S.$$

Si può quindi anche dire che una permutazione è una k-pla o una lista di k elementi distinti di S.

Esempio 1.10. Elencare tutte le permutazioni su 3 elementi di a, b, c, d rappresentando ciascuna permutazione come una lista di lettere.

Soluzione. Scrivendo le liste in ordine alfabetico si ha:

abc	*abd*	*acb*	*acd*	*adb*	*adc*
bac	*bad*	*bca*	*bcd*	*bda*	*bdc*
cab	*cad*	*cba*	*cbd*	*cda*	*cdb*
dab	*dac*	*dba*	*dbc*	*dca*	*dcb*

Notiamo che vi sono 24 permutazioni su 3 elementi scelti da un insieme di 4 elementi.

Definizione 1.11. Se S ha cardinalità n, una permutazione sugli n elementi di S è detta semplicemente una **permutazione di S.**

Esempio 1.12. Elencare tutte le permutazioni di $S = \{a, b, c\}$, rappresentandole come liste di lettere.

Soluzione.

abc	*acb*	*bac*	*bca*	*cab*	*cba*

Vi sono 6 permutazioni.

Ci poniamo ora il problema di contare il numero delle permutazioni. Per fare questo, cominciamo con la seguente definizione.

Definizione 1.13. Per ogni intero positivo n, il **fattoriale di n**, denotato con $n!$ e detto anche n **fattoriale**, è il numero

$$n! = n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1.$$

In altre parole $n!$ è il prodotto di tutti gli interi da 1 ad n. Se $n = 0$, per convenzione si pone

$$0! = 1.$$

La formula che calcola il numero di permutazioni è data dal

Teorema 1.14. *Dato un insieme S di cardinalità n, il numero delle permutazioni su k elementi di S è*

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

Inoltre il numero delle permutazioni di S è $n!$.

Dimostrazione. Ogni permutazione è una k-pla di elementi distinti di S del tipo

$$(f(1), f(2), \dots, f(k)),$$

quindi si deve provare che ci sono $n(n-1)(n-2)\dots(n-k+1)$ differenti k-ple di elementi distinti scelti nell'insieme S. Ci sono n modi di scegliere il primo elemento in tali k-ple. Scelto il primo elemento $f(1)$, ci sono $n-1$ modi di scegliere il secondo elemento in tali k-ple. Scelti i primi $t-1$ elementi di una k-pla ci sono $n-(t-1) = n-t+1$ restanti elementi tra cui scegliere il t-esimo elemento della k-pla.

Dunque, per il principio generalizzato del prodotto, il numero totale di k-ple è

$$\prod_{i=1}^{k}(n-i+1) = n(n-1)(n-2)\dots(n-k+1) = \frac{n!}{(n-k)!}.$$

Nel caso particolare $n = k$, poiché $(n-n)! = 0! = 1$, il numero totale di n-ple è $n!$ □

Il numero $\dfrac{n!}{(n-k)!}$ è spesso denotato con $(n)_k$ o anche con $P(n,k)$ ed è anche chiamato il **numero delle permutazioni di n elementi presi k alla volta**.

Nel calcolo di $\dfrac{n!}{(n-k)!}$ non è necessario calcolare $n!$ e $(n-k)!$ in quanto si può semplificare numeratore e denominatore della frazione per $(n-k)!$.

Esempio 1.15. In quanti modi in un insieme di 10 persone si possono scegliere 3 persone per farne un presidente, un vicepresidente ed un segretario, in modo che una stessa persona non ricopra più ruoli?

Soluzione. Si tratta di determinare il numero delle permutazioni di 10 elementi presi 3 alla volta. Questo numero è

$$(10)_3 = \frac{10!}{(10-3)!} = \frac{10!}{7!} = \frac{10\cdot 9\cdot 8\cdot 7\cdot\dots\cdot 2\cdot 1}{7\cdot\dots\cdot 2\cdot 1} = 10\cdot 9\cdot 8 = 720.$$

Teorema 1.16. *Il numero delle funzioni iniettive da un insieme X di cardinalità k ad un insieme S di cardinalità n è* $\dfrac{n!}{(n-k)!}$

Dimostrazione. Sia $X = \{a_1, a_2, \ldots, a_k\}$. Ad una funzione iniettiva $f : X \to S$ si associa la k-pla delle immagini distinte $(f(a_1), f(a_2), \ldots, f(a_k))$, cioè una permutazione di elementi di S e usando la formula del Teorema 1.14 abbiamo che queste sono in numero di $\dfrac{n!}{(n-k)!}$ □

Utilizzando il principio generalizzato del prodotto, è possibile calcolare il numero di tutte le funzioni e non solo di quelle iniettive (o permutazioni).

Teorema 1.17. *Il numero delle funzioni da un insieme X di cardinalità k ad uno S di cardinalità n è n^k.*

Dimostrazione. Sia $X = \{a_1, a_2, \ldots, a_k\}$. Ogni funzione $f : X \to S$ può essere descritta da una k-pla $(f(a_1), f(a_2), \ldots, f(a_k))$ di elementi di X, non necessariamente distinti. Ci sono n scelte per ciascun $f(a_i)$, quindi per il principio generalizzato del prodotto ci sono n^k funzioni. □

Qual è la differenza tra contare permutazioni (o funzioni iniettive) e contare funzioni? L'esempio seguente può aiutare a capire.

Esempio 1.18. Si hanno a disposizione 5 sedie e 7 colori. In quanti modi possiamo colorare le sedie se tutte le sedie devono avere colori diversi? In quanti modi possiamo colorare le sedie se lo stesso colore può essere usato per colorare più sedie (eventualmente tutte)?

Soluzione. Ogni colorazione assegna un colore ad una sedia, quindi si devono considerare le funzioni tra l'insieme delle sedie e quello dei colori. Se sedie diverse devono avere colori diversi dobbiamo considerare funzioni iniettive e quindi ci sono

$$P(7,5) = (7)_5 = \frac{7!}{(7-5)!} = 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 = 2520$$

colorazioni delle 5 sedie con 7 colori diversi. Se sedie diverse possono avere lo stesso colore, dobbiamo considerare tutte le funzioni e quindi ci sono $7^5 = 16807$ colorazioni di questo tipo.

2. Applicazioni dei principi base di enumerazione

Sottoinsiemi. Molti problemi possono essere enunciati in termini di sottoinsiemi di un insieme dato.

Esempio 2.1. Un insegnante chiede ad una classe di 12 studenti dei volontari per un progetto. Quanti diversi insiemi di volontari si possono formare?

Soluzione. Per descrivere i diversi insiemi di volontari indichiamo per ciascun studente se è o meno volontario nell'insieme scelto. Questo dà una funzione dall'insieme degli studenti all'insieme {volontario, non volontario}. Ciascuna diversa funzione rappresenta un insieme diverso di volontari. Quindi il numero degli insiemi di volontari è il numero delle funzioni che è $2^{12} = 4096$. Notiamo che tra i possibili insiemi di volontari c'è anche l'insieme vuoto (l'insieme che non ha volontari).

Il ragionamento appena fatto si può generalizzare.

Teorema 2.2. *Un insieme con n elementi ha 2^n sottoinsiemi.*

Dimostrazione. Per determinare un sottoinsieme di un insieme S di cardinalità n indichiamo per ciascun elemento di S se appartiene o no al sottoinsieme scelto utilizzando una funzione f da S all'insieme $\{0, 1\}$, dove $f(s) = 0$ se s non appartiene al sottoinsieme e $f(s) = 1$ se s appartiene al sottoinsieme. Questa funzione f è detta la **funzione caratteristica** del sottoinsieme. Quindi il numero dei sottoinsiemi è anche il numero delle funzioni da S in $\{0, 1\}$ che è 2^n. □

Esempio 2.3. Se $S = \{a, b, c\}$ si hanno i seguenti $8 = 2^3$ sottoinsiemi di S.

$$\emptyset, \quad \{a\}, \quad \{b\}, \quad \{c\}, \quad \{a, b\}, \quad \{a, c\}, \quad \{b, c\}, \quad \{a, b, c\}.$$

Spesso, quando trattiamo dei sottoinsiemi di S, dobbiamo considerare sottoinsiemi di cardinalità k fissata, detti anche **k-sottoinsiemi di S.** Per esempio, il presidente di un club deve nominare una commissione di 3 membri che si occupi di un certo problema. Il numero delle possibili commissioni è il numero dei 3-sottoinsiemi dell'insieme dei membri del club.

Il numero dei k-sottoinsiemi di un insieme S di cardinalità n è denotato con $\binom{n}{k}$ o $C(n, k)$. Per calcolare $\binom{n}{k}$ si utilizza il seguente teorema.

Teorema 2.4. *Il numero dei k-sottoinsiemi di un insieme S di cardinalità n è*

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

Dimostrazione. Cerchiamo una relazione tra il numero $\binom{n}{k}$ ed il numero $(n)_k$ delle permutazioni su k elementi di S. Ogni permutazione su k elementi di S è una k-pla di elementi distinti di S della forma

$$(f(1), f(2), \ldots, f(k)).$$

Questa k-pla individua un k-sottoinsieme

$$S' = \{f(1), f(2), \ldots f(k)\}$$

di S e il numero delle possibili permutazioni che individuano S' è $k!$ Quindi ad ogni sottoinsieme S' di S si associano $k!$ permutazioni, e il numero delle permutazioni su k elementi di S può essere anche contato come il prodotto del numero dei sottoinsiemi di S per $k!$. Si ha quindi:

$$\binom{n}{k} \cdot k! = (n)_k$$

Dividendo per $k!$ si ha:

$$\binom{n}{k} = \frac{n_k}{k!} = \frac{n!}{(n-k)! \cdot k!}$$

□

Esempio 2.5. Sia S un insieme con 16 elementi. Qual è il numero dei 2-sottoinsiemi di S? Qual è il numero dei 14-sottoinsiemi di S?

Soluzione. Il numero dei 2-sottoinsiemi è

$$\binom{16}{2} = \frac{16!}{2! \cdot (16-2)!} = \frac{16 \cdot 15 \cdot 14!}{2 \cdot 14!} = 120.$$

Il numero dei 14-sottoinsiemi è

$$\binom{16}{14} = \frac{16!}{14! \cdot (16-14)!} = \frac{16 \cdot 15 \cdot 14!}{14! \cdot 2} = 120.$$

Notiamo che si ottiene la stessa risposta in entrambi i casi. Questo poteva essere previsto notando che il numero di modi di scegliere 2 elementi nel sottoinsieme e 14 fuori dal sottoinsieme è uguale al numero di modi di scegliere 14 elementi nel sottoinsieme e 2 fuori.

Il numero $\binom{n}{k}$ è detto **coefficiente binomiale** per un motivo che vedremo tra poco. Si usa anche chiamare un sottoinsieme con k elementi una **combinazione di n elementi presi k alla volta.**

Quindi, per riassumere:

- $(n)_k$ è il numero delle *permutazioni* di n elementi presi k alla volta;
- $\binom{n}{k}$ è il numero delle *combinazioni* di n elementi presi k alla volta.

Nel primo caso si tratta di *funzioni iniettive*, nel secondo di *sottoinsiemi*.

Ritornando agli esempi sulle scelte di persone in un insieme dato, possiamo fornire l'esempio seguente che chiarisce la differenza tra permutazioni su k elementi, sottoinsiemi di k elementi e k-ple.

Esempio 2.6. 10 professori devono scegliere un presidente, un vicepresidente ed un segretario. Si deve anche scegliere una commissione di 3 membri. In quanti modi si possono fare queste scelte? Inoltre i 10 professori fanno da tutori ad un gruppo di allievi. Se 3 nuovi allievi si aggiungono al gruppo, in quanti modi diversi i 3 allievi possono scegliere i loro tutori?

Soluzione. Scegliere presidente, vice e segretario significa dare una terna o lista di 3 differenti persone nell'ordine delle cariche che devono ricoprire. Ogni lista diversa, anche con le stesse persone in ordine diverso, corrisponde ad una diversa scelta. Il numero di modi di scegliere la lista corrisponde al numero di permutazioni su 3 elementi di un insieme con 10 elementi, precisamente

$$(10)_3 = \frac{10!}{(10-3)!} = 10 \cdot 9 \cdot 8 = 720.$$

La scelta di 3 membri della commissione dà un sottoinsieme di 3 elementi, dove l'ordine in cui gli elementi compaiono è irrilevante. Scegliere le stesse persone in ordine diverso dà lo stesso sottoinsieme. Quindi il numero delle commissioni possibili è il numero dei sottoinsiemi (combinazioni) e cioè

$$\binom{10}{3} = \frac{10!}{3! \cdot (10-3)!} = \frac{10 \cdot 9 \cdot 8}{3 \cdot 2} = 120.$$

Per la scelta del tutore, ciascun allievo deve scegliere nell'insieme dei 10 professori (si noti che 2 o più ragazzi possono scegliere lo stesso tutore), quindi le 3 scelte danno una funzione dall'insieme dei 3 allievi nell'insieme dei 10 professori. Ci sono quindi $10^3 = 1000$ tali funzioni.

Un altro modo di vedere la differenza tra i diversi concetti è quello di scrivere le permutazioni su k elementi (o k-ple di elementi distinti), i sottoinsiemi con k elementi e le k-ple (di elementi anche coincidenti) scelti da un insieme.

Esempio 2.7. Si scrivano le permutazioni su 2 elementi, i sottoinsiemi con 2 elementi, e le 2-ple scelte dall'insieme $\{a, b, c\}$.

Soluzione.

permutazioni: ab, ac, ba, bc, ca, cb

sottoinsiemi: $\{a, b\}$, $\{a, c\}$, $\{b, c\}$

2-ple: (a, a), (a, b), (a, c), (b, a), (b, b), (b, c), (c, a), (c, b), (c, c).

Multi-insiemi. Alcuni problemi che sembrano richiedere l'uso di un insieme in verità utilizzano qualcosa di diverso che sarà chiamato un multi-insieme. Per esempio le lettere dell'alfabeto contenute nella parola *tetto* sono, in ordine alfabetico, $\{e, o, t, t, t\}$. La notazione con le parentesi graffe

ricorda quella di un insieme, ma un insieme non contiene elementi ripetuti. Siamo quindi in presenza di qualcosa di diverso.

Definizione 2.8. Un **multi-insieme** scelto da un insieme S è una funzione $m : S \to \mathbb{N}$ da S all'insieme degli interi non negativi. Per ogni $x \in S$, $m(x)$ è detta la **molteplicità** di x nel multi-insieme. La **cardinalità del multi-insieme** è la somma delle molteplicità degli elementi di S.

Esempio 2.9. Qual è la molteplicità di ciascuna lettera dell'alfabeto nella parola *tetto*? Qual è la cardinalità del multi-insieme delle lettere di *tetto*?

Soluzione. La molteplicità di e ed o è 1, di t è 3, di ogni altra lettera dell'alfabeto è 0. In simboli la funzione di molteplicità è data da $m(e) = m(o) = 1$, $m(t) = 3$, $m(a) = m(b) = m(c) = \cdots = 0$. La cardinalità del multi-insieme è $1 + 1 + 3 = 5$.

Notiamo che la funzione caratteristica di un sottoinsieme T di S è una funzione che assegna 0 o 1 a ciascun elemento di S e precisamente assegna 1 a ciascun elemento di T e 0 a tutti gli altri. Quindi si può pensare ad un sottoinsieme T come ad un multi-insieme scelto da S in cui le molteplicità sono date dalla funzione caratteristica. Si può allora intuire perché i coefficienti binomiali, che avevamo utilizzato per contare i k-sottoinsiemi di S, saranno anche utili per contare il numero dei multi-insiemi di k elementi scelti in S.

Teorema 2.10. *Il numero dei multi-insiemi di cardinalità k scelti da un insieme di n elementi è dato da*

$$\binom{n+k-1}{k}$$

Dimostrazione. Indichiamo con $x_1, x_2, \ldots, x_n$ gli elementi di S. Per ogni multi-insieme su S, cioè per ciascuna funzione m definita su S, possiamo definire una successione di $n + k - 1$ numeri 1 e 0 come segue:

si scrivono $m(x_1)$ numeri 1 seguiti da uno 0,

di seguito si scrivono $m(x_2)$ numeri 1 seguiti da uno 0,

... ...

si continua così fino ad avere scritto $m(x_{n-1})$ numeri 1 seguiti da uno 0,

alla fine si scrivono $m(x_n)$ numeri 1 e non si aggiunge lo 0 finale.

Ora $m(x_1) + m(x_2) + \cdots + m(x_n) = k$ è la cardinalità del multi-insieme e quindi una successione come quella precedente contiene k numeri 1 e $n-1$ numeri 0, quindi in totale $n + k - 1$ elementi.

Inoltre, data una successione come sopra, si individua un multi-insieme scelto da S utilizzando gli 0 per suddividere la successione in n gruppi di numeri 1, dove per ogni gruppo il numero di 1 che compaiono rappresenta la molteplicità dell'elemento corrispondente.

Quindi le successioni del tipo precedente sono tante quanti i multi-insiemi scelti da S. Poiché il numero di tali successioni è il numero di modi di scegliere i k elementi in cui posizionare i numeri 1, questo numero è anche dato da $\binom{n+k-1}{k}$ cioè dal numero dei sottoinsiemi con k elementi di un insieme con $n+k-1$ elementi e quindi questo è anche il numero dei multi-insiemi di cardinalità k scelti in S. □

Esempio 2.11. Dato $S = \{a, b, c\}$, e quindi $n = 3$, determiniamo tutti i multi-insiemi di cardinalità $k = 4$. Ve ne sono $\binom{n+k-1}{k} = \binom{6}{4} = 15$. Ogni multi-insieme determina una successione di $n + k - 1 = 6$ numeri 1 e 0 contenenti $k = 4$ numeri 1 ed $n - 1 = 2$ numeri 0. Le successioni sono:

111100, 111010, 110110, 101110, 011110, 111001, 110101, 101101,
011101, 110011, 101011, 011011, 100111, 010111, 001111.

Ad esempio alla successione 111100 si associa il multi-insieme $\{a, a, a, a\}$, a 101011 il multi-insieme $\{a, b, c, c\}$, e così via.

Esempio 2.12. Una pasticceria produce 5 tipi, a, b, c, d, e di paste ricoperte al cioccolato. In quanti modi diversi si può confezionare un vassoio con 8 di queste paste?

Soluzione. Ogni confezione di 8 paste può essere pensata come un multi-insieme di cardinalità 8 scelto da un insieme di cardinalità 5. Quindi ci sono

$$\binom{5+8-1}{8} = \binom{12}{8} = \frac{12!}{4! \cdot 8!} = \frac{12 \cdot 11 \cdot 10 \cdot 9}{4 \cdot 3 \cdot 2} = 11 \cdot 5 \cdot 9 = 495$$

confezioni diverse.

Se ad ogni confezione associamo la successione di 12 numeri 1 e 0 con 8 numeri 1, la confezione 001110110111 contiene 0 paste di tipo a, 0 paste di tipo b, 3 paste di tipo c, 2 paste di tipo d e 3 paste di tipo e.

Come risolvere problemi di enumerazione. Vediamo ora un esempio che ci aiuta a capire come scomporre e risolvere problemi di enumerazione.

Esempio 2.13. 10 professori devono scegliere 3 di loro per formare un direttivo composta da un presidente, un vice-presidente e un segretario. Inoltre devono scegliere una commissione di 3 membri. Se si vuole che i

membri del direttivo non facciano parte della commissione, in quanti modi diversi possono essere fatte entrambe le scelte?

Soluzione. Una volta scelto il direttivo, rimangono 7 professori tra i quali scegliere i 3 membri della commissione. Ci sono quindi $(10)_3$ modi di scegliere il direttivo e $\binom{7}{3}$ modi di scegliere la commissione. Quindi per il principio del prodotto si hanno

$$(10)_3 \cdot \binom{7}{3} = 720 \cdot 35 = 25200$$

modi diversi di fare entrambe le scelte. Si noti che si ha anche l'identità seguente

$$25200 = 120 \cdot 210 = \binom{10}{3} \cdot (7)_3.$$

Come si può ragionare per arrivare a quest'ultima identità?

L'esempio suggerisce la seguente strategia nella risoluzione dei problemi:

> *Un problema richiede l'uso del principio del prodotto se c'è una successione di decisioni da prendere (o opzioni da fare) e il numero dei modi di prendere una decisione non è influenzato dalle altre decisioni da prendere.*

Nell'esempio precedente, nonostante l'insieme delle persone disponibili per la commissione dipenda dalle persone scelte nel direttivo, il numero delle persone disponibili per la scelta della commissione non dipende da quali persone sono state scelte nel direttivo ed è sempre 7.

Quando la situazione per la successione delle decisioni da prendere è meno strutturata che nell'esempio precedente, è in genere necessario suddividere il problema in casi, trovare il numero dei modi di trattare i diversi casi e collegare i risultati insieme. Questo significa che stiamo applicando il principio della somma. Può essere utile, per visualizzare una tale successione di decisioni, utilizzare un diagramma chiamato **diagramma ad albero** delle decisioni.

Esempio 2.14. Lanciamo una moneta. Se viene testa, lanciamo un dado. Se viene croce lanciamo di nuovo la moneta e ci fermiamo. Quante possibilità diverse si hanno?

Soluzione. Ci sono 8 possibilità che sono la somma delle 6 possibilità quando viene testa e delle 2 quando viene croce. Provare a visualizzare la soluzione con un diagramma ad albero.

L'esempio precedente è particolarmente semplice e forse non richiederebbe l'uso del diagramma, ma in situazioni più complesse il diagramma ad albero può essere molto utile.

Esempio 2.15. Se abbiamo 4 monete da 10 centesimi, 3 monete da 20 centesimi e 2 monete da 50 centesimi, in quanti modi possiamo scegliere monete in modo da formare 1 euro?

Soluzione. La scelta dei 10 e 20 centesimi dipende dalla scelta dei 50 centesimi. Dividiamo il problema in tre casi:

a) nessuna moneta da 50 centesimi;

b) una moneta da 50 centesimi;

c) due monete da 50 centesimi.

Nel caso a), non abbiamo scelte e dobbiamo usare tutte le altre monete a disposizione.

Nel caso b) abbiamo due sottocasi, a seconda che utilizziamo 1 o 2 monete da 20 centesimi.

Nel caso c) non abbiamo da aggiungere altro.

In tutto quindi si hanno 4 modi di formare 1 euro con le monete a disposizione. Provare a visualizzare la soluzione con un diagramma ad albero.

3. Il Teorema del binomio e il triangolo di Pascal

Quando abbiamo introdotto i coefficienti binomiali $\binom{n}{k}$ non siamo stati in grado di giustificarne il nome. Si capisce il significato del nome se si considerano le potenze di particolari polinomi, detti *binomi*, che sono la somma di due termini, uno almeno dei quali contenente una variabile (*monomio*). Ad esempio sono binomi $x+y$, $2x+1$, $3-4x$, e $1+x^2$. Utilizzando la moltiplicazione tra polinomi, si possono calcolare le potenze di un binomio, ad esempio di $x+y$, e si ottiene:

$$(x+y)^0 = 1$$
$$(x+y)^1 = x+y$$
$$(x+y)^2 = x^2+2xy+y^2$$
$$(x+y)^3 = x^3+3x^2y+3xy^2+y^3$$
$$(x+y)^4 = x^4+4x^3y+6x^2y^2+4xy^3+y^4$$
$$(x+y)^5 = x^5+5x^4y+10x^3y^2+10x^2y^3+5xy^4+y^5$$

Se consideriamo, ad esempio, i coefficienti $1, 4, 6, 4, 1$ dei monomi che danno lo sviluppo di $(x+y)^4$, vediamo che questi numeri sono i coefficienti binomiali

$\binom{4}{k}$ per $k = 0, 1, 2, 3, 4$. Questa considerazione è vera in generale, vale infatti il seguente teorema.

