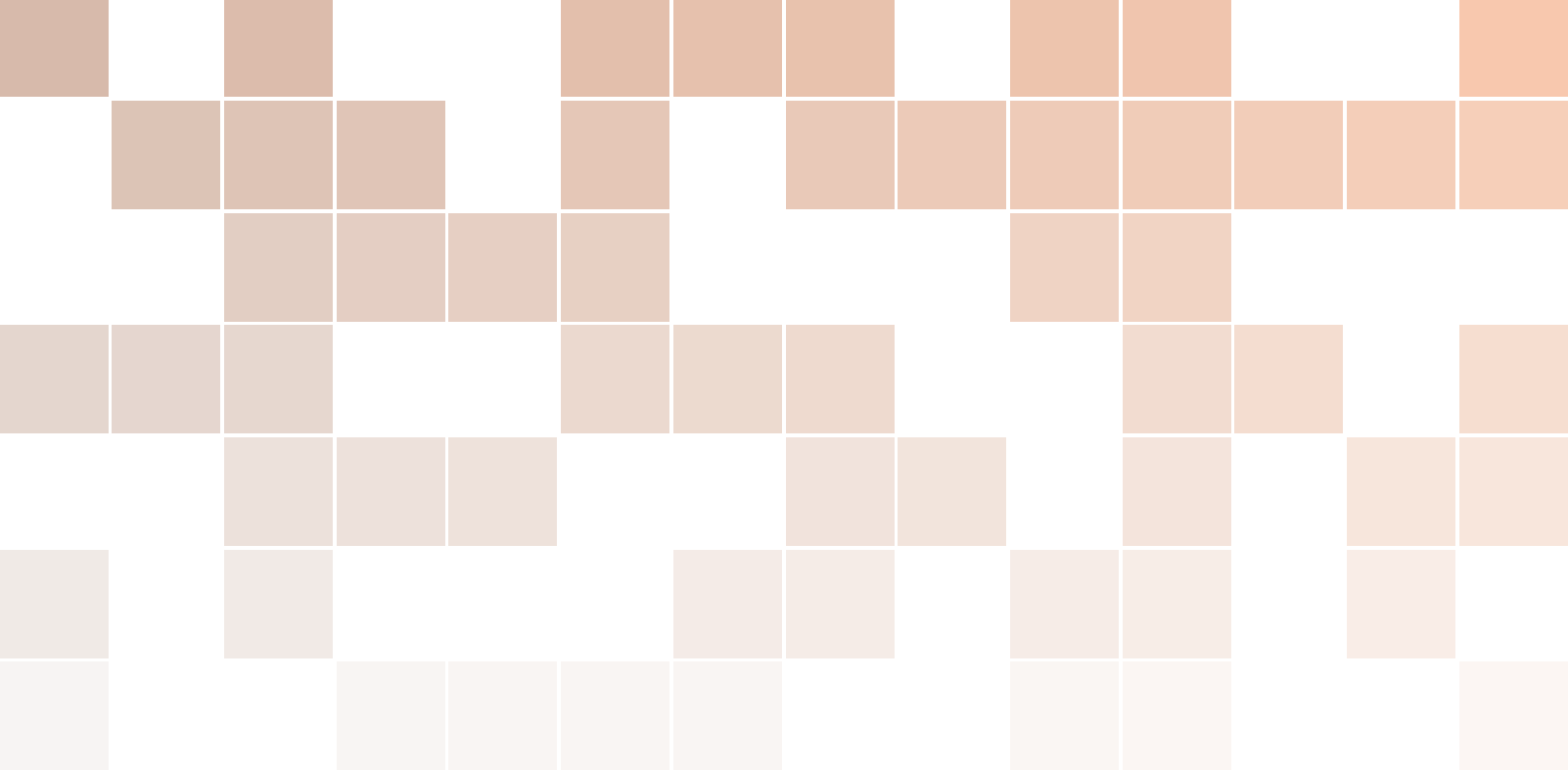
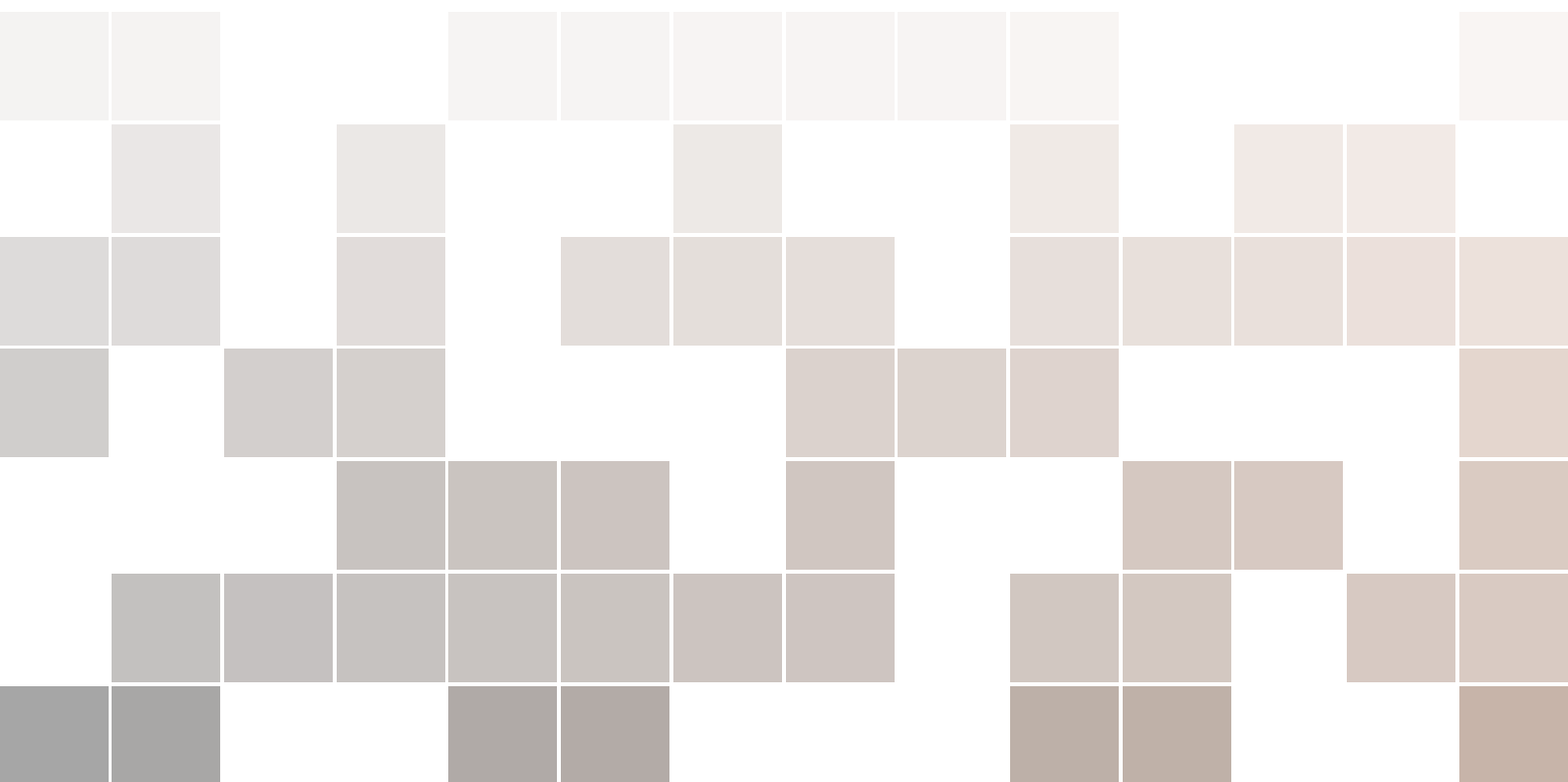




# Prerequisiti

**Enrico Bertolazzi**



Copyright © 2025 Enrico Bertolazzi

PUBLISHED BY PUBLISHER

[BOOK-WEBSITE.COM](https://book-website.com)

Licensed under the Creative Commons Attribution-NonCommercial 4.0 License (the “License”). You may not use this file except in compliance with the License. You may obtain a copy of the License at <https://creativecommons.org/licenses/by-nc-sa/4.0>. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an “AS IS” BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

*First printing, May 2025*



# Contents

|          |                                                                      |
|----------|----------------------------------------------------------------------|
| .....    | 7                                                                    |
| <b>1</b> | <b>I numeri complessi</b> ..... 9                                    |
| 1.1      | Operazioni di somma e prodotto su $\mathbb{R}^2$ ..... 9             |
| <b>2</b> | <b>Il campo dei numeri complessi</b> ..... 13                        |
| 2.1      | Proprietà dell'operazione di coniugazione ..... 14                   |
| 2.2      | Proprietà del modulo ..... 16                                        |
| 2.2.1    | Interpretazione geometrica della disuguaglianza triangolare ..... 16 |
| 2.3      | Forma polare dei numeri complessi ..... 17                           |
| 2.4      | Prodotto di numeri complessi in forma polare ..... 21                |
| 2.5      | Formula di de Moivre ..... 22                                        |
| 2.5.1    | Esempio di calcolo ..... 23                                          |
| <b>3</b> | <b>Esercizi Proposti</b> ..... 29                                    |
| <b>4</b> | <b>Matrici e vettori</b> ..... 33                                    |
| 4.1      | Notazioni ..... 33                                                   |
| 4.2      | Matrici ..... 33                                                     |
| 4.3      | Vettori ..... 34                                                     |
| 4.4      | Somma, differenza e prodotto per uno scalare ..... 35                |
| 4.5      | Operazioni con le matrici ..... 38                                   |
| <b>5</b> | <b>Norme, prodotti scalari ortogonalità</b> ..... 45                 |
| 5.1      | Norme di vettori ..... 45                                            |
| 5.2      | Diseguaglianze di Young, Hölder e Minkowski. .... 46                 |
| 5.3      | Alcune proprietà delle norme vettoriali ..... 48                     |
| 5.3.1    | Continuità delle norme ..... 48                                      |
| 5.3.2    | Equivalenza delle norme ..... 49                                     |
| 5.3.3    | Confronto tra norme standard ..... 49                                |

|          |                                                                                 |           |
|----------|---------------------------------------------------------------------------------|-----------|
| 5.4      | Prodotti scalari                                                                | 51        |
| 5.5      | Ortogonalità e angolo tra vettori                                               | 55        |
| 5.6      | Prodotto Vettoriale                                                             | 56        |
| 5.7      | Indipendenza Lineare e Basi in $\mathbb{K}^n$                                   | 58        |
| 5.8      | Indipendenza lineare dei vettori ortogonali                                     | 59        |
| 5.9      | Ortonormalizzazione di Gram-Schmidt                                             | 60        |
| <b>6</b> | <b>Determinanti</b>                                                             | <b>65</b> |
| 6.1      | Definizione Assiomatica                                                         | 65        |
| 6.2      | Alcune proprietà dei determinanti                                               | 67        |
| 6.3      | Esistenza ed unicità del determinante                                           | 70        |
| 6.4      | Determinanti del prodotto o teorema di Binet                                    | 71        |
| 6.5      | Determinante della matrice inversa                                              | 72        |
| 6.6      | Dipendenza e indipendenza lineare                                               | 73        |
| 6.7      | Indipendenza Lineare                                                            | 73        |
| 6.8      | Dimostrazione Theorem 7.2                                                       | 77        |
| 6.9      | Rango di una matrice                                                            | 78        |
| 6.10     | Cofattori di una matrice quadrata                                               | 78        |
| 6.11     | Cofattore                                                                       | 79        |
| 6.12     | Matrice Cofattore                                                               | 79        |
| 6.13     | Rappresentazione della matrice inversa                                          | 81        |
| 6.14     | Calcolo dei cofattori                                                           | 82        |
| 6.15     | Determinante della Trasposta                                                    | 84        |
| 6.16     | Determinante delle Matrici Diagonali a Blocchi                                  | 85        |
| <b>7</b> | <b>Sistemi lineari: regola di Cramer e Rouchè-Capelli</b>                       | <b>87</b> |
| 7.1      | Regola di Cramer                                                                | 87        |
| 7.2      | Il teorema di Rouchè-Capelli                                                    | 88        |
| 7.2.1    | Teorema di Rouchè-Capelli (versione estesa)                                     | 89        |
| 7.2.2    | Esistenza e unicità per sistemi lineari                                         | 89        |
| <b>8</b> | <b>Autovalori e Autovettori</b>                                                 | <b>91</b> |
| 8.1      | Polinomio Caratteristico                                                        | 91        |
| 8.2      | Autovettori                                                                     | 93        |
| 8.3      | Autovalori multipli                                                             | 93        |
| 8.3.1    | Molteplicità Geometrica                                                         | 93        |
| 8.3.2    | Molteplicità Algebrica                                                          | 93        |
| 8.4      | Matrici reali simmetriche e hermitiane                                          | 95        |
| 8.5      | Diagonalizzazione                                                               | 98        |
| 8.6      | Diagonalizzazione di matrici Hermitiane                                         | 99        |
| 8.7      | Spettro, raggio spettrale e localizzazione degli autovalori sul piano complesso | 99        |
| 8.8      | Teorema dei cerchi di Gerschgorin                                               | 100       |

---

|          |                                        |            |
|----------|----------------------------------------|------------|
| 8.9      | Matrici a Diagonale Dominante .....    | 102        |
| 8.10     | Matrici Definite Positive .....        | 102        |
| 8.11     | Autovalori di Matrici SPD .....        | 103        |
| 8.12     | Determinante di Matrici SPD .....      | 103        |
| 8.13     | Blocchi Diagonali di Matrici SPD ..... | 104        |
| 8.14     | Teorema di Sylvester .....             | 104        |
| <b>9</b> | <b>Norme di matrici .....</b>          | <b>107</b> |
| 9.1      | Cos'è la norma di una matrice? .....   | 107        |
| 9.2      | Definizione assiomatica di norma ..... | 110        |
| 9.3      | Norme compatibili .....                | 110        |
| 9.3.1    | Definizione di norme compatibili ..... | 110        |
| 9.4      | .....                                  | 110        |
|          | <b>References .....</b>                | <b>113</b> |





Ordinary differential equations are a powerful and essential tool in both theoretical and applied mathematics. Their ability to model dynamic systems, whether simple or complex, makes them indispensable in understanding the behavior of many real-world systems.

With advanced techniques and numerical methods, ODEs continue to play a crucial role in scientific research, engineering design, and beyond. Understanding ODEs is therefore central to many fields of study, from physics and engineering to biology and economics.

#### **Acknowledgement**



Figure 1: License: CC BY 4.0

This project is licensed under a [Creative Commons Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/). Some pages may contain materials that are subject to copyright, in which case you will see the copyright notice. make



# 1. I numeri complessi

## 1.1 Operazioni di somma e prodotto su $\mathbb{R}^2$

Consideriamo l'insieme  $\mathbb{R}^2$  di tutte le coppie ordinate di numeri reali:

$$\mathbb{R}^2 = \{(a, b) \mid a, b \in \mathbb{R}\}$$

Due coppie ordinate  $(a, b)$  e  $(c, d)$  sono **uguali** se e solo se:

$$(a, b) = (c, d) \iff a = c \text{ e } b = d$$

**Definiamo su  $\mathbb{R}^2$  due operazioni:**

- **Somma:** Dati due elementi  $(a, b)$  e  $(c, d)$  di  $\mathbb{R}^2$ , la loro **somma** è l'elemento  $(a + c, b + d)$  di  $\mathbb{R}^2$  definito da:

$$(a, b) + (c, d) = (a + c, b + d) \quad (1.1)$$

- **Prodotto:** Dati due elementi  $(a, b)$  e  $(c, d)$  di  $\mathbb{R}^2$ , il loro **prodotto** è l'elemento  $(ac - bd, ad + bc)$  di  $\mathbb{R}^2$  definito da:

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc) \quad (1.2)$$

**Osservazioni:**

- Le operazioni di somma e prodotto definite su  $\mathbb{R}^2$  estendono le usuali operazioni di somma e prodotto tra numeri reali, identificando un numero reale  $a$  con l'elemento  $(a, 0)$  di  $\mathbb{R}^2$ .
- Con queste operazioni, l'insieme  $\mathbb{R}^2$  forma un campo, noto come **campo dei numeri complessi**.

dove le operazioni di somma e di prodotto nei secondi membri delle equazioni Equation 1.1 e Equation 1.2 sono le operazioni standard definite nel campo dei numeri reali  $\mathbb{R}$ .

Esempio

$$(1, 2) + (-2, 3) = (-1, 5)$$

$$(2, 3) \cdot (1, -4) = (2 + 12, -8 + 3) = (14, -5)$$

### Proprietà dell'operazione di somma in $\mathbb{R}^2$

Consideriamo l'operazione di somma in  $\mathbb{R}^2$  definita da Equation 1.1.

Per ogni  $(a, b), (c, d), (e, f) \in \mathbb{R}^2$ , valgono le seguenti proprietà:

**1. Proprietà associativa:**

La somma è associativa, ovvero:

$$[(a, b) + (c, d)] + (e, f) = (a, b) + [(c, d) + (e, f)].$$

**2. Proprietà commutativa:**

La somma è commutativa, ovvero:

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

**3. Elemento neutro:**

Esiste un elemento neutro per la somma, ovvero un elemento  $(0, 0)$  tale che:

$$(a, b) + (0, 0) = (0, 0) + (a, b) = (a, b).$$

Indichiamo l'elemento neutro con il simbolo 0.

**4. Elemento opposto:**

Ogni elemento  $(a, b)$  ammette un opposto, ovvero un elemento  $(-a, -b)$  tale che:

$$(a, b) + (-a, -b) = (-a, -b) + (a, b) = (0, 0).$$

### Verifica delle proprietà:

Le proprietà sopra enunciate si verificano facilmente sfruttando le proprietà della somma dei numeri reali. Ad esempio, per dimostrare la proprietà associativa, si ha:

$$\begin{aligned} [(a, b) + (c, d)] + (e, f) &= (a + c, b + d) + (e, f) \\ &= ((a + c) + e, (b + d) + f) \\ &= (a + (c + e), b + (d + f)) \\ &= (a, b) + (c + e, d + f) \\ &= (a, b) + [(c, d) + (e, f)]. \end{aligned}$$

È facile verificare che tale opposto è costituito dalla coppia  $(-a, -b)$ .

Le altre proprietà si dimostrano in modo analogo.

### Proprietà del prodotto in $\mathbb{R}^2$

L'operazione di prodotto in  $\mathbb{R}^2$  definita in Equation 1.2 e qui ripetuta:

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc)$$

soddisfa le seguenti proprietà per ogni  $(a, b), (c, d), (e, f) \in \mathbb{R}^2$ :

**1. Proprietà associativa:**

Il prodotto è associativo, ovvero:

$$[(a, b) \cdot (c, d)] \cdot (e, f) = (a, b) \cdot [(c, d) \cdot (e, f)]$$

**2. Proprietà commutativa:**

Il prodotto è commutativo, ovvero:

$$(a, b) \cdot (c, d) = (c, d) \cdot (a, b)$$

**3. Elemento neutro:**

Esiste un elemento neutro per il prodotto, ovvero l'elemento  $(1, 0)$  tale che:

$$(a, b) \cdot (1, 0) = (1, 0) \cdot (a, b) = (a, b)$$

Indichiamo l'elemento neutro con il simbolo 1.

**4. Proprietà distributiva rispetto alla somma:**

Il prodotto è distributivo rispetto alla somma, ovvero:

$$[(a, b) + (c, d)] \cdot (e, f) = (a, b) \cdot (e, f) + (c, d) \cdot (e, f)$$

**Dimostrazione della proprietà commutativa**

Per dimostrare la proprietà commutativa, consideriamo:

$$\begin{aligned}(a, b) \cdot (c, d) &= (ac - bd, ad + bc) \\ &= (ca - db, cb + da) && \text{per la commutatività del} \\ &= (c, d) \cdot (a, b) && \text{prodotto e della somma in } \mathbb{R}\end{aligned}$$

Le altre proprietà si dimostrano in modo analogo, sfruttando le proprietà della somma e del prodotto in  $\mathbb{R}$ .



## 2. Il campo dei numeri complessi

Poiché l'insieme

$$\mathbb{R}' = \{(a, 0) \mid a \in \mathbb{R}\}$$

è contenuto in  $\mathbb{R}^2$ , ed è facilmente identificabile (mediante una corrispondenza biunivoca) con  $\mathbb{R}$ , è ragionevole e utile scrivere semplicemente  $a$  al posto di  $(a, 0)$  per ogni coppia di  $\mathbb{R}'$ . Definiamo inoltre l'**unità immaginaria**  $i$  come:

$$i \equiv (0, 1),$$

e possiamo quindi esprimere ogni coppia  $(a, b) \in \mathbb{R}^2$  nella forma:

$$(a, b) = (a, 0) + (0, b) = a + (0, b) \cdot (1, 0) = a + (b, 0) \cdot (0, 1) = a + ib.$$

Con questa notazione, osserviamo che  $i$  soddisfa la relazione:

$$i^2 = -1.$$

Possiamo ora definire l'insieme dei numeri complessi.

### 2.0.0.1 Definizione (Numeri complessi)

L'insieme dei numeri complessi, denotato con  $\mathbb{C}$ , è costituito dalle coppie di numeri reali su cui sono state introdotte le operazioni di somma e prodotto definite in Equation 1.1 e Equation 1.2:

$$\mathbb{C} = (\mathbb{R}^2, +, \cdot).$$

La notazione comunemente utilizzata per un numero complesso  $z \in \mathbb{C}$  è:

$$z = a + ib, \quad \text{con} \quad a, b \in \mathbb{R},$$

dove

$$a = \operatorname{Re}\{z\} \quad \text{è chiamato la parte reale di } z,$$

e

$$b = \operatorname{Im}\{z\} \quad \text{è chiamato la parte immaginaria di } z.$$

Notiamo che  $\mathbb{R} \subset \mathbb{C}$ , poiché ogni numero reale  $a$  può essere scritto come  $a = a + i \cdot 0$ . Inoltre, il prodotto di due numeri complessi  $z_1 = a + ib$  e  $z_2 = c + id$  segue le normali regole di moltiplicazione polinomiale. Infatti, il prodotto  $z_1 \cdot z_2$  può essere svolto sia in forma algebrica:

$$(a + ib) \cdot (c + id) = ac + iad + ibc + i^2 bd = ac - bd + i(ad + bc),$$

sia utilizzando la definizione di prodotto in  $\mathbb{R}^2$ , ottenendo lo stesso risultato.

Le proprietà delle operazioni di somma e prodotto introdotte in  $\mathbb{C}$  dimostrano che  $\mathbb{C}$  è un anello commutativo con identità moltiplicativa.

Ci proponiamo ora di dimostrare che  $\mathbb{C}$  è anche un **campo**, ovvero che ogni numero complesso diverso dallo zero ammette un inverso moltiplicativo. A tale scopo, introduciamo le nozioni di **coniugato** e di **modulo** di un numero complesso.

### 2.0.0.2 Definizione

Sia  $z = a + ib$ . Definiamo il **coniugato** di  $z$  come il numero complesso  $\bar{z} = a - ib$ .

#### Esempio

Alcuni esempi dell'operazione di coniugazione:

- $z = 1 + i \quad \bar{z} = 1 - i$
- $z = 3i \quad \bar{z} = -3i$
- $z = a \quad \bar{z} = a \quad \forall a \in \mathbb{R}$

## 2.1 Proprietà dell'operazione di coniugazione

Sia  $z = a + ib \in \mathbb{C}$ . Valgono le seguenti proprietà:

1.  $z + \bar{z} = 2a$ , da cui segue  $\Re\{z\} = \frac{z + \bar{z}}{2}$
2.  $z - \bar{z} = 2ib$ , da cui segue  $\Im\{z\} = \frac{z - \bar{z}}{2i}$
3.  $\overline{(\bar{z})} = z$
4.  $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$
5.  $\overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$
6.  $z = \bar{z} \iff z \in \mathbb{R}$
7.  $z \cdot \bar{z} = a^2 + b^2$  è un numero reale non negativo.

Dimostriamo tutte le proprietà elencate.

Siano  $z_1 = a + ib$  e  $z_2 = c + id$ . Consideriamo le seguenti proprietà:

1.  $z + \bar{z} = 2a$

Per  $z = a + ib$ , abbiamo:

$$z + \bar{z} = (a + ib) + (a - ib) = 2a.$$

$$\text{Quindi, } \Re\{z\} = \frac{z + \bar{z}}{2}.$$

2.  $z - \bar{z} = 2ib$

Sempre per  $z = a + ib$ :

$$z - \bar{z} = (a + ib) - (a - ib) = 2ib.$$

$$\text{Quindi, } \Im\{z\} = \frac{z - \bar{z}}{2i}.$$

$$3. \overline{\overline{z}} = z$$

Per  $z = a + ib$ , il coniugato di  $\overline{z} = a - ib$  è:

$$\overline{(\overline{z})} = \overline{(a - ib)} = a + ib = z.$$

$$4. \overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$$

Per  $z_1 = a + ib$  e  $z_2 = c + id$ :

$$\begin{aligned} \overline{z_1 + z_2} &= \overline{(a + ib) + (c + id)} \\ &= \overline{(a + c) + i(b + d)} \\ &= (a + c) - i(b + d) = \overline{z_1} + \overline{z_2}. \end{aligned}$$

$$5. \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$$

Per  $z_1 = a + ib$  e  $z_2 = c + id$ :

$$\begin{aligned} \overline{z_1 \cdot z_2} &= \overline{(a + ib) \cdot (c + id)} \\ &= \overline{ac + adi + bci - bd} \\ &= ac - bd - (ad + bc)i. \end{aligned}$$

Invece:

$$\begin{aligned} \overline{z_1} \cdot \overline{z_2} &= (a - ib) \cdot (c - id) \\ &= ac - adi - bci + bd \\ &= ac - bd - (ad + bc)i. \end{aligned}$$

Le due espressioni coincidono, quindi  $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$ .

$$6. z = \overline{z} \iff z \in \mathbb{R}$$

Se  $z = a + ib$  è uguale al suo coniugato:

$$z = \overline{z} \iff a + ib = a - ib \iff ib = -ib \iff b = 0.$$

Quindi,  $z$  è reale se e solo se  $b = 0$ , cioè  $z \in \mathbb{R}$ .

$$7. z \cdot \overline{z} = a^2 + b^2$$

Per  $z = a + ib$ :

$$\begin{aligned} z \cdot \overline{z} &= (a + ib) \cdot (a - ib) \\ &= a^2 - (ib)^2 \\ &= a^2 - (-b^2) \\ &= a^2 + b^2, \end{aligned}$$

che è un numero reale non negativo.

Per la proprietà 7, introduciamo la seguente definizione:

### 2.1.0.1 Definizione di modulo

Il **modulo** di un numero complesso  $z$  è definito come:

$$|z| = \sqrt{z \cdot \overline{z}}.$$

## 2.2 Proprietà del modulo

Sia  $z = a + ib \in \mathbb{C}$ . Le seguenti proprietà sono vere:

1. **Modulo del coniugato:**

$$|z| = |\bar{z}|.$$

2. **Non negatività:**

$|z|$  è un numero reale non negativo e

$$|z| = 0 \iff a^2 + b^2 = 0 \iff a = b = 0 \iff z = 0.$$

3. **Proprietà moltiplicativa:**

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2|.$$

4. **Disuguaglianza triangolare:**

$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

### 2.2.1 Interpretazione geometrica della disuguaglianza triangolare

I numeri complessi possono essere rappresentati graficamente come punti nel piano cartesiano. Per esempio, il numero  $z = a + ib$  è rappresentato dal punto con coordinate  $(a, b)$ . L'origine  $(0, 0)$  rappresenta il numero complesso 0, il punto  $(1, 0)$  rappresenta il numero complesso  $1 = 1 + 0i$ , e il punto  $(0, 1)$  rappresenta il numero complesso  $i = 0 + 1i$ .

Nel piano complesso:

- I punti sull'asse  $x$  corrispondono ai numeri reali  $(x, 0) \equiv x + 0i$ , e quest'asse è chiamato **asse reale**.
- I punti sull'asse  $y$  corrispondono ai numeri immaginari puri  $(0, y) \equiv 0 + yi$ , e quest'asse è chiamato **asse immaginario**.

Questa rappresentazione grafica aiuta a visualizzare la disuguaglianza triangolare nel contesto del piano complesso.

Geometricamente, la disuguaglianza triangolare riflette il fatto che in un triangolo la lunghezza di ogni lato è sempre minore o uguale alla somma delle lunghezze degli altri due lati (vedi figura Figure 2.2).

Dalla proprietà 2 del modulo, segue che per ogni numero complesso  $z \neq 0$ , è ben definito il numero  $1/|z|$ . Inoltre, è semplice verificare che l'inverso di  $z$  è dato da:

$$z^{-1} = \frac{\bar{z}}{|z|^2}.$$

#### Esempio (Calcolo dell'inverso di un numero complesso)

Consideriamo alcuni esempi per calcolare l'inverso di un numero complesso:

1. Per  $z = 3 - 2i$ :

$$|z|^2 = 3^2 + (-2)^2 = 13, \quad \text{quindi} \quad z^{-1} = \frac{\bar{z}}{|z|^2} = \frac{3 + 2i}{13}.$$

2. Per  $z = i$ :

$$|z|^2 = 1^2 = 1, \quad \text{quindi} \quad z^{-1} = \frac{\bar{z}}{|z|^2} = \frac{-i}{1} = -i.$$

3. Per  $z = a$ , dove  $a \in \mathbb{R}$ :

$$|z|^2 = a^2, \quad \text{quindi} \quad z^{-1} = \frac{1}{a} \quad \text{per ogni} \quad a \neq 0.$$

Questi esempi mostrano come calcolare l'inverso di numeri complessi in vari casi.

### Piano complesso

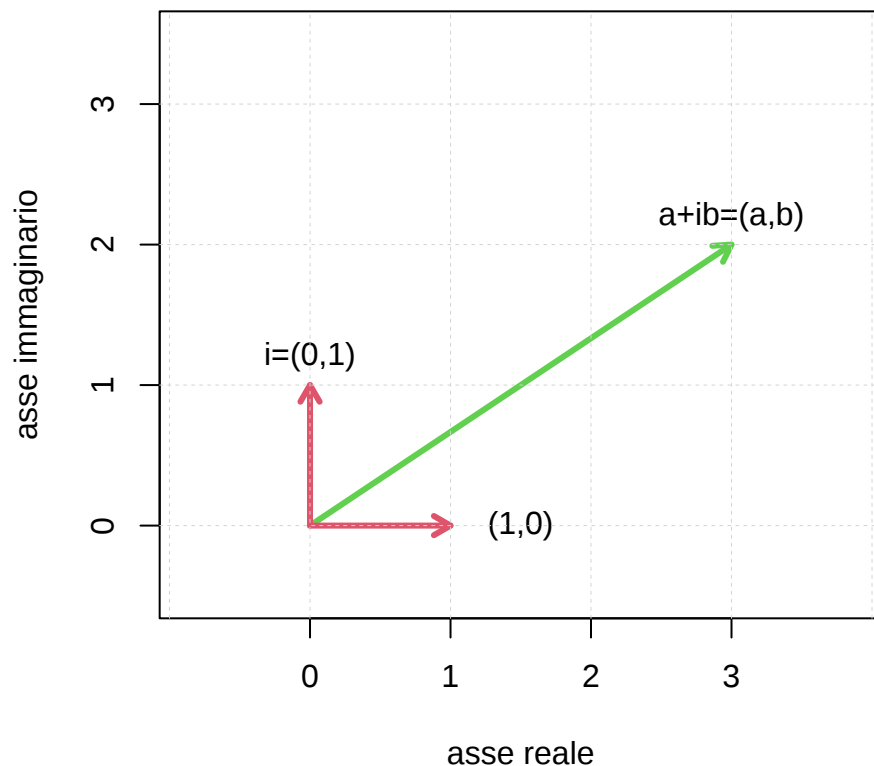


Figure 2.1: Il piano complesso.

Riferendoci all'ultimo esempio, notiamo che l'inverso di un numero reale, considerato come un elemento di  $\mathbb{C}$ , coincide con quello calcolato abitualmente in  $\mathbb{R}$ . Pertanto, la definizione di inverso in  $\mathbb{C}$  si estende naturalmente a quella di  $\mathbb{R}$ .

#### 2.2.1.1 Problema

Ora ci proponiamo di risolvere i seguenti problemi:

1. Calcolare  $1 + i^{2000}$ .
2. Trovare le soluzioni in  $\mathbb{C}$  dell'equazione  $z^n = w$ , con  $w \in \mathbb{C}$  fissato.

Per affrontare questi problemi, è utile introdurre la **rappresentazione trigonometrica** di un numero complesso.

## 2.3 Forma polare dei numeri complessi

Sia  $z$  un numero complesso scritto nella forma algebrica  $z = a + ib$  con  $a, b \in \mathbb{R}$ . Definiamo  $\rho$  come il modulo di  $z$  e  $\theta$  come l'angolo che il segmento dall'origine al punto  $(a, b)$  forma con l'asse  $x$ . Questi valori  $\rho$  e  $\theta$  sono chiamati *coordinate polari*:

$$\rho = |z| \quad (\text{modulo di } z)$$

$$\theta = \arg(z) \quad (\text{argomento di } z)$$

Utilizzando le proprietà dei triangoli rettangoli, otteniamo:

$$a = \rho \cos \theta \quad \text{e} \quad b = \rho \sin \theta$$

### Piano complesso

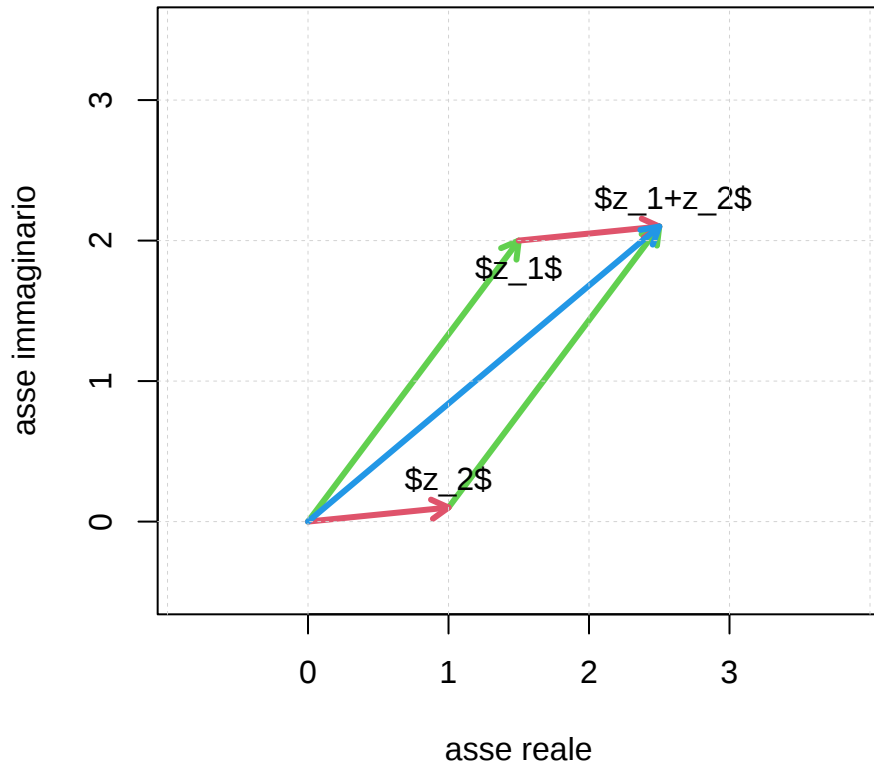


Figure 2.2: Somma di numeri complessi.

o, equivalente:

$$\rho = \sqrt{a^2 + b^2}, \quad \theta = \arctan\left(\frac{b}{a}\right)$$

Da ciò segue la *rappresentazione polare* di  $z$ :

$$z = a + ib = \rho \cos \theta + i \rho \sin \theta = \rho (\cos \theta + i \sin \theta)$$

Questa rappresentazione è illustrata nella figura Figure 2.3.

Warning: il pacchetto 'plotrix' è stato creato con R versione 4.5.2

#### 2.3.0.1 Osservazione

È importante notare che l'argomento di  $z$  è definito a meno di multipli interi di  $2\pi$ . In altre parole, variare l'angolo  $\theta$  di  $2\pi$  non modifica l'argomento di  $z$ . Mantenendo costante  $\rho$  e incrementando  $\theta$  di  $2\pi$ , si ottiene lo stesso numero complesso, poiché si completa un giro attorno alla circonferenza di centro  $(0,0)$  e raggio  $\rho$ , tornando così al punto di partenza.

