

Spazi vettoriali hermitiani.

1. Prodotto hermitiano, lunghezza e ortogonalità in \mathbf{C}^n .

Consideriamo lo spazio vettoriale

$$\mathbf{C}^n = \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, x_1, \dots, x_n \in \mathbf{C} \right\},$$

con la somma fra vettori e il prodotto di un vettore per uno scalare definiti rispettivamente da

$$\mathbf{x} + \mathbf{y} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \mathbf{x} := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \lambda \in \mathbf{C}.$$

Definizione. (*Prodotto hermitiano.*) Dati due vettori \mathbf{x} e \mathbf{y} in \mathbf{C}^n , il *prodotto hermitiano* $\mathbf{x} \cdot \mathbf{y}$ fra \mathbf{x} e \mathbf{y} è il numero complesso

$$\mathbf{x} \cdot \mathbf{y} := {}^t \mathbf{x} \bar{\mathbf{y}} = x_1 \bar{y}_1 + \dots + x_n \bar{y}_n.$$

Esempio. Dati $\mathbf{x} = \begin{pmatrix} 1+i \\ -2i \\ 2 \end{pmatrix}$ e $\mathbf{y} = \begin{pmatrix} 2-3i \\ 1-i \\ 3-i \end{pmatrix}$ in \mathbf{C}^3 , il prodotto hermitiano fra \mathbf{x} e \mathbf{y} è dato da

$$\mathbf{x} \cdot \mathbf{y} = (1+i \quad -2i \quad 2) \begin{pmatrix} 2+3i \\ 1+i \\ 3+i \end{pmatrix} = (1+i)(2+3i) + (-2i)(1+i) + 2(3+i) = 7 + 5i.$$

Esempio. Se due vettori $\mathbf{x}, \mathbf{y} \in \mathbf{C}^n$ hanno coordinate reali (ossia $\bar{\mathbf{x}} = \mathbf{x}$ e $\bar{\mathbf{y}} = \mathbf{y}$), allora il prodotto hermitiano $\mathbf{x} \cdot \mathbf{y}$ coincide col prodotto scalare reale fra \mathbf{x} e \mathbf{y} .

• Il prodotto hermitiano gode delle seguenti proprietà

(i) (*Proprietà di Hermitianità*) Per ogni $\mathbf{x}, \mathbf{y} \in \mathbf{C}^n$

$$\mathbf{x} \cdot \mathbf{y} = \overline{\mathbf{y} \cdot \mathbf{x}};$$

(ii) (*Proprietà distributiva*) Per ogni $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{C}^n$

$$(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}, \quad \mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z};$$

(iii) (*Sesquilinearità*) Per ogni $\mathbf{x}, \mathbf{y} \in \mathbf{C}^n$ ed ogni $\lambda \in \mathbf{C}$

$$(\lambda \mathbf{x}) \cdot \mathbf{y} = \lambda (\mathbf{x} \cdot \mathbf{y}) \quad \mathbf{x} \cdot (\lambda \mathbf{y}) = \bar{\lambda} (\mathbf{x} \cdot \mathbf{y});$$

(iv) (*Positività*) Per ogni $\mathbf{x} \in \mathbf{C}^n$

$$\mathbf{x} \cdot \mathbf{x} \geq 0,$$

$$\mathbf{x} \cdot \mathbf{x} = 0 \quad \text{se e soltanto se} \quad \mathbf{x} = \mathbf{0}.$$

Dimostrazione. La dimostrazione segue immediatamente dalle definizioni. Il punto (i) segue da

$$\mathbf{x} \cdot \mathbf{y} = x_1 \bar{y}_1 + \dots + x_n \bar{y}_n = \overline{y_1 \bar{x}_1 + \dots + y_n \bar{x}_n} = \overline{\mathbf{y} \cdot \mathbf{x}}.$$

Il punto (ii) segue da

$$\begin{aligned} (\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} &= (x_1 + y_1) \bar{z}_1 + \dots + (x_n + y_n) \bar{z}_n = \\ &= x_1 \bar{z}_1 + y_1 \bar{z}_1 + \dots + x_n \bar{z}_n + y_n \bar{z}_n = \\ &= \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}. \end{aligned}$$

$$\begin{aligned} \mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) &= x_1 \overline{(y_1 + z_1)} + \dots + x_n \overline{(y_n + z_n)} = \\ &= x_1 \bar{y}_1 + x_1 \bar{z}_1 + \dots + x_n \bar{y}_n + x_n \bar{z}_n = \\ &= \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}. \end{aligned}$$

Confrontando le quantità

$$\begin{aligned} \lambda(\mathbf{x} \cdot \mathbf{y}) &= \lambda(x_1 \bar{y}_1 + \dots + x_n \bar{y}_n) = \lambda x_1 \bar{y}_1 + \dots + \lambda x_n \bar{y}_n, \\ (\lambda \mathbf{x}) \cdot \mathbf{y} &= (\lambda x_1) \bar{y}_1 + \dots + (\lambda x_n) \bar{y}_n = \lambda x_1 \bar{y}_1 + \dots + \lambda x_n \bar{y}_n, \\ \mathbf{x} \cdot (\lambda \mathbf{y}) &= x_1 \overline{(\lambda y_1)} + \dots + x_n \overline{(\lambda y_n)} = \bar{\lambda} x_1 \bar{y}_1 + \dots + \bar{\lambda} x_n \bar{y}_n, \\ \bar{\lambda}(\mathbf{x} \cdot \mathbf{y}) &= \bar{\lambda}(x_1 \bar{y}_1 + \dots + x_n \bar{y}_n) = \bar{\lambda} x_1 \bar{y}_1 + \dots + \bar{\lambda} x_n \bar{y}_n \end{aligned}$$

otteniamo (iii). Per dimostrare (iv), osserviamo che

$$\mathbf{x} \cdot \mathbf{x} = |x_1|^2 + \dots + |x_n|^2.$$

Poiché i moduli di numeri complessi sono sempre reali non negativi, si ha che $\mathbf{x} \cdot \mathbf{x} \geq 0$. Se $\mathbf{x} = \mathbf{0}$, chiaramente $\mathbf{x} \cdot \mathbf{x} = 0$. Viceversa, se per un vettore $\mathbf{x} \in \mathbf{C}^n$ vale $\mathbf{x} \cdot \mathbf{x} = 0$, allora $|x_1|^2 + \dots + |x_n|^2 = 0$. Ciò è possibile solo se $x_1 = \dots = x_n = 0$.

