Dipartimento di matematica
Università di Milano

Introduzione all'algebra lineare (versione preliminare)

Giuseppe Canuto Ottavio G. Rizzo

Versione del 27 febbraio 2007

Indice

1 Gruppi, anelli, campi

1.1 Gruppi

1.2 Anelli

1.3 Campi

1.4 Il gruppo S_n

1.5 L'anello $K[X]$

2 Spazi vettoriali

2.1 Definizioni

Sia K il campo dei numeri reali o dei numeri complessi. Uno **spazio vettoriale** V/K è un'insieme V, i cui elementi sono detti **vettori**, su cui sono definite un'operazione di **somma** ed un'operazione di **prodotto di un vettore per uno scalare**:

$$\begin{array}{ccc} V \times V \longrightarrow V & \qquad & K \times V \longrightarrow V \\ (\mathbf{v}_1, \mathbf{v}_2) \longmapsto \mathbf{v}_1 + \mathbf{v}_2 & \qquad & (\alpha, \mathbf{v}) \longmapsto \alpha\mathbf{v} \end{array}$$

tali che, per ogni $\alpha, \alpha_1, \alpha_2 \in K$ e per ogni $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in V$:

1. $(\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3 = \mathbf{v}_1 + (\mathbf{v}_2 + \mathbf{v}_3)$
2. Esiste un vettore $\mathbf{0}_V$ tale che $\mathbf{v} + \mathbf{0}_V = \mathbf{0}_V + \mathbf{v} = \mathbf{v}$, per ogni $\mathbf{v} \in V$
3. Dato $\mathbf{v} \in V$, esiste un vettore $-\mathbf{v} \in V$ tale che $\mathbf{v} + (-\mathbf{v}) = -\mathbf{v} + \mathbf{v} = \mathbf{0}_V$
4. $\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_2 + \mathbf{v}_1$
5. $(\alpha_1 + \alpha_2)\mathbf{v} = \alpha_1\mathbf{v} + \alpha_2\mathbf{v}$
6. $\alpha(\mathbf{v}_1 + \mathbf{v}_2) = \alpha\mathbf{v}_1 + \alpha\mathbf{v}_2$
7. $\alpha_1(\alpha_2\mathbf{v}) = (\alpha_1\alpha_2)\mathbf{v}$
8. $1\mathbf{v} = \mathbf{v}$

Gli elementi del campo K saranno detti **scalari**. Il vettore $\mathbf{0}_V$ è detto **vettore nullo** di V, mentre $-\mathbf{v}$ è detto **opposto** di $\mathbf{v}$.

Le condizioni 1, 2, 3, 4 significano che V con l'operazione di somma è un gruppo commutativo, che indicheremo $V(+)$.

Le condizioni 5, 6 sono dette **proprietà distributive**.

Esempi

1. **Spazi K^n**

   Sia K^n lo spazio delle n-uple $(x_1, \ldots, x_n)$ con $x_i \in K$, per $i = 1, \ldots, n$. Se definiamo

   $$(x_1, \ldots, x_n) + (x_1', \ldots, x_n') = (x_1 + x_1', \ldots, x_n + x_n')$$
   $$\alpha(x_1, \ldots, x_n) = (\alpha x_1, \ldots, \alpha x_n)$$

   otteniamo uno spazio vettoriale su K che denotiamo K^n. Nel caso $n = 2$, useremo spesso la notazione (x, y) per gli elementi di K^2; nel caso $n = 3$, scriveremo (x, y, z).

2. **Polinomi** $K[X]$

   Sia $K[X]$ l'insieme dei polinomi in una variabile X a coefficienti in K: con le usuali operazioni di somma e di prodotto per una costante, forma uno spazio vettoriale su K.

3. **Funzioni continue**

   Sia I un intervallo della retta reale; l'insieme $C_{\mathbf{R}}(I)$ delle funzioni continue definite su I a valori in $\mathbf{R}$ è uno spazio vettoriale con le operazioni di somma di funzioni e di moltiplicazione per una costante.

Convenzione

Scriveremo semplicemente $\mathbf{v} - \mathbf{w}$ per indicare $\mathbf{v} + (-\mathbf{w})$.

Proprietà elementari. Abbiamo che:

1. $0\mathbf{v} = \mathbf{0}_V$, per ogni $\mathbf{v} \in V$

   Infatti $0\mathbf{v} = (0 + 0)\mathbf{v} = 0\mathbf{v} + 0\mathbf{v}$.

2. $\alpha\mathbf{0}_V = \mathbf{0}_V$, per ogni $\alpha \in K$

   Infatti $\alpha\mathbf{0}_V = \alpha(\mathbf{0}_V + \mathbf{0}_V) = \alpha\mathbf{0}_V + \alpha\mathbf{0}_V$.

3. Se $\alpha\mathbf{v} = \mathbf{0}_V$, allora $\alpha = 0$ oppure $\mathbf{v} = \mathbf{0}_V$

   Infatti se $\alpha \neq 0$, esiste α^{-1} e quindi $\mathbf{v} = (\alpha^{-1}\alpha)\mathbf{v} = \alpha^{-1}(\alpha\mathbf{v}) = \mathbf{0}_V$.

4. $-(\alpha\mathbf{v}) = (-\alpha)\mathbf{v} = \alpha(-\mathbf{v})$

   Verifichiamo la prima uguaglianza, mostrando che $\alpha\mathbf{v}$ è l'opposto di $(-\alpha)\mathbf{v}$: $\alpha\mathbf{v} + (-\alpha)\mathbf{v} = (\alpha - \alpha)\mathbf{v} = 0\mathbf{v} = \mathbf{0}_V$. Analogamente si verifica che il primo e il terzo vettore sono uguali.

5. $-\mathbf{v} = (-1)\mathbf{v}$

   Questo è il caso particolare del precedente, in cui $\alpha = 1$.

Sottospazi vettoriali

Sia V/K uno spazio vettoriale. Un sottoinsieme non vuoto Z di V è detto **sottospazio vettoriale** di V se sono verificate le due condizioni seguenti:

1. Se $\mathbf{v}_1, \mathbf{v}_2 \in Z$, allora $\mathbf{v}_1 + \mathbf{v}_2 \in Z$
2. Se $\mathbf{v} \in Z$ ed $\alpha \in K$, allora $\alpha\mathbf{v} \in Z$

Se Z è un sottospazio vettoriale di V, allora Z è uno spazio vettoriale con le operazione di somma e prodotto per uno scalare definite su V. Si osservi che dalla prima (o dalla seconda) segue subito che $\mathbf{0}_V$ appartiene ad ogni sottospazio vettoriale di V. Infatti, usando la seconda condizione: preso $\mathbf{v} \in Z$, allora $0\mathbf{v} = \mathbf{0}_V \in Z$.

Esempi

1. Sia $Z = \{\mathbf{0}_V\}$. Il sottoinsieme costituito dal solo vettore nullo è un sottospazio vettoriale.

2. L'insieme $K_d[X]$ dei polinomi di grado $\leqslant d$ è un sottospazio vettoriale di $K[X]$.

3. Il sottoinsieme $Z \subseteq \mathbf{R}^2$ formato dai vettori della forma $(x, 2x)$ è un sottospazio vettoriale di $\mathbf{R}^2$.

4. Se V_1 e V_2 sono due sottospazi vettoriali di uno spazio vettoriale V, allora $V_1 \cap V_2$ è un sottospazio vettoriale di V.

Combinazioni lineari

Siano $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ dei vettori in uno spazio vettoriale V. Un vettore $\mathbf{v}$ della forma $\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k$, con gli $\alpha_i \in K$ è detto **combinazione lineare** dei vettori $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$. Gli scalari $\alpha_1, \dots, \alpha_k$ sono anche detti **coefficienti** della combinazione lineare.

L'insieme delle combinazioni lineare di $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ è un sottospazio vettoriale di V, come si verifica immediatamente, che denotiamo $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$. Tale spazio è detto anche **spazio generato dai vettori** $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$; l'insieme $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ è detto **sistema di generatori** di V. Esso contiene ciascuno dei vettori $\mathbf{v}_i$ ed inoltre, se $Z \subseteq V$ è un sottospazio vettoriale contenente tutti i $\mathbf{v}_i$, allora $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) \subseteq Z$.

Uno spazio vettoriale V è detto **finitamente generato** se esiste un insieme finito $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ di vettori tale che $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$.

Osserviamo inoltre che, se $\mathbf{v}$ è combinazione lineare di $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$, allora $\mathcal{L}(\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k)$.

Esempi

1. K^n è finitamente generato. Infatti i vettori $(1, 0, \dots, 0)$, $(0, 1, 0, \dots, 0)$, $(0, \dots, 0, 1)$ formano un sistema di generatori. Tali vettori sono detti **versori** di K^n.

2. $K[X]$ non è finitamente generato su K. Infatti se così fosse, sia $\{P_1(X), \dots, P_k(X)\}$ un insieme di generatori dove $P_i(X)$ è un polinomio di grado d_i. Un polinomio $P(X)$ di grado $d > \max_{i=1}^{k}\{d_i\}$ non può essere combinazione lineare dei $\{P_i(X)\}_{i=1}^{k}$.

2.2 Applicazioni lineari

Definizione. Siano V e W due spazi vettoriali su K. Un'applicazione $\varphi\colon V \to W$ è detta **lineare** se valgono le seguenti proprietà:

$$\varphi(\mathbf{v}_1 + \mathbf{v}_2) = \varphi(\mathbf{v}_1) + \varphi(\mathbf{v}_2), \quad \text{per ogni } \mathbf{v}_1, \mathbf{v}_2 \in V$$
$$\varphi(\alpha\mathbf{v}) = \alpha\varphi(\mathbf{v}), \quad \text{per ogni } \alpha \in K, \mathbf{v} \in V$$

Se φ è un'applicazione lineare, allora $\varphi(\mathbf{o}_V) = \mathbf{o}_W$, cioè il vettore nullo del primo spazio vettoriale viene sempre mandato nel vettore nullo del secondo spazio vettoriale. Infatti abbiamo $\varphi(\mathbf{o}_V) = \varphi(\mathbf{o}_V + \mathbf{o}_V) = \varphi(\mathbf{o}_V) + \varphi(\mathbf{o}_V)$.

Data un'applicazione lineare φ ed un vettore $\alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k$, combinazione lineare dei $\mathbf{v}_i \in V$ con $\alpha_i \in K$, avremo: $\varphi(\alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k) = \alpha_1\varphi(\mathbf{v}_1) + \cdots + \alpha_k\varphi(\mathbf{v}_k)$.

L'insieme delle applicazioni lineari da V in W sarà denotato $\mathrm{Hom}_K(V, W)$. Tale insieme è esso stesso uno spazio vettoriale con le operazioni di somma e prodotto per uno scalare definite nel modo seguente: dati $\varphi, \psi \in \mathrm{Hom}_K(V, W)$, $\mathbf{x} \in V$ ed $\alpha \in K$:

$$(\varphi + \psi)(\mathbf{x}) = \varphi(\mathbf{x}) + \psi(\mathbf{x}), \qquad (\alpha\varphi)(\mathbf{x}) = \alpha\varphi(\mathbf{x})$$

Esempi

Lasciamo al lettore la semplice verifica che le seguenti applicazioni sono tutte lineari.

1. L'applicazione $V \to W$ che manda ogni $\mathbf{v} \in V$ in $\mathbf{o}_W$. Quest'applicazione viene detta talvolta **applicazione nulla**.
2. La proiezione $\pi: \mathbf{R}^3 \to \mathbf{R}^2$ definita da $\pi(x, y, z) = (x, y)$.
3. La derivazione $D : K[X] \to K[X]$. Se $P(X) = \sum_{k=0}^{n} a_k X^k$, allora $D\big(P(X)\big) = \sum_{k=1}^{n} k a_k X^{k-1}$.

Definizione. Sia $\varphi: V \to W$ un'applicazione lineare. Il **nucleo** di φ, denotato $\ker\varphi$ è il sottoinsieme di V che viene mandato in $\mathbf{o}_W$: $\ker\varphi = \{\mathbf{v} \in V : \varphi(\mathbf{v}) = \mathbf{o}_W\}$. L'**immagine** di φ, denotata $\mathrm{Im}\,\varphi$, è il sottoinsieme di W immagine di V.

Sia $\varphi: V \to W$ un'applicazione lineare. È allora immediato verificare che il nucleo $\ker\varphi$ è sottospazio vettoriale di V e che l'immagine $\mathrm{Im}\,\varphi$ è sottospazio vettoriale di W.

Esempi

1. Nell'esempio 1 di sopra, abbiamo che il nucleo è V mentre l'immagine è $\{\mathbf{o}_W\}$.
2. Nell'esempio 2, abbiamo $\ker\pi = \{(0, 0, z) \in \mathbf{R}^3 : z \in \mathbf{R}\}$, l'asse delle z.
3. Nell'esempio 3, il nucleo è l'insieme dei polinomi costanti, mentre l'immagine è tutto $K[X]$, cioè D è suriettiva.

Data un'applicazione lineare $\varphi: V \to W$, avremo che φ è suriettiva se $\mathrm{Im}\,\varphi = W$. Per quanto riguarda l'iniettività, abbiamo:

Proposizione 2.1

Un'applicazione lineare $\varphi: V \to W$ è iniettiva se e solo $\ker\varphi = \{\mathbf{o}_V\}$.

Dimostrazione. Come abbiamo visto, $\varphi(\mathbf{o}_V) = \mathbf{o}_W$.

Se φ è iniettiva, l'unico vettore la cui immagine è $\mathbf{o}_W$ è $\mathbf{o}_V$, cioè $\ker \varphi = \{\mathbf{o}_V\}$.

Viceversa, se $\ker \varphi = \{\mathbf{o}_V\}$, supponiamo che $\varphi(\mathbf{v}_1) = \varphi(\mathbf{v}_2)$: allora $\varphi(\mathbf{v}_1 - \mathbf{v}_2) = \varphi(\mathbf{v}_1) - \varphi(\mathbf{v}_2) = \mathbf{o}_W$ implica che $\mathbf{v}_1 - \mathbf{v}_2 \in \ker \varphi$, cioè $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{o}_V$. □

Definizione. Un'applicazione lineare biiettiva $\varphi \colon V \to W$ è detta **isomorfismo** (di V in W).

Proposizione 2.2
Se $\varphi \colon V \to W$ è un isomorfismo, allora l'applicazione inversa $\varphi^{-1} : W \to V$ è anch'essa lineare.

Dimostrazione. Esercizio. □

L'insieme degli isomorfismi da V in W è denotato $\mathrm{Isom}_K(V, W)$.

2.3 Indipendenza lineare e basi

Siano $\mathbf{v}_1, \dots, \mathbf{v}_n$ vettori dello spazio vettoriale V. Diciamo che essi sono **linearmente dipendenti** se esistono degli scalari $\alpha_1, \dots, \alpha_n$ non tutti nulli tali che $\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{o}_V$; sono detti **linearmente indipendenti** in caso contrario. In altre parole, i vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente indipendenti se il vettore $\mathbf{o}_V$ si ottiene come combinazione lineare dei $\mathbf{v}_1, \dots, \mathbf{v}_n$ in un solo modo: prendendo tutte le costanti $\alpha_i = 0$.

Si osservi che se uno dei vettori $\mathbf{v}_i$ è $\mathbf{o}_V$, allora essi sono linearmente dipendenti. Infatti, sia $\mathbf{v}_1 = \mathbf{o}_V$: se prendiamo $\alpha_1 = 1$, $\alpha_2 = \alpha_3 = \cdots = \alpha_n = 0$ otteniamo una combinazione a coefficienti non tutti nulli che dà il vettore $\mathbf{o}_V$.