Teorema 3.1. *Per ogni intero non negativo n, si ha*

$$\begin{aligned}(x+y)^n &= \sum_{k=0}^{n} \binom{n}{k} x^{n-k}y^k \\ &= x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{n-1}xy^{n-1} + y^n\end{aligned}$$

Dimostrazione. Quando operiamo per calcolare il prodotto

$$\underbrace{(x+y)(x+y)\cdots(x+y)}_{n \text{ fattori}} = (x+y)^n$$

iniziamo scegliendo x da tutti i fattori e, moltiplicandoli tra loro, otteniamo x^n. Poi scegliamo x da $n-1$ fattori e y dal fattore restante e, moltiplicandoli, otteniamo $x^{n-1}y$ che sommiamo con tutti gli altri addendi dello stesso tipo.

Quanti sono gli addendi del tipo $x^{n-1}y$? Ci sono $\binom{n}{1}$ modi di scegliere il fattore che contiene y. Quindi ci sono $\binom{n}{1}$ addendi del tipo $x^{n-1}y$.

Gli addendi del tipo $x^{n-2}y^2$ sono ottenuti scegliendo x in $n-2$ fattori e y nei due fattori restanti. In totale ci sono quindi $\binom{n}{2}$ addendi della forma $x^{n-2}y^2$ e in generale abbiamo $\binom{n}{k}$ modi di scegliere x in $n-k$ fattori e y nei restanti k fattori.

Quindi la somma degli addendi del tipo $x^{n-k}y^k$ è $\binom{n}{k}x^{n-k}y^k$. Facendo allora la somma in totale si ottiene:

$$\sum_{k=0}^{n} \binom{n}{k} x^{n-k}y^k = x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{n-1}xy^{n-1} + y^n.$$

□

Esempio 3.2. Scrivere lo sviluppo di $(x+y)^6$.

Soluzione. Si ha

$$\begin{aligned}(x+y)^6 &= \sum_{k=0}^{6} \binom{6}{k} x^{6-k}y^k \\ &= x^6 + \binom{6}{1}x^5y + \binom{6}{2}x^4y^2 + \cdots + \binom{6}{5}xy^5 + y^6 \\ &= x^6 + 6x^5y + 15x^4y^2 + 20x^3y^3 + 15x^2y^4 + 6xy^5 + y^6.\end{aligned}$$

Nell'esempio si può notare una simmetria dei coefficienti nel senso che si hanno gli stessi coefficienti da sinistra verso destra che da destra verso sinistra. Avevamo già notato questa simmetria provando che $\binom{16}{2} = \binom{16}{14}$. La regola generale che prova questa simmetria è data dal seguente teorema.

Teorema 3.3.

$$\binom{n}{k} = \binom{n}{n-k}, \quad \forall k = 0, 1, \cdots, n.$$

Dimostrazione. Per determinare un k-sottoinsieme di un insieme con n elementi si possono selezionare i k elementi **nel** sottoinsieme o, equivalentemente, gli $n-k$ elementi **fuori** dal sottoinsieme. Quindi il numero di modi di scegliere k elementi in n perché stiano in un sottoinsieme è uguale al numero di modi di sceglierne $n-k$ perché stiano fuori dal sottoinsieme. □

Esempio 3.4. Scrivere lo sviluppo di

$$(2x+3)^4$$

Soluzione. Si ha:

$$\begin{aligned}(2x+3)^4 &= \sum_{k=0}^{4} \binom{4}{k} (2x)^{4-k}3^k \\ &= \binom{4}{0}(2x)^4 + \binom{4}{1}(2x)^3 3^1 + \binom{4}{2}(2x)^2 3^2 + \binom{4}{3}(2x)3^3 + 3^4 \\ &= 1 \cdot 16x^4 + 4 \cdot 8x^3 \cdot 3 + 6 \cdot 4x^2 \cdot 9 + 4 \cdot 2x \cdot 27 + 1 \cdot 1 \cdot 81 \\ &= 16x^4 + 96x^3 + 216x^2 + 216x + 81.\end{aligned}$$

Esempio 3.5. Utilizzando il teorema del binomio provare che

$$\binom{6}{0} + \binom{6}{1} + \binom{6}{2} + \binom{6}{3} + \binom{6}{4} + \binom{6}{5} + \binom{6}{6} = 2^6$$

Soluzione. Utizzando il teorema si ha il seguente sviluppo:

$$(x+y)^6 = \binom{6}{0}x^6 + \binom{6}{1}x^5y + \binom{6}{2}x^4y^2 + \cdots + \binom{6}{5}xy^5 + \binom{6}{6}y^6$$

che, sostituendo $x = 1$ e $y = 1$, dà:

$$(1+1)^6 = \binom{6}{0}1^6 + \binom{6}{1}1^5 \cdot 1 + \binom{6}{2}1^4 \cdot 1^2 + \cdots + \binom{6}{5}1 \cdot 1^5 + \binom{6}{5}1^6$$

cioè

$$2^6 = \binom{6}{0} + \binom{6}{1} + \binom{6}{2} + \binom{6}{3} + \binom{6}{4} + \binom{6}{5} + \binom{6}{6}. \quad \square$$

Esempio 3.6. Quale è il coefficiente di x^4 in $(2-3x)^7$?
Soluzione. Il termine che contiene $4x^4$ è

$$\binom{7}{4}(2)^{7-4}(-3x)^4 = \frac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1} \cdot 8 \cdot 81 \cdot x^4.$$

Quindi il coefficiente di x^4 è: $7 \cdot 5 \cdot 8 \cdot 81 = 22680$.

Il teorema del binomio consente di sviluppare, ad esempio, $(x+y)^{12}$ senza dover moltiplicare i 12 fattori. C'è però da notare che il calcolo dei coefficienti binomiali $\binom{12}{k}$, con $k = 0, 1, \cdots, 12$ può richiedere molto tempo, anche se la simmetria dei coefficienti riduce di metà il lavoro.

Vediamo come il problema del calcolo dei coefficienti binomiali si può ulteriormente semplificare. Se dallo sviluppo di $(x+y)^n$ con $n = 0, 1, \cdots, 6$ evidenziamo solo i coefficienti, otteniamo il seguente triangolo, detto il **triangolo di Pascal**:

$$\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1 \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}$$

Dalla tabella è evidente la simmetria dei coefficienti e gli 1 lungo il bordo dipendono dal fatto che $\binom{n}{0}$ e $\binom{n}{n}$ sono sempre 1 (un insieme con n elementi ha esattamente un insieme vuoto e un n-sottinsieme). Un esame più attento della tabella mostra la seguente utile proprietà:

> *Ogni elemento nel triangolo di Pascal è la somma degli elementi alla sua sinistra e alla sua destra nella riga superiore.*

Chiamiamo questa proprietà la **proprietà di Pascal**. Nella tabella precedente si vede che, ad esempio, $\binom{6}{4} = 15$ è la somma di $\binom{5}{3} + \binom{5}{4} = 10+5$. Vale cioè in generale il seguente risultato.

Teorema 3.7. *Per tutti gli interi positivi n e per tutti gli interi k compresi tra 1 ed $n-1$ si ha:*

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

Dimostrazione. Contiamo i k-sottoinsiemi di $X = \{x_1, \ldots, x_n\}$ suddividendoli in

(1) k-sottoinsiemi che contengono x_n

(2) k-sottoinsiemi che non contengono x_n

Nel caso (1), ogni sottoinsieme è ottenuto aggiungendo x_n ad un $(k-1)$-sottoinsieme di $X' = \{x_1, \ldots, x_{n-1}\}$. Quindi il loro numero è il numero dei $(k-1)$-sottoinsiemi di X', e cioè $\binom{n-1}{k-1}$. Nel caso (2), ogni sottoinsieme è un k-sottoinsieme di X'. Quindi il loro numero è $\binom{n-1}{k}$.

Allora i k-sottoinsiemi di X sono in totale $\binom{n-1}{k-1} + \binom{n-1}{k}$. □

Osserviamo che il Teorema può anche essere dimostrato a partire dalla formula che definisce i coefficienti binomiali eseguendo le semplificazioni algebriche.

4. Il principio di inclusione-esclusione

Molti problemi di enumerazione trattano famiglie di sottoinsiemi di un insieme dato e richiedono di calcolare quanti elementi stanno nell'unione di questi sottoinsiemi oppure quanti elementi non stanno nell'unione (cioè non stanno in nessun sottoinsieme).

Per risolvere questo tipo di problemi bisogna sapere come sono fatte le intersezioni tra i diversi sottoinsiemi ed utilizzare un metodo detto il *principio di inclusione-esclusione*.

Esempio 4.1. Da un'indagine sugli esami sostenuti in Dicembre risulta che su 25 intervistati, 15 studenti hanno superato Logica, 12 Programmazione I, e 5 entrambi gli esami. Quanti hanno dato almeno uno dei due esami? Quanti non ne hanno dato nessuno dei due?

Soluzione. Sono note le cardinalità di U, insieme degli studenti intervistati, L, insieme degli studenti che hanno dato Logica, P, insieme degli studenti che hanno dato Programmazione I, e $L \cap P$, insieme degli studenti che hanno dato entrambi gli esami. Si ha infatti $|U| = 25$, $|L| = 15$, $|P| = 12$, $|L \cap P| = 5$.

Risolvere il problema significa determinare la cardinalità di $L \cup P$ e del complementare $\overline{L \cup P} = U - (L \cup P)$. Per calcolare $|L \cup P|$ possiamo pensare all'insieme $L \cup P$ come unione disgiunta dei due insiemi $L - P$ e P e, usando il principio della somma, si ha: $|L \cup P| = |L - P| + |P|$. Inoltre, sempre per il principio della somma si ha $|L| = |L - P| + |L \cap P|$ da cui $|L - P| = |L| - |L \cap P|$.

Otteniamo quindi:

$$|L \cup P| = |L| + |P| - |L \cap P| = 15 + 12 - 5 = 22,$$
$$|\overline{L \cup P}| = |U| - |L \cup P| = 25 - 22 = 3.$$

Il ragionamento appena usato può essere applicato a due qualunque insiemi, ottenendo quella che si chiama la *formula di inclusione-esclusione*, che generalizza il principio della somma al caso in cui gli insiemi non sono necessariamente disgiunti.

Teorema 4.2. *Se A e B sono due insiemi qualsiasi, si ha*

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Dimostrazione. Per contare gli elementi di $A \cup B$, notiamo che calcolando $|A|$, contiamo tutti gli elementi in A e, calcolando $|B|$, contiamo tutti gli elementi in B, e quindi gli elementi in $A \cap B$ vengono contati due volte. Dunque occorre sottrarre $|A \cap B|$ dalla somma $|A| + |B|$. □

La formula precedente si può estendere a 3 insiemi ottenendo:

Teorema 4.3. *Se A, B e C sono 3 insiemi si ha:*

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

Dimostrazione. Notiamo che se un elemento sta in A e non in B né C dà un contributo di 1 all'addendo $|A|$ e 0 ad ogni altro addendo della somma $|A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$. Analoga considerazione si può fare per gli elementi che stanno solo in B o solo in C.

Se un elemento sta in $A \cap B$ ma non in C, dà un contributo di 1 agli addendi $|A|$ e $|B|$, dà un contributo di -1 all'addendo $-|A \cap B|$ e contributo nullo ad ogni altro addendo. Quindi in totale viene conteggiato $1+1-1 = 1$ volta. Analoga considerazione si può fare per gli elementi che stanno in $B \cap C$ ma non in A o in $A \cap C$ ma non in B.

Se un elemento sta in $A \cap B \cap C$ dà un contributo di 1 agli addendi $|A|$, $|B|$, $|C|$, e $|A \cap B \cap C|$, dà un contributo di -1 agli addendi $-|A \cap B|$, $-|A \cap C|$ e $-|B \cap C|$, quindi in totale viene conteggiato $1+1+1-1-1-1+1=1$ volta.

Quindi, comunque scegliamo l'elemento in $A \cup B \cup C$ la somma precedente conta esattamente una volta ogni elemento, dovunque sia posizionato e dà quindi la cardinalità di $A \cup B \cup C$. □

Esempio 4.4. Sia $X = \{1, 2, \ldots, 100\}$. Quanti sono gli elementi di X che sono divisibili per 2, o per 5, o per 12?

Soluzione. Siano A l'insieme degli $x \in X$ che sono divisibili per 2, B l'insieme degli $x \in X$ che sono divisibili per 5, e C l'insieme degli $x \in X$ che sono divisibili per 12. Si ha $|A| = 50$, $|B| = 20$, e $|C| = 8$. Inoltre $A \cap B$ è l'insieme degli interi che sono divisibili per 2 e per 5 cioè per 10 e si ha $|A \cap B| = 10$. Analogamente si ottiene $|A \cap C| = 8$, $|B \cap C| = 1$ e $|A \cap B \cap C| = 1$. Segue $|A \cup B \cup C| = 50 + 20 + 8 - 10 - 1 - 8 + 1 = 60$.

Le formule precedenti sono casi particolari di una formula che vale per n insiemi. Siano S_1, S_2, …, S_n insiemi qualsiasi e poniamo $N = \{1, 2, \ldots, n\}$. Introduciamo le seguenti notazioni:

$$\sum_{i \in N} |S_i| = |S_1| + |S_2| + \cdots + |S_n|,$$

$$\sum_{i<j \in N} |S_i \cap S_j| = |S_1 \cap S_2| + |S_1 \cap S_3| + \cdots + |S_2 \cap S_3| + \ldots \\ \cdots + |S_i \cap S_j| + \cdots + |S_{n-1} \cap S_n|,$$

$$\sum_{i<j<k \in N} |S_i \cap S_j \cap S_k| = |S_1 \cap S_2 \cap S_3| + |S_1 \cap S_2 \cap S_4| + \ldots \\ \cdots + |S_i \cap S_j \cap S_k| + \cdots + |S_{n-2} \cap S_{n-1} \cap S_n|,$$

e così via.

La formula generale, di cui tralasciamo la dimostrazione, è la seguente (detta il **principio di inclusione-esclusione**).

Teorema 4.5. *Se S_1, S_2, …, S_n sono insiemi qualsiasi, allora*

$$|S_1 \cup S_2 \cdots \cup S_n| = \sum_{i \in N} |S_i| \; - \sum_{i<j \in N} |S_i \cap S_j| \; + \sum_{i<j<k \in N} |S_i \cap S_j \cap S_k| \; - \; \ldots$$

Esempio 4.6. Da un'indagine sul consumo di sostanze eccitanti, quali sigarette, alcool, caffè e tè su un gruppo di persone risulta la seguente tabella

S	A	C	T	SA	SC	ST	AC
10	12	13	17	6	8	7	7

AT	CT	SAC	SAT	SCT	ACT	$SACT$
8	7	5	4	5	4	3

dove S, A, C e T indicano le diverse sostanze, ed i numeri incolonnati corrispondenti rappresentano il numero di persone che fanno uso di dette sostanze.

Se ogni persona intervistata fa uso di almeno una tra le sostanze indicate, qual è il numero delle persone coinvolte nell'indagine?

Soluzione. Per le condizioni del problema ogni persona appartiene ad $S \cup A \cup C \cup T$. Quindi il numero delle persone coinvolte è dato dalla cardinalità dell'unione, che, per il principio di inclusione-esclusione, è:

$$|S \cup A \cup C \cup T| = 10+12+13+17-6-8-7-7-8-7+5+4+5+4-3 = 24.$$

Esempio 4.7. Un professore raccoglie i compiti in classe dati ai suoi n studenti, li mescola e li ridistribuisce agli studenti affinché ognuno corregga il compito di un compagno. Naturalmente non vuole che uno studente corregga il proprio compito. Quanti sono il numero di modi di distribuire i compiti in modo che nessun studente riceva indietro il proprio compito?

Soluzione. Indichiamo con $N = \{1, 2, \ldots, n\}$ l'insieme degli studenti. Ogni modo di distribuire i compiti può essere interpretato come una permutazione $(f(1), f(2), \ldots, f(n))$ sull'insieme N. Dire che i ha come immagine $f(i)$ significa dire che il compito dello studente i è corretto dallo studente $f(i)$.

Sia S_i l'insieme delle permutazioni che tengono fisso i, cioè l'insieme dei modi di distribuire i compiti in cui lo studente i riceve indietro il suo compito. Risolvere il problema significa determinare quante sono le permutazioni che non appartengono a nessun S_i. Questo numero si può calcolare sottraendo da $n!$, che è il numero di tutte le permutazioni possibili, la cardinalità dell'insieme $|S_1 \cup S_2 \cup \cdots \cup S_n|$.

Dobbiamo per prima cosa calcolare le cardinalità $|S_i|$, $|S_i \cap S_j|$, ...

Si ha :

$$\begin{aligned} |S_i| &= (n-1)! \\ |S_i \cap S_j| &= (n-2)! \quad \text{per } i \neq j \\ |S_i \cap S_j \cap S_k| &= (n-3)! \quad \text{per } i \neq j,\ i \neq k,\ j \neq k \\ &\cdots\cdots \end{aligned}$$

Poiché ci sono $\binom{n}{1}$ modi di scegliere i, ci sono $\binom{n}{1}$ addendi della forma $|S_i|$. Analogamente, poiché ci sono $\binom{n}{2}$ modi di scegliere i e j, ci sono

$\binom{n}{2}$ addendi della forma $|S_i \cap S_j|$, e così via. Per il principio di inclusione-esclusione si ha quindi:

$$\begin{aligned}|S_1 \cup S_2 \cdots \cup S_n| &= \sum_{i\in N} |S_i| \; - \sum_{i<j\in N} |S_i \cap S_j| \; + \sum_{i<j<k\in N} |S_i \cap S_j \cap S_k| - \dots \\ &= \binom{n}{1}(n-1)! - \binom{n}{2}(n-2)! + \binom{n}{3}(n-3)! - \dots \\ &= \sum_{i=1}^{n} (-1)^{i+1}\binom{n}{i}(n-i)! = \sum_{i=1}^{n} (-1)^{i+1}\frac{n!}{i!(n-i)!}(n-i)! \\ &= \sum_{i=1}^{n} (-1)^{i+1}\frac{n!}{i!}.\end{aligned}$$

Quindi il numero delle permutazioni che non tengono fisso nessun elemento è

$$n! - \sum_{i=1}^{n} (-1)^{i+1}\frac{n!}{i!} = \sum_{i=0}^{n} (-1)^{i}\frac{n!}{i!}.$$

Poiché il numero totale di permutazioni è $n!$, il rapporto tra il numero delle permutazioni che non tengono fisso alcun elemento e il numero totale delle permutazioni è:

$$\sum_{i=0}^{n} \frac{(-1)^i}{i!} = 1 - 1 + \frac{1}{2} - \frac{1}{6} + \dots.$$

Come si studia nel corso di Analisi questa somma infinita, o *serie numerica*, quando n diventa grande, si avvicina al valore $e^{-1} = 1/e$, dove $e = 2,718\dots$, e cioè ha come valore poco più di $1/3$. Questo significa che, se n è grande, poco più di $1/3$ delle volte e precisamente $1/e$ delle volte, restituendo i compiti, nessuno studente riceve indietro il proprio compito.

Esercizi

(1) Quanti interi pari con due cifre distinte ci sono tra 10 e 99? [R: 41]

(2) Una moneta è lanciata per 30 volte. Quante sono le possibili successioni di testa e croce? [R: 2^{30}]

(3) Quanti modi diversi ci sono di rispondere ad un test composto con 20 domande a risposte vero-falso? [R: 2^{20}]

(4) Quanti modi diversi ci sono di rispondere ad un test con 50 domande a risposta multipla, se 20 domande sono a 3 risposte e 30 a 5 risposte? [R: $3^{20} \cdot 5^{30}$]

(5) Quante targhe automobilistiche si possono formare con 2 lettere seguite da 3 numeri, seguiti ancora da altre 2 lettere, usando le 26 lettere dell'alfabeto e ricordando che lettere e numeri possono essere ripetuti? E quante se solo i numeri possono essere ripetuti mentre le lettere devono essere distinte? [R : $26^4 \cdot 10^3$; $26 \cdot 25 \cdot 24 \cdot 23 \cdot 10^3$]

(6) (a) Quanti sono gli interi compresi tra 100 e 999 che hanno le tre cifre distinte? (Suggerimento: contare prima i possibili valori per le centinaia, poi per le decine ed infine per le unità).
(b) Quanti tra i precedenti sono dispari? (Suggerimento: contare prima i possibili valori per le unità, poi per le centinaia ed infine per le decine).
[R: (a) 648; (b) 320]

(7) (a) Quanti sono i numeri telefonici che hanno 9 cifre e le cui 3 cifre iniziali sono 011?
(b) Quanti tra i precedenti non iniziano con 0110 o con 01199?
(c) Quanti tra i numeri in (a) non contengono né la cifra 8 né la cifra 9?
[R: (a) 10^6; (b) $10^6 - 10^5 - 10^4$; (c) 8^6]

(8) Sia S l'insieme degli interi n tali che $1.000 \leq n \leq 9.999$
(a) Quanti sono i numeri in S che hanno solo cifre dispari?
(b) Quanti sono i numeri in S che hanno solo cifre pari?
(c) Quanti sono i numeri in S che hanno cifre differenti tra loro?
[R: (a) 5^4; (b) $4 \cdot 5^3$; (c) $9^2 \cdot 8 \cdot 7$]

(9) Sia X un insieme di 10 elementi. Quanti sono i sottoinsiemi di X che contengono almeno 3 elementi? [R: 968]

(10) Siano A, B, C, D, E cinque persone.
(a) In quanti modi diversi possono sedersi in fila con A sempre al centro?
(b) In quanti modi diversi si possono sistemare intorno ad un tavolo rotondo, se sistemazioni che differiscono solo per rotazioni vanno identificate?
[R: (a) 4!; (b) 4!]

(11) Si considerino gli insiemi $X = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ e $Y = \{a, b\}$.
(a) Quante sono le funzioni di X in Y?
(b) Quante sono le funzioni iniettive di X in Y?
(c) Quante sono le funzioni suriettive di X in Y?
(d) Quante sono le funzioni di X in Y tali che a ammetta 3 controimmagini?
(e) Quante sono le funzioni di X in Y tali che a e b ammettano lo stesso numero di controimmagini?

[R: (a) 2^{10}; (b) 0; (c) $2^{10}-2$; (d) $\binom{10}{3}$; (e) 504]

(12) (a) Supponiamo di voler costruire fili di 20 palline colorate.
   (i) quanti fili diversi si possono ottenere utilizzando palline bianche, rosse e gialle?
   (ii) quanti fili diversi di palline bianche, rosse e gialle che iniziano e finiscono con una pallina rossa si possono costruire?

(b) Sia X l'insieme di tutti i possibili fili diversi di 20 palline bianche, rosse e gialle (come nel punto (i)). In X si stabilisca la seguente relazione:
$f_1 R f_2 \iff f_1$ e f_2 contengono lo stesso numero di palline rosse, di palline bianche e di palline gialle.
Verificato che R è una relazione d'equivalenza, si dica quanti elementi possiede la classe dei fili con 10 palline rosse e 10 palline bianche.