Inoltre, dato che le funzioni  $\sin \theta$  e  $\cos \theta$  sono  $2\pi$ -periodiche, si ha che:

$$z = \rho (\cos \theta + i \sin \theta) = \rho (\cos(\theta + 2k\pi) + i \sin(\theta + 2k\pi)), \quad k \in \mathbb{Z}.$$

#### 2.3.0.2 Osservazione

Consideriamo le espansioni in serie di Taylor per le funzioni  $e^{i\theta}$ ,  $\cos \theta$ , e  $\sin \theta$ . Abbiamo:

### Piano complesso

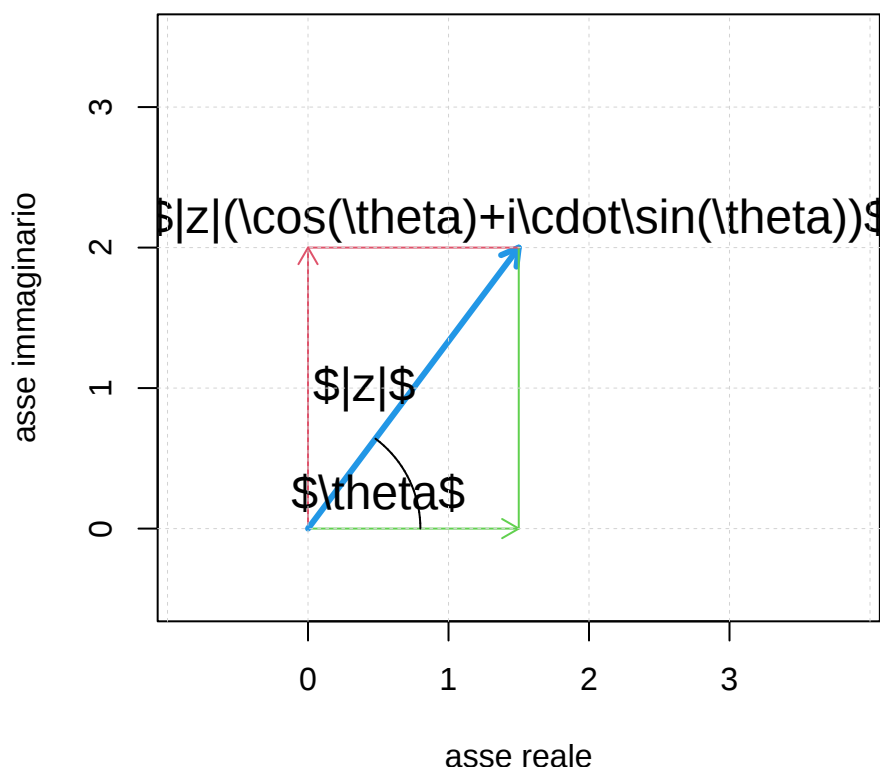


Figure 2.3: Forma polare di un numero complesso.

$$\begin{aligned}
 e^{i\theta} &= 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \dots \\
 &= 1 + i\theta - \frac{\theta^2}{2!} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + i\frac{\theta^5}{5!} + \dots \\
 &\quad + (-1)^k \frac{\theta^{2k}}{(2k)!} + i(-1)^k \frac{\theta^{2k+1}}{(2k+1)!} + \dots \\
 &= \left( 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \dots + (-1)^k \frac{\theta^{2k}}{(2k)!} + \dots \right) \\
 &\quad + i \left( \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \dots + (-1)^k \frac{\theta^{2k+1}}{(2k+1)!} + \dots \right) \\
 &= \cos \theta + i \sin \theta.
 \end{aligned}$$

Pertanto, il numero complesso  $\rho (\cos \theta + i \sin \theta)$  può essere scritto come:

$$\rho (\cos \theta + i \sin \theta) = \rho e^{i\theta}.$$

#### Esercizio

Consideriamo il numero complesso  $z = 1 - i$ . Determiniamo il suo modulo e il suo argomento. Dato che  $z = a + ib$  con  $a = 1$  e  $b = -1$ , calcoliamo:

- Il modulo di  $z$  è:

$$\rho = \sqrt{a^2 + b^2} = \sqrt{1^2 + (-1)^2} = \sqrt{2}.$$

- L'argomento di  $z$  è:

$$\theta = \arctan\left(\frac{b}{a}\right) = \arctan\left(\frac{-1}{1}\right) = -\frac{\pi}{4}.$$

Poiché l'argomento è tipicamente espresso in un intervallo da 0 a  $2\pi$ , aggiungiamo  $2\pi$  per ottenere un valore positivo:

$$\theta = -\frac{\pi}{4} + 2\pi = \frac{7\pi}{4}.$$

La forma polare di  $z$  è quindi:

$$z = \sqrt{2} \left( \cos\left(\frac{7\pi}{4}\right) + i \sin\left(\frac{7\pi}{4}\right) \right).$$

### Esercizio

Consideriamo il numero complesso dato da:

$$z = -2(\cos \alpha + i(-\sin \alpha)),$$

dove  $\alpha \in \mathbb{R}$ . Determiniamo il modulo e l'argomento di  $z$ .

Per esprimere  $z$  nella forma  $a + ib$ , calcoliamo:

- La parte reale è  $a = -2 \cos \alpha$ .
- La parte immaginaria è  $b = 2 \sin \alpha$ .

Il modulo di  $z$  è:

$$\begin{aligned} \rho &= \sqrt{a^2 + b^2} \\ &= \sqrt{(-2 \cos \alpha)^2 + (2 \sin \alpha)^2} \\ &= \sqrt{4 \cos^2 \alpha + 4 \sin^2 \alpha} \\ &= \sqrt{4(\cos^2 \alpha + \sin^2 \alpha)} \\ &= \sqrt{4} = 2. \end{aligned}$$

Per trovare l'argomento, consideriamo:

$$\begin{aligned} \theta &= \arctan\left(\frac{b}{a}\right) \\ &= \arctan\left(\frac{2 \sin \alpha}{-2 \cos \alpha}\right) \\ &= \arctan\left(-\frac{\sin \alpha}{\cos \alpha}\right) \\ &= \pi - \alpha. \end{aligned}$$

La forma polare di  $z$  è quindi:

$$z = 2(\cos(\pi - \alpha) + i \sin(\pi - \alpha)).$$

## 2.4 Prodotto di numeri complessi in forma polare

Per il prodotto di numeri complessi espressi in forma polare, la seguente proposizione è particolarmente utile e semplifica il calcolo delle potenze.

### 2.4.0.1 Proposizione

Il modulo del prodotto di due numeri complessi è uguale al prodotto dei loro moduli, mentre l'argomento del prodotto è uguale alla somma degli argomenti dei due numeri complessi.

Siano  $z_1$  e  $z_2$  due numeri complessi. Allora si ha:

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2|$$

e

$$\arg(z_1 \cdot z_2) = \arg(z_1) + \arg(z_2).$$

### 2.4.0.2 Dimostrazione

Consideriamo i numeri complessi  $z_1$  e  $z_2$  espressi nella loro forma trigonometrica:

$$z_1 = \rho(\cos \theta + i \sin \theta), \quad z_2 = \sigma(\cos \alpha + i \sin \alpha).$$

Il prodotto  $z_1 \cdot z_2$  si calcola come segue:

$$\begin{aligned} z_1 \cdot z_2 &= \rho \sigma (\cos \theta + i \sin \theta)(\cos \alpha + i \sin \alpha) \\ &= \rho \sigma [(\cos \theta \cos \alpha - \sin \theta \sin \alpha) + i(\cos \alpha \sin \theta + \sin \alpha \cos \theta)] \\ &= \rho \sigma [\cos(\theta + \alpha) + i \sin(\theta + \alpha)]. \end{aligned}$$

Nel passo finale abbiamo utilizzato le identità trigonometriche che esprimono il seno e il coseno della somma di due angoli in funzione dei seni e dei coseni dei singoli angoli.

### 2.4.0.3 Osservazione

Assumendo valide le proprietà della funzione esponenziale anche nel campo complesso, possiamo derivare la proposizione in modo più diretto. Consideriamo i numeri complessi in forma esponenziale:

$$z_1 = \rho e^{i\theta}, \quad z_2 = \sigma e^{i\alpha}.$$

Il prodotto di  $z_1$  e  $z_2$  è:

$$z_1 \cdot z_2 = \rho e^{i\theta} \cdot \sigma e^{i\alpha} = \rho \sigma e^{i(\theta + \alpha)}.$$

Quindi, il modulo del prodotto è il prodotto dei moduli e l'argomento è la somma degli argomenti, come affermato nella proposizione precedente.

Questa proposizione può essere estesa anche al calcolo del rapporto tra due numeri complessi, come descritto nel seguente corollario.

### 2.4.0.4 Corollario

Sia  $w = \frac{z_1}{z_2} \in \mathbb{C}$ . Allora:

$$|w| = \frac{|z_1|}{|z_2|}$$

e

$$\arg(w) = \arg(z_1) - \arg(z_2).$$

### 2.4.0.5 Dimostrazione

Scriviamo  $w$  come il prodotto di  $z_1$  e l'inverso di  $z_2$ :

$$w = z_1 \cdot \frac{1}{z_2} = z_1 \cdot \frac{\overline{z_2}}{|z_2|^2}.$$

Applicando la proposizione precedente, otteniamo:

$$\begin{aligned} |w| &= \left| z_1 \cdot \frac{\overline{z_2}}{|z_2|^2} \right| \\ &= |z_1| \cdot \left| \frac{\overline{z_2}}{|z_2|^2} \right| \\ &= |z_1| \cdot \frac{|\overline{z_2}|}{|z_2|^2} \\ &= |z_1| \cdot \frac{|z_2|}{|z_2|^2} \\ &= \frac{|z_1|}{|z_2|}. \end{aligned}$$

Per quanto riguarda l'argomento, abbiamo:

$$\arg(w) = \arg\left(z_1 \cdot \frac{\overline{z_2}}{|z_2|^2}\right) = \arg(z_1) + \arg\left(\frac{\overline{z_2}}{|z_2|^2}\right).$$

Notiamo che l'argomento di un numero complesso moltiplicato per un numero reale non cambia, quindi:

$$\arg\left(\frac{\overline{z_2}}{|z_2|^2}\right) = \arg(\overline{z_2}).$$

Dato che l'argomento di un numero complesso e del suo coniugato sono opposti:

$$\arg(\overline{z_2}) = -\arg(z_2),$$

segue che:

$$\arg(w) = \arg(z_1) - \arg(z_2).$$

Negli ultimi passaggi abbiamo tenuto conto del fatto che l'argomento di un numero complesso non varia se tale numero complesso viene moltiplicato per un qualsiasi numero reale, cioè

$$\arg(\alpha z) = \arg(z) \quad \forall z \in \mathbb{C} \quad \text{e} \quad \forall \alpha \in \mathbb{R}$$

e del fatto che

$$\arg(\overline{z}) = -\arg(z) \quad \forall z \in \mathbb{C}.$$

Tali verifiche sono lasciate per esercizio.

## 2.5 Formula di de Moivre

### 2.5.0.1 Proposizione

Sia  $\theta \in \mathbb{R}$  e  $n \in \mathbb{N}$ . Allora:

$$(\cos \theta + i \sin \theta)^n = \cos(n\theta) + i \sin(n\theta).$$

### 2.5.0.2 Dimostrazione

Consideriamo il numero complesso  $z = \cos \theta + i \sin \theta$ . Abbiamo che  $|z| = 1$ . Utilizzando le regole della moltiplicazione in forma polare, possiamo calcolare le potenze di  $z$ .

Per  $n = 2$ :

$$z^2 = (\cos \theta + i \sin \theta) \cdot (\cos \theta + i \sin \theta) = \cos^2 \theta - \sin^2 \theta + i(2 \cos \theta \sin \theta) = \cos(2\theta) + i \sin(2\theta).$$

Assumendo che la formula sia vera per  $n$ , cioè:

$$z^n = \cos(n\theta) + i \sin(n\theta),$$

dobbiamo dimostrare che è vera anche per  $n + 1$ . Consideriamo:

$$z^{n+1} = z^n \cdot z = (\cos(n\theta) + i \sin(n\theta)) \cdot (\cos \theta + i \sin \theta).$$

Utilizzando la formula per il prodotto di numeri complessi in forma trigonometrica, otteniamo:

$$z^{n+1} = \cos(n\theta) \cos \theta - \sin(n\theta) \sin \theta + i(\cos(n\theta) \sin \theta + \sin(n\theta) \cos \theta).$$

Semplificando, otteniamo:

$$z^{n+1} = \cos((n+1)\theta) + i \sin((n+1)\theta).$$

Quindi, per induzione, la formula è vera per ogni  $n \in \mathbb{N}$ .

### 2.5.0.3 Osservazione

Assumendo valide le proprietà della funzione esponenziale per l'elevamento a potenza anche nel campo complesso, la proposizione può essere ottenuta in modo più semplice. Se  $z = e^{i\theta}$ , allora:

$$z^n = (e^{i\theta})^n = e^{in\theta}.$$

### 2.5.1 Esempio di calcolo

Calcoliamo  $(1 + i)^{2000}$  utilizzando la formula di de Moivre. Prima, esprimiamo  $1 + i$  nella sua forma trigonometrica:

$$1 + i = \sqrt{2} \left( \cos \frac{\pi}{4} + i \sin \frac{\pi}{4} \right).$$

Applicando la formula di de Moivre:

$$\begin{aligned} (1 + i)^{2000} &= \left[ \sqrt{2} \left( \cos \frac{\pi}{4} + i \sin \frac{\pi}{4} \right) \right]^{2000} \\ &= (\sqrt{2})^{2000} \left( \cos \left( 2000 \cdot \frac{\pi}{4} \right) + i \sin \left( 2000 \cdot \frac{\pi}{4} \right) \right) \\ &= (\sqrt{2})^{2000} (\cos(500\pi) + i \sin(500\pi)) \\ &= 2^{1000} (\cos 0 + i \sin 0) \\ &= 2^{1000}. \end{aligned}$$

#### Esercizio

Dato un numero complesso  $z \in \mathbb{C}$  tale che  $|z| = 1$ , trova sul piano cartesiano i numeri complessi  $-z$ ,  $\bar{z}$ ,  $\frac{1}{z}$ ,  $iz$ , e  $z + iz$ .

#### Soluzione

Poiché  $z$  appartiene alla circonferenza unitaria, possiamo esprimerlo sia nella forma trigonometrica che in quella algebrica:

$$z = \cos \theta + i \sin \theta,$$

$$z = a + ib \quad \text{con} \quad \sqrt{a^2 + b^2} = 1.$$

Da questa espressione, otteniamo:

$$-z = -a - ib,$$

$$\bar{z} = a - ib,$$

$$\frac{1}{z} = \frac{\bar{z}}{|z|^2} = \bar{z}.$$

Il numero  $iz$  rappresenta il prodotto tra  $i$  e  $z$ . Utilizzando le regole del prodotto in forma polare, otteniamo:

$$|iz| = |i| \cdot |z| = 1 \cdot 1 = 1,$$

$$\arg(iz) = \arg(i) + \arg(z) = \frac{\pi}{2} + \theta.$$

Quindi,  $iz$  appartiene anch'esso alla circonferenza unitaria, con un argomento pari all'argomento di  $z$  aumentato di un angolo retto.

Per quanto riguarda  $z + iz$ , possiamo scrivere:

$$z + iz = z(1 + i).$$

Calcoliamo il modulo e l'argomento di  $z + iz$ :

$$|z + iz| = |z| \cdot |1 + i| = 1 \cdot \sqrt{2} = \sqrt{2},$$

$$\arg(z + iz) = \arg(z) + \arg(1 + i) = \theta + \frac{\pi}{4}.$$

Pertanto,  $z + iz$  appartiene a una circonferenza di raggio  $\sqrt{2}$  e il suo argomento è l'argomento di  $z$  aumentato di  $\frac{\pi}{4}$ .

### Esercizio

Trova i valori dei numeri complessi  $z$  che soddisfano l'equazione  $z^n = w$ , dove  $w$  è un numero complesso assegnato.

#### Soluzione

Siano  $z = \rho(\cos \theta + i \sin \theta)$  e  $w = \sigma(\cos \alpha + i \sin \alpha)$ , dove  $\rho$  e  $\theta$  sono incognite da determinare, mentre  $\sigma$  e  $\alpha$  sono dati. Imponendo la condizione  $z^n = w$ , otteniamo:

$$z^n = \rho^n(\cos(n\theta) + i \sin(n\theta)) = \sigma(\cos \alpha + i \sin \alpha).$$

Per uguagliare i moduli e gli argomenti, risolviamo:

$$\rho^n = \sigma \implies \rho = \sqrt[n]{\sigma},$$

$$n\theta = \alpha + 2k\pi \implies \theta_k = \frac{\alpha + 2k\pi}{n}, \quad k \in \mathbb{Z}.$$

Pertanto, tutte le soluzioni dell'equazione hanno lo stesso modulo  $\rho = \sqrt[n]{\sigma}$ . Gli angoli distintivi  $\theta$  sono determinati per valori  $k = 0, 1, \dots, n-1$ . Quindi, otteniamo  $n$  soluzioni distinte per l'equazione  $z^n = w$ :

$$\begin{aligned} z_0 &= \sqrt[n]{\sigma} \left( \cos \left( \frac{\alpha}{n} \right) + i \sin \left( \frac{\alpha}{n} \right) \right), \\ z_1 &= \sqrt[n]{\sigma} \left( \cos \left( \frac{\alpha + 2\pi}{n} \right) + i \sin \left( \frac{\alpha + 2\pi}{n} \right) \right), \\ z_2 &= \sqrt[n]{\sigma} \left( \cos \left( \frac{\alpha + 4\pi}{n} \right) + i \sin \left( \frac{\alpha + 4\pi}{n} \right) \right), \\ &\vdots \\ z_{n-1} &= \sqrt[n]{\sigma} \left( \cos \left( \frac{\alpha + 2(n-1)\pi}{n} \right) + i \sin \left( \frac{\alpha + 2(n-1)\pi}{n} \right) \right). \end{aligned}$$

Queste soluzioni sono chiamate **radici n-sime** di  $w$ .

### Esercizio

Trovare le soluzioni in  $\mathbb{C}$  dell'equazione  $z^3 = 1$ .

#### Soluzione

Con le notazioni adottate, abbiamo

$$1 = 1(\cos 0 + i \sin 0), \quad \rho = \sqrt[3]{1} = 1, \quad \theta_k = \frac{0 + 2k\pi}{3}, \quad k = 0, 1, 2,$$

per cui i valori distinti delle radici si ottengono in corrispondenza dei valori degli argomenti  $\theta_0 = 0, \theta_1 = \frac{2\pi}{3}, \theta_2 = \frac{4\pi}{3}$ :

$$\begin{aligned} z_0 &= \cos 0 + i \sin 0 = 1, \\ z_1 &= \cos \frac{2\pi}{3} + i \sin \frac{2\pi}{3} = -\frac{1}{2} + i \frac{\sqrt{3}}{2}, \\ z_2 &= \cos \frac{4\pi}{3} + i \sin \frac{4\pi}{3} = -\frac{1}{2} - i \frac{\sqrt{3}}{2}. \end{aligned}$$

#### Verifica delle Soluzioni

##### 1. Verifica di $z_0$

$$z_0 = 1$$

Calcoliamo  $z_0^3$ :

$$z_0^3 = 1^3 = 1$$

Che è uguale a 1, quindi  $z_0 = 1$  è una soluzione.

##### 2. Verifica di $z_1$

$$z_1 = -\frac{1}{2} + i \frac{\sqrt{3}}{2}$$

Calcoliamo  $z_1^3$ . Utilizziamo la forma trigonometrica:

$$z_1 = \cos \left( \frac{2\pi}{3} \right) + i \sin \left( \frac{2\pi}{3} \right)$$

Dove  $\cos\left(\frac{2\pi}{3}\right) = -\frac{1}{2}$  e  $\sin\left(\frac{2\pi}{3}\right) = \frac{\sqrt{3}}{2}$ .

Usando la formula di De Moivre:

$$z_1^3 = \left( \cos\left(\frac{2\pi}{3}\right) + i \sin\left(\frac{2\pi}{3}\right) \right)^3 = \cos(2\pi) + i \sin(2\pi) = 1$$

Quindi  $z_1^3 = 1$ , e  $z_1$  è una soluzione.

### 3. Verifica di $z_2$

$$z_2 = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$$

Calcoliamo  $z_2^3$ . Anche in questo caso, utilizziamo la forma trigonometrica:

$$z_2 = \cos\left(\frac{4\pi}{3}\right) + i \sin\left(\frac{4\pi}{3}\right)$$

Dove  $\cos\left(\frac{4\pi}{3}\right) = -\frac{1}{2}$  e  $\sin\left(\frac{4\pi}{3}\right) = -\frac{\sqrt{3}}{2}$ .

Usando la formula di De Moivre:

$$z_2^3 = \left( \cos\left(\frac{4\pi}{3}\right) + i \sin\left(\frac{4\pi}{3}\right) \right)^3 = \cos(4\pi) + i \sin(4\pi) = 1$$

Quindi  $z_2^3 = 1$ , e  $z_2$  è una soluzione.

In conclusione, tutte le soluzioni verificate soddisfano l'equazione  $z^3 = 1$ . Le soluzioni  $z_0$ ,  $z_1$ , e  $z_2$  sono corrette.

### Esercizio

Trovare le soluzioni in  $\mathbb{C}$  dell'equazione  $z^5 = 1 - i$ .

#### Soluzione

Innanzitutto, esprimiamo il numero complesso  $1 - i$  nella forma trigonometrica.

Calcoliamo il modulo e l'argomento di  $1 - i$ :

$$|1 - i| = \sqrt{1^2 + (-1)^2} = \sqrt{2}$$

L'argomento di  $1 - i$  è dato da:

$$\arg(1 - i) = \arctan\left(\frac{-1}{1}\right) = -\frac{\pi}{4}$$

Poiché l'argomento deve essere compreso nell'intervallo  $[0, 2\pi)$ , possiamo scrivere:

$$\arg(1 - i) = \frac{7\pi}{4}$$

Pertanto, possiamo esprimere  $1 - i$  come:

$$1 - i = \sqrt{2} \left( \cos \frac{7\pi}{4} + i \sin \frac{7\pi}{4} \right)$$

Ora, per trovare le soluzioni dell'equazione  $z^5 = 1 - i$ , dobbiamo trovare le radici quinta di  $1 - i$ . Utilizziamo la formula per le radici  $n$ -esime di un numero complesso:

$$\rho = \sqrt[5]{\sqrt{2}} = \sqrt[10]{2}$$

E per gli argomenti delle radici, abbiamo:

$$\theta_k = \frac{\frac{7\pi}{4} + 2k\pi}{5}, \quad k = 0, 1, 2, 3, 4$$

Quindi le soluzioni sono:

$$z_k = \sqrt[10]{2} (\cos \theta_k + i \sin \theta_k)$$

dove

$$\theta_k = \frac{\frac{7\pi}{4} + 2k\pi}{5}$$

per  $k = 0, 1, 2, 3, 4$ .

### Esercizio

Scrivere nella forma  $x + iy$  il numero complesso  $z = \frac{1+i}{1-i}$ .

#### Soluzione

Per scrivere il numero complesso  $z$  nella forma  $x + iy$ , iniziamo a semplificare la frazione:

$$z = \frac{1+i}{1-i}$$

Moltiplichiamo il numeratore e il denominatore per il coniugato del denominatore,  $1 + i$ :

$$z = \frac{(1+i)(1+i)}{(1-i)(1+i)}$$

Calcoliamo il denominatore:

$$(1-i)(1+i) = 1^2 - (i)^2 = 1 - (-1) = 1 + 1 = 2$$

Calcoliamo il numeratore:

$$(1+i)(1+i) = 1 + i + i + i^2 = 1 + 2i - 1 = 2i$$

Quindi:

$$z = \frac{2i}{2} = i$$

In forma algebrica, il numero complesso  $z$  è quindi:

$$z = i$$

Per confermare il risultato, possiamo anche considerare la forma trigonometrica dei numeri complessi nel numeratore e nel denominatore. Scriviamo:

$$1 + i = \sqrt{2} \left( \cos \frac{\pi}{4} + i \sin \frac{\pi}{4} \right)$$

$$1 - i = \sqrt{2} \left( \cos \frac{-\pi}{4} + i \sin \frac{-\pi}{4} \right)$$

La divisione di due numeri complessi in forma polare dà:

$$\begin{aligned}\frac{1+i}{1-i} &= \frac{\sqrt{2} \left( \cos \frac{\pi}{4} + i \sin \frac{\pi}{4} \right)}{\sqrt{2} \left( \cos \frac{-\pi}{4} + i \sin \frac{-\pi}{4} \right)} \\ &= \frac{\cos \frac{\pi}{4} + i \sin \frac{\pi}{4}}{\cos \frac{-\pi}{4} + i \sin \frac{-\pi}{4}} \\ &= \cos \frac{\pi}{4} - \frac{-\pi}{4} + i \sin \frac{\pi}{4} - \frac{-\pi}{4} \\ &= \cos \frac{\pi}{2} + i \sin \frac{\pi}{2} = i\end{aligned}$$

Quindi, confermiamo che:

$$z = i$$

### 3. Esercizi Proposti

- Scrivere i seguenti numeri complessi nella forma  $x + iy$  con  $x, y \in \mathbb{R}$

$$(-1 + 3i)^{-1};$$

$$(1 + i)(1 - i);$$

$$(1 + i)(2 - i);$$

$$(7 + \pi i)(\pi + i);$$

$$(2i + 1)\pi i;$$

$$(i + 1)(i - 2)(i + 3);$$

$$1 + i^3;$$

$$\frac{1 + 3i}{1 + i};$$

$$\frac{1 - 2i}{1 + i^3}.$$

- Scrivere nella forma  $x + iy$  l'inverso dei seguenti numeri complessi:

$$2 + i;$$

$$\frac{1}{i - 1};$$

$$i;$$

$$i^5 + i + 1.$$

- Trovare il luogo geometrico dei numeri complessi  $z$  tali che:

$$|z| = 1;$$

$$z + \bar{z} = 1;$$

$$|z| < 1;$$

$$|z - i| \leq 1;$$

$$|z - 2 + i| < 1;$$

$$z\bar{z} = z;$$

$$z - |z| = \bar{z};$$

$$|z|^2 - 2|z| - 3 = 0;$$

$$\arg(z) = \frac{\pi}{4}.$$

- Determinare modulo, argomento e rappresentazione trigonometrica dei numeri complessi  
 $z = (6 + i6\sqrt{3})^{12}$ ;  
 $\left(\cos \frac{\pi}{4} - i \sin \frac{\pi}{4}\right)^{20}$ .

- Esprimere in forma trigonometrica l'inverso del numero complesso  
 $z = 2 \left(\cos \frac{\pi}{4} + i \sin \frac{\pi}{4}\right) \cdot \frac{1}{3} \left(\cos \frac{\pi}{4} + i \sin \frac{\pi}{4}\right)$

- Determinare le soluzioni complesse delle seguenti equazioni:  
 $z^4 = i$ ;  
 $z^4 = 1 + i$ ;  
 $z^4 = 1$ ;  
 $z^3 = 1 - i\sqrt{3}$

- Per ognuno dei seguenti  $z \in \mathbb{C}$  calcolare  $\bar{z}$ ,  $z\bar{z}$ , e  $1/z$ .
- $z = 8 - 3i$   
 Soluzione:  $\bar{z} = 8 + 3i$ ;  $z\bar{z} = 73$ ;  $z^{-1} = \frac{8}{73} + \frac{3}{73}i$
- $z = 1 - 9i$   
 Soluzione:  $\bar{z} = 1 + 9i$ ;  $z\bar{z} = 82$ ;  $z^{-1} = \frac{1}{82} + \frac{9}{82}i$
- $z = 6 - 7i$   
 Soluzione:  $\bar{z} = 6 + 7i$ ;  $z\bar{z} = 85$ ;  $z^{-1} = \frac{6}{85} + \frac{7}{85}i$
- $z = -5 - 9i$   
 Soluzione:  $\bar{z} = -5 + 9i$ ;  $z\bar{z} = 106$ ;  $z^{-1} = -\frac{5}{106} + \frac{9}{106}i$
- $z = 7 + 6i$ .  
 Soluzione:  $\bar{z} = 7 - 6i$ ;  $z\bar{z} = 85$ ;  $z^{-1} = \frac{7}{85} - \frac{6}{85}i$

- Trovare  $x, y \in \mathbb{R}$  tali che:  
 $(-5 - 7i)x + (-1 + 2i)y = 2 + \frac{48}{5}i$ ;  
 Soluzione:  $x = -\frac{4}{5}$ ;  $y = 2$
- $(-4 + 1i)x + (8 - 5i)y = -\frac{22}{3} + \frac{19}{3}i$ ;  
 Soluzione:  $x = -\frac{7}{6}$ ;  $y = -\frac{3}{2}$
- $(-3 - 4i)x + (2 + 2i)y = -\frac{51}{5} - \frac{58}{5}i$ ;  
 Soluzione:  $x = \frac{7}{5}$ ;  $y = -3$
- $(9 + 7i)x + (6 + 8i)y = -12 - 26i$ ;  
 Soluzione:  $x = 2$ ;  $y = -5$
- $(-3 + 3i)x + (-9 + 1i)y = \frac{33}{5} - \frac{1}{5}i$ .

Soluzione:  $x = \frac{1}{5}; y = -\frac{4}{5}$

- Dimostrare l'identità  
 $|x+y|^2 + |x-y|^2 = 2(|x|^2 + |y|^2)$   
 si provi a darne una interpretazione geometrica.

- Sia  $z \in \mathbb{C}$  tale che  $|z| = 1$ . Per quale valore reale  $t$  si ha
  - $z = \bar{z} \left( \frac{1+ti}{1-ti} \right)$ , supposto  $z \neq i$ ?
  - $z = \frac{1+ti}{1-ti}$ , supposto  $z \neq -1$ ?
 Cosa succederebbe nel primo caso se fosse  $z = i$  e nel secondo caso se fosse  $z = -1$ ?

- Sia dato il numero complesso  $z = 1 + i$ . Calcolare utilizzando sia la forma algebrica che la forma polare  $z^n$  per  $n = 2, \dots, 10$ . Quanto vale  $z^{25}$ ?

- Siano  $z, w \in \mathbb{C}$ . In quale caso il modulo della somma
  - è uguale alla somma dei moduli degli addendi?
  - è uguale al valore assoluto della differenza dei moduli degli addendi?

- Utilizzando la forma polare, determinare  $z \in \mathbb{C}$  tale che
  - $|z| - z = 1 + 2i$ ;
  - $|z| + z = 2 + i$ .
 Dopo aver determinato il risultato, si proceda alla verifica.

- Sia dato il numero complesso  $z = 1 + \cos \alpha + i \sin \alpha$ , dove  $\alpha \in \mathbb{R}$ . Calcolare
  - $z^n$  per  $n \in \mathbb{N}$ ;
 Suggerimento: applicare la formula di de Moivre

- Sia  $z \in \mathbb{C}$  tale che  $z + \frac{1}{z} = 2 \cos \theta$  dove  $\theta = \arg(z)$ ; calcolare  
 $z^n + \frac{1}{z^n}$ ;  
 Suggerimento: applicare la formula di de Moivre  
 Soluzione:  $|z|$ .

- Utilizzando la formula di de Moivre, calcolare  
 $\cos\left(\frac{2\pi}{11}\right) + \cos\left(3\frac{2\pi}{11}\right) + \cos\left(5\frac{2\pi}{11}\right) + \cos\left(7\frac{2\pi}{11}\right) + \cos\left(9\frac{2\pi}{11}\right)$   
 Soluzione:  $-\frac{1}{2}$

- Utilizzando la formula di de Moivre, calcolare  
 $\cos\left(\frac{\pi}{11}\right) + \cos\left(3\frac{\pi}{11}\right) + \cos\left(5\frac{\pi}{11}\right) + \cos\left(7\frac{\pi}{11}\right) + \cos\left(9\frac{\pi}{11}\right)$

Soluzione:  $\frac{1}{2}$

< qrcode <https://bery2001.quarto.pub/complessi/> >

## 4. Matrici e vettori

### 4.1 Notazioni

In questi appunti, i vettori saranno indicati con lettere in grassetto minuscole, ad esempio:

$$\mathbf{a}, \quad \mathbf{b}, \quad \mathbf{c},$$

sono vettori. Le componenti dei vettori saranno indicate con la stessa lettera del vettore in corsivo, ad esempio:

$$a_1, \quad a_2, \quad c_i, \quad b_j,$$

sono componenti dei vettori  $\mathbf{a}$ ,  $\mathbf{b}$  e  $\mathbf{c}$ . Le matrici saranno indicate con lettere in grassetto maiuscole, ad esempio:

$$\mathbf{A}, \quad \mathbf{B}, \quad \mathbf{C},$$

sono matrici. Le componenti delle matrici saranno indicate con la stessa lettera in corsivo, ad esempio:

$$A_{1,3}, \quad B_{i,3}, \quad C_{j,2},$$

sono componenti delle matrici  $\mathbf{A}$ ,  $\mathbf{B}$  e  $\mathbf{C}$ . Gli scalari saranno normalmente indicati con lettere greche, ad esempio:

$$\alpha, \quad \beta, \quad \gamma, \quad \dots$$

Con il simbolo  $\mathbb{K}$  indicheremo sia il campo dei numeri reali  $\mathbb{R}$  sia il campo dei numeri complessi  $\mathbb{C}$ . Questo consente di usare  $\mathbb{K}$  al posto di  $\mathbb{R}$  o  $\mathbb{C}$ , evitando così duplicazioni nelle definizioni.

### 4.2 Matrici

**Definition 4.1 — Matrice.** Una matrice  $\mathbf{A} \in \mathbb{K}^{m \times n}$  è un insieme di  $m \times n$  numeri reali o complessi organizzati in  $m$  righe e  $n$  colonne. Ad esempio:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ -2 & 2 & 0 \end{pmatrix},$$

è una matrice  $2 \times 3$ , ossia una matrice con 2 righe e 3 colonne. Se la matrice  $\mathbf{A}$  è a valori

reali, si indica  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . Analogamente, se è a valori complessi, si usa  $\mathbf{A} \in \mathbb{C}^{m \times n}$ . In generale,  $\mathbf{A} \in \mathbb{K}^{m \times n}$  può rappresentare sia matrici a valori reali che complessi. Se non diversamente specificato, si assume che le matrici e i vettori siano a valori reali.

**Definition 4.2 — Matrice quadrata.** Una matrice si dice *matrice quadrata* se ha lo stesso numero di righe e di colonne. Le due matrici più semplici sono:

- La **matrice nulla**, indicata con  $\mathbf{0}$ , che ha tutti gli elementi uguali a zero:

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix},$$

- La **matrice identità**, indicata con  $\mathbf{I}$ , che è una matrice quadrata con tutti gli elementi nulli tranne quelli sulla diagonale principale, che sono tutti uguali a 1:

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

**Definition 4.3 — Matrici triangolari.** Una matrice quadrata  $\mathbf{A}$  è detta *triangolare superiore* se:

$$A_{i,j} = 0, \quad \text{per ogni } i > j,$$

dove  $i$  è l'indice di riga e  $j$  è l'indice di colonna. Questa matrice può essere visualizzata come:

$$\mathbf{A} = \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & * \end{pmatrix}$$

dove gli elementi contrassegnati con 0 sono detti *zeri strutturali*, mentre  $*$  rappresenta gli elementi che possono assumere valori diversi da zero, definiti come *non-zero*<sup>a</sup>. Analogamente, una matrice triangolare inferiore ha tutti gli zeri sopra la diagonale principale.

<sup>a</sup>Jacques Philippe Marie Binet (1786-1856)

#### Esempio

Le matrici

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

sono tutte triangolari superiori, la matrice  $\mathbf{C}$  è anche triangolare inferiore.

## 4.3 Vettori

**Definition 4.4 — Vettori riga e colonna.** Una matrice di dimensioni  $1 \times n$  è chiamata *vettore riga* e ha dimensione  $n$ . Ad esempio, un vettore riga con  $n$  elementi può essere rappresentato come:

$$\mathbf{a} = (a_1 \quad a_2 \quad \cdots \quad a_n).$$

Analogamente, una matrice di dimensioni  $m \times 1$  è chiamata *vettore colonna* e ha dimensione  $m$ . Un vettore colonna con  $m$  elementi può essere rappresentato come:

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Nel seguito, considereremo principalmente vettori colonna. Pertanto, quando faremo riferimento a un vettore senza specificare se si tratta di un vettore riga o colonna, intendiamo sempre un vettore colonna.

#### Esempio

Le due seguenti matrici:

$$\mathbf{a} = (1 \quad 1 \quad 2 \quad 3), \quad \mathbf{b} = \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix},$$

si possono considerare vettori:  $\mathbf{a}$  vettore riga di dimensione 4 e  $\mathbf{b}$  vettore colonna di dimensione 3.

Quando si indica una componente di un vettore riga o colonna, l'indice di riga o colonna è spesso omesso. Ad esempio:

- Per un vettore riga, la componente nella posizione 2 si scrive semplicemente come  $a_2$ , senza indicare l'indice di riga. Quindi, per un vettore riga  $\mathbf{a}$ , la seconda componente è  $a_2$ .
- Analogamente, per un vettore colonna, la componente nella posizione  $i$  è scritta come  $b_i$ , senza specificare l'indice di colonna.

In generale, per semplicità, si tende a usare solo l'indice della componente senza menzionare esplicitamente la riga o la colonna.

## 4.4 Somma, differenza e prodotto per uno scalare

**Definition 4.5 — Somma di matrici.** Date due matrici  $\mathbf{A}$  e  $\mathbf{B}$ , entrambe appartenenti allo spazio  $\mathbb{K}^{m \times n}$  (cioè matrici con  $m$  righe e  $n$  colonne), possiamo definire la loro somma  $\mathbf{C}$  come segue:

$$\mathbf{C} = \mathbf{A} + \mathbf{B}$$

La matrice  $\mathbf{C}$  è anch'essa una matrice in  $\mathbb{K}^{m \times n}$  e ogni suo elemento è dato dalla somma degli elementi corrispondenti di  $\mathbf{A}$  e  $\mathbf{B}$ . Più precisamente, per ogni elemento  $C_{i,j}$  della matrice  $\mathbf{C}$ , abbiamo:

$$C_{i,j} = A_{i,j} + B_{i,j}, \quad \text{per ogni} \quad \begin{matrix} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n. \end{matrix}$$

Analogamente, la differenza tra due matrici  $\mathbf{A}$  e  $\mathbf{B}$  è definita come:

$$\mathbf{D} = \mathbf{A} - \mathbf{B}$$

dove ogni elemento  $D_{i,j}$  della matrice  $\mathbf{D}$  è dato dalla differenza degli elementi corrispondenti di  $\mathbf{A}$  e  $\mathbf{B}$ :

$$D_{i,j} = A_{i,j} - B_{i,j}, \quad \text{per ogni} \quad \begin{matrix} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n. \end{matrix}$$

#### Esempio

Date le matrici  $2 \times 3$

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ -2 & 2 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & -1 & 1 \\ 4 & 3 & 2 \end{pmatrix},$$

otteniamo

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} 1 & 1 & 4 \\ 2 & 5 & 2 \end{pmatrix}, \quad \mathbf{A} - \mathbf{B} = \begin{pmatrix} 1 & 3 & 2 \\ -6 & -1 & -2 \end{pmatrix}.$$

**Definition 4.6 — Prodotto matrice scalare.** Data una matrice  $\mathbf{A} \in \mathbb{K}^{m \times n}$  e uno scalare  $\alpha \in \mathbb{K}$ , possiamo definire il prodotto dello scalare  $\alpha$  con la matrice  $\mathbf{A}$  come segue:

$$\mathbf{B} = \alpha \mathbf{A}$$

La matrice  $\mathbf{B}$  è anch'essa una matrice in  $\mathbb{K}^{m \times n}$  e ogni elemento di  $\mathbf{B}$  è ottenuto moltiplicando l'elemento corrispondente di  $\mathbf{A}$  per lo scalare  $\alpha$ . In altre parole, per ogni elemento  $B_{i,j}$  della matrice  $\mathbf{B}$ , abbiamo:

$$B_{i,j} = \alpha A_{i,j}, \quad \text{per ogni} \quad \begin{matrix} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n. \end{matrix}$$

#### Esempio

Data la matrice  $\mathbf{A} \in \mathbb{R}^{2 \times 3}$

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ -2 & 2 & 0 \end{pmatrix},$$

otteniamo

$$2\mathbf{A} = \begin{pmatrix} 2 & 4 & 6 \\ -4 & 4 & 0 \end{pmatrix}, \quad -2.5\mathbf{A} = \begin{pmatrix} -2.5 & -5 & -7.5 \\ 5 & -5 & 0 \end{pmatrix}.$$

**Definition 4.7 — Valore assoluto per componenti.** Data una matrice  $\mathbf{A} \in \mathbb{K}^{m \times n}$ , possiamo definire la matrice  $|\mathbf{A}|$  come la matrice i cui elementi sono i valori assoluti (o i moduli, se  $\mathbf{A}$  è a valori complessi) degli elementi di  $\mathbf{A}$ . In altre parole, se  $\mathbf{B} = |\mathbf{A}|$ , allora ogni elemento  $B_{i,j}$  della matrice  $\mathbf{B}$  è dato dal valore assoluto dell'elemento corrispondente  $A_{i,j}$  della matrice  $\mathbf{A}$ . Formalmente:

$$\mathbf{B} = |\mathbf{A}| \quad \text{implica} \quad B_{i,j} = |A_{i,j}|, \quad \text{per ogni} \quad \begin{matrix} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n. \end{matrix}$$

Dove  $|x|^a$  denota il valore assoluto di un numero reale  $x$  o il modulo di un numero complesso<sup>b</sup>.

