

# FORME BILINEARI SIMMETRICHE E FORME HERMITIANE

## 1. FORME BILINEARI SIMMETRICHE

Sia  $V$  uno spazio vettoriale sul campo  $\mathbb{K}$ .

**Definizione 1.1.** Si dice **forma bilineare simmetrica** sullo spazio vettoriale  $V$ , un'applicazione:

$$\phi : V \times V \longrightarrow \mathbb{K}$$

tale che:

1.  $\phi(u, v) = \phi(v, u)$ ,  $\forall u, v \in V$  (*simmetria*)
2.  $\phi(u, \lambda v + \mu w) = \lambda\phi(u, v) + \mu\phi(u, w)$ ,  $\forall u, v, w \in V$ ,  $\forall \lambda, \mu \in \mathbb{K}$  (*linearità nella seconda variabile*)

Combinando le due proprietà si ricava immediatamente la linearità anche nella prima variabile:

$$(1) \quad \phi(\lambda v + \mu w, u) = \lambda\phi(v, u) + \mu\phi(w, u), \quad \forall u, v, w \in V, \quad \forall \lambda, \mu \in \mathbb{K}$$

**Esercizio 1.2.** *Dimostrare la 1.*

**Esempio 1.3.** 1. *Il prodotto scalare standard nello spazio delle  $n$ -uple:*

$$\phi : \mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K}$$

$$\phi(X, Y) := X^t Y = \sum_{i=1}^n x_i y_i, \quad \forall X, Y \in \mathbb{K}^n$$

*rappresenta un primo esempio di forma bilineare simmetrica, grazie alle proprietà del prodotto righe per colonne.*

*N. B.  $X$  e  $Y$  indicano vettori colonna e l'apice  $t$  indica la trasposizione.*

2. *In  $\mathbb{R}^2$  la funzione  $\phi((x_1, x_2), (y_1, y_2)) := x_1 y_1 + x_2 y_1 + x_1 y_2$  è una forma bilineare simmetrica.*
3. *In  $\mathbb{R}^2$  la funzione  $\phi((x_1, x_2), (y_1, y_2)) := x_1 y_1 + 2x_2 y_1 + x_1 y_2$  **non** è una forma bilineare simmetrica.*

**Proposizione 1.4.** Sia  $A = (a_{ij}) \in \mathbb{K}^{n,n}$  una matrice quadrata di ordine  $n$  nel campo  $\mathbb{K}$ . Allora l'applicazione

$$\phi_A(X, Y) := X^t A Y = \sum_{i,j=1}^n a_{ij} x_i y_j$$

rappresenta una forma bilineare simmetrica se e solo se  $A$  è simmetrica ( $A^t = A$ ).

**Dimostrazione:** la condizione è necessaria in quanto  $a_{ij} = \phi_A(e_i, e_j)$  e  $a_{ji} = \phi_A(e_j, e_i)$ , dove  $e_i$  ( $1 \leq i \leq j$ ) indica l' $i$ -esimo vettore della base canonica.

Sufficienza:  $\phi_A(X, Y) = X^t A Y = (X^t A Y)^t = Y^t A^t X = Y^t A X = \phi_A(Y, X)$ . c.v.d.

La proposizione precedente mostra come si possa associare ad una matrice simmetrica una forma bilineare simmetrica in  $\mathbb{K}^n$ . In realtà **ogni** forma bilineare simmetrica si può rappresentare così, pur di fissare una base nello spazio vettoriale  $V$ . Infatti, dati i vettori  $u, v \in V$  e detti  $X, Y \in \mathbb{K}^n$  i vettori colonna che ne rappresentano le coordinate rispetto alla base fissata  $\mathcal{B} = \{e_1, \dots, e_n\}$ , si trova:

$$\phi(u, v) = \phi\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{i,j=1}^n \phi(e_i, e_j) x_i y_j = X^t A Y,$$

dove si è posto  $A = (a_{ij}) = (\phi(e_i, e_j))$ .

**Definizione 1.5.** Tale matrice  $A \in \mathbb{K}^{n,n}$  è detta **matrice associata** alla forma  $\phi$ , nella base  $\mathcal{B}$ .

**Esercizio 1.6.** 1. Determinare la matrice associata agli esempi 1.3, sia nella base canonica che nella base  $\mathcal{B} := \{(1, 1), (1, -1)\}$ .  
2. Determinare, per gli stessi esempi, le espressioni nelle diverse coordinate e verificare la coerenza dei risultati ottenuti.

## 2. FORMA QUADRATICA ASSOCIATA

Sia  $V$  uno spazio vettoriale sul campo  $\mathbb{K}$  e  $\phi$  una forma bilineare simmetrica su  $V$ .

**Definizione 2.1.** L'applicazione:

$$q_\phi : V \longrightarrow \mathbb{K}$$

definita da

$$q_\phi(u) = \phi(u, u)$$

si dice **forma quadratica associata** a  $\phi$ .

Si osservi che  $q_\phi(\lambda v) = \phi(\lambda v, \lambda v) = \lambda^2 \phi(v, v)$ ,  $\forall \lambda \in \mathbb{K}$  e  $\forall v \in V$ , proprietà che giustifica la terminologia scelta. In seguito, in assenza di ambiguità, si ometterà il pedice del simbolo  $q_\phi$ .

Il seguente, importante, Teorema mostra come si possa ricostruire la forma bilineare a partire dalla forma quadratica associata, purché il campo non sia quello binario.

**Teorema 2.2. (di polarizzazione)** *Se  $\text{char}\mathbb{K} \neq 2$  allora:*

$$\phi(u, v) = \frac{1}{2}(q(u+v) - q(u) - q(v)), \quad \forall u, v \in V$$

**Dimostrazione:**

$$\begin{aligned} q(u+v) &= \phi(u+v, u+v) = \phi(u, u) + \phi(v, v) + \phi(u, v) + \phi(v, u) = \\ &= q(u) + q(v) + 2\phi(u, v) \end{aligned}$$

e la tesi segue immediatamente. c.v.d.

**Esempio 2.3.** *Una semplice verifica mostra come, nel campo binario, la forma bilineare simmetrica associata alla matrice  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  dia luogo alla forma quadratica nulla. Quindi non c'è speranza di trovare, per il campo binario, un analogo al teorema di polarizzazione!*

*Osservazione 2.4. (forma quadratica in coordinate)* Se  $A \in \mathbb{K}^{n,n}$  è la matrice associata alla forma  $\phi$  e  $X \in \mathbb{K}^n$  la colonna delle coordinate di un vettore  $v \in V$  (rispetto ad una base fissata), avremo:

$$q(v) = X^t A X = \sum_{i,j=1}^n a_{ij} x_i x_j = \sum_{i=1}^n a_{ii} (x^i)^2 + 2 \sum_{i < j} a_{ij} x_i x_j.$$

### 3. PRODOTTI SCALARI

**Definizione 3.1.** Sia  $V$  uno spazio vettoriale reale. Una forma bilineare simmetrica  $\phi$  si dice *positiva* se

$$q(v) = \phi(v, v) \geq 0, \quad \forall v \in V.$$

Si dice *negativa* se

$$q(v) = \phi(v, v) \leq 0, \quad \forall v \in V.$$

Si dice *definita positiva* (o *negativa*) se, inoltre,

$$q(v) = 0 \iff v = 0$$

**Esercizio 3.2.** 1. *Dimostrare che la forma bilineare simmetrica di  $\mathbb{R}^2$ :*

$$\phi((x_1, x_2), (y_1, y_2)) = x_1y_1 + x_2y_1 + x_1y_2 + 4x_2y_2$$

*è definita positiva.*

2. *Dimostrare che la forma bilineare simmetrica:*

$$\phi((x_1, x_2), (y_1, y_2)) = x_1y_1 - x_2y_1 - x_1y_2 + x_2y_2$$

*è positiva ma non definita positiva.*

3. *Cosa si può dire della seguente forma bilineare simmetrica?*

$$\phi((x_1, x_2), (y_1, y_2)) = x_1y_1 + x_2y_1 + x_1y_2 - 4x_2y_2$$

**Definizione 3.3.** Una forma bilineare simmetrica definita positiva si dice anche **prodotto scalare**.

**Definizione 3.4.** Sia  $V$  uno spazio vettoriale reale dotato di un prodotto scalare  $\phi$ . La quantità

$$\|v\| := \sqrt{q(v)} = \sqrt{\phi(v, v)}$$

si dice **norma o lunghezza** del vettore  $v \in V$ .

**Teorema 3.5. (disuguaglianza di Cauchy-Schwarz)** *Sia  $V$  uno spazio vettoriale reale dotato di un prodotto scalare  $\phi$ . Allora*

$$|\phi(u, v)| \leq \|u\| \|v\|, \quad \forall u, v \in V$$

**Dimostrazione:** Poiché,  $\forall x \in \mathbb{R}$ ,

$$\begin{aligned} 0 \leq \|u + xv\|^2 &= q(u + xv) = x^2q(v) + 2x\phi(u, v) + q(u) = \\ &= \|v\|^2x^2 + 2\phi(u, v)x + \|u\|^2, \end{aligned}$$

il discriminante del trinomio  $\|v\|^2x^2 + 2\phi(u, v)x + \|u\|^2$  dev'essere  $\leq 0$ :

$$4\phi(u, v)^2 - 4\|u\|^2\|v\|^2 \leq 0.$$

Quindi  $\phi(u, v)^2 \leq \|u\|^2\|v\|^2$  e la tesi segue immediatamente. c.v.d.

**Teorema 3.6. (disuguaglianza triangolare)** *Sia  $V$  uno spazio vettoriale reale dotato di un prodotto scalare  $\phi$ . Allora*

$$\|u + v\| \leq \|u\| + \|v\|, \quad \forall u, v \in V$$

**Dimostrazione:**

$$\begin{aligned} \|u + v\|^2 &= \phi(u + v, u + v) = \phi(u, u) + \phi(v, v) + 2\phi(u, v) \leq \\ &\leq \|u\|^2 + \|v\|^2 + 2|\phi(u, v)| \leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| = (\|u\| + \|v\|)^2 \end{aligned}$$

e la tesi segue immediatamente. c.v.d.

#### 4. FORME HERMITIANE

Sia  $V$  uno spazio vettoriale sul campo  $\mathbb{C}$ .

**Definizione 4.1.** Si dice **forma hermitiana** sullo spazio vettoriale  $V$ , un'applicazione:

$$\phi : V \times V \longrightarrow \mathbb{C}$$

tale che:

1.  $\phi(u, v) = \overline{\phi(v, u)}$ ,  $\forall u, v \in V$  (*hermiticit *)
2.  $\phi(u, \lambda v + \mu w) = \lambda\phi(u, v) + \mu\phi(u, w)$ ,  $\forall u, v, w \in V, \forall \lambda, \mu \in \mathbb{K}$   
(*linearit  nella seconda variabile*)

Combinando le due propriet  si ricava immediatamente che una forma hermitiana   antilineare nella prima variabile:

$$\phi(\lambda v + \mu w, u) = \bar{\lambda}\phi(v, u) + \bar{\mu}\phi(w, u), \quad \forall u, v, w \in V, \forall \lambda, \mu \in \mathbb{C}$$

**Esempio 4.2.** *Il prodotto hermitiano standard in  $\mathbb{C}^n$ :*

$$\phi : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}$$

$$\phi(X, Y) := X^+Y = \sum_{i=1}^n \bar{x}_i y_i, \quad \forall X, Y \in \mathbb{C}^n$$

*rappresenta un primo esempio di forma hermitiana, grazie alle propriet  del prodotto righe per colonne.*

*Osservazione 4.3.* Si osservi come, dalla propriet  di hermiticit , segua

$$\phi(v, v) = \overline{\phi(v, v)} \in \mathbb{R}, \quad \forall v \in V.$$

Per le forme hermitiane vale una terminologia simile a quella usata per le forme simmetriche reali:

**Definizione 4.4.** 1. Sia  $V$  uno spazio vettoriale complesso. Una forma hermitiana  $\phi$  si dice *positiva* se

$$\phi(v, v) \geq 0, \quad \forall v \in V.$$

Si dice *negativa* se

$$\phi(v, v) \leq 0, \quad \forall v \in V.$$

Si dice *definita positiva (o negativa)* se, inoltre,

$$\phi(v, v) = 0 \iff v = 0$$

2. Una forma hermitiana definita positiva si dice anche **prodotto scalare hermitiano**.

3. Sia  $V$  uno spazio vettoriale complesso dotato di un prodotto scalare hermitiano  $\phi$ . La quantità

$$\|v\| := \sqrt{q(v)} = \sqrt{\phi(v, v)}$$

si dice **norma o lunghezza** del vettore  $v \in V$ .

**Definizione 4.5.** L' applicazione:

$$q : V \longrightarrow \mathbb{R}$$

definita da

$$q(v) = \phi(v, v)$$

si dice **forma quadratica associata** a  $\phi$ .

Si osservi che  $q(\lambda v) = \phi(\lambda v, \lambda v) = |\lambda|^2 \phi(v, v)$ ,  $\forall \lambda \in \mathbb{C}$ ,  $\forall v \in V$ .

## 5. MATRICE ASSOCIATA AD UNA FORMA HERMITIANA

**Definizione 5.1.** 1. Data una matrice complessa  $A \in \mathbb{C}^{m,n}$ , la matrice  $A^\dagger := \overline{A^t} \in \mathbb{C}^{n,m}$  si definisce matrice **aggiunta** oppure **hermitiana coniugata** di  $A$ .

2. Una matrice quadrata complessa  $A \in \mathbb{C}^{n,n}$  si dice **autoaggiunta** o **hermitiana** se coincide con la propria aggiunta:  $A = A^\dagger$ .

**Proposizione 5.2.** Sia  $A = (a_{ij}) \in \mathbb{C}^{n,n}$  una matrice quadrata di ordine  $n$  nel campo complesso. Allora l'applicazione

$$\phi_A(X, Y) := X^\dagger AY = \sum_{i,j=1}^n a_{ij} \overline{x_i} y_j$$

rappresenta una forma hermitiana se e solo se  $A$  è hermitiana.

**Dimostrazione:** la condizione è necessaria in quanto  $a_{ij} = \phi_A(e_i, e_j)$  e  $a_{ji} = \phi_A(e_j, e_i)$ , dove  $e_i$  ( $1 \leq i \leq j$ ) indica l'  $i$ -esimo vettore della base canonica.

Sufficienza:  $\phi_A(X, Y) = X^\dagger AY = (X^\dagger AY)^t = Y^t A^t \bar{X} = \overline{Y^\dagger A^\dagger X} = \overline{Y^\dagger AX} = \phi_A(Y, X)$ . c.v.d.

La proposizione precedente mostra come si possa associare ad una matrice autoaggiunta una forma hermitiana in  $\mathbb{C}^n$ . In realtà, come per le forma simmetriche, **ogni** forma hermitiana si può rappresentare così, pur di fissare una base nello spazio vettoriale  $V$ . Infatti, dati i vettori  $u, v \in V$  e detti  $X, Y \in \mathbb{C}^n$  i vettori colonna che ne rappresentano le coordinate rispetto alla base fissata  $\mathcal{B} = \{e_1, \dots, e_n\}$ , si trova:

$$\phi(u, v) = \phi\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{i,j=1}^n \phi(e_i, e_j) \bar{x}_i y_j = X^\dagger AY$$

dove si è posto  $A = (a_{ij}) = (\phi(e_i, e_j))$ .

**Definizione 5.3.** Tale matrice  $A \in \mathbb{C}^{n,n}$  è detta **matrice associata** alla forma  $\phi$ , nella base  $\mathcal{B}$ .

## CAMBIAMENTI DI COORDINATE

### 1. MATRICE DEL CAMBIAMENTO DI COORDINATE

Sia  $V$  uno spazio vettoriale sul campo  $\mathbb{K}$  e sia  $\mathcal{B} = \{e_1, \dots, e_n\}$  una base di  $V$ . Ogni vettore  $v \in V$  si esprimerà in modo unico come combinazione lineare di  $\mathcal{B}$ :

$$v = x_1 e_1 + \dots + x_n e_n = \sum_{i=1}^n x_i e_i, \quad x_i \in \mathbb{K}.$$

**Definizione 1.1.**  $X := (x_1, \dots, x_n)^t \in \mathbb{K}^n$  si dice **n-upla delle coordinate del vettore  $v$  rispetto alla base  $\mathcal{B}$** .

Supponiamo ora di avere fissato due basi,  $\mathcal{B} = \{e_1, \dots, e_n\}$  e  $\mathcal{B}' = \{e'_1, \dots, e'_n\}$ , in  $V$  e ricaviamo la relazione tra le coordinate del generico vettore rispetto alle due basi. Ogni vettore di  $\mathcal{B}$  potrà esprimersi come combinazione lineare di  $\mathcal{B}'$ :

$$\begin{cases} e_1 = m_{11}e'_1 + m_{21}e'_2 + \dots + m_{n1}e'_n \\ \vdots \\ e_n = m_{1n}e'_1 + m_{2n}e'_2 + \dots + m_{nn}e'_n \end{cases}$$

dove  $m_{ij} \in \mathbb{K}$ . Più sinteticamente:

$$e_j = \sum_{i=1}^n m_{ij} e'_i, \quad 1 \leq j \leq n.$$

**Definizione 1.2.** La matrice

$$M = M(\text{id}; \mathcal{B}, \mathcal{B}') = (m_{ij}) \in \mathbb{K}^{n,n}$$

si dice **matrice del cambiamento di coordinate** dalla base  $\mathcal{B}$  alla base  $\mathcal{B}'$ .

**Esercizio 1.3.** *Determinare la matrice del cambiamento di coordinate dalla base canonica alla base  $\mathcal{B} := \{(1, 1), (1, -1)\}$ , in  $\mathbb{R}^2$ .*

*Osservazione 1.4.* Si osservi che  $M$  rappresenta la matrice associata alla funzione identità, una volta fissata la base  $\mathcal{B}$  nel dominio e la base  $\mathcal{B}'$  nel codominio.

Per l'osservazione precedente se  $X \in \mathbb{K}^n$  ( $X' \in \mathbb{K}^n$ ) è l' $n$ -upla delle coordinate di un vettore  $v \in V$  rispetto alla base  $\mathcal{B}$  ( $\mathcal{B}'$ ), si avrà

$$(1) \quad X' = MX, \quad x'_i = \sum_{j=1}^n m_{ij}x_j.$$

La formula 1 può essere verificata direttamente:

$$\begin{aligned} v &= \sum_{i=1}^n x'_i e'_i = \sum_{j=1}^n x_j e_j = \\ &= \sum_{j=1}^n x_j \left( \sum_{i=1}^n m_{ij} e'_i \right) = \sum_{i=1}^n \left( \sum_{j=1}^n m_{ij} x_j \right) e'_i. \end{aligned}$$

## 2. CAMBIAMENTO DI COORDINATE PER UN ENDOMORFISMO

Sia  $f : V \rightarrow V$  un endomorfismo e sia  $\mathcal{B} = \{e_1, \dots, e_n\}$  una base dello spazio vettoriale  $V$ . Ricordiamo la seguente:

**Definizione 2.1.** Posto

$$f(e_j) = \sum_{i=1}^n f_{ij} e_i, \quad 1 \leq j \leq n,$$

la matrice  $F = (f_{ij}) = M(f; \mathcal{B}, \mathcal{B})$  si dice **matrice associata all'endomorfismo  $f$  nella base  $\mathcal{B}$** .