[R: (a) (i) 3^{20}; (ii) 3^{18}; (b) $\binom{20}{10}$]

(13) Qual è il numero dei possibili ambi su una stessa ruota nel gioco del Lotto? [R: $\binom{90}{2}$]

(14) Quante sono le possibili schedine del Totocalcio? [R: 3^{13}]

(15) Quanti numeri naturali di 8 cifre si possono scrivere senza usare lo 0? Tra questi sono in numeri maggiore quelli in cui compare la cifra 1 o quelli in cui non compare? [R: 9^8; in 8^8 non compare 1; in 9^8-8^8 compare 1]

(16) Dieci alpinisti si legano in cordata in modo che due di loro, inesperti, non sono né al primo né all'ultimo posto. In quanti modi diversi possono farlo? [R: $8\cdot 7\cdot 8!$]

(17) Quante permutazioni delle cifre 0, 1, ..., 9 contengono almeno una delle terne 289, 234 o 487? [R: $3\cdot 8!-6!$]

(18) Quanti sono gli anagrammi della parola BAR (anche privi di senso)? E della parola BABBO (anche privi di senso)? [R: 3!; 5!/3!]

(19) Un cesto contiene cassette per mangianastri di quattro marche differenti. In quanti modi possiamo scegliere 6 cassette dal cesto? [R: 84] (Nota Bene: cosa dobbiamo supporre in modo che 84 sia la risposta esatta?)

(20) Consideriamo la parola COMPUTER.
   (a) Quanti sono gli anagrammi?
   (b) In quanti anagrammi le vocali compaiono consecutivamente?
   (c) In quanti la lettera P sta a sinistra della lettera T?
   (d) In quanti vi sono esattamente 2 lettere tra M e C?

[R: (a) $8!$; (b) $6! \cdot 3!$; (c) $28 \cdot 6!$; (d) $10 \cdot 6!$]

(21) Quanti modi ci sono di sistemare le lettere AAABCDE in modo che due lettere consecutive non siano mai uguali? [R: $7!/3! - 6! + 5! = 10 \cdot 4! = 240$]

(22) Quante sono le mani di 5 carte da un mazzo di 52? In quanti modi si può fare poker (quattro carte di valore uguale)? Quante mani contengono esattamente una coppia? [R: $\binom{52}{5}$; $13 \cdot 48$; $13 \cdot \binom{4}{2} \cdot \binom{12}{3} \cdot 4^3$]

Capitolo 3

Equazioni ricorsive

1. Il concetto di ricorsione

In questa sezione vedremo come la soluzione di molti problemi di enumerazione può essere ottenuta con un ragionamento ricorsivo, cioè esprimendo i termini di una successione in relazione ai termini precedenti della successione stessa. Dopo aver visto una serie di esempi, daremo una definizione generale di equazione ricorsiva e cominceremo a risolvere alcuni casi particolari.

Equazioni ricorsive. Vi sono parecchi problemi la cui soluzione si formula in modo naturale con il linguaggio della ricorsione.

Ordinamento per selezione. Sia L una lista di n numeri, che vogliamo ordinare dal più piccolo al più grande. Un qualunque algoritmo per fare ciò dovrà compiere dei confronti fra coppie dei numeri della lista, e sia C_n il numero di tali confronti. Una stima del numero C_n darà una stima del tempo impiegato dall'algoritmo per ordinare la lista. L'algoritmo di selezione per prima cosa trova il più piccolo numero della lista, e per fare ciò occorrono $n-1$ confronti, scambia questo elemento con il primo della lista e poi riapplica lo stesso metodo alla lista di $n-1$ elementi dal secondo all'ultimo. Il numero di confronti che occorrono per ordinare una lista di $n-1$ elementi è, ovviamente, C_{n-1} e dunque

$$C_n = n - 1 + C_{n-1}$$

Sappiamo inoltre che $C_1 = 0$, perché per ordinare una lista di un solo elemento non sono necessari confronti.

Contare i sottoinsiemi. Sia $N = \{1, 2, \ldots, n\}$ un insieme di n elementi, e poniamo $S_n =$ numero di sottoinsiemi di N. Notiamo che sappiamo già la

risposta ($S_n = 2^n$) ma ora siamo interessati a formulare la soluzione del problema in linguaggio ricorsivo. L'osservazione fondamentale è la seguente: fissato l'elemento $n \in N$, un sottoinsieme di N o contiene n oppure non lo contiene. Ogni sottoinsieme che non contiene n è un sottoinsieme di $N' = \{1, 2, \ldots, n-1\}$, e perciò il numero dei sottoinsiemi di N che non contengono n è S_{n-1}. D'altra parte, un sottoinsieme che contiene n è semplicemente un sottoinsieme di N' a cui aggiungiamo l'elemento n. Dunque

$$\begin{aligned} S_n &= (\text{numero dei sottoinsiemi non contenenti } n) \\ &\qquad + (\text{numero dei sottoinsiemi contenenti } n) \\ &= S_{n-1} + S_{n-1} \\ &= 2S_{n-1} \end{aligned}$$

Sappiamo inoltre che $S_0 = 1$, perché l'insieme vuoto ha solo se stesso come sottoinsieme.

Contare le funzioni biiettive. Sia ancora $N = \{1, 2, \ldots, n\}$ e poniamo ora $F_n =$ numero di funzioni biiettive da N in un altro insieme X di n elementi. Notiamo che, se $f : N \to X$ è iniettiva, l'immagine $f(N)$ ha esattamente n elementi e quindi coincide con tutto X. Dunque la funzione f è suriettiva e perciò biiettiva. Concludiamo che, in questo caso, contare le funzioni biiettive è la stessa cosa che contare le funzioni iniettive, e quindi anche in questo caso conosciamo già la soluzione ($F_n = n!$). Quello che cerchiamo è una formulazione ricorsiva. Per definire una funzione biiettiva $f : N \to X$, abbiamo esattamente n scelte per $f(n)$, e dobbiamo poi mandare biiettivamente l'insieme $N' = \{1, 2, \ldots, n-1\}$ sull'insieme dei rimanenti $n-1$ elementi. Dunque

$$\begin{aligned} F_n &= (\text{numero di scelte per } f(n)) \\ &\qquad \times (\text{numero di biiezioni fra } N' \text{ e gli elementi rimanenti}) \\ &= nF_{n-1} \end{aligned}$$

Sappiamo inoltre che $F_1 = 1$, perché c'è una sola funzione biiettiva fra due insiemi con un solo elemento.

In tutti gli esempi precedenti il problema era determinare una successione di numeri e abbiamo trovato in ogni caso una equazione che ci consente di calcolare il termine n-esimo della successione conoscendo il termine $(n-1)$-esimo. Inoltre, abbiamo anche determinato il primo termine della successione. Questi corrispondono al passo induttivo e al passo iniziale di una dimostrazione per induzione, e dunque c'è una sola successione che soddisfa le condizioni date, e possiamo calcolare tutti i valori della successione a partire dal primo.

Esempio 1.1. Calcolare i primi 5 termini della successione $C_n = n - 1 + C_{n-1}$, ottenuta prima per l'algoritmo di ordinamento per selezione, cominciando dal termine 1.

Soluzione. Sappiamo che $C_1 = 0$. Applicando la relazione si ha

$$C_2 = 2 - 1 + C_1 = 1$$
$$C_3 = 3 - 1 + C_2 = 3$$
$$C_4 = 4 - 1 + C_3 = 6$$
$$C_5 = 5 - 1 + C_4 = 10$$

Esprimiamo ora in modo formale il concetto di equazione ricorsiva:

Definizione 1.2. Una equazione che consente di calcolare il termine n-esimo di una successione $\{a_n\}$ a partire dai termini precedenti (i termini a_{n-1}, a_{n-2}, ...) si dice **equazione ricorsiva** o anche **ricorrenza**.

In effetti negli esempi che abbiamo appena visto era possibile calcolare a_n usando solo a_{n-1} (e magari anche il valore n stesso) senza usare gli altri termini della successione. Una equazione ricorsiva si dice del **primo ordine** se consente di determinare a_n da a_{n-1}. Una equazione ricorsiva si dice di **ordine *r*** se consente di determinare a_n a partire dagli r termini precedenti $a_{n-1}, \ldots, a_{n-r}$.

Esempio 1.3. Qual è l'ordine delle ricorrenze seguenti?

(1) $a_n = 3a_{n-1} + n^2$
(2) $a_n = na_{n-1} + 2^n$
(3) $a_n = \sqrt{a_{n-1}} + a_{n-2}^3$
(4) $a_n = a_{n-1} + a_{n-2} + \cdots + a_1$

Soluzione. Le prime due sono del primo ordine, mentre la (3) è del secondo ordine. L'equazione (4) non ha un ordine, perché per calcolare un termine dobbiamo usare tutti i termini precedenti e non solo un numero prefissato.

Il problema che affronteremo nel nostro studio delle equazioni ricorsive è quello di trovare una soluzione, e cioè una formula che esprima, in funzione di n, tutti i valori della successione a_n data dalla ricorrenza.

Per esempio, sappiamo che $a_n = 2^n$ è una soluzione dell'equazione $a_n = 2a_{n-1}$, $a_0 = 1$; infatti sostituendo si ha $a_0 = 2^0 = 1$ e $a_n = 2^n = 2 \cdot 2^{n-1} = 2a_{n-1}$.

Allo stesso modo si può verificare che $a_n = n(n-1)/2$ è una soluzione dell'equazione ricorsiva $a_n = n - 1 + a_{n-1}$, $a_1 = 0$, equazione che descrive il numero dei confronti che l'algoritmo di selezione compie per ordinare una lista di n elementi.

Come per le equazioni ordinarie non si usa la stessa formula risolutiva per le equazioni di primo o di secondo grado (o di grado ancora più alto), così non esiste un unico metodo per ottenere la soluzione di tutte le equazioni ricorsive, ma piuttosto vari metodi che funzionano in casi particolari. Abbiamo già parlato dell'ordine di una equazione ricorsiva; c'è un secondo concetto che è importante nel classificare le equazioni.

Definizione 1.4. Una equazione ricorsiva del primo ordine si dice **lineare** se esistono funzioni $d(n)$ e $b(n)$ tali che l'equazione ha la forma

$$a_n = b(n)a_{n-1} + d(n).$$

In generale, una equazione ricorsiva si dice **lineare** se è della forma

$$a_n = b_{n-1}(n)a_{n-1} + b_{n-2}(n)a_{n-2} + \cdots + b_0(n)a_0 + d(n)$$

Lineare vuol dire, di solito, di primo grado; qui è importante notare che l'espressione di a_n in funzione dei termini precedenti è di primo grado nei termini della successione, ma magari i coefficienti non sono di primo grado.

Esempio 1.5. Vediamo quali delle equazioni dell'esempio 1.3 sono lineari:

(1) $a_n = 3a_{n-1} + n^2$: è lineare, con $b(n) = 3$ e $d(n) = n^2$;

(2) $a_n = na_{n-1} + 2^n$: è lineare, con $b(n) = n$ e $d(n) = 2^n$;

(3) $a_n = \sqrt{a_{n-1}} + a_{n-2}^3$: non è lineare, poiché a_{n-1} non compare a primo grado, bensì sotto radice (nemmeno a_{n-2} compare a primo grado);

(4) $a_n = a_{n-1} + a_{n-2} + \cdots + a_1$: è lineare, tutti i b_i sono funzioni costanti uguali a 1, e $d(n) = 0$.

Equazioni lineari del primo ordine omogenee. Le equazioni ricorsive che abbiamo ottenuto per il problema dei sottoinsiemi $S_n = 2S_{n-1}$ e per il problema delle funzioni biiettive $F_n = nF_{n-1}$ sono lineari, ma in entrambe abbiamo che $d(n) = 0$, e cioè il "termine noto" è nullo. Queste equazioni hanno un nome speciale:

Definizione 1.6. Una equazione ricorsiva del primo ordine lineare $a_n = b(n)a_{n-1} + d(n)$ si dice **omogenea** se $d(n) = 0$.

La definizione di omogenea è simile per le equazioni di ordine più alto (cioè $d(n) = 0$).

Esempio 1.7. Quale delle ricorrenze lineari seguenti è omogenea?

(1) $a_n = n^2 a_{n-1} + 2a_{n-3}$

(2) $a_n = na_{n-1} + 2^n$

(3) $a_n = a_{n-1} - a_{n-2}$

(4) $a_n = a_{n-2} - n$

Soluzione. Sono omogenee la (1), che è di terzo ordine, e la (3), che è di secondo ordine. La (2) è di primo ordine non omogenea perché $d(n) = 2^n$. La (4) è di secondo ordine non omogenea perché $d(n) = -n$.

Due dei tre esempi iniziali, la ricorrenza $S_n = 2S_{n-1}$ del problema dei sottoinsiemi e la ricorrenza $C_n = C_{n-1}+n-1$ del problema dei confronti nell'algoritmo dell'ordinamento hanno una proprietà in comune: il coefficiente del termine precedente è costante. Anche questa proprietà ha un nome:

Definizione 1.8. Una equazione ricorsiva del primo ordine lineare $a_n = b(n)a_{n-1} + d(n)$ si dice **a coefficienti costanti** se $b(n)$ è costante.

La definizione di equazione a coefficienti costanti è simile per le equazioni di ordine più alto (cioè tutti i coefficienti $b_{n-1}(n), \ldots, b_1(n)$ sono costanti).

Siamo ora pronti a enunciare il primo teorema, che dà una formula per risolvere le equazioni ricorsive del primo ordine, lineari, omogenee a coefficienti costanti.

Teorema 1.9. *L'equazione ricorsiva del primo ordine lineare omogenea a coefficienti costanti*

$$\begin{cases} a_n = ba_{n-1}, & \textit{valida per } n > m \\ a_m = c \end{cases}$$

ha come soluzione

$$a_n = cb^{n-m}, \qquad \textit{valida per } n \geq m$$

Dimostrazione. La dimostrazione avviene per induzione su n: se $n = m$ allora si ha $a_m = cb^{m-m} = cb^0 = c$.

Sia ora $n > m$. Poiché $a_n = cb^{n-m} = bcb^{n-1-m} = ba_{n-1}$, l'espressione $a_n = cb^{n-m}$ soddisfa l'equazione di ricorsione. □

Per esempio, per il problema dei sottoinsiemi abbiamo $S_n = 2S_{n-1}$ e $S_0 = 1$; in questo caso $b = 2$, $c = 1$ e $m = 0$ e dunque $S_n = 2^n$ per ogni $n \geq 0$.

È facile estendere la dimostrazione precedente al caso delle equazioni del primo ordine lineari omogenee a coefficienti non costanti, per esempio l'equazione $a_n = na_{n-1}$, $a_1 = 1$. Per scrivere la formula risolutiva abbiamo bisogno di un simbolo per il prodotto di vari fattori, analogo al simbolo di sommatoria.

Definizione 1.10. Se $b : \mathbb{N} \to \mathbb{R}$ è una funzione, poniamo

$$\prod_{i=m}^{n} b(i) = b(m) \cdot b(m+1) \cdots b(n)$$

Abbiamo allora

Teorema 1.11. *L'equazione ricorsiva del primo ordine lineare omogenea*

$$\begin{cases} a_n = b(n)a_{n-1}, & \text{valida per } n > m \\ a_m = c \end{cases}$$

ha come soluzione

$$a_n = c \prod_{i=m+1}^{n} b(i), \qquad \text{valida per } n > m$$

Dimostrazione. La dimostrazione è identica a quella del teorema precedente. □

Per esempio, per il problema delle funzioni biiettive abbiamo $F_n = nF_{n-1}$, $F_1 = 1$; in questo caso $b(n) = n$, $c = 1$ e $m = 1$ e dunque $F_n = 1 \cdot \prod_{i=2}^{n} i = 1 \cdot 2 \cdot 3 \cdots n = n!$ per ogni $n \geq 1$.

Esercizi svolti.

(1) Sia J_n il numero di modi in cui si possono suddividere n lavori fra 3 computers. Spiegare perché $J_n = 3J_{n-1}$.
Soluzione. Il lavoro n può essere assegnato a uno qualunque dei tre computers, e per ognuna di queste scelte dobbiamo poi assegnare i restanti $n-1$ lavori ai tre computers. Dunque, $J_n = 3J_{n-1}$.

(2) Qual è l'ordine (se esiste) delle seguenti equazioni ricorsive?
(a) $t_n = nt_{n-1}$
(b) $t_n = nt_{n-2} - n$
(c) $t_n = t_{n-1}^2 + n^2$
(d) $t_n = nt_{n-1} + (n-1)t_{n-2}$
(e) $t_n = 2t_2 + 3t_3 + \cdots + (n-1)t_{n-1}$
(f) $t_n = 2t_{n-1} + t_{n-2}^2$
Soluzione. Gli ordini sono:
primo ordine: la (a) e la (c);
secondo ordine: la (b), la (d) e la (f).
La (e) non ha ordine, perché non basta un numero fissato di elementi precedenti per determinare l'elemento t_n.

(3) Quali delle equazioni dell'esercizio (2) sono lineari?
Soluzione. Sono lineari la (a), (b), (d) e (e).

(4) Quali delle equazioni dell'esercizio (2) sono omogenee?
Soluzione. Sono omogenee la (a), (d), (e) e (f).

(5) Quali delle equazioni dell'esercizio (2) sono a coefficienti costanti?
Soluzione. Sono a coefficienti costanti la (c) e la (f).

(6) Risolvere l'equazione $a_n = 3a_{n-1}$, $a_0 = 2$.
Soluzione. L'equazione è del primo ordine, lineare, omogenea, a coefficienti costanti e si può usare la formula del Teorema 1.9 con $b = 3$, $c = 2$ e $m = 0$ ottenendo $a_n = 2(3^n)$ per $n \geq 0$.

(7) Una successione $\{a_n\}$ verifica $a_n = 2a_{n-1}$ per ogni $n \geq 0$ e sappiamo che $a_5 = 96$. Quanto vale a_0?
Soluzione. La successione è soluzione di un'equazione lineare omogena, e dunque $a_n = cb^n$ per ogni $n \geq 0$, dove $b = 2$ e $c = a_0$. Allora $96 = a_5 = a_0 2^5 = 32a_0$ e si ottiene $a_0 = 3$.

(8) Risolvere l'equazione $a_n = \sqrt{n}a_{n-1}$, $a_1 = 4$.
Soluzione. L'equazione è lineare omogenea ma i coefficienti non sono costanti e si può usare la formula del Teorema 1.11 con $b(n) = \sqrt{n}$, $c = 4$ e $m = 1$ ottenendo

$$a_n = 4 \prod_{i=2}^{n} \sqrt{i}, \qquad \text{per } n \geq 2.$$

2. Equazioni del primo ordine lineari

La soluzione delle equazioni del primo ordine lineari. Nel paragrafo precedente abbiamo imparato a risolvere le equazioni del primo ordine omogenee. Vediamo adesso di affrontare il problema della risoluzione di quelle non omogenee. Come abbiamo già visto, è utile trattare prima il caso delle equazioni a coefficienti costanti. Sia dunque

$$a_n = ba_{n-1} + d(n),$$

una tale equazione, valida per $n > 0$ e proviamo a calcolare alcuni termini, per vedere se riusciamo a trovare una forma generale in cui esprimere tali termini. Calcoliamo per esempio a_3: applicando la relazione ricorsiva abbiamo

$$\begin{aligned} a_3 &= ba_2 + d(3) \\ &= b\left(ba_1 + d(2)\right) + d(3) \\ &= b^2\left(ba_0 + d(1)\right) + bd(2) + d(3) \\ &= b^3 a_0 + b^2 d(1) + bd(2) + d(3) \\ &= b^3\left(a_0 + b^{-1}d(1) + b^{-2}d(2) + b^{-3}d(3)\right) \end{aligned}$$

dove nell'ultimo passaggio abbiamo raccolto b^3. Questi calcoli suggeriscono che per a_4 si potrebbe avere

$$a_4 = b^4 \left(a_0 + b^{-1}d(1) + b^{-2}d(2) + b^{-3}d(3) + b^{-4}d(4)\right)$$

e così via per i termini seguenti.

Enunciamo e dimostriamo allora il

Teorema 2.1. *L'equazione ricorsiva del primo ordine lineare a coefficienti costanti*

$$\begin{cases} a_n = ba_{n-1} + d(n), & \text{valida per } n > m \\ a_m = c \end{cases}$$

ha come soluzione

$$a_n = b^{n-m}\left(c + \sum_{i=1}^{n-m} d(m+i)b^{-i}\right), \qquad \text{valida per } n > m$$

Dimostrazione. Per $n = m+1$ la formula per la soluzione dà

$$a_{m+1} = b^{m+1-m}(c + d(m+1)b^{-1}) = bc + d(m+1) = ba_m + d(m+1),$$

che è quello che dà la ricorrenza. Supponiamo ora che sia $n > m+1$ e che

$$a_{n-1} = b^{n-1-m}\left(c + \sum_{i=1}^{n-1-m} d(m+i)b^{-i}\right).$$

Poiché $a_n = ba_{n-1} + d(n)$, sostituendo si ottiene

$$\begin{aligned} a_n &= b\left[b^{n-1-m}\left(c + \sum_{i=1}^{n-1-m} d(m+i)b^{-i}\right)\right] + d(n) \\ &= b^{n-m}\left(c + \sum_{i=1}^{n-1-m} d(m+i)b^{-i}\right) + b^{n-m}b^{-n+m}d(m+(n-m)) \\ &= b^{n-m}\left(c + \sum_{i=1}^{n-m} d(m+i)b^{-i}\right). \end{aligned}$$

Quindi, per il principio di induzione, la formula è valida per ogni $n > m$. □

Possiamo risolvere ora la ricorsione $C_n = C_{n-1} + n - 1$ che abbiamo trovato nel considerare il problema dell'algoritmo di ordinamento per selezione. In questo caso il primo termine della ricorsione è $C_1 = 0$, in quanto per una lista di un solo elemento non sono necessari confronti e dunque possiamo applicare la formula del Teorema 2.1 con $m = 1$, $b = 1$, $d(n) = n-1$, $C_1 = 0$ e quindi

$$C_n = 1^{n-1}\left(0 + \sum_{i=1}^{n-1}(1+i-1)1^{-i}\right) = \sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}.$$

Vediamo un altro esempio: risolviamo l'equazione

$$t_n = 2t_{n-1} + 2^n, \qquad t_0 = 1$$

In questo caso abbiamo $m = 0$, $b = 2$, $d(n) = 2^n$, $t_0 = 1$ e applicando la formula del Teorema 2.1 si ha

$$t_n = 2^n \left(1 + \sum_{i=1}^{n} 2^i 2^{-i}\right) = 2^n \left(1 + \sum_{i=1}^{n} 1\right) = 2^n(n+1).$$

A questo punto possiamo scrivere una formula valida per ogni equazione del primo ordine lineare, che comprende come casi speciali le formule già viste. Va però detto che la formula generale è complicata e quindi conviene usare quelle più semplici quando ciò sia possibile.

Teorema 2.2. *L'equazione ricorsiva del primo ordine lineare*

$$\begin{cases} a_n = b(n)a_{n-1} + d(n), & \textit{valida per } n > m \\ a_m = c \end{cases}$$

ha come soluzione

$$a_n = \left(\prod_{i=m+1}^{n} b(i)\right) \left[c + \sum_{i=m+1}^{n} \left(d(i) \prod_{j=m+1}^{i} \frac{1}{b(j)}\right)\right], \quad \textit{valida per } n > m.$$

La dimostrazione è simile a quella del Teorema 2.1 e non la riportiamo. Vediamo piuttosto un esempio di uso di questa formula.