Mediante il prodotto hermitiano definiamo in \mathbf{C}^n nozioni di lunghezza, distanza e ortogonalità.

Definizione. La norma o lunghezza $\|\mathbf{x}\|$ di un vettore $\mathbf{x} \in \mathbf{C}^n$ è definita da

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{|x_1|^2 + \dots + |x_n|^2}.$$

Esempio. La norma del vettore $\mathbf{x} = \begin{pmatrix} 1 - i \\ 2 + 3i \end{pmatrix} \in \mathbf{C}^2$ è data da

$$\|\mathbf{x}\| = \sqrt{|1 - i|^2 + |2 + 3i|^2} = \sqrt{(1 + 1) + (4 + 9)} = \sqrt{15}.$$

Definizione. La distanza fra i punti \mathbf{x} e \mathbf{y} in \mathbf{C}^n è definita da

$$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|.$$

In particolare, $\|\mathbf{x}\|$ coincide con la distanza di \mathbf{x} dall'origine.

• La norma gode delle seguenti proprietà:

(i) (*Omogeneità*) $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$, per ogni $\lambda \in \mathbf{C}$.

(ii) (*Disuguaglianza di Cauchy-Schwarz*) $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$, per ogni $\mathbf{x}, \mathbf{y} \in \mathbf{C}^n$.

(iii) (*Disuguaglianza triangolare*) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$, per ogni $\mathbf{x}, \mathbf{y} \in \mathbf{C}^n$.

Dimostrazione. Il punto (i) segue da

$$\|\lambda \mathbf{x}\| = \sqrt{|\lambda|^2 |x_1|^2 + \dots + |\lambda|^2 |x_n|^2} = |\lambda| \sqrt{|x_1|^2 + \dots + |x_n|^2} = |\lambda| \|\mathbf{x}\|.$$

(ii) Se $\mathbf{x} = 0$ oppure $\mathbf{y} = 0$ la disuguaglianza è chiaramente soddisfatta. Supponiamo adesso $\mathbf{x} \neq 0$ e $\mathbf{y} \neq 0$. Consideriamo un vettore della forma $\mathbf{z} = \alpha \mathbf{x} + \beta \mathbf{y}$, con $\alpha, \beta \in \mathbf{C}$. Per le proprietà (i)(ii)(iii)(iv) del prodotto hermitiano, abbiamo che

$$\mathbf{z} \cdot \mathbf{z} = |\alpha|^2 \|\mathbf{x}\|^2 + |\beta|^2 \|\mathbf{y}\|^2 + 2\operatorname{Re}(\alpha \bar{\beta} \mathbf{x} \cdot \mathbf{y}) \geq 0, \quad \forall \alpha, \beta \in \mathbf{C}.$$

In particolare, per $\alpha = \|\mathbf{y}\|^2$ e $\beta = -\mathbf{x} \cdot \mathbf{y}$, troviamo

$$\|\mathbf{y}\|^4 \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 |\mathbf{x} \cdot \mathbf{y}|^2 - 2\|\mathbf{y}\|^2 |\mathbf{x} \cdot \mathbf{y}|^2 = \|\mathbf{y}\|^4 \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2 |\mathbf{x} \cdot \mathbf{y}|^2 \geq 0.$$

Dividendo per $\|\mathbf{y}\|^2 \neq 0$, otteniamo

$$|\mathbf{x} \cdot \mathbf{y}|^2 \leq \|\mathbf{x}\|^2 \|\mathbf{y}\|^2$$

che è equivalente alla disuguaglianza cercata.

(iii) La disuguaglianza triangolare è equivalente alla disuguaglianza

$$\|\mathbf{x} + \mathbf{y}\|^2 \leq (\|\mathbf{x}\| + \|\mathbf{y}\|)^2.$$

Direttamente dalle definizioni e dalla disuguaglianza di Cauchy-Schwarz abbiamo

$$\|\mathbf{x} + \mathbf{y}\|^2 = (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\operatorname{Re}(\mathbf{x} \cdot \mathbf{y}) \leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\| \|\mathbf{y}\| = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2,$$

come richiesto.

Definizione. Due vettori $\mathbf{x}, \mathbf{y} \in \mathbf{C}^n$ si dicono *ortogonali* se $\mathbf{x} \cdot \mathbf{y} = 0$. Questo si indica con $\mathbf{x} \perp \mathbf{y}$.

• La *proiezione ortogonale* $\pi_{\mathbf{y}}(\mathbf{x})$ di un vettore \mathbf{x} su un vettore $\mathbf{y} \neq 0$ è un vettore $\pi_{\mathbf{y}}(\mathbf{x}) = c\mathbf{y}$ multiplo di \mathbf{y} per uno scalare complesso $c \in \mathbf{C}$, caratterizzato dalla proprietà

$$(\mathbf{x} - \pi_{\mathbf{y}}(\mathbf{x})) \cdot \mathbf{y} = 0.$$

In altre parole, la proiezione ortogonale di un vettore \mathbf{x} su un vettore $\mathbf{y} \neq 0$ determina una scomposizione del vettore \mathbf{x} nella somma di un vettore parallelo a \mathbf{y} e un vettore ortogonale a \mathbf{y}

$$\mathbf{x} = \mathbf{z} + \pi_{\mathbf{y}}(\mathbf{x}), \quad \mathbf{z} = \mathbf{x} - \pi_{\mathbf{y}}(\mathbf{x}), \quad \mathbf{z} \perp \mathbf{y}. \quad (2)$$