Supponiamo ora di avere fissato due basi,  $\mathcal{B} = \{e_1, \dots, e_n\}$  e  $\mathcal{B}' = \{e'_1, \dots, e'_n\}$ , in  $V$  e ricaviamo la relazione tra  $F = M(f; \mathcal{B}, \mathcal{B})$  e  $F' = M(f; \mathcal{B}', \mathcal{B}')$ :

(2)

$$F' = M(f; \mathcal{B}', \mathcal{B}') = M(id; \mathcal{B}, \mathcal{B}') M(f; \mathcal{B}, \mathcal{B}) M(id; \mathcal{B}', \mathcal{B}) = MFM^{-1}$$

dove  $M$  rappresenta la matrice del cambiamento di coordinate dalla base  $\mathcal{B}$  alla base  $\mathcal{B}'$ . L'espressione 2 può essere verificata direttamente, indicando con  $X$  ( $X'$ ) la  $n$ -upla delle coordinate di  $v \in V$  rispetto alla base  $\mathcal{B}$  ( $\mathcal{B}'$ ) e con  $Y$  ( $Y'$ ) la  $n$ -upla delle coordinate di  $f(v) \in V$  rispetto alla base  $\mathcal{B}$  ( $\mathcal{B}'$ ):

$$Y' = F'X' = F'MX$$

$$Y' = MY = MFX$$

quindi

$$F'MX = MFX, \quad \forall X \in \mathbb{K}^n$$

che è equivalente a

$$F' = MFM^{-1}.$$

**Definizione 2.2.** Due matrici quadrate  $F, F' \in \mathbb{K}^{n,n}$  si dicono **simili** se esiste una matrice invertibile  $M \in \mathbb{K}^{n,n}$  tale che  $F' = MFM^{-1}$ .

### 3. CAMBIAMENTO DI COORDINATE PER UNA FORMA SIMMETRICA O HERMITIANA

Sia  $V$  uno spazio vettoriale sul campo  $\mathbb{K}$ , sia  $\phi$  una forma bilineare simmetrica in  $V$  e siano  $\mathcal{B} = \{e_1, \dots, e_n\}$  e  $\mathcal{B}' = \{e'_1, \dots, e'_n\}$  due basi di  $V$ . Dati due vettori  $v, u \in V$  indichiamo con  $X$  e  $Y$  ( $X'$  e  $Y'$ ) le rispettive coordinate nella base  $\mathcal{B}$  ( $\mathcal{B}'$ ). Indichiamo inoltre con  $A$  ( $A'$ ) la matrice associata a  $\phi$  nella base  $\mathcal{B}$  ( $\mathcal{B}'$ ). Avremo:

$$\phi(v, u) = X^t AY = X'^t A' Y' = (MX)^t A' MY = X^t M^t A' MY,$$

dove  $M$  rappresenta la matrice del cambiamento di coordinate dalla base  $\mathcal{B}$  alla base  $\mathcal{B}'$ . Troviamo quindi:

$$X^t AY = X^t M^t A' MY, \quad \forall X, Y \in \mathbb{K}^n,$$

che è equivalente a

$$(3) \quad A = M^t A' M, \quad \text{oppure a} \quad A' = P^t A P \quad (P := M^{-1}).$$

**Definizione 3.1.** Due matrici quadrate  $A, A' \in \mathbb{K}^{n,n}$  si dicono **congruenti** se esiste una matrice invertibile  $P \in \mathbb{K}^{n,n}$ , tale che  $A' = P^t A P$ .

**Esercizio 3.2.** Verificare la 3 per la forma bilineare simmetrica in  $\mathbb{R}^2$ :

$$\phi((x_1, x_2), (y_1, y_2)) := x_1 y_1 + x_2 y_1 + x_1 y_2,$$

nel passaggio dalla base canonica alla base  $\mathcal{B} := \{(1, 1), (1, -1)\}$ .

Analogamente, sia  $V$  uno spazio vettoriale complesso, sia  $\phi$  una forma hermitiana in  $V$ . Nelle stesse notazioni, avremo:

$$\phi(v, u) = X^+ AY = X'^+ A' Y' = (MX)^+ A' MY = X^+ M^+ A' MY,$$

quindi

$$X^+ AY = X^+ M^+ A' MY, \quad \forall X, Y \in \mathbb{C}^n,$$

che è equivalente a

$$A = M^+ A' M \quad \text{oppure a} \quad A' = P^+ A P \quad (P := M^{-1}).$$

**Definizione 3.3.** Due matrici quadrate  $A, A' \in \mathbb{C}^{n,n}$  si dicono **congruenti in senso hermitiano** se esiste una matrice invertibile  $P \in \mathbb{C}^{n,n}$ , tale che  $A' = P^+AP$ .

## DIAGONALIZZAZIONE DI FORME BILINEARI SIMMETRICHE

### 1. DIAGONALIZZAZIONE DI UNA FORMA BILINEARE SIMMETRICA IN UN CAMPO DI CARATTERISTICA DISPARI

Sia  $V$  uno spazio vettoriale sul campo  $\mathbb{K}$ ,  $\phi$  una forma bilineare simmetrica su  $V$  e  $q$  la forma quadratica associata.

**Definizione 1.1.** Due vettori  $v, u \in V$  si dicono **ortogonali** rispetto a  $\phi$ , se  $\phi(v, u) = 0$ .

Ci proponiamo di stabilire se esiste una base  $\mathcal{B} = \{e_1, \dots, e_n\}$  rispetto alla quale la matrice associata a  $\phi$  risulti essere diagonale.

*Osservazione 1.2.* 1. Poiché la matrice associata a  $\phi$ , nella base  $\mathcal{B}$ , è  $A = (a_{ij}) = (\phi(e_i, e_j))$ , tale matrice sarà diagonale **se e solo se** la base  $\mathcal{B}$  è **ortogonale**.

2. L'espressione della forma quadratica in una base diagonale risulta essere:

$$q = \sum_{i=1}^n a_{ii} x_i^2.$$

**Definizione 1.3.** Un vettore  $v \in V$  si dice **isotropo** per  $\phi$ , se  $q(v) = \phi(v, v) = 0$ .

In altri termini un vettore isotropo è **ortogonale a sé stesso**, rispetto alla forma  $\phi$ .

**Esempio 1.4.** 1. Il vettore  $(1, 1) \in \mathbb{R}^2$  è isotropo per la forma

*bilineare simmetrica:*  $\phi((x_1, x_2), (y_1, y_2)) = x_1 y_1 - x_2 y_2$ .

2. Il vettore  $(1, 1) \in \mathbb{Z}_2^2$  è isotropo per la forma bilineare simmetrica:

$\phi((x_1, x_2), (y_1, y_2)) = x_1 y_1 + x_2 y_2$ .

3. Tutti i vettori di  $\mathbb{Z}_2^2$  sono isotropi per la forma bilineare sim-

*metrica:*  $\phi((x_1, x_2), (y_1, y_2)) = x_1 y_2 + x_2 y_1$ .

**Definizione 1.5.** Indicheremo con  $v^\perp$  l'insieme dei vettori ortogonali a  $v$ :

$$v^\perp := \{u \in V \mid \phi(v, u) = 0\}.$$

Si verifica facilmente che  $v^\perp$  è un sottospazio vettoriale di  $V$  di dimensione  $n$  (in tal caso  $v$  è isotropo) oppure  $n - 1$  (esercizio).

**Lemma 1.6.** *Se  $v \in V$  **non** è isotropo (cioè  $q(v) \neq 0$ ) allora  $V = L(v) \oplus v^\perp$ .*

**Dimostrazione:** è sufficiente dimostrare che  $L(v) \cap v^\perp = \{0\}$  (perché?).

Sia  $u \in L(v) \cap v^\perp$ . Abbiamo quindi:

1.  $u \in L(v) \Rightarrow u = \lambda v, \lambda \in \mathbb{K}$ .
2.  $u \in v^\perp \Rightarrow \phi(u, v) = 0$ .

Combinando si trova:

$$0 = \phi(u, v) = \phi(\lambda v, v) = \lambda \phi(v, v).$$

Poiché  $v$  non è isotropo si ha necessariamente  $\lambda = 0$ , quindi  $u = \lambda v = 0$ .  
c.v.d.

**Teorema 1.7. (di diagonalizzazione)** *Sia  $\phi$  una forma bilineare simmetrica definita in uno spazio vettoriale  $V$  sul campo  $\mathbb{K}$ . Se  $\text{char}\mathbb{K} \neq 2$  allora  $\phi$  ammette una base ortogonale.*

**Dimostrazione:** se  $\phi$  è la forma nulla non c'è nulla da dimostrare. Supponiamo quindi  $\exists u, v \in V$  tali che  $\phi(u, v) \neq 0$  e dimostriamo il teorema per induzione sulla dimensione,  $n$ , dello spazio vettoriale. Se  $\dim V = 1$  la tesi è ovvia. Dimostriamo il passo induttivo. Per il teorema di polarizzazione:

$$0 \neq \phi(u, v) = \frac{1}{2}(q(u+v) - q(u) - q(v)),$$

almeno uno dei tre vettori  $u, v, u+v$  (ad esempio  $v$ ) non è isotropo. Il lemma precedente ci consente di applicare l'ipotesi induttiva a  $v^\perp$  in quanto  $\dim v^\perp = n-1$ . Sia dunque  $\mathcal{B}_{v^\perp} = \{v_1, \dots, v_{n-1}\}$  una base ortogonale di  $v^\perp$ . Un secondo richiamo al lemma precedente ci consente di concludere che  $\mathcal{B} = \{v, v_1, \dots, v_{n-1}\}$  rappresenta una base ortogonale di  $V$ . c.v.d.

**Corollario 1.8.** *Ogni matrice simmetrica con coefficienti in un campo  $\mathbb{K}$ , tale che  $\text{char}\mathbb{K} \neq 2$ , è **congruente** ad una matrice diagonale.*

*Osservazione 1.9.* Poiché, nel campo binario, la forma quadratica associata alla matrice  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  è nulla e poichè tale matrice **non** può essere congruente alla matrice nulla (perché?), il teorema precedente è falso nel campo binario.

2. DIAGONALIZZAZIONE DI UNA FORMA BILINEARE SIMMETRICA  
NEL CAMPO COMPLESSO

**Teorema 2.1.** *Sia  $\phi$  una forma bilineare simmetrica definita in uno spazio vettoriale complesso  $V$ . Allora  $\phi$  ammette una base rispetto alla quale la matrice associata è del tipo:*

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

dove  $I_r$  indica la matrice identità,  $0$  matrici nulle di tipo opportuno e  $r$  il rango.

**Dimostrazione:** in base al teorema di diagonalizzazione dovrà esistere una base ortogonale  $\mathcal{B} = \{v_1, \dots, v_n\}$ . Non è restrittivo assumere:

1.  $q(v_i) \neq 0$ ,  $1 \leq i \leq r$
2.  $q(v_i) = 0$ ,  $r + 1 \leq i \leq n$ .

Allora la base  $\mathcal{B}' = \{e_1, \dots, e_n\}$  definita da

1.  $e_i = \frac{v_i}{\sqrt{q(v_i)}}$ ,  $1 \leq i \leq r$
2.  $e_i = v_i$ ,  $r + 1 \leq i \leq n$

soddisfa la tesi del teorema. c.v.d.

3. DIAGONALIZZAZIONE DI UNA FORMA BILINEARE SIMMETRICA  
NEL CAMPO REALE

**Teorema 3.1.** *(di Sylvester) Sia  $\phi$  una forma bilineare simmetrica definita in uno spazio vettoriale reale  $V$ . Allora  $\phi$  ammette una base rispetto alla quale la matrice associata è del tipo:*

$$\begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

dove  $p \leq r$  è indipendente dalla base scelta (le notazioni sono le stesse del teorema precedente).

**Dimostrazione:** in base al teorema di diagonalizzazione dovrà esistere una base ortogonale  $\mathcal{B} = \{v_1, \dots, v_n\}$ . Non è restrittivo assumere:

1.  $q(v_i) > 0$ ,  $1 \leq i \leq p$
2.  $q(v_i) < 0$ ,  $p + 1 \leq i \leq r$
3.  $q(v_i) = 0$ ,  $r + 1 \leq i \leq n$ .

Allora la base  $\mathcal{B}' = \{e_1, \dots, e_n\}$  definita da

1.  $e_i = \frac{v_i}{\sqrt{q(v_i)}}$ ,  $1 \leq i \leq p$
2.  $e_i = \frac{v_i}{\sqrt{-q(v_i)}}$ ,  $p+1 \leq i \leq r$
3.  $e_i = v_i$ ,  $r+1 \leq i \leq n$

soddisfa la tesi del teorema.

Resta da dimostrare che l'indice  $p$  non dipende dalla base scelta. Siano  $\mathcal{B}' = \{e_1, \dots, e_n\}$  e  $\mathcal{B}'' = \{f_1, \dots, f_n\}$  due basi del tipo appena introdotto e assumiamo che la forma quadratica del generico vettore  $v \in V$  si esprima nelle forme:

$$q(v) = \sum_{i=1}^p x_i^2 - \sum_{i=p+1}^r x_i^2 = \sum_{j=1}^q y_j^2 - \sum_{j=q+1}^r y_j^2 \quad (*)$$

dove le  $x_i$  ( $y_j$ ) rappresentano le coordinate del vettore  $v$  rispetto alla base  $\mathcal{B}'$  ( $\mathcal{B}''$ ). La tesi consiste nel provare  $p = q$ . Ragioniamo per assurdo e assumiamo  $p > q$ . Poniamo  $G := L(\{e_1, \dots, e_p\})$  e  $U := L(\{f_{q+1}, \dots, f_n\})$ . Si osservi che, in base alla (\*), la forma quadratica è definita positiva in  $G$  e negativa in  $U$ , quindi l'intersezione  $G \cap U$  può contenere solo il vettore nullo. Dal teorema di Grassmann si ricava:

$$\begin{aligned} \dim G \cap U &= \dim G + \dim U - \dim(G + U) = p + (n - q) - \dim(G + U) \geq \\ &\geq n + 1 - \dim(G + U) \geq 1 \end{aligned}$$

e si giunge ad un assurdo. c.v.d.

**Definizione 3.2.** 1. L'intero  $p$  del teorema precedente si dice **indice** di  $\phi$ .

2. La differenza  $p - (r - p) = 2p - r$  si dice **segnatura** di  $\phi$ .

#### 4. ALGORITMO DI GAUSS

Vogliamo stabilire un procedimento pratico, detto algoritmo di Gauss, per diagonalizzare una forma quadratica reale  $q = \sum_{i=1}^n a_{ii}x_i^2 + 2 \sum_{i < j} a_{ij}x_i x_j$ . Cerchiamo quindi di determinare un cambiamento di coordinate

$$y_i = L_i(x_1, \dots, x_n), \quad 1 \leq i \leq n,$$

affinché si abbia:

$$q = \sum_{i=1}^n a_{ii}x_i^2 + 2 \sum_{i < j} a_{ij}x_i x_j = \sum_{i=1}^n \lambda_i y_i^2.$$

Procediamo per induzione sulla dimensione e, poiché per  $n = 1$  non c'è nulla da dimostrare, occupiamoci direttamente del passo induttivo.

Distinguiamo due casi:

1.  $\exists i \mid a_{ii} \neq 0$ ;
2.  $a_{ii} = 0, \forall i$ .

1. Non è restrittivo assumere  $a_{11} \neq 0$ . Avremo dunque:

$$q = a_{11}x_1^2 + 2Rx_1 + S,$$

dove si è posto:

$$R = \sum_{j=2}^n a_{1j}x_j, \quad S = \sum_{i,j=2}^n a_{ij}x_ix_j.$$

Potremo quindi scrivere  $q$  nella forma:

$$q = a_{11}\left(x_1 + \frac{R}{a_{11}}\right)^2 - \left(\frac{R^2}{a_{11}}\right) + S.$$

Poiché  $-\left(\frac{R^2}{a_{11}}\right) + S$  è una forma quadratica nelle  $x_2, \dots, x_n$ , per ipotesi induttiva avremo:

$$-\left(\frac{R^2}{a_{11}}\right) + S = \sum_{i=2}^n \lambda_i y_i^2,$$

dove  $y_i = L_i(x_2, \dots, x_n)$ ,  $2 \leq i \leq n$ , rappresenta un opportuno cambiamento di coordinate. E' facile dimostrare (esercizio) che le forme lineari:

$$L_1(x_1, \dots, x_n) = x_1 + \frac{R}{a_{11}}, \quad L_2, \dots, L_n$$

sono linearmente indipendenti.

In conclusione:

$$q = a_{11}y_1^2 + \sum_{i=2}^n \lambda_i y_i^2, \quad \text{dove } y_i = L_i(x_1, \dots, x_n), \quad \forall i.$$

2. Non è restrittivo assumere  $a_{12} \neq 0$ . Avremo dunque:

$$q = 2a_{12}x_1x_2 + 2Rx_1 + 2Tx_2 + S,$$

dove si è posto:

$$R = \sum_{j=3}^n a_{1j}x_j, \quad T = \sum_{j=3}^n a_{2j}x_j, \quad S = \sum_{i,j=3}^n a_{ij}x_ix_j.$$

Potremo quindi scrivere  $q$  nella forma:

$$q = 2a_{12}\left(x_1 + \frac{T}{a_{12}}\right)\left(x_2 + \frac{R}{a_{12}}\right) - \left(\frac{2RT}{a_{12}}\right) + S = \alpha LL' + S' =$$

$$= \frac{\alpha}{4} \{(L + L')^2 - (L - L')^2\} + S',$$

dove si è posto  $\alpha = 2a_{12}$ ,  $L = (x_1 + \frac{T}{a_{12}})$ ,  $L' = (x_2 + \frac{R}{a_{12}})$  e  $S' = -(\frac{2RT}{a_{12}}) + S$ . Poiché  $S'$  è una forma quadratica nelle  $x_3, \dots, x_n$ , per ipotesi induttiva avremo:

$$S' = \sum_{i=3}^n \lambda_i y_i^2, \quad y_i = L_i(x_3, \dots, x_n), \quad 3 \leq i \leq n.$$

E' facile dimostrare (esercizio) che le forme lineari:

$$L_1 = L + L', \quad L_2 = L - L', \quad L_3, \dots, L_n,$$

sono linearmente indipendenti.

In conclusione:

$$q = \frac{\alpha}{4}(y_1^2 - y_2^2) + \sum_{i=3}^n \lambda_i y_i^2, \quad \text{dove } y_i = L_i(x_1, \dots, x_n), \quad \forall i.$$

**Esercizio 4.1.** *Diagonalizzare ciascuna delle seguenti forma quadratiche, determinandone indice e segnatura:*

1. in  $\mathbb{R}^2$ ,  $q(x, y) = x^2 + 6xy + 2y^2$ ;
2. in  $\mathbb{R}^2$ ,  $q(x, y) = 2xy - 2y^2$ ;
3. in  $\mathbb{R}^2$ ,  $q(x, y) = xy$ ;
4. in  $\mathbb{R}^3$ ,  $q(x, y, z) = x^2 + y^2 + z^2 - 4(xy + xz + yz)$ ;
5. in  $\mathbb{R}^3$ ,  $q(x, y, z) = x^2 + y^2 + z^2 - (xy + xz + yz)$ ;
6. in  $\mathbb{R}^3$ ,  $q(x, y, z) = x^2 + 6y^2 - 4xy + 8xz$ ;
7. in  $\mathbb{R}^3$ ,  $q(x, y, z) = 2(xy + xz + yz)$ ;
8. in  $\mathbb{R}^3$ ,  $q(x, y, z) = x^2 + 4y^2 - 4xy + 4yz$ .

# ORTOGONALITÀ

## 1. SOTTOSPAZIO ORTOGONALE

Sia  $V$  uno spazio vettoriale,  $\phi$  una forma bilineare simmetrica oppure una forma hermitiana su  $V$  e  $q$  la forma quadratica associata.

**Definizione 1.1.** Sia  $\emptyset \neq U \subset V$  un sottoinsieme. Si dice **ortogonale** di  $U$  il sottoinsieme:

$$U^\perp := \{u \in V \mid \phi(v, u) = 0, \forall v \in U\}.$$

**Lemma 1.2. (proprietà dell'ortogonale)**

1.  $U^\perp$  è un sottospazio vettoriale;
2.  $U \subset W \Rightarrow U^\perp \supset W^\perp$ ;
3.  $U \subset (U^\perp)^\perp$ ;
4. se  $U$  è un sottospazio e  $\mathcal{B}$  una base di  $U$  allora  $U^\perp = \mathcal{B}^\perp$ .

**Dimostrazione:** 1. Se  $u, u' \in U^\perp$  allora  $\phi(v, u) = \phi(v, u') = 0$ ,  $\forall v \in U$ . Quindi

$$\phi(v, \lambda u + \mu u') = \lambda \phi(v, u) + \mu \phi(v, u') = 0, \forall v \in U, \forall \lambda, \mu \in \mathbb{K}$$

cioè  $\lambda u + \mu u' \in U^\perp$ ,  $\forall \lambda, \mu \in \mathbb{K}$ .