Si osservi anche che, se $\mathbf{v}$ è combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_n$, allora i vettori $\mathbf{v}, \mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente dipendenti. Infatti se $\mathbf{v} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$, allora $\mathbf{o}_V = \mathbf{v} - \alpha_1\mathbf{v}_1 - \cdots - \alpha_n\mathbf{v}_n$: quest'ultima è una combinazione lineare a coefficienti non tutti nulli.

Definizione. Sia V uno spazio vettoriale su K finitamente generato. Un insieme $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ di vettori di V è una **base** di V se:

a) $V = \mathcal{L}(\mathbf{v}_1, \dots, \mathbf{v}_n)$, cioè $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è un sistema di generatori di V
b) Ogni vettore $\mathbf{v} \in V$ si esprime in modo unico come combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_n$

La condizione b) può essere sostituita con la condizione

b') I vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente indipendenti

Infatti, supponiamo che valgano le condizioni a) e b), e supponiamo che $\sum_{i=0}^n \alpha_i\mathbf{v}_i = \mathbf{o}_V$. Per l'unicità della scrittura di $\mathbf{o}_V$ come combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_n$, deve essere $\alpha_1 = \cdots = \alpha_n = 0$. Viceversa, supponiamo valgano a) e b'): se $\mathbf{v} = \sum_{i=0}^n \alpha_i\mathbf{v}_i = \sum_{i=0}^n \alpha'_i\mathbf{v}_i$, allora $\mathbf{o}_V = \sum_{i=0}^n (\alpha_i - \alpha'_i)\mathbf{v}_i$. Data l'indipendenza lineare dei vettori $\mathbf{v}_i$, otteniamo che $\alpha_i = \alpha'_i$ per ogni $i = 1, \dots, n$.

Esempio

In K^n i versori costituiscono una base, detta **base canonica** di K^n.

Proposizione 2.3 (Criterio di indipendenza)
Dati i vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$*, le due condizioni seguenti sono equivalenti:*

a) I vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ *sono indipendenti.*
b) $\mathbf{v}_1 \neq \mathbf{o}$ *e* $\mathbf{v}_k \notin \mathcal{L}(\mathbf{v}_1, \dots, \mathbf{v}_{k-1})$ *per ogni* $k = 2, \dots, n$.

Dimostrazione. Supponiamo che valga la prima condizione: in particolare nessuno di essi è $\mathbf{o}_V$. Inoltre, se $\mathbf{v}_k = \sum_{i=1}^{k-1} \alpha_i \mathbf{v}_i$ allora $\mathbf{v}_k - \sum_{i=1}^{k-1} \alpha_i \mathbf{v}_i = \mathbf{o}_V$ è, contro le ipotesi, una combinazione lineare a coefficienti non tutti nulli per il vettore $\mathbf{o}_V$.

Viceversa, supponiamo che $\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{o}_V$. Se $\alpha_n \neq 0$, possiamo moltiplicare per α_n^{-1} ed esprimere $\mathbf{v}_n$ come combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_n$, contro le ipotesi. Quindi $\alpha_n = 0$. Ripetendo il ragionamento si conclude che $\alpha_i = 0$ per ogni $i = n, \dots, 1$. □

Teorema 2.4 (Estrazione di una base)
Sia $\mathsf{G} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ *un sistema di generatori di* V *con* $V \neq \{\mathbf{o}_V\}$*. Allora esiste una base* $\mathsf{B} \subseteq \mathsf{G}$.

Dimostrazione. Ovviamente possiamo supporre che tutti i vettori di G siano diversi da $\mathbf{o}_V$. Eliminiamo il primo vettore che è combinazione lineare dei precedenti, se esiste: lo spazio generato dai rimanenti è ancora V. Ripetiamo il ragionamento finché non esistono più vettori combinazione lineare dei precedenti: l'insieme residuo B è ancora un sistema di generatori di V, ed è formato da vettori linearmente indipendenti per il criterio precedente. □

Teorema 2.5 (Completamento a una base)
Sia $\mathsf{C} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ *un insieme di vettori linearmente indipendenti dello spazio vettoriale* V *finitamente generato. Allora esiste una base* $\mathsf{B} \supseteq \mathsf{C}$.

Dimostrazione. Sia $\mathsf{E} = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ un sistema di generatori di V. Allora $\mathsf{C} \cup \mathsf{E}$ è un sistema di generatori di V. Ordiniamo i vettori di questo insieme mettendo prima i vettori di C e poi i vettori di E. Applicando la procedura del teorema 2.4 otteniamo una base $\mathsf{B} \subseteq \mathsf{C} \cup \mathsf{E}$. Per il modo in cui la base B è stata costruita, tutti i vettori di C sono contenuti in B. □

Dimensione

Senza dimostrazione

Sia V uno spazio vettoriale finitamente generato. Allora le basi di V hanno tutte la stessa cardinalità n. Definiamo la **dimensione** di V su K la cardinalità n di una sua base qualsiasi, e scriviamo $\dim_K V = n$.

Convenzione

Lo spazio vettoriale $V = \{\mathbf{o}_V\}$ ha dimensione zero.

Esempi

1. K^n è uno spazio vettoriale di dimensione n su K. Una sua base è quella canonica data dai versori.

2. $K_d[X]$, lo spazio vettoriale dei polinomi di grado $\leqslant d$, ha dimensione $d+1$. Una sua base è $\{1, X, \ldots, X^d\}$.

Componenti

Sia $\mathsf{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base dello spazio vettoriale V/K. Se $\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots \alpha_n\mathbf{v}_n$ è un vettore di V, diremo che $\alpha_1, \ldots, \alpha_n$ sono le **componenti** di $\mathbf{v}$ nella base B.

La scelta della base B determina un'isomorfismo

$$\begin{aligned} i_{\mathsf{B}} \colon V &\longrightarrow K^n \\ \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n &\longmapsto (\alpha_1, \ldots, \alpha_n) \end{aligned}$$

che associa ad un vettore $\mathbf{v}$ la n-upla delle sue componenti nella base B.

Proposizione 2.6

Sia V uno spazio vettoriale su K di dimensione n. Dati m vettori $\mathbf{v}_1, \ldots, \mathbf{v}_m$ di V, allora

a) *Se $V = \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ allora $m \geqslant n$.*
b) *Se $\mathbf{v}_1, \ldots, \mathbf{v}_m$ sono linearmente indipendenti, allora $m \leqslant n$.*

Dimostrazione. Le due affermazioni sono conseguenza immediata dei teoremi 2.4 e 2.5. □

Proposizione 2.7

Sia V uno spazio vettoriale su K di dimensione n, e sia $\mathsf{E} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un insieme formato da n vettori di V. Allora sono equivalenti:

a) *E è una base*
b) *I vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sono linearmente indipendenti*
c) *$V = \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$*

Dimostrazione. Esercizio □

Proposizione 2.8

Sia V/K uno spazio vettoriale di dimensione n, e W un suo sottospazio. Allora

a) *W è finitamente generato con* $\dim W \leqslant \dim V$
b) *$W = V$ se e solo se* $\dim W = \dim V$

Dimostrazione.

a) Supponiamo $W \neq \{\mathbf{o}_V\}$ (caso banale). Sia $\mathbf{w}_1 \in W$ con $\mathbf{w}_1 \neq \mathbf{o}_V$. Se $W = \mathcal{L}(\mathbf{w}_1)$, allora $\{\mathbf{w}_1\}$ è una base di W e $\dim W = 1$. Se $W \neq \mathcal{L}(\mathbf{w}_1)$, allora esiste $\mathbf{w}_2 \in W$ con $\mathbf{w}_2 \notin \mathcal{L}(\mathbf{w}_1)$. Se $W = \mathcal{L}(\mathbf{w}_1, \mathbf{w}_2)$, allora $\{\mathbf{w}_1, \mathbf{w}_2\}$ è una base di W e $\dim W = 2$. In caso contrario, possiamo continuare: dopo m passi avremo m vettori $\mathbf{w}_1, \dots, \mathbf{w}_m$ che, per il Criterio di indipendenza lineare, sono linearmente indipendenti. Per la proposizione precedente $m \leqslant n$, quindi il procedimento termina per un qualche $m_0 \leqslant n$. Questo significa che $\{\mathbf{w}_1, \dots, \mathbf{w}_{m_0}\}$ è una base di W e che $\dim W = m_0$.

b) In un senso l'affermazione è ovvia. Se invece $\dim W = \dim V$, sia $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ una base di W: questi vettori sono linearmente indipendenti e quindi costituiscono una base di V per la proposizione 2.7. Segue che $V = W$.

□

Proposizione 2.9

Sia $\varphi\colon V \to W$ un'applicazione lineare. Siano $\mathbf{v}_1, \dots, \mathbf{v}_n$ vettori di V. Allora

1. *Se $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ sono linearmente dipendenti, allora anche $\{\varphi(\mathbf{v}_1), \dots, \varphi(\mathbf{v}_n)\}$ lo sono (equivalentemente, se $\{\varphi(\mathbf{v}_1), \dots, \varphi(\mathbf{v}_n)\}$ sono indipendenti, anche $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ sono indipendenti).*
2. *Se φ è iniettiva e se $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ sono indipendenti, allora anche $\{\varphi(\mathbf{v}_1), \dots, \varphi(\mathbf{v}_n)\}$ sono indipendenti.*
3. *Se $V = \mathcal{L}(\mathbf{v}_1, \dots, \mathbf{v}_n)$, allora* $\operatorname{Im}\varphi = \mathcal{L}\big(\varphi(\mathbf{v}_1), \dots, \varphi(\mathbf{v}_n)\big)$.

Dimostrazione.

1. Se $\mathbf{o}_V = \sum_i \alpha_i \mathbf{v}_i$, allora $\mathbf{o}_W = \varphi(\mathbf{o}_V) = \sum_i \alpha_i \varphi(\mathbf{v}_i)$.
2. Se $\mathbf{o}_W = \sum_i \alpha_i \varphi(\mathbf{v}_i)$, allora $\varphi\big(\sum_i \alpha_i \mathbf{v}_i\big) = \mathbf{o}_W$ cioè $\sum_i \alpha_i \mathbf{v}_i \in \ker\varphi$. Per l'iniettività di φ, la somma vale $\mathbf{o}_V$; ma i vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono indipendenti, quindi $\alpha_1 = \dots = \alpha_n = 0$.
3. Se $\mathbf{w} \in \operatorname{Im}\varphi$, allora $\mathbf{w} = \varphi(\mathbf{v})$ per qualche $\mathbf{v} = \sum \alpha_i \mathbf{v}_i$, quindi $\mathbf{w} = \sum \alpha_i \varphi(\mathbf{v}_i)$.

□

Corollario 2.10

Sia $\varphi\colon V \to W$ un isomorfismo, allora

1. *I vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ di V sono linearmente indipendenti se e solo se i vettori $\varphi(\mathbf{v}_1), \dots, \varphi(\mathbf{v}_n)$ di W lo sono.*
2. *$V = \mathcal{L}(\mathbf{v}_1, \dots, \mathbf{v}_n)$ se e solo $W = \mathcal{L}\big(\varphi(\mathbf{v}_1), \dots, \varphi(\mathbf{v}_n)\big)$.*
3. *I vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono una base di V se e solo se i vettori $\varphi(\mathbf{v}_1), \dots, \varphi(\mathbf{v}_n)$ sono una base di W.*

Dimostrazione. Il corollario si ottiene dalla proposizione precedente applicata alla φ ed alla sua inversa φ^{-1}. □

2.4 Somma diretta e somma

Siano V_1 e V_2 due spazi vettoriali su K. Costruiamo un nuovo spazio vettoriale $V_1 \oplus V_2$, la **somma diretta** di V_1 e V_2. Gli elementi di $V_1 \oplus V_2$ sono le coppie $(\mathbf{v}_1, \mathbf{v}_2) \in V_1 \times V_2$. Le operazioni di somma e prodotto per uno scalare sono definite come segue:

$$\begin{aligned}(\mathbf{v}_1, \mathbf{v}_2) + (\mathbf{v}_1', \mathbf{v}_2') &= (\mathbf{v}_1 + \mathbf{v}_1', \mathbf{v}_2 + \mathbf{v}_2') \\ \alpha(\mathbf{v}_1, \mathbf{v}_2) &= (\alpha\mathbf{v}_1, \alpha\mathbf{v}_2)\end{aligned}$$

Se $V_1 = V_2 = K$, allora $K \oplus K = K^2$.

Possiamo iterare il procedimento: cioè dati tre spazi vettoriali V_1, V_2, V_3 su K, consideriamo lo spazio vettoriale $V_1 \oplus V_2 \oplus V_3$, lo spazio delle terne $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ con le operazioni di somma e prodotto per uno scalare definite nel modo ovvio. È chiaro che

$$\begin{aligned}(V_1 \oplus V_2) \oplus V_3 &\longrightarrow V_1 \oplus V_2 \oplus V_3 \\ \big((\mathbf{v}_1, \mathbf{v}_2), \mathbf{v}_3\big) &\longmapsto (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)\end{aligned}$$

definisce un isomorfismo. Dati k spazi vettoriali $V_1, V_2, \dots, V_k$ su K, indichiamo con $\bigoplus_{i=1}^k V_i$ la somma diretta di questi spazi.

Se $\mathsf{E}_\alpha = \{\mathbf{e}_1^\alpha, \dots \mathbf{e}_{n_\alpha}^\alpha\}$ è una base di V_α, con $\alpha = 1, \dots, k$, consideriamo i seguenti vettori di $\bigoplus_{i=1}^k V_i$:

$$\mathbf{v}_j^\alpha = (\mathbf{0}, \dots, \mathbf{0}, \mathbf{e}_j^\alpha, \mathbf{0}, \dots, \mathbf{0}) \tag{2.1}$$

la cui α-esima componente è un vettore della base E_α di V_α, mentre le altre componenti sono $\mathbf{0}$; dove $\alpha = 1, \dots, k$ ed $j = 1, \dots, n_\alpha$. L'insieme di tali vettori forma una base F di $\bigoplus_{i=1}^k V_i$; in particolare,

$$\dim \bigoplus_{i=1}^k V_i = n_1 + n_2 + \dots + n_k = \sum_{i=1}^k \dim V_i$$

Somma

Siano ora V_1, V_2 due sottospazi di uno spazio vettoriale V/K. Consideriamo l'insieme $V_1 + V_2 = \{\mathbf{v}_1 + \mathbf{v}_2 : \mathbf{v}_1 \in V_1, \mathbf{v}_2 \in V_2\} \subseteq V$. È immediato verificare che $V_1 + V_2$ è sottospazio vettoriale di V. Inoltre, $V_1, V_2 \subseteq V$; ed ogni sottospazio $Z \subseteq V$ contenente sia V_1 che V_2 contiene anche $V_1 + V_2$. Chiamiamo $V_1 + V_2$ la **somma** (in V) dei sottospazi V_1 e V_2.

Lo spazio $V_1 + V_2$ è caratterizzato dalla seguente proprietà: se W è un sottospazio di V che contiene sia V_1 che V_2, allora V contiene anche $V_1 + V_2$.

Se V_1 e V_2 sono sottospazi di V, consideriamo l'applicazione

$$\begin{aligned}\varphi\colon (V_1 \oplus V_2) &\longrightarrow V_1 + V_2 \\ (\mathbf{v}_1, \mathbf{v}_2) &\longmapsto \mathbf{v}_1 + \mathbf{v}_2\end{aligned}$$

Quest'applicazione è lineare e suriettiva; il suo nucleo è costituito dalle coppie $\{(\mathbf{v}_1, \mathbf{v}_2) : \mathbf{v}_1 = -\mathbf{v}_2\}$ cioè dalle coppie $\{(\mathbf{v}, -\mathbf{v}) : \mathbf{v} \in V_1 \cap V_2\}$. Il nucleo di φ è dunque isomorfo allo spazio $V_1 \cap V_2$ e l'applicazione φ è un isomorfismo se e solo se $V_1 \cap V_2 = \{\mathbf{o}_V\}$.