Risulta

$$\pi_{\mathbf{y}}(\mathbf{x}) = \frac{\mathbf{x} \cdot \mathbf{y}}{\mathbf{y} \cdot \mathbf{y}} \mathbf{y}.$$

Esempio. Siano $\mathbf{x} = \begin{pmatrix} i \\ 1+i \\ 0 \end{pmatrix}$ e $\mathbf{y} = \begin{pmatrix} i \\ -i \\ 1+3i \end{pmatrix}$. Poiché $\mathbf{x} \cdot \mathbf{y} = 2+i$ e $\mathbf{y} \cdot \mathbf{y} = 12$, troviamo

$$\pi_{\mathbf{y}}(\mathbf{x}) = \frac{2+i}{12} \begin{pmatrix} i \\ -i \\ 1+3i \end{pmatrix} = \begin{pmatrix} -1+2i \\ 1-2i \\ -1+7i \end{pmatrix}.$$

2. Basi ortonormali. Procedimento di ortonormalizzazione di Gram-Schmidt.

Definizione. Un sottoinsieme $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ di \mathbf{C}^n si dice un insieme ortogonale se i suoi elementi sono a due a due ortogonali fra loro.

Esercizio. Gli elementi di un insieme ortogonale sono linearmente indipendenti su \mathbf{C} .

Definizione. Una base ortonormale di \mathbf{C}^n è una base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ i cui elementi hanno norma uno e sono a due a due ortogonali:

$$\|\mathbf{e}_1\| = \dots = \|\mathbf{e}_n\| = 1, \quad \mathbf{e}_i \cdot \mathbf{e}_j = 0, \quad i \neq j.$$

Esempio. La base canonica di \mathbf{C}^n

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}$$

è una base ortonormale: ogni vettore $\mathbf{z} \in \mathbf{C}^n$ si scrive in modo unico come

$$\mathbf{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = z_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + z_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + z_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \quad z_1, z_2, \dots, z_n \in \mathbf{C}.$$

Inoltre si verifica facilmente che tutti i vettori di tale base hanno norma 1 e sono a due a due ortogonali fra loro.

Esempio. I vettori $\left\{ \begin{pmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{pmatrix}, \begin{pmatrix} i/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \right\}$ formano una base ortonormale di \mathbf{C}^2 .

Esempio. I vettori $\left\{ \begin{pmatrix} \cos \theta \\ \sin \theta \\ 0 \end{pmatrix}, \begin{pmatrix} -\sin \theta \\ \cos \theta \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ i \end{pmatrix} \right\}$ formano una base ortonormale di \mathbf{C}^3 .

Il metodo di *ortonormalizzazione di Gram-Schmidt* permette di ottenere una base ortonormale di \mathbf{C}^n a partire da una base qualsiasi.

Sia $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base di \mathbf{C}^n . Allora i vettori

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{v}_1 \\ \mathbf{u}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 \\ \mathbf{u}_3 &= \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2 \\ &\vdots \\ \mathbf{u}_n &= \mathbf{v}_n - \frac{\mathbf{v}_n \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 - \dots - \frac{\mathbf{v}_n \cdot \mathbf{u}_{n-1}}{\mathbf{u}_{n-1} \cdot \mathbf{u}_{n-1}} \mathbf{u}_{n-1} \end{aligned}$$

sono a due a due ortogonali (cfr. equazione (2)). Inoltre, poiché per ogni $1 \leq j \leq n$,

$$\text{span}_{\mathbf{C}}\{\mathbf{v}_1, \dots, \mathbf{v}_j\} = \text{span}_{\mathbf{C}}\{\mathbf{u}_1, \dots, \mathbf{u}_j\}, \quad (3)$$

anche i vettori $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ formano una base di \mathbf{C}^n . Infine i vettori

$$\{\mathbf{e}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \dots, \mathbf{e}_n = \frac{\mathbf{u}_n}{\|\mathbf{u}_n\|}\}$$

formano una base ortonormale di \mathbf{C}^n .

Esempio. Sia data la base $\{\mathbf{v}_1 = \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 0 \\ i \\ 0 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1+i \end{pmatrix}\}$ di \mathbf{C}^3 . Allora i vettori

$$\begin{aligned} \mathbf{u}_1 = \mathbf{v}_1 &= \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} \\ \mathbf{u}_2 &= \begin{pmatrix} 0 \\ i \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} = \begin{pmatrix} -1/2 \\ i/2 \\ 0 \end{pmatrix} \\ \mathbf{u}_3 = \mathbf{v}_3 &= \begin{pmatrix} 0 \\ 0 \\ 1+i \end{pmatrix} - 0 \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} - 0 \begin{pmatrix} 1/2 \\ i/2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1+i \end{pmatrix} \end{aligned}$$

formano una base ortogonale di \mathbf{C}^3 e i vettori

$$\mathbf{e}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \mathbf{e}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ i \\ 0 \end{pmatrix}, \mathbf{e}_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1+i \end{pmatrix}$$

formano una base ortonormale di \mathbf{C}^3 .

Osservazione. Sia $U \subset \mathbf{C}^n$ il sottospazio generato dai primi k vettori $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ della base $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ di \mathbf{C}^n . Dalla relazione (3) segue che i vettori $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$, ottenuti nel corso del procedimento di ortonormalizzazione di Gram-Schmidt, sono una base ortogonale di U .

Esercizio. Dato un sottospazio U in \mathbf{C}^n di dimensione k , esiste una base ortonormale di \mathbf{C}^n

$$\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$$

con la proprietà che $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ è una base ortonormale di U e $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ è una base ortonormale di U^\perp . In altre parole, data una base ortonormale di U , essa può essere completata ad una base ortonormale di \mathbf{C}^n .