2.

$$w \in W^\perp \Rightarrow \phi(v, w) = 0, \forall v \in W \Rightarrow \phi(v, w) = 0, \forall v \in U \Rightarrow w \in U^\perp.$$

3. Dalla definizione di  $U^\perp$  si ricava

$$\phi(v, u) = \phi(u, v) = 0, \forall v \in U, \forall u \in U^\perp,$$

quindi  $v \in (U^\perp)^\perp$ ,  $\forall v \in U$ .

4. L'inclusione  $U^\perp \subset \mathcal{B}^\perp$  segue dal punto 2..

Sia  $\mathcal{B} = \{e_1, \dots, e_k\}$ , sia  $u \in \mathcal{B}^\perp$  (quindi  $\phi(e_i, u) = 0$ ,  $1 \leq i \leq k$ ) e sia  $v \in U$ . La tesi consiste nel mostrare  $\phi(v, u) = 0$ . Poiché  $v \in U$  e  $\mathcal{B}$  è una base di  $U$  si avrà  $v = \sum_1^k \lambda_i e_i$ , per opportuni scalari  $\lambda_i$ . Ma allora

- (caso simmetrico)  $\phi(v, u) = \phi(\sum_1^k \lambda_i e_i, u) = \sum_1^k \lambda_i \phi(e_i, u) = 0$ ;
  - (caso hermitiano)  $\phi(v, u) = \phi(\sum_1^k \lambda_i e_i, u) = \sum_1^k \overline{\lambda_i} \phi(e_i, u) = 0$ .
- c.v.d.

**Definizione 1.3.** Il sottospazio

$$V^\perp = \{u \in V \mid \phi(v, u) = 0, \forall v \in V\}$$

si dice anche **nucleo** della forma  $\phi$  e si indica con  $\ker(\phi)$ .

Fissata una base nello spazio vettoriale e detta  $A$  la matrice associata a  $\phi$  in tale base, si vede immediatamente che il nucleo di  $\phi$  coincide con il nucleo della matrice  $A$ .

**Teorema 1.4. (regola del doppio ortogonale)** *Le seguenti condizioni sono tra loro equivalenti:*

1.  $U = (U^\perp)^\perp$  per ogni sottospazio  $U \subset V$ ;
2.  $\dim(U^\perp) = \dim V - \dim U$  per ogni sottospazio  $U \subset V$ ;
3.  $\ker(\phi) = \{0\}$  (oppure  $\det(A) \neq 0$ ).

**Dimostrazione:** 1.  $\Rightarrow$  3.. Ragioniamo per assurdo e poniamo  $\ker(\phi) = W \neq \{0\}$ . Allora avremo

$$(\{0\}^\perp)^\perp = V^\perp = \ker(\phi) = W \neq \{0\}$$

che contraddice l'ipotesi 1..

3.  $\Rightarrow$  2.. Assumiamo  $\phi$  simmetrica poichè nel caso hermitiano la dimostrazione è del tutto analoga. Fissiamo una base  $\mathcal{B} = \{e_1, \dots, e_n\}$  di  $V$ , una base  $\mathcal{B}' = \{f_1, \dots, f_k\}$  di  $U$  e indichiamo con  $X_1, \dots, X_k \in \mathbb{K}^n$  le colonne che rappresentano le coordinate dei vettori  $f_1, \dots, f_k$  nella base  $\mathcal{B}$ . Per il punto 4. del lemma precedente il sottospazio  $U^\perp$  è rappresentato, nella base  $\mathcal{B}$ , dell'insieme delle soluzioni del seguente sistema lineare omogeneo:

$$\begin{cases} X_1^t A Y = 0 \\ \vdots \\ X_k^t A Y = 0 \end{cases}$$

Poichè i vettori  $f_1, \dots, f_k$  sono indipendenti e poichè  $\det(A) \neq 0$ , il sistema precedente ha rango  $k$  e la tesi segue dal teorema di Rouché-Capelli.

2.  $\Rightarrow$  1.. Applicando due volte l'ipotesi si trova:

$$\dim(U^\perp)^\perp = \dim V - \dim U^\perp = \dim V - (\dim V - \dim U) = \dim U$$

e si conclude ricordando il punto 3. del lemma precedente. c.v.d.

Ricordiamo la seguente

**Definizione 1.5.** Un vettore  $v \in V$  si dice **isotropo** per  $\phi$ , se  $q(v) = \phi(v, v) = 0$ .

**Esempio 1.6.** 1. Ogni vettore appartenente al nucleo di  $\phi$  è isotropo;  
2. se  $u \in U \cap U^\perp$  allora  $u$  è isotropo.

**Teorema 1.7. (di decomposizione ortogonale)** Le seguenti condizioni sono tra loro equivalenti:

1.  $V = U \oplus U^\perp$  per ogni sottospazio  $U \subset V$ ;
2. non esistono in  $V$  vettori isotropi non nulli:  $q(v) = 0 \Rightarrow v = 0$ .

**Dimostrazione:** la condizione è necessaria in quanto  $q(v) = 0 \Rightarrow v \in L(v) \cap v^\perp$ .

Sufficienza: per il punto 2. dell'esempio precedente l'ipotesi implica  $U \cap U^\perp = \{0\}$ . Per il teorema di Grassmann basta quindi dimostrare che  $\dim U + \dim U^\perp = \dim V$ . Per il punto 1. dell'esempio precedente l'ipotesi implica  $\ker \phi = \{0\}$ , quindi siamo nelle ipotesi della regola del doppio ortogonale, che implica  $\dim U + \dim U^\perp = \dim V$ . c.v.d.

*Osservazione 1.8.* Per il teorema di Sylvester, se  $V$  è reale, l'ipotesi del teorema precedente ( $q(v) = 0 \Rightarrow v = 0$ ) è verificata soltanto per una forma definita positiva oppure definita negativa.

## 2. PROIEZIONI ORTOGONALI

D'ora in poi, in questo capitolo, supporremo  $V$  reale e  $\phi$  definita positiva. Sia  $U \subset V$  un sottospazio vettoriale. Il teorema di decomposizione ortogonale ci assicura che

$$\forall v \in V, \exists ! u \in U, \exists ! u^\perp \in U^\perp \mid v = u + u^\perp$$

**Esercizio 2.1.** Mostrare esistenza e l'unicità di tali vettori  $u \in U$  e  $u^\perp \in U^\perp$ .

**Definizione 2.2.** Il vettore  $u \in U$  si dice **proiezione ortogonale del vettore  $v$  nel sottospazio  $U$**  e si indica col simbolo  $\pi_U(v)$ .

La proiezione ortogonale  $\pi_U(v)$  è quindi caratterizzata dalle seguenti condizioni:

- $\pi_U(v) \in U$ ,
- $v - \pi_U(v) \in U^\perp$ ,

che si traducono in un sistema lineare di Cramer. Sia infatti  $\mathcal{B}_U = \{e_1, \dots, e_k\}$  una base del sottospazio  $U$ . La prima condizione dice che la proiezione cercata è combinazione lineare di  $\mathcal{B}_U$ :

$$\exists \lambda_1, \dots, \lambda_k \in \mathbb{R} \mid \pi_U(v) = \sum_{i=1}^k \lambda_i e_i.$$

La seconda condizione si esprime imponendo che il vettore  $v - \pi_U(v) = v - \sum_{i=1}^k \lambda_i e_i$  sia ortogonale alla base  $\mathcal{B}_U$  e ciò si traduce, appunto, nel seguente sistema lineare di Cramer (nelle incognite  $\lambda_i$ ):

$$\begin{cases} \phi(v - \pi_U(v), e_1) = 0 \\ \vdots \\ \phi(v - \pi_U(v), e_k) = 0 \end{cases}$$

cioè

$$\begin{cases} \sum_{i=1}^k \lambda_i \phi(e_i, e_1) = \phi(v, e_1) \\ \vdots \\ \sum_{i=1}^k \lambda_i \phi(e_i, e_k) = \phi(v, e_k) \end{cases}$$

**Esercizio 2.3.** *Determinare la proiezione ortogonale del vettore  $v = (1, 1, 1)$  nel sottospazio  $U = L((1, 0, 0), (0, 1, 0))$ , sia rispetto al prodotto scalare standard che rispetto al prodotto scalare:*

$$q(x, y, z) = x^2 + 2xy + 2y^2 + 4yz + 5z^2.$$

### 3. BASI ORTONORMALI E ALGORITMO DI GRAM-SCHMIDT.

**Definizione 3.1.** Una base  $\mathcal{B} = \{e_1, \dots, e_n\} \subset V$  si dice **ortonormale** per la forma  $\phi$  se

$$\phi(e_i, e_j) = \delta_{ij} \text{ (simbolo di Kronecker).}$$

Equivalentemente, la condizione di ortonormalità richiede che la matrice associata a  $\phi$  nella base  $\mathcal{B}$  sia la matrice identità ossia che, in coordinate rispetto a tale base, la forma si identifichi col prodotto scalare standard.

**Esercizio 3.2.** *Dimostrare che un prodotto scalare, anche hermitiano, ammette sempre una base ortonormale.*

E' particolarmente semplice calcolare le componenti di un vettore  $v \in V$  rispetto ad una base ortonormale:

$$\phi(e_i, v) = \phi(e_i, \sum_{k=1}^n \lambda_k e_k) = \sum_{k=1}^n \lambda_k \phi(e_i, e_k) = \lambda_i$$

quindi

$$\forall v \in V, \quad v = \sum_{i=1}^n \phi(e_i, v) e_i.$$

*Osservazione 3.3.* Sia  $U \subset V$  un sottospazio, sia  $\mathcal{B}' = \{e_1, \dots, e_v\}$  una base ortonormale di  $U$  e sia  $\mathcal{B} = \{e_1, \dots, e_n\}$  una base ortonormale di  $V$  contenente  $\mathcal{B}'$ . Il vettore  $\sum_{i=1}^k \phi(e_i, v) e_i$  soddisfa le due condizioni che caratterizzano la proiezione  $\pi_U(v)$ :

$$\pi_U(v) = \sum_{i=1}^k \phi(e_i, v) e_i.$$

**Esercizio 3.4.** Siano  $v, u \in V$  due vettori e supponiamo  $u \neq 0$ . Mostrare che:

$$\pi_{L(u)}(v) = \frac{\phi(v, u)}{\|u\|^2} u.$$

**Definizione 3.5.** Il vettore  $\frac{\phi(v, u)}{\|u\|^2} u =: \pi_u(v)$  si dice anche **proiezione** di  $v$  lungo  $u$ .

*Osservazione 3.6.* Si osservi che valgono le due condizioni:

- $\pi_u(v) \in L(u)$ ,
- $v - \pi_u(v) \perp u$ .

Sia  $\mathcal{B} = \{e_1, \dots, e_n\} \subset V$  una base qualunque. Applicando ripetutamente l'osservazione precedente, vede che i vettori

$$\begin{cases} v_1 = e_1 \\ v_2 = e_2 - \pi_{v_1}(e_2) \\ \vdots \\ v_n = e_n - \sum_{i=1}^{n-1} \pi_{v_i}(e_n) \end{cases}$$

sono tra loro ortogonali e, una volta normalizzati, danno luogo ad una base ortonormale. Tale procedimento di ortonormalizzazione va sotto il nome di **algoritmo di Gram-Schmidt**.

**Esercizio 3.7.** Si consideri  $\mathbb{R}^3$  dotato del prodotto scalare:

$$q(x, y, z) = x^2 + 2xy + 2y^2 + 4yz + 5z^2.$$

1. Usando l'algoritmo di Gram-Schmidt, determinare la proiezione ortogonale del vettore  $v = (1, 1, 1)$  nel sottospazio  $U = L((1, 0, 0), (0, 1, 0))$ .
2. Ortonormalizzare la base canonica usando l'algoritmo di Gram-Schmidt.

## ENDOMORFISMI SPECIALI

### 1. ISOMETRIE NEGLI SPAZI EUCLIDEI REALI

Sia  $V$  uno spazio vettoriale reale, dotato di un prodotto scalare  $\phi$  e sia  $q$  la forma quadratica associata.

**Definizione 1.1.** Un'applicazione lineare  $f : V \rightarrow V$  si dice **isometria** se conserva la norma di ciascun vettore:

$$(1) \quad \|v\| = \|f(v)\|, \quad \forall v \in V.$$

**Definizione 1.2.** Un'applicazione lineare  $f : V \rightarrow V$  si dice **ortogonale** se conserva tutti i prodotti scalari:

$$(2) \quad \phi(v, u) = \phi(f(v), f(u)), \quad \forall v, u \in V.$$

**Proposizione 1.3.** *Le due definizioni precedenti sono equivalenti.*

**Dimostrazione:** (1)  $\Rightarrow$  (2). Segue dall'identità di polarizzazione:

$$\begin{aligned} \phi(v, u) &= \frac{1}{2}(q(v+u) - q(v) - q(u)) = \frac{1}{2}(\|v+u\|^2 - \|v\|^2 - \|u\|^2) = \\ &= \frac{1}{2}(\|f(v+u)\|^2 - \|f(v)\|^2 - \|f(u)\|^2) = \frac{1}{2}(\|f(v)+f(u)\|^2 - \|f(v)\|^2 - \|f(u)\|^2) \\ &= \frac{1}{2}(q(f(v)+f(u)) - q(f(v)) - q(f(u))) = \phi(f(v), f(u)). \end{aligned}$$

$$(2) \Rightarrow (1).$$

$$\|v\| = \sqrt{\phi(v, v)} = \sqrt{\phi(f(v), f(v))} = \|f(v)\|.$$

c.v.d.

*Osservazione 1.4.* In base alla proposizione precedente se un endomorfismo conserva le lunghezze allora conserva anche gli angoli.

## 2. ISOMETRIE IN FORMA MATRICIALE

Vogliamo tradurre la definizione 1.2 in forma matriciale. Sia dunque  $\mathcal{B} = \{e_1, \dots, e_n\}$  una base di  $V$  e  $A$  la matrice associata a  $\phi$ , nella base  $\mathcal{B}$ . Fissati due vettori generici  $v, u \in V$  e indichiamo con  $X, Y \in \mathbb{R}^n$  le rispettive colonne di coordinate rispetto alla base  $\mathcal{B}$ , la (2) si traduce nella seguente condizione matriciale:

$$X^t AY = (FX)^t A(FY) = X^t F^t AFY, \quad \forall X, Y \in \mathbb{R}^n,$$

ossia, data l'arbitrarietà di  $X, Y \in \mathbb{R}^n$ , nella seguente equazione matriciale:

$$(3) \quad A = F^t AF.$$

In particolare, se  $\mathcal{B}$  è ortonormale (condizione equivalente a  $A = I_n$ ), la condizione precedente si traduce nella più semplice:

$$(4) \quad F^{-1} = F^t.$$

Poniamo dunque la seguente

**Definizione 2.1.** Una matrice  $F \in \mathbb{R}^{n,n}$  tale che  $F^{-1} = F^t$  si dice **ortogonale**.

Si può concludere:

**Teorema 2.2.** *Un endomorfismo è ortogonale se e solo se la sua matrice associata, **rispetto ad una base ortonormale**, è ortogonale.*

*Osservazione 2.3.* La definizione 2.1 esprime il fatto che le colonne della matrice sono ortonormali rispetto al prodotto scalare ordinario in  $\mathbb{R}^n$ .

## 3. ENDOMORFISMI SIMMETRICI

Sia  $V$  uno spazio vettoriale reale, dotato di un prodotto scalare  $\phi$ .

**Definizione 3.1.** Un endomorfismo  $f : V \rightarrow V$  si dice **simmetrico** se vale la seguente condizione:

$$(5) \quad \phi(f(v), u) = \phi(v, f(u)), \quad \forall v, u \in V.$$

Tale condizione, mantenendo le notazioni usuali, si traduce nella seguente condizione matriciale:

$$(FX)^t AY = X^t A(FY), \quad \forall X, Y \in \mathbb{R}^n,$$

ossia, data l'arbitrarietà di  $X, Y \in \mathbb{R}^n$ , nella seguente equazione matriciale:

$$(6) \quad F^t A = AF.$$

In particolare, se  $\mathcal{B}$  è ortonormale (condizione equivalente a  $A = I_n$ ), la condizione precedente si traduce nella più semplice:

$$(7) \quad F^t = F.$$

Si può quindi concludere:

**Teorema 3.2.** *Un endomorfismo è simmetrico se e solo se la sua matrice associata, **rispetto ad una base ortonormale**, è simmetrica.*

#### 4. ENDOMORFISMI UNITARI ED HERMITIANI

In questo paragrafo ci proponiamo di generalizzare i contenuti dei precedenti agli spazi vettoriali complessi. Sia  $V$  uno spazio vettoriale complesso, dotato di un prodotto scalare hermitiano  $\phi$ .

**Definizione 4.1.** Un endomorfismo  $f : V \rightarrow V$  si dice **unitario** se conserva tutti i prodotti scalari hermitiani:

$$(8) \quad \phi(v, u) = \phi(f(v), f(u)), \quad \forall v, u \in V.$$

Tale condizione, mantenendo le notazioni usuali, si si traduce nella seguente condizione matriciale:

$$X^+ A Y = (F X)^+ A (F Y) = X^+ F^+ A F Y, \quad \forall X, Y \in \mathbb{C}^n,$$

ossia, data l'arbitrarietà di  $X, Y \in \mathbb{C}^n$ , nella seguente equazione matriciale:

$$(9) \quad A = F^+ A F.$$

In particolare, se  $\mathcal{B}$  è ortonormale, la condizione precedente si traduce nella:

$$(10) \quad F^{-1} = F^+.$$

Poniamo dunque la seguente

**Definizione 4.2.** Una matrice  $F \in \mathbb{C}^{n,n}$  tale che  $F^{-1} = F^+$  si dice **unitaria**.

*Osservazione 4.3.* La definizione 4.2 esprime il fatto che le colonne della matrice sono ortonormali rispetto al prodotto hermitiano standard in  $\mathbb{C}^n$ .

**Definizione 4.4.** Un endomorfismo  $f : V \rightarrow V$  si dice **hermitiano** se vale la seguente condizione:

$$(11) \quad \phi(f(v), u) = \phi(v, f(u)), \quad \forall v, u \in V.$$

Tale condizione, mantenendo le notazioni usuali, si traduce nella seguente condizione matriciale:

$$(FX)^+AY = X^+A(FY), \quad \forall X, Y \in \mathbb{C}^n,$$

ossia, data l'arbitrarietà di  $X, Y \in \mathbb{C}^n$ , nella seguente equazione matriciale:

$$(12) \quad F^+A = AF.$$

In particolare, se  $\mathcal{B}$  è ortonormale, la condizione precedente si traduce nella:

$$(13) \quad F^+ = F.$$

Le formule 10 e 13 permettono di concludere:

**Teorema 4.5.** *Un endomorfismo è unitario (hermitiano) se e solo se la sua matrice associata, **rispetto ad una base ortonormale**, è unitaria (autoaggiunta).*

## 5. DIAGONALIZZAZIONE ORTOGONALE

**Teorema 5.1. (teorema spettrale)** *Sia  $V$  uno spazio vettoriale complesso,  $\phi$  un prodotto scalare hermitiano e sia  $f$  un endomorfismo hermitiano definito in  $V$ . Allora tutti gli autovalori di  $f$  sono reali.*

**Dimostrazione:** sia  $\lambda$  un autovalore e  $v$  un autovettore corrispondente. La proprietà 4.4 implica:

$$\phi(f(v), v) = \phi(v, f(v))$$

quindi

$$\phi(\lambda v, v) = \phi(v, \lambda v)$$

che equivale a

$$\bar{\lambda}\|v\|^2 = \lambda\|v\|^2.$$

Poiché  $v \neq 0$ , si trova  $\bar{\lambda} = \lambda$ . c.v.d.