Esempio 2.3. Risolvere l'equazione ricorsiva $a_n = na_{n-1} + n!$, $a_0 = 1$.

Soluzione. In questa equazione abbiamo $m = 0$, $a_0 = 1$, $b(i) = i$ e $d(i) = i!$. Sostituendo si ottiene

$$a_n = \left(\prod_{i=1}^{n} i\right) \left[1 + \sum_{i=1}^{n} \left(i! \prod_{j=1}^{i} \frac{1}{j}\right)\right] = n! \left[1 + \sum_{i=1}^{n} \frac{i!}{i!}\right] = n!(1+n) = (n+1)!$$

Algoritmi del tipo Divide et Impera. Un approccio alla soluzione di un problema può essere quello di dividere il problema in alcuni sottoproblemi più "piccoli", risolvere questi sottoproblemi e poi mettere insieme le soluzioni per ottenere la soluzione del problema iniziale. Un algoritmo che usi un simile metodo si dice **Algoritmo Divide et Impera** (dal famoso motto attribuito a Filippo il Macedone sul modo di conquistare e mantenere il potere) o **Algoritmo di Decomposizione**. Vi sono molti algoritmi che usano questo metodo e l'analisi del tempo impiegato da tali algoritmi conduce a equazioni ricorsive del primo ordine.

Ricerca binaria. Nell'algoritmo di **ricerca binaria**, cerchiamo una parola in una lista ordinata alfabeticamente (come per esempio un elenco telefonico) nel modo seguente: confrontiamo la parola con l'elemento a metà della lista, per vedere se la parola cercata è nella prima o nella seconda metà della lista. Se la parola a metà della lista non è quella cercata, cerchiamo ancora nella metà appropriata. In questo modo continuiamo a dimezzare la lunghezza della lista, fino ad ottenere lunghezza 1. A questo punto facciamo un solo confronto, che se risulta negativo indica che la parola cercata non appartiene alla lista. Per dividere il problema in metà è sufficiente un solo confronto e quando abbiamo trovato l'elemento il problema è risolto senza ulteriori passi. Dunque se c_n è il numero (massimo) di confronti necessari in una lista di lunghezza n si ha

$$\begin{cases} c_n = c_{n/2} + 1, & (n \text{ pari e } n > 0) \\ c_1 = 1 \end{cases}$$

Merging. Un altro importante algoritmo che usa la tecnica di decomposizione è un algoritmo di ordinamento noto come **merge sort**. Questo algoritmo ha la seguente descrizione, di stile "divide et impera":

> Se la lista ha un solo elemento, è già ordinata. Altrimenti, può essere divisa in due liste della stessa lunghezza e ognuna ordinata separatamente. Si riuniscono poi le due metà per ottenere una unica lista ordinata.

Per riunire due liste L_1 e L_2 ordinate in ordine crescente in una sola lista L ordinata in ordine crescente si può procedere come segue: si confronta il primo elemento di L_1 con il primo elemento di L_2, e si pone in L il minore dei due. Diciamo che il minore sia in L_1; si confronta ora il secondo elemento di L_1 con il primo elemento di L_2 e si pone in L il minore dei due. Si procede così fino a porre in L tutti gli elementi di L_1 e L_2. Poiché dopo ogni confronto un elemento di L_1 oppure di L_2 viene posto in L, il numero di confronti necessario per riunire le due liste è al più il numero degli elementi appartenti alle liste meno uno (dopo aver scelto il penultimo elemento, non è necessario un confronto per l'ultimo elemento).

Il numero di confronti necessario per ordinare la lista è il numero necessario per dividere la lista in metà più il numero necessario per ordinare le due metà più il numero necessario per riunire le due metà. Se indichiamo con s_n il numero massimo di confronti necessario ad ordinare una lista di lunghezza n, allora $s_{n/2}$ è il numero massimo di confronti necessari per ordinare una lista lunga la metà. L'algoritmo di riunione che abbiamo appena descritto compie $n-1$ confronti, e non sono necessari confronti per dividere

in due una lista. Otteniamo perciò:

$$\begin{cases} s_n = 2s_{n/2} + n - 1, & (n \text{ pari e } n > 0) \\ s_1 = 0 \end{cases}$$

Definizione 2.4. Una equazione ricorsiva si dice del tipo **divide et impera** se ha la forma

$$s_n = bs_{n/a} + d(n), \qquad \text{per un intero } a > 1.$$

Negli esempi che abbiamo appena visto si ha $a = 2$. Useremo questi esempi per illustrare il metodo di risoluzione di questo tipo di equazioni. Nei due esempi, le equazioni hanno senso solo se n è una potenza di 2, in modo da poter continuare a dividere per due; in generale, una equazione come nella Definizione 2.4 ha senso solo per n una potenza di a.

Poiché n è una potenza di 2, possiamo scrivere $n = 2^k$. Sostituendo si ottiene

$$\begin{aligned} c_{2^k} &= c_{2^{k-1}} + 1 \qquad \text{e} \\ s_{2^k} &= 2s_{2^{k-1}} + 2^k - 1 \end{aligned}$$

Ponendo $v_k = c_{2^k}$ e $u_k = s_{2^k}$, otteniamo le equazioni del primo ordine lineari:

$$\begin{aligned} v_k &= v_{k-1} + 1, \quad v_0 = 1 \qquad \text{e} \\ u_k &= 2u_{k-1} + 2^k - 1, \quad u_0 = 0 \end{aligned}$$

Poiché queste sono lineari del primo ordine a coefficienti costanti, possiamo scrivere le soluzioni usando il Teorema 2.1, ottenendo

$$\begin{aligned} v_k &= 1 + k \qquad \text{e} \\ u_k &= 2^k \left(0 + \sum_{i=1}^{k} (2^i - 1) \cdot 2^{-i} \right) \\ &= 2^k \left(k - \sum_{i=1}^{k} 2^{-i} \right) = 2^k \left(k - \sum_{i=1}^{k} \left(\frac{1}{2} \right)^i \right) \end{aligned}$$

Per calcolare la sommatoria, ricordiamo che se q è un qualunque numero, si ha

$$\begin{aligned}\sum_{i=1}^{k} q^i &= q + q^2 + \cdots + q^k \\ &= q(1 + \cdots + q^{k-1}) \\ &= q\frac{(1-q)(1 + \cdots + q^{k-1})}{1-q} \\ &= q\frac{1-q^k}{1-q}\end{aligned}$$

dove l'ultima uguaglianza è un prodotto notevole. Ponendo allora $q = 1/2$ e svolgendo i calcoli si ottiene

$$u_k = 2^k\left(k - 1 + \frac{1}{2^k}\right) = 2^k \cdot (k-1) + 1$$

Poiché le equazioni originali erano espresse in termini di n, dobbiamo eseguire la sostituzione inversa, cioè $k = \log_2(n)$. Si ha quindi

$$\begin{aligned}c_n &= 1 + \log_2(n) \qquad \text{e} \\ s_n &= n\log_2(n) - n + 1\end{aligned}$$

Queste soluzioni sono valide per valori di n che sono potenze di 2. Notiamo che entrambe le soluzioni sono espresse in termini di logaritmi, e poiché la funzione logaritmo ha un ordine di crescita molto lento, questo vuol dire che gli algoritmi del tipo *divide et impera* sono di solito piuttosto veloci. In particolare, l'algoritmo di merge sort impiega un numero di confronti dell'ordine di $n\log_2(n)$ per una lista di n elementi. Altri algoritmi, come l'ordinamento per selezione, eseguono un numero di confronti dell'ordine di n^2 e sono dunque molto più lenti per liste di notevoli dimensioni. Per esempio, $2^{10} = 1024$, e cioè $\log_2(1024) = 10$, e quindi per le proprietà dei logaritmi $\log_2(1.000.000)$ vale circa 20; dunque per una lista di un milione di elementi (per esempio, un elenco di indirizzi a cui spedire del materiale pubblicitario) l'algoritmo merge sort compie circa 20 milioni di confronti, mentre l'algoritmo di selezione ne compie circa mille miliardi.

Vediamo un altro esempio in dettaglio. Consideriamo l'equazione

$$t_n = 3t_{n/2} + n^2$$

Per convertirla in una equazione del primo ordine, dobbiamo usare la sostituzione $n = 2^k$, e ponendo $s_k = t_n$ avremo

$$\begin{aligned} s_k &= 3s_{k-1} + (2^k)^2 \\ &= 3s_{k-1} + 2^{2k} \\ &= 3s_{k-1} + 4^k \end{aligned}$$

Anche questa equazione è lineare, del primo ordine, a coefficienti costanti, e possiamo usare la formula del Teorema 2.1 per ottenere la soluzione: si ha

$$\begin{aligned} s_k &= 3^k \left(s_0 + \sum_{i=1}^{k} 3^{-i} 4^i \right) \\ &= 3^k \left(s_0 + \sum_{i=1}^{k} \left(\frac{4}{3} \right)^i \right) \end{aligned}$$

Calcoliamo la sommatoria come prima, ponendo $q = 4/3$; svolgendo tutti i calcoli si ottiene

$$s_k = 3^k s_0 - 4 \cdot 3^k + 4 \cdot 4^k.$$

Ricordando che $n = 2^k$, possiamo scrivere $4^k = (2^2)^k = 2^{2k} = (2^k)^2 = n^2$ e $3^k = (2^{\log_2(3)})^k = 2^{\log_2(3)k} = (2^k)^{\log_2(3)} = n^{\log_2(3)}$, ottenendo alla fine

$$t_n = (t_1 - 4) n^{\log_2(3)} + 4n^2.$$

Esercizi svolti.

(1) Risolvere le seguenti equazioni
   (a) $s_n = 3s_{n-1} + 3^n$, $s_0 = 1$;
   (b) $s_n = 3s_{n-1} + n3^n$, $s_1 = 4$;
   (c) $s_n = 2s_{n-1} + 1$, $s_3 = 1$;
   (d) $s_n = \frac{1}{2} s_{n-1} + 2^n$, $s_0 = 4$.

   Soluzione. Tutte le equazioni sono del primo ordine, lineari, a coefficienti costanti, e quindi possiamo usare la formula del Teorema 2.1.
   (a) Si ha $b = 3$, $d(n) = 3^n$, $m = 0$, $s_0 = 1$. La soluzione è

$$s_n = b^n \left(s_0 + \sum_{i=1}^{n} d(i) b^{-i} \right) = 3^n \left(1 + \sum_{i=1}^{n} 3^i 3^{-i} \right) = 3^n (n+1).$$

(b) Si ha $b = 3$, $d(n) = n3^n$, $m = 1$, $s_1 = 4$. La soluzione è

$$s_n = b^{n-1}\left(s_1 + \sum_{i=1}^{n-1} d(1+i)b^{-i}\right) = 3^{n-1}\left(4 + \sum_{i=1}^{n-1}(i+1)3^{i+1}3^{-i}\right)$$
$$= 3^{n-1}\left(4 + 3\sum_{i=1}^{n-1}(i+1)\right) = 3^{n-1}\left[4 + 3\left(\frac{n(n+1)}{2} - 1\right)\right]$$
$$= 3^{n-1}\left(1 + \frac{3n(n+1)}{2}\right).$$

(c) Si ha $b = 2$, $d(n) = 1$, $m = 3$, $s_3 = 1$. La soluzione è

$$s_n = b^{n-3}\left(s_3 + \sum_{i=1}^{n-3} d(3+i)b^{-i}\right) = 2^{n-3}\left(1 + \sum_{i=1}^{n-3} 1 \cdot 2^{-i}\right)$$
$$= 2^{n-3}\sum_{i=0}^{n-3}\left(\frac{1}{2}\right)^i = 2^{n-3}\frac{1-(1/2)^{n-2}}{1-1/2} = 2^{n-2}\left(1 - \frac{1}{2^{n-2}}\right)$$
$$= 2^{n-2} - 1.$$

(d) Si ha $b = 1/2$, $d(n) = 2^n$, $m = 0$, $s_0 = 4$. La soluzione è

$$s_n = b^n\left(s_0 + \sum_{i=1}^{n} d(i)b^{-i}\right) = \frac{1}{2^n}\left(4 + \sum_{i=1}^{n} 2^i \cdot \frac{1}{2^{-i}}\right)$$
$$= \frac{1}{2^n}\left(4 + \sum_{i=1}^{n} 2^{2i}\right) = \frac{1}{2^n}\left(4 + \sum_{i=1}^{n} 4^i\right) = \frac{1}{2^n}\left(4 + \frac{1-4^{n+1}}{1-4} - 1\right)$$
$$= \frac{1}{2^n}\left(3 + \frac{4^{n+1}-1}{3}\right) = \frac{1}{2^n}\frac{4^{n+1}+8}{3}$$
$$= \frac{4}{3}2^n + \frac{8}{3}2^{-n}.$$

(2) Risolvere le seguenti equazioni del tipo *divide et impera*
   (a) $a_n = a_{n/3} + 2$, $a_1 = 1$;
   (b) $a_n = 2a_{n/3} + n$, $a_1 = 1$;
   (c) $a_n = 3a_{n/3} + 1$, $a_1 = 1$.

Soluzione. Tutte le equazioni sono del tipo *divide et impera*, con $a = 3$. Ponendo $n = 3^k$ si trasformano in equazioni del primo ordine, lineari, a coefficienti costanti.

(a) Poniamo $n = 3^k$, $u_k = a_n$. Si ha l'equazione $u_k = u_{k-1} + 2$, $u_0 = 1$ e la soluzione è

$$u_k = 1^k\left(1 + \sum_{i=1}^{k} 2 \cdot 1^{-i}\right) = 1 + 2\sum_{i=1}^{k} 1$$
$$= 2k + 1$$

e dunque

$$a_n = 2\log_3 n + 1.$$

(b) Poniamo $n = 3^k$, $u_k = a_n$. Si ha l'equazione $u_k = 2u_{k-1} + 3^k$, $u_0 = 1$ e la soluzione è

$$\begin{aligned} u_k &= 2^k \left(1 + \sum_{i=1}^{k} 3^i \cdot 2^{-i}\right) = 2^k \left(1 + \sum_{i=1}^{k} \left(\frac{3}{2}\right)^i\right) \\ &= 2^k \left(\sum_{i=0}^{k} \left(\frac{3}{2}\right)^i\right) = 2^k \frac{1-(3/2)^{k+1}}{1-3/2} \\ &= -2^{k+1}\left(1-(3/2)^{k+1}\right) = 3^{k+1} - 2^{k+1} \end{aligned}$$

e dunque

$$a_n = 3n - 2 \cdot 2^{\log_3 n}.$$

(c) Poniamo $n = 3^k$, $u_k = a_n$. Si ha l'equazione $u_k = 3u_{k-1} + 1$, $u_0 = 1$ e la soluzione è

$$\begin{aligned} u_k &= 3^k \left(1 + \sum_{i=1}^{k} 1 \cdot 3^{-i}\right) = 3^k \sum_{i=0}^{k} \left(\frac{1}{3}\right)^i \\ &= 3^k \frac{1-(1/3)^{k+1}}{1-1/3} = \frac{1}{2}(3^{k+1} - 1) \\ &= \frac{1}{2}(3 \cdot 3^k - 1) \end{aligned}$$

e dunque

$$a_n = \frac{1}{2}(3n - 1).$$

3. Equazioni del secondo ordine lineari

Finora abbiamo incontrato solo esempi di ricorsione del primo ordine. Vediamo ora alcuni casi in cui la ricorsione è del secondo ordine. Il primo esempio è forse il più famoso di tutti: la successione di Fibonacci.

La successione di Fibonacci. Consideriamo una popolazione di conigli che, all'inizio, consiste in una coppia di conigli appena nati. Un coniglio di questa popolazione che ha meno di un mese di vita non si riproduce, ma nel secondo mese di vita diventa adulto e a partire dal secondo mese, e per ogni mese successivo, ogni coppia di conigli dà vita ad una nuova coppia di conigli. Quante coppie di conigli sono presenti alla fine del primo mese? Alla fine del secondo, terzo, quarto mese?

Possiamo riassumere la situazione nella tabella seguente:

Mese	Coppie di conigli appena nati alla fine del mese	Coppie di conigli adulti alla fine del mese	Numero totale di coppie di conigli alla fine del mese
0	1	0	1
1	0	1	1
2	1	1	2
3	1	2	3
4	2	3	5
5	3	5	8
.	.	.	.
.	.	.	.
.	.	.	.

Cominciamo con una coppia di conigli appena nati. Alla fine del mese 0 (cioè all'inizio) abbiamo una coppia di conigli non ancora adulti. Alla fine del primo mese, abbiamo sempre una coppia di conigli, che però è diventata adulta. Durante il secondo mese, la coppia adulta dà vita ad una nuova coppia di conigli, e quindi alla fine del secondo mese abbiamo due coppie di conigli, una adulta e una appena nata. Durante il terzo mese la coppia adulta dà vita ad una nuova coppia, e la coppia precedente diventa adulta. Dunque alla fine del terzo mese abbiamo due coppie adulte e una appena nata.

Poniamo f_n = numero dei conigli presenti alla fine del mese n. Osservando la tabella, si nota che la quantità f_n soddisfa la relazione

$$f_n = f_{n-1} + f_{n-2}, \quad \text{per } n \geq 2.$$

Il motivo è semplice da capire: alla fine del mese n si sono riprodotte le coppie adulte presenti alla fine del mese $n-1$, cioè tutte le coppie presenti alla fine del mese $n-2$, che sono f_{n-2}. A queste vanno aggiunte le coppie presenti alla fine del mese $n-1$, che sono f_{n-1}. Le condizioni iniziali sono $f_0 = 1$ e $f_1 = 1$, e questo determina completamente la successione di Fibonacci.

Un modello di crescita di piante. Questo esempio riguarda la crescita dei rami di una pianta. In generale, ogni anno i nuovi rami spuntano dai rami che sono cresciuti nell'anno precedente. Supponiamo che, in una certa varietà di piante, ogni anno spuntino esattamente due nuovi rami da ogni ramo che era spuntato l'anno precedente. Quanti rami sono presenti dopo un anno? Dopo due, tre, quattro anni?

Poniamo s_n = rami presenti dopo n anni. Durante l'anno n, i rami spuntati l'anno precedente sono $s_{n-1} - s_{n-2}$, (differenza fra i rami presenti dopo $n-1$ anni e quelli presenti dopo $n-2$ anni), e quindi durante l'anno n spunteranno $2(s_{n-1} - s_{n-2})$ nuovi rami. Aggiungendo a questi i rami vecchi,

che sono s_{n-1}, si ottiene la ricorsione

$$s_n = s_{n-1} + 2(s_{n-1} - s_{n-2}) = 3s_{n-1} - 2s_{n-2}.$$

Notiamo che per determinare completamente la successione s_n dobbiamo dare le condizioni iniziali s_0 = numero di rami presenti all'inizio e s_1 = numero di rami dopo un anno.

Scheduling di processi. Un computer può eseguire tre tipi di processi. Il processo 1 impiega un secondo per l'esecuzione, mentre i processi 2 e 3 impiegano due secondi. Quante sono le possibili sequenze di processi che possono essere eseguite in 1 secondo? In 2, 3, 4 secondi?

Rappresentiamo le sequenze come successioni delle cifre 1, 2 e 3. È chiaro che in un secondo è possibile eseguire solo il processo 1 una volta. In due secondi le sequenze possibili sono 11, 2, e 3. In tre secondi le sequenze possibili si ottengono eseguendo il processo 1 in coda ad una sequenza che impiega due secondi, e cioè 111, 21, 31, oppure eseguendo il processo 2 o il processo 3 in coda ad una sequenza che impiega un secondo, e cioè 12 e 13. In totale ci sono 5 sequenze possibili in tre secondi.

Poniamo t_n = numero di sequenze di processi che si possono eseguire in n secondi. Il ragionamento appena fatto mostra che la ricorsione è

$$t_n = t_{n-1} + 2t_{n-2}$$

poiché ogni sequenza di lunghezza $n-1$ produce una nuova sequenza aggiungendo il processo 1, mentre ogni sequenza di lunghezza $n-2$ produce due nuove sequenze, una aggiungendo il processo 2 e un'altra aggiungendo il processo 3.

Le condizioni iniziali sono, come abbiamo visto, $t_1 = 1$, $t_2 = 3$.

Le soluzioni di una ricorsione del secondo ordine. Tutti gli esempi di ricorsione che abbiamo visto possono essere scritti nella forma

$$a_n + ba_{n-1} + ca_{n-2} = 0$$

e cioè sono equazioni ricorsive **del secondo ordine, lineari, omogenee, a coefficienti costanti**. In quello che segue impareremo a risolvere equazioni di questo tipo. Notiamo che la risoluzione di una equazione non omogenea è notevolmente più difficile (e non ce ne occuperemo).

Per prima cosa osserviamo che una equazione omogenea ha sempre una soluzione del tipo $a_n = 0$ per ogni n, ma questa non è molto interessante. Vediamo perciò se è possibile trovare soluzioni non identicamente nulle. Nel paragrafo 1 abbiamo visto che la soluzione di un'equazione *del primo ordine,*

lineare, omogenea, a coefficienti costanti è data da una funzione esponenziale, del tipo $a_n = r^n$. Ci chiediamo quindi se questo può essere vero anche per le equazioni del secondo ordine.

Teorema 3.1. *La successione $a_n = r^n$ è una soluzione non identicamente nulla della ricorsione*

$$a_n + ba_{n-1} + ca_{n-2} = 0$$

se e solo se r è una soluzione dell'equazione $x^2 + bx + c = 0$.

Dimostrazione. Supponiamo che $a_n = r^n$ sia una soluzione dell'equazione $a_n + ba_{n-1} + ca_{n-2} = 0$. Sostituendo si ottiene $r^n + br^{n-1} + cr^{n-2} = 0$ per ogni n e poiché $r \neq 0$ (altrimenti $a_n = 0$ per ogni n) si può dividere per r^{n-2} ottenendo

$$r^2 + br + c = 0,$$

e quindi r è una soluzione dell'equazione $x^2 + bx + c = 0$.

Viceversa, supponiamo che r sia una soluzione dell'equazione $x^2 + bx + c = 0$. Moltiplicando per r^{n-2} si ottiene $r^n + br^{n-1} + cr^{n-2} = 0$ e questo mostra che $a_n = r^n$ è una soluzione dell'equazione ricorsiva. □

Il polinomio $x^2 + bx + c$ viene detto **polinomio caratteristico** dell'equazione $a_n + ba_{n-1} + ca_{n-2} = 0$. Dunque per trovare soluzioni di tipo esponenziale basta trovare le radici del polinomio caratteristico. Notiamo che le radici del polinomio potrebbero essere dei numeri reali oppure addirittura dei numeri complessi, se $\Delta = b^2 - 4c$ è negativo.

Osserviamo però che le soluzioni così trovate non sono necessariamente compatibili con le condizioni iniziali. Riprendiamo il terzo esempio visto: l'equazione ricorsiva $t_n - t_{n-1} - 2t_{n-2} = 0$ ha polinomio caratteristico $x^2 - x - 2$, le cui radici sono $r_1 = -1$ e $r_2 = 2$, come si ottiene risolvendo l'equazione di secondo grado. Dunque abbiamo due soluzioni, $t_n = (-1)^n$ e $t_n = 2^n$. Le condizioni iniziali sono $t_1 = 1$, $t_2 = 3$ e vediamo perciò che nessuna delle due soluzioni trovate soddisfa le condizioni iniziali.