• Sia $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base *ortogonale* di \mathbf{C}^n e sia $\mathbf{x} \in \mathbf{C}^n$. Allora le coordinate di \mathbf{x} in \mathcal{B} sono date da

$$x_1 = \frac{\mathbf{x} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}, \dots, x_n = \frac{\mathbf{x} \cdot \mathbf{v}_n}{\mathbf{v}_n \cdot \mathbf{v}_n}.$$

In particolare, se la base \mathcal{B} è *ortonormale*, le coordinate di \mathbf{x} in \mathcal{B} sono date da

$$x_1 = \mathbf{x} \cdot \mathbf{v}_1, \dots, x_n = \mathbf{x} \cdot \mathbf{v}_n.$$

Dim. Il vettore \mathbf{x} si scrive in modo unico come combinazione lineare degli elementi di \mathcal{B}

$$\mathbf{x} = x_1 \mathbf{v}_1 + \dots + x_n \mathbf{v}_n.$$

Per $1 \leq j \leq n$,

$$\mathbf{x} \cdot \mathbf{v}_j = x_1 \mathbf{v}_1 \cdot \mathbf{v}_j + \dots + x_n \mathbf{v}_n \cdot \mathbf{v}_j = x_j \mathbf{v}_j \cdot \mathbf{v}_j,$$

da cui

$$x_j = \frac{\mathbf{x} \cdot \mathbf{v}_j}{\mathbf{v}_j \cdot \mathbf{v}_j}.$$

Esempio. Le coordinate del vettore $\mathbf{x} = \begin{pmatrix} i \\ i \\ i \end{pmatrix}$ nella base ortonormale

$$\mathbf{e}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ i \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1+i \end{pmatrix}$$

sono date rispettivamente da $x_1 = \frac{1}{\sqrt{2}}(1+i)$, $x_2 = \frac{1}{\sqrt{2}}(1-i)$, $x_3 = \frac{1}{\sqrt{2}}(1+i)$.

3. Complementi ortogonali. Proiezioni ortogonali.

Sia U un sottoinsieme di \mathbf{C}^n .

Definizione. Un vettore $\mathbf{x} \in \mathbf{C}^n$ si dice ortogonale a U se è ortogonale a tutti gli elementi di U . L'ortogonale U^\perp di U è per definizione

$$U^\perp := \{\mathbf{x} \in \mathbf{C}^n \mid \mathbf{x} \cdot \mathbf{u} = 0, \forall \mathbf{u} \in U\}.$$

Proposizione. U^\perp è un sottospazio vettoriale di \mathbf{C}^n .

Dimostrazione. Siano $\mathbf{x}, \mathbf{y} \in U^\perp$, $\lambda \in \mathbf{C}$ e sia \mathbf{u} un arbitrario elemento di U . Dalle proprietà del prodotto hermitiano e dalle ipotesi segue che

$$(\mathbf{x} + \mathbf{y}) \cdot \mathbf{u} = \mathbf{x} \cdot \mathbf{u} + \mathbf{y} \cdot \mathbf{u} = 0 + 0 = 0, \quad (\lambda \mathbf{x}) \cdot \mathbf{u} = \lambda \mathbf{x} \cdot \mathbf{u} = \lambda 0 = 0.$$

In altre parole, $\mathbf{x} + \mathbf{y} \in U^\perp$ e $\lambda \mathbf{x} \in U^\perp$, per ogni $\lambda \in \mathbf{C}$, come richiesto.

• Se U è un sottospazio vettoriale di \mathbf{C}^n di dimensione complessa k e $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ è una base di U , allora

$$U^\perp = \left\{ \mathbf{x} \in \mathbf{C}^n \mid \begin{cases} \mathbf{x} \cdot \mathbf{u}_1 = 0 \\ \vdots \\ \mathbf{x} \cdot \mathbf{u}_k = 0 \end{cases} \right\}.$$

In questo caso, U^\perp è un sottospazio vettoriale di dimensione complessa $n - k$.

Dimostrazione. Sia $\mathbf{x} \in U^\perp$. Segue immediatamente dalla definizione che $\mathbf{x} \cdot \mathbf{u}_1 = \dots = \mathbf{x} \cdot \mathbf{u}_k = 0$. Viceversa, supponiamo che \mathbf{x} soddisfi il sistema $\mathbf{x} \cdot \mathbf{u}_1 = \dots = \mathbf{x} \cdot \mathbf{u}_k = 0$. Poiché un arbitrario elemento di U si scrive come $\mathbf{u} = \alpha_1 \mathbf{u}_1 + \dots + \alpha_k \mathbf{u}_k$, con $\alpha_i \in \mathbf{C}$, si ha

$$\mathbf{x} \cdot \mathbf{u} = \alpha_1 \mathbf{x} \cdot \mathbf{u}_1 + \dots + \alpha_k \mathbf{x} \cdot \mathbf{u}_k = 0.$$

In altre parole, $\mathbf{x} \in U^\perp$. Questa caratterizzazione esprime U^\perp come lo spazio delle soluzioni di un sistema lineare omogeneo di k equazioni indipendenti. Di conseguenza, U^\perp ha dimensione complessa $n - k$.

Esercizio. Verificare che $U \cap U^\perp = \{O\}$.

• Se U è un sottospazio vettoriale di \mathbf{C}^n , il sottospazio U^\perp si chiama complemento ortogonale di U . Lo spazio \mathbf{C}^n si decompone infatti come somma diretta di U e U^\perp

$$\mathbf{C}^n = U \oplus U^\perp \quad U \cap U^\perp = \{O\}.$$

In particolare, ogni elemento $\mathbf{x} \in \mathbf{C}^n$ si scrive *in modo unico* come somma di un elemento in U e un elemento in U^\perp

$$\mathbf{x} = \mathbf{x}_U + \mathbf{x}_{U^\perp}, \quad \mathbf{x}_U \in U, \quad \mathbf{x}_{U^\perp} \in U^\perp, \quad \text{e vale} \quad \|\mathbf{x}\|^2 = \|\mathbf{x}_U\|^2 + \|\mathbf{x}_{U^\perp}\|^2.$$

Definizione. Per definizione i vettori \mathbf{x}_U e \mathbf{x}_{U^\perp} sono rispettivamente le proiezioni ortogonali di \mathbf{x} su U e su U^\perp

$$\mathbf{x}_U = \pi_U(\mathbf{x}) \quad \mathbf{x}_{U^\perp} = \pi_{U^\perp}(\mathbf{x}).$$

Calcolo della proiezione ortogonale di un vettore su un sottospazio.