**Corollary 5.2.** *Sia  $V$  uno spazio vettoriale reale,  $\phi$  un prodotto scalare e sia  $f$  un endomorfismo simmetrico definito in  $V$ . Allora il polinomio caratteristico di  $f$  è completamente riducibile.*

**Teorema 5.3. (teorema della base spettrale)** *Sia  $V$  uno spazio vettoriale reale (complesso),  $\phi$  un prodotto scalare (hermitiano). Sia  $f$  un endomorfismo simmetrico (hermitiano) definito in  $V$ . Allora  $f$  ammette una base ortonormale di autovettori.*

**Dimostrazione:** Dimostriamo il teorema per induzione sulla dimensione,  $n$ .

Essendo ovvia la base d'induzione dimostriamo il passo induttivo. Sia  $v$  un autovettore che è lecito assumere normalizzato. Vogliamo applicare l'ipotesi induttiva a  $U := v^\perp$ . A tale scopo osserviamo

$$u \in U \Rightarrow \phi(f(u), v) = \phi(u, f(v)) = \phi(u, \lambda v) = \lambda \phi(u, v) = 0 \Rightarrow f(u) \in U.$$

Quindi la restrizione di  $f$  al sottospazio  $U$  è un endomorfismo simmetrico (hermitiano). Pertanto è lecito applicare l'ipotesi induttiva al sottospazio  $U$  che dovrà possedere una base ortonormale di autovettori:  $\mathcal{B}_U = \{e_1, \dots, e_{n-1}\}$ . Per concludere basterà aggiungere  $v$  alla suddetta base,  $\mathcal{B}_V = \{v, e_1, \dots, e_{n-1}\}$ , per ottenere una base ortonormale di autovettori in  $V$ . c.v.d.

**Definizione 5.4.** La diagonalizzazione tramite base ortonormale di autovettori si dice anche **diagonalizzazione ortogonale**.

**Esercizio 5.5.** 1. *Osservare che la diagonalizzazione ortogonale di una matrice simmetrica può essere interpretata sia come una diagonalizzazione per similitudine che come una diagonalizzazione per congruenza.*

2. *Dimostrare che diversi autospazi di un endomorfismo simmetrico o hermitiano sono tra loro ortogonali.*

3. *Diagonalizzare ortogonalmente le seguenti matrici:*

$$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}, \begin{pmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

## DECOMPOSIZIONE AI VALORI SINGOLARI (SVD)

### 1. PRELIMINARI

**Definizione 1.1.** Data una matrice  $M \in \mathbb{C}^{m,n}$  si dicono **matrici di Gram** associate a  $M$  le due matrici  $M^+M \in \mathbb{C}^{n,n}$  e  $MM^+ \in \mathbb{C}^{m,m}$ .

**Lemma 1.2.** *Le matrici di Gram sono hermitiane positive.*

**Dimostrazione:** L'hermiticit  segue immediatamente dalla definizione, inoltre:

$$\begin{aligned} X^+M^+MX &= (MX)^+MX \geq 0, \quad \forall X \in \mathbb{C}^n \\ Y^+MM^+Y &= (M^+Y)^+M^+Y \geq 0, \quad \forall Y \in \mathbb{C}^m. \end{aligned}$$

c.v.d.

*Remark 1.3.* In generale le matrici di Gram **non** sono definite positive.

**Lemma 1.4.**

$$\text{Ker}M = (\text{Im}M^+)^\perp, \quad \forall M \in \mathbb{C}^{m,n},$$

dove l'ortogonalit    intesa rispetto al prodotto hermitiano standard.

**Dimostrazione:**

$$\begin{aligned} X \in \text{Ker}M \subset \mathbb{C}^n &\iff MX = 0 \iff \\ Y^+MX = 0, \quad \forall Y \in \mathbb{C}^m &\iff X^+M^+Y = 0, \quad \forall Y \in \mathbb{C}^m \\ &\iff X \in (\text{Im}M^+)^\perp. \end{aligned}$$

c.v.d.

**Lemma 1.5.**

$$\text{Ker}M^+ = \text{Ker}MM^+, \quad \forall M \in \mathbb{C}^{m,n}.$$

**Dimostrazione:** Essendo ovvia l'inclusione  $\text{Ker}M^+ \subset \text{Ker}MM^+$ ,   sufficiente dimostrare l'inclusione opposta:

$$\begin{aligned} Y \in \text{Ker}MM^+ \subset \mathbb{C}^m &\iff MM^+Y = 0 \iff \\ X^+MM^+Y = 0, \quad \forall X \in \mathbb{C}^m &\implies Y^+MM^+Y = 0, \\ &\iff (M^+Y)^+M^+Y = 0 \iff Y \in \text{Ker}M^+. \end{aligned}$$

c.v.d.

*Remark 1.6.* Combinando il Lemma precedente con il teorema delle dimensioni si trova che le matrici di Gram hanno lo stesso rango di  $M$  e di  $M^+$ .

In base al Lemma 1.2 e al Teorema della base spettrale, le matrici di Gram sono diagonalizzabili ortogonalmente ed hanno autovalori non negativi.

**Esercizio 1.7.** *Dimostrare che un insieme ortonormale di vettori è indipendente.*

**Definizione 1.8.** Se  $B$  è una matrice quadrata e  $\lambda$  è un autovalore di  $B$ , indicheremo con  $V_\lambda^B = \text{Ker}(B - \lambda I)$  l'autospazio di  $B$  relativo a  $\lambda$ .

**Lemma 1.9.** *Se  $\lambda \neq 0$  è un autovalore positivo di  $MM^+$  e  $\{X_1, \dots, X_k\} \subset V_\lambda^{MM^+}$  è un insieme ortonormale di autovettori di  $MM^+$ , allora  $\{M^+ \frac{X_1}{\sqrt{\lambda}}, \dots, M^+ \frac{X_k}{\sqrt{\lambda}}\} \subset V_\lambda^{M^+M}$  è un insieme ortonormale di autovettori di  $M^+M$ .*

**Dimostrazione:**

$$M^+M(M^+ \frac{X_i}{\sqrt{\lambda}}) = M^+(MM^+ \frac{X_i}{\sqrt{\lambda}}) = \lambda M^+ \frac{X_i}{\sqrt{\lambda}} \implies M^+ \frac{X_i}{\sqrt{\lambda}} \in V_\lambda^{M^+M}.$$

Inoltre:

$$(M^+ \frac{X_i}{\sqrt{\lambda}})^+(M^+ \frac{X_j}{\sqrt{\lambda}}) = (\frac{X_i}{\sqrt{\lambda}})^+ MM^+ \frac{X_j}{\sqrt{\lambda}} = \frac{\lambda}{\lambda} X_i^+ X_j = \delta_{ij}.$$

c.v.d.

**Corollario 1.10.** *L'applicazione:*

$$X \mapsto \frac{1}{\sqrt{\lambda}} M^+ X$$

è un isomorfismo tra  $V_\lambda^{MM^+}$  e  $V_\lambda^{M^+M}$ , con inverso dato da:

$$X \mapsto \frac{1}{\sqrt{\lambda}} M X$$

**Dimostrazione:** segue direttamente dal Lemma precedente, osservando che la composizione tra le due applicazioni è l'identità. c.v.d.

*Remark 1.11.* 1. Il corollario precedente implica che le due matrici di Gram condividono gli autovalori **non nulli** e le rispettive molteplicità.  
2. Se  $m \neq n$  le matrici di Gram **non** hanno lo stesso polinomio caratteristico.

## 2. DECOMPOSIZIONE AI VALORI SINGOLARI

**Definizione 2.1.** Le radici quadrate degli autovalori non nulli delle due matrici di Gram  $M^+M \in \mathbb{C}^{n,n}$  e  $MM^+ \in \mathbb{C}^{m,m}$ , si dicono **valori singolari** della matrice  $M \in \mathbb{C}^{m,n}$ .

I valori singolari di una matrice, non necessariamente quadrata, possono essere interpretati come una sorta di “autovalori generalizzati”. Infatti vale il seguente:

**Teorema 2.2. (decomposizione ai valori singolari)**

Sia  $M \in \mathbb{C}^{m,n}$ , allora esistono due matrici unitarie,  $U \in \mathbb{C}^{m,m}$  e  $V \in \mathbb{C}^{n,n}$ , tali che:

$$UMV = D = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix},$$

dove  $S = \text{diag}(s_1, \dots, s_r)$  indica la matrice diagonale contenente i valori singolari di  $M$ , contati con le rispettive molteplicità, e 0 indica matrici nulle di tipo opportuno.

**Dimostrazione:** combinando il Lemma 1.2 con il teorema della base spettrale, sappiamo che  $MM^+$  ammette una base ortonormale di autovettori  $\mathcal{B} = \{u_1, \dots, u_m\} \subset \mathbb{C}^m$ . Pertanto rimarrà definita una matrice unitaria:

$$U^+ = (u_1, \dots, u_m) =: (U_r \mid U_k), \quad U_k = (u_{r+1}, \dots, u_m),$$

dove non è restrittivo assumere che i vettori colonna  $\{u_{r+1}, \dots, u_m\} \subset \mathbb{C}^m$  rappresentino una base ortonormale di  $\text{Ker}MM^+ = \text{Ker}M^+$  (si ricordi il Lemma 1.5).

Avremo dunque:

$$(1) \quad M^+U_k = 0.$$

La definizione della matrice  $U^+$  implica:

$$UMM^+U^+ = \begin{pmatrix} S^2 & 0 \\ 0 & 0 \end{pmatrix}$$

che, una volta scritta per blocchi, diventa:

$$\begin{aligned} (MM^+U_r \mid MM^+U_k) &= MM^+(U_r \mid U_k) = (U_r \mid U_k) \begin{pmatrix} S^2 & 0 \\ 0 & 0 \end{pmatrix} = \\ &= (U_r S^2 \mid 0), \end{aligned}$$

cioè:

$$(2) \quad MM^+U_r = U_rS^2.$$

Ricordando il Lemma 1.9, definiamo ora

$$(3) \quad V_r := M^+U_rS^{-1} \in \mathbb{C}^{n,r}.$$

Le colonne  $\{v_1, \dots, v_r\} \subset \mathbb{C}^n$  di tale matrice sono ortonormali:

$$V_r^+V_r = S^{-1}U_r^+MM^+U_rS^{-1} = S^{-1}U_r^+U_rS^2S^{-1} = S^{-1}S^2S^{-1} = I_r,$$

e, poiché  $M^+$  ha rango  $r$ , costituiscono una base ortonormale di  $ImM^+$ . Completando i vettori  $\{v_1, \dots, v_r\}$  ad una base ortonormale di  $\mathbb{C}^n$ ,  $\{v_1, \dots, v_n\} \subset \mathbb{C}^n$ , possiamo definire la matrice unitaria:

$$(4) \quad V = (v_1, \dots, v_n) =: (V_r \mid V_k), \quad V_k := (v_{r+1}, \dots, v_n).$$

Combinando 3 e 4 e ricordando il Lemma 1.4, si vede che le colonne di  $V_k$  costituiscono una base ortonormale di  $KerM$ , quindi:

$$(5) \quad MV_k = 0.$$

Concludendo:

$$\begin{aligned} U MV &= \begin{pmatrix} U_r^+ \\ U_k^+ \end{pmatrix} M(V_r \mid V_k) = \\ &= \begin{pmatrix} U_r^+ M V_r & U_r^+ M V_k \\ U_k^+ M V_r & U_k^+ M V_k \end{pmatrix} = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

in quanto:

- $U_r^+ M V_k = U_k^+ M V_k = 0$  per la 5,
- $U_k^+ M V_r = (V_r^+ M^+ U_k)^+ = 0$  per la 1,
- $U_r^+ M V_r = U_r^+ M M^+ U_r S^{-1} = U_r^+ U_r S^2 S^{-1} = S$  per la 2.

c.v.d.

*Remark 2.3.* Le colonne della matrice  $V$  costituiscono una base ortonormale di autovettori di  $M^+M$ . Infatti:

$$\begin{aligned} V^+ M^+ M V &= \begin{pmatrix} V_r^+ \\ V_k^+ \end{pmatrix} M^+ M (V_r \mid V_k) = \\ &= \begin{pmatrix} V_r^+ M^+ M V_r & V_r^+ M^+ M V_k \\ V_k^+ M^+ M V_r & V_k^+ M^+ M V_k \end{pmatrix} = \begin{pmatrix} S^2 & 0 \\ 0 & 0 \end{pmatrix}, \end{aligned}$$

in quanto:

- $V_r^+ M^+ M V_k = V_k^+ M^+ M V_r = V_k^+ M^+ M V_k$  per la 5,

$$\begin{aligned} \bullet V_r^+ M^+ M V_r &= S^{-1} U_r^+ M M^+ M M^+ U_r S^{-1} = \\ &= S^{-1} S^2 U_r^+ U_r S^2 S^{-1} = S^2 \text{ per la 2.} \end{aligned}$$

Quindi la SVD **diagonalizza simultaneamente le due matrici di Gram**.

La dimostrazione del teorema 2.2 è valida anche nel campo reale:

**Teorema 2.4.** *Sia  $M \in \mathbb{R}^{m,n}$ , allora esistono due matrici ortogonali,  $U \in \mathbb{R}^{m,m}$  e  $V \in \mathbb{R}^{n,n}$ , tali che:*

$$UMV = D = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix},$$

dove  $S = \text{diag}(s_1, \dots, s_r)$  indica la matrice diagonale contenente i valori singolari di  $M$ , contati con le rispettive molteplicità, e  $0$  indica matrici nulle di tipo opportuno.

**Esercizio 2.5.** *Determinare una decomposizione ai valori singolari per le seguenti matrici:*

$$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

# NORMA ED ESPONENZIALE DI UN OPERATORE

## 1. NORMA DI UN OPERATORE

Sia  $V$  uno spazio vettoriale reale o complesso dotato di una norma  $\|\cdot\|_V$ , ossia di una funzione:

$$\|\cdot\|_V : V \mapsto \mathbb{R}_{\geq 0},$$

che gode delle seguenti proprietà:

1.  $\|v\|_V = 0 \iff v = 0$ ;
2.  $\|\lambda v\|_V = |\lambda| \|v\|_V, \quad \forall \lambda \in \mathbb{K}, \forall v \in V$ ;
3.  $\|v + u\|_V \leq \|v\|_V + \|u\|_V, \forall v, u \in V$ , (disuguaglianza triangolare).

**Definizione 1.1.** Diremo che due norme,  $\|\cdot\|_1$  e  $\|\cdot\|_2$ , definite su uno spazio vettoriale, sono **equivalenti** se esistono due costanti positive,  $m_1$  e  $m_2$ , tali che:

$$\|v\|_1 \leq m_1 \|v\|_2 \leq m_2 \|v\|_1, \quad \forall v \in V.$$

Sussiste il seguente, importante,

**Teorema 1.2.** *Tutte le norme definite su uno spazio vettoriale **di dimensione finita** sono tra loro equivalenti.*

**Definizione 1.3.** Sia  $f \in \text{End}(V)$  un endomorfismo dello spazio normato  $V$ . Indicheremo con  $\|f\|$  la **norma spettrale** di  $f$ :

$$\|f\| := \sup_{v \neq 0} \frac{\|f(v)\|_V}{\|v\|_V}.$$

Tale definizione è ben posta in quanto:

$$\sup_{v \neq 0} \frac{\|f(v)\|_V}{\|v\|_V} = \sup_{v \neq 0} \left\| \frac{f(v)}{\|v\|_V} \right\|_V = \sup_{\|w\|_V=1} \|f(w)\|_V,$$

e l'ultima quantità è ben definita per il Teorema di Bolzano-Weierstrass, in quanto l'insieme  $\{w \in V : \|w\|_V = 1\}$  è compatto (in coordinate ortonormali esso è rappresentato da una sfera unitaria):

$$(1) \quad \sup_{\|w\|_V=1} \|f(w)\|_V = \max_{\|w\|_V=1} \|f(w)\|_V.$$

*Osservazione 1.4.* 1. L'espressione 1.3 implica:

$$\exists w \in V, \|w\|_V = 1, \text{ tale che: } \|f\| = \|f(w)\|_V.$$

2. Dalla definizione segue immediatamente:

$$\|f(v)\|_V \leq \|f\| \|v\|_V.$$

**Teorema 1.5.** *Le definizione 1.3, dà a  $EndV$  una struttura di spazio normato.*

**Dimostrazione:** dobbiamo verificare le tre proprietà che caratterizzano una norma.

1. Dalla definizione segue:

$$\begin{aligned} \|f\| = 0 &\iff \|f(v)\|_V = 0, \forall v \in V \iff \\ &\iff f(v) = 0, \forall v \in V \iff f = 0. \end{aligned}$$

2.

$$\begin{aligned} \|\lambda f\| &= \sup_{v \neq 0} \frac{\|\lambda f(v)\|_V}{\|v\|_V} = \sup_{v \neq 0} |\lambda| \frac{\|f(v)\|_V}{\|v\|_V} = \\ &= |\lambda| \sup_{v \neq 0} \frac{\|f(v)\|_V}{\|v\|_V} = |\lambda| \|f\|. \end{aligned}$$

3. Fissiamo  $f, g \in End(V)$ . Per l'osservazione 1.4, esiste un vettore  $v \in V$  tale che  $\|f + g\| = \|(f + g)(v)\|_V$ . Allora

$$\begin{aligned} \|f + g\| &= \|(f + g)(v)\|_V \leq \|f(v)\|_V + \|g(v)\|_V \leq \\ &\leq \max_{\|w\|_V=1} \|f(w)\|_V + \max_{\|w\|_V=1} \|g(w)\|_V = \|f\| + \|g\|. \end{aligned}$$

c.v.d.

**Teorema 1.6.**

$$\|f \circ g\| \leq \|f\| \|g\|, \quad \forall f, g \in End(V).$$

**Dimostrazione:** possiamo assumere  $f \circ g \neq 0$ , altrimenti l'enunciato è scontato. Per l'osservazione 1.4, esiste un vettore  $v \in V$  tale che  $\|f \circ g\| = \|(f \circ g)(v)\|_V = \|f(g(v))\|_V$ , quindi dovrà aversi  $g(v) \neq 0$  (altrimenti  $\|f \circ g\| = \|f(g(v))\|_V = 0$ , contrariamente all'ipotesi  $f \circ g \neq 0$ ). Allora

$$\begin{aligned} \|f \circ g\| &= \|(f \circ g)(v)\|_V = \|f(g(v))\|_V = \frac{\|g(v)\|_V}{\|g(v)\|_V} \|f(g(v))\|_V = \\ &= \|g(v)\|_V \|f\left(\frac{g(v)}{\|g(v)\|_V}\right)\|_V \leq (\max_{\|w\|_V=1} \|f(w)\|_V) \|g(v)\|_V. \end{aligned}$$

$$\cdot(\max_{\|w\|_V=1} \|f(w)\|_V) = \|f\| \|g\|.$$

c.v.d.

**Esercizio 1.7.** Fissato  $g \in \text{End}V$ , dimostrare la continuità dell' applicazione:

$$F_g : \text{End}V \longmapsto \text{End}V, \quad f \longmapsto F_g(f) = f \circ g.$$

**Definizione 1.8.** Analogamente, data  $F \in \mathbb{R}^{n,n}$  ( $F \in \mathbb{C}^{n,n}$ ), si pone:

$$\|F\| := \max_{\|X\|_S=1} \|FX\|_S,$$

dove  $\|\cdot\|_S$  indica il prodotto scalare (hermitiano) standard.

*Osservazione 1.9.* Fissata una base  $\mathcal{B}$  nello spazio vettoriale reale  $V$ , indicata con  $A$  la matrice associata al prodotto scalare nella base  $\mathcal{B}$  e indicata con  $F$  la matrice associata ad un endomorfismo  $f$  nella base  $\mathcal{B}$ , si trova:

$$\|f\| := \max_{\|v\|_V=1} \|f(v)\|_V = \max_{X^t A X=1} \sqrt{X^t F^t A F X}.$$

Pertanto, in generale,

$$\|f\| \neq \|F\| = \max_{X^t X=1} \sqrt{X^t F^t F X},$$

**a meno che la base non sia ortonormale.**

## 2. SERIE GEOMETRICA

**Teorema 2.1.** Sia  $f \in \text{End}V$  un endomorfismo tale che  $\|f\| < 1$ . Allora la serie:

$$\sum_{n=0}^{+\infty} f^n,$$

converge all'inverso di  $(f - id) \in \text{End}V$ .