<sup>a</sup>James Joseph Sylvester (1814-1897)

<sup>b</sup>Jacques Philippe Marie Binet (1786-1856)

### Esempio

Consideriamo le seguenti matrici:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & -2 & 1 \\ -2 & -3 & 2 \end{pmatrix},$$

$$\mathbf{B} = \begin{pmatrix} 1+i & 3+4i & 3 \\ 1 & -2 & 1 \\ -2 & -3-i & 1 \end{pmatrix},$$

Per ottenere le matrici dei valori assoluti (o moduli, nel caso di numeri complessi), calcoliamo:

$$|\mathbf{A}| = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & 3 & 2 \end{pmatrix},$$

$$|\mathbf{B}| = \begin{pmatrix} \sqrt{2} & 5 & 3 \\ 1 & 2 & 1 \\ 2 & 2 & 1 \end{pmatrix}.$$

Dove, per ogni elemento, abbiamo utilizzato il valore assoluto per i numeri reali e il modulo per i numeri complessi. Ad esempio:

- Per la matrice  $\mathbf{A}$ , il valore assoluto di ogni elemento è il numero stesso se è positivo o il suo valore assoluto se è negativo.
- Per la matrice  $\mathbf{B}$ , abbiamo calcolato il modulo di ogni elemento complesso. Per esempio, il modulo di  $1+i$  è  $\sqrt{2}$ , e il modulo di  $3+4i$  è 5.

**Definition 4.8 — Confronto di matrici.** Date due matrici  $\mathbf{A}$  e  $\mathbf{B}$  appartenenti a  $\mathbb{R}^{m \times n}$ , la notazione  $\mathbf{A} \geq \mathbf{B}$  indica che ogni elemento della matrice  $\mathbf{A}$  è maggiore o uguale al corrispondente elemento della matrice  $\mathbf{B}$ . Formalmente, abbiamo:

$$A_{i,j} \geq B_{i,j} \quad \text{per ogni} \quad \begin{matrix} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n. \end{matrix}$$

Questa definizione si applica anche ai vettori riga e colonna come casi particolari. In particolare:

- Un vettore riga di dimensione  $1 \times n$  può essere visto come una matrice  $1 \times n$ .
- Un vettore colonna di dimensione  $m \times 1$  può essere visto come una matrice  $m \times 1$ .

Quindi, per i vettori, la notazione  $\mathbf{a} \geq \mathbf{b}$  implica che ogni elemento del vettore  $\mathbf{a}$  è maggiore o uguale al corrispondente elemento del vettore  $\mathbf{b}$ .

In modo analogo con la scrittura  $\mathbf{A} > \mathbf{B}$  si intende

$$A_{i,j} > B_{i,j} \quad \text{per ogni} \quad \begin{matrix} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n. \end{matrix}$$

## Esempio

Date le matrici

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix},$$

abbiamo  $\mathbf{A} \geq \mathbf{B}$  e  $\mathbf{C} \geq \mathbf{B}$  mentre  $\mathbf{A}$  e  $\mathbf{C}$  non sono confrontabili.

## Esempio

Date le matrici

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

abbiamo  $\mathbf{A} > \mathbf{B}$ .

## 4.5 Operazioni con le matrici

**Definition 4.9 — Prodotto di matrici.** Dato che abbiamo due matrici  $\mathbf{A} \in \mathbb{K}^{m \times n}$  e  $\mathbf{B} \in \mathbb{K}^{n \times p}$ , possiamo definire il loro prodotto  $\mathbf{C} = \mathbf{AB}$  come una matrice  $\mathbb{K}^{m \times p}$ , dove ogni elemento  $C_{i,j}$  è calcolato come segue:

$$C_{i,j} = \sum_{k=1}^n A_{i,k} B_{k,j}, \quad i = 1, 2, \dots, m \quad \text{e} \quad j = 1, 2, \dots, p.$$

In altre parole,  $C_{i,j}$  è ottenuto sommando i prodotti degli elementi della  $i$ -esima riga della matrice  $\mathbf{A}$  con gli elementi corrispondenti della  $j$ -esima colonna della matrice  $\mathbf{B}$ . Questo processo è noto come *prodotto scalare* (verrà discusso in seguito in dettaglio) tra la  $i$ -esima riga di  $\mathbf{A}$  e la  $j$ -esima colonna di  $\mathbf{B}$ .

## Esempio

Date le matrici  $2 \times 3$  e  $3 \times 1$

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ -2 & 2 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix},$$

otteniamo

$$\mathbf{AB} = \begin{pmatrix} 1 & 2 & 3 \\ -2 & 2 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

## Osservazione

Non tutte le coppie di matrici possono essere moltiplicate tra loro. Perché il prodotto tra due matrici  $\mathbf{A}$  e  $\mathbf{B}$  sia definito, le matrici devono essere *compatibili*. In particolare, il prodotto  $\mathbf{AB}$  è possibile solo se il numero di colonne di  $\mathbf{A}$  è uguale al numero di righe di  $\mathbf{B}$ .

Lo stesso criterio vale per il prodotto inverso,  $\mathbf{BA}$ .

È importante notare che può accadere che il prodotto  $\mathbf{AB}$  sia definito, ma non il prodotto  $\mathbf{BA}$ ,

o viceversa. Inoltre, anche quando entrambi i prodotti sono definiti, le matrici risultanti  $\mathbf{AB}$  e  $\mathbf{BA}$  possono avere dimensioni diverse, con un diverso numero di righe e colonne. Nel prosieguo, ogni volta che si fa riferimento al prodotto tra matrici, sarà sempre sottinteso che esse sono compatibili, anche se non esplicitamente dichiarato.

#### Esempio

Un esempio utile per mostrare che la moltiplicazione tra matrici non è commutativa può essere fatto considerando le seguenti matrici  $2 \times 2$ :

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

In questo caso, possiamo calcolare sia il prodotto  $\mathbf{AB}$  che il prodotto  $\mathbf{BA}$ . Tuttavia, notiamo che  $\mathbf{AB} \neq \mathbf{BA}$ . Infatti, eseguendo i calcoli otteniamo:

$$\mathbf{AB} = \begin{pmatrix} 2 & 3 \\ -1 & 0 \end{pmatrix}, \quad \mathbf{BA} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}.$$

Questo esempio dimostra chiaramente che l'ordine con cui moltiplichiamo le matrici influisce sul risultato, mostrando che la moltiplicazione di matrici non è commutativa.

#### Esempio

Consideriamo un vettore riga  $\mathbf{a}$  e un vettore colonna  $\mathbf{b}$  definiti come:

$$\mathbf{a} = (1 \quad 2 \quad 3), \quad \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Il prodotto  $\mathbf{ab}$  è una matrice  $1 \times 1$ , che possiamo interpretare come uno scalare:

$$\mathbf{ab} = (1 \quad 2 \quad 3) \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 1 \cdot 1 + 2 \cdot 0 + 3 \cdot 1 = 4.$$

D'altra parte, il prodotto  $\mathbf{ba}$  produce una matrice  $3 \times 3$ :

$$\mathbf{ba} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} (1 \quad 2 \quad 3) = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 1 & 2 & 3 \end{pmatrix}.$$

Questo esempio mostra come il prodotto di un vettore riga per un vettore colonna risulti in uno scalare, mentre il prodotto di un vettore colonna per un vettore riga dia luogo a una matrice di dimensioni maggiori.

#### Osservazione

Per le matrici non vale la regola di annullamento. Questo significa che, se due matrici  $\mathbf{A}$  e  $\mathbf{B}$  soddisfano  $\mathbf{AB} = \mathbf{0}$ , non è necessariamente vero che  $\mathbf{A} = \mathbf{0}$  o  $\mathbf{B} = \mathbf{0}$ . Un esempio concreto può chiarire questo concetto:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

Calcolando il prodotto otteniamo:

$$\mathbf{AB} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{BA} = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}.$$

In questo esempio, il prodotto  $\mathbf{AB}$  è la matrice nulla, ma né  $\mathbf{A}$  né  $\mathbf{B}$  sono matrici nulle. Inoltre, osserviamo che  $\mathbf{AB} \neq \mathbf{BA}$ , il che conferma ancora una volta che la moltiplicazione di matrici non è commutativa.

### Esempio

Nell'insieme delle matrici, la matrice identità  $\mathbf{I}$  rappresenta l'elemento neutro rispetto alla moltiplicazione. In altre parole, per ogni matrice  $\mathbf{A} \in \mathbb{K}^{m \times n}$  e la matrice identità  $\mathbf{I} \in \mathbb{K}^{n \times n}$ , si ha:

$$\mathbf{AI} = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & A_{m-1,n-1} & A_{m-1,n} \\ A_{m,1} & \cdots & A_{m,n-1} & A_{m,n} \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Qui, gli elementi della matrice identità  $\mathbf{I}$  sono definiti dal simbolo di Kronecker<sup>1</sup>  $\delta_{i,j}$ , dove:

$$\delta_{i,j} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}.$$

Quindi, possiamo esprimere la moltiplicazione come:

$$(\mathbf{AI})_{i,j} = \sum_{k=1}^n A_{i,k} I_{k,j} = \sum_{k=1}^n A_{i,k} \delta_{k,j} = A_{i,j} \delta_{j,j} = A_{i,j}.$$

In modo del tutto analogo, si può dimostrare che  $\mathbf{IA} = \mathbf{A}$ , dove in questo caso  $\mathbf{I} \in \mathbb{K}^{m \times m}$ .

**Definition 4.10 — Inversa di una matrice.** La matrice inversa di una matrice *quadrata*  $\mathbf{A}$  è definita come la matrice quadrata  $\mathbf{B}$  (se esiste) tale che soddisfi la relazione:

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I},$$

dove  $\mathbf{I}$  è la matrice identità. Se una matrice  $\mathbf{A}$  possiede un'inversa, questa viene indicata con  $\mathbf{A}^{-1}$ , e la matrice  $\mathbf{A}$  si dice *invertibile*.

### Esempio

Consideriamo la matrice  $\mathbf{A} \in \mathbb{K}^{2 \times 2}$ :

$$\mathbf{A} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix},$$

<sup>1</sup>La notazione  $\mathcal{O}(g(n))$  rappresenta l'insieme delle funzioni  $h(n)$  tali che  $h(n) \leq C \cdot g(n)$  per  $n \geq n_0$ , dove  $C$  è una costante non negativa indipendente da  $n$ . L'espressione  $f = \mathcal{O}(g)$  indica che  $f$  è dell'ordine di  $g$ , e rappresenta un abuso notazionale per  $f \in \mathcal{O}(g)$ .

dove  $\alpha\delta \neq \beta\gamma$ . In questo caso, si può verificare che la matrice inversa, sia a destra che a sinistra, è data da:

$$\mathbf{A}^{-1} = \frac{1}{\alpha\delta - \beta\gamma} \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}.$$

Ma cosa accade se  $\alpha\delta = \beta\gamma$ ? In questo caso, il termine  $\alpha\delta - \beta\gamma$  si annulla, impedendo di calcolare l'inversa della matrice. Questo fenomeno è legato al **determinante** della matrice, un concetto che verrà introdotto in dettaglio più avanti.

La situazione dell'esempio precedente, in cui l'inversa destra e quella sinistra coincidono, è del tutto generale. Infatti, vale il seguente teorema.

**Theorem 4.1 — Unicità dell'inversa.** L'inversa di una matrice invertibile è unica.

*Proof.* Dimostriamo innanzitutto che l'inversa destra e l'inversa sinistra di una matrice invertibile coincidono.

Sia  $\mathbf{A}$  una matrice quadrata e siano  $\mathbf{B}$  e  $\mathbf{C}$  le sue inverse destra e sinistra, rispettivamente. Questo implica che:

$$\mathbf{AB} = \mathbf{I}.$$

Moltiplicando entrambi i membri dell'equazione a sinistra per  $\mathbf{C}$ , otteniamo:

$$\mathbf{CAB} = \mathbf{CI},$$

ma poiché  $\mathbf{CA} = \mathbf{I}$ , segue che:

$$\mathbf{IB} = \mathbf{CI} \Rightarrow \mathbf{B} = \mathbf{C}.$$

Ora dimostriamo l'unicità dell'inversa. Siano  $\mathbf{B}_1$  e  $\mathbf{B}_2$  due inverse destre di  $\mathbf{A}$ , ovvero supponiamo che:

$$\mathbf{AB}_1 = \mathbf{AB}_2 = \mathbf{I}.$$

Poiché l'inversa destra coincide con l'inversa sinistra, possiamo affermare che  $\mathbf{B}_1$  è anche l'inversa sinistra di  $\mathbf{A}$ , quindi:

$$\mathbf{B}_1\mathbf{A} = \mathbf{I} \quad \text{e} \quad \mathbf{AB}_2 = \mathbf{I}.$$

Da ciò segue che, essendo l'inversa destra e sinistra uguali, si ha:

$$\mathbf{B}_1 = \mathbf{B}_2.$$

Questo dimostra che l'inversa di una matrice invertibile è unica. ■

**Theorem 4.2 — Inversa del prodotto.** Siano  $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{n \times n}$  due matrici invertibili. Allora, anche il prodotto  $\mathbf{AB}$  è invertibile e vale la relazione:

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

*Proof.* La formula si dimostra tramite un calcolo diretto, verificando che essa rappresenta correttamente sia l'inversa destra che l'inversa sinistra. Infatti:

$$(\mathbf{AB})(\mathbf{AB})^{-1} = \mathbf{ABB}^{-1}\mathbf{A}^{-1} = \mathbf{I},$$

e allo stesso modo:

$$(\mathbf{AB})^{-1}(\mathbf{AB}) = \mathbf{B}^{-1}\mathbf{A}^{-1}\mathbf{AB} = \mathbf{I}.$$

L'unicità dell'inversa è garantita dal teorema precedente, che assicura che esiste una sola matrice che soddisfa queste proprietà. ■

**Definition 4.11 — Matrice trasposta.** Data una matrice  $\mathbf{A} \in \mathbb{K}^{m \times n}$ , la *matrice trasposta* di  $\mathbf{A}$ , denotata con  $\mathbf{A}^T$ , è la matrice  $\mathbb{K}^{n \times m}$  definita come segue:

$$(\mathbf{A}^T)_{i,j} = A_{j,i}, \quad \text{per ogni} \quad \begin{array}{l} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n. \end{array}$$

In altre parole, la trasposta di  $\mathbf{A}$  si ottiene scambiando le righe e le colonne della matrice originale.

#### Esempio

Data la matrice  $2 \times 3$

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 4 \end{pmatrix},$$

la sua trasposta è la seguente matrice  $3 \times 2$

$$\mathbf{A}^T = \begin{pmatrix} 1 & 2 \\ 2 & 0 \\ 3 & 4 \end{pmatrix}.$$

**Definition 4.12 — Matrice coniugata.** Data una matrice  $\mathbf{A} \in \mathbb{C}^{m \times n}$ , la *matrice coniugata* di  $\mathbf{A}$ , denotata con  $\overline{\mathbf{A}}$ , è la matrice  $\mathbb{C}^{m \times n}$  definita come segue:

$$(\overline{\mathbf{A}})_{i,j} = \overline{A_{i,j}}, \quad \text{per ogni} \quad \begin{array}{l} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n. \end{array}$$

In altre parole, la matrice coniugata si ottiene prendendo il coniugato complesso di ogni elemento della matrice  $\mathbf{A}$ . Se tutte le componenti di  $\mathbf{A}$  sono reali, allora  $\mathbf{A} = \overline{\mathbf{A}}$ .

#### Esempio

Data la matrice  $2 \times 2$  a valori complessi

$$\mathbf{A} = \begin{pmatrix} 1+2i & 2+4i \\ 2-i & 0 \end{pmatrix},$$

la sua coniugata è la seguente matrice  $2 \times 2$

$$\overline{\mathbf{A}} = \begin{pmatrix} 1-2i & 2-4i \\ 2+i & 0 \end{pmatrix}.$$

**Definition 4.13 — Matrice trasposta coniugata.** Data una matrice  $\mathbf{A} \in \mathbb{K}^{m \times n}$ , la *matrice trasposta coniugata* di  $\mathbf{A}$ , denotata con  $\mathbf{A}^H$ , è la matrice  $\mathbb{K}^{n \times m}$  definita come segue:

$$(\mathbf{A}^H)_{i,j} = \overline{A_{j,i}}, \quad \text{per ogni} \quad \begin{matrix} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n. \end{matrix}$$

In altre parole,  $\mathbf{A}^H$  si ottiene applicando il coniugato complesso a ciascun elemento della matrice trasposta di  $\mathbf{A}$ . Si ha quindi che:

$$\mathbf{A}^H = (\overline{\mathbf{A}})^T.$$

#### Esempio

Consideriamo la matrice  $2 \times 2$  a valori complessi

$$\mathbf{A} = \begin{pmatrix} 1+2i & 2+4i \\ 2-i & 0 \end{pmatrix},$$

la sua trasposta coniugata è data da

$$\mathbf{A}^H = \overline{\mathbf{A}}^T = \begin{pmatrix} 1-2i & 2+i \\ 2-4i & 0 \end{pmatrix}.$$

In altre parole, per ottenere la matrice trasposta coniugata  $\mathbf{A}^H$ , si applica il coniugato complesso a ciascun elemento della matrice trasposta di  $\mathbf{A}$ .

**Definition 4.14 — Matrice simmetrica.** Una matrice *quadrata*  $\mathbf{A} \in \mathbb{K}^{m \times m}$  è definita *simmetrica* se è uguale alla sua trasposta. In altre parole,

$$\mathbf{A} = \mathbf{A}^T.$$

Questo significa che ogni elemento della matrice soddisfa la condizione  $A_{i,j} = A_{j,i}$  per tutti  $i$  e  $j$ .

**Definition 4.15 — Matrice hermitiana.** Una matrice *quadrata*  $\mathbf{A} \in \mathbb{K}^{m \times m}$  è definita *hermitiana* se è uguale alla sua trasposta coniugata. In altre parole,

$$\mathbf{A} = \mathbf{A}^H = \overline{\mathbf{A}}^T.$$

Questo implica che ogni elemento della matrice soddisfa la condizione  $A_{i,j} = \overline{A_{j,i}}$  per tutti  $i$  e  $j$ .



## 5. Norme, prodotti scalari ortogonalità

### 5.1 Norme di vettori

Consideriamo un vettore  $\mathbf{x} \in \mathbb{R}^3$  con componenti  $x_1$ ,  $x_2$ , e  $x_3$ . Le componenti di questo vettore possono essere interpretate come le coordinate di un punto nello spazio tridimensionale  $\mathbb{R}^3$ . La distanza di questo punto dall'origine è la lunghezza del segmento che collega l'origine al punto  $\mathbf{x}$ . Questa lunghezza può essere vista come la lunghezza del vettore  $\mathbf{x}$ . Dal teorema di Pitagora<sup>1</sup>, la lunghezza del vettore  $\mathbf{x}$  è data da:

$$\sqrt{x_1^2 + x_2^2 + x_3^2}.$$

Possiamo estendere questa nozione di lunghezza a vettori in  $\mathbb{R}^n$  o  $\mathbb{C}^n$  con la seguente definizione. Per ogni vettore  $\mathbf{x} \in \mathbb{R}^n$  o  $\mathbb{C}^n$ , definiamo la sua *norma euclidea* come:

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}.$$

Questa misura è chiamata la *lunghezza del vettore* e possiede le seguenti proprietà:

1. **Non negatività:** La norma è sempre non negativa,

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} \geq 0.$$

Inoltre,  $\|\mathbf{x}\|_2 = 0$  solo se tutte le componenti sono zero, cioè  $\mathbf{x} = \mathbf{0}$ .

2. **Omogeneità:** Moltiplicare il vettore per uno scalare  $\alpha$  modifica la norma come segue:

$$\begin{aligned}\|\alpha\mathbf{x}\|_2 &= \sqrt{\sum_{i=1}^n |\alpha x_i|^2} \\ &= \sqrt{\sum_{i=1}^n |\alpha|^2 |x_i|^2} \\ &= |\alpha| \sqrt{\sum_{i=1}^n |x_i|^2} \\ &= |\alpha| \|\mathbf{x}\|_2.\end{aligned}$$

---

<sup>1</sup>Pitagora (580 a.C.-500 a.C.)

3. **Disuguaglianza triangolare:** Per ogni coppia di vettori  $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$ , si ha:

$$\|\mathbf{x} + \mathbf{y}\|_2 \leq \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2. \quad (5.1)$$

Questa è nota come *disuguaglianza triangolare*. Il termine “triangolare” si riferisce alla proprietà geometrica dei triangoli, dove la lunghezza di un lato è sempre minore o uguale alla somma delle lunghezze degli altri due lati. Questo concetto è illustrato nella figura Figure 5.1.

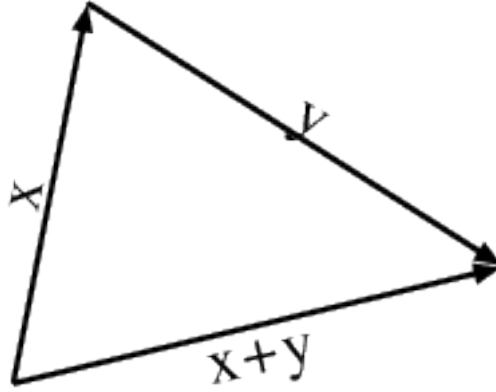


Figure 5.1: Disuguaglianza triangolare

## 5.2 Diseguaglianze di Young<sup>2</sup>, Hölder<sup>3</sup> e Minkowski<sup>4</sup>.

La dimostrazione della disuguaglianza triangolare (Equation 5.1) è un po' laboriosa e richiede l'uso di alcune disuguaglianze classiche, come quella di Young.

**Lemma 5.1 — Disuguaglianza di Young.** Dati due numeri reali  $p$  e  $q$  tali che

$$\frac{1}{p} + \frac{1}{q} = 1, \quad 1 < p, q < \infty,$$

allora per ogni coppia di numeri reali non negativi  $a$  e  $b$  si ha

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \quad (5.2)$$

Inoltre, la disuguaglianza diventa uguaglianza se  $a^p = b^q$ .

*Proof.* Consideriamo la funzione

$$f(t) = \frac{t}{p} - t^{1/p}. \quad (5.3)$$

Allora, calcoliamo la derivata prima della funzione:

$$f'(t) = \frac{1}{p} - \frac{t^{1/p-1}}{p} = \frac{1}{p} - \frac{t^{-1/q}}{p} = \frac{1}{p} \left( 1 - t^{-1/q} \right). \quad (5.4)$$

Poiché  $\frac{1}{q} < 1$ , abbiamo che  $f'(t) < 0$  per  $0 < t < 1$  e  $f'(t) > 0$  per  $t > 1$ . Quindi,  $t = 1$  è un punto di minimo per  $f(t)$  in  $(0, \infty)$ , e di conseguenza  $f(t) \geq f(1)$  per  $t > 0$ . Pertanto,

$$f(t) \geq f(1), \quad \text{cioè} \quad \frac{t}{p} - t^{1/p} \geq \frac{1}{p} - 1 = -\frac{1}{q}.$$

<sup>4</sup>Hermann Minkowski (1864-1909)

Da qui otteniamo:

$$t^{1/p} \leq \frac{1}{q} + \frac{t}{p}.$$

Osserviamo che se  $a = 0$  o  $b = 0$ , la disuguaglianza è banalmente vera. Consideriamo quindi  $a, b > 0$  e calcoliamo la disuguaglianza in  $t = a^p b^{-q}$ :

$$ab^{-q/p} \leq \frac{1}{q} + \frac{a^p b^{-q}}{p}.$$

Moltiplicando la disuguaglianza per  $b^q$  ed osservando che  $q - q/p = 1$ , otteniamo il risultato cercato:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Infine, osserviamo che se  $a^p = b^q$ , allora calcoliamo la disuguaglianza in  $t = 1$ , ottenendo l'uguaglianza:

$$ab = \frac{a^p}{p} + \frac{b^q}{q}.$$

■

Possiamo ora dimostrare la disuguaglianza di Hölder

**Theorem 5.1 — Disuguaglianza di Hölder.** Dati due numeri reali  $p$  e  $q$  tali che  $1 < p, q < \infty$  e che soddisfano  $1/p + 1/q = 1$ , e dati  $a_1, a_2, \dots, a_n \geq 0$  e  $b_1, b_2, \dots, b_n \geq 0$ , allora

$$\sum_{k=1}^n a_k b_k \leq \left( \sum_{k=1}^n a_k^p \right)^{1/p} \left( \sum_{k=1}^n b_k^q \right)^{1/q}.$$

*Proof.* Sia

$$A = \left( \sum_{k=1}^n a_k^p \right)^{1/p}, \quad B = \left( \sum_{k=1}^n b_k^q \right)^{1/q}.$$

Se  $AB = 0$ , allora o  $A = 0$  oppure  $B = 0$ <sup>5</sup>. Questo implica che tutti gli  $a_k$  sono zero, e quindi la disuguaglianza è banalmente vera. Lo stesso ragionamento si applica nel caso in cui  $B = 0$ .

Supponiamo quindi che  $AB > 0$ . Utilizzando la disuguaglianza Equation 5.2 del lemma Lemma 5.1, otteniamo per ogni  $k$ :

Supponiamo, per esempio, che  $A \neq 0$  e  $B \neq 0$ .

$$\frac{a_k}{A} \cdot \frac{b_k}{B} \leq \frac{a_k^p}{pA^p} + \frac{b_k^q}{qB^q}.$$

Sommando su tutti i  $k$ , abbiamo:

$$\frac{\sum_{k=1}^n a_k b_k}{AB} \leq \frac{\sum_{k=1}^n a_k^p}{pA^p} + \frac{\sum_{k=1}^n b_k^q}{qB^q}.$$

Osservando che

<sup>5</sup>O entrambe simultaneamente, ma basta considerare uno dei due casi per procedere!

$$A^p = \sum_{k=1}^n a_k^p \quad \text{e} \quad B^q = \sum_{k=1}^n b_k^q$$

possiamo scrivere:

$$\frac{\sum_{k=1}^n a_k b_k}{AB} \leq \frac{A^p}{pA^p} + \frac{B^q}{qB^q} = \frac{1}{p} + \frac{1}{q} = 1.$$

■

Infine, dimostriamo la disuguaglianza di Minkowski

**Theorem 5.2 — Disuguaglianza di Minkowski.** Sia  $1 \leq p < \infty$ ; e siano  $a_1, a_2, \dots, a_n \geq 0$  e  $b_1, b_2, \dots, b_n \geq 0$ . Allora si ha:

$$\left( \sum_{k=1}^n (a_k + b_k)^p \right)^{1/p} \leq \left( \sum_{k=1}^n a_k^p \right)^{1/p} + \left( \sum_{k=1}^n b_k^p \right)^{1/p}.$$

*Proof.* Il caso  $p = 1$  è immediato e risulta banale. Consideriamo ora il caso  $p > 1$ .

Iniziamo con l'espandere:

$$\sum_{k=1}^n (a_k + b_k)^p = \sum_{k=1}^n a_k (a_k + b_k)^{p-1} + \sum_{k=1}^n b_k (a_k + b_k)^{p-1}.$$

Applicando la disuguaglianza di Hölder a ciascuna delle due somme, con  $q$  definito dalla relazione  $1/p + 1/q = 1$ , otteniamo:

$$\begin{aligned} \sum_{k=1}^n (a_k + b_k)^p &\leq \left( \sum_{k=1}^n a_k^p \right)^{1/p} \left( \sum_{k=1}^n (a_k + b_k)^{q(p-1)} \right)^{1/q} \\ &\quad + \left( \sum_{k=1}^n b_k^p \right)^{1/p} \left( \sum_{k=1}^n (a_k + b_k)^{q(p-1)} \right)^{1/q}. \end{aligned}$$

Dividendo entrambi i lati dell'ineguaglianza per  $(\sum_{k=1}^n (a_k + b_k)^p)^{1/q}$  e notando che  $q(p-1) = p$ , otteniamo la disuguaglianza desiderata. ■

Osserviamo che la disuguaglianza di Minkowski per  $p = 2$  è proprio la Equation 5.1.

Possiamo estendere la nozione di lunghezza di un vettore attraverso una generalizzazione della funzione  $\|\cdot\|$  che mantenga le tre proprietà precedentemente menzionate.

## 5.3 Alcune proprietà delle norme vettoriali

### 5.3.1 Continuità delle norme

**Theorem 5.3** La norma vettoriale è una funzione uniformemente continua dallo spazio dei vettori  $\mathbb{K}^n$  in  $\mathbb{R}$ .

*Proof.* La continuità uniforme della norma segue immediatamente dalla disuguaglianza:

$$\left| \|\mathbf{x}\| - \|\mathbf{y}\| \right| \leq \|\mathbf{x} - \mathbf{y}\|,$$

che dimostra che la norma è uniformemente continua. ■

### 5.3.2 Equivalenza delle norme

**Theorem 5.4** Siano  $\|\cdot\|'$  e  $\|\cdot\|''$  due norme in  $\mathbb{K}^n$ . Esistono due costanti  $\alpha > 0$  e  $\beta > 0$  tali che, per ogni  $\mathbf{x} \in \mathbb{K}^n$ , vale:

$$\alpha \|\mathbf{x}\|'' \leq \|\mathbf{x}\|' \leq \beta \|\mathbf{x}\|''.$$

*Proof.* Se  $\mathbf{x} = \mathbf{0}$ , il teorema è ovvio. Per  $\mathbf{x} \neq \mathbf{0}$ , proviamo l'affermazione considerando il caso particolare in cui  $\|\cdot\|''$  è la norma  $\|\cdot\|_\infty$ . Il caso generale seguirà per confronto.

Definiamo l'insieme:

$$S = \{\mathbf{y} \in \mathbb{K}^n \mid \|\mathbf{y}\|_\infty = 1\}.$$

L'insieme  $S$  è chiuso e limitato, quindi compatto. Poiché la norma  $\|\cdot\|'$  è continua, essa raggiunge un minimo strettamente positivo  $\alpha$  e un massimo  $\beta$  su  $S$ .

Per ogni  $\mathbf{x} \in \mathbb{K}^n$ , possiamo scrivere:

$$\mathbf{y} = \frac{\mathbf{x}}{\|\mathbf{x}\|_\infty} \in S.$$

Di conseguenza, abbiamo:

$$\alpha \leq \|\mathbf{y}\|' = \left\| \frac{\mathbf{x}}{\|\mathbf{x}\|_\infty} \right\|' \leq \beta,$$

che implica:

$$\alpha \|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|' \leq \beta \|\mathbf{x}\|_\infty.$$

■

### 5.3.3 Confronto tra norme standard

Per ogni vettore  $\mathbf{x} \in \mathbb{K}^n$ , valgono le seguenti disuguaglianze:

$$\begin{aligned} (i) \quad & \|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n} \|\mathbf{x}\|_\infty, \\ (ii) \quad & \|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{n} \|\mathbf{x}\|_\infty, \\ (iii) \quad & \|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1 \leq n \|\mathbf{x}\|_\infty. \end{aligned}$$

*Proof.* Le disuguaglianze si ottengono nei seguenti modi:

- (i) Sia  $\|\cdot\|' = \|\cdot\|_2$ . Ogni vettore  $\mathbf{y} \in S = \{\mathbf{y} \in \mathbb{K}^n \mid \|\mathbf{y}\|_\infty = 1\}$  ha almeno una componente di modulo 1, diciamo la componente  $y_k$ , con  $1 \leq k \leq n$ , mentre le altre componenti soddisfano  $|y_i| \leq 1$  per  $i \neq k$ .

Pertanto, si ha:

$$\|\mathbf{y}\|_2^2 = 1 + \sum_{\substack{i=1 \\ i \neq k}}^n |y_i|^2 \quad \forall \mathbf{y} \in S,$$

da cui segue che:

$$\alpha = \min_{\mathbf{y} \in S} \|\mathbf{y}\|_2 = 1,$$

$$\beta = \max_{\mathbf{y} \in S} \|\mathbf{y}\|_2 = \sqrt{n}.$$

Quindi, si ottiene:

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n} \|\mathbf{x}\|_\infty.$$

- (ii) La prima disuguaglianza di (ii) si ottiene notando che, per ogni  $\mathbf{x} \in \mathbb{K}^n$ ,

$$\|\mathbf{x}\|_2^2 = \sum_{i=1}^n |x_i|^2 \leq \left( \sum_{i=1}^n |x_i| \right)^2 = \|\mathbf{x}\|_1^2.$$

La seconda disuguaglianza deriva dalla disuguaglianza di Cauchy-Schwarz:

$$|(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2,$$

applicata a un vettore ausiliario  $\mathbf{y}$  definito da:

$$y_i = \begin{cases} \frac{x_i}{|x_i|}, & \text{se } x_i \neq 0, \\ 0, & \text{se } x_i = 0. \end{cases}$$

Si osserva che:

$$|(\mathbf{x}, \mathbf{y})| = \sum_{i=1}^n |x_i| = \|\mathbf{x}\|_1,$$

e

$$\|\mathbf{y}\|_2 = \left( \sum_{i=1}^n |y_i|^2 \right)^{\frac{1}{2}} = \sqrt{n}.$$

Quindi, otteniamo:

$$\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{n} \|\mathbf{x}\|_\infty.$$

- (iii) Le disuguaglianze si ottengono combinando i risultati di (i) e (ii):

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1 \leq n \|\mathbf{x}\|_\infty.$$

■

**Definition 5.1 —  $p$ -Norma.** Utilizzando la disuguaglianza di Minkowski (Equation 5.1), si dimostra che per  $1 \leq p < \infty$ , la funzione definita come

$$\|\mathbf{x}\|_p = \left( \sum_{k=1}^n |x_k|^p \right)^{1/p}$$

è una norma. Questa norma è chiamata  $p$ -norma.

Due casi particolari di  $p$ -norma sono:

- **Norma 1:**  $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$ .
- **Norma 2:**  $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$ .

Inoltre, è possibile estendere la definizione della norma al caso limite  $p \rightarrow \infty$ :

$$\|\mathbf{x}\|_\infty = \max_{k=1, \dots, n} |x_k|,$$

e si può verificare facilmente che anche questa funzione è una norma.

## Esempio

Dati i vettori

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 1+i \\ i \\ 1 \\ -1 \end{pmatrix},$$

otteniamo

$$\|\mathbf{x}\|_{\infty} = \max\{|1|, |2|, |-2|\} = \max\{1, 2, 2\} = 2,$$

$$\|\mathbf{y}\|_{\infty} = \max\{|1+i|, |i|, |1|, |-1|\} = \max\{\sqrt{2}, 1, 1, 1\} = \sqrt{2},$$

$$\|\mathbf{x}\|_1 = |1| + |2| + |-2| = 1 + 2 + 2 = 5,$$

$$\|\mathbf{y}\|_1 = |1+i| + |i| + |1| + |-1| = \sqrt{2} + 1 + 1 + 1 = 3 + \sqrt{2},$$

$$\|\mathbf{x}\|_2 = \sqrt{|1|^2 + |2|^2 + |-2|^2} = \sqrt{1 + 4 + 4} = 3,$$

$$\|\mathbf{y}\|_2 = \sqrt{|1+i|^2 + |i|^2 + |1|^2 + |-1|^2} = \sqrt{2 + 1 + 1 + 1} = \sqrt{5}.$$

## 5.4 Prodotti scalari

Il prodotto scalare tra due vettori è molto usato in fisica ed ha la seguente definizione

**Definition 5.2 — Prodotto Scalare (Euclideo).** Il prodotto scalare (o prodotto dot) tra due vettori  $\mathbf{a}$  e  $\mathbf{b}$  è dato dalla formula:

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\|_2 \|\mathbf{b}\|_2 \cos \theta_{ab}, \quad (5.5)$$

dove  $\theta_{ab}$  rappresenta l'angolo compreso tra i due vettori  $\mathbf{a}$  e  $\mathbf{b}$ .

Dalla figura <sup>6</sup> Figure 5.2 possiamo derivare una formula per il prodotto scalare che non utilizza il coseno, ma solo la norma.

Consideriamo le seguenti definizioni:

$$L = \|\mathbf{b}\|_2 \cos \theta_{ab}, \quad H = \|\mathbf{b}\|_2 \sin \theta_{ab},$$

dove  $\theta_{ab}$  è l'angolo tra i vettori  $\mathbf{a}$  e  $\mathbf{b}$ . Applicando il teorema di Pitagora, otteniamo:

$$\|\mathbf{a} + \mathbf{b}\|_2^2 - H^2 = (\|\mathbf{a}\|_2 + L)^2 = \|\mathbf{a}\|_2^2 + L^2 + 2\|\mathbf{a}\|_2 L \quad (5.6)$$

e

$$\|\mathbf{a} - \mathbf{b}\|_2^2 - H^2 = (\|\mathbf{a}\|_2 - L)^2 = \|\mathbf{a}\|_2^2 + L^2 - 2\|\mathbf{a}\|_2 L \quad (5.7)$$

Sottraendo l'equazione Equation 5.7 da Equation 5.6, otteniamo:

$$\|\mathbf{a} + \mathbf{b}\|_2^2 - \|\mathbf{a} - \mathbf{b}\|_2^2 = 4\|\mathbf{a}\|_2 L = 4\|\mathbf{a}\|_2 \|\mathbf{b}\|_2 \cos \theta_{ab},$$

<sup>6</sup>Nella figura, l'angolo tra i vettori  $\mathbf{a}$  e  $\mathbf{b}$  è indicato con il simbolo  $\theta$  per ragioni tipografiche, anziché con  $\theta_{ab}$ .