Proposizione. Sia U un sottospazio di \mathbf{C}^n e sia $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ una qualunque base ortogonale di U . Sia $\mathbf{x} \in \mathbf{C}^n$ un vettore. Allora la proiezione ortogonale di \mathbf{x} su U è data da

$$\mathbf{x}_U = \pi_U(\mathbf{x}) = \frac{\mathbf{x} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \dots + \frac{\mathbf{x} \cdot \mathbf{u}_k}{\mathbf{u}_k \cdot \mathbf{u}_k} \mathbf{u}_k.$$

Dimostrazione. Sia $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ una base ortogonale di \mathbf{C}^n che completa $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$. In particolare, $\{\mathbf{u}_{k+1}, \dots, \mathbf{u}_n\}$ è una base ortogonale di U^\perp . In questa base,

$$\mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \dots + \frac{\mathbf{x} \cdot \mathbf{u}_k}{\mathbf{u}_k \cdot \mathbf{u}_k} \mathbf{u}_k + \frac{\mathbf{x} \cdot \mathbf{u}_{k+1}}{\mathbf{u}_{k+1} \cdot \mathbf{u}_{k+1}} \mathbf{u}_{k+1} + \dots + \frac{\mathbf{x} \cdot \mathbf{u}_n}{\mathbf{u}_n \cdot \mathbf{u}_n} \mathbf{u}_n.$$

Per l'unicità di \mathbf{x}_U e di \mathbf{x}_{U^\perp} , segue che

$$\mathbf{x}_U = \pi_U(\mathbf{x}) = \frac{\mathbf{x} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 + \dots + \frac{\mathbf{x} \cdot \mathbf{u}_k}{\mathbf{u}_k \cdot \mathbf{u}_k} \mathbf{u}_k$$

e analogamente

$$\mathbf{x}_{U^\perp} = \pi_{U^\perp}(\mathbf{x}) = \frac{\mathbf{x} \cdot \mathbf{u}_{k+1}}{\mathbf{u}_{k+1} \cdot \mathbf{u}_{k+1}} \mathbf{u}_{k+1} + \dots + \frac{\mathbf{x} \cdot \mathbf{u}_n}{\mathbf{u}_n \cdot \mathbf{u}_n} \mathbf{u}_n.$$

Osservazione. L'applicazione

$$\pi_U: \mathbf{C}^n \longrightarrow \mathbf{C}^n, \quad \mathbf{x} \mapsto \pi_U(\mathbf{x}),$$

che ad un vettore associa la sua proiezione ortogonale sul sottospazio U , è un'applicazione lineare. Vale infatti

$$\pi_U(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha \pi_U(\mathbf{x}) + \beta \pi_U(\mathbf{y}), \quad \forall \alpha, \beta \in \mathbf{C}, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbf{C}^n.$$

Inoltre, la proiezione ortogonale di \mathbf{x} su U è data dalla somma delle proiezioni di \mathbf{x} sui singoli vettori ortogonali $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$.

La proiezione ortogonale di un vettore \mathbf{x} su un sottospazio U è il punto di U più vicino ad \mathbf{x} .

Proposizione. Sia U un sottospazio di \mathbf{C}^n , sia $\mathbf{x} \in \mathbf{C}^n$ e sia \mathbf{x}_U la proiezione ortogonale di \mathbf{x} su U . Allora, per ogni $\mathbf{u} \in U$

$$d(\mathbf{x}, \mathbf{x}_U) \leq d(\mathbf{x}, \mathbf{u}).$$

Dimostrazione. Sia $\mathbf{u} \in U$ un elemento arbitrario. L'identità

$$\mathbf{x} - \mathbf{u} = \mathbf{x} - \mathbf{x}_U + \mathbf{x}_U - \mathbf{u}, \quad \text{con } \mathbf{x} - \mathbf{x}_U \in U^\perp, \quad \mathbf{x}_U - \mathbf{u} \in U$$

scomponi di $\mathbf{x} - \mathbf{u}$ come somma di due vettori ortogonali. In particolare implica

$$\|\mathbf{x} - \mathbf{u}\|^2 = \|\mathbf{x} - \mathbf{x}_U\|^2 + \|\mathbf{x}_U - \mathbf{u}\|^2$$

e

$$\|\mathbf{x} - \mathbf{x}_U\|^2 \leq \|\mathbf{x} - \mathbf{u}\|^2 \quad \Leftrightarrow \quad \|\mathbf{x} - \mathbf{x}_U\| \leq \|\mathbf{x} - \mathbf{u}\|$$

come richiesto.

Definizione. La distanza di un vettore \mathbf{x} da un sottospazio U è per definizione la distanza fra \mathbf{x} e la sua proiezione ortogonale su U

$$d(\mathbf{x}, U) = d(\mathbf{x}, \mathbf{x}_U).$$

In particolare, se $\mathbf{x} \in U$, allora $\mathbf{x}_U = \mathbf{x}$ e $d(\mathbf{x}, U) = 0$.

4. Applicazioni lineari unitarie.

Sia \mathbf{C}^n lo spazio delle ennuple complesse col prodotto hermitiano canonico.

Definizione. Un'applicazione lineare $F: \mathbf{C}^n \rightarrow \mathbf{C}^n$ si dice *unitaria* se

$$F(\mathbf{x}) \cdot F(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}, \quad \text{per ogni } \mathbf{x}, \mathbf{y} \in \mathbf{C}^n.$$

In altre parole, l'applicazione F è un'*isometria lineare* di \mathbf{C}^n .