**Dimostrazione:** dimostriamo la convergenza. Per il teorema 1.6, abbiamo

$$\|f^n\| \leq \|f\|^n, \quad \forall n > 0.$$

Ponendo  $\|f\| = \rho < 1$ , sappiamo che la serie  $\sum_{n=0}^{+\infty} \rho^n$  converge (al valore  $(1 - \rho)^{-1}$ ).

Dobbiamo dimostrare che  $\forall \epsilon > 0, \exists N \in \mathbb{N}$  tale che:

$$\left\| \sum_N^M f^n \right\| \leq \epsilon, \quad \forall M \geq N.$$

Ciò segue da:

$$\left\| \sum_N^M f^n \right\| \leq \sum_N^M \|f^n\| \leq \sum_N^M \|f\|^n = \sum_N^M \rho^n \leq \epsilon, \quad \forall M \geq N,$$

per un opportuno  $N$ , in virtù della convergenza della serie  $\sum_{n=0}^{\infty} \rho^n$ .

Dimostriamo ora che:

$$\sum_{n=0}^{+\infty} f^n = (id - f)^{-1},$$

ossia che:

$$(id - f) \circ \sum_{n=0}^{+\infty} f^n = id.$$

Infatti:

$$\begin{aligned} (id - f) \circ \sum_{n=0}^{+\infty} f^n &= \lim_{N \rightarrow +\infty} (id - f) \circ \sum_{n=0}^N f^n = \\ &= \lim_{N \rightarrow +\infty} \left( \sum_{n=0}^N f^n - \sum_{n=1}^{N+1} f^n \right) = \lim_{N \rightarrow +\infty} (id - f^{N+1}) = id, \end{aligned}$$

in quanto:

$$\|f^{N+1}\| \leq \|f\|^{N+1} = \rho^{N+1} \rightarrow 0, \quad \text{per } N \rightarrow +\infty.$$

c.v.d.

**Definizione 2.2.** Sia  $f \in \text{End}V$  un endomorfismo tale che  $\|f\| < 1$ .

La serie

$$\sum_{n=0}^{+\infty} f^n$$

si dice **serie geometrica** di  $f$ .

### 3. PROCEDIMENTI ITERATIVI

Sia  $A \in \mathbb{R}^{n,n}$  una matrice invertibile, e sia

$$(2) \quad AX = B$$

il sistema di Cramer associato ad  $A$  e alla colonna dei termini noti  $B \in \mathbb{R}^n$ . Assumiamo che la matrice  $A$  possa essere decomposta nella forma:

$$A = M - N,$$

con  $M \in \mathbb{R}^{n,n}$  invertibile. Indicata con  $Y = A^{-1}B \in \mathbb{R}^n$  la soluzione esatta del sistema 2, otteniamo:

$$(3) \quad Y = M^{-1}(NY + B).$$

**Definizione 3.1.** 1. Sia  $X_0 \in \mathbb{R}^n$  un vettore generico, la successione:

$$X_{m+1} = M^{-1}(NX_m + B)$$

si dice **procedimento iterativo** associato al problema 3 con **dato iniziale**  $X_0$ .

2. Il procedimento iterativo si dice **globalmente convergente** se, per ogni dato iniziale  $X_0 \in \mathbb{R}^n$ , tale successione converge alla soluzione esatta:

$$\lim_{m \rightarrow +\infty} X_m = Y.$$

**Teorema 3.2.** *Se  $\|M^{-1}N\| < 1$  allora il procedimento iterativo è globalmente convergente.*

**Dimostrazione:** per semplificare le notazioni poniamo  $P := M^{-1}N$ . Poiché

$$A = M - N = M(I_n - P),$$

e poiché  $\|P\| < 1$ , otteniamo:

$$(4) \quad A^{-1} = \left( \sum_{i=0}^{+\infty} P^i \right) M^{-1}$$

Definiamo

$$\Delta_m := X_m - Y = X_m - A^{-1}B,$$

che rappresenta la differenza tra l'  $m$ -esimo termine della successione e la soluzione esatta. La tesi è la seguente:

$$(5) \quad \lim_{m \rightarrow +\infty} \Delta_m = 0, \quad \forall X_0 \in \mathbb{R}^n.$$

Otteniamo:

$$\begin{aligned} \Delta_{m+1} &= X_{m+1} - A^{-1}B = PX_m + M^{-1}B - \left( \sum_{i=0}^{+\infty} P^i \right) M^{-1}B = \\ &= PX_m - \left( \sum_{i=1}^{+\infty} P^i \right) M^{-1}B = PX_m - P \left( \sum_{i=0}^{+\infty} P^i \right) M^{-1}B = \\ &= P \left( X_m - \left( \sum_{i=0}^{+\infty} P^i \right) M^{-1}B \right) = P(X_m - A^{-1}B) = P\Delta_m, \end{aligned}$$

quindi  $\Delta_{m+1} = P^{m+1}\Delta_0$  e

$$\|\Delta_{m+1}\|_S \leq \|P^{m+1}\| \|\Delta_0\|_S \leq \|P\|^{m+1} \|\Delta_0\|_S \rightarrow 0, \text{ se } m \rightarrow +\infty,$$

perché  $\|P\| < 1$ . c.v.d.

**Esercizio 3.3.** 1. **Metodo di Jacobi.** Scegliendo  $M$  coincidente con la parte diagonale di  $A$ ,  $M := \text{diag}(a_{11}, \dots, a_{nn})$ , si trova:

$$x_i^{(m+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j \neq i} a_{ij} x_j^{(m)} \right),$$

dove si è posto  $X_m = (x_1^{(m)}, x_2^{(m)}, \dots, x_n^{(m)})$  e  $B = (b_1, b_2, \dots, b_n)$ .

2. **Metodo di Gauss-Seidel.** Scegliendo  $M$  coincidente con la parte triangolare inferiore di  $A$ :

$$M := (m_{ij}) \mid m_{ij} = a_{ij} \text{ se } i \geq j, \quad m_{ij} = 0 \text{ se } i < j,$$

si trova:

$$x_i^{(m+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(m+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(m)} \right).$$

#### 4. PERTURBAZIONI DI UNA MATRICE DEFINITA POSITIVA

**Lemma 4.1.** Indichiamo con  $A \in \mathbb{R}^{n,n}$  una matrice simmetrica e con  $\lambda$  il minimo tra i suoi autovalori, allora:

$$\lambda = \min_{\|X\|_S=1} X^t A X.$$

**Dimostrazione:** per il teorema della base spettrale, esiste una matrice ortogonale  $U$  ( $U^t U = I_n$ ), tale che  $A = U^t D U$ ,  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$  ( $\lambda_i$  sono gli autovalori di  $A$ ). La matrice  $U$ , essendo ortogonale, rappresenta un'isometria di  $\mathbb{R}^n$ , pertanto avremo:

$$\begin{aligned} \min_{\|X\|_S=1} X^t A X &= \min_{\|X\|_S=1} X^t U^t D U X = \min_{\|X\|_S=1} (U X)^t D U X = \\ &= \min_{\|U X\|_S=1} (U X)^t D U X = \min_{\|Y\|_S=1} Y^t D Y = \min_{\|Y\|_S=1} \sum_1^n \lambda_i y_i^2. \end{aligned}$$

Quindi:

$$\min_{\|X\|_S=1} X^t A X \geq \min_{\|Y\|_S=1} \sum_1^n \lambda y_i^2 = \lambda \cdot \min_{\|Y\|_S=1} \sum_1^n y_i^2 = \lambda.$$

D'altra parte, se  $Y$  è un autovettore normalizzato relativo a  $\lambda$ , avremo infine:

$$\lambda = Y^t A Y \geq \min_{\|X\|_S=1} X^t A X.$$

c.v.d.

**Teorema 4.2.** *Se  $A \in \mathbb{R}^{n,n}$  è definita positiva e  $\lambda > 0$  è il minimo autovalore di  $A$ , allora ogni matrice simmetrica  $B \in \mathbb{R}^{n,n}$ , tale che  $\|B - A\| < \lambda$ , è definita positiva.*

**Dimostrazione:** In base al lemma precedente avremo:

$$\min_{\|X\|_S=1} X^t A X = \lambda > 0,$$

e basterà dimostrare:

$$\min_{\|X\|_S=1} X^t B X > 0.$$

Poiché

$$X^t B X = X^t A X + X^t (B - A) X,$$

si trova:

$$\begin{aligned} \min_{\|X\|_S=1} X^t B X &\geq \min_{\|X\|_S=1} X^t A X + \min_{\|X\|_S=1} X^t (B - A) X = \\ (6) \qquad &= \lambda + \min_{\|X\|_S=1} X^t (B - A) X. \end{aligned}$$

Per il teorema di Bolzano-Weierstrass, il valore  $\min_{\|X\|_S=1} X^t (B - A) X$  sarà assunto per un opportuno versore  $X_0$ :

$$(7) \qquad \min_{\|X\|_S=1} X^t (B - A) X = X_0^t (B - A) X_0.$$

Se  $X_0^t (B - A) X_0 \geq 0$  allora il teorema è provato, se  $X_0^t (B - A) X_0 < 0$  allora:

$$(8) \qquad X_0^t (B - A) X_0 = - | X_0^t (B - A) X_0 |$$

e, inoltre,

$$| X_0^t (B - A) X_0 | \leq \|X_0\|_S \| (B - A) X_0 \|_S \leq \|B - A\|,$$

per la disuguaglianza di Cauchy-Schwarz. Combinando l'ultima disuguaglianza con la 6, la 7 e la 8 troviamo:

$$\min_{\|X\|_S=1} X^t B X \geq \lambda - | X_0^t (B - A) X_0 | \geq \lambda - \|B - A\| > \lambda - \lambda = 0.$$

c.v.d.

## 5. ESPONENZIALE DI UN OPERATORE

**Definizione 5.1.** Dato  $f \in \text{End}V$ , definiamo la **serie esponenziale** di  $f$ :

$$e^f := \sum_{n=0}^{+\infty} \frac{f^n}{n!} = id + f + \frac{f \circ f}{2} + \dots$$

**Teorema 5.2.** *La serie esponenziale converge, per ogni  $f \in \text{End}V$ .*

**Dimostrazione:** ponendo  $\|f\| = \rho$ , sappiamo che la serie  $\sum_{n=0}^{+\infty} \frac{\rho^n}{n!}$  converge.

Dobbiamo dimostrare che  $\forall \epsilon > 0, \exists N \in \mathbb{N}$  tale che:

$$\left\| \sum_N^M \frac{f^n}{n!} \right\| \leq \epsilon, \quad \forall M \geq N.$$

Ciò segue da:

$$\left\| \sum_N^M \frac{f^n}{n!} \right\| \leq \sum_N^M \frac{\|f^n\|}{n!} \leq \sum_N^M \frac{\|f\|^n}{n!} = \sum_N^M \frac{\rho^n}{n!} \leq \epsilon, \quad \forall M \geq N,$$

per un opportuno  $N$ , in virtù della convergenza della serie  $\sum_{n=0}^{\infty} \frac{\rho^n}{n!}$ .  
c.v.d.

*Osservazione 5.3.* Dalla definizione segue:

$$e^0 = id.$$

**Definizione 5.4.** Analogamente, data una matrice quadrata  $F$ , a coefficienti reali oppure complessi, definiamo la **serie esponenziale** di  $F$ :

$$e^F := \sum_{n=0}^{+\infty} \frac{F^n}{n!}.$$

**Proposizione 5.5.** *Se  $f, g \in \text{End}V$ , con  $g$  invertibile, allora:*

$$g^{-1} \circ e^f \circ g = e^{g^{-1} \circ f \circ g}.$$

**Dimostrazione:** poiché

$$(g^{-1} \circ f \circ g)^n = (g^{-1} \circ f \circ g) \circ (g^{-1} \circ f \circ g) \circ \dots \circ (g^{-1} \circ f \circ g) = g^{-1} \circ f^n \circ g,$$

troviamo:

$$g^{-1} \circ e^f \circ g = \lim_{N \rightarrow +\infty} g^{-1} \circ \left( \sum_{n=0}^N \frac{f^n}{n!} \right) \circ g = \lim_{N \rightarrow +\infty} \sum_{n=0}^N g^{-1} \circ \frac{f^n}{n!} \circ g =$$

$$= \lim_{N \rightarrow +\infty} \sum_{n=0}^N (g^{-1} \circ \frac{f}{n!} \circ g)^n = e^{g^{-1} \circ f \circ g}.$$

c.v.d.

**Corollario 5.6.** *Se  $F$  è diagonalizzabile, allora anche  $e^F$  è diagonalizzabile dagli stessi autovettori di  $F$  e con autovalori uguali agli esponenziali degli autovalori di  $F$ .*

**Dimostrazione:** Se  $P$  è una matrice le cui colonne costituiscono una base di autovettori di  $F$ , avremo:

$$P^{-1}FP = D = \text{diag}(\lambda_1, \dots, \lambda_n),$$

dove  $\lambda_i$ ,  $1 \leq i \leq n$ , rappresentano gli autovalori di  $F$ .

Ma allora la proposizione precedente implica:

$$P^{-1}e^F P = e^{P^{-1}FP} = e^D = \sum_{n=0}^{+\infty} \frac{D^n}{n!} = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}).$$

c.v.d.

**Teorema 5.7.** *Siano  $f$  e  $g$  due operatori commutanti:  $f \circ g = g \circ f$ . Allora  $e^{f+g} = e^f \circ e^g$ .*

**Dimostrazione:** Poiché

$$e^f \circ e^g =: e^f e^g = \lim_{N \rightarrow +\infty} \sum_{n,m=0}^N \frac{f^n g^m}{n! m!},$$

è sufficiente dimostrare che la differenza

$$(9) \quad \sum_{n,m=0}^N \frac{f^n g^m}{n! m!} - \sum_{l=0}^N \frac{(f+g)^l}{l!}$$

tende a 0, se  $N$  tende a  $+\infty$ .

Poiché i due operatori commutano, abbiamo:

$$(f+g)^2 = f^2 + g^2 + fg + gf = f^2 + g^2 + 2fg$$

e, più in generale, vale la formula del binomio di Newton:

$$(10) \quad (f+g)^l = \sum_{k=0}^l \binom{l}{k} f^k g^{l-k},$$

pertanto la differenza 9 diventa:

$$\sum_{n,m=0}^N \frac{f^n g^m}{n! m!} - \sum_{l=0}^N \frac{1}{l!} \sum_{k=0}^l \binom{l}{k} f^k g^{l-k} =$$

$$\begin{aligned}
&= \sum_{n,m=0}^N \frac{f^n g^m}{n! m!} - \sum_{l=0}^N \sum_{k=0}^l \frac{f^k g^{l-k}}{k! (l-k)!} = \\
&= \sum_{n,m=0}^N \frac{f^n g^m}{n! m!} - \sum_{0 \leq n+m \leq N} \frac{f^n g^m}{n! m!} = \\
&= \sum_D \frac{f^n g^m}{n! m!}, \quad D := \{(n, m) \in \mathbb{N}^2 \mid n, m \leq N, n+m > N\}.
\end{aligned}$$

Concludendo troviamo:

$$\begin{aligned}
\left\| \sum_D \frac{f^n g^m}{n! m!} \right\| &\leq \sum_D \frac{\|f\|^n \|g\|^m}{n! m!} \leq \sum_D \frac{\|f\|^n \|g\|^m}{n! m!} \leq \\
&\leq \sum_{n+m > N} \frac{\|f\|^n \|g\|^m}{n! m!} = \sum_{l=N+1}^{+\infty} \sum_{k=0}^l \frac{\|f\|^k \|g\|^{l-k}}{k! (l-k)!} = \\
&= \sum_{l=N+1}^{+\infty} \frac{1}{l!} \sum_{k=0}^l \binom{l}{k} \|f\|^k \|g\|^{l-k} = \sum_{l=N+1}^{+\infty} \frac{1}{l!} (\|f\| + \|g\|)^l,
\end{aligned}$$

che tende a 0 (per  $N \rightarrow +\infty$ ), in virtù della convergenza della serie esponenziale. c.v.d.

**Corollario 5.8.** Per ogni  $f \in \text{End}V$ ,  $e^f$  è invertibile e, inoltre,  $(e^f)^{-1} = e^{-f}$ .

**Dimostrazione:**

$$e^f e^{-f} = e^{f-f} = e^0 = \text{id}.$$

c.v.d.

Più in generale:

**Corollario 5.9.** Per ogni  $f \in \text{End}V$  e per ogni coppia di scalari  $h, k$ :

$$e^{hf} e^{kf} = e^{(h+k)f}.$$

**Teorema 5.10.** Per ogni  $f \in \text{End}V$  la funzione

$$\mathbb{R} (\mathbb{C}) \longmapsto \text{End}V, \quad t \longmapsto e^{tf}$$

è derivabile in  $t$  e, più precisamente:

$$(e^{tf})' = f e^{tf} = e^{tf} f.$$

**Dimostrazione:** ovviamente  $f e^{tf} = e^{tf} f$ , quindi è sufficiente dimostrare  $(e^{tf})' = f e^{tf}$ .

Per cominciare, poniamo  $t = 0$  e dimostriamo:

$$(11) \quad (e^{tf})'_{t=0} = f,$$

cioè:

$$(12) \quad \lim_{t \rightarrow 0} \left\| \frac{e^{tf} - id}{t} - f \right\| = 0.$$

Poiché

$$\frac{e^{tf} - id}{t} - f = \frac{1}{t} \sum_{k=2}^{+\infty} \frac{(tf)^k}{k!},$$

otteniamo

$$\left\| \frac{e^{tf} - id}{t} - f \right\| = \left\| \frac{1}{t} \sum_{k=2}^{+\infty} \frac{(tf)^k}{k!} \right\| \leq \frac{1}{|t|} \sum_{k=2}^{+\infty} \frac{\|tf\|^k}{k!} = \frac{e^{\|tf\|} - 1}{|t|} - \|f\|,$$

che tende a 0 (per  $t \rightarrow 0$ ), in quanto  $(e^{t\|f\|})'_{t=0} = \|f\|$ . Pertanto la 12 (e la 11) è dimostrata.

Concludiamo la dimostrazione del teorema e mostriamo:

$$(e^{tf})'_{t=t_0} = f e^{t_0 f}.$$

Per il corollario 5.9:

$$e^{tf} = e^{(t-t_0)f} e^{t_0 f},$$

quindi:

$$\begin{aligned} (e^{tf})'_{t=t_0} &= \lim_{t \rightarrow t_0} \frac{e^{tf} - e^{t_0 f}}{t - t_0} = \lim_{t \rightarrow t_0} \frac{e^{(t-t_0)f} e^{t_0 f} - e^{t_0 f}}{t - t_0} = \\ &= \lim_{t \rightarrow t_0} \frac{e^{(t-t_0)f} - id}{t - t_0} e^{t_0 f} = \left( \lim_{t \rightarrow t_0} \frac{e^{(t-t_0)f} - id}{t - t_0} \right) e^{t_0 f} = \\ &= \left( \lim_{\tau \rightarrow 0} \frac{e^{\tau f} - id}{\tau} \right) e^{t_0 f} = f e^{t_0 f}. \end{aligned}$$

c.v.d.

## LA FORMA CANONICA DI UN OPERATORE

### 1. MATRICI DIAGONALI A BLOCCHI

Con delle linee orizzontali e verticali tratteggiate possiamo ripartire una matrice  $A$  in tante matrici più piccole, chiamate **blocchi** di  $A$ . La matrice  $A$  viene allora chiamata **matrice a blocchi**. Una data matrice può essere ripartita in blocchi in vari modi.