In generale, dati i sottospazi $V_1, \ldots, V_k$ di V indichiamo con $V_1 + \cdots + V_k = \sum_i V_i$ l'insieme dei vettori della forma $\mathbf{v}_1 + \cdots + \mathbf{v}_k$ con $\mathbf{v}_i \in V_1$. Tale insieme è un sottospazio vettoriale di V e, analogamente al caso $k = 2$, è il più piccolo sottospazio vettoriale di V contenente $V_1, \ldots, V_k$. Osserviamo che se E_i è un sistema di generatori di V_i, con $i = 1, \ldots, k$, cioè se $V_i = \mathcal{L}(\mathsf{E}_i)$, allora l'insieme $\mathsf{E} = E_1 \cup E_2 \cup \ldots E_k$ è un sistema di generatori di $\sum V_i$..

Se $V_1, \ldots, V_k$ sono sottospazi di V, analogamente al caso $k = 2$, esiste un'applicazione lineare e suriettiva

$$\begin{aligned} \varphi\colon \bigoplus_{1=1}^{k} V_i &\longrightarrow \sum_{1=1}^{k} V_i \\ (\mathbf{v}_1, \ldots, \mathbf{v}_k) &\longmapsto \mathbf{v}_1 + \cdots + \mathbf{v}_k \end{aligned}$$

Abbiamo il seguente criterio per la somma diretta che generalizza quanto visto nel caso $k = 2$:

Proposizione 2.11 (Criterio per la somma diretta)
Sia V uno spazio vettoriale su K, e siano $V_1, \ldots, V_k$ sottospazi vettoriali di V. Allora l'applicazione φ è un isomorfismo se e solo se $V_i \cap \sum_{j=1}^{i-1} V_j = \{\mathbf{o}\}$ per $i = 1, \ldots, k$.

Dimostrazione. L'applicazione φ è definita, analogamente al caso $k = 2$, da $\varphi(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \mathbf{v}_1 + \cdots + \mathbf{v}_k$. L'applicazione φ è un isomorfismo se e solo se $\ker \varphi = (\mathbf{o}, \ldots, \mathbf{o})$; cioè se l'unico modo per esprimere il vettore $\mathbf{o}_V$ come elemento di $\sum V_i$ è $\mathbf{o} + \cdots + \mathbf{o}$.

Equivalentemente, φ è un isomorfismo se e solo se ogni vettore della somma si scrive in modo unico come $\mathbf{v}_1 + \cdots + \mathbf{v}_k$ con $\mathbf{v}_i \in V_i$. Infatti, date due scritture

$$\mathbf{v}_1 + \cdots + \mathbf{v}_k = \mathbf{w}_1 + \cdots + \mathbf{w}_k, \qquad \text{con } \mathbf{v}_i, \mathbf{w}_i \in V_i, \text{ per ogni } i = 1, \ldots k \tag{2.2}$$

abbiamo

$$\mathbf{o} = (\mathbf{v}_1 - \mathbf{w}_1) + \cdots + (\mathbf{v}_k - \mathbf{w}_k)$$

Quindi $\mathbf{o}$ si scrive in modo unico se e solo se $\mathbf{v}_i = \mathbf{w}_i$ per ogni i.

Supponiamo dunque che φ sia un isomorfismo; se $\mathbf{v}_i \in V_i \cap \sum_{j=1}^{i-1} V_j$, allora

$$\mathbf{v}_i = \mathbf{v}_1 + \cdots + \mathbf{v}_{i-1}, \qquad \text{dove } \mathbf{v}_j \in V_j$$

cioè

$$\mathbf{o} = \mathbf{v}_i - \mathbf{v}_1 - \cdots - \mathbf{v}_{i-1}$$

Per l'unicità della scrittura del vettore $\mathbf{o}$, avremo che $\mathbf{v}_j = \mathbf{o}$ per ogni $j = 1, \ldots, i$. Viceversa supponiamo che valga l'equazione (2.2), allora posto $\mathbf{z}_k = \mathbf{v}_k - \mathbf{w}_k$, abbiamo

$$\mathbf{o} = \mathbf{z}_1 + \cdots + \mathbf{z}_k$$

Sia i è il massimo degli indici $1, \ldots, k$ tale che $\mathbf{z}_k \neq \mathbf{o}$: se $i \geqslant 1$ avremo

$$\mathbf{o} \neq \mathbf{z}_i = -\mathbf{z}_1 - \cdots - \mathbf{z}_{i-1} \in V_i \cap \sum_{j=1}^{i-1} V_j$$

contro l'ipotesi. □

Corollario 2.12

Con le notazione della proposizione precedente, siano E_i *una base di* V_i *ed* $n_i = \dim V_i$. *Allora* φ *è un isomorfismo se e solo se l'insieme* $\mathsf{E} = E_1 \cup \cdots \cup E_k$ *è formato da* $n_1 + \cdots + n_k$ *vettori linearmente indipendenti.*

Dimostrazione. Sappiamo che E è un sistema di generatori di $\sum V_i$. Se E è formato da $n_1 + \cdots + n_k$ vettori linearmente indipendenti, allora E è una base e la dimensione della somma è $n_1 + \cdots + n_k$.

Essendo φ un'applicazione suriettiva fra spazi vettoriali finiti della stessa dimensione, è un isomorfismo.

Viceversa, se φ è un isomorfismo, consideriamo la base F formata dai vettori $\mathbf{v}_j^\alpha$. Come si vede immediatamente, $\varphi(\mathsf{F}) = \mathsf{E}$: dato che φ è un isomorfismo, anche E è una base, formata da $n_1 + \cdots + n_k$ vettori indipendenti. dell'equazione (2.1) □

2.5 Nullità più rango

Teorema 2.13

Sia V *uno spazio vettoriale di dimensione finita, e sia* $\varphi\colon V \to W$ *un'applicazione lineare. Allora*

$$\dim V = \dim\ker\varphi + \dim\operatorname{Im}\varphi$$

Dimostrazione. Sia $n = \dim V$. Sia $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ una base di $\ker\varphi$; per il teorema di completamento ad una base, possiamo completare tale insieme ad una base $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$. Sia $Z = \mathcal{L}(\mathbf{v}_{k+1}, \dots, \mathbf{v}_n)$: chiaramente $\{\mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ è una sua base e quindi $\dim Z = n - k$. Abbiamo che $\ker\varphi \cap Z = \{\mathbf{o}_V\}$: infatti, se $\mathbf{v} \in \ker\varphi \cap Z$, allora $\mathbf{v} = \sum_{i=1}^k \alpha_i \mathbf{v}_i$ essendo $\mathbf{v} \in \ker\varphi$ ed anche $\mathbf{v} = \sum_{i=k+1}^n \alpha_i \mathbf{v}_i$ essendo $\mathbf{v} \in Z$; quindi $\mathbf{o}_V = \sum_{i=1}^k \alpha_i \mathbf{v}_i - \sum_{i=k+1}^n \alpha_i \mathbf{v}_i$; ma $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base di V, e quindi le costanti α_i sono tutte nulle.

Sia $\varphi|_Z\colon Z \to W$ l'applicazione ottenuta dalla restrizione di φ ai soli vettori di Z: tale applicazione è ovviamente lineare e $\operatorname{Im}\varphi = \operatorname{Im}\varphi|_Z$. Infatti, se $\mathbf{w} = \varphi(\mathbf{v})$ con $\mathbf{v} = \sum_{i=1}^n \beta_i \mathbf{v}_i$, allora

$$\mathbf{w} = \varphi\left(\sum_{i=1}^k \beta_i \mathbf{v}_i\right) + \varphi\left(\sum_{i=k+1}^n \beta_i \mathbf{v}_i\right) = \mathbf{o}_W + \varphi\left(\sum_{i=k+1}^n \beta_i \mathbf{v}_i\right) \in \operatorname{Im}\varphi|_Z$$

L'applicazione $\varphi|_Z\colon Z \to \operatorname{Im}\varphi$ è quindi iniettiva e suriettiva, cioè è un isomorfismo. Segue che $\dim\operatorname{Im}\varphi = \dim Z = n - k$. □

Corollario 2.14

Sia $\varphi\colon V \to W$ *un'applicazione lineare fra spazi vettoriali. Se* $\dim V = \dim W$, *allora*

$$\varphi \text{ è un isomorfismo} \iff \varphi \text{ è iniettiva} \iff \varphi \text{ è suriettiva}$$

Dimostrazione. Sia $n = \dim V = \dim W$. Abbiamo che φ è iniettiva se e solo se $\ker\varphi = \{\mathbf{o}_V\}$, mentre φ è suriettiva se e solo se $\dim \operatorname{Im}\varphi = n$. Tenuto conto del teorema 2.13, le due condizioni sono equivalenti. □

Corollario 2.15

Sia V_1 e V_2 sottospazi di V/K. Allora

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)$$

Dimostrazione. Abbiamo visto, infatti, che $\dim(V_1 \oplus V_2) = \dim V_1 + \dim V_2$ e che esiste un'applicazione lineare e suriettiva $V_1 \oplus V_2 \to V_1 + V_2$ con un nucleo isomorfo a $V_1 \cap V_2$. □

2.6 Spazi quozienti

Sia V uno spazio vettoriale su K e Z un sottospazio vettoriale di V. Introduciamo in V la seguente relazione di equivalenza: $\mathbf{x} \sim \mathbf{y}$ se $\mathbf{x} - \mathbf{y} \in Z$: la verifica che questa è una relazione di equivalenza è lasciata al lettore.

Indichiamo con V/Z l'insieme delle classi di equivalenza rispetto alla relazione precedente. Se $\mathbf{x} \in V$ la sua classe di equivalenza sarà indicata $[\mathbf{x}]$. L'insieme V/Z ha una struttura naturale di spazio vettoriale su K, definita come segue:

$$[\mathbf{x}] + [\mathbf{y}] = [\mathbf{x} + \mathbf{y}], \qquad \alpha[\mathbf{x}] = [\alpha\mathbf{x}]$$

Lasciamo al lettore la verifica che le operazione precedenti sono ben definite, cioè non dipendono dalla scelta del rappresentante nelle varie classi di equivalenza, e che V/Z è uno spazio vettoriale con le operazioni sopra definite. Tale spazio verrà detto **spazio vettoriale quoziente** di V rispetto a Z.

L'applicazione che associa al vettore $\mathbf{x} \in V$ la sua classe $[\mathbf{x}] \in V/Z$ è un'applicazione lineare e suriettiva che indichiamo con π.

Sia ora $\varphi\colon V \to W$ un'applicazione lineare. La φ induce un'applicazione

$$\begin{aligned} \psi\colon V/\ker\varphi &\longrightarrow \operatorname{Im}\varphi \\ [\mathbf{x}] &\longmapsto \varphi(\mathbf{x}) \end{aligned}$$

Tale applicazione è ben definita, cioè non dipende dalla scelta del rappresentante $\mathbf{x}$ della classe $[\mathbf{x}]$; è pure lineare e suriettiva, come si verifica immediatamente. La ψ è inoltre iniettiva perché $\psi([\mathbf{x}]) = \mathbf{o}_W$ se $\varphi(\mathbf{x}) = \mathbf{o}_W$, cioè se $\mathbf{x} \in \ker\varphi$, ovvero $[\mathbf{x}] = \mathbf{o}_{V/\ker\varphi}$. L'applicazione ψ è dunque un isomorfismo ed abbiamo il seguente diagramma di applicazioni lineari

$$\begin{tikzcd} V \arrow[d, "\pi"] \arrow[dr, "\varphi"] & \\ V/\ker\varphi \arrow[r, "\psi"'] & \operatorname{Im}\varphi \subseteq W \end{tikzcd}$$

Tale diagramma è **commutativo**, cioè $\varphi = \psi \circ \pi$.

Se Z è un sottospazio di V e $\pi\colon V \to V/Z$ è l'applicazione lineare suriettiva definita da $\pi(\mathbf{x}) = [\mathbf{x}]$, allora $\ker \pi = Z$ ed $\operatorname{Im} \pi = V/Z$. Quindi per il teorema 2.13

$$\dim V/Z = \dim V - \dim Z \tag{2.3}$$

2.7 Duale

Sia V/K uno spazio vettoriale. Indichiamo con V^* l'insieme $\operatorname{Hom}_K(V, K)$ delle applicazioni lineari di V in K; gli elementi di V^* sono detti anche **forme lineari**. Tale insieme è uno spazio vettoriale, detto **spazio vettoriale duale**, con le operazioni di somma e prodotto per uno scalare definite nel modo seguente: dati $\varphi, \psi \in V^*$, $\mathbf{x} \in V$ ed $\alpha \in K$:

$$(\varphi + \psi)(\mathbf{x}) = \varphi(\mathbf{x}) + \psi(\mathbf{x}), \qquad (\alpha\varphi)(\mathbf{x}) = \alpha\varphi(\mathbf{x})$$

Se $\mathsf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ è una base di V, allora una forma lineare $\varphi \in V^*$ è completamente determinata dal suo valore sugli elementi di E: infatti se $\mathbf{x} = \sum_i x_i \mathbf{e}_i$, allora $\varphi(\mathbf{x}) = \sum_i x_i \varphi(\mathbf{e}_i)$.

Sia $\mathbf{e}_i^*$ la forma lineare definita da $\mathbf{e}_i^*(\mathbf{e}_j) = \delta_{ij}$, con $1 \leqslant i, j \leqslant n$; quindi $\mathbf{e}_i^*\left(\sum_k x_k \mathbf{e}_k\right) = x_i$. Denotiamo $\mathsf{E}^* = \{\mathbf{e}_1^*, \dots, \mathbf{e}_n^*\}$ e verifichiamo che E^* è una base di V^*: data $\varphi \in V^*$, sia $\varphi(\mathbf{e}_i) = \alpha_i$, allora $\varphi = \sum_i \alpha_i \mathbf{e}_i^*$; infatti queste due forme lineari assumono lo stesso valore sugli elementi della base E. Inoltre, se $\sum \beta_i \mathbf{e}_i^* = \mathbf{0}_{V^*}$, allora, $0 = \left(\sum \beta_i \mathbf{e}_i^*\right)(\mathbf{e}_j) = \beta_j$, per ogni $j = 1, \dots, n$; quindi gli $\mathbf{e}_1^*, \dots, \mathbf{e}_n^*$ sono linearmente indipendenti. La base E^* è detta **base duale** di E.

3 Matrici

3.1 Definizioni

Una **matrice** $m \times n$ a coefficienti in K è una tabella A formata da $m \times n$ elementi di K

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

disposti su m righe ed n colonne. L'elemento a_{ij} che si trova sulla i-esima riga e la j-esima colonna sarà detto **elemento di posto** (i, j) della matrice A. Scriviamo anche $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$, o più semplicemente $A = (a_{ij})$.

L'insieme delle matrici $m \times n$ a coefficienti in K è denotato $\mathrm{M}_{m,n}(K)$ e possiede una struttura naturale di spazio vettoriale definita nel modo seguente: il vettore nullo è la matrice O i cui elementi sono tutti nulli date $A = (a_{ij}), B = (b_{ij}) \in \mathrm{M}_{m,n}(K)$, e $\alpha \in K$, siano

$$\begin{aligned} A + B &\text{ la matrice il cui elemento di posto } (i, j) \text{ è } a_{ij} + b_{ij} \text{ per ogni } i = 1, \dots, m;\ j = 1, \dots, n \\ \alpha A &\text{ la matrice il cui elemento di posto } (i, j) \text{ è } \alpha a_{ij} \text{ per ogni } i = 1, \dots, m;\ j = 1, \dots, n \end{aligned}$$

Sia E_{ij} la matrice avente l'elemento di posto (i, j) uguale ad 1 e tutti gli altri uguali a 0. Le matrici E_{ij}, con $i = 1, \dots, m$ e $j = 1, \dots, n$, formano una base di $\mathrm{M}_{m,n}(K)$: questo è quindi uno spazio vettoriale di dimensione mn. Infatti, data $A = (a_{ij})$, possiamo scrivere in modo unico $A = \sum_{i,j} a_{ij}E_{ij}$.