Osservazione. Direttamente dalla definizione segue che un'applicazione lineare unitaria conserva la norma dei vettori

$$\|F(\mathbf{x})\| = \|\mathbf{x}\|, \quad \text{per ogni } \mathbf{x} \in \mathbf{C}^n.$$

Questo fatto a sua volta implica che un'applicazione lineare unitaria è necessariamente iniettiva, e quindi suriettiva e biiettiva. Inoltre, sempre dalla definizione, segue che un'applicazione lineare unitaria manda basi ortonormali in basi ortonormali.

Sia $F: \mathbf{C}^n \rightarrow \mathbf{C}^n$ un'applicazione lineare unitaria e sia M la matrice rappresentativa di F nella base canonica (in dominio e codominio), così che

$$F(\mathbf{x}) = M\mathbf{x}, \quad \mathbf{x} \in \mathbf{C}^n.$$

Poiché per ogni $\mathbf{x}, \mathbf{y} \in \mathbf{C}^n$ vale

$$M\mathbf{x} \cdot M\mathbf{y} = {}^t(M\mathbf{x})\overline{M\mathbf{y}} = {}^t\mathbf{x}{}^tM\overline{M\mathbf{y}} = \mathbf{x} \cdot \mathbf{y}$$

la matrice M soddisfa la condizione ${}^tM \cdot \overline{M} = Id$, ossia $M^{-1} = {}^t\overline{M}$. Una matrice con questa proprietà si chiama *matrice unitaria*. Una matrice unitaria è anche caratterizzata dal fatto che le sue colonne (e le sue righe) formano una base ortonormale di \mathbf{C}^n .

Esercizio 4.1. Far vedere che

$$\begin{pmatrix} \cos \theta & i \sin \theta \\ i \sin \theta & \cos \theta \end{pmatrix}, \quad \begin{pmatrix} i \cos \theta & i \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

sono matrici unitarie, per ogni $\theta \in [0, 2\pi]$.

Esercizio 4.2. Sia M una matrice unitaria $n \times n$, ossia tale che ${}^t M \overline{M} = I_n$.

- (i) Far vedere che M^{-1} e ${}^t \overline{M}$ sono matrici unitarie.
- (ii) Far vedere che $|\det M| = 1$.
- (iii) Far vedere che se λ è un autovalore di M , allora $|\lambda| = 1$.
- (iv) Far vedere che il prodotto di due matrici unitarie è una matrice unitaria.

Sol. (i) dobbiamo verificare che ${}^t(M^{-1})\overline{M^{-1}} = I_n$:

$${}^t(M^{-1})\overline{M^{-1}} = ({}^t M^{-1})(\overline{M})^{-1} = \overline{M} {}^t M = {}^t M \overline{M} = I_n;$$

(nell'ultimo passaggio abbiamo usato che ${}^t M$ ed \overline{M} sono una inversa dell'altra e dunque commutano).

(ii) $1 = \det(I_n) = \det({}^t M \overline{M}) = \det({}^t M) \det(\overline{M}) = \det(M) \overline{\det(M)} = |\det(M)|^2$.

(iii) Sia λ un autovalore di M e sia $\mathbf{x} \neq 0$ un autovettore: $M\mathbf{x} = \lambda\mathbf{x}$. Dalle proprietà delle matrici unitarie segue

$$M\mathbf{x} \cdot M\mathbf{x} = \mathbf{x} \cdot \mathbf{x} = (\lambda\mathbf{x}) \cdot (\lambda\mathbf{x}) = \lambda\overline{\lambda}(\mathbf{x} \cdot \mathbf{x}) \quad \Leftrightarrow \quad \mathbf{x} \cdot \mathbf{x} = |\lambda|^2(\mathbf{x} \cdot \mathbf{x}).$$

Poiché $\mathbf{x} \cdot \mathbf{x} \neq 0$, deve essere $|\lambda|^2 = 1$, come richiesto.

(iv) Siano M ed N matrici unitarie.

$${}^t(MN) \overline{MN} = {}^t N {}^t M \overline{M} \overline{N} = {}^t N I_n \overline{N} = I_n.$$

5. Applicazioni lineari hermitiane.

Sia \mathbf{C}^n lo spazio delle ennuple complesse col prodotto hermitiano canonico.

Definizione. Un'applicazione lineare $F: \mathbf{C}^n \rightarrow \mathbf{C}^n$ si dice *hermitiana* se

$$F(\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot F(\mathbf{y}), \quad \text{per ogni } \mathbf{x}, \mathbf{y} \in \mathbf{C}^n.$$

Sia A la matrice rappresentativa di F nella base canonica (in dominio e codominio).

Poiché per ogni $\mathbf{x}, \mathbf{y} \in \mathbf{C}^n$ vale

$$\mathbf{Ax} \cdot \mathbf{y} = {}^t(\mathbf{Ax})\overline{\mathbf{y}} = {}^t \mathbf{x} {}^t A \overline{\mathbf{y}} = {}^t \mathbf{x} \overline{A\mathbf{y}} = {}^t \mathbf{x} \overline{A} \overline{\mathbf{y}} \cdot \mathbf{Ay} \quad (*)$$

la matrice A soddisfa la condizione ${}^t A = \overline{A}$. Una matrice con questa proprietà si chiama *matrice hermitiana*.

Esempio. Esempi di matrici hermitiane

$$M = \begin{pmatrix} 2 & 2i \\ -2i & -2 \end{pmatrix}, \quad N = \begin{pmatrix} 1 & 1-2i & 3 \\ 1+2i & 6 & 2i \\ 3 & -2i & 0 \end{pmatrix}.$$

Esempio. Una matrice hermitiana a coefficienti reali è una matrice reale simmetrica:

$$\begin{cases} M = \overline{M} \\ {}^t M = M \end{cases} \Rightarrow \begin{cases} M = \overline{M} \\ {}^t M = M \end{cases}.$$

Lemma 5.1. Sia $F: \mathbf{C}^n \rightarrow \mathbf{C}^n$ un'applicazione hermitiana. Sia U un sottospazio di \mathbf{C}^n e sia U^\perp il suo complemento ortogonale. Se $F(U) \subset U$, allora anche $F(U^\perp) \subset U^\perp$.