**Esempio 1.1.**

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 3 & 2 \\ 1 & 7 & 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & | & 3 & 4 & | & 5 \\ \hline 4 & 1 & | & 5 & 3 & | & 2 \\ \hline 1 & 7 & | & 2 & 3 & | & 1 \end{pmatrix} = \begin{pmatrix} \underline{1} & \underline{2} & \underline{3} & | & \underline{4} & \underline{5} \\ \hline \underline{4} & \underline{1} & \underline{5} & | & \underline{3} & \underline{2} \\ \hline 1 & 7 & 2 & | & 3 & 1 \end{pmatrix}.$$

*In questo esempio la matrice  $A$  è una matrice  $3 \times 5$ ; nella prima decomposizione  $A$  appare come una matrice a blocchi:  $2 \times 3$ , cioè appare nella forma:*

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix},$$

dove  $A_{11} = \begin{pmatrix} 1 & 2 \\ 4 & 1 \end{pmatrix}$ ,  $A_{12} = \begin{pmatrix} 3 & 4 \\ 5 & 3 \end{pmatrix}$ ,  $A_{13} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$ ,  $A_{21} = \begin{pmatrix} 1 & 7 \end{pmatrix}$ ,  
 $A_{22} = \begin{pmatrix} 2 & 3 \end{pmatrix}$ ,  $A_{23} = \begin{pmatrix} 1 \end{pmatrix}$ . Invece nella seconda decomposizione  $A$  appare come una matrice a blocchi  $3 \times 2$ .

La convenienza della partizione in blocchi consiste nel fatto che (almeno in certi casi) il risultato di operazioni su matrici a blocchi si può ottenere portando avanti il calcolo sui blocchi stessi, proprio come se essi fossero gli elementi effettivi della matrice. Ciò conduce ad una semplificazione dei calcoli. Questo avviene soprattutto in presenza di matrici diagonali a blocchi. Una matrice si dice **matrice diagonale a blocchi** se è una matrice quadrata, se i blocchi disposti lungo la diagonale principale sono quadrati, e se tutti i blocchi fuori della diagonale sono nulli.

**Esempio 1.2.** *La seguente matrice:*

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 & 6 \end{pmatrix}$$

è diagonale a blocchi, con i blocchi lungo la diagonale principale dati da  $A_{11} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ ,  $A_{22} = (3)$  ed  $A_{33} = \begin{pmatrix} 1 & 5 \\ 1 & 6 \end{pmatrix}$ . Si può rappresentare  $A$  anche sotto la forma

$$A = \begin{pmatrix} A_{11} & & \\ & A_{22} & \\ & & A_{33} \end{pmatrix},$$

in cui si indicano solo i blocchi disposti lungo la diagonale principale, e gli spazi bianchi stanno ad indicare i blocchi nulli.

Più in generale una matrice diagonale a blocchi si può indicare sotto la forma:

$$\begin{pmatrix} A_{11} & & & & \\ & A_{22} & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & A_{kk} \end{pmatrix}.$$

E' evidente che la somma di matrici diagonali a blocchi (con blocchi corrispondenti della stessa grandezza) è ancora una matrice diagonale a blocchi. Ciò vale anche per il prodotto per uno scalare. Inoltre se  $A_{11}, \dots, A_{kk}$  denotano i blocchi sulla diagonale di una matrice  $A$  diagonale a blocchi, allora il determinante di  $A$  è uguale al prodotto dei determinanti dei singoli blocchi, cioè  $\det(A) = \det(A_{11}) \cdots \det(A_{kk})$ . Infine  $A$  è invertibile se e solo se i singoli blocchi sono invertibili, e l'inversa di  $A$  è quella matrice diagonale a blocchi con i blocchi dati dalle matrici inverse dei blocchi di  $A$ .

**Esempio 1.3.** Con riferimento all'esempio precedente, abbiamo:

$$A^2 = \begin{pmatrix} A_{11}^2 & & & & \\ & A_{22}^2 & & & \\ & & A_{33}^2 & & \\ & & & & \\ & & & & \end{pmatrix} = \begin{pmatrix} 5 & 3 & 0 & 0 & 0 \\ 3 & 2 & 0 & 0 & 0 \\ 0 & 0 & 9 & 0 & 0 \\ 0 & 0 & 0 & 6 & 35 \\ 0 & 0 & 0 & 7 & 41 \end{pmatrix},$$

ed anche

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} & & & & \\ & A_{22}^{-1} & & & \\ & & A_{33}^{-1} & & \\ & & & & \\ & & & & \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 6 & -5 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}.$$

Infine

$$\det(A) = \det(A_{11}) \cdot \det(A_{22}) \cdot \det(A_{33}) = 1 \cdot 3 \cdot 1 = 3.$$

## 2. MATRICI DIAGONALI A BLOCCHI E SOTTOSPAZI INVARIANTI DI UN OPERATORE LINEARE

Ora vogliamo mostrare che in certe situazioni la matrice rappresentativa di un operatore si può presentare come una matrice diagonale a blocchi. Il concetto che consente di tradurre la nozione di matrice a blocchi nel contesto degli operatori lineari è quello di **sottospazio invariante**.

**Definizione 2.1.** Sia  $f : V \rightarrow V$  un endomorfismo di uno spazio vettoriale  $V$ . Un sottospazio  $U$  di  $V$  si dice **sottospazio invariante** per  $f$  se i vettori di  $U$  sono trasformati tramite  $f$  ancora in vettori di  $U$ , cioè se  $f(U) \subset U$ .

In tal caso l'operatore  $f$  induce per restrizione un operatore di  $U$ , cioè l'operatore  $f_U : u \in U \rightarrow f(u) \in U$ . Si osservi anche che sia  $V$  che  $\{0\}$  sono sempre invarianti. Altri esempi di sottospazi invarianti sono gli autospazi. La seguente osservazione è utile per riconoscere se un sottospazio è oppure no invariante.

*Osservazione 2.2.* Supponiamo che  $U$  sia generato dai vettori  $u_1, \dots, u_h$ . Allora  $U$  è invariante per  $f$  se e solo se i vettori  $f(u_1), \dots, f(u_h)$  appartengono ancora ad  $U$ .

Quanto segue mostra il legame con le matrici diagonali a blocchi. Supponiamo che  $V$  sia somma diretta di due sottospazi invarianti per  $f$ ,  $U$  e  $W$ :

$$V = U \oplus W.$$

Denotiamo con  $\mathcal{B}_U = \{u_1, \dots, u_h\}$  una base per  $U$  e con  $\mathcal{B}_W = \{w_1, \dots, w_k\}$  una base per  $W$ . Sappiamo che  $\mathcal{B} := \mathcal{B}_U \cup \mathcal{B}_W$  è una base per  $V$ . Sia  $f_U : U \rightarrow U$  l'operatore indotto da  $f$  su  $U$ , e sia  $f_W : W \rightarrow W$  quello indotto su  $W$ .

**Esercizio 2.3.** *La matrice rappresentativa di  $f$  rispetto alla base  $\mathcal{B}$ ,  $M(f; \mathcal{B}, \mathcal{B})$ , è diagonale a blocchi, con blocchi dati dalla matrice  $M(f_U; \mathcal{B}_U, \mathcal{B}_U)$  di  $f_U$  rispetto ad  $\mathcal{B}_U$  e dalla matrice  $M(f_W; \mathcal{B}_W, \mathcal{B}_W)$  di  $f_W$  rispetto ad  $\mathcal{B}_W$ , cioè:*

$$M(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} M(f_U; \mathcal{B}_U, \mathcal{B}_U) & 0 \\ 0 & M(f_W; \mathcal{B}_W, \mathcal{B}_W) \end{pmatrix}.$$

In altre parole, la conoscenza di una decomposizione di  $V$  in sottospazi invarianti permette di rappresentare il dato operatore con una matrice diagonale a blocchi. Ciò consente di semplificare lo studio dell'operatore. Ad esempio il polinomio caratteristico di  $f$  si decompone nel prodotto del polinomio di  $f_U$  e del polinomio di  $f_W$ , quindi per calcolare gli autovalori di  $f$  ci si riconduce a calcolare le radici di polinomi aventi grado più piccolo di quello di  $f$ . Una ovvia generalizzazione si ha quando  $V$  ammette una decomposizione in più di due sottospazi invarianti.

**Esempio 2.4.** *Sia  $f : \mathbb{C}^4 \rightarrow \mathbb{C}^4$  l'operatore lineare definito ponendo  $f(x, y, z, t) := (x + y + z, t, -y, x + z)$ . Consideriamo i seguenti sottospazi:  $U := L(2e_1 + e_2 - e_3 + e_4)$ ,  $W := L(e_1 - e_3, e_2, e_4)$ . Poniamo  $\mathcal{B}_U := \{2e_1 + e_2 - e_3 + e_4\}$  e  $\mathcal{B}_W := \{e_1 - e_3, e_2, e_4\}$ . E' chiaro che  $\mathcal{B}_U$  è una base per  $U$ ,  $\mathcal{B}_W$  è una base per  $W$ , e che  $\mathcal{B} := \mathcal{B}_U \cup \mathcal{B}_W$  è una base per  $\mathbb{C}^4$ . Dunque  $\mathbb{C}^4 = U \oplus W$ . Proviamo che  $U$  e  $W$  sono invarianti per  $f$ . Infatti  $f(2e_1 + e_2 - e_3 + e_4) = 2e_1 + e_2 - e_3 + e_4 \in U$ , e ciò comporta che  $U$  è invariante. Analogamente abbiamo:  $f(e_1 - e_3) = 0 \in W$ ,  $f(e_2) = e_1 - e_3 \in W$  e  $f(e_4) = e_2 \in W$ , quindi anche  $W$  è invariante. Andiamo a studiare gli operatori indotti per restrizione. Studiare l'operatore indotto su  $U$  significa calcolare la matrice rappresentativa di  $f_U$  rispetto alla base  $\mathcal{B}_U$ . Nel nostro caso  $M(f_U; \mathcal{B}_U, \mathcal{B}_U) = \begin{pmatrix} 1 \end{pmatrix}$*

(in particolare  $f_U = id_U$ ). Invece  $M(f_W; \mathcal{B}_W, \mathcal{B}_W) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

(quindi se  $(y', z', t')$  sono le coordinate di  $W$  rispetto alla base  $\mathcal{B}_W$  si ha  $\psi(y', z', t') = (z', t', 0)$ ). Quindi:

$$M(f; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Quindi se  $(x', y', z', t')$  sono le coordinate di  $\mathbb{C}^4$  rispetto alla base  $\mathcal{B}$  si ha  $f(x', y', z', t') = (x', z', t', 0)$ , ed inoltre  $p_f(x) = p_{f_U}(x) \cdot p_{f_W}(x) = (x-1)x^3$ .

### 3. BLOCCHI DI JORDAN E STRINGHE DI UN OPERATORE LINEARE

**Definizione 3.1.** Fissiamo un numero complesso  $\lambda \in \mathbb{C}$  ed un intero  $p \in \mathbb{N}$ . Si definisce **blocco di Jordan di ordine  $p$  relativo all'autovalore  $\lambda$**  la matrice quadrata  $J_{\lambda,p}$  di ordine  $p$  che ha sulla diagonale principale tutte le entrate uguali a  $\lambda$ , sopra la diagonale principale tutte le entrate uguali ad 1, e tutte le rimanenti entrate nulle.

**Esempio 3.2.**

$$J_{\lambda,1} = \begin{pmatrix} \lambda \end{pmatrix}, \quad J_{\lambda,2} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad J_{\lambda,3} = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix},$$

$$J_{\lambda,4} = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}.$$

Si osservi che  $J_{\lambda,p} = \lambda I_p + J_{0,p}$ . Quindi, come diremo meglio in seguito, lo studio di un blocco di Jordan si può sempre ricondurre a quello di un blocco relativo all'autovalore 0. Il polinomio caratteristico di  $J_{\lambda,p}$  è  $p_{J_{\lambda,p}}(x) = (-1)^p(x-\lambda)^p$ . Dunque  $J_{\lambda,p}$  ha soltanto l'autovalore  $\lambda$ . Inoltre, quando  $p \geq 2$   $J_{\lambda,p}$  non è diagonalizzabile. Infatti  $a_\lambda = p$  mentre  $g_\lambda = p - rk(J_{\lambda,p} - \lambda I_p) = p - (p-1) = 1$ .

La nozione di blocco di Jordan corrisponde, nell'ambito degli operatori lineari, a quella di **stringa**:

**Definizione 3.3.** Se  $f : V \rightarrow V$  è un operatore lineare ed  $\mathcal{C} = \{u_1, \dots, u_p\}$  è un sistema di  $p$  vettori di  $V$ , allora  $\mathcal{C}$  si dice che è una  $\lambda$ -stringa di lunghezza  $p$  per  $f$  se  $u_1 \neq 0$  ed inoltre:

$$\begin{cases} f(u_p) = \lambda u_p + u_{p-1} \\ f(u_{p-1}) = \lambda u_{p-1} + u_{p-2} \\ \dots \\ f(u_1) = \lambda u_1. \end{cases}$$

Il vettore  $u_p$  si dice **leader** della stringa.

Possiamo rappresentare una stringa nel seguente modo:

$$u_p \xrightarrow{f-\lambda id_V} u_{p-1} \xrightarrow{f-\lambda id_V} u_{p-2} \xrightarrow{f-\lambda id_V} \dots \xrightarrow{f-\lambda id_V} u_1 \xrightarrow{f-\lambda id_V} 0.$$

**Esercizio 3.4.** *Mostrare che la condizione  $u_1 \neq 0$  comporta che i vettori di una stringa siano linearmente indipendenti.*

*Osservazione 3.5.* 1. Il vettore  $u_1 \neq 0$  è un autovettore relativo all'autovalore  $\lambda$ .

2. Una stringa di lunghezza 1 altro non è che un autovettore.

**Esercizio 3.6.** *Sia  $f : V \rightarrow V$  un operatore lineare e  $\mathcal{C}$  un sistema di vettori di  $V$ . Allora  $\mathcal{C}$  è una  $\lambda$ -stringa di lunghezza  $p$  per  $f$  se e solo se  $\mathcal{C}$  è linearmente indipendente,  $U = L(\mathcal{C})$  è invariante per  $f$ , e la matrice  $M(f_U; \mathcal{C}, \mathcal{C})$  è il blocco di Jordan  $J_{\lambda, p}$ .*

**Esempio 3.7.** 1. *Nell'esempio 2.4 sia  $\mathcal{B}_U$  che  $\mathcal{B}_W$  sono stringhe per  $f$ . La prima è una stringa di lunghezza 1 relativa all'autovalore 1, mentre la seconda è una stringa di lunghezza 3 relativa all'autovalore 0;*

2. *La base canonica di  $\mathbb{C}^4$  è una stringa per l'operatore  $f(x, y, z, t) = (\lambda x + y, \lambda y + z, \lambda z + t, \lambda t)$ .*

#### 4. IL TEOREMA DI JORDAN PER UN OPERATORE NILPOTENTE

**Definizione 4.1.** Un operatore  $f \in \text{End}V$  si dice **nilpotente** se esiste un esponente  $p \in \mathbb{N}$  tale che  $f^p = 0$ . Il minimo di tali esponenti si dice **indice di nilpotenza** di  $f$ .

**Esercizio 4.2.** *L'unico autovalore di un operatore nilpotente è lo 0.*



1. I vettori  $\{u_1, \dots, u_k\} \subset U$  si dicono **indipendenti modulo**  $W$  se sono indipendenti e se  $L(\{u_1, \dots, u_k\}) \cap W = \{0\}$ :

$$\lambda_1 u_1 + \dots + \lambda_k u_k \in W \iff \lambda_1 = \dots = \lambda_k = 0.$$

2. I vettori  $\{u_1, \dots, u_k\} \subset U$  si dicono **una base di  $U$  modulo  $W$**  se, inoltre,  $U = L(\{u_1, \dots, u_k\}) \oplus W$  (in altri termini  $L(\{u_1, \dots, u_k\})$  e  $W$  sono **complementari** in  $U$ ).

*Osservazione 4.7.* Se i vettori  $\{u_1, \dots, u_k\} \subset U$  sono una base di  $U$  modulo  $W$  allora si otterrà una base di  $U$  aggiungendo ad essi una qualunque base di  $W$ .

**Lemma 4.8.** *Sia  $f$  un operatore nilpotente di indice  $p$ , se i vettori  $\{u_1, \dots, u_k\} \subset \text{Ker } f^i$  ( $i \leq p$ ) sono indipendenti modulo  $\text{Ker } f^{i-1}$  allora i vettori  $\{f(u_1), \dots, f(u_k)\} \subset \text{Ker } f^{i-1}$  sono indipendenti modulo  $\text{Ker } f^{i-2}$ .*

**Dimostrazione:** Se  $u \in \text{Ker } f^i$  allora  $f^{i-1}f(u) = f^i(u) = 0$ , quindi  $f(u) \in \text{Ker } f^{i-1}$ . Inoltre:

$$\begin{aligned} \lambda_1 f(u_1) + \dots + \lambda_k f(u_k) \in \text{Ker } f^{i-2} &\iff f^{i-2}(\lambda_1 f(u_1) + \dots + \lambda_k f(u_k)) = 0 \\ \iff f^{i-1}(\lambda_1 u_1 + \dots + \lambda_k u_k) = 0 &\iff \lambda_1 u_1 + \dots + \lambda_k u_k \in \text{Ker } f^{i-1} \iff \\ &\iff \lambda_1 = \dots = \lambda_k = 0. \end{aligned}$$

c.v.d.

**Teorema 4.9. (Teorema di Jordan per un operatore nilpotente)** *Sia  $f : V \rightarrow V$  un operatore nilpotente (di indice  $p$ ) definito su uno spazio vettoriale complesso  $V$ . Allora esiste una base  $\mathcal{B}$  di  $V$  formata da stringhe, detta base a stringhe, tale che la matrice rappresentativa  $J = M(f; \mathcal{B}, \mathcal{B})$  è una matrice diagonale a blocchi, con blocchi di Jordan relativi all'autovalore nullo. Tale matrice si chiama la **forma canonica di Jordan** di  $f$ . Essa è unica a meno dell'ordine con cui figurano i blocchi. La molteplicità geometrica dell'autovalore nullo è uguale al numero dei blocchi che appaiono in  $J$ .*

**Dimostrazione:** Poniamo

$$n_i := \dim \text{Ker } f^i, \quad q_i = n_i - n_{i-1}, \quad 1 \leq i \leq p \quad (n_0 = 0).$$

Vogliamo definire opportunamente delle basi  $\{u_1^i, \dots, u_{q_i}^i\} \subset \text{Ker } f^i$  di  $\text{Ker } f^i$  modulo  $\text{Ker } f^{i-1}$ . Procediamo per induzione decrescente a

partire da  $V = Ker f^p$ , scegliendo una base  $\{u_1^p, \dots, u_{q_p}^p\} \subset Ker f^p$  di  $Ker f^p$  modulo  $Ker f^{p-1}$ . Tali vettori rappresenteranno i leaders delle stringhe di lunghezza  $p$ .

Definiamo poi:

$$u_k^{p-1} := f(u_k^p) \in Ker f^{p-1}, \quad 1 \leq k \leq q_p.$$

In virtù del lemma 4.8, tali vettori sono indipendenti modulo  $Ker f^{p-2}$  e possono quindi essere completati ad una base  $\{u_1^{p-1}, \dots, u_{q_{p-1}}^{p-1}\} \subset Ker f^{p-1}$  di  $Ker f^{p-1}$  modulo  $Ker f^{p-2}$ . I vettori “aggiunti” rappresenteranno i leaders delle stringhe di lunghezza  $p - 1$  (e sono in numero  $q_{p-1} - q_p \geq 0$ ).

Più in generale, avendo definito una base  $\{u_1^{i+1}, \dots, u_{q_{i+1}}^{i+1}\} \subset Ker f^{i+1}$  di  $Ker f^{i+1}$  modulo  $Ker f^i$ , poniamo:

$$u_k^i := f(u_k^{i+1}) \in Ker f^i, \quad 1 \leq k \leq q_{i+1}.$$

In virtù del lemma 4.8, tali vettori sono indipendenti modulo  $Ker f^{i-1}$  e possono quindi essere completati ad una base  $\{u_1^i, \dots, u_{q_i}^i\} \subset Ker f^i$  di  $Ker f^i$  modulo  $Ker f^{i-1}$ . I vettori “aggiunti” rappresenteranno i leaders delle stringhe di lunghezza  $i$  (e sono in numero  $q_i - q_{i+1} \geq 0$ ).