Prodotto

Siano

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in \mathrm{M}_{m,n}(K), \qquad B = (b_{jl})_{\substack{j=1,\dots,n \\ l=1,\dots,p}} \in \mathrm{M}_{n,p}(K),$$

Il **prodotto** AB è la matrice $C \in \mathrm{M}_{m,p}(K)$ il cui elemento di posto (i, l) è

$$c_{il} = \sum_{j=1}^{n} a_{ij}b_{jl}$$

Tale prodotto è detto **prodotto righe per colonne**.

Se $m = n$, indichiamo con $\mathrm{M}_n(K)$ lo spazio vettoriale $\mathrm{M}_{n,n}(K)$ delle **matrici quadrate** $n \times n$.

Sia $I_n \in \mathrm{M}_n(K)$ la matrice $n \times n$ tale che il suo elemento di posto (i, j) è δ_{ij}; cioè avente tutti gli elementi al di fuori della diagonale uguali a 0, e uguali ad 1 quelli sulla diagonale; chiamiamo I_n la **matrice identica** $n \times n$.

Data una matrice $A \in \mathrm{M}_{m,n}(K)$, è immediato verificare che $AI_n = I_m A = A$.

Se $A \in \mathrm{M}_{m,n}(K)$, la **trasposta** ${}^tA$ di A è la matrice $B = (b_{ji}) \in M_{n,m}(K)$ dove $b_{ji} = a_{ij}$.

3.2 Matrice associata ad un'applicazione lineare

Siano V e W due spazi vettoriali su K di dimensione rispettivamente n ed m, e $\varphi\colon V \to W$ un'applicazione lineare. Fissiamo una base $\mathsf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di V ed una base $\mathsf{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ di W. In questa situazione possiamo associare a φ una matrice $A \in \mathrm{M}_{m,n}(K)$, detta **matrice rappresentativa** di φ (o **associata** a φ) nelle basi E, F, denotata anche $\mathrm{M}_{\mathsf{E}}^{\mathsf{F}}(\varphi)$. La matrice A è costituita dalle colonne $C_1, \dots, C_n$ definite nel modo seguente: la colonna C_j ha come elementi le componenti del vettore $\varphi(\mathbf{e}_j)$ espresse nella base F. Cioè, se $A = (a_{ij})$ allora a_{ij} è la i-esima componente di $\varphi(\mathbf{e}_j)$ nella base F.

Se identifichiamo lo spazio W con K^m attraverso l'isomorfismo i_{F}, allora C_i è il vettore di K^m dato da $i_{\mathsf{F}} \circ \varphi(\mathbf{e}_i)$. Quindi $\mathcal{L}(C_1, \dots, C_n) = \mathrm{Im}(i_{\mathsf{F}} \circ \varphi) \subseteq K^m$.

Essendo i_{F} un isomorfismo in cui $\mathrm{Im}\,\varphi$ e $\mathrm{Im}(i_{\mathsf{F}} \circ \varphi)$ si corrispondono, avremo che

$$\dim \mathrm{Im}\,\varphi = \dim \mathrm{Im}(i_{\mathsf{F}} \circ \varphi) = \dim \mathcal{L}(C_1, \dots, C_n)$$

Tale dimensione è quindi uguale al numero massimo di colonne indipendenti di A.

Si noti che, data φ, la matrice A che la rappresenta *dipende* dalla scelta delle basi E ed F: studieremo in seguito il modo in cui essa varia al variare delle basi.

Con le notazioni di sopra, sia $\mathbf{v} = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n \in V$ e $\mathbf{w} = \varphi(\mathbf{v}) = y_1\mathbf{f}_1 + \dots + y_m\mathbf{f}_m$. Avremo allora

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = y_1 \tag{3.1}$$
$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = y_2 \tag{3.2}$$
$$\vdots \tag{3.3}$$
$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = y_m \tag{3.4}$$

Se indichiamo con X la matrice $n \times 1$ ${}^t(x_1, \dots, x_n)$ e con Y la matrice ${}^t(y_1, \dots, y_m)$, allora (3.1) può essere scritto in modo compatto nella forma

$$AX = Y$$

dove AX è da intendersi come prodotto di matrici.

Proposizione 3.1

Siano V, W, T spazi vettoriali su K di dimensione rispettiva n, m, p. Siano fissate delle basi

$\mathsf{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, $\mathsf{F} = \{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$, $\mathsf{G} = \{\mathbf{g}_1, \ldots, \mathbf{g}_p\}$ *rispettivamente di* V, W, T. *Siano infine* $\varphi: V \to W$ *e* $\psi: W \to T$ *applicazioni lineari. Allora, se* $A = M_{\mathsf{E}}^{\mathsf{F}}(\varphi)$ *e* $B = M_{\mathsf{F}}^{\mathsf{G}}(\psi)$, *si ha*

$$BA = M_{\mathsf{E}}^{\mathsf{G}}(\psi \circ \varphi)$$

cioè

$$M_{\mathsf{E}}^{\mathsf{G}}(\psi \circ \varphi) = M_{\mathsf{F}}^{\mathsf{G}}(\psi) M_{\mathsf{E}}^{\mathsf{F}}(\varphi)$$

Dimostrazione. Siano $\mathbf{v} \in V$, $\mathbf{w} = \varphi(\mathbf{v}) \in W$, $\mathbf{t} = \psi(\mathbf{w}) \in T$. Se

$$\mathbf{v} = \sum x_i \mathbf{e}_i, \quad \mathbf{w} = \sum y_j \mathbf{f}_j, \quad \mathbf{t} = \sum z_k \mathbf{g}_k$$

posto

$$X = {}^t(x_1, \ldots, x_n), \quad Y = {}^t(y_1, \ldots, y_m), \quad Z = {}^t(z_1, \ldots, z_p)$$

avremo

$$Y = AX, \qquad Z = BY \quad \Longrightarrow \quad Z = BAX$$

Segue che BA è la matrice rappresentatica di $\psi \circ \varphi$ rispetto alle basi E e G. □

3.3 Sistemi di equazioni lineari

Un'**equazione lineare** in n variabili su K è un'equazione della forma

$$a_1 x_1 + \ldots a_n x_n = b$$

dove $a_1, \ldots, a_n, b \in K$. Una soluzione è una n-upla $(\alpha_1, \ldots, \alpha_n)$ con $\alpha_1, \ldots, \alpha_n \in K$ tale che $a_1\alpha_1 + \ldots a_n\alpha_n = b$. Gli a_i sono detti **coefficienti** dell'equazione lineare, mentre b è detto **termine noto**.

Un **sistema** di m equazioni lineari in n incognite è un insieme di m equazioni lineari in n incognite

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{cases} \tag{3.5}$$

Una soluzione del sistema è una n-upla $(\alpha_1, \ldots, \alpha_n)$ soluzione di ciascuna delle m equazioni. Il sistema è detto **risolubile** se esiste almeno una soluzione $(\alpha_1, \ldots, \alpha_n)$.

La matrice $A = (a_{ij}) \in M_{m,n}(K)$ è detta **matrice dei coefficienti** del sistema.

Se indichiamo con X la matrice $n \times 1$ ${}^t(x_1, \ldots, x_n)$ e con B la matrice ${}^t(b_1, \ldots, b_m)$, il sistema (3.5) può essere scritto in modo compatto nella forma

$$AX = B$$

dove AX è da intendersi come prodotto di matrici.

Se $C_1, \ldots, C_n$ sono le colonne di A, allora il sistema (3.5) può essere letto in K^m come

$$x_1C_1 + x_2C_2 + \ldots x_nC_n = B \tag{3.6}$$

Infine, se $R_1, \ldots, R_m$ sono le righe di A, il sistema (3.5) può essere scritto sotto la forma

$$R_1X = 0, \ldots, R_mX = 0 \tag{3.7}$$

(notiamo che R_iX è un numero ottenuto dal prodotto di una matrice $1 \times n$ e di una $n \times 1$)

Un sistema $AX = O$ è detto **omogeneo**: tale sistema ammette sempre la soluzione $(0, \ldots, 0)$, detta soluzione banale. Due sistemi omogeni $AX = O$ e $A'X = O$ nelle stesse incognite si dicono **equivalenti** se hanno le stesse soluzioni.

Per esempio, se A è sistema $m \times n$ (3.5), ed R_i è una riga combinazione lineare delle altre, allora il sistema omogeneo $AX = O$ è equivalente al sistema $A'X = O$ in cui A' è la matrice ottenuta da A cancellando la riga R_i. Infatti il sistema $AX = O$ può essere scritto sotto la forma (3.7); quindi, supponendo che $R_i = \sum_{j \neq i} \alpha_j R_j$, abbiamo che: se $R_jX = 0$ per ogni $j \neq i$, allora $R_iX = \sum_{j \neq i} \alpha_j R_j X = 0$; cioè se X è soluzione di $A'X = O$, allora X è soluzione di $AX = O$. Il viceversa è ovvio.

Osserviamo che, se $AX = O$ è un sistema $m \times n$ ed $A'X = O$ è un sistema $p \times n$, e se $\varphi\colon K^n \to K^m$ è l'applicazione lineare associata ad A rispetto alle basi standard, analogamente per $\varphi' : K^n \to K^p$, i due sistemi sono equivalenti se e solo se $\ker \varphi = \ker \varphi'$.

3.4 Rango

Sia A una matrice $m \times n$; sia c il numero massimo di colonne di A linearmente indipendenti ed r il numero massimo di righe linearmente indipendenti. Sia $\varphi\colon K^n \to K^m$ l'applicazione lineare espressa da A nelle basi canoniche dei versori. Allora $c = \dim \operatorname{Im} \varphi$ mentre per il abbiamo:

$$n = c + \dim \ker \varphi \tag{3.8}$$

Scegliamo ora r righe di A linearmente indipendenti: possiamo supporre senza perdita di generalità che esse siano le prime r. Ogni altra riga di A è pertanto combinazione lineare delle prime r. Sia A' la matrice $r \times n$ sottomatrice di A formata dalle prime r righe di A; sia φ' l'applicazione lineare $K^n \to K^r$ espressa da A' nelle basi canoniche dei versori. Posto $c' = \dim \operatorname{Im} \varphi'$, avremo come sopra:

$$n = c' + \dim \ker \varphi' \tag{3.9}$$

ed inoltre $c' \leqslant r$. I sistemi lineari $AX = O$ ed $A'X = O$ sono equivalenti perché tutte le righe di A si esprimono come combinazione lineare delle prime r. Ciò significa che $\ker \varphi = \ker \varphi'$. Dalle (3.8) e (3.9) otteniamo $c = c'$ e quindi $c \leqslant r$.

Ragionando come sopra con la matrice ${}^tA$ si ottiene $r \leqslant c$ e quindi $r = c$.

Abbiamo dimostrato

Proposizione-definizione 3.2

Sia A una matrice $m \times n$, allora il numero massimo di righe linearmente indipendenti è uguale al numero massimo di colonne linearmente indipendenti. Tale numero è detto **rango** *della matrice A ed è denotato $\rho(A)$. Segue che $\rho(A) \leqslant \min\{m, n\}$.*

Proposizione 3.3

Siano V uno spazio vettoriale di dimensione n e W di dimensione m e sia $\varphi: V \to W$ un'applicazione lineare. Fissiamo due basi E di V ed F di W, e sia A la matrice rappresentativa di φ in queste basi. Allora

1. *φ è suriettiva se e solo se $\rho(A) = m$*
2. $\dim\ker\varphi = n - \rho(A)$, *in particolare φ è iniettiva se e solo se $\rho(A) = n$*

Dimostrazione. Sappiamo che $\rho(A)$ è esattamente la dimensione di $\mathrm{Im}(\varphi)$: quindi φ è suriettiva se e solo se $\rho(A) = m$.

Per il , $\dim\ker\varphi = n - \dim\mathrm{Im}(f) = n - \rho(A)$; d'altra parte, φ è iniettiva se e solo se $\ker\varphi = \{\mathbf{0}_V\}$, da cui segue che $\ker\varphi = \{\mathbf{0}_V\}$ se e solo se $n = \rho$. □

Corollario 3.4

Sia A una matrice quadrata $n \times n$; allora A è invertibile se e solo se $\rho(A) = n$.

Dimostrazione. Consideriamo A come matrice rappresentativa di un endomorfismo φ di K^n rispetto alla base standard dei versori. L'applicazione φ è invertibile se e solo se la matrice A lo è; sappiamo che φ è invertibile se e solo se è suriettiva. Quindi A è invertibile se e solo se $\rho(A) = n$. □

Proposizione 3.5 (Rouché-Capelli)

Sia $AX = B$ un sistema lineare di m equazioni in n incognite. Indichiamo con $(A|B)$ la matrice $m \times (n+1)$ ottenuta aggiungendo alle colonne di A la colonna di B.

Allora il sistema $AX = B$ è risolubile se e solo se $\rho(A) = \rho\big((A|B)\big)$.

Dimostrazione. Se $C_1, \dots, C_n$ sono le colonne di A, allora il sistema (3.5) può essere letto in K^m come $x_1C_1 + x_2C_2 + \dots x_nC_n = B$. Quindi il sistema $AX = B$ è risolubile se e solo se il vettore colonna B è combinazione lineare dei vettori $C_1, \dots, C_n$; in altre parole lo spazio vettoriale $\mathcal{L}(C_1, \dots, C_n)$ è uguale allo spazio $\mathcal{L}(C_1, \dots, C_n, B)$, cioè $\rho(A) = \rho\big((A|B)\big)$. □

Definizione. La matrice $(A|B)$ è detta **matrice completa** del sistema $AX = B$.

Osservazione 3.6. Nel caso dei sistemi omogenei $AX = O$, il criterio precedente non dà alcuna informazione; infatti la colonna $B = O$ ed il criterio è banalmente verificato. D'altra parte sappiamo che un tale sistema ammette sempre la soluzione banale.

3.5 Cambiamenti di base

Indichiamo con $\mathrm{GL}_n(K)$ l'insieme delle matrici $n \times n$ invertibili a coefficienti in K, detto **gruppo lineare generale** (di ordine n) su K. In esso l'elemento neutro è costituito dalla matrice I_n. Identifichiamo $\mathrm{GL}_1(K)$ con K^*, cioè col gruppo degli elementi di K diversi da 0: in particolare è un gruppo commutativo. Se $n \geqslant 2$, allora $\mathrm{GL}_n(K)$ non è commutativo; infatti le matrici

$$A = \left(\begin{array}{cc|ccc} 1 & 1 & & O & \\ 0 & 1 & & & \\ \hline & & 1 & & \\ & O & & \ddots & \\ & & & & 1 \end{array}\right), \qquad B = \left(\begin{array}{cc|ccc} 1 & 1 & & O & \\ 1 & 0 & & & \\ \hline & & 1 & & \\ & O & & \ddots & \\ & & & & 1 \end{array}\right)$$

non commutano.