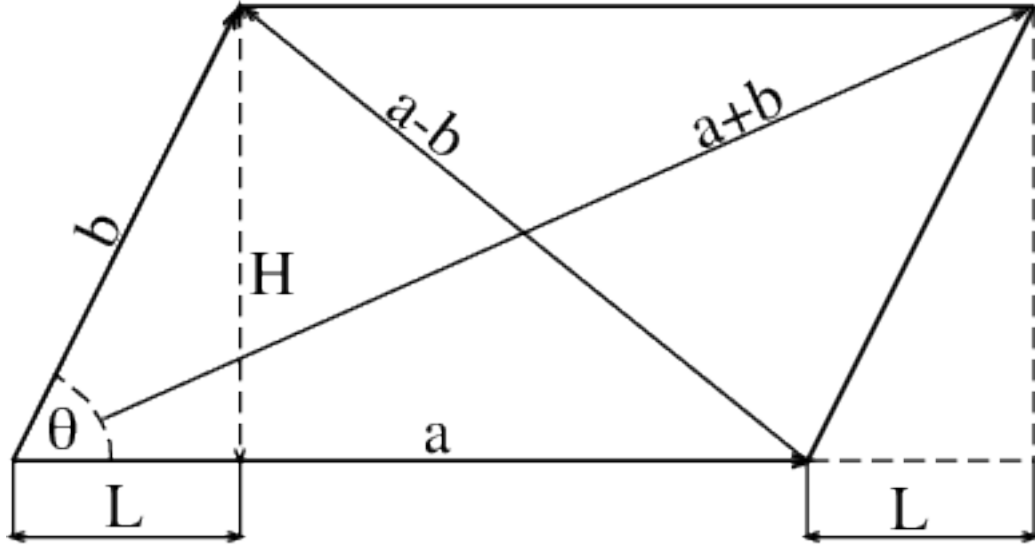


Figure 5.2: prodotto scalare e parallelogrammo associato

e quindi, utilizzando la formula Equation 5.5, possiamo scrivere:

$$\mathbf{a} \cdot \mathbf{b} = \frac{\|\mathbf{a} + \mathbf{b}\|_2^2 - \|\mathbf{a} - \mathbf{b}\|_2^2}{4}.$$

Inoltre, osserviamo che:

$$\|\mathbf{a} + \mathbf{b}\|_2^2 - \|\mathbf{a} - \mathbf{b}\|_2^2 = \sum_{k=1}^n [|a_k + b_k|^2 - |a_k - b_k|^2] = 4 \sum_{k=1}^n a_k b_k,$$

da cui il prodotto scalare si esprime come:

$$\mathbf{a} \cdot \mathbf{b} = \sum_{k=1}^n a_k b_k. \quad (5.8)$$

Questa formula è valida per vettori reali. Per un vettore complesso  $\mathbf{a}$ , la formula Equation 5.8 non restituisce direttamente il quadrato della lunghezza, poiché se  $a_k$  sono numeri complessi:

$$\mathbf{a} \cdot \mathbf{a} = \sum_{k=1}^n a_k^2 \neq \sum_{k=1}^n |a_k|^2.$$

Tuttavia, è possibile modificare la definizione di prodotto scalare per ottenere una formula che, applicata a vettori reali, sia equivalente a Equation 5.5 e che, nel caso di vettori complessi, restituisca  $\mathbf{a} \cdot \mathbf{a} = \|\mathbf{a}\|_2^2$ .

**Definition 5.3 — Prodotto Scalare nel Campo Complesso.** Definiamo il prodotto scalare tra due vettori  $\mathbf{x}$  e  $\mathbf{y}$  di dimensione  $n$  come:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i \bar{y}_i, \quad (5.9)$$

dove  $\bar{z}$  denota il coniugato complesso di un numero complesso  $z$ . Ricordiamo che:

$$\overline{a + ib} = a - ib$$

e il coniugato complesso soddisfa le seguenti proprietà:

$$\begin{aligned}\bar{z} \cdot z &= (a + ib)(a - ib) = a^2 + b^2 = |z|^2, \\ \bar{\bar{z}} &= \overline{a - ib} = a + ib = z, \\ \overline{z + w} &= \bar{z} + \bar{w},\end{aligned}$$

oltre a:

$$\begin{aligned}\overline{zw} &= \overline{(a + ib)(c + id)} = \overline{(ac - bd) + (bc + ad)i} \\ &= (ac - bd) - (bc + ad)i, \\ \bar{z} \cdot \bar{w} &= (a - ib)(c - id) \\ &= ac - bd - (bc + ad)i,\end{aligned}$$

da cui segue:

$$\overline{zw} = \bar{z} \cdot \bar{w}.$$

Per un numero complesso  $z = a + ib$ , abbiamo:

$$\Re(z) = \frac{z + \bar{z}}{2} = a, \quad \Im(z) = \frac{z - \bar{z}}{2i} = b.$$

La formula Equation 5.9 è nota come *prodotto scalare euclideo*.

Questa funzione soddisfa le seguenti proprietà:

1. **Non negatività e nullità:**

$$\mathbf{x} \cdot \mathbf{x} = \sum_{i=1}^n x_i \bar{x}_i = \sum_{i=1}^n |x_i|^2 \geq 0.$$

Inoltre,  $\mathbf{x} \cdot \mathbf{x} = 0$  se e solo se  $x_i = 0$  per tutti  $i$ , il che implica  $\mathbf{x} = \mathbf{0}$ .

2. **Simmetria:**

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i \bar{y}_i = \overline{\sum_{i=1}^n \bar{x}_i y_i} = \overline{\mathbf{y} \cdot \mathbf{x}}.$$

3. **Linearità:**

$$(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \sum_{i=1}^n (x_i + y_i) \bar{z}_i = \sum_{i=1}^n x_i \bar{z}_i + \sum_{i=1}^n y_i \bar{z}_i = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}.$$

4. **Omogeneità:**

$$(\alpha \mathbf{x}) \cdot \mathbf{y} = \sum_{i=1}^n \alpha x_i \bar{y}_i = \alpha \sum_{i=1}^n x_i \bar{y}_i = \alpha (\mathbf{x} \cdot \mathbf{y}).$$

Queste proprietà (1)-(4) possono essere utilizzate per definire il prodotto scalare in modo assiomatico.

**Definition 5.4 — Prodotto Scalare in generale.** Una funzione  $\langle \cdot, \cdot \rangle : \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}$  è un *prodotto scalare* se soddisfa le seguenti proprietà per ogni  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{K}^n$  e per ogni  $\alpha \in \mathbb{K}$ :

1. **Non negatività e nullità:**

$$\langle \mathbf{x}, \mathbf{x} \rangle \geq 0 \quad \text{e} \quad \langle \mathbf{x}, \mathbf{x} \rangle = 0 \text{ se e solo se } \mathbf{x} = \mathbf{0}.$$

2. **Simmetria Coniugata:**

$$\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle},$$

dove  $\bar{\cdot}$  denota l'operazione di coniugazione complessa.

## 3. Linearità nella prima variabile:

$$\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle.$$

## 4. Omogeneità:

$$\langle \alpha \mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle.$$

Queste proprietà definiscono il concetto di prodotto scalare in uno spazio vettoriale.

## Osservazione

Nella definizione assiomatica del prodotto scalare, abbiamo utilizzato la notazione  $\langle \mathbf{x}, \mathbf{y} \rangle$  per indicare il prodotto scalare tra due vettori  $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$ . Nella discussione precedente, invece, abbiamo utilizzato la notazione  $\mathbf{x} \cdot \mathbf{y}$ .

È importante sottolineare che la scelta della notazione non influisce sulla definizione del prodotto scalare né sulle sue proprietà.

Esiste anche una terza notazione di uso comune,  $\mathbf{x}^T \mathbf{y}$ . In questa notazione, si assume implicitamente che tutti i vettori siano vettori colonna, e l'apice  $T$  denota l'operazione di trasposizione, che trasforma un vettore colonna in un vettore riga.

La definizione generale di *trasposto* di una matrice e di un vettore, di cui abbiamo già anticipato l'idea, sarà introdotta tra poco.

## Esempio

E' facile verificare che anche la seguente funzione

$$[\mathbf{x}, \mathbf{y}] = \sum_{k=1}^n k x_k \overline{y_k},$$

definisce un prodotto scalare per ogni  $k$  reale positivo.

## Osservazione

La funzione  $\|\cdot\|_2$  può essere espressa tramite il prodotto scalare euclideo come segue:

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x} \cdot \mathbf{x}}.$$

In modo analogo, dato un prodotto scalare generico  $(\cdot, \cdot)$ , è possibile definire l'applicazione:

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}},$$

che soddisfa le proprietà di una norma. Questa applicazione è conosciuta come *norma indotta dal prodotto scalare*.

**Theorem 5.5 — Disuguaglianza di Cauchy<sup>a</sup>-Schwarz<sup>b</sup>.** Per un prodotto scalare generico, la disuguaglianza di Cauchy-Schwarz è espressa come:

$$(\mathbf{x}, \mathbf{y}) \leq \|\mathbf{x}\| \|\mathbf{y}\| \quad (5.10)$$

dove  $\|\cdot\|$  rappresenta la norma indotta dal prodotto scalare. La disuguaglianza è stretta solo se  $\mathbf{x}$  e  $\mathbf{y}$  sono allineati, cioè se esiste uno scalare  $\gamma$  tale che  $\mathbf{x} = \gamma \mathbf{y}$ .

<sup>a</sup>Augustin Louis Cauchy (1789-1857)

<sup>b</sup>Karl Herman Amandus Schwarz (1843-1921).

*Proof.* La disuguaglianza è ovvia se uno dei vettori è nullo. Supponiamo quindi che entrambi i vettori siano non nulli. Consideriamo il vettore  $\mathbf{x} - \alpha\mathbf{y}$  e applicando la proprietà (1) della definizione di prodotto scalare alla sua norma, otteniamo:

$$(\mathbf{x} - \alpha\mathbf{y}, \mathbf{x} - \alpha\mathbf{y}) \geq 0.$$

Sviluppando l'espressione, abbiamo:

$$\begin{aligned} 0 &\leq (\mathbf{x} - \alpha\mathbf{y}, \mathbf{x} - \alpha\mathbf{y}) \\ &= (\mathbf{x}, \mathbf{x}) - \alpha(\mathbf{y}, \mathbf{x}) - \overline{\alpha}(\mathbf{x}, \mathbf{y}) + \alpha\overline{\alpha}(\mathbf{y}, \mathbf{y}) \\ &= (\mathbf{x}, \mathbf{x}) - \alpha\overline{(\mathbf{x}, \mathbf{y})} - \overline{\alpha}[(\mathbf{x}, \mathbf{y}) - \alpha(\mathbf{y}, \mathbf{y})]. \end{aligned} \quad (5.11)$$

Scegliendo  $\alpha$  in modo da annullare l'espressione tra parentesi quadre, otteniamo:

$$\alpha = \frac{(\mathbf{x}, \mathbf{y})}{(\mathbf{y}, \mathbf{y})}.$$

Sostituendo questo valore di  $\alpha$ , si ottiene:

$$0 \leq \|\mathbf{x}\|^2 - \frac{|(\mathbf{x}, \mathbf{y})|^2}{\|\mathbf{y}\|^2},$$

che è equivalente alla disuguaglianza di Cauchy-Schwarz espressa in Equation 5.10. Se  $\mathbf{x} - \alpha\mathbf{y} \neq \mathbf{0}$ , allora la disuguaglianza è stretta, e quindi anche Equation 5.10 è stretta. ■

## 5.5 Ortogonalità e angolo tra vettori

Il concetto di prodotto scalare permette di definire l'ortogonalità e l'angolo tra due vettori. Utilizzando il prodotto scalare euclideo espresso in Equation 5.9 e la formula per il prodotto scalare in Equation 5.5, possiamo calcolare l'angolo  $\theta_{ab}$  tra due vettori  $\mathbf{a}$  e  $\mathbf{b}$  con la seguente formula:

$$\theta_{ab} = \arccos\left(\frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\|_2 \|\mathbf{b}\|_2}\right).$$

### Esempio

I vettori  $\mathbf{a}$  e  $\mathbf{b}$  definiti come segue

$$\mathbf{a} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

formano un angolo di circa  $30^\circ$  o circa 0.5236 radianti, infatti

$$\cos \theta_{ab} = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\|_2 \|\mathbf{b}\|_2} = \frac{3}{\sqrt{12}}.$$

Se l'angolo tra i vettori è  $90^\circ$  allora  $\cos 90^\circ = 0$  implica che il loro prodotto scalare è nullo. Questo suggerisce la seguente definizione.

**Definition 5.5 — Ortogonalità.** Due vettori  $\mathbf{a}$  e  $\mathbf{b}$  si dicono ortogonali se il loro prodotto scalare è nullo. In tal caso, scriviamo  $\mathbf{a} \perp \mathbf{b}$ , indicando che:

$$\mathbf{a} \cdot \mathbf{b} = 0.$$

## Osservazione

La definizione dell'angolo tra vettori è specifica per vettori reali, poiché utilizza il concetto di coseno. Tuttavia, la definizione di ortogonalità è valida anche per vettori complessi e per qualsiasi prodotto scalare. Questo perché l'ortogonalità è una proprietà puramente algebrica, che dipende solo dal prodotto scalare tra i vettori.

## Esempio

I vettori

$$\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix},$$

sono ortogonali (cioè  $\mathbf{a} \perp \mathbf{b}$ ) infatti

$$\mathbf{a} \cdot \mathbf{b} = 1 \cdot 1 + 2 \cdot (-2) + 3 \cdot 1 = 0,$$

analogamente i vettori

$$\mathbf{x} = \begin{pmatrix} 1+i \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 1+i \\ -1 \\ -1 \end{pmatrix},$$

sono ortogonali infatti

$$\mathbf{x} \cdot \mathbf{y} = (1+i)(\overline{1+i}) + 1 \cdot (-1) + 1 \cdot (-1) = 0.$$

## 5.6 Prodotto Vettoriale

Consideriamo due vettori  $\mathbf{a}$  e  $\mathbf{b}$  nello spazio tridimensionale  $\mathbb{R}^3$ . Vogliamo trovare un terzo vettore  $\mathbf{x}$  che sia ortogonale a entrambi i vettori dati. Algebricamente, il problema può essere formulato come:

$$\text{Trovare } \mathbf{x} \in \mathbb{R}^3 \text{ tale che: } \begin{cases} \mathbf{a} \cdot \mathbf{x} = 0, \\ \mathbf{b} \cdot \mathbf{x} = 0. \end{cases}$$

Espresso in termini delle componenti dei vettori, questo diventa:

$$\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 = 0, \\ b_1x_1 + b_2x_2 + b_3x_3 = 0. \end{cases} \quad (5.12)$$

Una soluzione a questo sistema di equazioni è data dalle componenti:

$$\begin{cases} x_1 = a_2b_3 - a_3b_2, \\ x_2 = a_3b_1 - a_1b_3, \\ x_3 = a_1b_2 - a_2b_1. \end{cases} \quad (5.13)$$

Questa soluzione è conosciuta come *prodotto vettoriale* e si indica con:

$$\mathbf{x} = \mathbf{a} \wedge \mathbf{b}.$$

Inoltre, possiamo verificare la seguente relazione:

$$(\|\mathbf{a} \wedge \mathbf{b}\|_2)^2 + (\mathbf{a} \cdot \mathbf{b})^2 = (\|\mathbf{a}\|_2)^2 (\|\mathbf{b}\|_2)^2, \quad (5.14)$$

da cui, utilizzando la formula sopra, otteniamo:

$$\|\mathbf{a} \wedge \mathbf{b}\|_2 = \|\mathbf{a}\|_2 \|\mathbf{b}\|_2 \sin \theta_{ab}.$$

Il prodotto vettoriale è molto utile in geometria nello spazio. Per esempio, per trovare l'equazione di un piano che passa per tre punti distinti  $\mathbf{a}$ ,  $\mathbf{b}$ , e  $\mathbf{c}$ , possiamo usare il prodotto vettoriale. Definiamo i vettori:

$$\mathbf{v} = \mathbf{b} - \mathbf{a}, \quad \mathbf{w} = \mathbf{c} - \mathbf{a},$$

che sono vettori complanari al piano. Il vettore normale al piano è dato da:

$$\mathbf{N} = \mathbf{v} \wedge \mathbf{w}.$$

L'equazione del piano può quindi essere scritta come:

$$\mathbf{N} \cdot \mathbf{x} = \mathbf{N} \cdot \mathbf{a}.$$

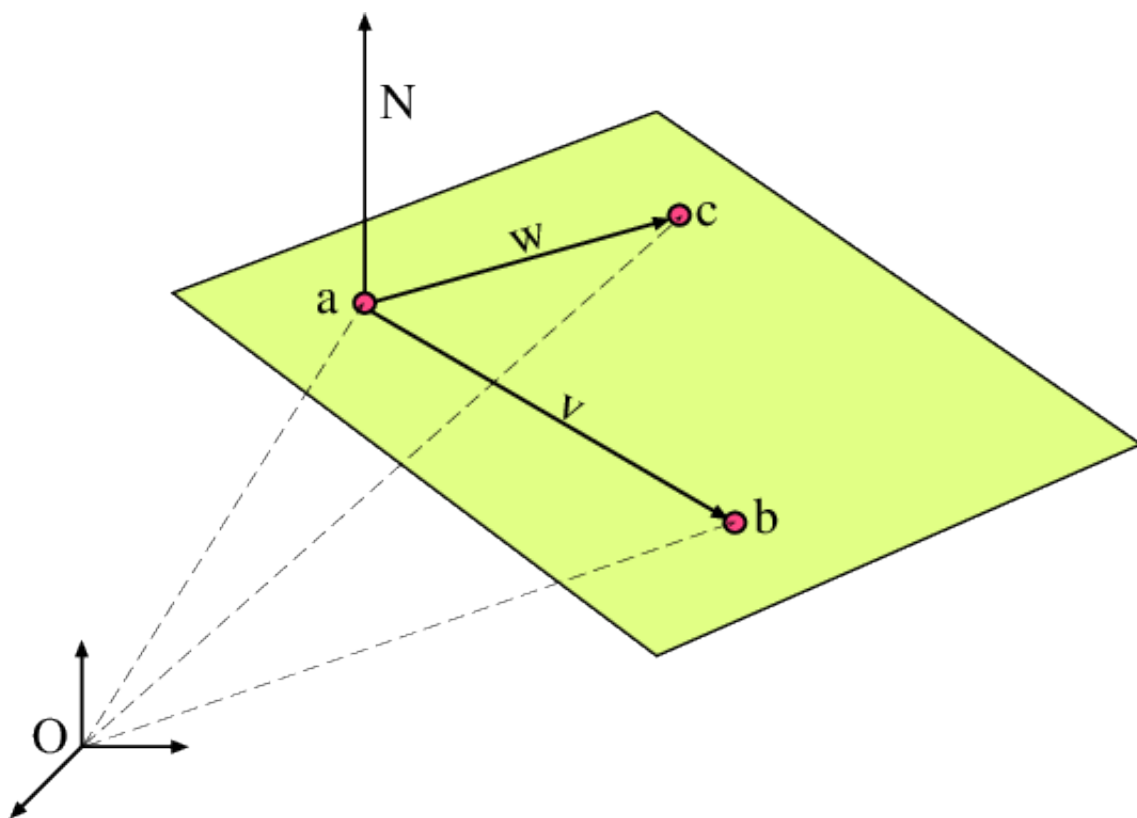


Figure 5.3: piano per 3 punti

#### Esempio

Dati i punti

$$\mathbf{a} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},$$

trovare il piano passante per  $\mathbf{a}$ ,  $\mathbf{b}$  e  $\mathbf{c}$ . Calcoliamo innanzitutto

$$\mathbf{v} = \mathbf{b} - \mathbf{a} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix},$$

$$\mathbf{w} = \mathbf{c} - \mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix},$$

da cui

$$\mathbf{N} = \mathbf{v} \wedge \mathbf{w} = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ -2 \end{pmatrix},$$

e infine

$$\mathbf{N} \cdot \mathbf{x} = 4x_1 + 2x_2 - 2x_3, \quad \mathbf{N} \cdot \mathbf{a} = 2,$$

ponendo  $\mathbf{x} = (x, y, z)^T$  otteniamo l'equazione del piano

$$2x + y - z = 1.$$

## 5.7 Indipendenza Lineare e Basi in $\mathbb{K}^n$

Il concetto di dipendenza e indipendenza lineare è fondamentale in algebra lineare.

**Definition 5.6 — Indipendenza Lineare.** Consideriamo  $k$  vettori non nulli  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$  in uno spazio vettoriale. Questi vettori sono detti *linearmente dipendenti* se esistono degli scalari  $\alpha_1, \alpha_2, \dots, \alpha_k$ , con almeno uno di essi non nullo, tali che:

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_k \mathbf{x}_k = \mathbf{0}.$$

Se, invece, non esistono tali scalari e l'unica soluzione all'equazione sopra è quella triviale, cioè tutti gli scalari sono zero, allora i vettori  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$  sono detti *linearmente indipendenti*.

Consideriamo  $k$  vettori *linearmente indipendenti* in  $\mathbb{K}^n$ , denotati con  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ . Dato un vettore  $\mathbf{w} \in \mathbb{K}^n$ , è possibile che  $\mathbf{w}$  possa essere scritto come una combinazione lineare di questi vettori. In altre parole, esistono scalari  $\alpha_1, \alpha_2, \dots, \alpha_k$  tali che:

$$\mathbf{w} = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_k \mathbf{x}_k.$$

Questi scalari  $\alpha_i$  sono scelti in modo tale che l'equazione sia soddisfatta.

**Definition 5.7 — Base.** Se i vettori  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$  sono tali che ogni vettore  $\mathbf{w} \in \mathbb{K}^n$  può essere espresso come una combinazione lineare di essi, allora diremo che questi vettori formano una *base* di  $\mathbb{K}^n$ . Affinché ciò accada, è necessario che il numero di vettori  $k$  sia uguale alla dimensione dello spazio, cioè  $k = n$ .

Inoltre, questa condizione è anche sufficiente. Infatti, se si scelgono  $n$  vettori qualsiasi in  $\mathbb{K}^n$  che siano linearmente indipendenti, allora questi vettori formano sempre una base di  $\mathbb{K}^n$ .

■ **Example 5.1 — Esempio.** Consideriamo i vettori  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  in  $\mathbb{K}^n$ , definiti come segue:

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Questi vettori sono chiaramente linearmente indipendenti. Infatti, se consideriamo la combinazione lineare

$$\alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 + \dots + \alpha_n \mathbf{e}_n = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix},$$

questa sarà nulla se e solo se tutti i coefficienti  $\alpha_i$  sono zero, cioè  $\alpha_i = 0$  per  $i = 1, 2, \dots, n$ . ■

#### Esempio

I vettori  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ , come definiti nell'esempio Example 5.1, formano una base di  $\mathbb{K}^n$ . Per ogni vettore  $\mathbf{w} \in \mathbb{K}^n$  con componenti  $w_1, w_2, \dots, w_n$ , possiamo esprimere  $\mathbf{w}$  come una combinazione lineare di questi vettori:

$$\mathbf{w} = w_1 \mathbf{e}_1 + w_2 \mathbf{e}_2 + \dots + w_n \mathbf{e}_n.$$

Questa base è chiamata *base canonica* di  $\mathbb{K}^n$ .

#### Esempio

Consideriamo i vettori

$$\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} -1 \\ 6 \\ 1 \end{pmatrix}.$$

Questi vettori *non sono* linearmente indipendenti. Infatti, è facile verificare che

$$\mathbf{a} - 2\mathbf{b} + \mathbf{c} = \mathbf{0}.$$

Questo implica che esiste una combinazione lineare non banale (con coefficienti diversi da zero) che dà come risultato il vettore nullo.

## 5.8 Indipendenza lineare dei vettori ortogonali

**Theorem 5.6** I  $k$  vettori  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$  che sono ortogonali a coppie,

$$\mathbf{x}_i \perp \mathbf{x}_j \quad \text{per } i \neq j,$$

sono necessariamente linearmente indipendenti.

*Proof.* Supponiamo che esistano  $k$  scalari  $\alpha_1, \alpha_2, \dots, \alpha_k$  tali che

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_k \mathbf{x}_k = \mathbf{0}.$$

Per dimostrare che i vettori sono linearmente indipendenti, consideriamo il prodotto scalare di entrambi i lati dell'equazione con il vettore  $\mathbf{x}_i$  per  $i = 1, 2, \dots, k$ . Utilizzando l'ortogonalità dei vettori, otteniamo:

$$\begin{aligned} 0 &= \mathbf{x}_i \cdot (\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_k \mathbf{x}_k), \\ &= \alpha_1 (\mathbf{x}_i \cdot \mathbf{x}_1) + \alpha_2 (\mathbf{x}_i \cdot \mathbf{x}_2) + \dots + \alpha_i (\mathbf{x}_i \cdot \mathbf{x}_i) + \dots + \alpha_k (\mathbf{x}_i \cdot \mathbf{x}_k). \end{aligned}$$

Poiché  $\mathbf{x}_i \cdot \mathbf{x}_j = 0$  per  $i \neq j$ , rimane

$$0 = \alpha_i (\mathbf{x}_i \cdot \mathbf{x}_i).$$

Dato che  $\mathbf{x}_i \cdot \mathbf{x}_i > 0$  (perché  $\mathbf{x}_i$  è un vettore non nullo), segue che  $\alpha_i = 0$ .

Quindi, tutti i coefficienti  $\alpha_i$  devono essere zero, il che dimostra che i vettori  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$  sono linearmente indipendenti. ■

## 5.9 Ortonormalizzazione di Gram<sup>7</sup>-Schmidt<sup>8</sup>

**Definition 5.8 — Vettori ortogonali.** Dati  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ , diremo che gli stessi formano un *sistema ortogonale* se sono a due a due ortogonali, cioè

$$\mathbf{v}_i \perp \mathbf{v}_j, \quad i \neq j.$$

**Definition 5.9 — Vettori ortonormali.** Dati  $k$  vettori  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ , diremo che gli stessi formano un *sistema ortonormale* se sono a due a due ortogonali e di norma 1, cioè

$$\|\mathbf{u}_i\|_2 = 1, \quad \mathbf{u}_i \perp \mathbf{u}_j, \quad i \neq j.$$

**Definition 5.10 — Span.** Dati  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ , definiremo con  $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$  lo spazio vettoriale generato dalle loro combinazioni lineari

$$\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = \{\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_k \mathbf{v}_k \mid \alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{K}\}$$

Dati  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ , linearmente indipendenti è possibile costruire  $k$  vettori  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$  a due a due ortogonali e di norma unitaria tali che

$$\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = \text{span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k).$$

### Theorem 5.7 Ortonormalizzazione di Gram-Schmidt

Consideriamo  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  che sono linearmente indipendenti. Possiamo costruire una nuova sequenza di  $k$  vettori ortonormali  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$  che soddisfano le seguenti proprietà:

1.  $\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|_2}$ :  
Il primo vettore  $\mathbf{u}_1$  è ottenuto normalizzando  $\mathbf{v}_1$ .
2.  $\mathbf{u}_i \perp \mathbf{u}_j$  per ogni  $i \neq j$ :  
I vettori  $\mathbf{u}_i$  sono ortogonali tra loro.
3.  $\|\mathbf{u}_i\|_2 = 1$  per ogni  $i = 1, 2, \dots, k$ :  
Ogni vettore  $\mathbf{u}_i$  è normalizzato, cioè ha lunghezza unitaria.
4.  $\mathbf{V}_j = \text{span}(\mathbf{u}_1, \dots, \mathbf{u}_j)$  per ogni  $j = 1, 2, \dots, k$ :  
Lo span dei primi  $j$  vettori  $\mathbf{v}_i$  è lo stesso dello span dei primi  $j$  vettori ortonormali  $\mathbf{u}_i$ , dove

$$\mathbf{V}_j = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j)$$

<sup>8</sup>Erhard Schmidt (1876-1959)

e

$$U_j = \text{span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_j)$$

.

*Proof.* Dati  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  che sono linearmente indipendenti, notiamo che  $\|\mathbf{v}_i\|_2 \neq 0$  per ogni  $i = 1, 2, \dots, k$ . Quindi, possiamo sempre normalizzare il primo vettore  $\mathbf{v}_1$  definendo  $\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|_2}$ . La dimostrazione procede per induzione.

- **Passo 1:**

Per  $k = 1$ , il teorema è ovviamente vero, poiché abbiamo  $\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|_2}$ , che soddisfa tutte le condizioni richieste.

- **Passo 2:**

Supponiamo che il teorema sia vero per  $k - 1$  vettori. Ovvero, assumiamo di avere già trovato  $k - 1$  vettori ortonormali  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}$  tali che:

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|_2}, \quad \mathbf{V}_j = \mathbf{U}_j \text{ per } j = 1, 2, \dots, k - 1$$

Definiamo ora il vettore ausiliario  $\mathbf{w}_k$  e il vettore ortonormale  $\mathbf{u}_k$  come segue:

$$\mathbf{w}_k = \mathbf{v}_k - \sum_{i=1}^{k-1} \beta_i \mathbf{u}_i,$$

$$\mathbf{u}_k = \alpha \mathbf{w}_k,$$

dove i coefficienti  $\alpha$  e  $\beta_i$  devono essere scelti in modo che  $\mathbf{u}_k \perp \mathbf{u}_i$  per ogni  $i = 1, 2, \dots, k - 1$  e  $\|\mathbf{u}_k\|_2 = 1$ .

Calcoliamo il prodotto scalare di  $\mathbf{w}_k$  con  $\mathbf{u}_j$ :

$$\mathbf{w}_k \cdot \mathbf{u}_j = \mathbf{v}_k \cdot \mathbf{u}_j - \sum_{i=1}^{k-1} \beta_i \mathbf{u}_i \cdot \mathbf{u}_j = \mathbf{v}_k \cdot \mathbf{u}_j - \beta_j \quad \text{per } j = 1, 2, \dots, k - 1$$

Impostando  $\mathbf{w}_k \cdot \mathbf{u}_j = 0$ , otteniamo:

$$\beta_j = \mathbf{v}_k \cdot \mathbf{u}_j \quad \text{per } j = 1, 2, \dots, k - 1$$

Per determinare  $\alpha$ , imponiamo che  $\|\mathbf{u}_k\|_2 = 1$ :

$$1 = \|\mathbf{u}_k\|_2^2 = \mathbf{u}_k \cdot \mathbf{u}_k = \alpha^2 (\mathbf{w}_k \cdot \mathbf{w}_k) = \alpha^2 \|\mathbf{w}_k\|_2^2$$

Da cui si ottiene  $\alpha = \frac{1}{\|\mathbf{w}_k\|_2}$ . È necessario che  $\mathbf{w}_k \neq \mathbf{0}$ ; altrimenti, avremmo:

$$\mathbf{0} = \mathbf{v}_k - \sum_{i=1}^{k-1} \beta_i \mathbf{u}_i$$

Poiché  $\mathbf{U}_{k-1} = \mathbf{V}_{k-1}$ , esisterebbero  $k - 1$  scalari  $\gamma_i$  per cui:

$$\mathbf{v}_k = \sum_{i=1}^{k-1} \beta_i \mathbf{u}_i = \sum_{i=1}^{k-1} \gamma_i \mathbf{v}_i$$

Questo contraddice l'indipendenza lineare dei vettori  $\mathbf{v}_i$ .

- **Passo 3:**

Infine, dobbiamo verificare che  $\mathbf{U}_k = \mathbf{V}_k$ . Consideriamo una combinazione lineare generica dei vettori  $\mathbf{v}_i$ :

$$\mathbf{z} = \sum_{i=1}^k \eta_i \mathbf{v}_i$$

Mostriamo che  $\mathbf{z} \in \mathbf{U}_k$ . Per l'ipotesi induttiva, esistono  $k-1$  scalari  $\zeta_i$  tali che:

$$\sum_{i=1}^{k-1} \eta_i \mathbf{v}_i = \sum_{i=1}^{k-1} \zeta_i \mathbf{u}_i$$

Utilizzando la relazione:

$$\mathbf{v}_k = \frac{\mathbf{u}_k}{\alpha} + \sum_{i=1}^{k-1} \beta_i \mathbf{u}_i$$

Otteniamo:

$$\mathbf{z} = \frac{\eta_k}{\alpha} \mathbf{u}_k + \eta_k \sum_{i=1}^{k-1} \beta_i \mathbf{u}_i + \sum_{i=1}^{k-1} \zeta_i \mathbf{u}_i = \frac{\eta_k}{\alpha} \mathbf{u}_k + \sum_{i=1}^{k-1} (\zeta_i + \eta_k \beta_i) \mathbf{u}_i$$

Quindi,  $\mathbf{z} \in \mathbf{U}_k$ , e poiché  $\mathbf{z}$  è arbitrario, abbiamo  $\mathbf{V}_k \subset \mathbf{U}_k$ .

Viceversa, se  $\mathbf{z} \in \mathbf{U}_k$ , possiamo scrivere:

$$\mathbf{z} = \sum_{i=1}^k \zeta_i \mathbf{u}_i$$

e utilizzando la relazione:

$$\mathbf{z} = \zeta_k \alpha \left( \mathbf{v}_k - \sum_{i=1}^{k-1} \beta_i \mathbf{u}_i \right) + \sum_{i=1}^{k-1} \zeta_i \mathbf{u}_i = \zeta_k \alpha \mathbf{v}_k - \sum_{i=1}^{k-1} (\zeta_i - \zeta_k \alpha \beta_i) \mathbf{u}_i$$

Per l'ipotesi induttiva, esistono  $k-1$  scalari  $\omega_i$  tali che:

$$\sum_{i=1}^{k-1} (\zeta_i - \zeta_k \alpha \beta_i) \mathbf{u}_i = \sum_{i=1}^{k-1} \omega_i \mathbf{v}_i$$

Quindi,  $\mathbf{z} \in \mathbf{V}_k$ . Poiché  $\mathbf{z}$  è arbitrario, abbiamo  $\mathbf{U}_k \subset \mathbf{V}_k$ , e quindi  $\mathbf{U}_k = \mathbf{V}_k$ . ■

Questo teorema porta al seguente algoritmo per l'ortonormalizzazione di un insieme di vettori:

- **Input:**  $n$  vettori linearmente indipendenti  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$
- **Passo 1:** Calcola il primo vettore ortonormale:
  - $\mathbf{u}_1 \leftarrow \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|_2}$
- **Passo 2:** Per ogni  $k$  da 2 a  $n$ :
  - Calcola il vettore ausiliario  $\mathbf{w}_k$ :

$$\mathbf{w}_k \leftarrow \mathbf{v}_k - \sum_{i=1}^{k-1} (\mathbf{v}_k \cdot \mathbf{u}_i) \mathbf{u}_i$$

- Normalizza  $\mathbf{w}_k$  per ottenere il vettore ortonormale  $\mathbf{u}_k$ :

$$\mathbf{u}_k \leftarrow \frac{\mathbf{w}_k}{\|\mathbf{w}_k\|_2}$$

- **Output:** I vettori  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$  sono ortonormali.

## Osservazione

Consideriamo la matrice  $\mathbf{Q} \in \mathbb{R}^{m \times k}$  definita come

$$\mathbf{Q} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k],$$

dove le colonne  $\mathbf{u}_j$  sono i vettori ortonormali ottenuti tramite il procedimento di Gram-Schmidt. L'ortonormalità di questi vettori è espressa dalla seguente proprietà della matrice  $\mathbf{Q}$ :

$$\mathbf{Q}^T \mathbf{Q} = \mathbf{I} \in \mathbb{R}^{k \times k},$$

dove  $\mathbf{I}$  è la matrice identità di dimensione  $k \times k$ .

È importante notare che, se  $k < m$ , la matrice  $\mathbf{Q}\mathbf{Q}^T$  non è in generale la matrice identità di dimensione  $m \times m$ . In realtà, la matrice  $\mathbf{Q}\mathbf{Q}^T$  soddisfa le seguenti proprietà:

$$\mathbf{P} = \mathbf{Q}\mathbf{Q}^T \quad \text{con} \quad \mathbf{P}^2 = \mathbf{P} \quad \text{e} \quad \mathbf{P}^T = \mathbf{P},$$

dove  $\mathbf{P}$  è una matrice di proiezione ortogonale. Quindi, la matrice  $\mathbf{Q}\mathbf{Q}^T$  è un *proiettore ortogonale*.

## Osservazione

Data una base di  $k$  vettori linearmente indipendenti

$$\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\} \text{ in } \mathbb{K}^n$$

che sono ortonormali a coppie, è sempre possibile estendere questo insieme a una base ortonormale completa di  $\mathbb{K}^n$ .

In particolare, è possibile trovare  $n - k$  vettori aggiuntivi

$$\{\mathbf{u}_{k+1}, \mathbf{u}_{k+2}, \dots, \mathbf{u}_n\}$$

in modo che l'insieme completo

$$\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$$

sia una base ortonormale di  $\mathbb{K}^n$ . Infatti, partendo da una base qualsiasi, è sufficiente rimuovere i vettori linearmente dipendenti dai vettori  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ . I vettori rimanenti, insieme ai vettori iniziali, formeranno una base ortonormale. Utilizzando il procedimento di ortonormalizzazione, otteniamo una base ortonormale completa, senza modificare i vettori  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$  originali.



## 6. Determinanti

### 6.1 Definizione Assiomatica

In questa sezione, introdurremo il concetto di determinante in modo assiomatico. Definiremo il determinante come una funzione

$$|\cdot| : \mathbb{K}^{n \times n} \rightarrow \mathbb{K},$$

che associa a ogni matrice quadrata  $\mathbf{A} \in \mathbb{K}^{n \times n}$  uno scalare, denotato come  $|\mathbf{A}|$ . La funzione determinante sarà definita in modo che rispetti una serie di proprietà che enunceremo come assiomi.

Per semplificare l'esposizione, utilizzeremo una notazione matriciale basata sulle colonne. Indicheremo con  $\mathbf{A}_{\bullet,j}$  la colonna  $j$ -esima della matrice  $\mathbf{A}$ :

$$\mathbf{A} = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & A_{n-1,n} \\ A_{n,1} & \cdots & A_{n,n-1} & A_{n,n} \end{pmatrix},$$
$$\mathbf{A}_{\bullet,j} = \begin{pmatrix} A_{1,j} \\ A_{2,j} \\ \vdots \\ A_{n,j} \end{pmatrix},$$

in modo da pensare alla matrice  $\mathbf{A}$  come *partizionata per colonne*:

$$\mathbf{A} = (\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}).$$

Ogni colonna  $\mathbf{A}_{\bullet,j}$  della matrice  $\mathbf{A}$  è un vettore colonna e può essere espressa come combinazione lineare dei vettori della base canonica di  $\mathbb{K}^n$ :

$$\mathbf{A}_{\bullet,j} = \sum_{k=1}^n A_{k,j} \mathbf{e}_k,$$

dove i coefficienti sono le componenti della matrice nella colonna considerata. La dipendenza della funzione determinante dalle colonne della matrice può essere scritta come:

$$|\mathbf{A}| := |\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}|.$$

Enunceremo ora le tre proprietà fondamentali che definiscono assiomaticamente la funzione determinante.

**Definition 6.1 — Determinante.** Definiamo il determinante con le seguenti proprietà fondamentali:

1. **Linearità nelle colonne:** Il determinante è una funzione *multi-lineare* rispetto alle colonne della matrice. Ciò significa che:

$$\begin{aligned} |\dots, \lambda \mathbf{a}, \dots| &= \lambda |\dots, \mathbf{a}, \dots|, \\ |\dots, \mathbf{a} + \mathbf{b}, \dots| &= |\dots, \mathbf{a}, \dots| + |\dots, \mathbf{b}, \dots|. \end{aligned}$$

2. **Proprietà di nullità:** Il determinante è nullo se due colonne della matrice sono uguali:

$$|\dots, \mathbf{a}, \mathbf{a}, \dots| = 0.$$

3. **Determinante della matrice identità:** Il determinante della matrice identità è pari a 1:

$$|\mathbf{I}| = |\mathbf{e}_1, \dots, \mathbf{e}_n| = 1,$$

dove  $\mathbf{e}_i$  sono i vettori della base canonica in  $\mathbb{K}^n$ .

Queste tre proprietà sono sufficienti per garantire l'esistenza e l'unicità della funzione determinante e implicano numerose conseguenze importanti, che esploreremo nel seguito. Tuttavia, molte di queste conseguenze dipendono solo dalle prime due proprietà e non dalla terza.