Alla fine del procedimento si otterrà una base di autovettori  $\{u_1^1, \dots, u_{q_1}^1\} \subset Ker f$  (si osservi che  $q_1 = n_1$ ). Per l'osservazione 4.7, l'unione  $\{u_1^2, \dots, u_{q_2}^2\} \cup \{u_1^1, \dots, u_{q_1}^1\}$  rappresenta una base di  $Ker f^2$  e, più in generale,

$$\bigcup_{k=1}^i \{u_1^k, \dots, u_{q_k}^k\}$$

rappresenta una base di  $Ker f^i$ . I vettori

$$\bigcup_{k=1}^p \{u_1^k, \dots, u_{q_k}^k\}$$

rappresentano quindi una base a stringhe di  $Ker f^p = V$ . c.v.d

*Osservazione 4.10.* Un operatore è diagonalizzabile se e solo se tutti i blocchi che appaiono nella forma canonica hanno ordine 1.

**Esercizio 4.11.** *Trovare una base a stringhe per le seguenti matrici nilpotenti:*

$$\begin{pmatrix} 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 0 & 1 & 0 & -1 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}.$$

## 5. IL TEOREMA DI JORDAN SULLA FORMA CANONICA

Sia  $f$  un operatore definito in  $V$ ,  $\lambda$  un autovalore di  $f$  e  $V_\lambda = \text{Ker}(f - \lambda id)$  l'autospazio corrispondente.

Come nell'esercizio 4.4, se

$$\text{Ker}(f - \lambda id)^i = \text{Ker}(f - \lambda id)^{i+1}$$

allora

$$\text{Ker}(f - \lambda id)^{i+1} = \text{Ker}(f - \lambda id)^{i+2}.$$

**Definizione 5.1.** Se  $n_\lambda$  è il minimo tra gli esponenti  $i$  che verificano  $\text{Ker}(f - \lambda id)^i = \text{Ker}(f - \lambda id)^{i+1}$ , allora il sottospazio:

$$\tilde{V}_\lambda := \text{Ker}(f - \lambda id)^{n_\lambda},$$

si dice **autospazio generalizzato** associato all'autovalore  $\lambda$ .

Avremo dunque una catena di inclusioni strette:

(2)

$$V_\lambda = \text{Ker}(f - \lambda id) \subset \text{Ker}(f - \lambda id)^2 \subset \dots \subset \text{Ker}(f - \lambda id)^{n_\lambda} = \tilde{V}_\lambda.$$

*Osservazione 5.2.* Per il teorema delle dimensioni, avremo anche la seguente sequenza di inclusioni strette:

$$(3) \quad \text{Im}(f - \lambda id) \supset \text{Im}(f - \lambda id)^2 \supset \dots \supset \text{Im}(f - \lambda id)^{n_\lambda}.$$

**Esercizio 5.3.** *Osservare che*

$$f(f - \lambda id)^i = (f - \lambda id)^i f, \quad \forall i \in \mathbb{N}.$$

**Lemma 5.4.** *I sottospazi  $\tilde{V}_\lambda$  e  $Im(f - \lambda id)^{n_\lambda}$  sono invarianti e complementari:*

$$V = \tilde{V}_\lambda \oplus Im(f - \lambda id)^{n_\lambda}.$$

**Dimostrazione.** Dimostriamo che sono invarianti:

$$\begin{aligned} v \in Ker(f - \lambda id)^{n_\lambda} &\implies (f - \lambda id)^{n_\lambda}(f(v)) = f((f - \lambda id)^{n_\lambda}(v)) = 0 \implies \\ &\implies f(v) \in Ker(f - \lambda id)^{n_\lambda}; \end{aligned}$$

$$\begin{aligned} v \in Im(f - \lambda id)^{n_\lambda} &\implies v = (f - \lambda id)^{n_\lambda}(u) \implies \\ f(v) &= f((f - \lambda id)^{n_\lambda}(u)) = (f - \lambda id)^{n_\lambda}(f(u)) \implies f(v) \in Im(f - \lambda id)^{n_\lambda}. \end{aligned}$$

Per concludere, ricordando il teorema delle dimensioni, è sufficiente dimostrare:

$$\tilde{V}_\lambda \cap Im(f - \lambda id)^{n_\lambda} = \{0\}.$$

Sia  $v \in \tilde{V}_\lambda \cap Im(f - \lambda id)^{n_\lambda}$ . Poiché  $v \in Im(f - \lambda id)^{n_\lambda}$ , esiste  $u \in V$  tale che  $v = (f - \lambda id)^{n_\lambda}(u)$ . Poiché  $v \in \tilde{V}_\lambda$ ,

$$0 = (f - \lambda id)^{2n_\lambda}(v) = (f - \lambda id)^{2n_\lambda}(u),$$

quindi  $u \in Ker(f - \lambda id)^{2n_\lambda} = Ker(f - \lambda id)^{n_\lambda}$  e  $v = (f - \lambda id)^{n_\lambda}(u) = 0$ .  
c.v.d.

Siamo in condizione di enunciare il il seguente Teorema, che è il risultato principale di questo capitolo.

**Teorema 5.5. (Teorema di Jordan)** *Sia  $f : V \rightarrow V$  un operatore lineare definito su uno spazio vettoriale complesso  $V$ . Allora esiste una base  $\mathcal{B}$  di  $V$ , formata da stringhe, e detta base a stringhe, tale che la matrice rappresentativa  $J = M(f; \mathcal{B}, \mathcal{B})$  è una matrice diagonale a blocchi, con blocchi di Jordan. Tale matrice si chiama la **forma canonica di Jordan** di  $f$ . Essa è unica a meno dell'ordine con cui figurano i blocchi. Gli autovalori di  $f$  coincidono con gli autovalori dei blocchi di  $J$ . Per ciascun autovalore  $\lambda$  di  $f$ ,  $g_\lambda$  è uguale al numero dei blocchi relativi a  $\lambda$  che appaiono in  $J$ , mentre  $a_\lambda$  è uguale alla somma delle grandezze dei blocchi relativi a  $\lambda$  che appaiono in  $J$  e coincide con la dimensione dell'autospazio generalizzato  $\tilde{V}_\lambda$ .*

**Dimostrazione.** Per il lemma 5.4, la restrizione di  $f$  al sottospazio invariante  $Im(f - \lambda id)^{n_\lambda}$  non ammette  $\lambda$  come autovalore. Inoltre, se  $u \in V_\mu$  è un autovettore con autovalore  $\mu \neq \lambda$ , allora

$$(f - \lambda id)(u) = (\mu - \lambda)u$$



$a_{-4} = 1$ ,  $g_{-4} = 1$ . Infine il polinomio caratteristico di  $f$  e'  $p_f(x) = -x^9(x - \sqrt{5})^6(x + 4)$ .

**Esercizio 5.8.** Determinare una base a stringhe per i seguenti endomorfismi:

1. in  $\mathbb{C}^3$ :  $f(x, y, z) = (-4x + z, -2x - 3y + 2z, -x - 2z)$ ;
2. in  $\mathbb{C}^4$ :  $f(x, y, z, t) = (x + y + z, t, -y, x + z)$ ;
3. in  $\mathbb{C}^4$ :  $f(x, y, z, t) = (2z + t, 2y + z + t, 2z + t, 2t)$ .

*Osservazione 5.9.* E' possibile determinare la forma canonica di Jordan **senza conoscere una base a stringhe**. Infatti le dimensioni degli autospazi generalizzati sono individuate dal polinomio caratteristico. Inoltre, usando le notazioni del teorema 4.9 e riferendole ai singoli autospazi generalizzati, ogni differenza  $q_i - q_{i+1}$  rappresenta il numero di stringhe di lunghezza  $i$ .

**Esercizio 5.10.** 1. Determinare la forma canonica di Jordan della matrice

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

senza ricercare una base a stringhe.

2. Sapendo che una matrice  $A \in \mathbb{C}^{10,10}$  ha indice di nilpotenza  $p = 4$  e sapendo che  $A^2$  ha rango 3, calcolare le possibili forme canoniche per  $A$ .

## 6. TEOREMA DI CAYLEY-HAMILTON

Una interessante conseguenza del Teorema di Jordan è il famoso Teorema di Cayley-Hamilton. Prima di enunciare tale teorema ci occorre qualche nozione preliminare.

Innanzitutto fissiamo un polinomio  $\varphi(t) = a_0 + a_1t + \dots + a_h t^h$ . Assegnato un numero  $c \in \mathbb{C}$  possiamo valutare il polinomio in  $c$ . Si ottiene un altro numero, che si indica con  $\varphi(c)$ . Ad esempio, se  $\varphi(t) = 3 + t - 2t^2$  allora  $\varphi(-2) = -7$ . In modo analogo possiamo valutare il polinomio in una matrice quadrata  $A$ , ponendo

$$\varphi(A) := a_0I + a_1A + \dots + a_h A^h.$$

E, assegnato un operatore  $f : V \rightarrow V$ , possiamo definire

$$\varphi(f) := a_0 id_V + a_1 f + \cdots + a_h f^h$$

come quell'operatore  $\varphi(f) : V \rightarrow V$  che al vettore  $u \in V$  associa il vettore

$$\varphi(f)(u) = a_0 u + a_1 f(u) + \cdots + a_h f^h(u).$$

Si osservi che se  $A$  è una matrice che rappresenta  $f$  rispetto a qualche base, allora  $\varphi(A)$  rappresenta  $\varphi(f)$  rispetto alla stessa base. Inoltre, comunque si assegnino polinomi  $\varphi(t)$  e  $\psi(t)$ , una matrice  $A$  ed un operatore  $f$ , valgono le seguenti proprietà di calcolo:

$$(\varphi + \psi)(A) = \varphi(A) + \psi(A), \quad (\varphi \cdot \psi)(A) = \varphi(A) \cdot \psi(A) = \psi(A) \cdot \varphi(A)$$

$$(\varphi + \psi)(f) = \varphi(f) + \psi(f), \quad (\varphi \cdot \psi)(f) = \varphi(f) \cdot \psi(f) = \psi(f) \cdot \varphi(f).$$

Sia  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  una qualunque matrice  $2 \times 2$ . Il polinomio caratteristico di  $A$  è:

$$p_A(x) = \det \begin{pmatrix} a-x & b \\ c & d-x \end{pmatrix} = ad - bc - (a+d)x + x^2.$$

Si trova:

$$\begin{aligned} p_A(A) &= (ad - bc) \cdot I - (a+d)A + A^2 = \\ &= (ad - bc) \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - (a+d) \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 = \\ &= \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} - \begin{pmatrix} a^2 + ad & ab + db \\ ac + dc & ad + d^2 \end{pmatrix} + \begin{pmatrix} a^2 + bc & ab + bd \\ ac + dc & bc + d^2 \end{pmatrix} = \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Più in generale, vale il seguente:

**Teorema 6.1. (di Cayley-Hamilton)** Sia  $f : V \rightarrow V$  un operatore lineare definito su uno spazio vettoriale complesso  $V$ . Sia  $p_f(x)$  il polinomio caratteristico di  $f$ . Allora  $p_f(f)$  è l'operatore nullo.

**Dimostrazione.** la dimostrazione si può vedere come una applicazione del Teorema di Jordan. Più precisamente come una conseguenza della decomposizione di  $V$  come somma dei suoi autospazi generalizzati.

Siano  $\lambda_1, \lambda_2, \dots, \lambda_k$  gli autovalori distinti di  $f$ , e

$$p_f(X) = (-1)^n (x - \lambda_1)^{a_1} \cdot (x - \lambda_2)^{a_2} \cdot \dots \cdot (x - \lambda_k)^{a_k}$$

la corrispondente decomposizione di  $p_f(x)$ , dove  $a_i = a_{\lambda_i}$ . Consideriamo anche la decomposizione di  $V$  nella somma diretta dei suoi autospazi generalizzati:

$$V = \tilde{V}_{\lambda_1} \oplus \dots \oplus \tilde{V}_{\lambda_k}.$$

Sia  $u$  un qualunque vettore di  $V$ . Noi vogliamo provare che  $p_f(f)(u) = 0$ . In virtù della decomposizione in autospazi generalizzati, possiamo scrivere

$$u = u_1 + \dots + u_k$$

secondo opportuni (ed unici) vettori  $u_i \in \tilde{V}_{\lambda_i}$ . Poiché  $p_f(f)(u) = p_f(f)(u_1) + \dots + p_f(f)(u_k)$  sarà sufficiente provare che  $p_f(f)(u_i) = 0$ , per ogni  $i = 1, \dots, k$ .

Adesso, tenendo presente che  $\tilde{V}_{\lambda_i} = \ker(f - \lambda_i \text{id})^{a_i}$ , si ha:

$$\begin{aligned} p_f(f)(u_i) &= (f - \lambda_1 \text{id})^{a_1} \cdot (f - \lambda_2 \text{id})^{a_2} \cdot \dots \cdot (f - \lambda_k \text{id})^{a_k}(u_i) \\ &= \left( \prod_{j \neq i} (f - \lambda_j \text{id})^{a_j} \right) \cdot (f - \lambda_i \text{id})^{a_i}(u_i) = 0. \end{aligned}$$

Ciò conclude la dimostrazione del Teorema di Cayley-Hamilton. c.v.d.



è una soluzione del sistema:

$$\begin{cases} \dot{x}_1(t) = 3x_1(t) + t \\ \dot{x}_2(t) = x_2(t) + 1 \\ \dot{x}_3(t) = x_2(t) + x_3(t). \end{cases}$$

Fissato un istante  $t_0 \in I$  (che generalmente sarà  $t_0 = 0$ ) ed un vettore costante  $c \in \mathbb{R}^n$ , una scrittura del tipo

$$(2) \quad \begin{cases} \dot{x}(t) = Ax(t) + f(t) \\ x(t_0) = c \end{cases}$$

dicesi **problema di Cauchy**. Una soluzione di tale problema 2 è una soluzione  $y(t)$  del problema 1 che in più soddisfa la condizione iniziale  $y(t_0) = c$ . Vale il seguente importante teorema:

**Teorema 1.2. (Teorema di esistenza ed unicità)** *Assegnato un problema di Cauchy, esso ammette un'unica soluzione.*

**Esempio 1.3.** 1. *La funzione vettoriale  $y(t) := (\frac{e^{3t}}{9} - \frac{t}{3} - \frac{1}{9}, -1, 1)$  è l'unica soluzione del problema di Cauchy:*

$$\begin{cases} \dot{x}_1(t) = 3x_1(t) + t \\ \dot{x}_2(t) = x_2(t) + 1 \\ \dot{x}_3(t) = x_2(t) + x_3(t) \\ x(0) = (0, -1, 1)^t. \end{cases}$$

2. *La funzione identicamente nulla  $y(t) \equiv 0$  è l'unica soluzione del problema di Cauchy:*

$$\begin{cases} \dot{x}(t) = Ax(t) \\ x(0) = 0. \end{cases}$$

Consideriamo nuovamente il sistema 1. L'equazione:

$$(3) \quad \dot{x}(t) = Ax(t)$$

si dice **equazione omogenea associata alla 1**. Sia  $\mathcal{V}$  l'insieme delle soluzioni di 3, definite nell'intervallo  $I$ . Se  $y(t)$  e  $z(t)$  sono due soluzioni di 3 e  $c$  è uno scalare, allora  $y(t) + z(t)$  e  $cy(t)$  sono ancora soluzioni

dell'equazione omogenea. In altre parole  $\mathcal{V}$  ha una naturale struttura di spazio vettoriale. Come conseguenza del Teorema di esistenza ed unicità possiamo provare che  $\mathcal{V}$  ha dimensione  $n$ , dove  $n$  è l'ordine della matrice  $A$ . Infatti, consideriamo un vettore  $c \in \mathbb{R}^n$  e indichiamo con  $y_c(t)$  l'unica soluzione di 3 che soddisfa la condizione iniziale  $y_c(0) = c$ . Risulta pertanto ben definita l'applicazione:

$$\psi : c \in \mathbb{R}^n \rightarrow y_c(t) \in \mathcal{V}.$$

L'applicazione  $\psi$  è biiettiva e lineare. Ne risulta che  $\mathcal{V}$  è isomorfo ad  $\mathbb{R}^n$  e quindi  $\mathcal{V}$  ha dimensione  $n$ . La funzione  $\psi$  dicesi **integrale generale** dell'equazione omogenea 3. Un altro modo di presentare l'integrale generale è il seguente. Detta  $\{y_1(t), \dots, y_n(t)\}$  una base per  $\mathcal{V}$  allora ogni altra soluzione  $y(t)$  di (3) si scrive come combinazione lineare

$$(4) \quad y(t) = c_1 y_1(t) + c_2 y_2(t) + \dots + c_n y_n(t), \quad c_i \in \mathbb{R}.$$

Anche la scrittura 4 prende il nome di integrale generale di 3. Si osservi che le funzioni  $y_i(t)$  si possono ottenere risolvendo, per ogni  $i = 1, \dots, n$ , il problema di Cauchy

$$\begin{cases} \dot{x}(t) = Ax(t) \\ x(0) = e_i \end{cases}$$

dove  $e_i$  denota l' $i$ -esimo vettore canonico di  $\mathbb{R}^n$ . Tale base, ottenuta in corrispondenza della base canonica di  $\mathbb{R}^n$ , sarà ancora chiamata **base canonica di  $\mathcal{V}$** . Con la scelta della base canonica per  $\mathcal{V}$ , la soluzione  $y(t)$  che appare nella 4 assume all'istante iniziale proprio il valore  $c = (c_1, c_2, \dots, c_n)$ , cioè  $y(0)$  coincide con il vettore delle coordinate di  $y(t)$  rispetto alla base canonica di  $\mathcal{V}$ .

Ritornando al problema 1, sia  $u(t)$  una qualunque soluzione di 1.

**Esercizio 1.4.** *Provare che tutte e sole le soluzioni di 1 si ottengono sommando a  $u(t)$  le soluzioni dell'equazione omogenea associata 3.*

L'integrale generale del problema 1 assume la forma:

$$(5) \quad y(t) = u(t) + c_1 y_1(t) + c_2 y_2(t) + \dots + c_n y_n(t), \quad c_i \in \mathbb{R}.$$

In altre parole per determinare tutte e sole le soluzioni di 1 occorre conoscerne almeno una, tutte le altre si ottengono sommando a questa

le soluzioni dell'equazione omogenea associata 3. Come nel caso omogeneo, se si sceglie  $u(t)$  la soluzione che vale 0 all'istante iniziale e la base canonica di  $\mathcal{V}$ , allora, nell'integrale generale 5,  $y(0)$  coincide con le coordinate di  $y(t) - u(t)$  rispetto alla base canonica di  $\mathcal{V}$ .

Per risolvere esplicitamente l'equazione 1, cioè per trovare l'integrale generale, si possono usare due metodi: la forma canonica di Jordan, oppure la Trasformata di Laplace. Cominceremo a vedere come si usa la forma canonica.