Dim. Sia \mathbf{w} un arbitrario elemento di U^\perp . Dobbiamo verificare che $F(\mathbf{w}) \in U^\perp$, ossia che $F(\mathbf{w}) \cdot \mathbf{u} = 0$ per ogni $\mathbf{u} \in U$. Poiché per ipotesi $F(\mathbf{u}) \in U$, dalla (*) segue che

$$F(\mathbf{w}) \cdot \mathbf{u} = \mathbf{w} \cdot F(\mathbf{u}) = 0, \quad \forall \mathbf{u} \in U.$$

Dunque $F(\mathbf{w}) \in U^\perp$ come richiesto.

Per le matrici hermitiane vale un teorema di diagonalizzazione, di cui il teorema di diagonalizzazione per le matrici simmetriche reali è un caso particolare. Ad esempio la dimostrazione del fatto che gli autovalori di una matrice simmetrica reale sono reali si ottiene come caso particolare dell'analogo risultato per le matrici hermitiane. È curioso che questo metodo sia più semplice di una dimostrazione diretta.

Teorema spettrale per matrici hermitiane. Sia A una matrice hermitiana $n \times n$.

- (i) Sia λ un autovalore di A . Allora $\lambda = \bar{\lambda}$, cioè λ è reale.
- (ii) Autospazi relativi ad autovalori distinti sono ortogonali.
- (iii) Sia λ un autovalore di A di molteplicità algebrica k e sia V_λ l'autospazio di λ . Allora $\dim V_\lambda = k$.

Dim. (i) Sia $\mathbf{x} \in V_\lambda$ un autovettore di autovalore λ . Per definizione \mathbf{x} è un vettore non nullo, tale che $A\mathbf{x} = \lambda\mathbf{x}$. Sfruttando la relazione (*) troviamo

$$A\mathbf{x} \cdot \mathbf{x} = (\lambda\mathbf{x}) \cdot \mathbf{x} = \lambda(\mathbf{x} \cdot \mathbf{x}) = \mathbf{x} \cdot A\mathbf{x} = \mathbf{x} \cdot (\lambda\mathbf{x}) = \bar{\lambda}(\mathbf{x} \cdot \mathbf{x}),$$

da cui segue che $\lambda(\mathbf{x} \cdot \mathbf{x}) = \bar{\lambda}(\mathbf{x} \cdot \mathbf{x})$. Dividendo ambo i termini per $\mathbf{x} \cdot \mathbf{x} \neq 0$, troviamo $\lambda = \bar{\lambda}$, come richiesto.

(ii) Siano λ e μ autovalori distinti di A e siano V_λ e V_μ i rispettivi autospazi. Facciamo vedere che elementi arbitrari $\mathbf{x} \in V_\lambda$ e $\mathbf{y} \in V_\mu$ sono ortogonali fra loro. Dalla (*) si ha

$$A\mathbf{x} \cdot \mathbf{y} = (\lambda\mathbf{x}) \cdot \mathbf{y} = \lambda(\mathbf{x} \cdot \mathbf{y}) = \mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot (\mu\mathbf{y}) = \mu(\mathbf{x} \cdot \mathbf{y}), \quad (\mu \text{ è reale})$$

da cui segue che $(\lambda - \mu)(\mathbf{x} \cdot \mathbf{y}) = 0$. Poiché $\lambda \neq \mu$, deve valere $\mathbf{x} \cdot \mathbf{y} = 0$, cioè $\mathbf{x} \perp \mathbf{y}$ come richiesto.

(iii) Siano $\lambda_1, \dots, \lambda_k$ gli autovalori di A e siano $V_{\lambda_1}, \dots, V_{\lambda_k}$ gli autospazi corrispondenti. Consideriamo il seguente sottospazio di \mathbf{C}^n

$$U = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k} = \{X \in \mathbf{C}^n \mid \exists \lambda \in \mathbf{C} : AX = \lambda X\}.$$

Sia $L_A: \mathbf{C}^n \rightarrow \mathbf{C}^n$, $\mathbf{x} \mapsto A\mathbf{x}$ l'applicazione lineare data dalla moltiplicazione per A . È chiaro dalla definizione di U che $L_A(U) \subset U$. Dal Lemma 5.1 segue che $L_A(U^\perp) \subset U^\perp$; pertanto la restrizione ad U^\perp definisce un'applicazione lineare hermitiana $L_A|_{U^\perp}: U^\perp \rightarrow U^\perp$. Questa applicazione ha almeno un autovalore, per cui esistono $\sigma \in \mathbf{C}$ (per la precisione $\sigma \in \mathbf{R}$) e un vettore $\mathbf{x} \in U^\perp$ tali che $L_A(\mathbf{x}) = \sigma\mathbf{x}$. Questo contraddice la definizione di U , e implica $U^\perp = \{0\}$. In particolare, $\mathbf{C}^n = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k}$ e ogni autospazio di A ha dimensione massima, uguale alla molteplicità algebrica dell'autovalore corrispondente.

Direttamente da fatti (i)(ii)(iii) segue che

- (iv) Esiste una base ortonormale di \mathbf{C}^n formata da autovettori di A .
- (v) La matrice A è diagonalizzabile mediante una matrice unitaria, ossia esiste una matrice unitaria M tale che

$$M^{-1}AM = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & \dots & \dots & \lambda_n \end{pmatrix}, \quad (3)$$

dove $\lambda_1, \dots, \lambda_n$ sono gli autovalori di A .