Per chiarire questa distinzione, introdurremo una notazione alternativa per una funzione determinante più generale, denotata come  $\mathcal{D}(\mathbf{A})$ . Questa funzione soddisfa le proprietà 1 e 2, ma non necessariamente la proprietà 3, e può assumere un qualsiasi valore non nullo. Utilizzeremo il simbolo  $\mathcal{D}(\mathbf{A})$  invece di  $|\mathbf{A}|$  ogni volta che non è necessaria la proprietà 3.

#### Osservazione

In alcuni casi particolari, è possibile definire facilmente una formula che esprime una funzione con le proprietà di un determinante. Consideriamo i seguenti esempi:

- **Caso  $n = 1$ :** Per una matrice  $1 \times 1$ , il determinante è semplicemente l'elemento della matrice:

$$|A_{1,1}| = A_{1,1}.$$

- **Caso  $n = 2$ :** Per una matrice  $2 \times 2$ , il determinante può essere calcolato con la formula:

$$\begin{vmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{vmatrix} = A_{1,1}A_{2,2} - A_{2,1}A_{1,2}.$$

- **Caso  $n = 3$ :** Per una matrice  $3 \times 3$ , il determinante è dato da:

$$\begin{vmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{vmatrix} = \begin{aligned} &A_{1,1}A_{2,2}A_{3,3} + A_{1,2}A_{2,3}A_{3,1} \\ &+ A_{2,1}A_{3,2}A_{1,3} - A_{1,3}A_{2,2}A_{3,1} \\ &- A_{1,2}A_{2,1}A_{3,3} - A_{1,1}A_{2,3}A_{3,2}. \end{aligned}$$

- **Matrici triangolari superiori:** Per una matrice triangolare superiore  $n \times n$ , il determinante è il prodotto degli elementi sulla diagonale principale:

$$\begin{vmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,n} \\ 0 & A_{2,2} & \dots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_{n,n} \end{vmatrix} = A_{1,1}A_{2,2} \cdots A_{n,n}.$$

Questa formula è valida anche per le matrici triangolari inferiori e per le matrici diagonali.

Si invita il lettore a verificare che tutte e tre le proprietà assiomatiche dei determinanti sono soddisfatte dalle formule sopra esposte.

## 6.2 Alcune proprietà dei determinanti

**Lemma 6.1 — Prodotto per uno scalare.** Dalla proprietà 1 segue immediatamente che per ogni matrice  $\mathbf{A} \in \mathbb{K}^{n \times n}$  e per ogni scalare  $\lambda$ , vale la seguente relazione:

$$\mathcal{D}(\lambda \mathbf{A}) = \lambda^n \mathcal{D}(\mathbf{A})$$

*Proof.* Consideriamo una matrice  $\mathbf{A}$  con colonne  $\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}$ . Applicando la proprietà multi-lineare del determinante, otteniamo:

$$\begin{aligned} \mathcal{D}(\lambda \mathbf{A}_{\bullet,1}, \lambda \mathbf{A}_{\bullet,2}, \lambda \mathbf{A}_{\bullet,3}, \dots, \lambda \mathbf{A}_{\bullet,n}) \\ &= \lambda \mathcal{D}(\mathbf{A}_{\bullet,1}, \lambda \mathbf{A}_{\bullet,2}, \lambda \mathbf{A}_{\bullet,3}, \dots, \lambda \mathbf{A}_{\bullet,n}) \\ &= \lambda^2 \mathcal{D}(\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \lambda \mathbf{A}_{\bullet,3}, \dots, \lambda \mathbf{A}_{\bullet,n}) \\ &= \dots \\ &= \lambda^n \mathcal{D}(\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}) \end{aligned}$$

Da qui, il lemma segue immediatamente. ■

**Osservazione: Somma di matrici**

Dalla multilinearità del determinante, espressa nella proprietà 1 di Definition 6.1, segue un risultato importante ma negativo riguardo al determinante della somma di due matrici. In particolare, non è vero che

$$\mathcal{D}(\mathbf{A} + \mathbf{B}) = \mathcal{D}(\mathbf{A}) + \mathcal{D}(\mathbf{B}).$$

Infatti, consideriamo la somma di due matrici  $\mathbf{A}$  e  $\mathbf{B}$ , che può essere scritta in forma partizionata per colonne come segue:

$$\mathbf{A} + \mathbf{B} = [\mathbf{A}_{\bullet,1} + \mathbf{B}_{\bullet,1}, \mathbf{A}_{\bullet,2} + \mathbf{B}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n} + \mathbf{B}_{\bullet,n}].$$

Applicando la proprietà di multilinearità del determinante, otteniamo:

$$\begin{aligned} \mathcal{D}(\mathbf{A} + \mathbf{B}) &= \mathcal{D}(\mathbf{A}_{\bullet,1} + \mathbf{B}_{\bullet,1}, \mathbf{A}_{\bullet,2} + \mathbf{B}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n} + \mathbf{B}_{\bullet,n}) \\ &= \sum_{j=1}^n \mathcal{D}(\mathbf{A}_{\bullet,1}, \dots, \mathbf{A}_{\bullet,j} + \mathbf{B}_{\bullet,j}, \dots, \mathbf{A}_{\bullet,n} + \mathbf{B}_{\bullet,n}) \\ &= \mathcal{D}(\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}) + \left( \begin{array}{c} \text{altri termini derivanti} \\ \text{dalla combinazione con } \mathbf{B} \end{array} \right) \\ &\neq \mathcal{D}(\mathbf{A}) + \mathcal{D}(\mathbf{B}). \end{aligned}$$

In generale, il determinante della somma di due matrici è una combinazione molto più complessa dei determinanti delle singole matrici. Invitiamo il lettore a costruire un controesempio per illustrare questo fatto.

Sempre dalla proprietà 1 di Definition 6.1 segue immediatamente il seguente lemma.

**Lemma 6.2 — Determinante colonna nulla.** Se una colonna è nulla, il determinante è nullo.

*Proof.* Se  $\mathbf{v}_i = \mathbf{0}$ , il vettore nullo, allora avremo

$$\begin{aligned}\mathcal{D}(\dots, \mathbf{0}, \dots) &= \mathcal{D}(\dots, 0 \cdot \mathbf{0}, \dots), \\ &= 0 \cdot \mathcal{D}(\dots, \mathbf{0}, \dots), \\ &= 0.\end{aligned}$$

■

Dalla proprietà 2 di Definition 6.1 e combinando insieme le proprietà 1 e 2, possiamo derivare diverse conseguenze riguardanti il comportamento del determinante sotto lo scambio di colonne.

**Lemma 6.3 — Scambio di colonne Consecutive.** Se due colonne consecutive di una matrice vengono scambiate, il determinante cambia segno.

*Proof.* Consideriamo una matrice  $\mathbf{A}$  con colonne  $\dots, \mathbf{w}, \mathbf{z}, \dots$ . Per la proprietà 2 di Definition 6.1, sappiamo che:

$$\mathcal{D}(\dots, \mathbf{w}, \mathbf{w}, \dots) = 0 \quad \text{e} \quad \mathcal{D}(\dots, \mathbf{z}, \mathbf{z}, \dots) = 0.$$

Inoltre, per la multilinearità del determinante, abbiamo:

$$\begin{aligned}\mathcal{D}(\dots, \mathbf{w} + \mathbf{z}, \mathbf{w} + \mathbf{z}, \dots) &= \mathcal{D}(\dots, \mathbf{w}, \mathbf{w} + \mathbf{z}, \dots) + \mathcal{D}(\dots, \mathbf{z}, \mathbf{w} + \mathbf{z}, \dots) \\ &= \mathcal{D}(\dots, \mathbf{w}, \mathbf{w}, \dots) + \mathcal{D}(\dots, \mathbf{w}, \mathbf{z}, \dots) \\ &\quad + \mathcal{D}(\dots, \mathbf{z}, \mathbf{w}, \dots) + \mathcal{D}(\dots, \mathbf{z}, \mathbf{z}, \dots).\end{aligned}$$

Sappiamo che:

$$\mathcal{D}(\dots, \mathbf{w}, \mathbf{w}, \dots) = 0 \quad \text{e} \quad \mathcal{D}(\dots, \mathbf{z}, \mathbf{z}, \dots) = 0,$$

quindi:

$$\mathcal{D}(\dots, \mathbf{w} + \mathbf{z}, \mathbf{w} + \mathbf{z}, \dots) = \mathcal{D}(\dots, \mathbf{w}, \mathbf{z}, \dots) + \mathcal{D}(\dots, \mathbf{z}, \mathbf{w}, \dots).$$

Ma poiché:

$$\mathcal{D}(\dots, \mathbf{w} + \mathbf{z}, \mathbf{w} + \mathbf{z}, \dots) = 0,$$

ne segue che:

$$0 = \mathcal{D}(\dots, \mathbf{w}, \mathbf{z}, \dots) + \mathcal{D}(\dots, \mathbf{z}, \mathbf{w}, \dots),$$

ovvero:

$$\mathcal{D}(\dots, \mathbf{w}, \mathbf{z}, \dots) = -\mathcal{D}(\dots, \mathbf{z}, \mathbf{w}, \dots).$$

Quindi, scambiando due colonne consecutive, il determinante cambia segno. ■

Questa proprietà si estende anche allo scambio di due colonne in qualsiasi posizione, non solo quelle adiacenti. Prima di approfondire questo, è utile considerare il seguente risultato fondamentale.

**Lemma 6.4 — Colonne Uguali.** Se due colonne di una matrice sono identiche, il determinante è nullo.

*Proof.* Supponiamo che le colonne  $\mathbf{v}_i$  e  $\mathbf{v}_j$  siano uguali, con  $i < j$ . Possiamo scambiare la colonna  $\mathbf{v}_i$  con le colonne adiacenti fino a portarla accanto a  $\mathbf{v}_j$ .

$$\begin{aligned} \mathcal{D}(\dots, \mathbf{v}_i, \mathbf{v}_{i+1}, \mathbf{v}_{i+2}, \dots, \mathbf{v}_j, \dots) \\ &= (-1) \mathcal{D}(\dots, \mathbf{v}_{i+1}, \mathbf{v}_i, \mathbf{v}_{i+2}, \dots, \mathbf{v}_j, \dots), \\ &= (-1)^2 \mathcal{D}(\dots, \mathbf{v}_{i+1}, \mathbf{v}_{i+2}, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots), \\ &= \dots \\ &= \sigma \mathcal{D}(\dots, \mathbf{v}_{i+1}, \mathbf{v}_{i+2}, \dots, \mathbf{v}_i, \mathbf{v}_j, \dots) \end{aligned}$$

dove  $\sigma = (-1)^{j-i}$ , quindi può essere solo  $\pm 1$ . Sostituendo  $\mathbf{v}_i = \mathbf{v}_j = \mathbf{a}$ , secondo la Proprietà 2 di Definition 6.1, otteniamo:

$$\mathcal{D}(\dots, \mathbf{v}_{i+1}, \dots, \mathbf{a}, \mathbf{a}, \dots) = 0.$$

■

**Lemma 6.5 — Scambio di Colonne.** Se due colonne qualsiasi, ad esempio la colonna  $i$ -esima e la colonna  $j$ -esima (con  $i \neq j$ ), vengono scambiate, il determinante cambia segno.

*Proof.* Il ragionamento è simile a quello utilizzato per le colonne consecutive. Consideriamo:

$$\begin{aligned} 0 &= \mathcal{D}(\dots, \mathbf{w} + \mathbf{z}, \dots, \mathbf{w} + \mathbf{z}, \dots), \\ &= \mathcal{D}(\dots, \mathbf{w}, \dots, \mathbf{w} + \mathbf{z}, \dots) + \mathcal{D}(\dots, \mathbf{z}, \dots, \mathbf{w} + \mathbf{z}, \dots), \\ &= \mathcal{D}(\dots, \mathbf{w}, \dots, \mathbf{w}, \dots) + \mathcal{D}(\dots, \mathbf{w}, \dots, \mathbf{z}, \dots) \\ &\quad + \mathcal{D}(\dots, \mathbf{z}, \dots, \mathbf{w}, \dots) + \mathcal{D}(\dots, \mathbf{z}, \dots, \mathbf{z}, \dots). \end{aligned}$$

Il primo e l'ultimo termine sono nulli per il Lemma Lemma 6.4, quindi otteniamo:

$$0 = \mathcal{D}(\dots, \mathbf{w}, \dots, \mathbf{z}, \dots) + \mathcal{D}(\dots, \mathbf{z}, \dots, \mathbf{w}, \dots),$$

dove  $\mathbf{w}$  e  $\mathbf{z}$  possono essere in qualsiasi posizione, non necessariamente adiacenti. ■

Combinando il Lemma Lemma 6.4 con la Proprietà 1 di Definition 6.1, otteniamo un risultato molto interessante che sarà utilizzato nel seguito.

**Lemma 6.6 — Combinazione Lineare di Colonne.** Se ad una colonna si somma una qualsiasi combinazione lineare delle altre colonne (escludendo quella in esame), il valore del determinante rimane invariato.

*Proof.* Sia

$$\mathbf{b} = \sum_{\substack{j=1 \\ j \neq i}}^n \beta_j \mathbf{v}_j,$$

con  $\beta_1, \dots, \beta_n$  scalari qualsiasi. Consideriamo:

$$\begin{aligned}
& \mathcal{D}(\dots, \mathbf{v}_{i-1}, \mathbf{v}_i + \mathbf{b}, \mathbf{v}_{i+1}, \dots) \\
&= \mathcal{D}\left(\dots, \mathbf{v}_{i-1}, \mathbf{v}_i + \sum_{\substack{j=1 \\ j \neq i}}^n \beta_j \mathbf{v}_j, \mathbf{v}_{i+1}, \dots\right), \\
&= \mathcal{D}(\dots, \mathbf{v}_{i-1}, \mathbf{v}_i, \mathbf{v}_{i+1}, \dots) \\
&+ \sum_{\substack{j=1 \\ j \neq i}}^n \beta_j \mathcal{D}(\dots, \mathbf{v}_{i-1}, \mathbf{v}_j, \mathbf{v}_{i+1}, \dots).
\end{aligned}$$

Poiché, per il Lemma Lemma 6.5, abbiamo

$$\mathcal{D}(\dots, \mathbf{v}_{i-1}, \mathbf{v}_j, \mathbf{v}_{i+1}, \dots) = 0 \quad \text{per } j \neq i,$$

segue che

$$\mathcal{D}(\dots, \mathbf{v}_{i-1}, \mathbf{v}_i + \mathbf{b}, \mathbf{v}_{i+1}, \dots) = \mathcal{D}(\dots, \mathbf{v}_{i-1}, \mathbf{v}_i, \mathbf{v}_{i+1}, \dots).$$

■

### 6.3 Esistenza ed unicità del determinante

**Theorem 6.1 — Esistenza ed Unicità.** Esiste una e una sola funzione determinante che soddisfa le proprietà 1, 2 e 3 di Definition 6.1, e tale funzione è denotata con il simbolo  $|\cdot|$ .

*Proof.* Dato che ogni colonna della matrice può essere espressa come combinazione lineare dei vettori della base canonica:

$$\mathbf{A}_{\bullet, j} = \sum_{k=1}^n A_{k,j} \mathbf{e}_k,$$

e utilizzando la multilinearità del determinante, otteniamo:

$$\begin{aligned}
\mathcal{D}(\mathbf{A}_{\bullet, 1}, \dots, \mathbf{A}_{\bullet, n}) &= \mathcal{D}\left(\sum_{i_1=1}^n A_{i_1, 1} \mathbf{e}_{i_1}, \sum_{i_2=1}^n A_{i_2, 2} \mathbf{e}_{i_2}, \dots, \sum_{i_n=1}^n A_{i_n, n} \mathbf{e}_{i_n}\right), \\
&= \sum_{i_1=1}^n A_{i_1, 1} \mathcal{D}\left(\mathbf{e}_{i_1}, \sum_{i_2=1}^n A_{i_2, 2} \mathbf{e}_{i_2}, \dots, \sum_{i_n=1}^n A_{i_n, n} \mathbf{e}_{i_n}\right), \\
&= \sum_{i_1=1}^n A_{i_1, 1} \sum_{i_2=1}^n A_{i_2, 2} \cdots \sum_{i_n=1}^n A_{i_n, n} \mathcal{D}(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \dots, \mathbf{e}_{i_n}).
\end{aligned} \tag{6.1}$$

Notiamo che la sommatoria finale in Equation 6.1 coinvolge  $n^n$  termini, ma solo  $n!$  di questi sono non nulli. Infatti, i termini con colonne uguali sono nulli per il lemma Lemma 6.4. Gli unici termini non nulli sono quelli in cui gli indici  $i_1, \dots, i_n$  sono distinti e formano una permutazione dei primi  $n$  numeri interi.

Ogni permutazione può essere ottenuta tramite una serie di scambi, e il lemma Lemma 6.5 implica che:

$$\mathcal{D}(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \dots, \mathbf{e}_{i_n}) = \sigma(i_1, i_2, \dots, i_n) \mathcal{D}(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n),$$

dove  $\sigma(i_1, i_2, \dots, i_n)$  è il segno della permutazione, che può essere  $+1$  o  $-1$  e indica il numero di scambi necessari per trasformare la sequenza  $1, 2, \dots, n$  in  $i_1, i_2, \dots, i_n$ .

Quindi, possiamo riscrivere Equation 6.1 come:

$$\begin{aligned} \mathcal{D}(\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}) \\ = \sum_{(i_1, i_2, \dots, i_n) \in \Pi(n)} \sigma(i_1, i_2, \dots, i_n) A_{i_1,1} A_{i_2,2} \cdots A_{i_n,n} \mathcal{D}(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n). \end{aligned} \quad (6.2)$$

Poiché  $\mathcal{D}(\mathbf{I}) \neq 0$ , possiamo definire:

$$|\mathbf{A}| = \frac{\mathcal{D}(\mathbf{A})}{\mathcal{D}(\mathbf{I})} = \sum_{(i_1, i_2, \dots, i_n) \in \Pi(n)} \sigma(i_1, i_2, \dots, i_n) A_{i_1,1} A_{i_2,2} \cdots A_{i_n,n}. \quad (6.3)$$

Le tre proprietà assiomatiche dei determinanti sono verificate dalla funzione  $|\mathbf{A}|$ . Le prime due proprietà derivano dal fatto che  $|\mathbf{A}|$  coincide con  $\mathcal{D}(\mathbf{A})$  fino a una costante moltiplicativa, mentre la terza proprietà è garantita dalla normalizzazione rispetto a  $\mathcal{D}(\mathbf{I})$ . ■

Si noti che la dimostrazione è costruttiva, poiché suggerisce una formula per il calcolo del determinante di una matrice, anche se non particolarmente pratica. Questo fatto è enunciato come il seguente corollario.

**Corollary 6.1 — Corollario.** Per ogni matrice  $\mathbf{A} \in \mathbb{K}^n$ , il determinante può essere calcolato come:

$$|\mathbf{A}| = \sum_{(i_1, i_2, \dots, i_n) \in \Pi(n)} \sigma(i_1, i_2, \dots, i_n) A_{i_1,1} A_{i_2,2} \cdots A_{i_n,n},$$

dove  $\sigma(i_1, i_2, \dots, i_n)$  è il segno della permutazione  $(i_1, i_2, \dots, i_n)$ , definito come nella dimostrazione del teorema Theorem 6.1.

Inoltre, dalla dimostrazione del teorema si ricava anche un'altra proprietà interessante che sarà utilizzata successivamente.

**Corollary 6.2 — Corollario.** Se  $\mathcal{D}(\mathbf{A})$  soddisfa le proprietà 1 e 2 di Definition 6.1 allora

$$\mathcal{D}(\mathbf{A}) = |\mathbf{A}| \mathcal{D}(\mathbf{I}).$$

## 6.4 Determinanti del prodotto o teorema di Binet<sup>1</sup>

**Theorem 6.2 — Teorema di Binet.** Siano  $\mathbf{A}$  e  $\mathbf{B}$  due matrici quadrate dello stesso ordine. Allora:

$$|\mathbf{AB}| = |\mathbf{A}| |\mathbf{B}|.$$

*Proof.* Consideriamo la matrice prodotto  $\mathbf{C} = \mathbf{AB}$ . Ogni colonna della matrice  $\mathbf{C}$  può essere espressa come:

$$\mathbf{C}_{\bullet,j} = \mathbf{A} \mathbf{B}_{\bullet,j},$$

dove  $\mathbf{B}_{\bullet,j}$  è la  $j$ -esima colonna della matrice  $\mathbf{B}$ . Quindi:

$$\mathbf{C} = \mathbf{AB} = \mathbf{A}[\mathbf{B}_{\bullet,1}, \dots, \mathbf{B}_{\bullet,n}] = [\mathbf{A} \mathbf{B}_{\bullet,1}, \dots, \mathbf{A} \mathbf{B}_{\bullet,n}].$$

Il determinante della matrice prodotto  $\mathbf{C}$  è:

$$|\mathbf{C}| = |\mathbf{AB}| = |\mathbf{A} \mathbf{B}_{\bullet,1}, \mathbf{A} \mathbf{B}_{\bullet,2}, \dots, \mathbf{A} \mathbf{B}_{\bullet,n}|.$$

Definiamo una funzione:

<sup>1</sup>Jacques Philippe Marie Binet (1786-1856)

$$\mathcal{D}_A(\mathbf{v}_1, \dots, \mathbf{v}_n) = |\mathbf{A}\mathbf{v}_1, \dots, \mathbf{A}\mathbf{v}_n|.$$

Perciò, possiamo scrivere:

$$|\mathbf{A}\mathbf{B}| = \mathcal{D}_A(\mathbf{B}).$$

Ora verifichiamo che  $\mathcal{D}_A$  soddisfa le proprietà di multilinearità e alternanza:

**1. Multilinearità:**

- Se moltiplichiamo una colonna  $\mathbf{v}_k$  per uno scalare  $\lambda$ , otteniamo:

$$\begin{aligned} \mathcal{D}_A(\mathbf{v}_1, \dots, \lambda \mathbf{v}_k, \dots, \mathbf{v}_n) &= |\mathbf{A}\mathbf{v}_1, \dots, \lambda \mathbf{A}\mathbf{v}_k, \dots, \mathbf{A}\mathbf{v}_n|, \\ &= \lambda |\mathbf{A}\mathbf{v}_1, \dots, \mathbf{A}\mathbf{v}_k, \dots, \mathbf{A}\mathbf{v}_n|, \\ &= \lambda \mathcal{D}_A(\mathbf{v}_1, \dots, \mathbf{v}_k, \dots, \mathbf{v}_n). \end{aligned}$$

- Se sommiamo una colonna  $\mathbf{v}_k$  a una combinazione lineare delle altre colonne, otteniamo:

$$\begin{aligned} \mathcal{D}_A(\mathbf{v}_1, \dots, \mathbf{v}_k + \mathbf{w}_k, \dots, \mathbf{v}_n) &= |\mathbf{A}\mathbf{v}_1, \dots, \mathbf{A}(\mathbf{v}_k + \mathbf{w}_k), \dots, \mathbf{A}\mathbf{v}_n|, \\ &= |\mathbf{A}\mathbf{v}_1, \dots, \mathbf{A}\mathbf{v}_k + \mathbf{A}\mathbf{w}_k, \dots, \mathbf{A}\mathbf{v}_n|, \\ &= \mathcal{D}_A(\mathbf{v}_1, \dots, \mathbf{v}_k, \dots, \mathbf{v}_n) + \mathcal{D}_A(\mathbf{v}_1, \dots, \mathbf{w}_k, \dots, \mathbf{v}_n). \end{aligned}$$

**2. Alternanza:**

- Se due colonne  $\mathbf{v}_k$  e  $\mathbf{v}_l$  sono uguali, allora:

$$\begin{aligned} \mathcal{D}_A(\mathbf{v}_1, \dots, \mathbf{v}_k, \dots, \mathbf{v}_k, \dots, \mathbf{v}_n) &= |\mathbf{A}\mathbf{v}_1, \dots, \mathbf{A}\mathbf{v}_k, \dots, \mathbf{A}\mathbf{v}_k, \dots, \mathbf{A}\mathbf{v}_n| \\ &= 0. \end{aligned}$$

Quindi,  $\mathcal{D}_A$  soddisfa le proprietà 1 e 2 del determinante. Dalla formula del determinante e considerando che:

$$\mathcal{D}_A(\mathbf{I}) = |\mathbf{A}\mathbf{e}_1, \dots, \mathbf{A}\mathbf{e}_n| = |\mathbf{A}|,$$

otteniamo:

$$\mathcal{D}_A(\mathbf{B}) = |\mathbf{B}| \cdot \mathcal{D}_A(\mathbf{I}) = |\mathbf{A}| \cdot |\mathbf{B}|.$$

Pertanto, abbiamo dimostrato che:

$$|\mathbf{A}\mathbf{B}| = |\mathbf{A}| \cdot |\mathbf{B}|.$$

■

## 6.5 Determinante della matrice inversa

La formula del prodotto di determinanti permette di correlare immediatamente il determinante di una matrice con il determinante della matrice inversa. Infatti si ha che

$$1 = |\mathbf{I}| = |\mathbf{A}\mathbf{A}^{-1}| = |\mathbf{A}| |\mathbf{A}^{-1}|,$$

da cui si deduce che

$$|\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|}.$$

## 6.6 Dipendenza e indipendenza lineare

I determinanti possono essere usati per vedere se un insieme di vettori è o non è linearmente dipendente. Consideriamo prima il caso di  $n$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  in  $\mathbb{K}^n$ .

In tal caso possiamo calcolare il determinante della matrice costituita dalle colonne  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  e vale il teorema

## 6.7 Indipendenza Lineare

Un insieme di  $n$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  in  $\mathbb{K}^n$  è linearmente indipendente se e solo se il determinante di questi vettori è diverso da zero:

$$\mathcal{D}(\mathbf{v}_1, \dots, \mathbf{v}_n) \neq 0.$$

*Proof.* Supponiamo che i vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  siano linearmente dipendenti. Questo significa che esistono coefficienti  $\beta_1, \beta_2, \dots, \beta_n$ , non tutti nulli, tali che:

$$\mathbf{0} = \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2 + \dots + \beta_n \mathbf{v}_n.$$

Senza perdita di generalità, possiamo supporre che il coefficiente  $\beta_i$  sia diverso da zero. Da questa ipotesi, possiamo esprimere  $\mathbf{v}_i$  come:

$$\begin{aligned} \mathbf{v}_i &= -\frac{\beta_1}{\beta_i} \mathbf{v}_1 - \dots - \frac{\beta_{i-1}}{\beta_i} \mathbf{v}_{i-1} - \frac{\beta_{i+1}}{\beta_i} \mathbf{v}_{i+1} - \dots - \frac{\beta_n}{\beta_i} \mathbf{v}_n \\ &= -\frac{1}{\beta_i} \sum_{j \neq i} \beta_j \mathbf{v}_j. \end{aligned}$$

Poiché possiamo aggiungere una qualsiasi combinazione lineare delle altre colonne della matrice  $\mathbf{A}$  alla  $i$ -esima colonna senza modificare il valore del determinante, come indicato nel lemma~Lemma 6.6, possiamo scrivere:

$$\begin{aligned} \mathcal{D}(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n) &= \mathcal{D}\left(\mathbf{v}_1, \dots, \mathbf{v}_i + \frac{1}{\beta_i} \sum_{j \neq i} \beta_j \mathbf{v}_j, \dots, \mathbf{v}_n\right) \\ &= \mathcal{D}(\mathbf{v}_1, \dots, \mathbf{0}, \dots, \mathbf{v}_n) \\ &= 0. \end{aligned}$$

Viceversa, se i vettori  $\mathbf{v}_1, \dots, \mathbf{v}_n$  sono linearmente indipendenti, allora essi formano una base per  $\mathbb{K}^n$ . Di conseguenza, ogni vettore della base canonica può essere espresso come combinazione lineare dei vettori  $\mathbf{v}_1, \dots, \mathbf{v}_n$ . In particolare, per ogni  $i = 1, \dots, n$  abbiamo:

$$\mathbf{e}_i = \sum_{k=1}^n A_{i,k} \mathbf{v}_k.$$

Consideriamo il determinante della matrice formata dalle colonne  $\mathbf{e}_i$ :

$$\begin{aligned}
1 &\neq \mathcal{D}(\mathbf{I}) \\
&= \mathcal{D}(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) \\
&= \mathcal{D}\left(\sum_{k_1=1}^n A_{1,k_1} \mathbf{v}_{k_1}, \sum_{k_2=1}^n A_{2,k_2} \mathbf{v}_{k_2}, \dots, \sum_{k_n=1}^n A_{n,k_n} \mathbf{v}_{k_n}\right) \\
&= \sum_{k_1, k_2, \dots, k_n=1}^n A_{1,k_1} A_{2,k_2} \cdots A_{n,k_n} \mathcal{D}(\mathbf{v}_{k_1}, \mathbf{v}_{k_2}, \dots, \mathbf{v}_{k_n}) \\
&= \sum_{k_1, k_2, \dots, k_n \in S} \sigma(k_1, k_2, \dots, k_n) A_{1,k_1} A_{2,k_2} \cdots A_{n,k_n} \mathcal{D}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) \\
&= C \mathcal{D}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n),
\end{aligned}$$

dove  $C$  è definito come:

$$C = \sum_{k_1, \dots, k_n \in S} \sigma(k_1, \dots, k_n) A_{1,k_1} \cdots A_{n,k_n},$$

e rappresenta una quantità finita, essendo una somma di prodotti finiti. Dato che  $C \neq 0$ , segue che  $\mathcal{D}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) \neq 0$ . ■

È possibile estendere il risultato precedente anche al caso in cui il numero di vettori sia inferiore a  $n$ . Tuttavia, questa generalizzazione richiede l'introduzione di strumenti e concetti aggiuntivi. l'introduzione di qualche strumento ulteriore.

**Definition 6.2 — Sottomatrice.** Data una matrice  $\mathbf{A}$  (che non deve necessariamente essere quadrata), possiamo ottenere una *sottomatrice* selezionando solo gli elementi che appartengono all'intersezione di un insieme specifico di righe e colonne di  $\mathbf{A}$ .

Una sottomatrice può essere indicata con la notazione:

$$\mathbf{A}_{i_1, i_2, \dots, i_s; j_1, j_2, \dots, j_r},$$

dove gli indici  $i_1, i_2, \dots, i_s$  corrispondono alle righe selezionate e  $j_1, j_2, \dots, j_r$  alle colonne selezionate della matrice  $\mathbf{A}$ .

Ad esempio dalla matrice

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 3 & 0 \\ 0 & 2 & 2 & -1 \\ 1 & 1 & -5 & 9 \\ 3 & 5 & 7 & 0 \end{pmatrix},$$

possiamo estrarre la sottomatrice

$$\mathbf{A}_{1,2;1,4} = \begin{pmatrix} 1 & 3 \\ 3 & 5 \end{pmatrix}.$$

Ovviamente una sottomatrice è ancora una matrice e quindi può essere indicata al solito da una qualsiasi lettera maiuscola.

**Definition 6.3 — Minore.** Sia  $\mathbf{A} \in \mathbb{K}^{n \times n}$  una matrice. Il determinante di una qualsiasi sottomatrice di  $\mathbf{A}$  di ordine  $k \times k$ , ottenuta selezionando  $k$  righe e  $k$  colonne, è chiamato *minore* di ordine  $k$  di  $\mathbf{A}$ . La sottomatrice di  $\mathbf{A}$  utilizzata per calcolare questo determinante è denotata come:

$$\mathbf{A}_{i_1, i_2, \dots, i_k; j_1, j_2, \dots, j_k} \in \mathbb{K}^{k \times k}.$$

Ci proponiamo ora di dimostrare il seguente teorema.

**Theorem 6.3 — Minori e dipendenza lineare.** Se i  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  sono linearmente dipendenti, allora tutti i minori di ordine  $k$  della matrice  $\mathbf{A} \in \mathbb{K}^{n \times k}$ , definita come

$$\mathbf{A} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k),$$

sono nulli, a condizione che  $n \geq k^a$ .

<sup>a</sup>Perché è necessario che  $n \geq k$ ? Cosa accadrebbe se invece  $n < k$ ?

*Proof.* Consideriamo un generico minore di ordine  $k$ , che può essere scritto come:

$$\mathbf{A}_{i_1, i_2, \dots, i_k; 1, 2, \dots, k} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k),$$

dove  $\mathbf{w}_i = (\mathbf{v}_i)_{i_1, i_2, \dots, i_k}$  rappresenta il vettore ottenuto dalle componenti con indici di riga  $i_1, i_2, \dots, i_k$  del vettore  $\mathbf{v}_i$ .

Poiché i vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  sono linearmente dipendenti, esiste una combinazione lineare non banale che annulla il vettore zero:

$$\mathbf{0} = \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2 + \dots + \beta_k \mathbf{v}_k.$$

Applicando questa combinazione lineare componente per componente alle righe con indici  $i_1, i_2, \dots, i_k$ , otteniamo:

$$\mathbf{0} = \beta_1 \mathbf{w}_1 + \beta_2 \mathbf{w}_2 + \dots + \beta_k \mathbf{w}_k.$$

Pertanto, i vettori  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$  sono linearmente dipendenti. Di conseguenza, il minore corrispondente è nullo, come dimostra il lemma Lemma 6.6. ■

Possiamo *invertire* l'enunciato di questo teorema<sup>2</sup>. Se esiste almeno un minore di ordine  $k$  non nullo, allora i  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  sono necessariamente linearmente indipendenti.

Questo fornisce un criterio utile per verificare l'indipendenza lineare di un insieme di  $k$  vettori in  $\mathbb{K}^n$ .

Vogliamo ora dimostrare che la relazione tra l'esistenza di almeno un minore di ordine  $k$  non nullo per  $k$  vettori colonna in  $\mathbb{K}^n$  e la loro indipendenza lineare è in realtà un'equivalenza. A tal fine, consideriamo il seguente teorema.

**Theorem 6.4 — Minori e Indipendenza Lineare.** Se i  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in \mathbb{K}^n$  sono linearmente indipendenti e  $k \leq n$ , allora esiste almeno un minore di ordine  $k$  non nullo della matrice  $\mathbf{A} \in \mathbb{K}^{n \times k}$  con  $n \geq k$ , definita come

$$\mathbf{A} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k).$$

La dimostrazione di questo teorema è leggermente più complessa e richiede l'uso di alcuni risultati intermedi, che enunciamo nel lemma seguente.

**Lemma 6.7 — Lemma Intermedio.** Se i  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  sono linearmente indipendenti e sostituiamo il  $i$ -esimo vettore con

$$\mathbf{z} = \mathbf{v}_i + \sum_{\substack{j=1 \\ j \neq i}}^k \beta_j \mathbf{v}_j,$$

per una qualsiasi scelta degli scalari  $\beta_1, \dots, \beta_k$ , allora i  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{z}, \mathbf{v}_{i+1}, \dots,$

<sup>2</sup>Ricordiamo che dalla logica elementare se la proposizione  $\mathcal{A}$  implica  $\mathcal{B}$ , allora la negazione di  $\mathcal{B}$  implica la negazione di  $\mathcal{A}$ .

$\mathbf{v}_k$  rimangono ancora linearmente indipendenti.

*Proof.* Consideriamo la seguente combinazione lineare dei vettori

$$\mathbf{0} = \gamma_1 \mathbf{v}_1 + \gamma_2 \mathbf{v}_2 + \cdots + \gamma_{i-1} \mathbf{v}_{i-1} + \gamma_i \mathbf{z} + \gamma_{i+1} \mathbf{v}_{i+1} + \cdots + \gamma_k \mathbf{v}_k.$$

Sostituendo  $\mathbf{z}$  nella combinazione lineare, otteniamo

$$\begin{aligned} \mathbf{0} &= \gamma_1 \mathbf{v}_1 + \gamma_2 \mathbf{v}_2 + \cdots + \gamma_{i-1} \mathbf{v}_{i-1} + \gamma_i \left( \mathbf{v}_i + \sum_{\substack{j=1 \\ j \neq i}}^k \beta_j \mathbf{v}_j \right) + \gamma_{i+1} \mathbf{v}_{i+1} + \cdots + \gamma_k \mathbf{v}_k \\ &= (\gamma_1 + \gamma_i \beta_1) \mathbf{v}_1 + (\gamma_2 + \gamma_i \beta_2) \mathbf{v}_2 + \cdots + (\gamma_{i-1} + \gamma_i \beta_{i-1}) \mathbf{v}_{i-1} \\ &\quad + \gamma_i \mathbf{v}_i + (\gamma_{i+1} + \gamma_i \beta_{i+1}) \mathbf{v}_{i+1} + \cdots + (\gamma_k + \gamma_i \beta_k) \mathbf{v}_k. \end{aligned}$$

Poiché i vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  sono linearmente indipendenti, i coefficienti delle loro combinazioni devono essere zero. Questo implica

$$\gamma_s + \gamma_i \beta_s = 0 \quad \text{per } s \neq i$$

e

$$\gamma_i = 0.$$

Di conseguenza, si ha che

$$\gamma_s = 0 \quad \text{per } s = 1, 2, \dots, k,$$

ovvero, i vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{i-1}, \mathbf{z}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k$  sono linearmente indipendenti. ■

**Lemma 6.8 — Minori e Combinazioni Lineari.** Sia  $\mathbf{A}$  una matrice le cui colonne sono date dai vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ . I minori di ordine  $k$  rimangono invariati se al  $i$ -esimo vettore si aggiunge una combinazione lineare degli altri vettori. In altre parole, se sostituiamo il vettore  $\mathbf{v}_i$  con il vettore

$$\mathbf{z} = \mathbf{v}_i + \sum_{\substack{j=1 \\ j \neq i}}^k \beta_j \mathbf{v}_j,$$

i minori di ordine  $k$  della matrice risultante rimangono invariati.

*Proof.* Consideriamo un generico minore di ordine  $k$  della matrice trasformata. Questo minore può essere scritto come

$$\mathbf{B} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{i-1}, \mathbf{x}, \mathbf{w}_{i+1}, \dots, \mathbf{w}_k),$$

dove  $\mathbf{w}_i$  è il vettore  $\mathbf{v}_i$  con le righe selezionate e  $\mathbf{x}$  è il vettore  $\mathbf{z}$  con le stesse righe selezionate. Osserviamo che

$$\mathbf{x} = \mathbf{w}_i + \sum_{\substack{j=1 \\ j \neq i}}^k \beta_j \mathbf{w}_j.$$

Poiché il determinante di un minore di ordine  $k$  è invariato quando si sostituisce un vettore con una sua combinazione lineare degli altri vettori, il determinante della matrice  $\mathbf{B}$  rimane lo stesso. Questo risultato è confermato dal lemma Lemma 6.6. ■

**Lemma 6.9 — Minori e Combinazioni Lineari.** Sia  $\mathbf{A}$  una matrice le cui colonne sono i vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ , che sono linearmente indipendenti. Allora, mediante opportune combinazioni lineari dei vettori colonna che non alterano il valore dei minori di ordine  $k$ , è possibile ridurre  $\mathbf{A}$  alla forma:

$$\mathbf{A} = \begin{pmatrix} \mathbf{T} \\ \mathbf{M} \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} T_{1,1} & 0 & \cdots & 0 \\ T_{2,1} & T_{2,2} & & \vdots \\ \vdots & & \ddots & 0 \\ T_{k,1} & T_{k,2} & \cdots & T_{k,k} \end{pmatrix},$$

dove  $T_{i,i} \neq 0$  per  $i = 1, 2, \dots, k$ , a meno di permutazioni delle righe.