Si pone dunque il problema di trovare altre soluzioni per un'equazione del secondo ordine, in modo da poter soddisfare le condizioni iniziali. A questo proposito vale il seguente teorema, che dipende in modo essenziale dal fatto che stiamo considerando equazioni **lineari** ed **omogenee**. Enunceremo (e dimostreremo) il teorema solo per equazioni ricorsive del secondo ordine, ma in effetti si potrebbe generalizzare ad equazioni di ordine qualunque.

Teorema 3.2. *Data l'equazione ricorsiva* **lineare** *ed* **omogenea**

$$a_n + ba_{n-1} + ca_{n-2} = 0,$$

siano a_n e a'_n due soluzioni. Allora, per ogni scelta di numeri C_1 e C_2, la successione

$$a''_n = C_1 a_n + C_2 a'_n$$

è una soluzione dell'equazione.

Dimostrazione. È una semplice verifica. Basta sostituire l'espressione data nell'equazione e verificare che si ha una identità. Infatti, mettendo in evidenza C_1 e C_2 si ha:

$$\begin{aligned} a''_n + ba''_{n-1} + ca''_{n-2} = \\ (C_1 a_n + C_2 a'_n) + b(C_1 a_{n-1} + C_2 a'_{n-1}) + c(C_1 a_{n-2} + C_2 a'_{n-2}) = \\ C_1(a_n + ba_{n-1} + ca_{n-2}) + C_2(a'_n + ba'_{n-1} + ca'_{n-2}) = 0 \end{aligned}$$

poiché, essendo a_n e a'_n soluzioni dell'equazione ricorsiva, le quantità in parentesi nell'ultima riga sono nulle per ogni n. □

Tornando all'esempio precedente, il Teorema 3.2 afferma che, comunque siano scelti C_1 e C_2, la successione

$$t_n = C_1(-1)^n + C_2 2^n$$

è una soluzione dell'equazione $t_n - t_{n-1} - 2t_{n-2} = 0$. Abbiamo dunque ottenuto **infinite** soluzioni dell'equazione ricorsiva, e fra queste cerchiamo l'**unica** che soddisfi le condizioni iniziali. Per fare questo, basta imporre che sia $t_1 = 1$, $t_2 = 3$. Si ottiene il sistema

$$\begin{cases} -C_1 + 2C_2 = 1 \\ C_1 + 4C_2 = 3 \end{cases}$$

che ha per soluzioni $C_1 = 1/3$, $C_2 = 2/3$. Concludiamo che l'equazione ricorsiva che descrive le possibili sequenze di processi del terzo esempio

$$\begin{cases} t_n - t_{n-1} - 2t_{n-2} = 0, \quad n > 2 \\ t_1 = 1, \quad t_2 = 3 \end{cases}$$

ha per soluzione

$$t_n = \frac{1}{3}\left[(-1)^n + 2^{n+1}\right], \quad n \geq 1.$$

Vediamo un esempio differente: consideriamo l'equazione

$$\begin{cases} a_n - 4a_{n-1} + 4a_{n-2} = 0, \quad n > 2 \\ a_1 = 1, \quad a_2 = 1 \end{cases}$$

Il polinomio caratteristico è $x^2 - 4x + 4 = (x-2)^2$ e dunque vi è una sola radice, $r = 2$. Il Teorema 3.2 afferma che $a_n = C2^n$ è una soluzione per ogni scelta di C, ma si vede subito che non si può scegliere C in modo che sia contemporaneamente $a_1 = 2C = 1$, $a_2 = 4C = 1$. Dobbiamo dunque trovare

un'altra soluzione, indipendente da $a_n = 2^n$ per poter avere **due** costanti C_1 e C_2 da scegliere in modo indipendente.

Il Teorema seguente spiega come trovare un'altra soluzione.

Teorema 3.3. *Data l'equazione ricorsiva*

$$a_n + ba_{n-1} + ca_{n-2} = 0,$$

supponiamo che il polinomio caratteristico x^2+bx+c abbia una sola radice r, di molteplicità due. Allora le successioni

$$a_n = r^n$$
$$a'_n = n \cdot r^n$$

sono entrambe soluzioni dell'equazione ricorsiva.

Dimostrazione. Il polinomio caratteristico ha una sola soluzione se e solo se $\Delta = b^2 - 4c = 0$ e cioè se e solo se $x^2 + bx + c = (x - r)^2$. Ma allora si ha $b = -2r$, $c = r^2$.

Sappiamo già che $a_n = r^n$ è una soluzione dell'equazione ricorsiva (Teorema 3.1). Dobbiamo perciò solo verificare che anche a'_n è una soluzione. Si ha

$$\begin{aligned} nr^n + b(n-1)r^{n-1} + c(n-2)r^{n-2} = \\ nr^n - 2r(n-1)r^{n-1} + r^2(n-2)r^{n-2} = \\ nr^n - 2nr^n + 2r^n + nr^n - 2r^n = 0 \end{aligned}$$

e dunque anche a'_n è soluzione. □

Ritorniamo all'equazione

$$\begin{cases} a_n - 4a_{n-1} + 4a_{n-2} = 0, \quad n > 2 \\ a_1 = 1, \quad a_2 = 1 \end{cases}$$

Per i Teoremi precedenti abbiamo che, comunque siano scelti C_1 e C_2, la successione

$$a_n = C_1 2^n + C_2 n 2^n$$

è una soluzione dell'equazione. Per trovare la soluzione che soddisfa le condizioni iniziali, imponiamo $a_1 = 2C_1 + 2C_2 = 1$ e $a_2 = 4C_1 + 8C_2 = 1$, e risolvendo il sistema si ottiene $C_1 = 3/4$, $C_2 = -1/4$. Abbiamo perciò che l'equazione ha soluzione

$$a_n = \frac{3}{4}2^n - \frac{1}{4}n \cdot 2^n = (3 - n)2^{n-2}, \quad n \geq 1.$$

Possiamo a questo punto riassumere quello che abbiamo fatto, dando un metodo per risolvere tutte le equazioni del secondo ordine, lineari, omogenee, a coefficienti costanti.

Sia data l'equazione

$$\begin{cases} a_n + ba_{n-1} + ca_{n-2} = 0, \quad n > m+1 \\ a_m = A, \quad a_{m+1} = B \end{cases}$$

Per prima cosa si risolve l'equazione $x^2 + bx + c = 0$, trovando così le radici del polinomio caratteristico. Vi sono due casi:

(1) Le radici r_1 e r_2 sono distinte (reali o complesse). Allora, per i Teoremi 3.1 e 3.2

$$a_n = C_1 r_1^n + C_2 r_2^n$$

è una soluzione dell'equazione, per ogni scelta dei numeri C_1 e C_2.

(2) Le radici sono coincidenti $r_1 = r_2 = r$. Allora per i Teoremi 3.2 e 3.3,

$$a_n = C_1 r^n + C_2 n r^n$$

è una soluzione dell'equazione, per ogni scelta dei numeri C_1 e C_2.

A questo punto si devono determinare le costanti C_1 e C_2 in modo da soddisfare le condizioni iniziali. Per fare ciò si risolve il sistema

$$\begin{cases} a_m = A \\ a_{m+1} = B \end{cases}$$

nelle incognite C_1 e C_2. Con le soluzioni ottenute, si trova la soluzione dell'equazione ricorsiva.

Applichiamo questo metodo per risolvere l'equazione di Fibonacci

$$\begin{cases} f_n - f_{n-1} - f_{n-2} = 0, \quad n > 1 \\ f_0 = 1, \quad f_1 = 1 \end{cases}$$

Il primo passo è risolvere l'equazione $x^2 - x - 1 = 0$ che ha le due soluzioni reali e distinte

$$r_1 = \frac{1+\sqrt{5}}{2}, \quad r_2 = \frac{1-\sqrt{5}}{2}.$$

Dunque la successione di Fibonacci ha la forma

$$f_n = C_1 \left(\frac{1+\sqrt{5}}{2}\right)^n + C_2 \left(\frac{1-\sqrt{5}}{2}\right)^n.$$

Il secondo passo è determinare C_1 e C_2: risolviamo il sistema

$$\begin{cases} f_0 = C_1 \left(\frac{1+\sqrt{5}}{2}\right)^0 + C_2 \left(\frac{1-\sqrt{5}}{2}\right)^0 \\ f_1 = C_1 \left(\frac{1+\sqrt{5}}{2}\right)^1 + C_2 \left(\frac{1-\sqrt{5}}{2}\right)^1 \end{cases}$$

e cioè

$$\begin{cases} C_1 + C_2 = 1 \\ \left(\frac{1+\sqrt{5}}{2}\right) C_1 + \left(\frac{1-\sqrt{5}}{2}\right) C_2 = 1 \end{cases}$$

Dalla prima equazione si ricava $C_2 = 1 - C_1$, e sostituendo nella seconda e svolgendo i calcoli si trova

$$C_1 = \frac{1+\sqrt{5}}{2\sqrt{5}} = \frac{1}{\sqrt{5}} \frac{1+\sqrt{5}}{2}$$
$$C_2 = 1 - C_1 = -\frac{1}{\sqrt{5}} \frac{1-\sqrt{5}}{2}$$

Abbiamo dunque

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

Notiamo che nonostante la soluzione sia espressa in termini di frazioni e radici, per ogni valore di n la quantità f_n è un numero intero. Da questa espressione possiamo anche farci un'idea della grandezza dei numeri di Fibonacci. Valutando numericamente le radici dell'equazione caratteristica si ottiene

$$r_1 = \frac{1+\sqrt{5}}{2} = 1.618\ldots, \quad r_2 = \frac{1-\sqrt{5}}{2} = -0.618\ldots$$

e quindi la prima parentesi cresce come un esponenziale di base r_1 mentre la seconda tende a zero, poiché la base dell'esponenziale è minore di 1 in valore assoluto. Concludiamo che i numeri di Fibonacci crescono approssimativamente come $(1.618\ldots)^n$.

Risolviamo anche l'equazione del modello di crescita dei rami

$$\begin{cases} s_n - 3s_{n-1} + 2s_{n-2} = 0, \quad n > 1 \\ s_0 = 2, \quad s_1 = 3 \end{cases}$$

L'equazione caratteristica è $x^2 - 3x + 2 = 0$, ed ha le due soluzioni distinte

$$r_1 = 1, \quad r_2 = 2.$$

La soluzione generale è quindi

$$s_n = C_1 1^n + C_2 2^n$$

e il sistema da risolvere per determinare C_1 e C_2 è

$$\begin{cases} s_0 = C_1 1^0 + C_2 2^0 = C_1 + C_2 = 2 \\ s_1 = C_1 1^1 + C_2 2^1 = C_1 + 2C_2 = 3 \end{cases}$$

La soluzione è $C_1 = C_2 = 1$ e dunque l'equazione ricorsiva ha soluzione

$$s_n = 1 \cdot 1^n + 1 \cdot 2^n = 2^n + 1, \quad n \geq 0.$$

Esercizi

(1) Sia f_n il numero di funzioni da un insieme di n elementi ad un insieme di m elementi. Spiegare perché $f_n = mf_{n-1}$.

(2) Qual è l'ordine (se esiste) delle seguenti equazioni ricorsive?
   (a) $t_n = t_{n-1} + t_{n-2}$
   (b) $t_n = 3t_{n-1} + n$
   (c) $t_n = (t_{n-1} + t_{n-2})(t_{n-1} - t_{n-2})$
   (d) $t_n = n^2 t_{n-1} + n^3 t_{n-2}$
   (e) $t_n = t_{n-3} - t_{n-2}$
   (f) $t_n = t_{n-1}^2 + \sqrt{t_{n-3}}$

(3) Quali delle equazioni dell'esercizio (2) sono lineari?

(4) Quali delle equazioni dell'esercizio (2) sono omogenee?

(5) Quali delle equazioni dell'esercizio (2) sono a coefficienti costanti?

(6) Risolvere le seguenti equazioni
   (a) $a_n = \frac{1}{2}a_{n-1}$, $a_0 = 4$
   (b) $a_n = 4a_{n-1}$, $a_1 = 2$
   (c) $a_n = -2a_{n-1}$, $a_0 = 1$
   (d) $a_n = 5a_{n-1}$, $a_0 = 0$
   (e) $a_n = n^2 a_{n-1}$, $a_2 = 3$
   (f) $a_n = (n-1)a_{n-1}$, $a_1 = 1$

(7) Una successione $\{a_n\}$ verifica $a_n = \frac{1}{2}a_{n-1}$ per ogni $n \geq 0$ e sappiamo che $a_6 = \frac{3}{2}$. Quanto vale a_0?

(8) Un gruppo di $2n$ persone deve essere diviso in coppie per partecipare ad una gara di scacchi. Trovare una equazione ricorsiva per il numero d_n dei modi di suddividere $2n$ persone in coppie.

(9) Stesso problema di prima, ma dividendo $4n$ persone in gruppi di quattro, per un torneo di bridge.

Sia $C(n,k) = \binom{n}{k}$ il coefficiente binomiale che esprime il numero di sottoinsiemi di k elementi di un insieme di n elementi. La relazione

$$C(n,k) = C(n-1,k-1) + C(n-1,k)$$

è una equazione ricorsiva in due variabili, n e k. Nei due esercizi seguenti si chiede di determinare (ma non di risolvere) relazioni ricorsive in due variabili simili a questa.

(10) Trovare una equazione ricorsiva in due variabili per il numero $F(n,m)$ di funzioni suriettive da un insieme di n elementi ad un insieme di m elementi. (deve essere $n \geq m$, $m \geq 1$).

(11) Trovare una equazione ricorsiva in due variabili per il numero $G(n,k)$ di sottoinsiemi di k elementi dell'insieme $\{1,2,\ldots,n\}$ che non contengono numeri consecutivi. (deve essere $1 \leq k \leq (n+1)/2$, $n \geq 3$).

(12) Risolvere le seguenti equazioni
   (a) $s_n = 2s_{n-1} + 2^n$, $s_0 = 3$;
   (b) $s_n = 2s_{n-1} + n2^n$, $s_3 = 4$;
   (c) $s_n = 3s_{n-1} + 1$, $s_0 = 2$;
   (d) $s_n = 3s_{n-1} + \dfrac{1}{3^n}$, $s_0 = \dfrac{1}{3}$.

(13) Risolvere le seguenti equazioni del tipo *divide et impera*
   (a) $a_n = a_{n/4} + 3$, $a_1 = 1$;
   (b) $a_n = 3a_{n/4} + n$, $a_1 = 1$;
   (c) $a_n = 2a_{n/3} + 1$, $a_1 = 1$.

(14) Calcolare la potenza n-esima del numero intero i facendo moltiplicazioni successive per i richiede $n-1$ moltiplicazioni. Supponendo che $n = 2^k$, descrivere un algoritmo del tipo *divide et impera* tale che, se s_n è il numero di moltiplicazioni necessarie per calcolare la potenza n-esima, si abbia

$$s_n = s_{n/2} + 1.$$

Risolvere poi questa ricorrenza per determinare il numero di moltiplicazioni richieste da questo metodo.

(15) Calcolare il prodotto di una lista di n numeri interi facendo moltiplicazioni successive richiede $n-1$ moltiplicazioni. Supponendo che $n = 2^k$, descrivere un algoritmo del tipo *divide et impera* tale che, se m_n è il numero di moltiplicazioni necessarie per calcolare il prodotto degli n numeri nella lista, si abbia

$$m_n = 2m_{n/2} + 1.$$

Risolvere poi questa ricorrenza per determinare il numero di moltiplicazioni richieste da questo metodo.

Capitolo 4

Sistemi di equazioni lineari e matrici

1. Introduzione ai sistemi di equazioni lineari

Cominciamo con l'introdurre un po' di terminologia e discutere un metodo per la risoluzione dei sistemi lineari.

Una retta nel piano xy può essere rappresentata algebricamente con un'equazione della forma

$$ax + by = c.$$

Un'equazione di questo tipo viene detta equazione lineare nelle incognite x e y. Più in generale, possiamo dare la definizione seguente.

Definizione 1.1. Una **equazione lineare** nelle n incognite x_1, x_2, ..., x_n è una equazione che può essere espressa nella forma

$$a_1x_1 + a_2x_2 + \ldots a_nx_n = b,$$

dove a_1, a_2, ..., a_n, b sono delle costanti reali.

Esempio 1.2. Le seguenti sono equazioni lineari:

$$x + 3y = 7,$$
$$x_1 - 2x_2 - 3x_3 + x_4 = 7,$$
$$y = \frac{1}{2}x + 3z + 1,$$
$$x_1 + x_2 + \ldots x_n = 1.$$

Osserviamo che una equazione lineare non contiene potenze o radici delle incognite. Tutte le incognite appaiono a primo grado e non sono argomenti di

funzioni trigonometriche, logaritmiche o esponenziali. Le seguenti equazioni non sono lineari:

$$x + 3y^2 = 7,$$
$$y - \operatorname{sen} x = 0,$$
$$\sqrt{x_1} + 2x_2 + x_3 = 1.$$

Definizione 1.3. Una **soluzione** dell'equazione lineare $a_1x_1+\cdots+a_nx_n = b$ è una sequenza di n numeri $s_1, \ldots, s_n$ tale che l'equazione sia soddisfatta sostituendo $x_1 = s_1, \ldots, x_n = s_n$. L'insieme di tutte le soluzioni dell'equazione viene detto **insieme delle soluzioni**.

Esempio 1.4. Trovare l'insieme delle soluzioni delle seguenti equazioni:

$$\text{(i) } 4x - 2y = 1, \qquad \text{(ii) } x_1 - 4x_2 + 7x_3 = 5.$$

Soluzione. Per trovare le soluzioni della (i), possiamo assegnare un valore arbitrario ad x e ricavare la y, oppure possiamo assegnare un valore arbitrario ad y e ricavare la x. Seguiamo il primo procedimento: assegniamo ad x il valore arbitrario t, e otteniamo

$$x = t, \qquad y = 2t - \frac{1}{2}.$$

Queste formule descrivono l'insieme delle soluzioni in termini di un parametro arbitrario t. Soluzioni particolari possono essere ottenute assegnando valori particolari a t: per esempio, $t = 1$ dà la soluzione $x = 1$, $y = 3/2$, mentre $t = 1/2$ dà la soluzione $x = 1/2$, $y = 1/2$.

Se seguiamo il secondo procedimento e assegniamo ad y il valore arbitrario t, otteniamo

$$x = \frac{1}{2}t + \frac{1}{4}, \qquad y = t.$$

Sebbene queste formule siano diverse dalle precedenti, al variare di t esse descrivono lo stesso insieme di soluzioni. Per esempio, le formule precedenti davano la soluzione $x = 1$, $y = 3/2$ per il valore $t = 1$, mentre ora occorre considerare $t = 3/2$ per ottenere la stessa soluzione.

Per risolvere la (ii), possiamo assegnare valori arbitrari a due delle incognite e ricavare la terza. Per esempio, otteniamo

$$x_1 = 5 + 4s - 7t, \qquad x_2 = s, \qquad x_3 = t.$$

Definizione 1.5. Un insieme di equazioni lineari nelle incognite $x_1, \ldots, x_n$ viene detto **sistema di equazioni lineari** o semplicemente **sistema lineare**. Una sequenza di numeri $s_1, \ldots, s_n$ viene detta **soluzione** del sistema se $x_1 = s_1, \ldots, x_n = s_n$ è una soluzione di ogni equazione del sistema.

Per esempio, $x_1 = 1$, $x_2 = 2$, $x_3 = -1$ è una soluzione del sistema

$$\begin{cases} 4x_1 - x_2 + 3x_3 = -1 \\ 3x_1 + x_2 + 9x_3 = -4 \end{cases}$$

poiché i valori soddisfano entrambe le equazioni. Invece, $x_1 = 1$, $x_2 = 8$, $x_3 = 1$ non è una soluzione poiché questi valori soddisfano la prima equazione ma non la seconda.

Non tutti i sistemi lineari hanno soluzioni. Per esempio, il sistema

$$\begin{cases} x + y = 4 \\ 2x + 2y = 6 \end{cases}$$

non ha soluzione: questo è evidente se si moltiplica la prima equazione per 2.

Un sistema che non ha soluzione viene detto **incompatibile**. Se invece un sistema ammette soluzioni viene detto **compatibile**. Per comprendere le varie possibilità che si presentano risolvendo i sistemi lineari, consideriamo il caso del generico sistema di due equazioni in due incognite:

$$\begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases}$$

I grafici di queste equazioni sono rette nel piano, siano esse l_1 e l_2. Poiché un punto (x, y) appartiene ad una retta se e solo se i numeri x e y sono una soluzione dell'equazione corrispondente, le soluzioni del sistema corrispondono ai punti di intersezione delle due rette. Vi sono tre possibilità:

(1) le rette l_1 e l_2 sono parallele: allora non vi sono intersezioni e quindi il sistema è incompatibile;

(2) le rette l_1 e l_2 si incontrano in un punto: allora il sistema ha esattamente una soluzione;

(3) le rette l_1 e l_2 coincidono: allora il sistema ha infinite soluzioni.

Sebbene abbiamo considerato un caso semplice, vedremo in seguito che queste possibilità sono quelle che si verificano in generale:

> *ogni sistema di equazioni lineari o non ha soluzioni o ne ha esattamente una oppure ne ha infinite.*

Un sistema arbitrario di m equazioni in n incognite si scrive

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots a_{2n}x_n = b_2 \\ \vdots \quad\quad \vdots \quad\quad\quad \vdots \quad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots a_{mn}x_n = b_m \end{cases}$$

dove $x_1, \dots, x_n$ sono le incognite e tutte le a e b sono costanti.

Il doppio indice che si usa per i coefficienti delle equazioni del sistema serve ad identificare la posizione del coefficiente: per esempio, a_{23} si trova nella seconda equazione ed è il coefficiente dell'incognita x_3.

Un sistema di equazioni può essere rappresentato in forma simbolica da una tabella rettangolare di numeri nel modo seguente:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix}$$

e cioè scriviamo solo i coefficienti e i termini noti delle equazioni, sottointendendo i simboli $+$, $=$ e le incognite. Questa tabella viene detta **matrice completa** del sistema. Per esempio, la matrice completa del sistema

$$\begin{cases} x_1 & & & + & 2x_3 & = & 9 \\ 2x_1 & + & 4x_2 & - & 3x_3 & = & 1 \\ 3x_1 & + & 6x_2 & - & 5x_3 & = & 0 \end{cases}$$

è

$$\begin{bmatrix} 1 & 0 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{bmatrix}$$

Osserviamo che quando si scrive la matrice completa di un sistema, i coefficienti delle incognite vanno scritti nello stesso ordine in cui compaiono nel sistema. Inoltre, se una incognita non compare in una equazione, nel posto corrispondente si scrive 0, poiché non comparire in un'equazione è come comparire con coefficiente 0.

Il metodo che si usa per risolvere un sistema di equazioni è di sostituire il sistema dato con un altro, che abbia lo stesso insieme di soluzioni, ma che sia più facile da risolvere. Il nuovo sistema si ottiene in una serie di passi, ognuno dei quali consiste in una delle seguenti operazioni:

(1) Moltiplicare una equazione per una costante non nulla.

(2) Scambiare due equazioni fra loro.

(3) Aggiungere un multiplo di una equazione ad un'altra.