Definizione. Sia V uno spazio vettoriale su K, e siano $\mathsf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, $\mathsf{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ due basi di V. La **matrice di passaggio** da E ad F è la matrice $P = (p_{ij})$, $n \times n$, la cui colonna j-esima è formata dalle componenti di $\mathbf{f}_j$ nella base E. Cioè

$$\begin{aligned} \mathbf{f}_1 &= p_{11}\mathbf{e}_1 + p_{21}\mathbf{e}_2 + \cdots + p_{n1}\mathbf{e}_n \\ \mathbf{f}_2 &= p_{12}\mathbf{e}_1 + p_{22}\mathbf{e}_2 + \cdots + p_{n2}\mathbf{e}_n \\ \vdots\ &\ \vdots \\ \mathbf{f}_n &= p_{1n}\mathbf{e}_1 + p_{2n}\mathbf{e}_2 + \cdots + p_{nn}\mathbf{e}_n \end{aligned}$$

Interpretazioni della matrice di passaggio. La matrice $P = (p_{ij})$ di passaggio da E ad F ha due interpretazioni:

a) Sia $\varphi: V \to V$ definita da $\varphi(\mathbf{e}_i) = \mathbf{f}_i$. Allora $P = M^{\mathsf{E}}_{\mathsf{E}}(\varphi)$.
b) La matrice P è anche la matrice rappresentativa dell'applicazione identica $\mathrm{id}_V : V \to V$ nella coppia di basi F, E: $P = M^{\mathsf{E}}_{\mathsf{F}}(\mathrm{id}_V)$.

Proposizione 3.7

Con le notazioni di sopra, sia P la matrice di passaggio da E ad F; allora P è invertibile e P^{-1} è la matrice di passaggio da F ad E.

Dimostrazione. Sia $Q = M^{\mathsf{F}}_{\mathsf{E}}(\mathrm{id}_V)$, allora

$$QP = M^{\mathsf{F}}_{\mathsf{F}}(\mathrm{id}_V) = I_n$$

Quindi P è invertibile e $Q = P^{-1}$. □

Proposizione 3.8

Sia V uno spazio vettoriale su K ed $\mathsf{E} = \{\mathbf{e}_1, \dots \mathbf{e}_n\}$ una base di V. Sia $P = (p_{ij}) \in \mathrm{GL}_n(K)$.

Poniamo

$$\begin{aligned}
\mathbf{f}_1 &= p_{11}\mathbf{e}_1 + p_{21}\mathbf{e}_2 + \cdots + p_{n1}\mathbf{e}_n \\
\mathbf{f}_2 &= p_{12}\mathbf{e}_1 + p_{22}\mathbf{e}_2 + \cdots + p_{n2}\mathbf{e}_n \\
\vdots\ &\vdots \\
\mathbf{f}_n &= p_{1n}\mathbf{e}_1 + p_{2n}\mathbf{e}_2 + \cdots + p_{nn}\mathbf{e}_n
\end{aligned}$$

Allora $\mathsf{F} = \{\mathbf{f}_1, \dots \mathbf{f}_n\}$ *è una base di V e P è la matrice di passaggio da* E *ad* F.

Dimostrazione. Sia φ l'endomorfismo di V definito da $\varphi(\mathbf{e}_i) = \mathbf{f}_i$. Allora $P = M^{\mathsf{E}}_{\mathsf{E}}(\varphi)$. Essendo P invertibile, le sue colonne sono linearmente indipendenti, quindi i vettori $\mathbf{f}_1, \dots \mathbf{f}_n$ costituiscono una base. □

Proposizione 3.9

Con le notazioni della proposizione 3.7, sia $\mathbf{v}$ *un vettore di V. Sia* $\mathbf{v} = \sum x_i\mathbf{f}_i = \sum y_j\mathbf{e}_j$. *Se* $X = {}^t(x_1, \dots, x_n)$ *e* $Y = {}^t(y_1, \dots, y_n)$, *allora* $Y = PX$.

Dimostrazione. Basta ricordare che P è la matrice rappresentativa dell'applicazione identica nella coppia di basi F, E. □

Definizione. Siano A, B due matrici $m \times n$. Diciamo che A e B sono **equivalenti** se esistono $R \in \mathrm{GL}_m(K)$ e $P \in \mathrm{GL}_n(K)$ tali che

$$B = RAP$$

Proposizione 3.10

Siano V uno spazio vettoriale di dimensione n e W uno spazio vettoriale di dimensione m; siano poi E *una base di V e* F *una base di W. Sia φ un'applicazione lineare da V a W ed* $A = M^{\mathsf{F}}_{\mathsf{E}}(\varphi)$ *la sua matrice rappresentativa nelle basi* E *ed* F. *Allora*

1. *Se* E' *è una base di V ed* F' *una base di W, allora* $B = M^{\mathsf{F}'}_{\mathsf{E}'}(\varphi)$ *è equivalente ad A.*
2. *Viceversa, se B è una matrice equivalente ad A, allora B rappresenta φ in una coppia opportuna di basi* E' *di V e* F' *di W.*

Dimostrazione.

1. Siano $P = M^{\mathsf{E}}_{\mathsf{E}'}(\mathrm{id}_V)$, la matrice di cambiamento di base da E a E'; e $Q = M^{\mathsf{F}}_{\mathsf{F}'}(\mathrm{id}_W)$, la matrice di cambiamento di base da F a F'. Allora $\varphi = \mathrm{id}_W^{-1} \circ \varphi \circ \mathrm{id}_V$, quindi

$$M^{\mathsf{F}'}_{\mathsf{E}'}(\varphi) = M^{\mathsf{F}'}_{\mathsf{F}}(\mathrm{id}_W^{-1}) \cdot M^{\mathsf{F}}_{\mathsf{E}}(\varphi) \cdot M^{\mathsf{E}}_{\mathsf{E}'}(\mathrm{id}_V)$$

Cioè

$$B = Q^{-1}AP$$

.

2. Se $B = RAP$, poniamo $Q = R^{-1}$. Siano

E'	la base di V tale che	$P = M^{\mathsf{E}}_{\mathsf{E}'}(\mathrm{id}_V)$
F'	la base di W tale che	$Q = M^{\mathsf{F}}_{\mathsf{F}'}(\mathrm{id}_V)$

Allora

$$M^{\mathsf{F}'}_{\mathsf{E}'}(\varphi) = Q^{-1}AP = RAP = B$$

□

Corollario 3.11

Se B è equivalente ad A, allora $\rho(A) = \rho(B)$.

Dimostrazione. Data A, sia $\varphi\colon K^n \to K^m$ l'applicazione lineare rappresentata da A nelle basi standard dei versori di K^n e K^m.

Per la proposizione precedente, B rappresenta φ in una opportuna coppia di basi di K^n e K^m.

Ne segue che $\rho(A) = \dim \operatorname{Im} \varphi = \rho(B)$. □

Definizione. Siano A e B matrici quadrate $n \times n$. Diciamo che B è **simile** ad A se esiste $P \in \mathrm{GL}_n(K)$ tale che $B = P^{-1}AP$.

Osservazione 3.12. La relazione di similitudine è una relazione di equivalenza, come si verifica facilmente.

Matrici simili sono, in particolare, equivalenti; quindi hanno lo stesso rango.

4 Determinante

4.1 Introduzione

Partiamo da un esempio. Consideriamo $\mathbf{R}^2$ come spazio vettoriale su $\mathbf{R}$ dotato della base canonica $\mathbf{e}_1 = (1, 0)$, $\mathbf{e}_2 = (0, 1)$. Definiamo l'applicazione $\Phi \colon \mathbf{R}^2 \to R$ nel modo seguente. Dati

$$\begin{aligned} \mathbf{v}_1 &= a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2 \\ \mathbf{v}_2 &= a_{21}\mathbf{e}_1 + a_{22}\mathbf{e}_2 \end{aligned}$$

allora

$$\Phi(\mathbf{v}_1, \mathbf{v}_2) = a_{11}a_{22} - a_{12}a_{21}$$

Il valora assoluto $|\Phi(\mathbf{v}_1, \mathbf{v}_2)|$ è l'area del parallelogramma definito dai vettori $\mathbf{v}_1$ e $\mathbf{v}_2$.

L'applicazione Φ ha le seguenti caratteristiche, di immediata verifica.

1. Φ è **bilineare**, cioè è lineare in ciascuna delle variabili indipendentemente. Ciò significa che per ogni $\alpha, \beta \in \mathbf{R}$ e per ogni $\mathbf{v}_1, \mathbf{v}_1', \mathbf{v}_2, \mathbf{v}_2' \in \mathbf{R}^2$,
$$\begin{aligned} \Phi(\alpha\mathbf{v}_1 + \beta\mathbf{v}_1', \mathbf{v}_2) &= \alpha\Phi(\mathbf{v}_1, \mathbf{v}_2) + \beta\Phi(\mathbf{v}_1', \mathbf{v}_2) \\ \Phi(\mathbf{v}_1, \alpha\mathbf{v}_2 + \beta\mathbf{v}_2') &= \alpha\Phi(\mathbf{v}_1, \mathbf{v}_2) + \beta\Phi(\mathbf{v}_1, \mathbf{v}_2') \end{aligned}$$
2. Φ è **alternante**, cioè
$$\Phi(\mathbf{v}_1, \mathbf{v}_2) = -\Phi(\mathbf{v}_2, \mathbf{v}_1)$$
(in particolare $\Phi(\mathbf{v}, \mathbf{v}) = 0$)
3. $\Phi(\mathbf{e}_1, \mathbf{e}_2) = 1$.

Proposizione 4.1
$\Phi(\mathbf{v}_1, \mathbf{v}_2) = 0$ *se e solo se* $\mathbf{v}_1$ *e* $\mathbf{v}_2$ *sono linearmente dipendenti.*

Dimostrazione. L'affermazione è ovvia per il significato geometrico del $|\Phi|$, e possiamo darne una dimostrazione diretta: se $\mathbf{v}_2 = \lambda\mathbf{v}_1$ con $\lambda \in \mathbf{R}$, allora

$$\Phi(\mathbf{v}_1, \mathbf{v}_2) = \Phi(\mathbf{v}_1, \lambda\mathbf{v}_1) = \lambda\Phi(\mathbf{v}_1, \mathbf{v}_1) = 0$$

Viceversa, sia $\Phi(\mathbf{v}_1, \mathbf{v}_2) = 0$. Se i due vettori non sono entrambi nulli, possiamo supporre $\mathbf{v}_1 \neq \mathbf{0}$ e quindi, ad esempio, $a_{11} \neq 0$. Poniamo $\lambda = a_{21}/a_{11}$, allora

$$0 = \Phi(\mathbf{v}_1, \mathbf{v}_2) = a_{11}a_{22} - \lambda a_{12}a_{11} = a_{11}(a_{22} - \lambda a_{12})$$

Quindi $a_{22} = \lambda a_{12}$ e $\mathbf{v}_2 = \lambda\mathbf{v}_1$. □

Generalizziamo la definizione di Φ a quella di forma multilineare alternante.

4.2 Forme multilineari

Siano $V_1, \ldots, V_r, F$ spazi vettoriali su K. Un'applicazione $\Phi\colon V_1 \times \cdots \times V_r \to F$ è detta **multilineare** se è lineare in ciascuna variabile separatamente; cioè se per ogni $\alpha, \beta \in K$, per ogni $\mathbf{v}_1, \mathbf{v}_1' \in V_1, \ldots, \mathbf{v}_r, \mathbf{v}_r' \in V_r$ e per ogni $i = 1, \ldots, r$:

$$\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \alpha\mathbf{v}_i + \beta\mathbf{v}_i', \mathbf{v}_{i+1}, \ldots, \mathbf{v}_r) = \alpha\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_r) + \beta\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_i', \ldots, \mathbf{v}_r)$$

L'insieme di tali applicazione con le operazioni usuali di somma e moltiplicazione per uno scalare forma uno spazio vettoriale su K che indichiamo con $L_r(V_1, \ldots, V_r; F)$.

Se $X = (\mathbf{v}_1, \ldots, \mathbf{v}_r)$ è un elemento di $V_1 \times \cdots \times V_r$, scriveremo $\Phi(X)$ anziché $\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r)$

Nel caso in cui $V_1 = \cdots = V_r = V$, scriveremo semplicemente $L_r(V; F)$. Se inoltre $F = K$ scriveremo $L_r(V)$: gli elementi di quest'ultimo spazio sono detti **forme** r**-multilineari** su V.

Una forma $\Phi \in L_r(V)$ è detta **simmetrica** se, per ogni σ appartenente al gruppo delle sostituzioni S_r si ha $\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r) = \Phi(\mathbf{v}_{\sigma(1)}, \ldots, \mathbf{v}_{\sigma(r)})$. L'insieme delle forme r-multilineari simmetriche è un sottospazio vettoriale di $L_r(V)$ denotato $L_r^s(V)$.

Una forma $\Phi \in L_r(V)$ è detta **alternante** se, per ogni σ appartenente al gruppo delle sostituzioni S_r si ha $\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r) = \varepsilon(\sigma)\Phi(\mathbf{v}_{\sigma(1)}, \ldots, \mathbf{v}_{\sigma(r)})$. L'insieme delle forme r-multilineari alternanti è un sottospazio vettoriale di $L_r(V)$ denotato $L_r^a(V)$.

Chiamiamo **forma nulla** la forma Φ che assume sempre il valore 0.

Proposizione 4.2

Sia $\Phi \in L_r(V)$. Allora $\Phi \in L_r^a(V)$ se e solo se $\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r) = 0$ ogni volta che esistono $i \neq j$ tali che $\mathbf{v}_i = \mathbf{v}_j$.

Dimostrazione. Supponiamo che Φ sia alternante. Se esistono $i \neq j$ con $\mathbf{v}_i = \mathbf{v}_j$, allora sia τ la trasposizione $(i, j) \in S_r$. Quindi $\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r) = \Phi(\mathbf{v}_{\tau(1)}, \ldots, \mathbf{v}_{\tau(r)}) = \varepsilon(\tau)\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r) = -\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r)$; perciò $\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r) = 0$.

Reciprocamente, supponiamo dapprima che $r = 2$. Allora l'unica permutazione di S_2 diversa dall'identità è la trasposizione $\tau = (1, 2)$. Abbiamo $0 = \Phi(\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2) = \Phi(\mathbf{v}_1, \mathbf{v}_1) + \Phi(\mathbf{v}_1, \mathbf{v}_2) + \Phi(\mathbf{v}_2, \mathbf{v}_1) + \Phi(\mathbf{v}_2, \mathbf{v}_2) = \Phi(\mathbf{v}_1, \mathbf{v}_2) + \Phi(\mathbf{v}_2, \mathbf{v}_1)$ Dunque $\Phi(\mathbf{v}_{\tau(1)}, \mathbf{v}_{\tau(2)}) = \Phi(\mathbf{v}_2, \mathbf{v}_1) = -\Phi(\mathbf{v}_1, \mathbf{v}_2) = \varepsilon(\tau)\Phi(\mathbf{v}_1, \mathbf{v}_2)$. Se $r > 2$, consideriamo una trasposizione $\tau = (i, j)$ con $i < j$. Se $(\mathbf{v}_1, \ldots, \mathbf{v}_r) \in V^r$, poniamo

$$\Psi(\mathbf{t}, \mathbf{w}) = \Phi(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{t}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_{j-1}, \mathbf{w}, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_r)$$

Segue che Ψ è una forma bilineare su V con $\Psi(\mathbf{v}, \mathbf{v}) = 0$ per ogni $v \in V$. Per quanto visto nel caso $r = 2$, avremo $\Psi(\mathbf{v}_j, \mathbf{v}_i) = -\Psi(\mathbf{v}_i, \mathbf{v}_j)$, cioè

$$\Phi(\mathbf{v}_{\tau(1)}, \ldots, \mathbf{v}_{\tau(r)}) = \varepsilon(\tau)\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r)$$

Consideriamo ora un $\sigma \in S_r$ qualsiasi: sappiamo che $\sigma = \tau_1 \cdots \tau_k$ dove i τ_i sono trasposizioni. Poniamo $\chi = \tau_1 \cdots \tau_{k-1}$.