*Proof.* La dimostrazione si basa su una procedura costruttiva. Supponiamo inizialmente che

$$(\mathbf{v}_1)_1 \neq 0.$$

Se così non fosse, possiamo scambiare la prima riga con una riga  $i$  in cui

$$(\mathbf{v}_i)_1 \neq 0,$$

esiste almeno una tale riga  $i$ , altrimenti il vettore  $\mathbf{v}_1$  sarebbe nullo e i vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  sarebbero linearmente dipendenti.

Sostituiamo quindi  $\mathbf{v}_i$  con

$$\mathbf{z}_i = \mathbf{v}_i - \frac{(\mathbf{v}_i)_1}{(\mathbf{v}_1)_1} \mathbf{v}_1 = (0, (\mathbf{z}_i)_2, (\mathbf{z}_i)_3, \dots, (\mathbf{z}_i)_k)^T,$$

e otteniamo una nuova matrice  $\mathbf{B}$  nella forma:

$$\mathbf{B} = \begin{pmatrix} (\mathbf{v}_1)_1 & 0 & \cdots & 0 \\ (\mathbf{v}_1)_2 & & & \\ \vdots & \mathbf{C} & & \\ (\mathbf{v}_1)_n & & & \end{pmatrix},$$

Dove  $\mathbf{C}$  è una matrice che può essere ulteriormente trattata. Ripetiamo il procedimento per le sottomatrici ottenute, osservando che una combinazione lineare di zeri resta zero, e quindi la struttura della matrice viene semplificata progressivamente. ■

A questo punto la dimostrazione del teorema Theorem 7.2 è immediata.

## 6.8 Dimostrazione Theorem 7.2

*Proof.* Si consideri il minore  $\mathbf{T}$  del lemma Lemma 6.9 ■

### Osservazione

Il lettore attento<sup>3</sup> noterà che, senza esplicitarlo, stiamo impiegando un processo di eliminazione di Gauss sulla matrice  $\mathbf{B}^T$ . La triangolarizzazione mediante scambio di righe può essere effettuata in  $k$  passi proprio perché  $\mathbf{B}$  è di rango massimo, garantendo così la presenza di un elemento pivotale non nullo in ogni passo. Questo procedimento fornisce anche una tecnica pratica per determinare il rango di una matrice, alternativa alla *metodologia ingenua* di cercare manualmente la sottomatrice quadrata non singolare di dimensioni massime.

Questi risultati si riassumono nel teorema finale.

**Theorem 6.5 — Indipendenza lineare.** I vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  sono linearmente indipendenti se e solo se la matrice  $\mathbf{A}$ , formata dalle loro componenti rispetto alla base canonica, possiede almeno un minore di ordine  $k$  non nullo.

## 6.9 Rango di una matrice

**Definition 6.4 — Rango.** Il rango di una matrice rettangolare  $\mathbf{A}$  è definito come il massimo numero di vettori colonna linearmente indipendenti presenti in  $\mathbf{A}$ . Si denota con  $\text{rank}(\mathbf{A})$ .

I teoremi finora dimostrati offrono un metodo per determinare il rango di una matrice.

**Theorem 6.6 — Rango.** Per una matrice  $\mathbf{A}$ , il rango della matrice è uguale all'ordine del minore più grande che sia non nullo.

*Proof.* Sia  $\mathbf{A}$  una matrice con colonne  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$ . Consideriamo  $k$  colonne della matrice,  $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \dots, \mathbf{a}_{i_k}$ .

1. **Se i vettori  $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \dots, \mathbf{a}_{i_k}$  sono linearmente indipendenti**, allora per il teorema Theorem 6.5, esiste almeno un minore di ordine  $k$  formato da queste colonne che ha determinante non nullo. Questo dimostra che il rango della matrice  $\mathbf{A}$  è almeno  $k$ .
2. **Viceversa, se esiste un minore di ordine  $k$  con determinante non nullo**, formato dalle colonne  $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \dots, \mathbf{a}_{i_k}$ , allora i vettori  $\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \dots, \mathbf{a}_{i_k}$  sono linearmente indipendenti per il teorema Theorem 6.5. Questo implica che il rango della matrice  $\mathbf{A}$  è almeno  $k$ .

In sintesi, il rango di una matrice è il numero massimo di colonne linearmente indipendenti, e questo corrisponde all'ordine del più grande minore non nullo della matrice. ■

## 6.10 Cofattori di una matrice quadrata

Sia  $\mathbf{A}$  una matrice quadrata di ordine  $n$ , che possiamo rappresentare come segue:

$$\mathbf{A} = (\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}),$$

dove  $\mathbf{A}_{\bullet,i}$  denota la  $i$ -esima colonna di  $\mathbf{A}$ .

Il determinante di  $\mathbf{A}$  può essere espresso come una combinazione lineare dei determinanti di matrici ottenute sostituendo ogni colonna della matrice  $\mathbf{A}$  con la base canonica  $\mathbf{e}_i$ . In particolare, possiamo scrivere:

$$\begin{aligned} |\mathbf{A}| &= |\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}|, \\ &= \left| \sum_{i=1}^n A_{i,1} \mathbf{e}_i, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n} \right|, \\ &= \sum_{i=1}^n A_{i,1} |\mathbf{e}_i, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}|. \end{aligned}$$

Dove:

- $A_{i,1}$  sono gli elementi della prima colonna della matrice  $\mathbf{A}$ .
- $\mathbf{e}_i$  è il vettore colonna della base canonica con 1 nella  $i$ -esima posizione e 0 altrove.
- $|\mathbf{e}_i, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}|$  è il minore che si ottiene eliminando la prima colonna e sostituendo la  $i$ -esima colonna con  $\mathbf{e}_i$ .

Questa formula esprime il determinante di  $\mathbf{A}$  come una somma pesata dei determinanti dei minori ottenuti con la sostituzione della colonna  $\mathbf{A}_{\bullet,1}$  con vettori della base canonica.

<sup>3</sup>Si riferisce al lettore con una certa familiarità con le tecniche di algebra lineare.

## 6.11 Cofattore

Definiamo i *cofattori* di una matrice quadrata come i determinanti delle matrici ottenute sostituendo una colonna con un vettore della base canonica.

Per una matrice quadrata  $\mathbf{A}$  di ordine  $n$ , il cofattore associato all'elemento  $\mathbf{A}_{i,j}$  è il determinante della matrice che si ottiene sostituendo la  $j$ -esima colonna della matrice  $\mathbf{A}$  con il vettore colonna della base canonica  $\mathbf{e}_i$ . Indichiamo questo cofattore con il simbolo  $\alpha_{i,j}$ :

$$\alpha_{i,j} = |\mathbf{A}_{\bullet,1}, \dots, \mathbf{A}_{\bullet,j-1}, \mathbf{e}_i, \mathbf{A}_{\bullet,j+1}, \dots, \mathbf{A}_{\bullet,n}|,$$

dove  $\mathbf{A}_{\bullet,k}$  rappresenta la  $k$ -esima colonna della matrice  $\mathbf{A}$ .

## 6.12 Matrice Cofattore

Data una matrice quadrata  $\mathbf{A}$ , la *matrice cofattore* di  $\mathbf{A}$ , indicata con  $\text{cofatt}(\mathbf{A})$ , è una matrice composta dai cofattori  $\alpha_{i,j}$  di  $\mathbf{A}$ . Essa si costruisce sostituendo ogni elemento  $\mathbf{A}_{i,j}$  della matrice originale con il corrispondente cofattore  $\alpha_{i,j}$ . La matrice cofattore è quindi definita come segue:

$$\text{cofatt}(\mathbf{A}) = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,n} \end{pmatrix},$$

dove ciascun elemento  $\alpha_{i,j}$  è il cofattore relativo all'elemento  $\mathbf{A}_{i,j}$  della matrice originale  $\mathbf{A}$ . Introduciamo una notazione conveniente per i cofattori:

$$\alpha_{i,j} = |\mathbf{A}_{\bullet,1}, \dots, \mathbf{A}_{\bullet,j-1}, \mathbf{e}_i, \mathbf{A}_{\bullet,j+1}, \dots, \mathbf{A}_{\bullet,n}| = |\mathbf{A} \stackrel{j}{\leftarrow} \mathbf{e}_i|,$$

dove  $\mathbf{A} \stackrel{j}{\leftarrow} \mathbf{v}$  denota la matrice ottenuta da  $\mathbf{A}$  sostituendo la  $j$ -esima colonna con il vettore colonna  $\mathbf{v}$ .

Con questa notazione, la formula di Cramer può essere espressa come segue:

$$\mathbf{Ax} = \mathbf{b} \quad \Rightarrow \quad x_k = \frac{|\mathbf{A} \stackrel{k}{\leftarrow} \mathbf{b}|}{|\mathbf{A}|}.$$

In questa formula,  $\mathbf{A} \stackrel{k}{\leftarrow} \mathbf{b}$  indica la matrice ottenuta da  $\mathbf{A}$  sostituendo la  $k$ -esima colonna con il vettore  $\mathbf{b}$ .

La multilinearità del determinante può essere espressa come segue:

1. Se sostituiamo la  $k$ -esima colonna di  $\mathbf{A}$  con  $\lambda \mathbf{v}$ , otteniamo

$$|\mathbf{A} \stackrel{k}{\leftarrow} \lambda \mathbf{v}| = \lambda |\mathbf{A} \stackrel{k}{\leftarrow} \mathbf{v}|.$$

2. Se sostituiamo la  $k$ -esima colonna di  $\mathbf{A}$  con la somma  $\mathbf{v} + \mathbf{w}$ , otteniamo

$$|\mathbf{A} \stackrel{k}{\leftarrow} (\mathbf{v} + \mathbf{w})| = |\mathbf{A} \stackrel{k}{\leftarrow} \mathbf{v}| + |\mathbf{A} \stackrel{k}{\leftarrow} \mathbf{w}|.$$

Utilizzando queste proprietà di linearità, possiamo calcolare il determinante attraverso lo *sviluppo rispetto alla prima colonna*:

$$|\mathbf{A}| = \sum_{i=1}^n A_{i,1} |\mathbf{A} \stackrel{1}{\leftarrow} \mathbf{e}_i| = \sum_{i=1}^n A_{i,1} \alpha_{i,1},$$

oppure, in modo simile, attraverso lo *sviluppo rispetto alla  $k$ -esima colonna*:

$$|\mathbf{A}| = \sum_{i=1}^n A_{i,k} |\mathbf{A} \leftarrow^k \mathbf{e}_i| = \sum_{i=1}^n A_{i,k} \alpha_{i,k}.$$

Queste proprietà e formule sono essenziali per il calcolo del determinante e saranno utilizzate nelle prossime sezioni.

Presentiamo ora una proprietà utile per il seguito.

**Theorem 6.7 — Prodotto con la Matrice Cofattore.** Sia  $\mathbf{A}$  una matrice quadrata e  $\alpha_{i,j}$  i suoi cofattori. Allora vale:

$$\delta_{j,k} |\mathbf{A}| = \sum_{i=1}^n A_{i,k} \alpha_{i,j},$$

dove  $\delta_{j,k}$  è il simbolo di Kronecker, che è uguale a 1 se  $j = k$  e 0 altrimenti.

*Proof.* Quando  $j = k$ , il teorema fornisce esattamente la formula dello sviluppo del determinante rispetto alla colonna  $k$ .

Per il caso  $j \neq k$ , consideriamo:

$$\begin{aligned} \sum_{i=1}^n A_{i,k} \alpha_{i,j} &= \sum_{i=1}^n A_{i,k} |\mathbf{A} \leftarrow^j \mathbf{e}_i| \\ &= \sum_{i=1}^n |\mathbf{A} \leftarrow^k A_{i,j} \mathbf{e}_i| \\ &= |\mathbf{A} \leftarrow^k \left( \sum_{i=1}^n A_{i,j} \mathbf{e}_i \right)| \\ &= |\mathbf{A} \leftarrow^k \mathbf{A}_{\bullet,j}| = 0. \end{aligned}$$

Infatti, la matrice  $\mathbf{A} \leftarrow^k \mathbf{A}_{\bullet,j}$  si ottiene sostituendo la colonna  $j$ -esima della matrice  $\mathbf{A}$  con la colonna  $k$ -esima. Se  $j = k$ , otteniamo nuovamente la matrice  $\mathbf{A}$ , altrimenti otteniamo una matrice con due colonne identiche, il cui determinante è nullo, come stabilito nel lemma Lemma 6.4. ■

Enunciamo ora una proprietà che sarà utile nel seguito.

**Theorem 6.8 — Prodotto con la Matrice Cofattore.** Sia  $\mathbf{A}$  una matrice quadrata e  $\alpha_{i,j}$  i suoi cofattori. Allora vale:

$$\delta_{j,k} |\mathbf{A}| = \sum_{i=1}^n A_{i,k} \alpha_{i,j},$$

dove  $\delta_{j,k}$  è il simbolo di Kronecker, che è uguale a 1 se  $j = k$  e 0 altrimenti.

*Proof.* Quando  $j = k$ , il teorema fornisce esattamente la formula dello sviluppo del determinante rispetto alla colonna  $k$ .

Per il caso  $j \neq k$ , consideriamo:

$$\begin{aligned}
\sum_{i=1}^n A_{i,k} \alpha_{i,j} &= \sum_{i=1}^n A_{i,k} |\mathbf{A} \stackrel{j}{\leftarrow} \mathbf{e}_i| \\
&= \sum_{i=1}^n |\mathbf{A} \stackrel{k}{\leftarrow} A_{i,j} \mathbf{e}_i| \\
&= |\mathbf{A} \stackrel{k}{\leftarrow} \left( \sum_{i=1}^n A_{i,j} \mathbf{e}_i \right)| \\
&= |\mathbf{A} \stackrel{k}{\leftarrow} \mathbf{A}_{\bullet,j}| = 0.
\end{aligned}$$

Infatti, la matrice  $\mathbf{A} \stackrel{k}{\leftarrow} \mathbf{A}_{\bullet,j}$  si ottiene sostituendo la colonna  $j$ -esima della matrice  $\mathbf{A}$  con la colonna  $k$ -esima. Se  $j = k$ , otteniamo nuovamente la matrice  $\mathbf{A}$ , altrimenti otteniamo una matrice con due colonne identiche, il cui determinante è nullo, come stabilito nel lemma Lemma 6.4. ■

## 6.13 Rappresentazione della matrice inversa

Il risultato precedente permette di definire una formula per la rappresentazione formale<sup>4</sup> dell'inversa di una matrice quadrata.

Consideriamo, per iniziare, il seguente teorema.

**Lemma 6.10 — Matrice Cofattore.** Sia  $\mathbf{A}$  una matrice quadrata di ordine  $n$  e  $\text{cofatt}(\mathbf{A})$  la sua matrice cofattore. Allora si ha:

$$\text{cofatt}(\mathbf{A})^T \mathbf{A} = |\mathbf{A}| \mathbf{I}.$$

*Proof.* Consideriamo il prodotto della trasposta della matrice cofattore di  $\mathbf{A}$  per la matrice  $\mathbf{A}$ .

Per calcolare l'elemento  $(j, k)$  della matrice prodotto  $\text{cofatt}(\mathbf{A})^T \mathbf{A}$ , dobbiamo considerare la somma:

$$(\text{cofatt}(\mathbf{A})^T \mathbf{A})_{jk} = \sum_{i=1}^n (\text{cofatt}(\mathbf{A}))_{i,j} A_{i,k}.$$

Dove  $(\text{cofatt}(\mathbf{A}))_{i,j} = \alpha_{i,j}$  è il cofattore dell'elemento  $A_{i,j}$  della matrice  $\mathbf{A}$ . Quindi possiamo scrivere:

$$\sum_{i=1}^n (\text{cofatt}(\mathbf{A}))_{i,j} A_{i,k} = \sum_{i=1}^n \alpha_{i,j} A_{i,k}.$$

Applicando il lemma Theorem 6.8, otteniamo:

$$\sum_{i=1}^n \alpha_{i,j} A_{i,k} = \delta_{jk} |\mathbf{A}|.$$

Pertanto, l'elemento  $(j, k)$  della matrice prodotto è:

$$(\text{cofatt}(\mathbf{A})^T \mathbf{A})_{jk} = \delta_{jk} |\mathbf{A}|,$$

che corrisponde all'elemento  $(j, k)$  della matrice  $|\mathbf{A}| \mathbf{I}$ . Di conseguenza:

$$\text{cofatt}(\mathbf{A})^T \mathbf{A} = |\mathbf{A}| \mathbf{I}.$$

■

<sup>4</sup>Si noti bene che si sta parlando di *rappresentazione formale* e non di *calcolo* della matrice inversa.

Il teorema seguente fornisce l'espressione dell'inversa di una matrice  $\mathbf{A}$  con determinante non nullo.

**Theorem 6.9 — Matrice Inversa.** Sia  $\mathbf{A}$  una matrice quadrata di ordine  $n$  con determinante diverso da zero. Allora  $\mathbf{A}$  è invertibile e la sua inversa è data da:

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \text{cofatt}(\mathbf{A})^T.$$

*Proof.* Utilizzando il lemma Lemma 6.10, abbiamo:

$$\text{cofatt}(\mathbf{A})^T \mathbf{A} = |\mathbf{A}| \mathbf{I}.$$

Dividendo entrambi i membri per  $|\mathbf{A}|$ , otteniamo:

$$\frac{1}{|\mathbf{A}|} \text{cofatt}(\mathbf{A})^T \mathbf{A} = \mathbf{I}.$$

Questo dimostra che:

$$\frac{1}{|\mathbf{A}|} \text{cofatt}(\mathbf{A})^T$$

è una matrice che, moltiplicata per  $\mathbf{A}$ , dà l'identità  $\mathbf{I}$ , quindi è l'inversa di  $\mathbf{A}$ .

Poiché l'inversa di una matrice è unica, abbiamo che:

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \text{cofatt}(\mathbf{A})^T.$$

■

#### Osservazione

Questa formula fornisce una rappresentazione esplicita della matrice inversa. È importante notare che l'inversa di una matrice può essere calcolata solo se il determinante della matrice è diverso da zero, il che implica che la matrice  $\mathbf{A}$  deve avere rango massimo.

Questa osservazione è simile a quella fatta per la regola di Cramer. Sebbene la formula dell'inversa sia teoricamente interessante, in pratica risulta poco utile per grandi matrici. Il motivo è che il calcolo di  $n^2$  cofattori, e quindi di minori di ordine  $n-1$ , è computazionalmente oneroso.

Come anticipato per la regola di Cramer, esistono tecniche numeriche più efficienti per calcolare l'inversa di una matrice, che richiedono un costo computazionale significativamente inferiore.

## 6.14 Calcolo dei cofattori

**Theorem 6.10 — Matrice cofattore.** Sia  $\mathbf{A}$  una matrice quadrata di ordine  $n$  e  $\alpha_{i,j}$  il cofattore associato all'elemento  $A_{i,j}$  di  $\mathbf{A}$ . Allora, il cofattore  $\alpha_{i,j}$  è dato da

$$\alpha_{i,j} = (-1)^{i+j} |\mathbf{A}_{-i,-j}| = (-1)^{i-j} |\mathbf{A}_{-i,-j}|,$$

dove  $\mathbf{A}_{-i,-j}$  è la sottomatrice di  $\mathbf{A}$  ottenuta eliminando la riga  $i$  e la colonna  $j$ .

*Proof.* Consideriamo il determinante del seguente blocco di matrice:

$$\alpha_{i,j} = \begin{vmatrix} A_{1,1} & \cdots & A_{1,j-1} & 0 & A_{1,j+1} & \cdots & A_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ A_{i-1,1} & \cdots & A_{i-1,j-1} & 0 & A_{i-1,j+1} & \cdots & A_{i-1,n} \\ A_{i,1} & \cdots & A_{i,j-1} & 1 & A_{i,j+1} & \cdots & A_{i,n} \\ A_{i+1,1} & \cdots & A_{i+1,j-1} & 0 & A_{i+1,j+1} & \cdots & A_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,j-1} & 0 & A_{n,j+1} & \cdots & A_{n,n} \end{vmatrix}$$

Possiamo semplificare questa matrice tramite opportune combinazioni lineari delle colonne:

$$\alpha_{i,j} = \begin{vmatrix} A_{1,1} & \cdots & A_{1,j-1} & 0 & A_{1,j+1} & \cdots & A_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ A_{i-1,1} & \cdots & A_{i-1,j-1} & 0 & A_{i-1,j+1} & \cdots & A_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ A_{i+1,1} & \cdots & A_{i+1,j-1} & 0 & A_{i+1,j+1} & \cdots & A_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,j-1} & 0 & A_{n,j+1} & \cdots & A_{n,n} \end{vmatrix}$$

Definiamo i vettori

$$\mathbf{v}_1 = \begin{pmatrix} A_{1,1} \\ \vdots \\ A_{i-1,1} \\ A_{i+1,1} \\ \vdots \\ A_{n,1} \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} A_{1,2} \\ \vdots \\ A_{i-1,2} \\ A_{i+1,2} \\ \vdots \\ A_{n,2} \end{pmatrix}, \dots, \quad \mathbf{v}_{j-1} = \begin{pmatrix} A_{1,j-1} \\ \vdots \\ A_{i-1,j-1} \\ A_{i+1,j-1} \\ \vdots \\ A_{n,j-1} \end{pmatrix},$$

$$\mathbf{v}_{j+1} = \begin{pmatrix} A_{1,j+1} \\ \vdots \\ A_{i-1,j+1} \\ A_{i+1,j+1} \\ \vdots \\ A_{n,j+1} \end{pmatrix}, \quad \mathbf{v}_{j+2} = \begin{pmatrix} A_{1,j+2} \\ \vdots \\ A_{i-1,j+2} \\ A_{i+1,j+2} \\ \vdots \\ A_{n,j+2} \end{pmatrix}, \dots, \quad \mathbf{v}_n = \begin{pmatrix} A_{1,n} \\ \vdots \\ A_{i-1,n} \\ A_{i+1,n} \\ \vdots \\ A_{n,n} \end{pmatrix}.$$

Si verifica facilmente che  $\alpha_{i,j}$  può essere espressa come funzione dei vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n$ . Quindi possiamo scrivere:

$$\alpha_{i,j} = f_{ij}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n),$$

dove la funzione  $f_{ij}$  soddisfa i primi due assiomi sui determinanti. Definiamo il simbolo  $\mathbf{I}_n \in \mathbb{K}^{n \times n}$  come la matrice identità di dimensioni  $n \times n$ . Per il corollario del teorema Theorem 6.1 abbiamo:

$$\begin{aligned} \alpha_{i,j} &= f_{ij}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n), \\ &= |\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n| f_{ij}(\mathbf{I}_{n-1}), \end{aligned}$$

e osserviamo che

$$\begin{aligned}
& |\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n| \\
&= \begin{vmatrix} A_{1,1} & \cdots & A_{1,j-1} & A_{1,j+1} & \cdots & A_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{i-1,1} & \cdots & A_{i-1,j-1} & A_{i-1,j+1} & \cdots & A_{i-1,n} \\ A_{i+1,1} & \cdots & A_{i+1,j-1} & A_{i+1,j+1} & \cdots & A_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,j-1} & A_{n,j+1} & \cdots & A_{n,n} \end{vmatrix}, \\
&= |\mathbf{A}_{-i,-j}|.
\end{aligned}$$

Supponendo  $i < j$ , otteniamo:

$$f_{ij}(\mathbf{I}_{n-1}) = |\mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_{j-1}, \mathbf{e}_i, \mathbf{e}_j, \dots, \mathbf{e}_n|,$$

dove  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  è la base canonica di  $\mathbb{K}^{n \times n}$ . Dopo  $|i - j|$  scambi, otteniamo:

$$\begin{aligned}
& \mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_{j-1}, \mathbf{e}_i, \mathbf{e}_j, \dots, \mathbf{e}_n, \\
& \Rightarrow \\
& \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n.
\end{aligned}$$

Pertanto:

$$\begin{aligned}
f_{ij}(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{n-1}) &= |\mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_{j-1}, \mathbf{e}_i, \mathbf{e}_j, \dots, \mathbf{e}_n|, \\
&= (-1)^{|i-j|} |\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n|, \\
&= (-1)^{|i-j|}.
\end{aligned}$$

■

**Corollary 6.3 — Corollario.** Il determinante ha quindi il seguente sviluppo (detto di Laplace<sup>a</sup>)

$$|\mathbf{A}| = \sum_{i=1}^n (-1)^{i+j} A_{i,j} |\mathbf{A}_{-i,-j}|, \quad j = 1, 2, \dots, n.$$

<sup>a</sup>Pierre-Simon Laplace (1749-1827)

## 6.15 Determinante della Trasposta

Ci chiediamo ora che valore assume il determinante della matrice trasposta. Abbiamo già visto lo sviluppo del determinante lungo le colonne; cerchiamo ora una formula analoga per lo sviluppo per righe.

**Theorem 6.11 — Sviluppo per Riga.** Il determinante può essere sviluppato lungo la riga  $i$  come segue:

$$|\mathbf{A}| = \sum_{j=1}^n (-1)^{i+j} A_{i,j} |\mathbf{A}_{-i,-j}|, \quad i = 1, 2, \dots, n. \quad (6.4)$$

*Proof.* Consideriamo il determinante di  $\mathbf{A}$ :

$$\begin{aligned}
|\mathbf{A}| &= |\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}|, \\
&= |A_{1,1}\mathbf{e}_1 + \mathbf{b}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}|,
\end{aligned}$$

dove

$$\mathbf{b} = \begin{pmatrix} 0 \\ A_{2,1} \\ \vdots \\ A_{n,1} \end{pmatrix},$$

e applicando la multilinearità del determinante otteniamo:

$$\begin{aligned} |\mathbf{A}| &= |A_{1,1}\mathbf{e}_1, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}| + |\mathbf{b}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}|, \\ &= A_{1,1}|\mathbf{e}_1, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}| + |\mathbf{b}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}|, \\ &= A_{1,1}\alpha_{1,1} + |\mathbf{b}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}|, \end{aligned}$$

Proseguendo in modo analogo, otteniamo:

$$\begin{aligned} |\mathbf{A}| &= A_{1,1}\alpha_{1,1} + |\mathbf{b}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}|, \\ &= A_{1,1}\alpha_{1,1} + A_{1,2}\alpha_{1,2} + |\mathbf{c}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}|, \\ &\vdots \\ &= A_{1,1}\alpha_{1,1} + \dots + A_{1,n}\alpha_{1,n} + |\mathbf{0}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}|, \\ &= A_{1,1}\alpha_{1,1} + \dots + A_{1,n}\alpha_{1,n}, \end{aligned}$$

e utilizzando il teorema Theorem 6.10, otteniamo l'espressione Equation 6.4. ■

Siamo in grado ora di dimostrare il seguente teorema

**Theorem 6.12 — Determinante della Trasposta.** Sia  $\mathbf{A}$  una matrice quadrata di ordine  $n$ . Allora:

$$|\mathbf{A}| = |\mathbf{A}^T|.$$

*Proof.* La dimostrazione si può effettuare per induzione sull'ordine della matrice.

- **Caso base:** Per  $n = 1$ , non c'è nulla da dimostrare poiché il determinante di una matrice  $1 \times 1$  è triviale.
- **Passo induttivo:** Supponiamo che il teorema sia vero per una matrice di ordine  $n - 1$ . Dobbiamo dimostrare che vale anche per una matrice di ordine  $n$ . Consideriamo il determinante di  $\mathbf{A}$  sviluppato lungo la prima colonna e il determinante di  $\mathbf{A}^T$  sviluppato lungo la prima riga. In entrambi i casi, otteniamo una somma di termini in cui ogni termine è il prodotto di un elemento della colonna di  $\mathbf{A}$  o della riga di  $\mathbf{A}^T$  (che coincidono) e i corrispondenti cofattori. Questi cofattori sono determinanti di sottomatrici di ordine  $n - 1$ . Le sottomatrici di  $\mathbf{A}$  e  $\mathbf{A}^T$  sono trasposte l'una dell'altra, quindi, per l'ipotesi induttiva, i determinanti delle sottomatrici sono uguali. Pertanto, il determinante di  $\mathbf{A}$  e il determinante di  $\mathbf{A}^T$  sono uguali, e il teorema è dimostrato. ■

## 6.16 Determinante delle Matrici Diagonali a Blocchi

**Theorem 6.13** Sia  $\mathbf{A}$  una matrice quadrata partizionata a blocchi come segue:

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{pmatrix},$$

dove  $\mathbf{B} \in \mathbb{K}^{p \times p}$  e  $\mathbf{C} \in \mathbb{K}^{q \times q}$ . Allora:

$$|A| = |B||C|.$$

*Proof.* Fissata la matrice  $C$ , il determinante di  $A$  è una funzione delle colonne della matrice  $B$ . Pertanto, possiamo scrivere:

$$|A| = \mathcal{D}_C(B_{\bullet,1}, B_{\bullet,2}, \dots, B_{\bullet,p}).$$

È immediato verificare che  $\mathcal{D}_C(B_{\bullet,1}, B_{\bullet,2}, \dots, B_{\bullet,p})$  soddisfa i primi due assiomi dei determinanti. Pertanto, per il corollario del teorema Theorem 6.1, otteniamo:

$$\begin{aligned} \mathcal{D}_C(B_{\bullet,1}, B_{\bullet,2}, \dots, B_{\bullet,p}) &= |B| \mathcal{D}_C(I_{\bullet,1}, I_{\bullet,2}, \dots, I_{\bullet,p}), \\ &= |B| \begin{vmatrix} I & 0 \\ 0 & C \end{vmatrix}. \end{aligned}$$

In maniera analoga, otteniamo:

$$\begin{aligned} \begin{vmatrix} I & 0 \\ 0 & C \end{vmatrix} &= \mathcal{D}_I(C_{\bullet,1}, C_{\bullet,2}, \dots, C_{\bullet,q}), \\ &= |C| |I_{\bullet,1}, I_{\bullet,2}, \dots, I_{\bullet,q}|, \\ &= |C| \begin{vmatrix} I & 0 \\ 0 & I \end{vmatrix}, \\ &= |C|. \end{aligned}$$

Da cui segue subito:

$$|A| = |B| \begin{vmatrix} I & 0 \\ 0 & C \end{vmatrix} = |B||C|.$$

■

## 7. Sistemi lineari: regola di Cramer e Rouchè-Capelli

### 7.1 Regola di Cramer

La regola di Cramer<sup>1</sup>, è un metodo che consente di risolvere *teoricamente* un sistema lineare con  $n$  equazioni e  $n$  incognite, a condizione che la matrice dei coefficienti sia non singolare.

Consideriamo un sistema lineare che può essere scritto in forma matriciale compatta come:

$$\mathbf{Ax} = \mathbf{b}$$

Il vettore  $\mathbf{b}$  può essere espresso come una combinazione lineare delle colonne di  $\mathbf{A}$ :

$$\mathbf{b} = \mathbf{Ax} = \mathbf{A}_{\bullet,1}x_1 + \mathbf{A}_{\bullet,2}x_2 + \cdots + \mathbf{A}_{\bullet,n}x_n = \sum_{j=1}^n \mathbf{A}_{\bullet,j}x_j$$

dove  $\mathbf{x}$  è il vettore delle incognite e  $\mathbf{A}_{\bullet,j}$  rappresenta la  $j$ -esima colonna di  $\mathbf{A}$ .

Per calcolare l'incognita  $x_i$ , consideriamo il determinante della matrice ottenuta sostituendo la  $i$ -esima colonna di  $\mathbf{A}$  con  $\mathbf{b}$ :

$$\begin{aligned} & \mathcal{D}(\mathbf{A}_{\bullet,1}, \dots, \mathbf{A}_{\bullet,i-1}, \mathbf{b}, \mathbf{A}_{\bullet,i+1}, \dots, \mathbf{A}_{\bullet,n}) \\ &= \mathcal{D}\left(\mathbf{A}_{\bullet,1}, \dots, \mathbf{A}_{\bullet,i-1}, \sum_{j=1}^n \mathbf{A}_{\bullet,j}x_j, \mathbf{A}_{\bullet,i+1}, \dots, \mathbf{A}_{\bullet,n}\right) \\ &= \sum_{j=1}^n x_j \mathcal{D}(\mathbf{A}_{\bullet,1}, \dots, \mathbf{A}_{\bullet,i-1}, \mathbf{A}_{\bullet,j}, \mathbf{A}_{\bullet,i+1}, \dots, \mathbf{A}_{\bullet,n}) \\ &= x_i \mathcal{D}(\mathbf{A}_{\bullet,1}, \dots, \mathbf{A}_{\bullet,i-1}, \mathbf{A}_{\bullet,i}, \mathbf{A}_{\bullet,i+1}, \dots, \mathbf{A}_{\bullet,n}) \end{aligned}$$

Da questa espressione, possiamo ricavare  $x_i$  come:

$$x_i = \frac{\mathcal{D}(\mathbf{A}_{\bullet,1}, \dots, \mathbf{A}_{\bullet,i-1}, \mathbf{b}, \mathbf{A}_{\bullet,i+1}, \dots, \mathbf{A}_{\bullet,n})}{\mathcal{D}(\mathbf{A}_{\bullet,1}, \dots, \mathbf{A}_{\bullet,n})}$$

La regola di Cramer è applicabile solo se la matrice  $\mathbf{A}$  è non singolare, cioè se:

$$\mathcal{D}(\mathbf{A}) = \mathcal{D}(\mathbf{A}_{\bullet,1}, \dots, \mathbf{A}_{\bullet,n}) \neq 0$$

---

<sup>1</sup>James Joseph Sylvester (1814-1897)

## Osservazione

La regola di Cramer richiede il calcolo di  $n + 1$  determinanti, ognuno dei quali implica  $n!$  somme e prodotti, per un totale di  $(n + 1)!$  operazioni aritmetiche. In confronto, esistono metodi di risoluzione dei sistemi lineari che hanno un costo computazionale di  $\mathcal{O}(n^3)$  operazioni<sup>2</sup>.

Per valori di  $n$  più grandi, la crescita fattoriale del costo rende la regola di Cramer impraticabile, limitandone l'uso a scopi essenzialmente *teorici*.

## 7.2 Il teorema di Rouchè<sup>3</sup>-Capelli<sup>4</sup>

Utilizzando i teoremi introdotti finora, possiamo definire dei criteri per determinare se un sistema di equazioni lineari ammette soluzioni. Consideriamo il sistema:

$$\mathbf{A}\mathbf{x} = \mathbf{b},$$

dove  $\mathbf{A} \in \mathbb{K}^{m \times n}$  è una matrice,  $\mathbf{x} \in \mathbb{K}^n$  è un vettore di variabili incognite e  $\mathbf{b} \in \mathbb{K}^m$  è un vettore noto.

Supponiamo che  $\mathbf{x}$  sia una soluzione del sistema (non preoccupiamoci per ora della sua unicità). In tal caso, possiamo esprimere il vettore  $\mathbf{b}$  come una combinazione lineare delle colonne di  $\mathbf{A}$ :

$$\mathbf{b} = x_1 \mathbf{A}_{\bullet,1} + x_2 \mathbf{A}_{\bullet,2} + \cdots + x_n \mathbf{A}_{\bullet,n},$$

dove  $x_1, x_2, \dots, x_n$  sono i componenti del vettore  $\mathbf{x}$ , e  $\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}$  sono le colonne della matrice  $\mathbf{A}$ .

Questo significa che  $\mathbf{b}$  è una combinazione lineare delle colonne di  $\mathbf{A}$ . Pertanto, il numero di vettori linearmente indipendenti tra le colonne  $\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}$  è lo stesso di quello delle colonne  $\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}, \mathbf{b}$ .

In simboli, possiamo scrivere:

$$\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}, \mathbf{b}),$$

dove  $\text{rank}(\mathbf{A}, \mathbf{b})$  indica il rango della matrice estesa

$$(\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}, \mathbf{b}),$$

ossia la matrice formata dalle colonne di  $\mathbf{A}$  con in aggiunta il vettore colonna  $\mathbf{b}$ .

Viceversa, se il sistema non ammette soluzioni, il vettore  $\mathbf{b}$  non può essere scritto come una combinazione lineare delle colonne di  $\mathbf{A}$ . In tal caso,  $\mathbf{b}$  è linearmente indipendente dalle colonne di  $\mathbf{A}$ , e quindi il numero di vettori colonna linearmente indipendenti di  $\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}$  è inferiore a quello delle colonne  $\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}, \mathbf{b}$ .

In simboli, ciò si esprime come:

$$\text{rank}(\mathbf{A}) < \text{rank}(\mathbf{A}, \mathbf{b}).$$

Questi risultati possono essere riassunti nel seguente teorema:

<sup>2</sup>La notazione  $\mathcal{O}(g(n))$  rappresenta l'insieme delle funzioni  $h(n)$  tali che  $h(n) \leq C \cdot g(n)$  per  $n \geq n_0$ , dove  $C$  è una costante non negativa indipendente da  $n$ . L'espressione  $f = \mathcal{O}(g)$  indica che  $f$  è dell'ordine di  $g$ , e rappresenta un abuso notazionale per  $f \in \mathcal{O}(g)$ .

<sup>4</sup>Alfredo Capelli (1855-1910)

**Theorem 7.1 — Rouché-Capelli versione compatta.** Sia dato il sistema

$$\mathbf{Ax} = \mathbf{b},$$

dove  $\mathbf{A} \in \mathbb{K}^{m \times n}$ ,  $\mathbf{x} \in \mathbb{K}^n$  e  $\mathbf{b} \in \mathbb{K}^m$ . Il sistema ammette soluzioni se e solo se:

$$\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}, \mathbf{b}).$$

I risultati sui determinanti permettono di dare una forma operativa del teorema Theorem 7.1, infatti

### 7.2.1 Teorema di Rouché-Capelli (versione estesa)

Sia dato il sistema

$$\mathbf{Ax} = \mathbf{b},$$

dove  $\mathbf{A} \in \mathbb{K}^{m \times n}$ ,  $\mathbf{x} \in \mathbb{K}^n$  e  $\mathbf{b} \in \mathbb{K}^m$ . Allora il sistema ha soluzione se e solo se il massimo ordine dei minori non nulli della matrice  $\mathbf{A}$  è lo stesso della matrice  $(\mathbf{A}, \mathbf{b})$ .