Poiché le righe della matrice completa corrispondono alle equazioni del sistema, le tre operazioni corrispondono alle tre operazioni seguenti sulle righe della matrice associata:

(1) Moltiplicare una riga per una costante non nulla.

(2) Scambiare due righe fra loro.

(3) Aggiungere un multiplo di una riga ad un'altra.

Queste sono dette **operazioni elementari (fra righe)**. È chiaro che se si eseguono le operazioni (1) e (2), l'insieme delle soluzioni non cambia. Dimostriamo che questo vale anche per l'operazione (3). Poiché in una operazione di tipo (3) intervengono solo due equazioni, possiamo considerare il sistema seguente:

$$\begin{array}{lccccccc} R_1: & a_{11}x_1 & + & a_{12}x_2 & + & \dots & a_{1n}x_n & = & b_1 \\ R_2: & a_{21}x_1 & + & a_{22}x_2 & + & \dots & a_{2n}x_n & = & b_2 \end{array}$$

Se c è un numero reale, il sistema che si ottiene aggiungendo c volte la riga R_1 alla riga R_2 è:

$$\begin{array}{rccccccc} R_1: & a_{11}x_1 & + & \dots & a_{1n}x_n & = & b_1 \\ R_2+cR_1: & (a_{21}+ca_{11})x_1 & + & \dots & (a_{2n}+ca_{1n})x_n & = & b_2+cb_1 \end{array}$$

Se supponiamo che $s_1, \dots, s_n$ sia una soluzione del primo sistema, le

$$\begin{array}{ccccccccc} a_{11}s_1 & + & a_{12}s_2 & + & \dots & a_{1n}s_n & = & b_1 \\ a_{21}s_1 & + & a_{22}s_2 & + & \dots & a_{2n}s_n & = & b_2 \end{array}$$

sono uguaglianze di numeri. È quindi immediato osservare che $s_1, \dots, s_n$ è anche una soluzione del secondo sistema.

Viceversa, sia $w_1, \dots, w_n$ una soluzione del secondo sistema. Allora le

$$\begin{array}{ccccccccc} a_{11}w_1 & + & a_{12}w_2 & + & \dots & a_{1n}w_n & = & b_1 \\ (a_{21}+ca_{11})w_1 & + & (a_{22}+ca_{12})w_2 & + & \dots & (a_{2n}+ca_{1n})w_n & = & b_2+cb_1 \end{array}$$

sono uguaglianze di numeri. Osserviamo quindi che $w_1, \dots, w_n$ è una soluzione di R_1 e inoltre, moltiplicando per c la prima uguaglianza e sottraendola dalla seconda si ottiene che $w_1, \dots, w_n$ è una soluzione anche di R_2.

2. Il metodo di eliminazione di Gauss

In questo paragrafo svilupperemo un metodo sistematico per la risoluzione dei sistemi lineari; questo metodo si base sull'idea di modificare la matrice completa associata al sistema, mediante operazioni elementari, fino ad ottenere un sistema semplice abbastastanza da poter essere risolto direttamente.

Definizione 2.1. Una matrice si dice **ridotta per righe** se verifica le seguenti proprietà:

(1) Se vi sono righe interamente nulle, queste sono raggruppate al fondo della matrice.

(2) Se una riga non è interamente nulla, allora il primo numero non nullo (da sinistra) è 1. (Questo viene detto **1 iniziale**).

(3) Se due righe consecutive non sono interamente nulle, l'1 iniziale della seconda riga appare più a destra di quello della prima riga.

Per esempio, le matrici seguenti sono ridotte per righe:

$$A = \begin{bmatrix} 1 & 4 & 7 & 3 \\ 0 & 1 & 6 & 4 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 8 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

È immediato verificare che, in una matrice ridotta per righe, sotto un 1 iniziale vi sono solo zeri.

Il sistema lineare associato ad una matrice ridotta per righe si risolve in modo elementare; per convincercene, risolviamo i sistemi associati alle matrici precedenti.

Il sistema associato alla matrice A è:

$$\begin{cases} x_1 + 4x_2 + 7x_3 = 3 \\ x_2 + 6x_3 = 4 \\ x_3 = 0 \end{cases}$$

La soluzione è evidente: si comincia a risolvere dall'ultima equazione per ottenere $x_3 = 0$; sostituendo nella penultima si ha $x_2 = 4$ e quindi, sostituendo nella prima, $x_1 = -13$.

Il sistema associato alla matrice B è:

$$\begin{cases} x_1 + x_2 = 0 \\ x_2 = 0 \end{cases}$$

L'ultima riga, interamente nulla, corrisponde ad un'equazione della forma $0 = 0$ e cioè ad una identità sempre soddisfatta, quindi non si deve riportare nel sistema. Anche in questo caso la soluzione è evidente: dalla seconda equazione si ha $x_2 = 0$ e sostituendo nella prima si ottiene $x_1 = 0$.

Il sistema associato alla matrice C è:

$$\begin{cases} x_2 + 2x_3 + x_4 = 0 \\ x_3 + 8x_4 = 2 \\ 0 = 1 \end{cases}$$

Poiché la prima colonna è nulla, la prima incognita che compare è x_2 e non x_1. In questo sistema la terza equazione non può essere verificata per nessun valore delle incognite e quindi il sistema è incompatibile.

Vediamo ancora un esempio. Sia D la matrice:

$$D = \begin{bmatrix} 1 & 0 & 2 & 3 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Il sistema associato è:

$$\begin{cases} x_1 + & + 2x_3 + 3x_4 = 0 \\ & x_2 + x_3 + 2x_4 = 0 \\ & x_3 + x_4 = 0 \end{cases}$$

In questo caso, risolvendo l'ultima equazione si ottiene $x_3 = -x_4$, e cioè questa equazione ha infinite soluzioni che dipendono da un parametro libero, x_4. Sostituendo nella penultima si ottiene $x_2 = -x_3 - 2x_4 = -x_4$ e infine $x_1 = -2x_3 - 3x_4 = -x_4$. Quindi questo sistema ammette infinite soluzioni, che si ottengono facendo variare il parametro libero x_4.

È ormai chiaro che un sistema la cui matrice sia ridotta per righe è sempre facilmente risolubile. Il **metodo di eliminazione di Gauss**, che ora presenteremo, è un algoritmo che permette di passare da una matrice qualunque ad una ridotta per righe mediante una sequenza di operazioni elementari fra righe, dunque senza cambiare le soluzioni del sistema associato.

Seguiamo la successione dei passi dell'algoritmo in un esempio. Consideriamo la matrice

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Passo 1. Determinare la prima colonna da sinistra non nulla. Se non vi sono colonne non nulle, l'algoritmo termina.

In questo caso, è la prima colonna della matrice.

Passo 2. Se necessario, scambiare la prima riga con un'altra riga, in modo che il primo elemento della prima riga sia diverso da zero.

$$\begin{bmatrix} 2 & 4 & -10 & 6 & 12 & 28 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Abbiamo scambiato la prima riga con la seconda.

Passo 3. Dividere la prima riga per il suo primo elemento, in modo da ottenere un 1 iniziale nella prima riga.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Abbiamo diviso la prima riga per 2.

Passo 4. Ad ogni riga tranne la prima sommiamo un multiplo della prima riga, in modo da ottenere degli zeri nella colonna che contiene l'1 iniziale.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$

*La seconda riga è invariata, perché sotto l'*1 *iniziale della prima riga vi era già uno zero. Invece abbiamo sottratto* 2 *volte la prima riga alla terza, in modo da ottenere uno zero nella prima colonna.*

Passo 5. A questo punto ripetiamo i passi 1–4 usando la sottomatrice che si ottiene coprendo la prima riga. Si continua in questo modo fino a che non vi sono più righe da considerare.

Nel nostro esempio, coprendo la prima riga si ha che la prima colonna non nulla è la terza. Dividiamo la seconda riga per -2 *in modo da ottenere un* 1 *iniziale*

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -7/2 & -6 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$

e poi sottraiamo 5 *volte la seconda riga alla terza per avere degli zeri sotto l'*1 *iniziale della seconda riga:*

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -7/2 & -6 \\ 0 & 0 & 0 & 0 & 1/2 & 1 \end{bmatrix}$$

ora dobbiamo considerare la sottomatrice formata dalla sola terza riga, la prima colonna non nulla è la quinta, e moltiplichiamo per 2 *la terza riga per avere un* 1 *iniziale:*

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -7/2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

A questo punto la matrice è ridotta per righe.

Diamo ora una regola generale per risolvere i sistemi di equazioni le cui matrici associate siano ridotte per righe. Per prima cosa consideriamo l'equazione associata all'ultima riga non nulla: se la riga è

$$0 \quad 0 \quad 0 \quad \dots \quad 0 \quad 1$$

allora l'equazione corrispondente è $0 = 1$ e cioè il sistema è incompatibile. Altrimenti il sistema ha soluzioni. Procediamo come segue. L'incognita x_i è detta **incognita iniziale** se la colonna i contiene un 1 iniziale, altrimenti viene detta incognita non iniziale. Risolviamo il sistema lasciando, in ogni equazione, le incognite iniziali a primo membro e portando a secondo

membro tutto il resto. Partendo dall'ultima equazione, sostituiamo ogni equazione in tutte quelle che si trovano sopra ad essa. A questo punto le incognite non iniziali, che si trovano a secondo membro, sono detti i **parametri liberi**. Da essi dipendono le soluzioni, e il sistema è completamente risolto. Si ha la formula:

> numero dei parametri liberi = numero delle incognite − numero delle righe non nulle della matrice ridotta.

Infatti ogni riga non nulla determina esattamente una incognita iniziale.

Notiamo che un sistema compatibile ha soluzione unica se e solo se non ci sono parametri liberi, ossia il numero di righe non nulle della matrice ridotta è pari al numero di incognite.

Per esempio, se il sistema ridotto è

$$\begin{cases} x_1 + 3x_2 - 2x_3 + 2x_5 + 2x_6 = 0 \\ \quad + x_3 + 2x_4 + 3x_6 = 1 \\ \quad\quad\quad\quad\quad\quad x_6 = 1 \end{cases}$$

risolvendo rispetto alle incognite iniziali si ottiene

$$\begin{aligned} x_1 &= -3x_2 + 2x_3 - 2x_5 - 2x_6 \\ x_3 &= 1 - 2x_4 - 3x_6 \\ x_6 &= 1 \end{aligned}$$

Effettuando la sostituzione si ottiene

$$\begin{aligned} x_1 &= -6 - 3x_2 - 4x_4 - 2x_5 \\ x_3 &= -2 - 2x_4 \\ x_6 &= 1 \end{aligned}$$

I parametri liberi sono quindi x_2, x_4 e x_5, e il sistema è completamente risolto.

3. Operazioni fra matrici

Dato un sistema lineare, abbiamo associato ad esso una **matrice**, cioè una tabella di numeri che rappresenta l'informazione data dai coefficienti e dai termini noti del sistema. Abbiamo formulato il metodo di riduzione compiendo operazioni sulla matrice, poiché è più semplice operare su tabelle di numeri che su equazioni contenenti lettere (le incognite). Programmare il metodo di riduzione è abbastanza semplice, mentre scrivere un programma che gestisce il calcolo letterale è ben più complicato.

In molte altre parti della matematica si trova lo stesso fenomeno: l'informazione può essere organizzata in una tabella di numeri, e le operazioni

che si compiono sono spesso le stesse. Conviene quindi studiare in generale le matrici e le operazioni algebriche che si possono definire fra di esse. Fra le molte operazioni possibili, nella pratica se ne incontrano tre, che sono quelle che definiremo ora.

Definizione 3.1. Una **matrice di tipo** (m, n) è una tabella rettangolare di numeri, con m righe e n colonne.

Le matrici si denotano in genere con lettere maiuscole, A, B, ...; se A è una matrice di tipo (m, n), si dice anche che A ha dimensione (m, n), oppure che A è una matrice $m \times n$ (che si legge: A è una matrice m per n). Se le dimensioni sono uguali, $m = n$, la matrice si dice **quadrata di ordine** n. Gli elementi di una matrice si indicano in genere con una lettera minuscola uguale al nome della matrice, e con due indici, per indicare la posizione di riga e di colonna dell'elemento nella matrice. Una generica matrice A di tipo (m, n) si scrive dunque

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Se non si vuole scrivere tutta la tabella, ma si vuole indicare che A è una matrice, e dare un nome ai suoi elementi, si usa anche la notazione

$$A = (a_{ij}), \quad 1 \le i \le m, 1 \le j \le n$$

e spesso si omettono le diseguaglianze, se il tipo della matrice è chiaro dal contesto. Le parentesi intorno alla tabella sono tonde o anche quadre (come abbiamo fatto nei paragrafi precedenti), ma mai graffe. Invece nella notazione $A = (a_{ij})$, le parentesi sono sempre tonde.

Definizione 3.2. Siano $A = (a_{ij})$ e $B = (b_{ij})$ due matrici di tipo (m, n). La **matrice somma** di A e B è la matrice $C = (c_{ij})$ di tipo (m, n) i cui elementi sono definiti da

$$c_{ij} = a_{ij} + b_{ij}, \quad 1 \le i \le m, \ 1 \le j \le n.$$

Per indicare che C è la somma di A e B si usa la notazione $A + B = C$.

Notiamo che il simbolo + che compare nelle formule $c_{ij} = a_{ij} + b_{ij}$ è il simbolo della usuale somma fra numeri, mentre nella formula $A + B = C$ rappresenta la nuova operazione appena definita, la somma di matrici.

Notiamo anche che è possibile sommare fra loro solo matrici dello stesso tipo. Non ha senso, per esempio, sommare una matrice 2×3 con una matrice 3×5.

Il significato dell'operazione di somma fra matrici è abbastanza semplice: si sommano fra loro gli elementi corrispondenti (cioè nella stessa posizione) delle matrici A e B, e si mette la somma degli elementi nella stessa posizione. Per esempio

$$\begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & -1 \end{pmatrix} = \begin{pmatrix} 2+1 & 3+2 & 1+0 \\ -1+2 & 0+2 & 4-1 \end{pmatrix} = \begin{pmatrix} 3 & 5 & 1 \\ 1 & 2 & 3 \end{pmatrix}$$

La somma fra matrici gode di proprietà simili alla somma fra numeri. Sia $\mathcal{M}(m,n)$ l'insieme di tutte le matrici di tipo (m,n). Fra queste c'è una matrice particolare, detta **matrice nulla**, e indicata con $\mathbf{0}_{m,n}$, che è la matrice i cui elementi sono tutti nulli. Per le matrici di $\mathcal{M}(m,n)$ valgono le seguenti proprietà:

(1) La somma è associativa: $(A+B)+C = A+(B+C)$, per ogni A, B, $C \in \mathcal{M}(m,n)$.

(2) La somma è commutativa: $A+B = B+A$, per ogni A, $B \in \mathcal{M}(m,n)$.

(3) $\mathbf{0}_{m,n}$ è l'elemento neutro: $A+\mathbf{0}_{m,n} = A$, per ogni $A \in \mathcal{M}(m,n)$.

(4) Esiste l'opposto: per ogni $A \in \mathcal{M}(m,n)$ esiste $B \in \mathcal{M}(m,n)$ tale che $A+B = \mathbf{0}_{m,n}$

Le prime tre proprietà si dimostrano direttamente dalla definizione di somma fra matrici, notando che la somma di numeri gode delle stesse tre proprietà. Per la (4), basta osservare che la matrice opposta di A si ottiene prendendo la matrice che ha come elementi gli opposti degli elementi di A.

Passiamo ora alla moltiplicazione. Vi sono due tipi di moltiplicazione che coinvolgono le matrici. Nel primo, si moltiplicano **numeri** con **matrici**.

Definizione 3.3. Sia $A = (a_{ij})$ una matrice di tipo (m,n) e sia $c \in \mathbb{R}$ un numero. Il **prodotto di** c **per** A è la matrice $B = (b_{ij})$ di tipo (m,n) i cui elementi sono definiti da

$$b_{ij} = c \cdot a_{ij}, \quad 1 \le i \le m,\ 1 \le j \le n,$$

e si scrive $B = cA$.

Notiamo che in questo caso i fattori della moltiplicazione sono elementi di genere diverso: il primo è un numero, il secondo è una matrice. Il risultato è una matrice, dello stesso tipo della matrice che si moltiplica. Notiamo anche che si scrive sempre $B = cA$ e mai $B = Ac$, anche se le operazioni che definiscono gli elementi di B sono commutative.

Anche in questo caso, il significato dell'operazione è semplice: tutti gli elementi di A vengono moltiplicati per c e dunque si può dire che cA è un **multiplo** della matrice A. Questo spiega perché si scrive cA e non Ac: il triplo di A si scrive $3A$ e non $A3$.

Siamo arrivati alla terza operazione: il prodotto fra matrici. Questa è forse la più importante di tutte ed è certamente quella che ha la definizione più complicata. Inoltre vedremo che è quella che ha le proprietà più sorprendenti.

Definizione 3.4. Sia $A = (a_{ij})$ di tipo (m, n) e $B = (b_{ij})$ di tipo (n, p). Il **prodotto di A per B** è la matrice $C = (c_{ij})$ di tipo (m, p) i cui elementi sono definiti da

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \quad 1 \leq i \leq m,\ 1 \leq j \leq n,$$

e si scrive $C = AB$.

Perché sia possibile moltiplicare A e B è necessario che il numero di *colonne* di A sia uguale al numero di *righe* di B. Dunque è possibile che sia definito il prodotto AB ma non il prodotto BA. Per esempio, se A è di tipo $(2, 3)$ e B è di tipo $(3, 4)$, allora AB esiste, ed è di tipo $(2, 4)$, ma BA non esiste.

Anche quando entrambi i prodotti sono definiti, non è detto che siano uguali. Per esempio, se A è di tipo $(2, 3)$ e B è di tipo $(3, 2)$, allora AB esiste ed è di tipo $(2, 2)$, mentre BA è di tipo $(3, 3)$ e dunque sono certamente diverse.

C'è un caso in cui sia AB che BA sono entrambe definite ed è possibile che siano uguali: quando A e B sono quadrate dello stesso ordine. In questo caso AB e BA sono ancora quadrate dello stesso ordine, ma non è detto che siano uguali. Per esempio, se

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}$$

allora

$$AB = \begin{pmatrix} 4 & 6 \\ 4 & 0 \end{pmatrix}, \quad BA = \begin{pmatrix} 2 & 4 \\ 7 & 2 \end{pmatrix}$$

e quindi sono diverse. Può però capitare che i due prodotti siano uguali. Se

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 2 \\ 2 & 1 \end{pmatrix}$$

allora

$$AB = \begin{pmatrix} 6 & 4 \\ 4 & 4 \end{pmatrix}, \quad BA = \begin{pmatrix} 6 & 4 \\ 4 & 4 \end{pmatrix}$$

e dunque $AB = BA$, in questo caso. In generale quindi la moltiplicazione fra matrici non è commutativa.

L'interpretazione del prodotto di matrici è più complicata di quella della somma, ma è molto utile. Il prodotto per matrici viene anche detto prodotto **righe per colonne** poiché nella moltiplicazione $AB = C$ si usano le *righe* della matrice A e le *colonne* di B. Infatti, leggendo la formula che esprime il prodotto, si vede che l'elemento c_{ij} del prodotto, che sta nella riga i e nella colonna j della matrice C, si ottiene moltiplicando gli elementi della riga i-esima della matrice A per gli elementi corrispondenti della colonna j-esima della matrice B e sommando tutti i prodotti.

Abbiamo già incontrato questo tipo di prodotto, anche se non lo abbiamo messo in evidenza. Sia infatti

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \dots & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & a_{mn}x_n & = & b_m \end{cases}$$

un sistema di m equazioni in n incognite, e poniamo

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

La matrice A è di tipo (m, n), la matrice X è di tipo $(n, 1)$ e la matrice B è di tipo $(m, 1)$. Eseguendo la moltiplicazione fra matrici (esercizio!) si vede che il sistema di equazioni corrisponde alla uguaglianza di matrici

$$AX = B$$

Non possiamo non notare la similitudine fra questa scrittura e l'equazione $ax = b$ di primo grado in una incognita. La soluzione di questa equazione è $x = \dfrac{b}{a} = a^{-1}b$. Ci si può chiedere se si possa risolvere un sistema lineare come si risolve una equazione, e cioè "dividendo" per A. La risposta è sì, se si interpreta correttamente il significato di "dividere" per una matrice, e però questo sarà possibile solo per alcune matrici.

Analizzando le proprietà del prodotto di matrici come abbiamo fatto per la somma, si può dimostrare che il prodotto è associativo, cioè $(AB)C = A(BC)$, ma abbiamo visto che in generale non è commutativo. Esiste un elemento neutro, e cioè l'analogo del numero 1 per la moltiplicazione? Introduciamo la **matrice unità**, anche detta **matrice identità**, I_n, definita come segue: $I_n = (a_{ij})$ è una matrice quadrata di ordine n i cui

elementi sono dati da

$$a_{ii} = 1, \quad a_{ij} = 0 \text{ se } i \neq j.$$

Per esempio,

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Se A è una matrice di tipo (m, n), usando la definizione di prodotto di matrici si verifica che

$$I_m \cdot A = A \cdot I_n = A$$

e cioè le matrici I_n si comportano come il numero 1 rispetto alla moltiplicazione. Notiamo però che a sinistra e a destra dobbiamo usare un "1" diverso, altrimento non si può moltiplicare. C'è un caso in cui a sinistra e a destra abbiamo la stessa matrice unità: è quando la matrice A è quadrata di ordine n. Si ha allora $I_n A = A I_n = A$, e cioè I_n si comporta come una vera unità per le matrici di $\mathcal{M}(n, n)$, quadrate di ordine n.

La divisione fra numeri è definita mediante la moltiplicazione: $a/b = c$ vuol dire che $bc = a$. Lo stesso è per il reciproco (o inverso) di un numero: $a^{-1} = 1/a = b$ vuol dire che $ab = 1$. Inoltre dividere è come moltiplicare per l'inverso: $a/b = a \cdot (1/b)$. Possiamo usare le stesse definizioni per le matrici, ma solo per le matrici quadrate. Infatti solo per queste abbiamo un analogo del numero 1.

Definizione 3.5. Sia A una matrice quadrata di ordine n. La **matrice inversa** di A, se esiste, è la matrice B quadrata di ordine n tale che

$$AB = BA = I_n$$

L'inversa della matrice A si indica con A^{-1}.

Poiché la moltiplicazione non è commutativa, dobbiamo chiedere che AB e BA siano entrambi uguali a I_n. Nella definizione si parla della matrice inversa mentre potrebbe capitare che, data la matrice A, ci sia più di una matrice inversa di A. In effetti questo non capita, e cioè la matrice inversa è unica. Possiamo dimostrarlo come conseguenza dell'associatività del prodotto fra matrici.

Teorema 3.6. *Sia A una matrice quadrata di ordine n. Se B, C sono tali che $AB = BA = AC = CA = I_n$ allora $B = C$, cioè la matrice inversa, se esiste, è unica.*

Dimostrazione.

$$AB = I_n \implies C(AB) = CI_n = C \implies (CA)B = C \implies B = I_nB = C$$

□

Notiamo che sia nella Definizione 3.5 che nel Teorema 3.6 c'è la cautela:"se esiste". In effetti, non tutti i numeri hanno inverso: per esempio, non esiste l'inverso di 0, ma questo è l'unico numero che non ha inverso. Invece è possibile che non esista la matrice inversa di una matrice non nulla. Studieremo questa questione, e anche il modo per calcolare la matrice inversa, quando esiste, nel paragrafo 5.