## 2. ESPONENZIALE E FORMA CANONICA

Ricordiamo la serie esponenziale di una matrice quadrata  $A$ :

$$e^A := \sum_{h=0}^{+\infty} \frac{A^h}{h!} := I + \frac{A}{1!} + \frac{A^2}{2!} + \cdots + \frac{A^h}{h!} + \cdots$$

Vediamo come si calcola esplicitamente l'esponenziale, utilizzando la forma canonica. Poichè saremo interessati a matrici del tipo  $tA$  (cioè a matrici moltiplicate per un parametro  $t$ ), calcoleremo l'esponenziale della matrice  $tA$  (che si riduce all'esponenziale di  $A$  per  $t = 1$ ). In tale contesto si usa mettere lo scalare  $t$  a destra, cioè scriveremo  $At$  invece di  $tA$ . Sia  $J$  la forma canonica di Jordan di  $A$ . Sappiamo allora che esiste qualche matrice invertibile  $P$  tale che  $A = PJP^{-1}$ . Quindi

$$(6) \quad e^{At} = \sum_{h=0}^{+\infty} \frac{A^h t^h}{h!} = \sum_{h=0}^{+\infty} \frac{(PJP^{-1})^h t^h}{h!} = P \left( \sum_{h=0}^{+\infty} \frac{J^h t^h}{h!} \right) P^{-1} = P e^{Jt} P^{-1}.$$

Tale formula ci mostra che **per calcolare l'esponenziale di  $At$  è sufficiente conoscere una base a stringhe per  $A$  (cioè le colonne della matrice  $P$ ), e saper calcolare l'esponenziale di una matrice del tipo  $Jt$ , dove  $J$  è una matrice a blocchi di Jordan.** Nel capitolo precedente abbiamo imparato come si calcola  $P$ . Ci rimane solo da vedere come si calcola  $e^{Jt}$ . Innanzitutto osserviamo che se  $J_1, \dots, J_k$  sono i blocchi di Jordan di cui è costituita  $J$  allora dalla

stessa definizione di matrice esponenziale segue che:

$$(7) \quad e^{Jt} = \begin{pmatrix} e^{J_1 t} & 0 & \dots & \dots & 0 \\ 0 & e^{J_2 t} & \dots & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \dots & e^{J_k t} \end{pmatrix}.$$

In altre parole **la matrice esponenziale  $e^{Jt}$ , con  $J$  matrice a blocchi di Jordan, è una matrice a blocchi, con blocchi dati dalle matrici esponenziali dei blocchi di  $Jt$** . Quindi per calcolare  $e^{Jt}$  ci si può ricondurre al caso in cui  $J$  sia un singolo blocco di Jordan. Supponiamo allora che

$$J = \lambda I + N$$

sia un blocco di Jordan di ordine  $p$  con autovalore  $\lambda$ . Qui  $N$  rappresenta il blocco di Jordan relativo all'autovalore 0. Poiché  $\lambda I$  ed  $N$  commutano, allora  $e^{Jt} = e^{I\lambda t} e^{Nt}$ . Poiché  $e^{I\lambda t}$  è la matrice diagonale con  $e^{\lambda t}$  sulla diagonale principale, allora

$$e^{Jt} = e^{\lambda t} e^{Nt}.$$

Quindi per calcolare  $e^{Jt}$  è sufficiente saper calcolare  $e^{Nt}$ . Ma  $N$  è nilpotente con indice di nilpotenza  $p$ , per cui

$$e^{Nt} = I + Nt + \frac{N^2 t^2}{2!} + \dots + \frac{N^{p-1} t^{p-1}}{(p-1)!}.$$

Questa matrice si calcola facilmente (si vedano gli esempi successivi), e ciò conclude il calcolo esplicito della matrice esponenziale tramite la forma canonica di Jordan.

**Esempio 2.1.** *Facciamo degli esempi su come si calcola l'esponenziale di un singolo blocco.*

$$e^{\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} t} = e^{\lambda t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & t e^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix};$$

$$e^{\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} t} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & t e^{\lambda t} & \frac{t^2}{2} e^{\lambda t} \\ 0 & e^{\lambda t} & t e^{\lambda t} \\ 0 & 0 & e^{\lambda t} \end{pmatrix};$$

$$e^{\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix} t} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \frac{t^3}{6} \\ 0 & 1 & t & \frac{t^2}{2} \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2}e^{\lambda t} & \frac{t^3}{6}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2}e^{\lambda t} \\ 0 & 0 & e^{\lambda t} & te^{\lambda t} \\ 0 & 0 & 0 & e^{\lambda t} \end{pmatrix}.$$

In generale, se  $J$  è il blocco di Jordan di ordine  $p$  relativo all'autovalore  $\lambda$ ,  $e^{Jt}$  è quella matrice quadrata di ordine  $p$  che sulla diagonale  $a_{1j}, a_{2,j+1}, \dots, a_{p-j+1,p}$  ha tutte le entrate uguali a  $e^{\lambda t} \frac{t^{j-1}}{(j-1)!}$ .

**Esempio 2.2.** Calcoliamo l'esponenziale della matrice  $At$ , dove

$$A := \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Una base a stringhe per  $A$  è formata dai vettori

$\mathcal{S} := \{(1, 0, 0), (0, 0, 1), (0, 1, 0)\}$ . Quindi  $A = PJP^{-1}$ , dove  $P$  è la matrice che ha per colonne i vettori di  $\mathcal{S}$ , e  $J$  è la matrice a blocchi di Jordan:

$$J := \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

In base alle formule precedenti possiamo dire che:

$$\begin{aligned} e^{At} = Pe^{Jt}P^{-1} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} e^{3t} & 0 & 0 \\ 0 & e^t & te^t \\ 0 & 0 & e^t \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \\ &= \begin{pmatrix} e^{3t} & 0 & 0 \\ 0 & e^t & 0 \\ 0 & te^t & e^t \end{pmatrix}. \end{aligned}$$

### 3. CALCOLO ESPlicito DELL'INTEGRALE GENERALE

Ritornando ai sistemi di equazioni differenziali, andiamo a vedere come si applicano le considerazioni precedenti sull'esponenziale per il calcolo esplicito dell'integrale generale. A tale proposito vale il seguente:

**Teorema 3.1.** *Il problema di Cauchy:*

$$\begin{cases} \dot{x}(t) = Ax(t) \\ x(0) = c \end{cases}$$

ammette come unica soluzione la funzione

$$y(t) = e^{At}c.$$

In particolare la base canonica per lo spazio delle soluzioni dell'equazione omogenea  $\dot{x}(t) = Ax(t)$  è data dalle colonne della matrice  $e^{At}$ . Inoltre la soluzione particolare  $u(t)$  dell'equazione  $\dot{x}(t) = Ax(t) + f(t)$  soddisfacente la condizione iniziale  $u(0) = 0$  è data dalla formula

$$u(t) = \int_0^t e^{A(t-s)} \cdot f(s) ds.$$

La dimostrazione del teorema precedente è immediata.

Vediamo con degli esempi come si applica questo teorema.

**Esempio 3.2.** *Calcoliamo l'integrale generale dell'equazione*

$$\begin{cases} \dot{x}_1 = 3x_1 + t \\ \dot{x}_2 = x_2 + 1 \\ \dot{x}_3 = x_2 + x_3, \end{cases}$$

cioè dell'equazione

$$\dot{x}(t) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \cdot x(t) + \begin{pmatrix} t \\ 1 \\ 0 \end{pmatrix}.$$

Dal Teorema 3.1 e dall'esempio precedente segue che l'integrale generale dell'equazione omogenea associata è

$$y(t) = c_1 \begin{pmatrix} e^{3t} \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ e^t \\ te^t \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ 0 \\ e^t \end{pmatrix}.$$

Poi sappiamo anche che la soluzione particolare  $u(t)$  si può calcolare nel seguente modo:

$$u(t) = \int_0^t e^{A(t-s)} \cdot f(s) ds = \int_0^t \begin{pmatrix} e^{3(t-s)} & 0 & 0 \\ 0 & e^{t-s} & 0 \\ 0 & (t-s)e^{t-s} & e^{t-s} \end{pmatrix} \cdot \begin{pmatrix} s \\ 1 \\ 0 \end{pmatrix} ds$$

$$= \int_0^t \begin{pmatrix} se^{3(t-s)} \\ e^{t-s} \\ (t-s)e^{t-s} \end{pmatrix} ds = \begin{pmatrix} \int_0^t se^{3(t-s)} ds \\ \int_0^t e^{t-s} ds \\ \int_0^t (t-s)e^{t-s} ds \end{pmatrix} = \begin{pmatrix} \frac{e^{3t}}{9} - \frac{t}{3} - \frac{1}{9} \\ e^t - 1 \\ 1 - e^t + te^t \end{pmatrix}.$$

In conclusione l'integrale generale dell'equazione assegnata è

$$y(t) = \begin{pmatrix} \frac{e^{3t}}{9} - \frac{t}{3} - \frac{1}{9} \\ e^t - 1 \\ 1 - e^t + te^t \end{pmatrix} + c_1 \begin{pmatrix} e^{3t} \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ e^t \\ te^t \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ 0 \\ e^t \end{pmatrix}.$$

**Esempio 3.3.** Risolviamo il seguente problema di Cauchy:

$$(8) \quad \begin{cases} \dot{x}_1 = 3x_1 + t \\ \dot{x}_2 = x_2 + 1 \\ \dot{x}_3 = x_2 + x_3, \end{cases} \quad \begin{cases} x_1(0) = 1 \\ x_2(0) = 1 \\ x_3(0) = 1. \end{cases}$$

Nell'esempio precedente abbiamo calcolato l'integrale generale, esplicitato tramite la base canonica dell'equazione omogenea associata, e la soluzione particolare  $u(t)$  con  $u(0) = 0$ . Sappiamo che in tal caso la soluzione del problema di Cauchy assegnato si ottiene ponendo nell'integrale generale  $(c_1, c_2, c_3) = (1, 1, 1)$ . Quindi il problema assegnato ammette la soluzione

$$y(t) = \begin{pmatrix} x \frac{10}{9} e^{3t} - \frac{1}{3}t - \frac{1}{9} \\ 2e^t - 1 \\ 2te^t + 1 \end{pmatrix}.$$

Possiamo anche verificare che il risultato è esatto. Infatti  $y(0) = (1, 1, 1)^t$ , ed inoltre si ha:

$$\begin{aligned} \dot{y}(t) &= \begin{pmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \\ \dot{y}_3(t) \end{pmatrix} = \\ &= \begin{pmatrix} \frac{10}{3}e^{3t} - \frac{1}{3} \\ 2e^t \\ 2(t+1)e^t \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{10}{9}e^{3t} - \frac{1}{3}t - \frac{1}{9} \\ 2e^t - 1 \\ 2te^t + 1 \end{pmatrix} + \begin{pmatrix} t \\ 1 \\ 0 \end{pmatrix}, \end{aligned}$$

cioè  $y(t)$  soddisfa la 8.

**Esempio 3.4.** Calcolare l'integrale generale del sistema:

$$\begin{cases} \dot{x} = -4y \\ \dot{y} = x. \end{cases}$$

Il polinomio caratteristico della matrice  $A$  del sistema è  $t^2 + 4 = (t - 2i)(t + 2i)$ . Quindi  $A$  è diagonalizzabile ed una base di autovettori è formata dai vettori  $(2i, 1)$ ,  $(2i, -1)$ . Pertanto, tenuto conto delle formule di Eulero, abbiamo:

$$\begin{aligned} e^{At} &= \frac{i}{4} \begin{pmatrix} 2i & 2i \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} e^{2ti} & 0 \\ 0 & e^{-2ti} \end{pmatrix} \cdot \begin{pmatrix} -1 & -2i \\ -1 & 2i \end{pmatrix} = \\ &= \begin{pmatrix} \cos 2t & -2\operatorname{sen} 2t \\ \frac{1}{2}\operatorname{sen} 2t & \cos 2t \end{pmatrix}. \end{aligned}$$

Quindi l'integrale generale richiesto è:

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1 \begin{pmatrix} \cos 2t \\ \frac{1}{2}\operatorname{sen} 2t \end{pmatrix} + c_2 \begin{pmatrix} -2\operatorname{sen} 2t \\ \cos 2t \end{pmatrix}.$$

**Esempio 3.5.** Risolvere il seguente problema di Cauchy:

$$\begin{cases} \dot{x} = -4y \\ \dot{y} = x \end{cases} \quad \begin{cases} x(0) = 1 \\ y(0) = -1. \end{cases}$$

Possiamo mettere  $c_1 = 1$  e  $c_2 = -1$  nell'integrale generale calcolato nell'esempio precedente. La soluzione cercata è:

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \cos 2t + 2\operatorname{sen} 2t \\ \frac{1}{2}\operatorname{sen} 2t - \cos 2t \end{pmatrix}.$$

#### 4. CENNI SULL'USO DELLA TRASFORMATA DI LAPLACE NELLA RISOLUZIONE DEI SISTEMI DI EQUAZIONI DIFFERENZIALI

Per risolvere un problema di Cauchy si può anche utilizzare la **Trasformata di Laplace**. A tale proposito consideriamo il problema di Cauchy:

$$\begin{cases} \dot{x}(t) = Ax(t) + f(t) \\ x(0) = c. \end{cases}$$

Applicando ad entrambi i membri la trasformazione di Laplace  $\mathcal{L}$  otteniamo:

$$sX(s) - c = AX(s) + F(s),$$

dove con  $X(s)$  e  $F(s)$  abbiamo denotato le trasformate di  $x(t)$  e di  $f(t)$ .

Si deduce la seguente formula:

$$X(s) = (sI - A)^{-1}(c + F(s))$$

che riconduce il calcolo della trasformata di Laplace della soluzione  $x(t)$  che stiamo cercando ad un calcolo algebrico. Ad esempio, nel caso omogeneo (cioè quando  $f(t) = 0$ ) per conoscere  $X(s)$  è sufficiente calcolare la matrice inversa  $(sI - A)^{-1}$ , che in generale si presenta come una matrice con entrate funzioni razionali di  $s$  (cioè rapporti di polinomi nella variabile  $s$ ). Una volta calcolata  $X(s)$ , per conoscere la soluzione del problema di Cauchy  $x(t)$  occorre calcolare l'antitrasformata di  $X(s)$ :

$$x(t) = \mathcal{L}^{-1}(X(s)).$$

Chiaramente è qui che si incontrano le maggiori difficoltà nei calcoli.

**Esempio 4.1.** *Come esempio, riprendiamo il seguente problema di Cauchy, che abbiamo già risolto utilizzando l'esponenziale di una matrice, e vediamo come si risolve utilizzando la Trasformata di Laplace:*

$$\begin{cases} \dot{x}_1 = 3x_1 + t \\ \dot{x}_2 = x_2 + 1 \\ \dot{x}_3 = x_2 + x_3, \end{cases} \quad \begin{cases} x_1(0) = 1 \\ x_2(0) = 1 \\ x_3(0) = 1. \end{cases}$$

Abbiamo appena visto che

$$\begin{aligned} X(s) &= \left( \begin{pmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & s \end{pmatrix} - \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \right)^{-1} \left( \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \mathcal{L} \begin{pmatrix} t \\ 1 \\ 0 \end{pmatrix} \right) \\ &= \begin{pmatrix} s-3 & 0 & 0 \\ 0 & s-1 & 0 \\ 0 & -1 & s-1 \end{pmatrix}^{-1} \left( \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{1}{s^2} \\ \frac{1}{s} \\ 0 \end{pmatrix} \right) = \begin{pmatrix} \frac{s^2+1}{s^2(s-3)} \\ \frac{s+1}{s(s-1)} \\ \frac{s^2+1}{s(s-1)^2} \end{pmatrix}. \end{aligned}$$

Adesso per calcolare la soluzione cercata  $x(t)$  occorre antitrasformare le entrate di  $X(s)$ , cioè occorre calcolare:

$$x(t) = \mathcal{L}^{-1} \begin{pmatrix} \frac{s^2+1}{s^2(s-3)} \\ \frac{s+1}{s(s-1)} \\ \frac{s^2+1}{s(s-1)^2} \end{pmatrix} = \begin{pmatrix} \mathcal{L}^{-1} \left( \frac{s^2+1}{s^2(s-3)} \right) \\ \mathcal{L}^{-1} \left( \frac{s+1}{s(s-1)} \right) \\ \mathcal{L}^{-1} \left( \frac{s^2+1}{s(s-1)^2} \right) \end{pmatrix}.$$

Sospendiamo momentaneamente lo svolgimento dell'esempio, che riprenderemo tra poco, per vedere come si calcola in generale l'antitrasformata di una funzione razionale. Tale calcolo si basa sulle seguenti osservazioni. Innanzitutto ricordiamo che è nota l'antitrasformata di una

funzione razionale "semplice", cioè del tipo

$$\frac{1}{(s - \lambda)^{n+1}},$$

dove  $n$  è un intero  $\geq 0$  e  $\lambda \in \mathbb{R}$  (una tale funzione è detta anche **fratto semplice**). Infatti si ha la seguente formula:

$$\mathcal{L}^{-1} \left( \frac{1}{(s - \lambda)^{n+1}} \right) = \frac{t^n}{n!} e^{\lambda t}.$$

L'altra osservazione consiste nella seguente proposizione che consente di ricondurre una qualunque funzione razionale ad una somma di fratti semplici.

**Teorema 4.2.** *Siano  $P(s)$  e  $Q(s)$  polinomi non nulli. Supponiamo che il grado di  $P(s)$  sia strettamente minore del grado di  $Q(s)$ . Sia*

$$Q(s) = c(s - \lambda_1)^{m_1} \cdots (s - \lambda_h)^{m_h}$$

la fattorizzazione di  $Q(s)$  in polinomi lineari. Allora esistono  $m_1 + m_2 + \cdots + m_h$  costanti opportune  $a_{11}, \dots, a_{1m_1}, a_{21}, \dots, a_{2m_2}, \dots, a_{m_h 1}, \dots, a_{m_h m_h}$  tali che valga la seguente uguaglianza:

$$\frac{P(s)}{Q(s)} = \frac{a_{11}}{s - \lambda_1} + \frac{a_{12}}{(s - \lambda_1)^2} + \cdots + \frac{a_{1m_1}}{(s - \lambda_1)^{m_1}} + \cdots + \frac{a_{m_h 1}}{s - \lambda_h} + \cdots + \frac{a_{m_h m_h}}{(s - \lambda_h)^{m_h}}.$$

L'espressione precedente è detta anche **decomposizione in fratti semplici di  $\frac{P(s)}{Q(s)}$** . Inoltre si osservi che l'ipotesi sui gradi di  $P(s)$  e  $Q(s)$  è sempre soddisfatta nel caso dei sistemi di equazioni differenziali, come andremo ora a vedere.

**Esempio 4.3.** *Infatti riprendiamo l'esempio precedente. Cominciamo col calcolare*

$$\mathcal{L}^{-1} \left( \frac{s^2 + 1}{s^2(s - 3)} \right).$$

Per fare ciò innanzitutto decomponiamo in fratti semplici la frazione assegnata, cioè cerchiamo costanti  $A, B, C$  tali che

$$\frac{s^2 + 1}{s^2(s - 3)} = \frac{A}{s} + \frac{B}{s^2} + \frac{C}{s - 3}.$$

Sommando le frazioni al secondo membro otteniamo:

$$\frac{s^2 + 1}{s^2(s - 3)} = \frac{(A + C)s^2 + (B - 3A)s - 3B}{s^2(s - 3)}.$$

Le due frazioni hanno lo stesso denominatore. Quindi la loro uguaglianza equivale all'uguaglianza dei rispettivi numeratori, cioè al fatto che

$$s^2 + 1 = (A + C)s^2 + (B - 3A)s - 3B.$$

Ricordando che due polinomi sono uguali se e solo se sono uguali i rispettivi coefficienti, l'uguaglianza precedente equivale alle seguenti condizioni:

$$\begin{cases} A + C = 1 \\ -3A + B = 0 \\ -3B = 1. \end{cases}$$

Questo sistema lineare consente di calcolare le costanti. Si trova così che  $A = -\frac{1}{9}$ ,  $B = -\frac{1}{3}$  e  $C = \frac{10}{9}$ . Quindi resta calcolata la decomposizione in fratti semplici:

$$\frac{s^2 + 1}{s^2(s - 3)} = \frac{-1/9}{s} + \frac{-1/3}{s^2} + \frac{10/9}{s - 3}.$$

Ora tenuto conto della linearità dell'antitransformata siamo in grado di calcolare la prima componente  $x_1(t)$  di  $x(t)$ , che è:

$$\begin{aligned} x_1(t) &= \mathcal{L}^{-1} \left( \frac{s^2 + 1}{s^2(s - 3)} \right) = \mathcal{L}^{-1} \left( \frac{-1/9}{s} + \frac{-1/3}{s^2} + \frac{10/9}{s - 3} \right) \\ &= -\frac{1}{9} \mathcal{L}^{-1} \left( \frac{1}{s} \right) - \frac{1}{3} \mathcal{L}^{-1} \left( \frac{1}{s^2} \right) + \frac{10}{9} \mathcal{L}^{-1} \left( \frac{1}{s - 3} \right) = -\frac{1}{9} - \frac{t}{3} + \frac{10}{9} e^{3t}. \end{aligned}$$

In modo analogo si calcolano  $x_2(t)$  ed  $x_3(t)$ , trovando:

$$x(t) = \begin{pmatrix} \frac{10}{9} e^{3t} - \frac{1}{3} t - \frac{1}{9} \\ 2e^t - 1 \\ 2te^t + 1 \end{pmatrix},$$

che è la stessa soluzione che avevamo trovato in precedenza utilizzando l'esponenziale di una matrice.