*Proof.* Basta applicare il teorema seguente dimostrato negli appunti sui determinanti. ■

**Theorem 7.2 — Minori e Indipendenza Lineare.** Se i  $k$  vettori  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in \mathbb{K}^n$  sono linearmente indipendenti e  $k \leq n$ , allora esiste almeno un minore di ordine  $k$  non nullo della matrice  $\mathbf{A} \in \mathbb{K}^{k \times k}$  con  $n \geq k$ , definita come

$$\mathbf{A} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k).$$

Sia  $\mathbf{x} \in \mathbb{K}^n$  una soluzione del sistema lineare  $\mathbf{Ax} = \mathbf{b}$ , con  $\mathbf{A} \in \mathbb{K}^{m \times n}$  e  $\mathbf{b} \in \mathbb{K}^m$ , con  $\mathbf{b} \neq \mathbf{0}$ , e  $\tilde{\mathbf{x}} \in \mathbb{K}^n$  una soluzione del sistema lineare omogeneo  $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{0}$ .

È interessante notare che  $\mathbf{x} + \tilde{\mathbf{x}}$  è ancora soluzione del sistema lineare non omogeneo:

$$\mathbf{A}(\mathbf{x} + \tilde{\mathbf{x}}) = \mathbf{Ax} + \mathbf{A}\tilde{\mathbf{x}} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

Quindi, data una soluzione *particolare*  $\mathbf{x}$  del sistema lineare non omogeneo —  $\mathbf{b} \neq \mathbf{0}$  — per ogni soluzione  $\tilde{\mathbf{x}}$  del sistema lineare omogeneo si ottiene una diversa soluzione del problema non omogeneo della forma  $\mathbf{x} + \tilde{\mathbf{x}}$ .

È intuitivo che la soluzione di un sistema lineare non omogeneo — una volta che ne sia ammessa l'esistenza — è *unica* solo nei casi in cui possiamo garantire che sia sempre  $\tilde{\mathbf{x}} = \mathbf{0}$ , cioè che l'unica soluzione possibile del sistema lineare omogeneo  $\mathbf{A}\tilde{\mathbf{x}} = \mathbf{0}$  sia il vettore nullo.

Formalizziamo questa considerazione nel seguente teorema per matrici quadrate.

### 7.2.2 Esistenza e unicità per sistemi lineari

**Theorem 7.3** Sia  $\mathbf{A} \in \mathbb{K}^{n \times n}$ , tale che  $\text{rank}(\mathbf{A}) = n$ . Allora la soluzione  $\mathbf{x} \in \mathbb{K}^n$  del problema  $\mathbf{Ax} = \mathbf{b}$ , con  $\mathbf{b} \in \mathbb{K}^n$  vettore assegnato, *esiste ed è unica*.

*Proof.* Dimostriamo separatamente prima l'esistenza e poi l'unicità.

- (i) **Esistenza.** Poiché in  $\mathbb{K}^n$  possiamo avere al massimo  $n$  vettori linearmente indipendenti,  $\text{rank}(\mathbf{A}) = n$  implica che  $\text{rank}(\mathbf{A}, \mathbf{b}) = n$ , e l'esistenza segue dal teorema Theorem 7.3
- (ii) **Unicità.** Essendo la matrice  $\mathbf{A}$  quadrata di ordine  $n$  uguale al suo rango, si deduce immediatamente che le sue colonne  $\mathbf{A}_{\bullet,1}, \mathbf{A}_{\bullet,2}, \dots, \mathbf{A}_{\bullet,n}$ , sono vettori linearmente indipendenti. Supponiamo ora di avere due soluzioni possibili del sistema lineare non

omogeneo, che indichiamo con  $\mathbf{x}_1$  e  $\mathbf{x}_2$ , per cui possiamo scrivere  $\mathbf{A}\mathbf{x}_1 = \mathbf{A}\mathbf{x}_2 = \mathbf{b}$ . Per differenza si ottiene  $\mathbf{A}(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$ , da cui segue che  $\mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0}$ , cioè  $\mathbf{x}_1 = \mathbf{x}_2$ . ■

Che cosa succede se la matrice è rettangolare?

#### Osservazione

Se la matrice  $\mathbf{A}$  è rettangolare con più colonne che righe ( $m < n$ ), l'unicità della soluzione non è garantita. Infatti, il rango massimo della matrice è  $m$ , il che implica che almeno  $n - m$  colonne sono linearmente dipendenti dalle altre.

Di conseguenza, esistono almeno  $n - m$  vettori linearmente indipendenti in  $\mathbb{K}^n$  che rappresentano combinazioni lineari delle colonne di  $\mathbf{A}$ . Questi vettori costituiscono lo spazio delle soluzioni del problema omogeneo associato, rendendo la soluzione non unica.

#### Osservazione

Nel caso di matrici rettangolari con più righe che colonne ( $m > n$ ), il massimo rango possibile è ovviamente  $n$ . Tuttavia, il sistema lineare ammette soluzioni solo se la condizione

$$\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}, \mathbf{b})$$

è soddisfatta. In caso contrario, il sistema è *sovradeterminato*.

Questo significa che possiamo ridurre il sistema eliminando un certo numero di equazioni fino a ottenere una matrice quadrata o rettangolare con più colonne che righe. Se  $\text{rank}(\mathbf{A}) = n$ , possiamo eliminare opportunamente  $m - n$  equazioni per arrivare a una situazione che soddisfa il teorema @thm-sl:7b.

Se  $\text{rank}(\mathbf{A}) < n$ , dovremo eliminare un numero maggiore di equazioni e ci troveremo nella situazione descritta nell'osservazione precedente.

## 8. Autovalori e Autovettori

Sia  $\mathbf{A} \in \mathbb{K}^{n \times n}$  una matrice quadrata con elementi nel campo  $\mathbb{K}$ . Consideriamo la matrice  $\mathbf{M}(\lambda)$ , che dipende dallo scalare  $\lambda \in \mathbb{K}$ , definita dalla relazione

$$\mathbf{M}(\lambda) = \mathbf{A} - \lambda \mathbf{I},$$

dove  $\mathbf{I}$  rappresenta la matrice identità di dimensione  $n$ .

Ci poniamo la seguente domanda: **per quali valori dello scalare  $\lambda$  la matrice  $\mathbf{M}(\lambda)$  e la sua trasposta  $\mathbf{M}^T(\lambda)$  sono singolari?**

Una matrice è singolare se il suo determinante è nullo, cioè se non è di rango massimo. Quindi, la condizione di singolarità per  $\mathbf{M}(\lambda)$  e  $\mathbf{M}^T(\lambda)$  si traduce nella richiesta che il loro determinante sia nullo:

$$\det(\mathbf{M}(\lambda)) = 0 \quad \text{e} \quad \det(\mathbf{M}^T(\lambda)) = 0.$$

Poiché vale la proprietà

$$\det(\mathbf{M}(\lambda)) = \det(\mathbf{M}^T(\lambda)),$$

ne segue che i valori di  $\lambda$  che rendono singolare  $\mathbf{M}(\lambda)$  coincidono con quelli che rendono singolare  $\mathbf{M}^T(\lambda)$ . In altre parole, gli **scalari  $\lambda$  per cui  $\det(\mathbf{M}(\lambda)) = 0$**  sono gli stessi per entrambe le matrici.

**Definition 8.1 — Definizione di Autovalore.** I valori di  $\lambda$  che rendono singolare la matrice  $\mathbf{M}(\lambda)$  – e quindi anche  $\mathbf{M}^T(\lambda)$  – sono detti **autovalori** di  $\mathbf{A}$ .

### 8.1 Polinomio Caratteristico

Il determinante della matrice  $\mathbf{M}(\lambda)$  è un polinomio in  $\lambda$  di grado  $n$ . Questo polinomio si chiama **polinomio caratteristico** di  $\mathbf{A}$ .

**Lemma 8.1 — struttura del polinomio Caratteristico.** Il determinante della matrice  $\mathbf{M}(\lambda) = \mathbf{A} - \lambda \mathbf{I}$  è un **polinomio in  $\lambda$  di grado  $n$** .

La struttura ricorsiva del determinante può essere descritta come segue. Per una matrice  $\mathbf{A} - \lambda \mathbf{I}$  di ordine  $n$ , possiamo sviluppare il determinante lungo una riga o una colonna. Sviluppando lungo la prima riga, si ottiene:

$$\det(\mathbf{M}(\lambda)) = \sum_{j=1}^n (-1)^{1+j} (\mathbf{A}_{1,j} - \lambda \delta_{1,j}) \det(\mathbf{A}_{-1,-j} - \lambda \mathbf{I}_{-1,-j}),$$

dove:

- $\mathbf{A}_{-1,-j}$  è la matrice ottenuta eliminando la prima riga e la  $j$ -esima colonna da  $\mathbf{A}$ ,
- $\mathbf{I}_{-1,-j}$  è la matrice identità corrispondente di ordine ridotto.

Il determinante di ciascuna sottomatrice  $\mathbf{A}_{-1,-j} - \lambda \mathbf{I}_{-1,-j}$  può essere sviluppato ulteriormente allo stesso modo. Questo processo ricorsivo termina quando il determinante si riduce a una matrice  $1 \times 1$ , il cui determinante è semplicemente l'elemento  $\mathbf{A}_{i,i} - \lambda$ .

Dato che ogni passo dello sviluppo del determinante introduce termini lineari in  $\lambda$ , il risultato finale è un **polinomio in  $\lambda$  di grado al massimo  $n$** .

Poiché una matrice è singolare quando il suo determinante è nullo, la condizione di singolarità per  $\mathbf{M}(\lambda)$  ci porta alla seguente definizione.

**Definition 8.2 — Definizione: Polinomio Caratteristico.** Il polinomio

$$p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}),$$

è detto **polinomio caratteristico** della matrice  $\mathbf{A}$ .

Osservazione

Gli zeri del polinomio caratteristico di  $\mathbf{A}$  sono gli autovalori della matrice  $\mathbf{A}$ .

Per il teorema fondamentale dell'algebra il polinomio caratteristico si fattorizza come prodotto di binomi elevati ad opportuni esponenti  $m_{a,i}$

$$p_{\mathbf{A}}(\lambda) = (-1)^n (\lambda - \lambda_1)^{m_{a,1}} (\lambda - \lambda_2)^{m_{a,2}} \dots (\lambda - \lambda_s)^{m_{a,s}}, \quad (8.1)$$

dove  $m_{a,1} + m_{a,2} + \dots + m_{a,s} = n$  e  $\lambda_i \neq \lambda_j$  se  $i \neq j$ .

**Definition 8.3 — Molteplicità algebrica.** Se il polinomio caratteristico di  $\mathbf{A}$  ammette  $s \leq n$  zeri distinti che sono gli  $s$  autovalori, e si fattorizza come sopra, ad ogni autovalore  $\lambda_i$  compete una *molteplicità algebrica*  $m_{a,i}$ .

Se la matrice  $\mathbf{M}(\lambda)$  è singolare per un dato valore dello scalare  $\lambda$ , allora deve esistere almeno un vettore  $\mathbf{x} \neq \mathbf{0}$  tale che

$$\mathbf{M}(\lambda)\mathbf{x} = (\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}.$$

Analogamente se la matrice  $\mathbf{M}(\lambda)^T$  è singolare per un dato valore dello scalare  $\lambda$ , allora deve esistere almeno un vettore  $\mathbf{y} \neq \mathbf{0}$  tale che

$$\mathbf{M}(\lambda)^T \mathbf{y} = (\mathbf{A}^T - \lambda \mathbf{I})\mathbf{y} = \mathbf{0},$$

cioè, trasponendo,

$$\mathbf{y}^T \mathbf{M}(\lambda) = \mathbf{y}(\mathbf{A} - \lambda \mathbf{I}) = \mathbf{0}^T.$$

In generale si ha che  $\mathbf{x} \neq \mathbf{y}$ , anche se si riferiscono allo stesso scalare  $\lambda$  che rende singolari  $\mathbf{M}(\lambda)$  e  $\mathbf{M}^T(\lambda)$ .

## 8.2 Autovettori

**Definition 8.4 — Autovettore destro.** Un vettore colonna  $\mathbf{u} \neq \mathbf{0}$  è un autovettore *destro* della matrice  $\mathbf{A}$  rispetto ad uno scalare  $\lambda$  se vale la relazione

$$\mathbf{A}\mathbf{u} = \lambda\mathbf{u}.$$

**Definition 8.5 — Autovettore sinistro.** Un vettore riga  $\mathbf{w}^T \neq \mathbf{0}$  è un autovettore *sinistro* della matrice  $\mathbf{A}$  rispetto ad uno scalare  $\lambda$  se vale la relazione

$$\mathbf{w}^T \mathbf{A} = \lambda \mathbf{w}^T.$$

cioè  $\mathbf{w}$  è autovettore di  $\mathbf{A}^T$ .

### Osservazione

Ricordando la definizione di *kernel* di una matrice, possiamo affermare che gli **autovettori destri** di  $\mathbf{A}$  associati a un autovalore  $\lambda \in \mathbb{K}$  sono esattamente gli elementi di  $\ker(\mathbf{A} - \lambda\mathbf{I})$ . Allo stesso modo, gli **autovettori sinistri** di  $\mathbf{A}$  rispetto allo stesso autovalore  $\lambda$  sono tutti e soli gli elementi di  $\ker(\mathbf{A}^T - \lambda\mathbf{I})$ .

## 8.3 Autovalori multipli

**Lemma 8.2 — Autovalori Multipli.** Se  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$  sono  $k$  autovettori associati allo stesso autovalore  $\lambda_i$ , allora ogni loro combinazione lineare è anch'essa un autovettore corrispondente allo stesso autovalore.

*Proof.* Siano  $\beta_1, \beta_2, \dots, \beta_k$  dei scalari arbitrari. Allora, per ogni combinazione lineare degli autovettori  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ , si ha:

$$\mathbf{A} \left( \sum_{j=1}^k \beta_j \mathbf{x}_j \right) = \sum_{j=1}^k \beta_j \mathbf{A} \mathbf{x}_j = \sum_{j=1}^k \beta_j \lambda_i \mathbf{x}_j = \lambda_i \left( \sum_{j=1}^k \beta_j \mathbf{x}_j \right).$$

Quindi, la combinazione lineare  $\sum_{j=1}^k \beta_j \mathbf{x}_j$  è un autovettore rispetto allo stesso autovalore  $\lambda_i$ . ■

### 8.3.1 Molteplicità Geometrica

**Definition 8.6 — Definizione: Molteplicità Geometrica.** La *molteplicità geometrica* di un autovalore  $\lambda_i$ , indicata con  $m_{g,i}$ , è il massimo numero di autovettori linearmente indipendenti che possono essere scelti nello spazio  $\ker(\mathbf{A} - \lambda_i \mathbf{I})$  associato all'autovalore  $\lambda_i$ .

### 8.3.2 Molteplicità Algebrica

**Definition 8.7 — Definizione: Molteplicità Algebrica.** La *molteplicità algebrica* di un autovalore  $\lambda_i$  di una matrice  $\mathbf{A}$  è definita come la molteplicità dell'autovalore  $\lambda_i$  come radice del polinomio caratteristico di  $\mathbf{A}$ . In altre parole, è il numero di volte che  $\lambda_i$  compare come radice del polinomio caratteristico

$$p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}),$$

dove  $\mathbf{I}$  è la matrice identità della stessa dimensione di  $\mathbf{A}$ . La molteplicità algebrica di  $\lambda_i$  è quindi il grado di  $\lambda_i$  come radice del polinomio caratteristico.

**Theorem 8.1 — Molteplicità confronto.** La molteplicità algebrica  $m_{a,i}$  di un autovalore  $\lambda_i$  è sempre maggiore o uguale alla corrispondente molteplicità geometrica  $m_{g,i}$ , cioè:

$$m_{a,i} \geq m_{g,i}.$$

*Proof.* Siano  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{m_{g,i}}$  gli autovettori corrispondenti a  $\lambda_i$ , che possiamo sempre considerare *ortonormali* (se necessario, possiamo ortonormalizzarli usando il processo di Gram-Schmidt).

Se  $m_{g,i} < n$ , completiamo con altri  $n - m_{g,i}$  vettori per ottenere una base ortonormale. Definiamo  $\mathbf{T}$  come la matrice le cui colonne sono questi vettori:

$$\mathbf{T} = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_{m_{g,i}} \quad \mathbf{u}_{m_{g,i}+1} \quad \cdots \quad \mathbf{u}_n),$$

dove  $\mathbf{T}^T \mathbf{T} = \mathbf{I}$  implica che  $\mathbf{T}^T = \mathbf{T}^{-1}$ .

Consideriamo la matrice  $\mathbf{A}$  trasformata dalla base ortonormale:

$$\mathbf{AT} = (\lambda_i \mathbf{u}_1 \quad \lambda_i \mathbf{u}_2 \quad \cdots \quad \lambda_i \mathbf{u}_{m_{g,i}} \quad \mathbf{w} \quad \cdots \quad \mathbf{z}).$$

Pertanto, possiamo scrivere:

$$\mathbf{T}^{-1} \mathbf{AT} = \mathbf{T}^T \mathbf{AT} = \begin{pmatrix} \lambda_i & & & & & \\ & \ddots & & & & \\ & & \lambda_i & & & \\ & & & \mathbf{B} & & \\ & & & & \mathbf{C} & \\ & \mathbf{0} & & & & \end{pmatrix}.$$

Dove  $\mathbf{B}$  e  $\mathbf{C}$  rappresentano blocchi della matrice. Pertanto, il determinante del polinomio caratteristico è:

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}) &= \det(\mathbf{T}^{-1}(\mathbf{A} - \lambda \mathbf{I})\mathbf{T}) \\ &= \det(\mathbf{T}^{-1} \mathbf{AT} - \lambda \mathbf{I}) \\ &= \begin{vmatrix} \lambda_i - \lambda & & & & & \\ & \ddots & & & & \\ & & \lambda_i - \lambda & & & \\ & & & \mathbf{B} & & \\ & & & & \mathbf{C} - \lambda \mathbf{I} & \\ & \mathbf{0} & & & & \end{vmatrix} \\ &= (\lambda_i - \lambda)^{m_{g,i}} \det(\mathbf{C} - \lambda \mathbf{I}). \end{aligned}$$

Pertanto,  $\lambda_i$  è una radice del polinomio caratteristico con molteplicità almeno  $m_{g,i}$ , ma può essere superiore se  $\det(\mathbf{C} - \lambda_i \mathbf{I}) = 0$ . Concludiamo che  $m_{a,i} \geq m_{g,i}$ . ■

#### Osservazione

Se un autovalore, per esempio l' $i$ -esimo, ha molteplicità  $m_{a,i} > m_{g,i}$ , allora, poiché  $m_{a,1} + m_{a,2} + \cdots + m_{a,s} = n$ , non ci sono abbastanza autovettori linearmente indipendenti per formare una base.

#### Osservazione

Una matrice  $\mathbf{A}$  è non singolare se e solo se non ammette lo zero come autovalore.

#### Osservazione

Se  $\mathbf{A}$  è a valori reali e  $\lambda$  è un numero reale allora possiamo trovare una base di  $\ker(\mathbf{A} - \lambda \mathbf{I})$  i cui vettori hanno componenti reali.

## 8.4 Matrici reali simmetriche e hermitiane

Ricordiamo le definizioni già introdotte nel capitolo riguardante vettori e matrici. Una matrice quadrata  $\mathbf{A}$  si dice *simmetrica* se soddisfa la condizione  $\mathbf{A} = \mathbf{A}^T$ , dove  $\mathbf{A}^T$  è la trasposta di  $\mathbf{A}$ . Allo stesso modo, una matrice quadrata è *ermitiana* se  $\mathbf{A} = \mathbf{A}^H$ , dove  $\mathbf{A}^H$  denota la trasposizione coniugata (o matrice di Hermite).

### Osservazione

Per matrici reali (tutte le componenti sono reali), i concetti di simmetria e hermitianeità coincidono, mentre differiscono per matrici complesse (almeno una componente è complessa).

### Esempio

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 3 & 1 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 2 & 3 \\ -2 & 1 & 1 \\ -3 & 1 & 0 \end{pmatrix},$$

$\mathbf{A}$  è una matrice simmetrica, mentre  $\mathbf{B}$  non lo è.

### Esempio

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 3 & 1 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 2+i & 3-4i \\ 2-i & 1 & 1 \\ 3+4i & 1 & 0 \end{pmatrix},$$

$$\mathbf{C} = \begin{pmatrix} 1 & 2-i & 3+4i \\ 2-i & 1 & 1 \\ 3+4i & 1 & 0 \end{pmatrix},$$

$\mathbf{A}$  e  $\mathbf{B}$  sono matrici hermitiane,  $\mathbf{C}$  non lo è.  $\mathbf{A}$  e  $\mathbf{C}$  sono matrici simmetriche,  $\mathbf{B}$  non lo è.

**Definition 8.8 — Matrice normale.** Una matrice quadrata  $\mathbf{A}$  si dice *normale* se  $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$ .

### Esempio

$$\mathbf{A} = \begin{pmatrix} 1 & 2-i & 0 \\ 2+i & 4 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

$\mathbf{A}$  è una matrice normale (è anche hermitiana).

Le proprietà di simmetria per matrici reali o di ermiticità per matrici complesse impongono condizioni molto **forti** sugli autovalori e sugli autovettori. In particolare, il seguente teorema è valido:

**Theorem 8.2** Una matrice reale simmetrica o ermitiana ha solo autovalori reali.

*Proof.* Consideriamo una matrice reale  $\mathbf{A}$  simmetrica. Poiché ogni matrice reale simmetrica è anche ermitiana, il caso della matrice ermitiana è sufficientemente generico. Supponiamo quindi che  $\mathbf{A}$  sia ermitiana e che  $\lambda$  sia un autovalore con il corrispondente autovettore  $\mathbf{u} \neq \mathbf{0}$ . Moltiplicando a sinistra per  $\mathbf{u}^H$ , otteniamo:

$$\mathbf{A}\mathbf{u} = \lambda \mathbf{u}$$

da cui segue:

$$\mathbf{u}^H \mathbf{A}\mathbf{u} = \lambda \mathbf{u}^H \mathbf{u} = \lambda \|\mathbf{u}\|_2^2.$$

Isoliamo l'autovalore  $\lambda$ :

$$\mathbf{u}^H \mathbf{A}\mathbf{u} = \lambda \|\mathbf{u}\|_2^2 \quad \Rightarrow \quad \lambda = \frac{\mathbf{u}^H \mathbf{A}\mathbf{u}}{\|\mathbf{u}\|_2^2}.$$

Dalla definizione di norma e autovettore, il denominatore  $\|\mathbf{u}\|_2^2$  è sicuramente un numero reale *strettamente* positivo. Inoltre, il numeratore  $\mathbf{u}^H \mathbf{A}\mathbf{u}$  è reale, come dimostrato dalla seguente sequenza di uguaglianze:

$$\begin{aligned} \mathbf{u}^H \mathbf{A}\mathbf{u} &= (\mathbf{u}^H \mathbf{A}\mathbf{u})^T, \\ &= \overline{(\mathbf{u}^H \mathbf{A}\mathbf{u})^T}, \\ &= \overline{(\mathbf{u}^H \mathbf{A}\mathbf{u})^H}, \\ &= \overline{\mathbf{u}^H \mathbf{A}^H \mathbf{u}}, \\ &= \overline{\mathbf{u}^H \mathbf{A}\mathbf{u}}. \end{aligned}$$

Poiché  $\mathbf{A} = \mathbf{A}^H$ , risulta che  $\mathbf{u}^H \mathbf{A}\mathbf{u} \in \mathbb{R}$ . Essendo  $\lambda$  un rapporto tra numeri reali, anche  $\lambda$  è reale. ■

**Theorem 8.3 — Esistenza degli Autovettori Reali.** Per ogni autovalore di una matrice reale simmetrica esiste almeno un autovettore che ha solo componenti reali.

*Proof.* Sia  $\mathbf{A}$  una matrice reale simmetrica e  $\lambda$  un autovalore di  $\mathbf{A}$ . Supponiamo che  $\mathbf{u} = \mathbf{A} + i\mathbf{b}$  sia un autovettore associato a  $\lambda$ , dove  $\mathbf{A}$  e  $\mathbf{b}$  sono rispettivamente le componenti reali e complesse di  $\mathbf{u}$ . Allora abbiamo:

$$\mathbf{A}(\mathbf{A} + i\mathbf{b}) = \lambda(\mathbf{A} + i\mathbf{b}).$$

Separando le parti reale e immaginaria, otteniamo:

$$\mathbf{A}\mathbf{A} + i\mathbf{A}\mathbf{b} = \lambda\mathbf{A} + i\lambda\mathbf{b}.$$

Eguagliando le parti reali e immaginarie, otteniamo:

$$\mathbf{A}\mathbf{A} = \lambda\mathbf{A} \quad \text{e} \quad \mathbf{A}\mathbf{b} = \lambda\mathbf{b}.$$

Poiché  $\mathbf{u} \neq \mathbf{0}$ , almeno uno dei vettori  $\mathbf{A}$  o  $\mathbf{b}$  deve essere non nullo. Pertanto, almeno uno dei vettori reali  $\mathbf{A}$  o  $\mathbf{b}$  è un autovettore di  $\mathbf{A}$  rispetto all'autovalore  $\lambda$ . ■

**Theorem 8.4 — Ortogonalità degli Autovettori.** Sia  $\mathbf{A}$  una matrice reale simmetrica o ermitiana, e siano  $\lambda$  e  $\mu$  due autovalori distinti di  $\mathbf{A}$  con i corrispondenti autovettori  $\mathbf{u}$  e  $\mathbf{v}$ . Allora, gli autovettori  $\mathbf{u}$  e  $\mathbf{v}$  sono ortogonali, cioè:

$$\mathbf{u} \cdot \mathbf{v} = 0.$$

*Proof.* Consideriamo le seguenti equazioni:

$$\begin{aligned} \mathbf{A}\mathbf{u} &= \lambda\mathbf{u}, \\ \mathbf{A}\mathbf{v} &= \mu\mathbf{v}. \end{aligned}$$

Moltiplichiamo la prima equazione a sinistra per  $\mathbf{v}^H$  e la seconda per  $\mathbf{u}^H$ :

$$\begin{aligned} \mathbf{v}^H\mathbf{A}\mathbf{u} &= \lambda\mathbf{v}^H\mathbf{u}, \\ \mathbf{u}^H\mathbf{A}\mathbf{v} &= \mu\mathbf{u}^H\mathbf{v}. \end{aligned}$$

Poiché  $\mathbf{A}$  è simmetrica o ermitiana, abbiamo  $\mathbf{v}^H\mathbf{A}\mathbf{u} = \mathbf{u}^H\mathbf{A}\mathbf{v}$ . Quindi:

$$\lambda\mathbf{v}^H\mathbf{u} = \mu\mathbf{u}^H\mathbf{v}.$$

Da cui segue:

$$\lambda\mathbf{v}^H\mathbf{u} = \mu\mathbf{v}^H\mathbf{u} \quad \Rightarrow \quad (\lambda - \mu)\mathbf{v}^H\mathbf{u} = 0.$$

Poiché  $\lambda \neq \mu$ , abbiamo:

$$\mathbf{v}^H\mathbf{u} = 0,$$

cioè  $\mathbf{u}$  e  $\mathbf{v}$  sono ortogonali. ■

Infine vale la seguente osservazione.

#### Osservazione

Se  $\mathbf{A}$  è una matrice simmetrica (o ermitiana), allora ogni autovettore destro è anche un autovettore sinistro. Questo segue direttamente dalle definizioni. Per esempio, se  $\mathbf{A}$  è simmetrica, abbiamo:

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \lambda\mathbf{x} \\ \Rightarrow (\mathbf{A}\mathbf{x})^T &= \lambda\mathbf{x}^T && [\text{trasposizione}] \\ \Rightarrow \mathbf{x}^T\mathbf{A}^T &= \lambda\mathbf{x}^T && [\mathbf{A} = \mathbf{A}^T] \\ \Rightarrow \mathbf{x}^T\mathbf{A} &= \lambda\mathbf{x}^T. \end{aligned}$$

Quindi,  $\mathbf{x}$  è un autovettore sinistro con lo stesso autovalore  $\lambda$ .

Infine, una proprietà fondamentale delle matrici reali simmetriche o ermitiane è che possono essere diagonalizzate mediante opportune trasformazioni chiamate *trasformazioni di similitudine*. In questo caso particolare, tali trasformazioni coinvolgono matrici *ortogonali* o *unitarie*. Di seguito forniamo la definizione di queste matrici.

**Definition 8.9 — Matrice ortogonale.** Una matrice  $\mathbf{A}$  è *ortogonale* se la sua trasposta coincide con la sua inversa. Formalmente,

$$\mathbf{A}^T = \mathbf{A}^{-1}.$$

**Definition 8.10 — Matrice unitaria.** Una matrice  $\mathbf{A}$  è *unitaria* se la sua trasposta coniugata coincide con la sua inversa. Formalmente,

$$\mathbf{A}^H = \mathbf{A}^{-1}.$$

## Osservazione

Se la matrice  $\mathbf{A}$  è a valori reali, allora i concetti di unitarietà e ortogonalità coincidono.

## 8.5 Diagonalizzazione

**Theorem 8.5 — Teorema di Diagonalizzazione.** Sia  $\mathbf{A}$  una matrice reale simmetrica. Allora esiste una matrice  $\mathbf{U}$  ortogonale tale che:

$$\mathbf{U}^T \mathbf{A} \mathbf{U} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

dove  $\lambda_1, \lambda_2, \dots, \lambda_n$  sono gli autovalori della matrice  $\mathbf{A}$ .

*Proof.* La dimostrazione si basa su un argomento per induzione.

- **Base dell'induzione:** Il teorema è vero per le matrici  $1 \times 1$ .
- **Passo induttivo:** Sia  $\lambda_1$  un autovalore di  $\mathbf{A}$  (reale per il Teorema Theorem 8.2) e  $\mathbf{w}_1$  un autovettore corrispondente di norma 1 (che possiamo assumere a valori reali per il teorema Theorem 8.3).

Possiamo completare  $\mathbf{w}_1$  a una base ortonormale  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$  (vedi note su norme e prodotti scalari). Sia  $\mathbf{U}$  la matrice le cui colonne sono i vettori  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ .

Allora  $\mathbf{U}^T \mathbf{U} = \mathbf{I}$ , infatti:

$$(\mathbf{U}^T \mathbf{U})_{i,j} = \sum_{k=1}^n U_{k,i} U_{k,j} = \mathbf{U}_{\bullet,i} \cdot \mathbf{U}_{\bullet,j} = \delta_{i,j}.$$

Inoltre, dato che  $\mathbf{A} \mathbf{U}_{\bullet,1} = \lambda_1 \mathbf{U}_{\bullet,1}$ , otteniamo:

$$\mathbf{A} \mathbf{U} = \mathbf{U} \begin{pmatrix} \lambda_1 & \boldsymbol{\eta} \\ \mathbf{0} & \mathbf{B} \end{pmatrix},$$

dove  $\boldsymbol{\eta}$  è un vettore riga e  $\mathbf{B}$  è una matrice quadrata non ancora specificata. Moltiplicando a sinistra per  $\mathbf{U}^T$ , otteniamo:

$$\mathbf{U}^T \mathbf{A} \mathbf{U} = \begin{pmatrix} \lambda_1 & \boldsymbol{\eta}^T \\ \mathbf{0} & \mathbf{B} \end{pmatrix}.$$

Dato che  $\mathbf{A}$  è simmetrica, segue che  $\mathbf{U}^T \mathbf{A} \mathbf{U}$  è simmetrica. Pertanto, la matrice a blocchi a destra è simmetrica, quindi  $\boldsymbol{\eta} = \mathbf{0}$  e  $\mathbf{B}$  è simmetrica.

Applicando l'induzione, esiste una matrice  $\mathbf{S}$  ortogonale tale che:

$$\mathbf{S}^T \mathbf{B} \mathbf{S} = \begin{pmatrix} \lambda_2 & & & \\ & \lambda_3 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

Pertanto, abbiamo:

$$\begin{aligned}
\begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S}^T \end{pmatrix} \mathbf{U}^T \mathbf{A} \mathbf{U} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} &= \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S}^T \end{pmatrix} \begin{pmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S}^T \mathbf{B} \mathbf{S} \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & & \mathbf{0}^T \\ & \begin{pmatrix} \lambda_2 & & \\ & \lambda_3 & \\ & & \ddots \\ \mathbf{0} & & & \lambda_n \end{pmatrix} & \\ & & & \end{pmatrix}.
\end{aligned}$$

Quindi, ponendo:

$$\mathbf{U} = \mathbf{U} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{pmatrix},$$

otteniamo che  $\mathbf{U}$  è la matrice ortogonale cercata. ■

## 8.6 Diagonalizzazione di matrici Hermitiane

**Theorem 8.6 — Teorema di Diagonalizzazione per matrici Hermitiane.** Sia  $\mathbf{A}$  una matrice hermitiana. Allora esiste una matrice unitaria  $\mathbf{U}$  tale che:

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{pmatrix},$$

dove  $\lambda_1, \lambda_2, \dots, \lambda_n$  sono gli autovalori della matrice  $\mathbf{A}$ .

*Proof.* La dimostrazione di questo teorema segue un argomento simile a quello del teorema di diagonalizzazione per matrici simmetriche, con le seguenti modifiche:

1. **Sostituzione dei Termini:** Sostituiamo i termini “simmetrico” e “ortogonale” con “hermitiano” e “unitario”, rispettivamente.
2. **Trasposizione e Coniugazione:** Tutte le trasposizioni di matrici reali, indicate con  $(\cdot)^T$ , devono essere reinterpretate come trasposizioni con coniugazione, indicate con  $(\cdot)^H$ .  
In altre parole, se  $\mathbf{A}$  è hermitiana, per ogni autovalore  $\lambda$  e autovettore  $\mathbf{u}$  della matrice  $\mathbf{A}$ , si ha:

$$\mathbf{A} \mathbf{u} = \lambda \mathbf{u}$$

Applicando la trasposizione con coniugazione, otteniamo:

$$\mathbf{u}^H \mathbf{A} \mathbf{u} = \lambda \|\mathbf{u}\|_2^2.$$

Da ciò, possiamo dedurre che l’autovalore  $\lambda$  è reale e, quindi, la matrice  $\mathbf{A}$  può essere diagonalizzata mediante una matrice unitaria  $\mathbf{U}$ , che riporta gli autovalori lungo la diagonale.

Quindi, il processo di diagonalizzazione per matrici hermitiane è simile a quello per matrici simmetriche, con l’aggiunta della coniugazione complessa. ■

## 8.7 Spettro, raggio spettrale e localizzazione degli autovalori sul piano complesso

**Definition 8.11 — Spettro.** L'insieme degli autovalori di una matrice  $\mathbf{A}$  si chiama *spettro* della matrice e si indica usualmente col simbolo  $\sigma(\mathbf{A})$ .

**Definition 8.12 — Raggio spettrale.** Data una matrice quadrata  $\mathbf{A}$  con autovalori  $\lambda_1, \lambda_2, \dots, \lambda_n$  il numero

$$\rho(\mathbf{A}) = \max\{|\lambda_i| : i = 1, 2, \dots, n\},$$

è detto *raggio spettrale* della matrice  $\mathbf{A}$ .

#### Esempio

Determinare il raggio spettrale della seguente matrice

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

Calcoliamo gli autovalori come zeri del polinomio caratteristico

$$\begin{aligned} p_{\mathbf{A}}(\lambda) &= |\mathbf{A} - \lambda \mathbf{I}| = \dots \\ &= (1 - \lambda)(1 - \lambda + i)(1 - \lambda - i), \end{aligned}$$

Determiniamo il raggio spettrale dalla definizione

$$\rho(\mathbf{A}) = \max\{|1|, |1 + i|, |1 - i|\} = \sqrt{2}.$$

## 8.8 Teorema dei cerchi di Gerschgorin

**Theorem 8.7 — Teorema dei cerchi di Gerschgorin.** Sia  $\mathbf{A} \in \mathbb{K}^{n \times n}$ . Definiamo l'insieme dei numeri complessi:

$$S = \bigcup_{i=1}^n \left\{ z \in \mathbb{C} : |A_{i,i} - z| \leq \sum_{j=1, j \neq i}^n |A_{i,j}| \right\}.$$

Allora, ogni autovalore di  $\mathbf{A}$  appartiene all'insieme  $S$ .