4. L'algoritmo di moltiplicazione di Strassen

In questo paragrafo analizzeremo l'operazione di moltiplicazione di matrici dal punto di vista della complessità del calcolo, e cioè cercheremo di capire quale sia il metodo più efficiente per moltiplicare due matrici, in termini del numero di operazioni fra numeri necessarie per la moltiplicazione delle matrici. Ci occuperemo solo di matrici quadrate.

Siano dunque $A = (a_{ij})$ e $B = (b_{ij})$ due matrici quadrate di ordine n e il problema è quello di calcolare il prodotto $C = AB$, prodotto dato dalla definizione 3.4. Se usiamo le formule scritte allora, è semplice contare il numero di operazioni necessarie: dobbiamo calcolare tutti gli elementi c_{ij} della matrice C e ce ne sono n^2. Per ogni elemento facciamo n moltiplicazioni fra gli elementi della riga di A e la colonna di B e poi facciamo $n-1$ somme per sommare insieme gli n prodotti ottenuti. In totale ci sono dunque

$$\begin{array}{ll} n^2 \cdot n = n^3 & \text{moltiplicazioni} \\ n^2(n-1) = n^3 - n^2 & \text{addizioni} \end{array}$$

Per esempio, per moltiplicare due matrici di tipo $(2,2)$ ci vogliono 8 moltiplicazioni e 4 addizioni. Nel 1969, Volker Strassen ha scoperto un metodo per moltiplicare due matrici di tipo $(2,2)$ che usa solo 7 moltiplicazioni (ma 18 addizioni). La cosa può sembrare non molto interessante, ma il metodo di Strassen può essere usato ricorsivamente per matrici di ordine 4, 8, 16 e così via (una semplice generalizzazione consente di usare il metodo per tutte le matrici, non solo quelle con ordine potenza di 2), portando ad un algoritmo di moltiplicazione che necessita di un numero totale di operazioni, somme e moltiplicazioni, dell'ordine di $7 \cdot n^{2.807\ldots}$ (l'esponente è esattamente $\log_2(7)$), che è inferiore, per n grande, all'algoritmo precedente, che invece usa circa $2n^3$ operazioni.

Per prima cosa, descriviamo il metodo di Strassen per matrici di tipo $(2,2)$. Poniamo

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

e sia

$$C = AB = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

Calcoliamo le espressioni:

$$\begin{aligned} d_1 &= (a_{11} + a_{22}) \cdot (b_{11} + b_{22}) \\ d_2 &= (a_{21} + a_{22}) \cdot b_{11} \\ d_3 &= a_{11} \cdot (b_{12} - b_{22}) \\ d_4 &= a_{22} \cdot (-b_{11} + b_{21}) \\ d_5 &= (a_{11} + a_{12}) \cdot b_{22} \\ d_6 &= (-a_{11} + a_{21}) \cdot (b_{11} + b_{12}) \\ d_7 &= (a_{12} - a_{22}) \cdot (b_{21} + b_{22}) \end{aligned}$$

e poi abbiamo

$$\begin{aligned} c_{11} &= d_1 + d_4 - d_5 + d_7 \\ c_{12} &= d_3 + d_5 \\ c_{21} &= d_2 + d_4 \\ c_{22} &= d_1 + d_3 - d_2 + d_6 \end{aligned}$$

Per calcolare i d_i servono 7 moltiplicazioni e 10 addizioni, e poi vi sono altre 8 addizioni nel calcolo finale dei c_{ij}. In totale, 7 moltiplicazioni e 18 addizioni. Notiamo che le formule non sono per niente ovvie da trovare, ma soprattutto è sorprendente che si possano fare solo 7 moltiplicazioni invece di 8. Abbiamo però usato molte più addizioni (18 contro solo 4) e si può pensare che tutto sommato il metodo non sia efficiente.

L'importanza delle formule appena viste è però un'altra: se le osserviamo con attenzione, ci accorgiamo che in nessun punto viene usata la proprietà commutativa della moltiplicazione fra numeri. Infatti, le semplificazioni che portano al calcolo dei termini c_{ij} a partire dai termini d_i non sfruttano la proprietà commutativa della moltiplicazione. Possiamo pensare allora di generalizzare queste formule a matrici di ordine superiore, sostituendo matrici a numeri e utilizzando la somma e il prodotto fra matrici al posto delle operazioni fra numeri.

Consideriamo il caso di matrici di tipo $(4,4)$. Siano

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

dove le matrici A_{ij} e B_{ij} sono matrici di tipo $(2,2)$. È facile vedere che, se scriviamo

$$AB = C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

possiamo usare le stesse formule di prima per calcolare le sottomatrici C_{ij}, dove eseguiamo le somme fra matrici nel modo ovvio, cioè sommando gli elementi corrispondenti. In generale, possiamo ripetere la stessa procedura se A e B sono matrici di ordine $n = 2^k$, e le sottomatrici A_{ij} e B_{ij} hanno ordine $n/2 = 2^{k-1}$.

Poiché la procedura è ricorsiva, del tipo *divide et impera*, possiamo scrivere equazioni ricorsive che determinano il numero di moltiplicazioni e di addizioni. Poniamo m_n = numero di moltiplicazioni fra numeri necessario per moltiplicare due matrici di ordine n, e s_n = numero di addizioni fra numeri necessario per moltiplicare due matrici di ordine n.

L'equazione per m_n è:

$$\begin{cases} m_n = 7m_{n/2}, & (n \text{ pari e } n > 2) \\ m_2 = 7 \end{cases}$$

dove m_2 corrisponde alla moltiplicazione fra matrici di ordine 2. La spiegazione è immediata: usando l'algoritmo di Strassen, per moltiplicare due matrici di ordine n occorrono 7 moltiplicazioni fra matrici di ordine $n/2$. Con la sostituzione $n = 2^k$, che si usa per questo tipo di equazioni, e ponendo $u_k = m_n$, la ricorsione diventa

$$\begin{cases} u_k = 7u_{k-1}, & (k \geq 2) \\ u_1 = 7 \end{cases}$$

ed essendo questa una equazione del primo ordine omogenea, la soluzione è immediata ed è $u_k = 7^k$. Poniamo $\alpha = \log_2(7)$ e cioè $7 = 2^\alpha$. Poiché $k = \log_2(n)$, la soluzione in termini di m_n è:

$$m_n = 7^{\log_2(n)} = (2^\alpha)^{\log_2(n)} = 2^{\alpha \log_2(n)} = (2^{\log_2(n)})^\alpha = n^\alpha = n^{\log_2(7)}.$$

Poiché $\log_2(7) = 2.807\cdots < 3$, l'ordine di crescita dal numero delle moltiplicazioni dell'algoritmo di Strassen è inferiore all'ordine dell'algoritmo tradizionale.

Consideriamo ora il caso delle addizioni: in questo caso s_n rappresenta il numero di addizioni necessarie per *moltiplicare* due matrici con l'algoritmo

di Strassen. L'equazione per s_n è

$$\begin{cases} s_n = 7s_{n/2} + 18 \cdot \left(\frac{n}{2}\right)^2, & (n \text{ pari e } n > 2) \\ s_2 = 18 \end{cases}$$

Infatti per moltiplicare due matrici di ordine n eseguiamo 7 *moltiplicazioni* di matrici di ordine $n/2$ e in ognuna di queste moltiplicazioni occorrono $s_{n/2}$ addizioni. Inoltre, eseguiamo 18 addizioni di matrici di ordine $n/2$ e per addizionare due tali matrici occorrono $(n/2)^2$ addizioni, e cioè una addizione per ogni elemento delle matrici.

Ponendo come prima $v_k = s_n$, $n = 2^k$ la ricorsione diventa lineare del primo ordine

$$\begin{cases} v_k = 7v_{k-1} + 18 \cdot (2^{k-1})^2, & (k \geq 2) \\ v_1 = 18 \end{cases}$$

Notiamo che possiamo scrivere $(2^{k-1})^2 = (2^2)^{k-1} = 4^{k-1}$. L'equazione è del primo ordine, lineare, a coefficienti costanti, non omogenea e possiamo usare la formula del Teorema 2.1 del Capitolo 3, ottenendo (attenzione: $m = 1$)

$$\begin{aligned} v_k &= 7^{k-1}\left(18 + \sum_{i=1}^{k-1} 18 \cdot 4^i 7^{-i}\right) = 7^{k-1} \cdot 18\left(1 + \sum_{i=1}^{k-1}\left(\frac{4}{7}\right)^i\right) \\ &= 7^{k-1} \cdot 18 \sum_{i=0}^{k-1}\left(\frac{4}{7}\right)^i = 7^{k-1} \cdot 18\frac{1-(4/7)^k}{1-(4/7)} \\ &= 7^{k-1} \cdot 18 \cdot \frac{7}{3}\left(1 - \frac{4^k}{7^k}\right) = 6 \cdot 7^k - 6 \cdot 4^k \end{aligned}$$

Se come prima poniamo $\alpha = \log_2(7)$ e ricordiamo che $n = 2^k$ otteniamo

$$s_n = 6n^{\log_2(7)} - 6n^2$$

Il numero totale di operazioni compiute dall'algoritmo di Strassen per moltiplicare due matrici quadrate di ordine $n = 2^k$ è quindi

$$m_n + s_n = n^{\log_2(7)} + 6n^{\log_2(7)} - 6n^2 = 7n^{\log_2(7)} - 6n^2$$

Poiché l'algoritmo tradizionale impiega $2n^3 - n^2$ operazioni, vediamo che per n abbastanza grande l'algoritmo di Strassen è migliore. Quale è il primo numero n per cui conviene usare l'algoritmo di Strassen? Occorre risolvere la disequazione

$$7n^{\log_2(7)} - 6n^2 < 2n^3 - n^2$$

e la risposta è, almeno per numeri della forma $n = 2^k$, che il primo caso in cui l'algoritmo di Strassen è migliore è per $k = 10$, e cioè matrici quadrate di ordine $n = 2^{10} = 1024$. In questo caso l'algoritmo di Strassen esegue 1 971 035 287 operazioni, mentre l'algoritmo usuale ne esegue 2 146 435 072.

Sembra dunque che l'algoritmo di Strassen non abbia una importanza pratica, poiché è certamente raro dover usare matrici di tali dimensioni. Questo è vero, ma l'importanza della scoperta di Strassen è un'altra: per centinaia di anni nessuno aveva mai pensato alla possibilità di moltiplicare due matrici se non usando la formula tradizionale della definizione. L'algoritmo di Strassen pone una domanda: se c'è un altro modo, migliore di quello tradizionale, quale sarà il migliore di tutti? Dato un algoritmo per la moltiplicazione di due matrici quadrate, il numero di operazioni necessarie sarà dell'ordine di n^a, dove l'esponente a dipende dall'algoritmo usato. Per esempio, la formula della definizione ha $a = 3$, mentre l'algoritmo di Strassen ha $a = \log_2(7) = 2.807\ldots$ Naturalmente, minore è l'esponente a e migliore è l'algoritmo. Qual è il valore minimo possibile per a? È abbastanza chiaro che $a \geq 2$, cioè occorrono almeno n^2 operazioni per moltiplicare due matrici di ordine n: infatti, dobbiamo calcolare gli n^2 elementi della matrice risultato, e per ogni elemento dovremo fare almeno una operazione.

Al giorno d'oggi (33 anni dopo la scoperta di Strassen) l'algoritmo migliore conosciuto ha $a = 2.376\ldots$, già un buon miglioramento rispetto all'algoritmo di Strassen. Inoltre, il metodo usato per ottenere questo risultato (da D. Coppersmith e S. Winograd, 1990) sembra indicare che il valore minimo di a è in effetti 2, e cioè ci dovrebbe essere un algoritmo che compie circa n^2 operazioni per moltiplicare due matrici di ordine n ma questo algoritmo, se veramente esiste, non è stato ancora trovato.

5. Inversione di matrici

In questo paragrafo vedremo come il metodo di riduzione di Gauss che abbiamo visto nel §2 consenta di calcolare l'inversa di una matrice quadrata, quando esiste.

Il problema di determinare l'inversa di una matrice si può porre nel modo seguente: data una matrice quadrata di ordine n, $A = (a_{ij})$, determinare gli elementi della matrice $B = (b_{ij})$ tali che

$$AB = I_n$$

In questo problema, gli elementi a_{ij} sono noti mentre gli elementi b_{ij} sono le incognite. Abbiamo dunque n^2 incognite. Scriviamo l'equazione $AB = I_n$ più in dettaglio:

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \ldots & b_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

Effettuando la moltiplicazione di matrici, ci accorgiamo che le incognite $b_{11}, b_{21}, \ldots, b_{n1}$, cioè gli elementi della *prima colonna* di B, compaiono in un sistema che non contiene le altre incognite, e precisamente nel sistema

$$\begin{cases} a_{11}b_{11} & + & a_{12}b_{21} & + & \ldots & a_{1n}b_{n1} & = & 1 \\ a_{21}b_{11} & + & a_{22}b_{21} & + & \ldots & a_{2n}b_{n1} & = & 0 \\ \vdots & & \vdots & & & \vdots & & \vdots \\ a_{n1}b_{11} & + & a_{n2}b_{21} & + & \ldots & a_{nn}b_{n1} & = & 0 \end{cases}$$

Anche le incognite $b_{12}, b_{22}, \ldots, b_{n2}$, cioè gli elementi della *seconda colonna* di B, compaiono in un sistema che non contiene le altre incognite, e precisamente nel sistema

$$\begin{cases} a_{11}b_{12} & + & a_{12}b_{22} & + & \ldots & a_{1n}b_{n2} & = & 0 \\ a_{21}b_{12} & + & a_{22}b_{22} & + & \ldots & a_{2n}b_{n2} & = & 1 \\ \vdots & & \vdots & & & \vdots & & \vdots \\ a_{n1}b_{12} & + & a_{n2}b_{22} & + & \ldots & a_{nn}b_{n2} & = & 0 \end{cases}$$

Continuando in questo modo, ci accorgiamo che anche per le altre colonne di B vale la stessa cosa. Possiamo formulare quello che abbiamo trovato come segue: se indichiamo con B_i la i-esima colonna della matrice B e con E_i la i-esima colonna della matrice unità I_n, e cioè E_i è la colonna con tutti gli elementi nulli tranne l'i-esimo che vale 1, allora l'equazione

$$AB = I_n$$

equivale agli n sistemi lineari

$$A \cdot B_1 = E_1, \quad A \cdot B_2 = E_2, \ldots, \quad A \cdot B_n = E_n.$$

Per determinare la matrice inversa di A dobbiamo dunque risolvere questi sistemi, ed è immediato osservare che i sistemi hanno tutti la stessa matrice dei coefficienti, e differiscono solo per la colonna dei termini noti.

In generale, il metodo di riduzione consiste nel considerare la matrice *completa* del sistema, cioè la matrice formata dai coefficienti e dalla colonna dei termini noti, fare in modo da avere gli 1 iniziali nella parte della matrice dei coefficienti mediante una opportuna sequenza di operazioni elementari fra righe, e applicare queste operazioni anche alla colonna dei termini noti.

Il metodo di riduzione può essere applicato facendo in modo di avere elementi nulli non solo **sotto** gli 1 iniziali, ma anche **sopra** gli 1 iniziali. Questo è possibile, compiendo delle operazioni fra righe aggiuntive rispetto a quanto descritto in precedenza: nel Passo 4 dell'algoritmo di riduzione sommiamo un multiplo della riga in cui abbiamo ottenuto l'1 iniziale non solo alle righe sotto ma anche a quelle sopra. Risolvendo un sistema in questo modo, alla fine della riduzione, **se la matrice dei coefficienti si è trasformata nella matrice identità**, nella colonna dei termini noti si

leggono le soluzioni del sistema. Vediamo questa ultima affermazione con un esempio. Risolviamo il sistema:

$$\begin{cases} 2x & + & 3y & + & z & = & 1 \\ x & + & 2y & - & z & = & 0 \\ x & + & y & + & z & = & 0 \end{cases}$$

La matrice completa del sistema è

$$\begin{bmatrix} 2 & 3 & 1 & 1 \\ 1 & 2 & -1 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

e i passi di riduzione sono (indichiamo in grassetto gli 1 iniziali)

$$\begin{bmatrix} 2 & 3 & 1 & 1 \\ 1 & 2 & -1 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} \mathbf{1} & 1 & 1 & 0 \\ 1 & 2 & -1 & 0 \\ 2 & 3 & 1 & 1 \end{bmatrix} \xrightarrow[R_3 = R_3 - 2R_1]{R_2 = R_2 - R_1} \begin{bmatrix} \mathbf{1} & 1 & 1 & 0 \\ 0 & \mathbf{1} & -2 & 0 \\ 0 & 1 & -1 & 1 \end{bmatrix}$$

$$\xrightarrow{R_3 = R_3 - R_2} \begin{bmatrix} \mathbf{1} & 1 & 1 & 0 \\ 0 & \mathbf{1} & -2 & 0 \\ 0 & 0 & \mathbf{1} & 1 \end{bmatrix}$$

A questo punto la matrice è ridotta per righe, ma possiamo continuare mettendo degli 0 sopra gli 1 iniziali, sempre usando operazioni elementari fra righe:

$$\begin{bmatrix} \mathbf{1} & 1 & 1 & 0 \\ 0 & \mathbf{1} & -2 & 0 \\ 0 & 0 & \mathbf{1} & 1 \end{bmatrix} \xrightarrow[R_1 = R_1 - R_3]{R_2 = R_2 + 2R_3} \begin{bmatrix} \mathbf{1} & 1 & 0 & -1 \\ 0 & \mathbf{1} & 0 & 2 \\ 0 & 0 & \mathbf{1} & 1 \end{bmatrix} \xrightarrow{R_1 = R_1 - R_2} \begin{bmatrix} \mathbf{1} & 0 & 0 & -3 \\ 0 & \mathbf{1} & 0 & 2 \\ 0 & 0 & \mathbf{1} & 1 \end{bmatrix}$$

Il sistema si è trasformato in

$$\begin{cases} x & = & -3 \\ y & = & 2 \\ z & = & 1 \end{cases}$$

e cioè è risolto, e le soluzioni si leggono nella colonna dei termini noti. Se poniamo

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & -1 \\ 1 & 1 & 1 \end{pmatrix},$$

allora il sistema che abbiamo appena risolto dà la prima colonna della matrice inversa di A. Per trovare le altre due colonne dell'inversa, dobbiamo risolvere i sistemi

$$\begin{cases} 2x & + & 3y & + & z & = & 0 \\ x & + & 2y & - & z & = & 1 \\ x & + & y & + & z & = & 0 \end{cases}, \quad \begin{cases} 2x & + & 3y & + & z & = & 0 \\ x & + & 2y & - & z & = & 0 \\ x & + & y & + & z & = & 1 \end{cases}$$

ma ci rendiamo subito conto che dobbiamo ripetere gli stessi passi di riduzione appena compiuti, usando però colonne dei termini noti diverse. Possiamo compiere simultanenamente la riduzione nel seguente modo: consideriamo la matrice

$$\begin{bmatrix} 2 & 3 & 1 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} = [A \mid I_3]$$

e riduciamo per righe, fino ad ottenere (se possibile) la matrice identità nella prima parte. Notiamo che i passi sono gli stessi compiuti in precedenza.

$$\begin{bmatrix} 2 & 3 & 1 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 2 & 3 & 1 & 1 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow[R_3 = R_3 - 2R_1]{R_2 = R_2 - R_1} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & -2 & 0 & 1 & -1 \\ 0 & 1 & -1 & 1 & 0 & -2 \end{bmatrix}$$

$$\xrightarrow{R_3 = R_3 - R_2} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & -2 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{bmatrix}$$

$$\xrightarrow[R_1 = R_1 - R_3]{R_2 = R_2 + 2R_3} \begin{bmatrix} 1 & 1 & 0 & -1 & 1 & 2 \\ 0 & 1 & 0 & 2 & -1 & -3 \\ 0 & 0 & 1 & 1 & -1 & 1 \end{bmatrix}$$

$$\xrightarrow{R_1 = R_1 - R_2} \begin{bmatrix} 1 & 0 & 0 & -3 & 2 & 5 \\ 0 & 1 & 0 & 2 & -1 & -3 \\ 0 & 0 & 1 & 1 & -1 & 1 \end{bmatrix}$$

Poiché otteniamo la matrice identità nella prima parte, nella seconda parte leggiamo la matrice inversa: infatti sulle colonne ci sono la soluzione dei tre sistemi che danno l'inversa, e dunque è proprio la matrice inversa. Abbiamo perciò:

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & -1 \\ 1 & 1 & 1 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} -3 & 2 & 5 \\ 2 & -1 & -3 \\ 1 & -1 & 1 \end{pmatrix}.$$

Possiamo riassumere quello che abbiamo ottenuto come segue:

Teorema 5.1. *Se una matrice quadrata A di ordine n può essere ridotta per righe alla matrice identità, allora la matrice inversa A^{-1} esiste e può essere determinata come segue:*

(1) *si considera la matrice di tipo* $(n, 2n)$ $C = [A \mid I_n]$*, formata a sinistra da* A *e a destra da* I_n*;*
(2) *si riduce per righe la matrice* C *fino ad ottenere la matrice* $C' = [I_n \mid B]$*;*
(3) *si ha* $A^{-1} = B$*, la matrice che compare a destra nella matrice* C'*.*

La dimostrazione si ottiene ragionando come nell'esempio precedente, ma supponendo che A sia una matrice quadrata generica di ordine n. Ci si rende subito conto che non cambia niente dal caso 3×3 che abbiamo visto nell'esempio. Si può anche dimostrare (ma noi non lo faremo) che se esiste la matrice inversa A^{-1} allora è possibile ridurre mediante operazioni elementari fra righe la matrice A alla matrice identità I_n.

Enunciamo infine il teorema sulla risoluzione dei sistemi lineari quadrati:

Teorema 5.2. *Se una matrice* A *quadrata di ordine* n *è invertibile, allora per ogni matrice* B *di tipo* $(n, 1)$ *(cioè per ogni colonna di termini noti), il sistema di equazioni lineari*

$$AX = B$$

ha una e una sola soluzione, data da $X = A^{-1}B$*.*

Dimostrazione. Dall'equazione $AX = B$, moltiplicando a sinistra per la matrice A^{-1} si ha:

$$AX = B \iff A^{-1}AX = A^{-1}B \iff I_nX = A^{-1}B \iff X = A^{-1}B$$

Dunque X è soluzione se e solo se $X = A^{-1}B$, e poiché l'inversa è unica per il Teorema 3.6, anche la soluzione X trovata è l'unica possibile. □

Osserviamo che la formula $X = A^{-1}B$ consente di risolvere sistemi lineari quando si sappia calcolare la matrice A^{-1} inversa della matrice A dei coefficienti. Questo può essere utile se si devono risolvere molti sistemi che abbiano la stessa matrice dei coefficienti ma diverse colonne di termini noti: si calcola la matrice A^{-1} una sola volta e le soluzioni si ottengono per moltiplicazione.

Se la matrice A non ammette inversa, allora il sistema $AX = B$ può avere infinite soluzioni oppure nessuna, ma mai soluzione unica (non dimostreremo questo fatto).