Dimostriamo che $\Phi(\mathbf{v}_{\sigma(1)}, \ldots, \mathbf{v}_{\sigma(r)}) = \varepsilon(\sigma)\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r)$ per induzione su k. Per $k = 1$ la formula è stata già dimostrata. Supponiamo che sia vera per $k-1$, quindi che $\Phi(\mathbf{v}_{\chi(1)}, \ldots, \mathbf{v}_{\chi(r)}) = \varepsilon(\chi)\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r)$. Abbiamo

$$\begin{aligned}\Phi(\mathbf{v}_{\sigma(1)}, \ldots, \mathbf{v}_{\sigma(r)}) &= \Phi(\mathbf{v}_{\chi\tau_k(1)}, \ldots, \mathbf{v}_{\chi\tau_k(r)}) = \varepsilon(\chi)\Phi(\mathbf{v}_{\tau_k(1)}, \ldots, \mathbf{v}_{\tau_k(r)}) \\ &= \varepsilon(\chi)\varepsilon(\tau_k)\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r) = \varepsilon(\sigma)\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r)\end{aligned}$$

□

Proposizione 4.3

Sia $\Phi \in L_r^a(V)$. Se $\mathbf{v}_1, \ldots, \mathbf{v}_r$ sono linearmente dipendenti, allora $\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r) = 0$.

Dimostrazione. Possiamo supporre, senza perdità di generalità, che $\mathbf{v}_r$ sia combinazione lineare dei precedenti; cioè $\mathbf{v}_r = \sum_{i=1}^{r-1} \alpha_i \mathbf{v}_i$. Quindi

$$\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r) = \Phi\left(\mathbf{v}_1, \ldots, \mathbf{v}_{r-1}, \sum_{i=1}^{r-1} \alpha_i \mathbf{v}_i\right) = \sum_{i=1}^{r-1} \alpha_i \Phi\left(\mathbf{v}_1, \ldots, \mathbf{v}_{r-1}, \mathbf{v}_i\right)$$

Per la proposizione precedente ogni addendo è nullo.

□

Proposizione 4.4

Sia $\Phi \in L_r^a(V)$. Siano $\mathbf{w}_j = \sum_{i=1}^r a_{ji}\mathbf{v}_i$, con $j = 1, \ldots, r$. Allora

$$\Phi(\mathbf{w}_1, \ldots, \mathbf{w}_r) = \left(\sum_{\sigma \in S_r} \varepsilon(\sigma) a_{1,\sigma(1)} \cdots a_{r,\sigma(r)}\right) \Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r)$$

Dimostrazione. Abbiamo che

$$\Phi(\mathbf{w}_1, \ldots, \mathbf{w}_r) = \Phi\left(\sum_{i=1}^r a_{1i}\mathbf{v}_i, \ldots, \sum_{i=1}^r a_{ri}\mathbf{v}_i\right) = \sum_{i=1}^r \sum_{k_i=1}^r a_{1k_1} \cdots a_{rk_r} \Phi(\mathbf{v}_{k_1}, \ldots, \mathbf{v}_{k_r})$$

Notiamo che $\Phi(\mathbf{v}_{k_1}, \ldots, \mathbf{v}_{k_r}) = 0$ ogni volta che due indici sono uguali, per la proposizione 4.2. Nella sommatoria rimangono quindi i soli termini in cui $\{k_1, \ldots, k_r\}$ è una permutazione di $\{1, \ldots, r\}$. Se $k_1 = \sigma(1), \ldots, k_r = \sigma(r)$ con $\sigma \in S_r$, avremo

$$\begin{aligned}\Phi(\mathbf{w}_1, \ldots, \mathbf{w}_r) &= \sum_{\sigma \in S_r} a_{1,\sigma(1)} \cdots a_{r,\sigma(r)} \Phi(\mathbf{v}_{\sigma(1)}, \ldots, \mathbf{v}_{\sigma(r)}) \\ &= \left(\sum_{\sigma \in S_r} \varepsilon(\sigma) a_{1,\sigma(1)} \cdots a_{r,\sigma(r)}\right) \Phi(\mathbf{v}_1, \ldots, \mathbf{v}_r)\end{aligned}$$

□

Teorema 4.5

Sia V uno spazio vettoriale di dimensione n su K. Allora lo spazio $L_n^a(V)$ delle forme n-multilineari alternanti ha dimensione uno. Se $\mathsf{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ è una base di V, per ogni $\alpha \in K$, esiste una ed una sola $\Phi \in L_n^a(V)$ tale che $\Phi(\mathbf{e}_1, \ldots, \mathbf{e}_n) = \alpha$.

Dimostrazione. Dati $\mathbf{v}_j = \sum_{i=1}^n a_{ji}\mathbf{e}_i$, con $j = 1, \ldots, n$, poniamo

$$\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \left(\sum_{\sigma \in S_r} \varepsilon(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \right) \alpha$$

Bisogna verificare che la forma Φ così definita è multilineare e alternante. L'unicità della Φ è conseguenza della proposizione 4.4. Segue che $\dim L_n^a(V) = 1$ e che, fissata una base E di V, esiste un isomorfismo canonico $L_n^a(V) \to K$ che associa a Φ il valore $\alpha = \Phi(\mathsf{E})$. □

Corollario 4.6

Sia V uno spazio vettoriale di dimensione n su K e sia $\Phi \in L_n^a(V)$. Sia $\mathsf{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ una base di V. Allora Φ non è nulla se e solo se $\Phi(\mathsf{E}) \neq 0$.

Dimostrazione. È conseguenza immediata del teorema precedente. □

Proposizione 4.7

Sia V uno spazio vettoriale di dimensione n su K e sia $\Phi \in L_n^a(V)$. Supponiamo Φ non nulla; se $X = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, allora $\Phi(X) \neq 0$ se e solo se X è una base di V.

Dimostrazione. Se X è una base di V, $\Phi(X) \neq 0$ per il corollario precedente. D'altra parte, se X non è una base, i vettori di X sono linearmente dipendenti e quindi, per la proposizione 4.3, $\Phi(X) = 0$. □

4.3 Determinante

Definizione. Sia V uno spazio vettoriale su K di dimensione n, e sia $\mathsf{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ una base di V. Dati n vettori $\mathbf{v}_j = \sum_{i=1}^n a_{ji}\mathbf{e}_i$, definiamo il **determinante** $\det_{\mathsf{E}}$ nella base E come la forma n-multilineare alternante

$$\det\nolimits_{\mathsf{E}}(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \sum_{\sigma \in S_r} \varepsilon(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

Per quanto visto nel paragrafo precedente, $\det_{\mathsf{E}}$ è l'unica forma n-multilineare alternante tale che $\det_{\mathsf{E}}(\mathsf{E}) = 1$.

Proposizione 4.8

Siano $\mathsf{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ ed $\mathsf{E}' = \{\mathbf{e}'_1, \ldots, \mathbf{e}'_n\}$ due basi di V. Allora: $\det_{\mathsf{E}}(\mathsf{E}') \det_{\mathsf{E}'}(\mathsf{E}) = 1$.

Definizione. Sia $\varphi\colon V \to V$ un endomorfismo ed E una base fissata come sopra. Definiamo il **determinante** di φ come lo scalare

$$\det(\varphi) = \det_{\mathsf{E}}\big(\varphi(\mathbf{e}_1), \dots, \varphi(\mathbf{e}_n)\big)$$

Tale definizione è ben posta: cioè, non dipende dalla scelta della base per la seguente proposizione

Proposizione 4.9

Siano $\mathsf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ *ed* $\mathsf{E}' = \{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ *due basi di* V*. Allora, dato un endomorfismo* φ *di* V *si ha*

$$\det_{\mathsf{E}}\big(\varphi(\mathbf{e}_1), \dots, \varphi(\mathbf{e}_n)\big) = \det_{\mathsf{E}'}\big(\varphi(\mathbf{e}'_1), \dots, \varphi(\mathbf{e}'_n)\big)$$

4.4 Proprietà del determinante

Proposizione 4.10

Sia A *una matrice quadrata* $n \times n$*. Allora* $\det A = \det {}^t\!A$.

Osservazione 4.11. Il determinante di una matrice A è quindi una forma multilineare alternata nelle colonne di A.

Proposizione 4.12

Sia M *una matrice quadrata* $n \times n$ *della forma*

$$M = \left(\begin{array}{c|c} A & C \\ \hline O & B \end{array}\right), \quad \textit{dove } A \in \mathrm{M}_p(K),\ B \in \mathrm{M}_q(K),\ n = p + q$$

Allora $\det M = \det A \det B$.

Dimostrazione. Sia C fissata, al variare di X in $\mathrm{M}_p(K)$ ed Y in $\mathrm{M}_q(K)$. Poniamo

$$M(X, Y) = \left(\begin{array}{c|c} X & C \\ \hline O & Y \end{array}\right)$$

In particolare $M = M(A, B)$. Definiamo

$$\begin{aligned} \theta\colon \mathrm{M}_p(K) \times \mathrm{M}_q(K) &\longrightarrow K \\ (X, Y) &\longmapsto \det M(X, Y) \end{aligned}$$

La funzione θ è p-multilineare alternante nelle colonne di X, dunque

$$\theta(X, Y) = \det(X)\varphi(Y), \quad \text{dove } \varphi \text{ è una funzione che dipende solo da } Y$$

Se poniamo $X = I_p$ avremo

$$\theta(I_p, Y) = \det \left(\begin{array}{c|c} I_p & C \\ \hline O & Y \end{array}\right)$$

cioè $\varphi(Y) = \theta(I_p, Y)$ per ogni Y. La funzione $\theta(I_p, Y)$ è q-multilineare alternata nelle righe di Y; perciò

$$\theta(I_p, Y) = \varphi(Y) = \alpha \det Y$$

per un opportuna costante $\alpha \in K$. Avremo $\theta(I_p, I_q) = \alpha \det(I_q) = \alpha$. Segue che

$$\theta(X, Y) = \det(X)\det(Y)\theta(I_p, I_q)$$

La matrice $M(I_p, I_q)$ è ottenuta dalla matrice I_n modificando ciascuna delle ultime q colonne di I_n con combinazioni lineari delle prime p colonne di I_n. Quindi $\alpha = \det M(I_p, I_q) = \det I_n = 1$. □

Data una matrice quadrata $A = (a_{ij})$ $n \times n$, indichiamo con

- A_{ij} la matrice quadrata $(n-1) \times (n-1)$ ottenuta da A cancellando la i-esima riga e la j-esima colonna
- $\Delta_{ij}(A) = \det A_{ij}$
- $C_{ij}(A) = (-1)^{i+j}\Delta_{ij} = (-1)^{i+j}\det A_{ij}$

Se non vi è possibilità di equivoco, scriveremo Δ_{ij} per $\Delta_{ij}(A)$ e C_{ij} per $C_{ij}(A)$.

Proposizione 4.13 (Sviluppo del determinante secondo gli elementi di una colonna)
Sia A una matrice quadrata $n \times n$. Allora abbiamo, per ogni $j = 1, \dots, n$:

$$\det A = \sum_{i=1}^{n} a_{ij} C_{ij}$$

Dimostrazione. Ci riduciamo a considerare il caso $j = 1$. Infatti, indicate con $C_1, C_2, \dots, C_n$ le colonne di A, cioè $A = (C_1, C_2, \dots, C_n)$; sia B la matrice ottenuta da A portando la j-esima colonna al primo posto, cioè $B = (C_j, C_1, C_2, \dots, C_{j-1}, C_{j+1}, \dots, C_n)$. Poiché B è ottenuta da A con $j-1$ scambi di colonne, si ha $\det B = (-1)^{j-1}\det A$. Se $B = (b_{ij})$, avremo $b_{i1} = a_{ij}$ e $B_{i1} = A_{ij}$; perciò, se la proposizione è vera per $j = 1$ avremo $\det B = \sum_i (-1)^{i+1} b_{i1}\Delta_{i1}(B)$. Quindi

$$\det A = (-1)^{j-1}\sum_i (-1)^{i+1} a_{ij}\Delta_{ij}(A) = \sum_i (-1)^{i+j} a_{ij}\Delta_{ij}(A)$$

Supponiamo quindi $j = 1$. Scriviamo

$$C_1 = a_{11}E_1 + \cdots + a_{n1}E_n, \quad \text{dove } E_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \longleftarrow \textit{i-esimo posto}$$

Sia A_i la matrice ottenuta dalla matrice A sostituendo la prima colonna C_1 con $a_{i1}E_1$. Per la linearità del determinante nella prima colonna, avremo $\det A = \sum_i \det A_i$.

Sia ora A'_i la matrice ottenuta da A_i portando la riga i-esima al primo posto, cioè

$$A'_i = \left(\begin{array}{c|ccc} a_{i1} & a_{i2} & \dots & a_{in} \\ \hline 0 & & & \\ \vdots & & A_{i1} & \\ 0 & & & \end{array}\right)$$

Per la proposizione precedente, $\det A'_i = a_{i1} \det A_{i1}$. Dato che la matrice A'_i è ottenuta dalla matrice A_i con $i - 1$ scambi di righe, avremo

$$\det A_i = (-1)^{i-1} \det A'_i = (-1)^{i+1} a_{i1} \det A_{i1} \tag{4.1}$$

Quindi $\det A = \sum_i (-1)^{i+1} a_{i1} \det A_{i1}$. Questo dimostra la proposizione nel caso $j = 1$. □

Proposizione 4.14 (Sviluppo del determinante secondo gli elementi di una riga)
Sia A una matrice quadrata $n \times n$. Allora abbiamo, per ogni $i = 1, \dots, n$:

$$\det A = \sum_{j=1}^{n} a_{ij} C_{ij}$$

Dimostrazione. Basta osservare che $\det A = \det {}^t\!A$ e che le righe di A sono le colonne di ${}^t\!A$. □

Proposizione 4.15
Sia A una matrice quadrata $n \times n$. Allora

1. *Se $k \neq j$, allora* $\sum_{i=1}^{n} a_{ik} C_{ij} = 0$
2. *Se $k \neq i$, allora* $\sum_{j=1}^{n} a_{ij} C_{kj} = 0$

Dimostrazione. Dimostriamo il primo punto, essendo la dimostrazione del secondo analoga. Fissiamo un indice j. Siano $C_1, \dots, C_n$ le colonne della matrice A: scriviamo $A = (C_1, \dots, C_n)$. Sia $B = (b_{lm})$ la matrice le cui colonne $D_1, \dots, D_n$ sono definite nel modo seguente:

$$\begin{cases} D_h = C_h & \text{se } h \neq j \\ D_j = C_k & \text{dove } k \neq j \end{cases}$$

La matrice B ha due colonne uguali e quindi il suo determinante è zero. Inoltre, per ogni i,

$$b_{ij} = a_{ik}, \qquad \Delta_{ij}(B) = \Delta_{ij}(A)$$

Segue che, sviluppando il determinante di B secondo gli elementi della j-esima colonna,

$$0 = \det B = \sum_{i=1}^{n} (-1)^{i+j} b_{ij} \Delta_{ij}(B) = \sum_{i=1}^{n} (-1)^{i+j} a_{ik} \Delta_{ij}(A) = \sum_{i=1}^{n} a_{ik} C_{ij}$$

□

Proposizione 4.16

Sia $A = (a_{ij}) \in \mathrm{M}_n(K)$. *Sia* $C = \big(C_i j(A)\big)$, *la matrice (detta dei* ***cofattori*** *di A) il cui elemento di posto* (i, j) *sia* $C_{ij}(A)$. *Allora*

$$A \cdot {}^t\!C = \det(A) \cdot I_n$$

Dimostrazione. Se $D = {}^t\!C = (d_{hk})$ avremo $d_{hk} = C_{kh}$, quindi nel prodotto AD l'elemento di posto (i, j) è

$$\sum_{k=1}^{n} a_{ik} d_{kj} = \sum_{k=1}^{n} a_{ik} C_{jk} = \begin{cases} \det(A) & \text{se } i = j \\ 0 & \text{se } i = j \end{cases}$$

per le proposizioni precedenti. □

Proposizione 4.17 (Teorema di Binet)

Siano A, B *due matrici quadrate* $n \times n$. *Allora*

$$\det AB = \det(A)\det(B)$$

Dimostrazione. Siano $A = (a_{ij})$ ed $\mathsf{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base dei versori di K^n. Sia Φ l'unica forma n-multilineare alternante tale che $\Phi(\mathsf{E}) = 1$. Se $\mathbf{v}_k = \sum_j a_{kj}\mathbf{e}_j$ con $k = 1, \ldots, n$, allora, per la proposizione 4.4,

$$\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \det(A)\Phi(\mathsf{E}) = \det(A)$$

Sia ora $B = (b_{lm})$. Se $\mathbf{w}_i = \sum_k b_{ik}\mathbf{v}_k$ con $i = 1, \ldots, n$, allora

$$\Phi(\mathbf{w}_1, \ldots, \mathbf{w}_n) = \det(B)\Phi(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \det(B)\det(A)$$

Ma

$$\mathbf{w}_i = \sum_k b_{ik}\mathbf{v}_k = \sum_{k,j} b_{ik} a_{kj}\mathbf{e}_j = \sum_j c_{ij}\mathbf{e}_j$$

dove $C = (c_{ij}) = BA$. Quindi

$$\Phi(\mathbf{w}_1, \ldots, \mathbf{w}_n) = \det(C)\Phi(\mathsf{E}) = \det(C)$$

Segue che $\det(BA) = \det(B)\det(A)$. □

Proposizione 4.18

Sia A una matrice $n \times n$ *a coefficienti in K. Allora*

$$A \text{ è invertibile se e solo se } \det(A) \neq 0$$

Dimostrazione. Se esiste A^{-1}, allora $A \cdot A^{-1} = I_n$, quindi per il teorema di Binet, $\det(A)\det(A^{-1}) = \det(I_n) = 1$; di conseguenza, $\det(A) \neq 0$.