*Proof.* Sia  $\lambda$  un autovalore di  $\mathbf{A}$  con il corrispondente autovettore  $\mathbf{u}$ . Sia  $k$  l'indice della componente di  $\mathbf{u}$  con il modulo massimo:

$$|u_k| = \max_{i=1, \dots, n} |u_i|.$$

Consideriamo l'equazione caratteristica:

$$\mathbf{A}\mathbf{u} - \lambda\mathbf{u} = \mathbf{0}.$$

Analizzando la  $k$ -esima componente, otteniamo:

$$\sum_{j=1}^n A_{k,j} u_j - \lambda u_k = 0.$$

Isoliamo il termine  $\lambda u_k$ :

$$(A_{k,k} - \lambda) u_k = - \sum_{j=1, j \neq k}^n A_{k,j} u_j.$$

Calcoliamo il modulo di entrambi i membri dell'equazione e dividiamo per  $|u_k| > 0$ :

$$|A_{k,k} - \lambda| \leq \sum_{j=1, j \neq k}^n |A_{k,j}| \left| \frac{u_j}{u_k} \right|.$$

Poiché  $\left| \frac{u_j}{u_k} \right| \leq 1$ , otteniamo:

$$|A_{k,k} - \lambda| \leq \sum_{j=1, j \neq k}^n |A_{k,j}|.$$

Quindi,  $\lambda$  soddisfa l'inequazione che definisce il cerchio di Gerschgorin centrato in  $A_{k,k}$  con raggio  $\sum_{j=1, j \neq k}^n |A_{k,j}|$ . Pertanto,  $\lambda$  appartiene all'insieme  $S$ . ■

#### Esempio

$$\mathbf{A} = \begin{pmatrix} 1+i & \frac{1}{2} & \frac{1}{2} \\ 1 & 2 & 0 \\ \frac{1}{4} & -\frac{1}{4} & 6+i \end{pmatrix}, \quad \sigma(\mathbf{A}) = \begin{cases} \lambda_1 = 0.69 + 0.82i, \\ \lambda_2 = 2.29 + 0.18i, \\ \lambda_3 = 6.02 + 1.01i \end{cases}$$

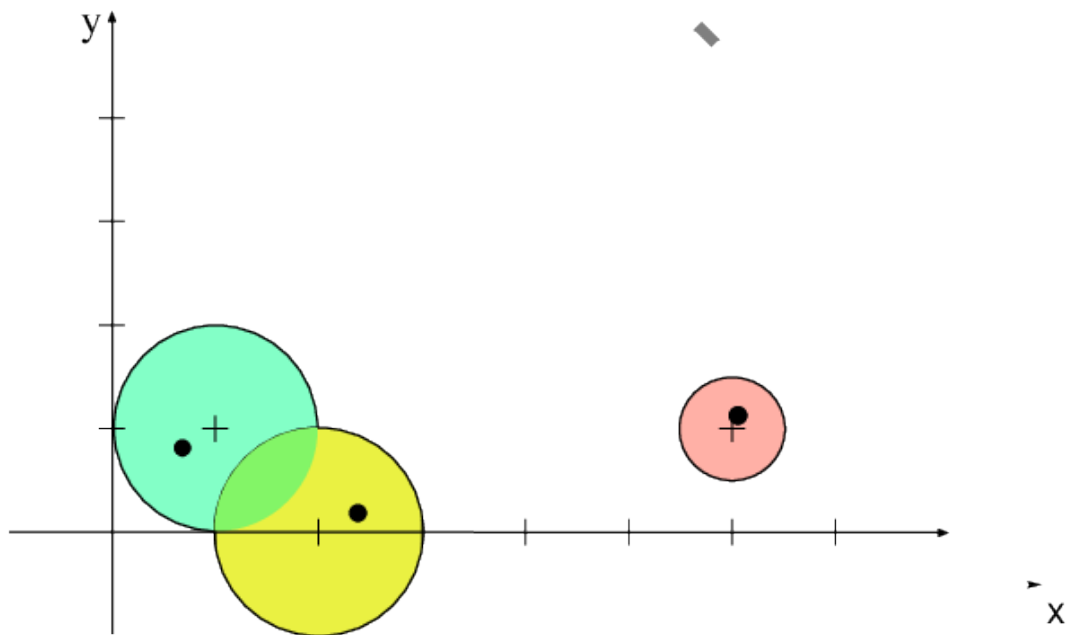


Figure 8.1: Gerschgorin circles

#### Esempio

Possiamo avere anche situazioni *insolite*. Per esempio,

$$\mathbf{A} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & -2 & 2 \end{pmatrix}, \quad \sigma(\mathbf{A}) = \begin{cases} \lambda_1 = 4, \\ \lambda_2 = 2(1+i), \\ \lambda_3 = 2(1-i) \end{cases}$$

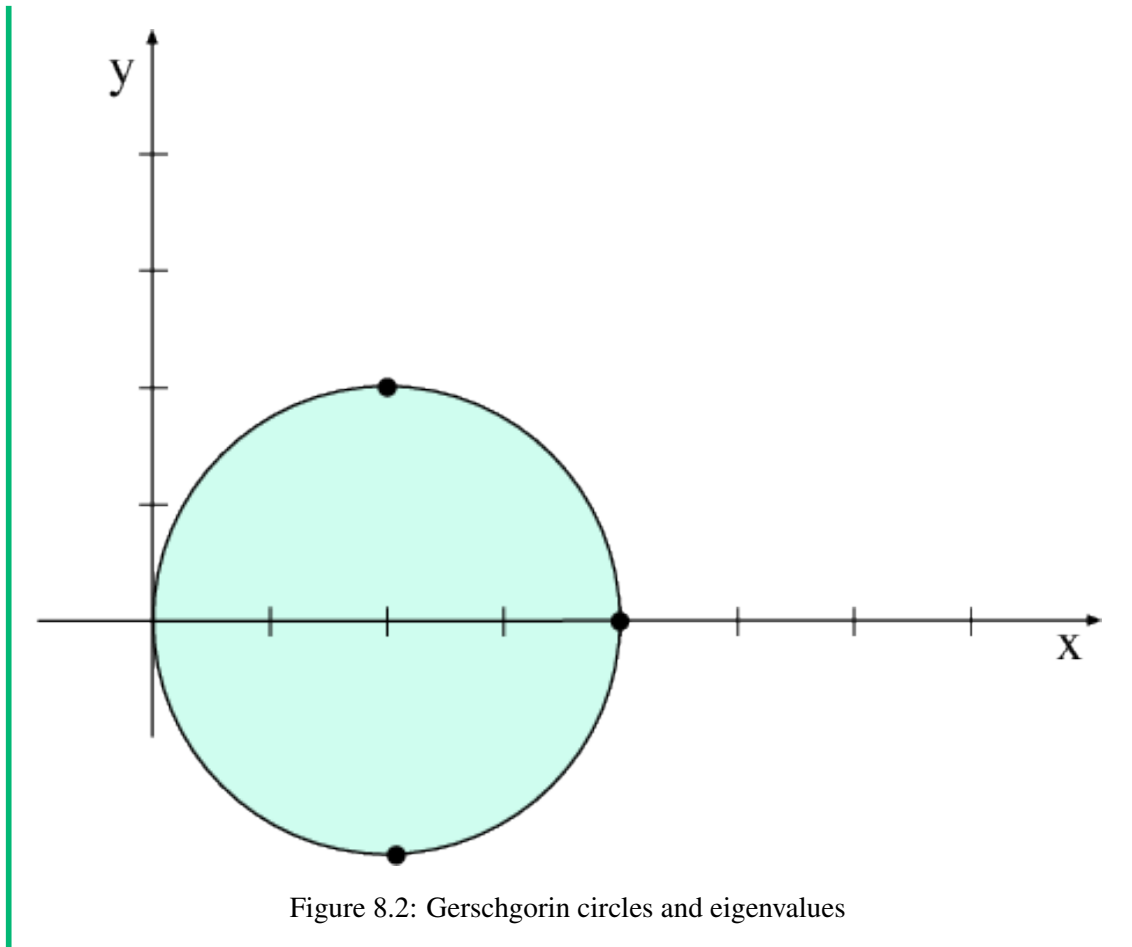


Figure 8.2: Gerschgorin circles and eigenvalues

## 8.9 Matrici a Diagonale Dominante

**Definition 8.13 — Dominanza Diagonale.** Una matrice quadrata  $\mathbf{A}$  è chiamata *diagonale dominante* se soddisfa la seguente condizione per ogni riga  $i$ :

$$|A_{i,i}| \geq \sum_{j=1, j \neq i}^n |A_{i,j}|,$$

e la disuguaglianza deve essere stretta per *almeno* un indice  $i$ . Se la disuguaglianza è stretta per *ogni* indice  $i$ , allora la matrice  $\mathbf{A}$  è detta *diagonale strettamente dominante*.

**Theorem 8.8 — Teorema sulla Diagonale Dominante.** Se una matrice  $\mathbf{A}$  è a diagonale strettamente dominante, allora è non singolare.

*Proof.* Per dimostrare che una matrice a diagonale strettamente dominante è non singolare, consideriamo il fatto che il cerchio di Gerschgorin centrato in  $A_{i,i}$  con raggio  $\sum_{j=1, j \neq i}^n |A_{i,j}|$  non può contenere l'origine se la matrice è a diagonale strettamente dominante. Questo implica che lo zero non può essere un autovalore della matrice  $\mathbf{A}$ . Di conseguenza, la matrice  $\mathbf{A}$  non è singolare. ■

## 8.10 Matrici Definite Positive

**Definition 8.14 — Matrice Definita Positiva.** Una matrice quadrata  $\mathbf{A}$  è definita *positiva* se soddisfa le seguenti condizioni per ogni vettore  $\mathbf{u} \in \mathbb{K}^n$ :

1.  $\mathbf{u}^T \mathbf{A} \mathbf{u} > 0$  per ogni vettore  $\mathbf{u} \neq \mathbf{0}$

2.  $\mathbf{u}^T \mathbf{A} \mathbf{u} = 0$  solo se  $\mathbf{u} = \mathbf{0}$

Se la seconda condizione non è soddisfatta e  $\mathbf{u}^T \mathbf{A} \mathbf{u} \geq 0$  per ogni  $\mathbf{u}$ , allora la matrice  $\mathbf{A}$  è definita *semipositiva*.

**Definition 8.15 — Matrice Simmetrica e Definita Positiva.** Quando una matrice definita positiva è anche simmetrica, essa è detta *simmetrica e definita positiva*. Questa proprietà è abbreviata in *SPD*, dall'inglese **S**ymmetric **P**ositive **D**efinite.

## 8.11 Autovalori di Matrici SPD

**Theorem 8.9 — Teorema degli Autovalori di Matrici SPD.** Se  $\mathbf{A}$  è una matrice definita positiva, allora tutti i suoi autovalori sono positivi.

*Proof.* Sia  $\lambda$  un autovalore di  $\mathbf{A}$ , e sia  $\mathbf{u}$  un autovettore associato, cioè:

$$\mathbf{A} \mathbf{u} = \lambda \mathbf{u}$$

Moltiplicando entrambi i lati dell'equazione per  $\mathbf{u}^T$ , otteniamo:

$$\mathbf{u}^T \mathbf{A} \mathbf{u} = \lambda \mathbf{u}^T \mathbf{u} = \lambda \|\mathbf{u}\|_2^2$$

Poiché  $\mathbf{A}$  è definita positiva,  $\mathbf{u}^T \mathbf{A} \mathbf{u} > 0$  per ogni  $\mathbf{u} \neq \mathbf{0}$ , il che implica che:

$$\lambda = \frac{\mathbf{u}^T \mathbf{A} \mathbf{u}}{\|\mathbf{u}\|_2^2} > 0$$

Pertanto, ogni autovalore  $\lambda$  di  $\mathbf{A}$  è positivo. ■

## 8.12 Determinante di Matrici SPD

**Theorem 8.10 — Determinante di una Matrice SPD.** Se  $\mathbf{A}$  è una matrice simmetrica e definita positiva (SPD), allora il suo determinante è positivo:

$$|\mathbf{A}| > 0.$$

*Proof.* Dato che  $\mathbf{A}$  è una matrice simmetrica e definita positiva, esiste una matrice ortogonale reale  $\mathbf{U}$  e una matrice diagonale  $\mathbf{\Lambda}$  con autovalori positivi, tali che:

$$\mathbf{A} = \mathbf{U}^T \mathbf{\Lambda} \mathbf{U},$$

dove  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$  e  $\lambda_i > 0$  sono gli autovalori di  $\mathbf{A}$ . Utilizzando la proprietà del determinante per la decomposizione in fattori ortogonali:

$$\begin{aligned} |\mathbf{A}| &= |\mathbf{U}^T \mathbf{\Lambda} \mathbf{U}| \\ &= |\mathbf{U}^T| |\mathbf{\Lambda}| |\mathbf{U}| \\ &= |\mathbf{U}|^T |\mathbf{\Lambda}| |\mathbf{U}| \\ &= |\mathbf{\Lambda}| \\ &= \lambda_1 \lambda_2 \cdots \lambda_n > 0, \end{aligned}$$

dove  $|\mathbf{U}|^T |\mathbf{U}| = |\mathbf{U}^T \mathbf{U}| = |\mathbf{I}| = 1$ .

Poiché  $\lambda_i > 0$  per tutti  $i$ , il determinante  $|\mathbf{A}|$  è positivo. ■

### 8.13 Blocchi Diagonali di Matrici SPD

**Theorem 8.11 — Proprietà dei Blocchi Diagonali in Matrici SPD.** Sia  $\mathbf{M} \in \mathbb{K}^{(n+m) \times (n+m)}$  una matrice SPD della forma:

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{C} \end{pmatrix},$$

dove  $\mathbf{A} \in \mathbb{K}^{n \times n}$  e  $\mathbf{C} \in \mathbb{K}^{m \times m}$ , allora sia  $\mathbf{A}$  che  $\mathbf{C}$  sono anch'esse SPD.

*Proof.* Consideriamo un vettore  $\boldsymbol{\omega}$  della forma:

$$\boldsymbol{\omega} = \begin{pmatrix} \mathbf{u} \\ \mathbf{0} \end{pmatrix},$$

dove  $\mathbf{u} \in \mathbb{K}^n$ . Calcoliamo:

$$\begin{aligned} 0 &\leq \boldsymbol{\omega}^T \mathbf{M} \boldsymbol{\omega} \\ &= (\mathbf{u}^T \quad \mathbf{0}^T) \begin{pmatrix} \mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{C} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{0} \end{pmatrix} \\ &= \mathbf{u}^T \mathbf{A} \mathbf{u}. \end{aligned}$$

Poiché  $\boldsymbol{\omega}^T \mathbf{M} \boldsymbol{\omega} = \mathbf{u}^T \mathbf{A} \mathbf{u}$  è sempre non negativo e uguale a zero solo se  $\mathbf{u} = \mathbf{0}$ , segue che  $\mathbf{A}$  deve essere SPD.

Analogamente, considerando un vettore della forma:

$$\boldsymbol{\omega} = \begin{pmatrix} \mathbf{0} \\ \mathbf{u} \end{pmatrix},$$

si mostra che  $\mathbf{C}$  è SPD. ■

### 8.14 Teorema di Sylvester<sup>1</sup>

**Theorem 8.12 — Teorema di Sylvester.** Sia  $\mathbf{M} \in \mathbb{R}^{n \times n}$  una matrice con la seguente partizione:

$$\mathbf{M} = \begin{pmatrix} \mathbf{A}^{(k)} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{C} \end{pmatrix},$$

dove  $\mathbf{A}^{(k)} \in \mathbb{R}^{k \times k}$  e  $\mathbf{C} \in \mathbb{R}^{(n-k) \times (n-k)}$ . Allora, le seguenti affermazioni sono equivalenti:

1.  $\mathbf{M}$  è simmetrica e definita positiva (SPD).
2.  $|\mathbf{A}^{(k)}| > 0$  per ogni  $k = 1, 2, \dots, n$ .

*Proof.*

- [1  $\Rightarrow$  2]

Se  $\mathbf{M}$  è SPD, allora ogni blocco principale  $\mathbf{A}^{(k)}$  deve essere SPD. Questo implica che  $|\mathbf{A}^{(k)}| > 0$  per ogni  $k$ .

- [2  $\Rightarrow$  1]

La dimostrazione si effettua per induzione sul numero di righe e colonne della matrice.

- **Caso base** ( $n = 1$ ): Quando  $n = 1$ , la matrice  $\mathbf{M}$  è uno scalare, e il teorema è triviale poiché  $\mathbf{M}$  è SPD se e solo se il suo valore è positivo.

<sup>1</sup>James Joseph Sylvester (1814-1897)

- **Passo induttivo** ( $n > 1$ ): Supponiamo che il teorema sia vero per matrici di dimensione  $n - 1$  (ipotesi induttiva). Consideriamo una matrice  $\mathbf{M}$  di dimensione  $n$  e partizioniamola come segue:

$$\mathbf{M} = \begin{pmatrix} \mathbf{A}^{(n-1)} & \mathbf{b} \\ \mathbf{b}^T & A_{n,n} \end{pmatrix},$$

dove  $\mathbf{A}^{(n-1)}$  è una matrice  $(n-1) \times (n-1)$ , e  $\mathbf{b}$  è un vettore colonna di dimensione  $n-1$ .

Definiamo la matrice di blocco  $\mathbf{L}$  come:

$$\mathbf{L} = \begin{pmatrix} \mathbf{I} & -(\mathbf{A}^{(n-1)})^{-T} \mathbf{b} \\ \mathbf{0} & 1 \end{pmatrix},$$

e la sua trasposta:

$$\mathbf{L}^T = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{b}^T (\mathbf{A}^{(n-1)})^{-1} & 1 \end{pmatrix}.$$

Calcoliamo:

$$\mathbf{L}^T \mathbf{M} \mathbf{L} = \begin{pmatrix} \mathbf{A}^{(n-1)} & \mathbf{0} \\ \mathbf{0} & A_{n,n} - \mathbf{b}^T \mathbf{A}^{(n-1)} \mathbf{b} \end{pmatrix}.$$

Dal momento che  $|\mathbf{L}| = |\mathbf{L}^T| = 1$ , abbiamo:

$$\begin{aligned} |\mathbf{M}| &= |\mathbf{L}^{-T} \mathbf{L}^T \mathbf{M} \mathbf{L} \mathbf{L}^{-1}| \\ &= |\mathbf{L}^{-T}| |\mathbf{L}^T \mathbf{M} \mathbf{L}| |\mathbf{L}^{-1}| \\ &= \begin{vmatrix} \mathbf{A}^{(n-1)} & \mathbf{0} \\ \mathbf{0} & A_{n,n} - \mathbf{b}^T \mathbf{A}^{(n-1)} \mathbf{b} \end{vmatrix} \\ &= |\mathbf{A}^{(n-1)}| \cdot (A_{n,n} - \mathbf{b}^T \mathbf{A}^{(n-1)} \mathbf{b}). \end{aligned}$$

Per ipotesi induttiva,  $\mathbf{A}^{(n-1)}$  è SPD, quindi  $|\mathbf{A}^{(n-1)}| > 0$ . Inoltre,  $|\mathbf{M}| > 0$  implica che:

$$A_{n,n} - \mathbf{b}^T \mathbf{A}^{(n-1)} \mathbf{b} > 0.$$

Consideriamo ora un vettore  $\mathbf{z}$  generico e poniamo  $\boldsymbol{\omega} = \mathbf{L}^{-1} \mathbf{z}$ , con la partizione:

$$\boldsymbol{\omega} = \begin{pmatrix} \mathbf{u} \\ \alpha \end{pmatrix},$$

dove  $\mathbf{u}$  è un vettore di dimensione  $n-1$  e  $\alpha$  è uno scalare. Allora:

$$\begin{aligned} \mathbf{z}^T \mathbf{M} \mathbf{z} &= \boldsymbol{\omega}^T \mathbf{L}^T \mathbf{M} \mathbf{L} \boldsymbol{\omega} \\ &= (\mathbf{u}^T \quad \alpha) \begin{pmatrix} \mathbf{A}^{(n-1)} & \mathbf{0} \\ \mathbf{0} & A_{n,n} - \mathbf{b}^T \mathbf{A}^{(n-1)} \mathbf{b} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \alpha \end{pmatrix} \\ &= \mathbf{u}^T \mathbf{A}^{(n-1)} \mathbf{u} + \alpha^2 \cdot (A_{n,n} - \mathbf{b}^T \mathbf{A}^{(n-1)} \mathbf{b}). \end{aligned}$$

Poiché  $\mathbf{A}^{(n-1)}$  è SPD per ipotesi induttiva, abbiamo che:

$$\mathbf{z}^T \mathbf{M} \mathbf{z} \geq 0.$$

Se  $\mathbf{z}^T \mathbf{M} \mathbf{z} = 0$ , allora  $\alpha = 0$  e  $\mathbf{u} = \mathbf{0}$ , quindi  $\mathbf{z} = \mathbf{L}^{-1} \mathbf{0} = \mathbf{0}$ . Di conseguenza,  $\mathbf{M}$  è SPD. ■



## 9. Norme di matrici

Desideriamo estendere la nozione di norma, che abbiamo introdotto per i vettori, al caso delle matrici. Poiché l'idea di *lunghezza di una matrice* non è immediatamente intuitiva, affronteremo il problema in modo formale.

Per iniziare, esamineremo alcune proprietà interessanti delle norme vettoriali, che ci aiuteranno a comprendere e definire le norme per le matrici.

### 9.1 Cos'è la norma di una matrice?

Consideriamo una matrice  $\mathbf{A} \in \mathbb{K}^{m \times n}$  e un vettore colonna  $\mathbf{x} \in \mathbb{K}^n$ . Il prodotto  $\mathbf{Ax}$  risulta essere un vettore colonna di dimensione  $m$ . Ora, introduciamo due norme vettoriali, che indichiamo con  $\|\cdot\|_a$  e  $\|\cdot\|_b$ , rispettivamente definite negli spazi  $\mathbb{K}^m$  e  $\mathbb{K}^n$ . L'obiettivo è trovare la più piccola costante  $K$  (ammesso che esista) tale che:

$$\|\mathbf{Ax}\|_a \leq K \|\mathbf{x}\|_b, \quad \forall \mathbf{x} \in \mathbb{K}^n \quad (9.1)$$

Per  $\mathbf{x} = \mathbf{0}$ , la disuguaglianza precedente è soddisfatta banalmente per qualsiasi valore di  $K$ . Tuttavia, possiamo restringere la nostra attenzione ai vettori  $\mathbf{x} \neq \mathbf{0}$ . Per questi, si può scrivere:

$$\|\mathbf{Ax}\|_a = \frac{\|\mathbf{Ax}\|_a}{\|\mathbf{x}\|_b} \|\mathbf{x}\|_b \leq \left[ \sup_{\mathbf{y} \neq \mathbf{0}} \frac{\|\mathbf{Ay}\|_a}{\|\mathbf{y}\|_b} \right] \|\mathbf{x}\|_b$$

Questa espressione ci permette di identificare  $K$  come l'estremo superiore del rapporto  $\frac{\|\mathbf{Ay}\|_a}{\|\mathbf{y}\|_b}$ , dove  $\mathbf{y}$  è un vettore non nullo.

Nel seguito dimostreremo che, nelle condizioni considerate, tale estremo superiore esiste ed è finito.

Osserviamo che questa quantità dipende unicamente dalla matrice  $\mathbf{A}$  e dalla scelta delle norme vettoriali  $\|\cdot\|_a$  e  $\|\cdot\|_b$ , come è ragionevole aspettarsi.

Pertanto, questa quantità può essere vista come una caratteristica della matrice  $\mathbf{A}$  e si comporta come una norma, nel senso che soddisfa le tre proprietà fondamentali richieste per la definizione assiomatica di norma.

**Theorem 9.1 — Norma Indotta.** Per ogni matrice  $\mathbf{A} \in \mathbb{K}^{m \times n}$ , vale la seguente relazione:

$$\sup_{\mathbf{y} \neq \mathbf{0}} \frac{\|\mathbf{Ay}\|_a}{\|\mathbf{y}\|_b} = \max_{\|\mathbf{y}\|_b=1} \|\mathbf{Ay}\|_a.$$

*Proof.* Partiamo dalla seguente espressione:

$$\frac{\|\mathbf{A}\mathbf{y}\|_a}{\|\mathbf{y}\|_b} = \|\mathbf{A}\left(\frac{\mathbf{y}}{\|\mathbf{y}\|_b}\right)\|_a = \|\mathbf{A}\mathbf{z}\|_a, \quad \text{con } \mathbf{z} = \frac{\mathbf{y}}{\|\mathbf{y}\|_b}.$$

Dato che  $\|\mathbf{z}\|_b = 1$ , possiamo riscrivere il supremum come:

$$\sup_{\mathbf{y} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{y}\|_a}{\|\mathbf{y}\|_b} = \sup_{\|\mathbf{z}\|_b=1} \|\mathbf{A}\mathbf{z}\|_a.$$

Per dimostrare che questo supremum esiste ed è finito, procediamo in modo analogo alla dimostrazione dell'equivalenza delle norme. Introduciamo l'insieme

$$S = \{\mathbf{z} \in \mathbb{K}^n : \|\mathbf{z}\|_b = 1\}.$$

Essendo  $S$  un insieme compatto in  $\mathbb{K}^n$  e la funzione  $\|\mathbf{A}\mathbf{z}\|_a$  continua (essendo una composizione di funzioni continue), tale funzione assume un massimo su  $S$ . Sia  $\mathbf{z}_M \in S$  il vettore che massimizza  $\|\mathbf{A}\mathbf{z}\|_a$ , ossia:

$$\|\mathbf{A}\mathbf{z}_M\|_a = \max_{\|\mathbf{z}\|_b=1} \|\mathbf{A}\mathbf{z}\|_a.$$

Poiché il supremum rappresenta il maggiorante più piccolo possibile, e non può essere strettamente maggiore del massimo trovato, concludiamo che:

$$\max_{\|\mathbf{z}\|_b=1} \|\mathbf{A}\mathbf{z}\|_a = \|\mathbf{A}\mathbf{z}_M\|_a \leq \sup_{\|\mathbf{z}\|_b=1} \|\mathbf{A}\mathbf{z}\|_a \leq \max_{\|\mathbf{z}\|_b=1} \|\mathbf{A}\mathbf{z}\|_a.$$

Questo dimostra che il supremum coincide con il massimo. ■

**Theorem 9.2 — Norma di Matrice.** La quantità definita come

$$\|\mathbf{A}\| = \sup_{\mathbf{y} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{y}\|_a}{\|\mathbf{y}\|_b}$$

soddisfa le tre proprietà assiomatiche delle norme. Per questo motivo, possiamo usare il simbolo  $\|\mathbf{A}\|$  per indicare la norma della matrice  $\mathbf{A}$ .

*Proof.* Dimostriamo che la funzione che associa a ogni matrice  $\mathbf{A} \in \mathbb{K}^{m \times n}$  il valore  $\|\mathbf{A}\|$  soddisfa le tre proprietà fondamentali delle norme.

**1. Non negatività:**

Per definizione,  $\|\mathbf{A}\|$  è non negativa. Verifichiamo ora che  $\|\mathbf{A}\| = 0$  se e solo se  $\mathbf{A} = \mathbf{0}$  (la matrice nulla).

- Se  $\mathbf{A} = \mathbf{0}$ , allora  $\mathbf{A}\mathbf{x} = \mathbf{0}$  per ogni vettore  $\mathbf{x} \in \mathbb{K}^n$ . Questo implica:

$$\frac{\|\mathbf{A}\mathbf{x}\|_a}{\|\mathbf{x}\|_b} = 0 \quad \forall \mathbf{x} \neq \mathbf{0},$$

da cui segue che  $\|\mathbf{A}\| = 0$ .

- Viceversa, dimostriamo che  $\|\mathbf{A}\| = 0$  implica  $\mathbf{A} = \mathbf{0}$ . Ragioniamo per contrapposizione: supponiamo che  $\mathbf{A} \neq \mathbf{0}$  e mostriamo che in questo caso  $\|\mathbf{A}\| > 0$ . Sia  $\mathbf{e}_i$  l' $i$ -esimo vettore della base canonica. Se  $\mathbf{A} \neq \mathbf{0}$ , allora esiste almeno un indice  $i$  per cui  $\mathbf{A}\mathbf{e}_i \neq \mathbf{0}$ . Pertanto:

$$\frac{\|\mathbf{A}\mathbf{e}_i\|_a}{\|\mathbf{e}_i\|_b} > 0,$$

il che implica  $\|\mathbf{A}\| > 0$ .

**2. Subadditività (Disuguaglianza triangolare):**

Siano  $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{m \times n}$ . Dobbiamo verificare che:

$$\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|.$$

Per ogni  $\mathbf{x} \in \mathbb{K}^n$ , vale:

$$\frac{\|(\mathbf{A} + \mathbf{B})\mathbf{x}\|_a}{\|\mathbf{x}\|_b} \leq \frac{\|\mathbf{A}\mathbf{x}\|_a}{\|\mathbf{x}\|_b} + \frac{\|\mathbf{B}\mathbf{x}\|_a}{\|\mathbf{x}\|_b}.$$

Prendendo il supremum su tutti i vettori  $\mathbf{x} \neq \mathbf{0}$ , otteniamo:

$$\sup_{\mathbf{y} \neq \mathbf{0}} \frac{\|(\mathbf{A} + \mathbf{B})\mathbf{y}\|_a}{\|\mathbf{y}\|_b} \leq \sup_{\mathbf{y} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{y}\|_a}{\|\mathbf{y}\|_b} + \sup_{\mathbf{y} \neq \mathbf{0}} \frac{\|\mathbf{B}\mathbf{y}\|_a}{\|\mathbf{y}\|_b},$$

ovvero:

$$\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|.$$

**3. Omogeneità:**

Sia  $\lambda \in \mathbb{K}$  uno scalare arbitrario. Verifichiamo che:

$$\|\lambda \mathbf{A}\| = |\lambda| \|\mathbf{A}\|.$$

Abbiamo:

$$\|\lambda \mathbf{A}\| = \sup_{\mathbf{y} \neq \mathbf{0}} \frac{\|\lambda \mathbf{A}\mathbf{y}\|_a}{\|\mathbf{y}\|_b}.$$

Poiché  $\|\lambda \mathbf{A}\mathbf{y}\|_a = |\lambda| \|\mathbf{A}\mathbf{y}\|_a$ , otteniamo:

$$\|\lambda \mathbf{A}\| = |\lambda| \sup_{\mathbf{y} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{y}\|_a}{\|\mathbf{y}\|_b} = |\lambda| \|\mathbf{A}\|.$$

■

**Osservazione**

Siano  $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{m \times m}$  e supponiamo di utilizzare la stessa norma per  $\mathbb{K}^m$ , cioè  $\|\cdot\|_a = \|\cdot\|_b$ . Consideriamo la norma del prodotto delle matrici  $\mathbf{AB}$ , definita come:

$$\|\mathbf{AB}\| = \sup_{\mathbf{y} \neq \mathbf{0}} \frac{\|\mathbf{AB}\mathbf{y}\|}{\|\mathbf{y}\|}. \quad (9.2)$$

A partire da Equation 9.2 possiamo derivare la seguente disuguaglianza:

$$\frac{\|\mathbf{AB}\mathbf{y}\|}{\|\mathbf{y}\|} \leq \frac{\|\mathbf{A}\| \|\mathbf{B}\mathbf{y}\|}{\|\mathbf{y}\|} = \|\mathbf{A}\| \frac{\|\mathbf{B}\mathbf{y}\|}{\|\mathbf{y}\|} \leq \|\mathbf{A}\| \|\mathbf{B}\|.$$

Poiché questa disuguaglianza è valida per ogni  $\mathbf{y} \in \mathbb{K}^n$ , possiamo concludere che:

$$\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|.$$

Questa relazione dimostra che la norma del prodotto di due matrici non supera mai il prodotto delle norme delle singole matrici.

## 9.2 Definizione assiomatica di norma

L'ultima osservazione suggerisce di introdurre la seguente definizione per le norme di matrice.

**Definition 9.1** Siano  $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{m \times n}$  e  $\alpha$  uno scalare. La norma di una matrice deve soddisfare le seguenti proprietà:

1. **Positività:**  $\|\mathbf{A}\| \geq 0$  e  $\|\mathbf{A}\| = 0$  se e solo se  $\mathbf{A} = \mathbf{0}$ . Questo significa che la norma è sempre un numero non negativo e che è nulla solo nel caso in cui la matrice sia la matrice nulla.
2. **Omogeneità rispetto a scalari:**  $\|\alpha \mathbf{A}\| = |\alpha| \|\mathbf{A}\|$ . La norma di una matrice moltiplicata per uno scalare è pari al valore assoluto dello scalare moltiplicato per la norma della matrice.
3. **Disuguaglianza triangolare:**  $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ . La norma della somma di due matrici non supera la somma delle norme delle singole matrici.
4. **Submoltiplicatività** (se  $m = n$ ):  $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|$ . Nel caso di matrici quadrate, la norma del prodotto di due matrici è minore o uguale al prodotto delle loro norme.

## 9.3 Norme compatibili

Nel contesto delle matrici e dei vettori, è spesso utile considerare come le norme delle matrici e dei vettori interagiscono tra loro. In particolare, una norma matriciale può essere definita in modo da essere *compatibile* con le norme dei vettori sui quali agisce la matrice. Questo concetto di compatibilità garantisce che la norma di un vettore trasformato da una matrice non cresca in maniera incontrollata rispetto alla norma del vettore originale.

### 9.3.1 Definizione di norme compatibili

**Definition 9.2** Una norma matriciale  $\|\cdot\|$  si dice *compatibile* con due norme vettoriali  $\|\cdot\|_a$  e  $\|\cdot\|_b$  se, per ogni matrice  $\mathbf{M} \in \mathbb{K}^{m \times n}$  e per ogni vettore  $\mathbf{x} \in \mathbb{K}^n$ , si verifica la seguente disuguaglianza:

$$\|\mathbf{M}\mathbf{x}\|_a \leq \|\mathbf{M}\| \|\mathbf{x}\|_b.$$

Questa proprietà esprime il fatto che l'operazione di moltiplicare un vettore per una matrice non "amplifica" troppo la norma del vettore. In altre parole, la norma della matrice funge da limite superiore per l'effetto di  $\mathbf{M}$  sulla norma del vettore  $\mathbf{x}$ , mantenendo una coerenza tra le norme usate per la matrice e quelle per i vettori.

Un caso rilevante per l'analisi degli algoritmi che verranno discussi in seguito si presenta quando le norme  $\|\cdot\|, \|\cdot\|_a$  e  $\|\cdot\|_b$  appartengono allo stesso tipo. In particolare, ciò avviene quando consideriamo le norme  $\|\cdot\|, \|\cdot\|_1, \|\cdot\|_2$  e  $\|\cdot\|_\infty$ .

Le corrispondenti norme applicate alle matrici vengono denotate con gli stessi simboli.

## 9.4

**Theorem 9.3** La norma matriciale  $\|\cdot\|_\infty : \mathbb{K}^{m \times n} \mapsto \mathbb{R}$  definita da

$$\sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty}, \quad (9.3)$$

vale

$$\|\mathbf{A}\|_\infty = \max_{i=1}^m \left( \sum_{j=1}^n |A_{i,j}| \right).$$

*Proof.* Osserviamo che se  $\mathbf{A} \in \mathbb{K}^{m \times n}$  vale

$$\begin{aligned} \|\mathbf{Ax}\|_\infty &= \max_{i=1}^m \left| \sum_{j=1}^n A_{i,j} x_j \right| \\ &\leq \max_{i=1}^m \sum_{j=1}^n |A_{i,j} x_j| \\ &\leq \max_{i=1}^m \left( \max_{j=1}^n |x_j| \right) \sum_{j=1}^n |A_{i,j}|, \\ &\leq \|\mathbf{x}\|_\infty \max_{i=1}^m \sum_{j=1}^n |A_{i,j}|, \end{aligned}$$

quindi

$$\frac{\|\mathbf{Ax}\|_\infty}{\|\mathbf{x}\|_\infty} \leq \max_{i=1}^m \sum_{j=1}^n |A_{i,j}|,$$

sia ora  $k$  la riga per cui

$$\max_{i=1}^m \sum_{j=1}^n |A_{i,j}| = \sum_{j=1}^n |A_{k,j}|,$$

e sia  $\mathbf{x}$  il vettore definito da

$$x_j = \begin{cases} +1 & A_{k,j} > 0 \\ -1 & A_{k,j} \leq 0 \end{cases},$$

allora avremo  $\|\mathbf{x}\|_\infty = 1$  e

$$\begin{aligned} (\mathbf{Ax})_k &= \sum_{j=1}^n A_{k,j} x_j \\ &= \sum_{j=1}^n |A_{k,j}| \\ &= \max_{i=1}^m \sum_{j=1}^n |A_{i,j}|, \\ &\Rightarrow \\ \|\mathbf{Ax}\|_\infty &\geq \max_{i=1}^m \sum_{j=1}^n |A_{i,j}|, \end{aligned}$$

e questo implica la Equation 9.3. ■

In modo analogo si dimostra il seguente teorema:

**Definition 9.3 — Norma 1.**

$$\sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{Ax}\|_1}{\|\mathbf{x}\|_1},$$

vale

$$\|\mathbf{A}\|_1 = \max_{j=1}^n \left( \sum_{i=1}^m |A_{i,j}| \right).$$

**Definition 9.4 — Norma-2.** La norma  $\|\cdot\|_2 : \mathbb{K}^{m \times n} \mapsto \mathbb{R}$  è definita come segue

$$\|\mathbf{A}\|_2 = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{Ax}\|_2}{\|\mathbf{x}\|_2}$$

**Theorem 9.4 — Norma-2 caratterizzazione.** Vale la seguente uguaglianza

$$\|\mathbf{A}\|_2 = \sqrt{\rho(\mathbf{A}^H \mathbf{A})}.$$

*Proof.* Osserviamo che

$$\|\mathbf{A}\|_2^2 = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{Ax}\|_2^2}{\|\mathbf{x}\|_2^2} = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^H \mathbf{A}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} \quad (9.4)$$

La matrice  $\mathbf{A}^H \mathbf{A}$  è una matrice hermitiana e quindi per il teorema di diagonalizzazione di matrici Hermitiane esiste una matrice unitaria  $\mathbf{U}$  che la diagonalizza

$$\mathbf{U}^H \mathbf{A}^H \mathbf{A} \mathbf{U} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}, \quad (9.5)$$

dove  $\lambda_1, \lambda_2, \dots, \lambda_n$  sono gli autovalori della matrice  $\mathbf{A}^H \mathbf{A}$ . Dalla Equation 9.4 segue

$$\begin{aligned} \|\mathbf{A}\|_2^2 &= \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^H \mathbf{U} \mathbf{U}^H \mathbf{A}^H \mathbf{A} \mathbf{U} \mathbf{U}^H \mathbf{x}}{\mathbf{x}^H \mathbf{U} \mathbf{U}^H \mathbf{x}} \\ &= \sup_{\mathbf{y} \neq \mathbf{0}} \frac{\mathbf{y}^H \mathbf{A}^H \mathbf{A} \mathbf{y}}{\mathbf{y}^H \mathbf{y}} \end{aligned}$$

dalla Equation 9.5

$$\begin{aligned} \|\mathbf{A}\|_2^2 &= \sup_{\mathbf{y} \neq \mathbf{0}} \frac{\sum_{i=1}^n \lambda_i y_i^2}{\sum_{i=1}^n y_i^2} \\ &= \max\{\lambda_1, \lambda_2, \dots, \lambda_n\} \\ &= \rho(\mathbf{A}^H \mathbf{A}) \end{aligned}$$

■

#### Osservazione

È da notare che nel caso  $n = m$  le matrici  $\mathbf{A}^H \mathbf{A}$  e  $\mathbf{A} \mathbf{A}^H$  sono simili, infatti

$$\mathbf{A}^{-H} (\mathbf{A}^H \mathbf{A}) \mathbf{A}^H = \mathbf{A} \mathbf{A}^H$$

e quindi

$$\|\mathbf{A}\|_2 = \sqrt{\rho(\mathbf{A}^H \mathbf{A})} = \sqrt{\rho(\mathbf{A} \mathbf{A}^H)}. \quad (9.6)$$

Nel caso  $n \neq m$  le matrici  $\mathbf{A} \mathbf{A}^H$  e  $\mathbf{A}^H \mathbf{A}$  non possono essere simili in quanto  $\mathbf{A} \mathbf{A}^H \in \mathbb{K}^{n \times n}$  e  $\mathbf{A}^H \mathbf{A} \in \mathbb{K}^{m \times m}$ . Si può comunque provare che la Equation 9.6 vale anche in questo caso.



## References

- [1] Roberto Bevilacqua, Dario Bini, Milvio Capovani, and Ornella Menchi. 1992. *Metodi numerici*. Zanichelli.
- [2] L. A. Hageman and D. M. Young. 1986. *Applied iterative methods*. Accademic Press.
- [3] W. W. Hager. 1988. *Applied numerical linear algebra*. Prentice Hall.
- [4] R. A. Horn and C. R. Johnson. 1985. *Matrix analysis*. Cambridge University Press.
- [5] S. Lang and I. Mencattini. 2014. *Algebra lineare*. Bollati Boringhieri.
- [6] J. M. Ortega. 1987. *Matrix theory (a second course)*. Plenum Press.
- [7] J. M. Ortega. 1987. *Matrix theory (a second course)*. Plenum Press.
- [8] J. M. Ortega and W. C. Rheinboldt. 1970. *Iterative solution of nonlinear equation in several variables*. Accademic Press.
- [9] A. Quarteroni, F. Saleri, and P. Gervasio. 2017. *Calcolo scientifico: Esercizi e problemi risolti con MATLAB e octave*. Springer Milan.
- [10] J. Stoer and R. Bulirsch. 1975. *Introduzione all'analisi numerica*. Zanichelli.
- [11] R. S. Varga. 1962. *Matrix iterative analysis*. Prentice Hall.
- [12] Z. I. Woznicki. 1994. Nonnegative splitting theory. *Japan J. Indust. Appl. Math* 11, (1994), 289–342.