6. Determinanti

Il metodo visto nel paragrafo precedente per determinare la matrice inversa di una matrice data A non consente di rispondere alla domanda "A è invertibile?" senza determinare l'inversa. Inoltre, la matrice inversa è determinata

dall'algoritmo e non da una formula. Se nella pratica avere un algoritmo efficiente è una buona cosa, avere una formula può essere più utile per studiare le proprietà della matrice inversa, senza dover tutte le volte eseguire tutti i calcoli.

La teoria dei determinanti risponde a queste esigenze. In termini di determinante di una matrice c'è un criterio semplice per distinguere le matrici invertibili da quelle non invertibili, e possiamo anche dare una formula per gli elementi della matrice inversa in funzione degli elementi della matrice originale. Questo è utile se interessa calcolare solo alcuni elementi dell'inversa. Questo con l'algoritmo di riduzione non è in genere possibile: certi elementi si possono calcolare solo dopo averne calcolati altri, che magari non interessano. Un'altra formula utile che si può scrivere mediante i determinanti è la Formula di Cramer, per la soluzione dei sistemi lineari quadrati che hanno soluzione unica.

Vi sono molti modi di definire il determinante di una matrice. Noi abbiamo scelto di usare una definizione ricorsiva, e di usare la Prima Regola di Laplace come giustificazione per la definizione. In altri libri si trova spesso una definizione in termini di permutazioni, e la Prima Regola di Laplace è una formula per il calcolo del determinante, piuttosto che una definizione. Naturalmente le due definizioni sono equivalenti, e si può vedere un qualunque libro di Algebra Lineare per avere maggiori informazioni al riguardo.

Cominciamo con il definire il determinante di una matrice 2×2.

Definizione 6.1. Sia

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Il **determinante** di A è il numero

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Notiamo che la notazione precedente, gli elementi di una matrice racchiusi tra barre verticali, e non tra parentesi tonde o quadre, significa sempre il deteminante della matrice.

Una prima osservazione che si può fare sul significato del determinante è la seguente: se A è una matrice 2×2, allora $\det(A) = 0$ se e solo se le righe di A sono multiple l'una dell'altra. Infatti

$$A = \begin{pmatrix} a & b \\ ha & hb \end{pmatrix} \implies \det(A) = ahb - hab = 0.$$

Viceversa, se $\det(A) = ad - bc = 0$ allora $ad = bc$. Se $a = b = 0$, allora la matrice A ha una riga nulla, e quindi questa riga è un multiplo, nullo,

dell'altra. Se $a \neq 0$ e $b = 0$, allora deve essere $d = 0$, e ancora le righe sono multiple, perché c sarà un multiplo di a. Se $a \neq 0$ e $b \neq 0$, allora $ad = bc \implies \dfrac{d}{b} = \dfrac{c}{a} = h$ e quindi $c = ha$, $d = hb$, cioè la seconda riga è multipla della prima.

Vediamo ora come si definisce il determinante di una matrice 3×3.

Definizione 6.2. Sia

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Il **determinante** di A è il numero

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \cdot \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \cdot \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Il significato della formula è il seguente: si considerano successivamente gli elementi di una riga, la prima, si moltiplicano ognuno per il determinante della matrice 2×2 che si ottiene cancellando la riga e la colonna su cui sta l'elemento considerato, e poi si sommano questi prodotti, alcuni con il loro segno, altri cambiati di segno. Per scrivere la formula in un modo più facile da ricordare, abbiamo bisogno di una notazione per i determinanti che vi compaiono. Poniamo allora

Definizione 6.3. Sia A una matrice quadrata. Il numero

$$A_{ij} = (-1)^{i+j} \cdot \det(\text{matrice ottenuta cancellando la riga } i \text{ e la colonna } j)$$

viene detto **complemento algebrico** dell'elemento a_{ij}.

Con questa notazione la formula del determinante di una matrice 3×3 si scrive

$$\det(A) = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}$$

Dunque il determinante di una matrice 3×3 si calcola a partire dagli elementi della matrice stessa e da determinanti di matrici 2×2. Allo stesso modo, definiamo ricorsivamente il determinante di una matrice $n \times n$.

Definizione 6.4. Sia A una matrice quadrata di ordine n. Il **determinante** di A è il numero

$$\det(A) = a_{11}A_{11} + a_{12}A_{12} + \ldots a_{1n}A_{1n} = \sum_{j=1}^{n} a_{1j}A_{1j}$$

Le quantità A_{1j} che compaiono nella formula sono determinanti di matrici quadrate di ordine $n-1$ e dunque si calcolano con la stessa formula utilizzando determinanti di matrici di ordine $n-2$ e così via fino ad arrivare a matrici di ordine 2, dove usiamo la Definizione 6.1. La formula appena data viene spesso detta **sviluppo di Laplace del determinante secondo la prima riga**.

Dire che il determinante si calcola mediante lo sviluppo secondo la *prima* riga fa venire in mente che ci sono altri modi di sviluppare, secondo un'altra riga. Infatti, perché usare la prima riga e non la seconda o la terza o una qualunque altra? O addirittura, perché non usare una colonna? Ci dovrebbero essere tanti determinanti possibili di una matrice, uno per ogni riga e uno per ogni colonna. Il fatto importante è che tutti questi "determinanti" sono uguali, e questo è il contenuto della Prima Regola di Laplace.

Teorema 6.5 (Prima Regola di Laplace)**.** *Sia A una matrice quadrata di ordine n. Allora il determinante di A può essere calcolato sviluppando secondo una riga o una colonna qualsiasi, cioè il determinante di A è uguale alla somma dei prodotti degli elementi di una riga (o una colonna) qualsiasi per i rispettivi complementi algebrici. In formule:*

$$\det(A) = \sum_{j=1}^{n} a_{ij}A_{ij} = \sum_{j=1}^{n} a_{ji}A_{ji}, \quad i = 1, \ldots, n$$

Non daremo la dimostrazione di questo importante teorema, ma vediamo almeno un esempio.

Esempio 6.6. Calcolare il determinante della matrice

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 3 & 0 \\ 2 & -1 & 2 \end{pmatrix}.$$

Sviluppiamo secondo la prima riga. Si ha:

$$\begin{aligned} \det(A) &= \begin{vmatrix} 1 & 2 & 3 \\ -1 & 3 & 0 \\ 2 & -1 & 2 \end{vmatrix} = 1 \cdot \begin{vmatrix} 3 & 0 \\ -1 & 2 \end{vmatrix} - 2 \cdot \begin{vmatrix} -1 & 0 \\ 2 & 2 \end{vmatrix} + 3 \cdot \begin{vmatrix} -1 & 3 \\ 2 & -1 \end{vmatrix} \\ &= 1 \cdot (3 \cdot 2 - 0 \cdot (-1)) - 2 \cdot (-1 \cdot 2 - 0 \cdot 2) + 3 \cdot (-1 \cdot (-1) - 3 \cdot 2) \\ &= 6 + 4 - 15 = -5 \end{aligned}$$

Se sviluppiamo secondo la seconda riga si ha:

$$\begin{aligned}\det(A) = \begin{vmatrix} 1 & 2 & 3 \\ -1 & 3 & 0 \\ 2 & -1 & 2 \end{vmatrix} &= -(-1) \cdot \begin{vmatrix} 2 & 3 \\ -1 & 2 \end{vmatrix} + 3 \cdot \begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix} \\ &= 1 \cdot (2 \cdot 2 - 3 \cdot (-1)) + 3 \cdot (1 \cdot 2 - 3 \cdot 2) + 0 \\ &= 7 - 12 + 0 = -5\end{aligned}$$

Se sviluppiamo secondo la prima colonna si ha:

$$\begin{aligned}\det(A) = \begin{vmatrix} 1 & 2 & 3 \\ -1 & 3 & 0 \\ 2 & -1 & 2 \end{vmatrix} &= 1 \cdot \begin{vmatrix} 3 & 0 \\ -1 & 2 \end{vmatrix} - (-1) \cdot \begin{vmatrix} 2 & 3 \\ -1 & 2 \end{vmatrix} + 2 \cdot \begin{vmatrix} 2 & 3 \\ 3 & 0 \end{vmatrix} \\ &= 1 \cdot (3 \cdot 2 - 0 \cdot (-1)) + 1 \cdot (2 \cdot 2 - 3 \cdot (-1)) + 2 \cdot (2 \cdot 0 - 3 \cdot 3) \\ &= 6 + 7 - 18 = -5\end{aligned}$$

e il determinante è sempre lo stesso. Notiamo che in questo caso conviene sviluppare secondo la seconda riga (oppure la terza colonna), poiché non dobbiamo calcolare il determinante 2×2 che è moltiplicato per l'elemento 0 in posizione $(2, 3)$. In generale, conviene sviluppare secondo la riga o la colonna che contiene il maggior numero di 0.

Vediamo ora un elenco di proprietà dei determinanti. Ognuna è una conseguenza abbastanza semplice delle precedenti, e tutte discendono dalla Prima Regola di Laplace. Non tutte sono ugualmente importanti: alcune sono semplicemente dei passaggi intermedi per giungere più facilmente alle successive.

Proprietà 1. *Se gli elementi di una riga (o di una colonna) sono tutti nulli, allora*

$$\det(A) = 0.$$

Infatti, basta sviluppare seconda la riga tutta nulla.

Proprietà 2. *Se la matrice A' si ottiene dalla matrice A moltiplicando tutti gli elementi di una riga (o di una colonna) per il numero c, allora*

$$\det(A') = c \cdot \det(A)$$

Supponiamo di avere moltiplicato per c la riga i della matrice $A = (a_{ij})$. Allora gli elementi della matrice $A' = (a'_{ij})$ sono

$$\begin{aligned} a'_{ij} &= c \cdot a_{ij}, \quad j = 1, \dots, n \\ a'_{kj} &= a_{kj}, \quad k \neq i, j = 1, \dots, n \end{aligned}$$

Poiché i complementi algebrici della riga i delle matrici A e A' si calcolano senza utilizzare gli elementi della riga i si ha $A'_{ij} = A_{ij}$, in quanto al di fuori

della riga i le due matrici sono uguali. Calcoliamo allora il determinante di A' sviluppando secondo la riga i:

$$\det(A') = \sum_{j=1}^{n} a'_{ij} A'_{ij} = \sum_{j=1}^{n} c \cdot a_{ij} A_{ij} = c \cdot \sum_{j=1}^{n} a_{ij} A_{ij} = c \cdot \det(A)$$

Ponendo $c = 0$ si osserva che la Proprietà 1 è un caso particolare della Proprietà 2.

Proprietà 3. *Se la matrice A' si ottiene dalla matrice A scambiando fra loro due righe (oppure due colonne), allora*

$$\det(A') = -\det(A).$$

Se la matrice A è di ordine 2, è una verifica immediata: si ha

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A' = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$$

e dunque

$$\det(A') = cb - da = -(ad - bc) = -\det(A).$$

Supponiamo allora la proprietà dimostrata per tutte le matrici di ordine $n-1$, e dimostriamo la proprietà per le matrici di ordine n. Per induzione, avremo che la proprietà è vera per le matrici di ogni ordine. Sia dunque A una matrice quadrata di ordine $n > 2$, e sia A' ottenuta da A scambiando le righe i e j. Calcoliamo il determinante di A' sviluppando secondo una riga k, diversa da i e j. Allora $a'_{kl} = a_{kl}$, cioè la riga k è uguale nella matrice A e nella matrice A', mentre i complementi algebrici A'_{kl} sono determinanti di matrici di ordine $n-1$ che sono ottenute dalle corrispondenti sottomatrici di A scambiando la riga i con la riga j. Dunque, per ipotesi induttiva, $A'_{kl} = -A_{kl}$ e allora

$$\det(A') = \sum_{l=1}^{n} a'_{kl} A'_{kl} = \sum_{l=1}^{n} a_{kl} \cdot (-A_{kl}) = -\sum_{l=1}^{n} a_{kl} A_{kl} = -\det(A)$$

e la proprietà è dimostrata.

Proprietà 4. *Il determinante di una matrice con due righe uguali (o due colonne uguali) è nullo.*

Scambiando le due righe uguali la matrice non cambia e dunque nemmeno il determinante, ma questo deve cambiare segno per la Proprietà 3. L'unica possibilità è che il determinante sia nullo.

Una importante conseguenza è la Seconda Regola di Laplace.

Seconda regola di Laplace. *Sia A una matrice quadrata di ordine n. La somma dei prodotti degli elementi di una riga (o una colonna) qualsiasi per i complementi algebrici di un'altra riga (o di un'altra colonna) è 0. In formule:*

$$\sum_{j=1}^{n} a_{ij}A_{kj} = \sum_{j=1}^{n} a_{ji}A_{jk} = 0, \quad i \neq k$$

Infatti, consideriamo la matrice A' che si ottiene dalla matrice A ripetendo, nella riga k, gli elementi della riga i. Poiché A' ha le due righe i e k uguali il suo determinante è nullo. D'altra parte, sviluppando il determinante di A' secondo la riga k si ottiene

$$0 = \det(A') = \sum_{j=1}^{n} a_{kj}A'_{kj} = \sum_{j=1}^{n} a_{ij}A_{kj}$$

poiché $a_{kj} = a_{ij}$ in quanto nella riga k abbiamo ripetuto la riga i e $A'_{kj} = A_{kj}$ in quanto questi complementi algebrici si calcolano non considerando la riga k, che è l'unica differenza fra le matrici A e A'.

Le Proprietà 2 e 3 descrivono il comportamento del determinante quando su una matrice si esegue la prima o la seconda operazione elementare che abbiamo usato nel metodo di riduzione. Ci possiamo chiedere cosa accade quando si usa la terza operazione.

Proprietà 5. *Se la matrice A' si ottiene dalla matrice A sostituendo la riga i con la riga i più un multiplo della riga k, allora*

$$\det(A') = \det(A)$$

Gli elementi di A' sono

$$\begin{aligned} a'_{ij} &= a_{ij} + c \cdot a_{kj}, \quad j = 1, \dots, n \\ a'_{lj} &= a_{lj}, \quad l \neq i, j = 1, \dots, n \end{aligned}$$

e per i complementi algebrici si ha $A'_{ij} = A_{ij}$ in quanto al di fuori dalla riga i le due matrici coincidono. Sviluppando il determinante di A' secondo la riga i si ha:

$$\begin{aligned} \det(A') &= \sum_{j=1}^{n} a'_{ij}A'_{ij} = \sum_{j=1}^{n} (a_{ij} + c \cdot a_{kj})A_{ij} \\ &= \sum_{j=1}^{n} a_{ij}A_{ij} + c\sum_{j=1}^{n} a_{kj}A_{ij} = \det(A) \end{aligned}$$

in quanto nell'ultima somma il primo addendo vale $\det(A)$ per la Prima Regola di Laplace e il secondo addendo vale 0 per la Seconda Regola di Laplace.

Possiamo dunque pensare di calcolare il determinante di una matrice effettuando la riduzione per righe, tenendo conto ad ogni passo dei cambiamenti del determinante usando le Proprietà 2, 3 e 5 e poi calcolando il determinante della matrice ridotta. È immediato osservare che una matrice quadrata ridotta per righe è **triangolare**, cioè tutti gli elementi sotto la diagonale principale sono nulli. È facile calcolare il determinante di una tale matrice.

Proprietà 6. *Sia A una matrice triangolare. Allora il determinante di A è il prodotto degli elementi della diagonale principale.*

La matrice A ha la forma

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

e si può calcolare il determinante sviluppando secondo la prima colonna, e proseguire calcolando i determinanti sempre secondo la prima colonna. Si ottiene

$$\det(A) = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn}$$

Le dimostrazioni delle ultime due proprietà sono più difficili, e non le daremo. Abbiamo visto all'inizio che per una matrice A di ordine 2 la condizione $\det(A)$ è equivalente ad avere le righe multiple l'una dell'altra. Questo non è più vero in generale, e la corretta generalizzazione è la seguente.

Proprietà 7. *Sia A una matrice quadrata di ordine n. Allora* $\det(A) = 0$ *se e solo se una riga (oppure una colonna) di A è combinazione lineare delle altre.*

Se una riga, per esempio la prima, è combinazione lineare delle altre, e cioè

$$R_1 = c_2 R_2 + \dots c_n R_n$$

dove abbiamo indicato con $R_1, \dots, R_n$ le righe della matrice A, allora non è difficile, usando ripetutamente la Seconda Regola di Laplace, dimostrare che $\det(A) = 0$. Il viceversa è più difficile.

Proprietà 8 (Teorema di Binet)**.** *Se A e B sono due matrici quadrate di ordine n, allora*

$$\det(AB) = \det(A) \cdot \det(B)$$

Questo teorema è sorprendente e non esiste nessuna formula analoga per la *somma* di matrici. Notiamo che una conseguenza del teorema è che, sebbene in generale $AB \neq BA$, tuttavia è sempre $\det(AB) = \det(BA)$, fatto che forse non è da aspettarsi.

Usando le due Regole di Laplace ed il Teorema di Binet possiamo finalmente caratterizzare le matrici che ammettono inversa e anche dare una formula per il calcolo della matrice inversa.

Teorema 6.7. *Sia A una matrice quadrata di ordine n. La matrice A è invertibile se e solo se* $\det(A) \neq 0$. *Sia $B = (b_{ij})$ la matrice inversa. Allora*

$$b_{ij} = \frac{A_{ji}}{\det(A)}$$

Notiamo che nella definizione di b_{ij} non c'è errore: gli indici si scambiano.

Dimostrazione. Se A è invertibile e B è la matrice inversa, allora $AB = I_n$. Dal Teorema di Binet si ha:

$$\det(A) \cdot \det(B) = \det(AB) = \det(I_n) = 1$$

e dunque $\det(A)$ non può essere 0. Notiamo anche che $\det(B) = 1/\det(A)$ e cioè che il determinante della matrice inversa è l'inverso del determinante.

Sia ora A tale che $\det(A) \neq 0$. Allora basta dimostrare che la matrice B definita nell'enunciato del teorema è in effetti la matrice inversa di A. Per fare ciò, basta calcolare i prodotti AB e BA e dimostrare che sono entrambi uguali a I_n. Dimostriamo in dettaglio che $AB = I_n$ e lasciamo la verifica dell'altra condizione per esercizio.

Sia $AB = C$, e siano c_{ij} gli elementi di C. Se $i = j$, cioè siamo sulla diagonale principale, dalla definizione di prodotto di matrici si ha:

$$c_{ii} = \sum_{k=1}^{n} a_{ik} b_{ki} = \sum_{k=1}^{n} a_{ik} \frac{A_{ik}}{\det(A)} = \frac{1}{\det(A)} \sum_{k=1}^{n} a_{ik} A_{ik} = 1$$

perché l'ultima somma è $\det(A)$ per la Prima Regola di Laplace.

Sia ora $i \neq j$. Si ha

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = \sum_{k=1}^{n} a_{ik} \frac{A_{jk}}{\det(A)} = \frac{1}{\det(A)} \sum_{k=1}^{n} a_{ik} A_{jk} = 0$$

usando la Seconda Regola di Laplace. Abbiamo perciò che $C = I_n$. □

Esempio 6.8. Calcoliamo l'inversa della matrice

$$A = \begin{pmatrix} 3 & -1 & 0 \\ -2 & 1 & 1 \\ 2 & -1 & 4 \end{pmatrix}$$

Si ha $\det(A) = 5$, e i complementi algebrici degli elementi di A sono

$$\begin{array}{lll} A_{11} = 5, & A_{12} = 10, & A_{13} = 0 \\ A_{21} = 4, & A_{22} = 12, & A_{23} = 1 \\ A_{31} = -1, & A_{32} = -3, & A_{33} = 1 \end{array}$$

Possiamo allora scrivere la matrice inversa, ricordando che nella posizione (i, j) va il valore $A_{ji}/\det(A)$. Si ha

$$A^{-1} = \begin{pmatrix} 1 & 4/5 & -1/5 \\ 2 & 12/5 & -3/5 \\ 0 & 1/5 & 1/5 \end{pmatrix}$$

Usando il Teorema 6.7, possiamo scrivere la formula risolutiva per i sistemi lineari quadrati con determinante diverso da 0, cioè quelli per cui la soluzione è unica. La formula, ben nota nei casi di sistemi 2×2 e 3×3, è la Formula di Cramer.

Sia $AX = B$ un sistema di equazioni lineari, dove A è una matrice quadrata di ordine n. Sia A_i la matrice che si ottiene dalla matrice A sostituendo la *colonna* i con la colonna B dei termini noti. Allora

Formula di Cramer. *Se* $\det(A) \neq 0$ *allora il sistema ha soluzione unica, e la soluzione è*

$$x_i = \frac{\det A_i}{\det(A)}, \quad i = 1, 2, \ldots, n$$

Dimostrazione. La matrice A è invertibile per il Teorema 6.7, e per il Teorema 5.2 sappiamo che il sistema ha soluzione unica, data da $X = A^{-1}B$. Basta dunque calcolare tale prodotto. Il valore x_i è l'elemento di posto $(i, 1)$ nella colonna delle incognite X e quindi si ottiene moltiplicando la riga i della matrice A^{-1} per la colonna B dei termini noti. Si ha

$$x_i = \frac{A_{1i}}{\det(A)} b_1 + \frac{A_{2i}}{\det(A)} b_2 + \cdots + \frac{A_{ni}}{\det(A)} b_n = \frac{1}{\det(A)} \sum_{k=1}^{n} A_{ki} b_k = \frac{\det A_i}{\det(A)}$$

dove l'ultima uguaglianza è lo sviluppo del determinante di A_i secondo la colonna i. □

Esempio 6.9. Illustriamo la Formula di Cramer con un esempio. Risolviamo il sistema

$$\begin{cases} 2x + 3y + z = 2 \\ x + 2y - z = 3 \\ x + y + z = 1 \end{cases}$$

La matrice A, e le matrici A_1, A_2 e A_3 sono

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$A_1 = \begin{pmatrix} 2 & 3 & 1 \\ 3 & 2 & -1 \\ 1 & 1 & 1 \end{pmatrix} \quad A_2 = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & -1 \\ 1 & 1 & 1 \end{pmatrix} \quad A_3 = \begin{pmatrix} 2 & 3 & 2 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}$$

Calcolando i determinanti si ha

$$\det(A) = -1, \quad \det(A_1) = -5, \quad \det(A_2) = 2, \quad \det(A_3) = 2$$

e dunque la soluzione del sistema è $x = 5$, $y = -2$, $z = -2$, come si può verificare sostituendo nelle equazioni del sistema.

Esercizi

Risolvere i seguenti sistemi lineari:

(1)
$$\begin{cases} 2x_1 - 3x_2 = -2 \\ 2x_1 + x_2 = 1 \\ 3x_1 + 2x_2 = 1 \end{cases}$$

(2)
$$\begin{cases} 4x_1 - 8x_2 = 12 \\ 3x_1 - 6x_2 = 9 \\ -3x_1 + 4x_2 = -6 \end{cases}$$

(3)
$$\begin{cases} 3x_1 + 2x_2 - x_3 = -15 \\ 5x_1 + 3x_2 + 2x_3 = 0 \\ 3x_1 + x_2 + 3x_3 = 11 \\ 11x_1 + 7x_2 = -30 \end{cases}$$

(4)
$$\begin{cases} 5x_1 + 2x_2 + 6x_3 = 0 \\ -2x_1 + x_2 + 3x_3 = 0 \end{cases}$$

(5)
$$\begin{cases} x_1 - 2x_2 + x_3 - 4x_4 = 1 \\ x_1 + 3x_2 + 7x_3 + 2x_4 = 2 \\ x_1 - 12x_2 - 11x_3 - 16x_4 = 5 \end{cases}$$