Viceversa, se $\det A \neq 0$, dalla relazione $A \cdot {}^t\!C = \det(A) \cdot I_n$ dove C è la matrice dei cofattori di A, ricaviamo che

$$A^{-1} = \frac{1}{\det A}\,{}^t\!C$$

□

5 Autovalori e autovettori

5.1 Autovalori, autovettori e polinomio caratteristico

Sia V uno spazio vettoriale su K di dimensione finita, e φ un endomorfismo di W. Ci chiediamo se esiste un sottospazio W di V tale che $\varphi|_W$ sia un endomorfismo di W di tipo scalare, cioè $\varphi|_W(\mathbf{v}) = \lambda\mathbf{v}$ per ogni $\mathbf{v} \in W$ con λ scalare fisso indipendente da $\mathbf{v}$. Diamo dunque la seguente

Definizione. Un elemento $\lambda \in K$ è detto **autovalore** di φ se esiste un vettore $\mathbf{v} \neq \mathbf{o}$ con $\varphi(\mathbf{v}) = \lambda\mathbf{v}$. Un tale $\mathbf{v}$ è detto **autovettore** di φ relativo all'autovalore λ.

Si noti che, se $\mathbf{v} = \mathbf{o}$, allora $\varphi(\mathbf{v}) = \lambda\mathbf{v}$ è verificato per ogni $\lambda \in K$.

Se indichiamo con $\lambda\,\mathrm{id}$ l'endomorfismo di V dato dalla moltiplicazione per λ, vediamo che λ è un autovalore se e solo se esiste un vettore $\mathbf{v} \neq \mathbf{o}$ in V tale $(\varphi - \lambda\,\mathrm{id})(\mathbf{v}) = \mathbf{o}$. Quindi λ è un autovalore di φ se e solo se $\ker(\varphi - \lambda\,\mathrm{id}) \neq \{\mathbf{o}\}$. Nel seguito, indicheremo con φ_λ l'endomorfismo $\varphi - \lambda\,\mathrm{id}$. Possiamo perciò dire che λ è un autovalore di φ se e solo se $\ker\varphi_\lambda \neq \{\mathbf{o}\}$. In particolare, $\lambda = 0$ è un autovalore se e solo $\ker\varphi \neq \{\mathbf{o}\}$.

Fissato un endomorfismo φ di V, e dato $\lambda \in K$, poniamo $V_\lambda = \ker\varphi_\lambda$. Osserviamo che λ è un autovalore se e solo se $\dim V_\lambda \geqslant 1$.

Definizione. Con le notazioni di sopra, sia λ un autovalore di φ. Il sottospazio V_λ di V è detto **autospazio** di λ. V_λ è costituito dal vettore $\mathbf{o}$ e da tutti gli autovettori di V; per questo sarà detto anche **spazio degli autovettori** di λ.

Fissiamo ora una base E di V, e sia $A = (a_{ij})$ la matrice rappresentativa $M^{\mathsf{E}}_{\mathsf{E}}(\varphi)$ di φ nella base E. La matrice rappresentantiva, nella stessa base, di $\varphi_\lambda = \varphi - \lambda\,\mathrm{id}$ è allora $A_\lambda = A - \lambda I$. Quindi λ è un autovalore se e solo se $\det A_\lambda = 0$.

Gli autovalori dell'endomorfismo φ sono allora le radici in K del polinomio $P_\varphi(x) = \det(A - xI)$.

Definizione. Il polinomonio $P_\varphi(x) = \det(A - xI)$ è detto **polinomio caratteristico** di φ.

Osserviamo che il polinomio dipende soltanto dall'endomorfismo φ e non dalla matrice rappresentativa. Infatti, posto $P_A(x) = \det(A - xI)$, se $B = Q^{-1}AQ$ è la matrice rappresentativa di φ in una base F, avremo

$$\begin{aligned} P_B(x) &= \det(B - xI) = \det(Q^{-1}AQ - xI) = \det(Q^{-1}AQ - Q^{-1}xIQ) \\ &= \det\big(Q^{-1}(A - xI)Q\big) = \det(A - xI) = P_A(x) \end{aligned}$$

Abbiamo visto dunque che matrici simili hanno lo stesso polinomio caratteristico.

Il polinomio caratteristico $P_\varphi(x)$ è un polinomio in x di grado n, in cui il coefficiente del termine di grado massimo è $(-1)^n$. Il termine noto è $P_\varphi(0) = \det(A)$. Il coefficiente del termine di grado $n-1$ è $(-1)^{n-1}\operatorname{Tr} A$; in particolare se $n = 2$, abbiamo che $P_\varphi(x) = x^2 - \operatorname{Tr}(A)x + \det(A)$.

Osserviamo che in generale, un endomorfismo φ può non avere alcun autovalore. Ad esempio, se consideriamo $V = \mathbf{R}^2$ dotato della base canonica E dei versori $\{\mathbf{e}_1, \mathbf{e}_2\}$ sia φ l'endomorfismo rappresentato dalla matrice $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$; tale endomorfismo è la rotazione di $\pi/2$ in senso antiorario. È chiaro, quindi, che nessun vettore non nullo è mandato in un multiplo di se stesso: infatti, calcolando $P_\varphi(x) = \det(A - xI)$ otteniamo $x^2 + 1$, che non ha radici reali.

Notiamo che la dimensione dell'autospazio V_λ sarà

$$\dim V_\lambda = \dim\ker\varphi_\lambda = n - \rho(A_\lambda) = n - \rho(A - \lambda I)$$

5.2 Polinomio minimo

Sia $K[x]$ l'insieme dei polinomi in x a coefficienti in K; sappiamo che $K[x]$ con le usuali operazioni di somma e moltiplicazione di polinomi è un anello commutativo (cioè la moltiplicazione è commutativa) con un elemento neutro (per la moltiplicazione). Sappiamo anche che, con le operazioni di somma e moltiplicazione per uno scalare, $K[x]$ è uno spazio vettoriale su K.

Nell'insieme $\operatorname{End}_K(V)$, dove V è uno spazio vettoriale su K, sono definite le operazioni di somma, composizione di due endomorfismi e moltiplicazione per uno scalare. Se definiamo la moltiplicazione $\psi \cdot \varphi$ come $\psi \circ \varphi$, l'insieme $\operatorname{End}_K(V)$ è un anello con elemento neutro moltiplicativo id_V. Lo stesso insieme, con le operazioni di somma e moltiplicazione per uno scalare, è uno spazio vettoriale su K. Se V è di dimensione finita n, ed E è una sua base fissata, possiamo identificare $\operatorname{End}_V(K)$ con $M_n(K)$ associando ad ogni φ la sua matrice rappresentativa $M^{\mathsf{E}}_{\mathsf{E}}(\varphi)$. Sappiamo che in questa corrispondenza $M^{\mathsf{E}}_{\mathsf{E}}(\psi\varphi) = M^{\mathsf{E}}_{\mathsf{E}}(\psi)M^{\mathsf{E}}_{\mathsf{E}}(\varphi)$. In particolare $\operatorname{End}_K(V)$ non è commutativo se $\dim V > 1$.

Sia ora V uno spazio vettoriale di dimensione finita n, e φ un suo endomorfismo fissato. Consideriamo l'applicazione

$$\begin{aligned} S\colon K[x] &\longrightarrow \operatorname{End}_K(V) \\ x &\longmapsto \varphi \\ F(x) &\longmapsto F(\varphi) \end{aligned}$$

dove, se $F(x) = a_0 + a_1x + \cdots + a_nx^n$, l'endomorfismo $F(\varphi)$ è $a_0\,\mathrm{id} + a_1\varphi + \cdots + a_n\varphi^n$; notiamo che per la notazione introdotta sopra, $\varphi^k = \varphi \circ \cdots \circ \varphi$ applicato k volte; mentre per convenzione $\varphi^0 = \mathrm{id}$.

L'immagine di S è l'insieme degli endomorfismi di V esprimibili come polinomi in φ: essi costituiscono un sottoanello commutativo di $\operatorname{End}_K(V)$ che viene indicato con $K[\varphi]$. L'applicazione S rispetta la somma, la moltiplicazione ed in particolare la moltiplicazione per uno scalare: è quindi un omomorfismo fra anelli ed un'applicazione lineare fra spazi vettoriali. Se E è una base di V, posto $A = M_{\mathsf{E}}^{\mathsf{E}}(\varphi)$ e identificando $\mathrm{End}_K(V)$ con $M_n(K)$, possiamo interpretare S come

$$\begin{aligned} S\colon K[x] &\longrightarrow M_n(K) \\ x &\longmapsto A \\ F(x) &\longmapsto F(A) \end{aligned}$$

Notiamo ancora che $k[x]$ è uno spazio vettoriale di dimensione infinita, mentre $M_n(K)$ ha dimensione finita n^2. Consideriamo l'insieme $\ker S$ formato da tutti i polinomi $F(x)$ tali che $F(\varphi)$ è l'endomorfismo nullo: questo insieme ha le seguenti caratteristiche, di immediata verifica:

a) Se $F(x), G(x) \in \ker S$, allora $(F+G)(x) \in \ker S$
b) Se $F(x) \in \ker S$ ed $R(x)$ è un polinomio qualsiasi in $K[x]$, allora $(RF)(x) \in \ker S$

Definizione. Un sottoinsieme J di $K[x]$ si dice **ideale** di $K[x]$ se soddisfa le due proprietà precedenti, cioè

a) Se $F(x), G(x) \in J$, allora $F(x) + G(x) \in J$
b) Se $F(x) \in J$ ed $R(x)$ è un polinomio qualsiasi in $K[x]$, allora $R(x)F(x) \in J$

Si osservi che se $F(x) \in J$ e $\lambda \in K$, allora anche $\lambda F(x) \in J$; perciò, per le due proprietà, J è un sottospazio vettoriale di $K[x]$.

Proposizione 5.1

Sia J un ideale di $K[x]$, e sia $G(x)$ un polinomio in J di grado minimo. Allora J è l'insieme di tutti i polinomi multipli di $G(x)$.

Dimostrazione. Se $F(x) \in J$, allora possiamo scrivere $F(x) = Q(x)G(x) + R(x)$ dove $\partial R(x) < \partial G(x)$. Per la proprietà b), $Q(x)G(x) \in J$, e quindi, per la proprietà a), abbiamo $R(x) = F(x) - Q(x)G(x) \in J$. Per la minimalità di $\partial G(x)$, segue che $R(x) = 0$. □

Osservazione 5.2. Dalla dimostrazione segue che i polinomi di grado minimo di J differiscono per una costante moltiplicativa.

Definizione. Un polinomio di grado minimo di J è detto **generatore** di J. Scriveremo $J = \big(G(x)\big)$.

Come abbiamo osservato, i generatori differiscono per una costante moltiplicativa: ne esiste pertanto uno solo monico, cioè quello avente come coefficiente del termine di grado massimo il numero 1.

Ritorniamo ora all'insieme $\ker S$ definito più sopra.

Definizione. Il generatore monico di $\ker S$ è detto il **polinomio minimo** di φ, e sarà indicato come $Q_\varphi(x)$.

Il polinomio $Q_\varphi(x)$ è quindi il polinomio monico di grado minimo tale che $Q_\varphi(\varphi)$ sia l'endomorfismo nullo. Il polinomio minimo $Q_\varphi(x)$ ed il polinomio caratteristico $P_\varphi(x)$ sono legati dall'importante teorema di Cayley-Hamilton, di cui esistono due versioni, la seconda più precisa della prima.

Teorema 5.3 (Cayley-Hamilton I)

Il polinomio caratteristico $P_\varphi(X) \in \ker S$; equivalentemente, $P_\varphi(\varphi)$ è l'endomorfismo nullo; equivalentemente, il polinomio minimo $Q_\varphi(x)$ divide il polinomio caratteristico.

Teorema 5.4 (Cayley-Hamilton II)

Il polinomio minimo $Q_\varphi(x)$ divide il polinomio caratteristico.$P_\varphi(x)$ e questi due polinomi hanno gli stessi fattori irriducibili.

Osservazione 5.5. I fattori irriducibili di $Q_\varphi(x)$ compaiono quindi in $P_\varphi(x)$ con una molteplicità maggiore o uguale a quella con cui compaiono in $Q_\varphi(x)$.

Osservazione 5.6. Dato φ, sappiamo che $\partial P_\varphi(x) = n$. In generale avremo $\partial Q_\varphi(x) \leqslant n$ e tale grado dipende da φ. In generale, $P_\varphi(x) \neq Q_\varphi(x)$. Per esempio, se $\varphi = \lambda \operatorname{id}$, allora è chiaro che $P_\varphi(x) = (1-x)^n$ mentre $Q_\varphi(x) = x - 1$.

5.3 Molteplicità algebrica e geometrica

Sia V uno spazio vettoriale su K di dimensione n e sia φ un suo endomorfismo. Sia poi λ un autovalore di φ ed m_λ la molteplicità di λ come radice del polinomio caratteristico $P_\varphi(x)$. Questo significa che $P_\varphi(x) = (\lambda - x)^{m_\lambda} G(x)$ con $G(\lambda) \neq 0$. Sia, ancora, V_λ l'autospazio relativo a λ.

Definizione. L'intero m_λ è detto **molteplicità algebrica** di λ. La dimensione $r_\lambda = \dim(V_\lambda)$ è detta **molteplicità geometrica** di λ.

Segue dalle definizioni che

$$1 \leqslant m_\lambda \leqslant n, \qquad 1 \leqslant r_\lambda \leqslant n$$

Proposizione 5.7

Con le notazioni di cui sopra, $r_\lambda \leqslant m_\lambda$.