Una matrice M che soddisfa la relazione (3) è una qualunque matrice che ha per colonne una base ortonormale di \mathbf{C}^n formata da autovettori di A . In altre parole, se

$$\left\{ \begin{pmatrix} v_{11} \\ \vdots \\ v_{n1} \end{pmatrix}, \dots, \begin{pmatrix} v_{1n} \\ \vdots \\ v_{nn} \end{pmatrix} \right\}$$

è una qualunque base ortonormale di \mathbf{C}^n formata da autovettori di A (di autovalori $\lambda_1, \dots, \lambda_n$ rispettivamente), allora la matrice unitaria

$$M = \begin{pmatrix} v_{11} & \dots & v_{1n} \\ \vdots & & \vdots \\ v_{n1} & \dots & v_{nn} \end{pmatrix}$$

soddisfa la relazione (3).

Esercizio. Sia A una matrice $n \times n$ antihermitiana, cioè che soddisfa la relazione ${}^t A = -\bar{A}$. Verificare che:

- (i) Sia λ un autovalore di A . Allora $\lambda = -\bar{\lambda}$, cioè λ è immaginario puro.
 - (ii) Autospazi relativi ad autovalori distinti sono ortogonali.
 - (iii) Sia λ un autovalore di A di molteplicità algebrica k e sia V_λ l'autospazio di λ . Allora $\dim V_\lambda = k$.
 - (iv) Esiste una base ortonormale di \mathbf{C}^n formata da autovettori di A .
 - (v) La matrice A è diagonalizzabile mediante una matrice unitaria.
- (suggerimento: ripercorrere la dimostrazione del Lemma 5.1 e del teorema spettrale per matrici hermitiane.....).

Concludiamo questa sezione con un teorema di diagonalizzazione per matrici unitarie. Anche per le matrici unitarie vale un analogo del Lemma 5.1.

Lemma 5.2. Sia $F: \mathbf{C}^n \rightarrow \mathbf{C}^n$ un'applicazione lineare unitaria. Sia $W \subset \mathbf{C}^n$ un sottospazio tale che $F(W) = W$. Verificare che vale anche $F(W^\perp) = W^\perp$.

Dim. Sia $\mathbf{w} \in W^\perp$. Dobbiamo dimostrare che $F(\mathbf{w}) \in W^\perp$, ossia che $F(\mathbf{w}) \cdot \mathbf{u} = 0$, per ogni $\mathbf{u} \in W$. Dalle proprietà delle applicazioni unitarie abbiamo

$$F(\mathbf{w}) \cdot F(\mathbf{u}) = \mathbf{w} \cdot \mathbf{u} = 0, \quad \forall \mathbf{u} \in W.$$

Ricordiamo che un'applicazione unitaria è necessariamente biettiva. In particolare al variare di $\mathbf{u} \in W$, anche $F(\mathbf{u})$ copre tutti i vettori di W . Dunque il lemma è dimostrato.

Teorema spettrale per matrici unitarie. Sia U una matrice unitaria $n \times n$.

- (i) Autospazi relativi ad autovalori distinti sono ortogonali.
- (ii) Sia λ un autovalore di U di molteplicità algebrica k e sia V_λ l'autospazio di λ . Allora $\dim V_\lambda = k$.

Dim. (i) Siano λ e μ autovalori distinti di U e siano V_λ e V_μ i rispettivi autospazi. Facciamo vedere che elementi arbitrari $\mathbf{x} \in V_\lambda$ e $\mathbf{y} \in V_\mu$ sono ortogonali fra loro. Per le proprietà delle matrici unitarie abbiamo

$$\mathbf{x} \cdot \mathbf{y} = U\mathbf{x} \cdot U\mathbf{y} = (\lambda\mathbf{x}) \cdot (\mu\mathbf{y}) = \lambda\bar{\mu}(\mathbf{x} \cdot \mathbf{y}) \quad \Leftrightarrow \quad (1 - \lambda\bar{\mu})(\mathbf{x} \cdot \mathbf{y}) = 0.$$

Ricordiamo che gli autovalori di una matrice unitaria hanno modulo uno e che per un numero complesso μ di modulo uno vale $\mu^{-1} = \bar{\mu}$. Dunque per $\lambda \neq \mu$, si ha $(1 - \lambda\bar{\mu}) = (\mu - \lambda) \neq 0$, da cui segue che $\mathbf{x} \cdot \mathbf{y} = 0$, come richiesto.

(iii) Siano $\lambda_1, \dots, \lambda_k$ gli autovalori di U e siano $V_{\lambda_1}, \dots, V_{\lambda_k}$ gli autospazi corrispondenti. Consideriamo il seguente sottospazio di \mathbf{C}^n

$$W = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k} = \{\mathbf{x} \in \mathbf{C}^n \mid \exists \lambda \in \mathbf{C} : U\mathbf{x} = \lambda\mathbf{x}\}.$$

Sia $L_U: \mathbf{C}^n \rightarrow \mathbf{C}^n$, $\mathbf{x} \mapsto U\mathbf{x}$ l'applicazione lineare data dalla moltiplicazione per U . È chiaro dalla definizione di W che $L_U(W) \subset W$. Dal Lemma 5.2 segue che $L_U(W^\perp) \subset W^\perp$; pertanto la restrizione ad W^\perp definisce un'applicazione lineare *unitaria* $L_U|_{W^\perp}: W^\perp \rightarrow W^\perp$. Questa applicazione ha almeno un autovalore, per cui esistono $\sigma \in \mathbf{C}$ e un vettore $\mathbf{x} \in W^\perp$ tali che $L_U(\mathbf{x}) = \sigma\mathbf{x}$. Questo contraddice la definizione di W , e implica $W^\perp = \{0\}$. In particolare, $\mathbf{C}^n = V_{\lambda_1} \oplus \dots \oplus V_{\lambda_k}$ e ogni autospazio di U ha dimensione massima, uguale alla molteplicità algebrica dell'autovalore corrispondente.

Come nel caso hermitiano, dai fatti (i) e (ii) segue che

- (iii) *Esiste una base ortonormale di \mathbf{C}^n formata da autovettori di U .*
- (iv) *La matrice U è diagonalizzabile mediante una matrice unitaria.*