Dimostrazione. Poniamo $r = r_\lambda$ ed $s = n - r$. Sia $\{\mathbf{e}_1, \dots, \mathbf{e}_r\}$ una base di V_λ che completiamo ad una base $\mathsf{E} = \{\mathbf{e}_1, \mathbf{e}_r, \mathbf{e}_{r+1}, \dots, \mathbf{e}_n\}$ di V. La matrice rappresentativa $A = M_{\mathsf{E}}^{\mathsf{E}}(\varphi)$ è quindi della forma

$$A = \begin{pmatrix} \lambda & & & \\ & \ddots & & B \\ & & \lambda & \\ & O & & C \end{pmatrix}$$

cioè, contiene in alto a sinistra un blocco quadrato diagonale $r \times r$ con tutti λ sulla diagonale, al di sotto del quale vi sono solo zeri. Abbiamo così

$$A - xI_n = \begin{pmatrix} \lambda - x & & & \\ & \ddots & & B \\ & & \lambda - x & \\ & O & & C - xI_s \end{pmatrix}$$

Perciò per la proposizione ??, $P_\varphi(x) = \det(A - xI_n) = (\lambda - x)^r \det(C - xI_s)$. Vediamo quindi che λ è una radice di $P_\varphi(x)$ con molteplicità algebrica almeno $r = r_\lambda$. In altre parole $m_\lambda \geqslant r_\lambda$. $\square$

Esempio

Sia $\varphi\colon \mathbf{R}^2 \to \mathbf{R}^2$ definito rispetto alla base standard dei versori da $\varphi(\mathbf{e}_1) = \mathbf{e}_1$, $\varphi(\mathbf{e}_2) = \mathbf{e}_1 + \mathbf{e}_2$. Abbiamo $P_\varphi(x) = (x-1)^2$, cioè c'è un solo autovalore $\lambda = 1$ con $m_\lambda = 2$. L'autospazio V_1, corrispondente all'unico autovalore, ha invece dimensione $r_1 = 1$.

5.4 Indipendenza di autovettori

Siano V uno spazio vettoriale su K di dimensione n, e φ un endomorfismo di V. Si hanno allora le due seguenti proposizioni:

Proposizione 5.8

Siano $\lambda_1, \dots, \lambda_k$ autovalori distinti di φ (appartenenti a K); per ogni $i = 1, \dots, k$, sia $\mathbf{v}_i$ un autovettore non nullo relativo a λ_i. Allora i vettori $\mathbf{v}_1, \dots, \mathbf{v}_k$ sono linearmente indipendenti.

Dimostrazione. Se fossero linearmente dipendenti, esisterebbe un indice $i \geqslant 1$ tale che i vettori $\mathbf{v}_1, \dots, \mathbf{v}_i$ sono linearmente indipendenti, mentre $\mathbf{v}_{i+1}$ è combinazione lineare dei precedenti, cioè

$$\mathbf{v}_{i+1} = a_1\mathbf{v}_1 + \dots a_i\mathbf{v}_i \tag{5.1}$$

dove i coefficienti a_i non sono tutti nulli perché $\mathbf{v}_{i+1} \neq \mathbf{o}$. Applicando la φ alla (5.1), otteniamo

$$\lambda_{i+1}\mathbf{v}_{i+1} = a_1\lambda_1\mathbf{v}_1 + \dots a_i\lambda_i\mathbf{v}_i \tag{5.2}$$

Usando la (5.1), possiamo riscrivere la (5.2) come

$$\lambda_{i+1}\sum_{k=1}^{i} a_k\mathbf{v}_k = \sum_{h=1}^{i} a_h\lambda_h\mathbf{v}_k$$

cioè

$$a_1(\lambda_{i+1} - \lambda_1)\mathbf{v}_1 + a_2(\lambda_{i+2} - \lambda_2)\mathbf{v}_2 + \dots a_i(\lambda_{i+1} - \lambda_i)\mathbf{v}_i = 0 \tag{5.3}$$

Dato che $\mathbf{v}_1, \dots, \mathbf{v}_i$ sono linearmente indipendenti, i coefficienti della combinazione lineare (5.3) sono tutti nulli. Poiché $\lambda_{i+1} \neq \lambda_j$ per ogni $j = 1, \dots, i$, segue che $a_j = 0$ per ogni $j = 1, \dots, i$ contro l'ipotesi che tali coefficienti non fossero tutti nulli. □

Proposizione 5.9

Sia, con le stesse ipotesi e notazioni della proposizione precedente, V_{λ_i} l'autospazio relativo a λ_i, con $i = 1, \dots, k$. Allora

$$\sum_{i=1}^{k} V_i \simeq \bigoplus_{i=1}^{k} V_i$$

Dimostrazione. In caso contrario, per la proposizione 2.11, esisterebbe un vettore $\mathbf{v}_i \neq \mathbf{o}$ tale che

$$\mathbf{v}_i \in V_{\lambda_i} \cap \bigoplus_{k=1}^{i-1} V_{\lambda_k}$$

cioè $\mathbf{v}_i = \mathbf{v}_1 + \cdots + \mathbf{v}_{i-1}$ dove $\mathbf{v}_k \in V_{\lambda_k}$ ed i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$ non sono tutti nulli (non lo è $\mathbf{v}_i$).

Possiamo quindi supporre senza perdita di generalità che i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$ siano tutti non nulli. Otteniamo così una relazione di dipendenza lineare

$$\mathbf{v}_i - \mathbf{v}_1 - \cdots - \mathbf{v}_{i-1} = \mathbf{o}$$

contro la proposizione precedente. □

5.5 Endomorfismi triangolabili

Definizione. Sia φ un'endomorfismo di uno spazio vettoriale V di dimensione n su K. Diciamo che φ è triangolabile se esiste una base in cui la matrice rappresentantiva A di φ è triangolare superiore, cioè della forma $a_{ij} = 0$ per ogni $i > j$.

Teorema 5.10

Un endomorfismo φ di V spazio vettoriale su K è triangolabile se e solo se il suo polinomio caratteristico $P_\varphi(X)$ si fattorizza completamente in $K[X]$.

(Questo è sempre vero se $K = \mathbf{C}$, essendo $\mathbf{C}$ algebricamente chiuso)

Dimostrazione. Se $A = (a_{ij})$ è una matrice triangolare superiore che rappresenta φ in una base E, allora

$$P_\varphi(X) = \prod_i (a_{ii} - X)$$

Quindi $P_\varphi(X)$ è completamente fattorizzato in $K[X]$ e gli autovalori sono esattamente $\lambda_i = a_{ii}$.

Viceversa, supponiamo che $P_\varphi(X)$ sia completamente fattorizzato in $K[X]$ e dimostriamo che φ è triangolabile per induzione sulla dimensione n di V: se $n = 1$, non vi è nulla da dimostrare.

Supponiamo che l'affermazione sia vera per spazi vettoriali di dimensione minore o uguale a $n-1$, dove $\dim V = n$. Sia λ_1 un autovalore e $\mathbf{v}_1 \neq \mathbf{o}$ un autovettore relativo a λ_1. Sia poi $\mathsf{E} = \{\mathbf{v}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ una base ottenuta per completamento a partire da $\mathbf{v}_1$. Denotiamo con W lo spazio vettoriale, di dimensione $n-1$, $\mathcal{L}(\mathbf{w}_2, \ldots, \mathbf{w}_n)$. L'insieme $\mathsf{E}' = \{\mathbf{w}_2, \ldots, \mathbf{w}_n\}$ è una base di W

La matrice A che rappresenta φ nella base E è della forma

$$A = \left(\begin{array}{c|ccc} \lambda_1 & a_{12} & \ldots & a_{1n} \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right)$$

Sia ψ l'endomorfismo di W la cui matrice rappresentativa nella base E' è B.

Se paragoniamo l'immagine di $\mathbf{w}_i$, con $i = 2, \ldots, n$, attraverso le due applicazioni φ e ψ vediamo che $\varphi(\mathbf{w}_i) = a_{1i}\mathbf{v}_1 + \psi(\mathbf{w}_i)$ cioè $\varphi(\mathbf{w}_i) - \psi(\mathbf{w}_i) \in \mathcal{L}(\mathbf{v}_1)$. Data la linearità di φ e ψ avremo che, per ogni $\mathbf{w} \in W$, $\varphi(\mathbf{w}) - \psi(\mathbf{w}) \in \mathcal{L}(\mathbf{v}_1)$. Il polinomio caratteristico dell'endomorfismo φ è

$$P_\varphi(X) = \det(A - XI_n) = (\lambda_1 - X)\det(B - XI_{n-1}) = (\lambda_1 - X)P_\psi(X)$$

Poichè $P_\varphi(X)$ si fattorizza completamente in $K[X]$, anche il polinomio $P_\psi(X)$ si fattorizza completamente in $K[X]$.

Per ipotesi induttiva, esiste quindi una base $\mathsf{F}' = \{\mathbf{v}_2, \dots, \mathbf{v}_n\}$ di W in cui ψ è rappresentato da una matrice triangolare superiore C. Prendiamo allora l'insieme $\mathsf{F} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$: esso forma una base di V. In questa base la matrice rappresentativa di φ è triangolare superiore perché, per ogni $i = 2, \dots, n$, $\varphi(\mathbf{v}_i)$ differisce da $\psi(\mathbf{v}_i)$ soltanto per un vettore in $\mathcal{L}(\mathbf{v}_1)$. Infatti avremo

$$M_{\mathsf{F}}^{\mathsf{F}}(\varphi) = D = \left(\begin{array}{c|ccc} \lambda_1 & d_{12} & \dots & d_{1n} \\ \hline 0 & & & \\ \vdots & & C & \\ 0 & & & \end{array}\right)$$

□

5.6 Teorema di Cayley-Hamilton

Dimostriamo il primo teorema di Cayley-Hamilton nell'ipotesi che il campo K sia algebricamente chiuso (per esempio, $\mathbf{C}$).

Teorema 5.11 (Primo teorema di Cayley-Hamilton)
Sia V uno spazio vettoriale di dimensione n su K, dove K è algebricamente chiuso. Sia φ un endomorfismo di V e $P_\varphi(X)$ il suo polinomio caratteristico. Allora $P_\varphi(\varphi)$ è l'endomorfismo nullo.

Dimostrazione. L'ipotesi che il campo sia algebricamente chiuso implica che $P_\varphi(X)$ si fattorizza completamente in $K[X]$, quindi in una opportuna base $\mathsf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di V φ è rappresentato da una matrice triangolare

$$\varphi \longleftrightarrow \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Inoltre $P_\varphi(X) = \prod_{i=1}^n (\lambda_i - X)$, e

$$\varphi(\mathbf{e}_i) = \lambda_i \mathbf{e}_i + \mathbf{v}_i, \qquad \text{con } \mathbf{v}_i \in \mathcal{L}(\mathbf{e}_1, \dots, \mathbf{e}_{i-1})$$

Se poniamo $\varphi_{\lambda_i} = \varphi - \lambda_i \,\mathrm{id}_V$, allora $\varphi_{\lambda_i}(\mathbf{e}_i) = \mathbf{v}_i \in \mathcal{L}(\mathbf{e}_1, \dots, \mathbf{e}_{i-1})$. Osserviamo che gli endomorfismi φ_{λ_i} e φ_{λ_j} commutano, come si può verificare direttamente (in generale, due endomorfismi esprimibili come polinomi in φ commutano). Dato che $\varphi_{\lambda_1}(\mathbf{e}_1) = \mathbf{0}$, allora φ_{λ_1} ristretto a $\mathcal{L}(\mathbf{e}_1)$ è l'endomorfismo nullo.

Verifichiamo che $\varphi_{\lambda_2} \circ \varphi_{\lambda_1}$ ristretto a $\mathcal{L}(\mathbf{e}_1, \mathbf{e}_2)$ è zero: infatti

$$(\varphi_{\lambda_2} \circ \varphi_{\lambda_1})(\mathbf{e}_1) = \varphi_{\lambda_2}\big(\varphi_{\lambda_1}(\mathbf{e}_1)\big) = \varphi_{\lambda_2}(\mathbf{0}) = \mathbf{0}$$
$$(\varphi_{\lambda_2} \circ \varphi_{\lambda_1})(\mathbf{e}_2) = (\varphi_{\lambda_1} \circ \varphi_{\lambda_2})(\mathbf{e}_2) = \varphi_{\lambda_1}\big(\varphi_{\lambda_2}(\mathbf{e}_2)\big) = \varphi_{\lambda_1}(\mathbf{v}_2) = \mathbf{0}$$

L'ultima uguaglianza segue dal fatto che $\mathbf{v}_2 \in \mathcal{L}(\mathbf{e}_1)$.

Consideriamo ora $\varphi_{\lambda_3} \circ \varphi_{\lambda_2} \circ \varphi_{\lambda_1}$: in modo analogo al caso precedente si verifica che $\varphi_{\lambda_3} \circ \varphi_{\lambda_2} \circ \varphi_{\lambda_1}$ ristretto a $\mathcal{L}(\mathbf{e}_1, \mathbf{e}_2), \mathbf{e}_3$ è zero. Analogamente si verifica che, per ogni $i = 1, \ldots, n$ l'endomorfismo $\varphi_{\lambda_i} \circ \varphi_{\lambda_{i-1}} \circ \ldots \varphi_{\lambda_1}$ si annulla su $\mathcal{L}(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_i)$. In particolare, per $i = n$, l'endomorfismo $\varphi_{\lambda_n} \circ \varphi_{\lambda_{n-1}} \circ \varphi_{\lambda_1}$ è nullo su tutto V. Ora,

$$\begin{aligned} P_\varphi(\varphi) &= (\lambda_1 \operatorname{id} - \varphi) \circ (\lambda_2 \operatorname{id} - \varphi) \circ \cdots \circ (\lambda_n \operatorname{id} - \varphi) \\ &= (-1)^n \varphi_{\lambda_1} \circ \varphi_{\lambda_2} \circ \cdots \circ \varphi_{\lambda_n} \\ &= (-1)^n \varphi_{\lambda_n} \circ \cdots \circ \varphi_{\lambda_2} \circ \varphi_{\lambda_1} \end{aligned}$$

Quindi $P_\varphi(\varphi)$ è l'endomorfismo nullo di V. □

5.7 Endomorfismi diagonalizzabili

Definizione. Sia V uno spazio vettoriale su K di dimensione n, e sia φ un endomorfismo di V. Diciamo che φ è **diagonalizzabile** se esiste una base E di V in cui la matrice rappresentativa di φ è diagonale.

Si ha il seguente

Teorema 5.12 (Criterio di diagonalizzabilità)
Le condizioni seguenti sono equivalenti

1. *φ è diagonalizzabile*
2. a) *Il polinomio caratteristico $P_\varphi(X)$ si fattorizza completamente in $K[X]$*
   b) *Per ogni autovalore λ, la molteplicità algebrica m_λ è uguale a quella geometrica r_λ*
3. *Se $\lambda_1, \ldots, \lambda_k$ sono tutti gli autovalori distinti (in K) di φ, allora $V = \bigoplus_{i=1}^k V_{\lambda_i}$.*

Esempio

Caso degli autovalori distinti.

5.8 Endomorfismi nilpotenti

6 Spazi euclidei ed unitari

6.1 Prodotti scalari reali e spazi euclidei

Definizione. Sia V uno spazio vettoriale su $\mathbf{R}$. Un **prodotto scalare** su V è un'applicazione

$$\begin{aligned} V \times V &\longrightarrow V \\ (\mathbf{v}, \mathbf{w}) &\longmapsto \langle \mathbf{v}, \mathbf{w} \rangle \end{aligned}$$

che soddisfa le seguenti condizioni, per ogni $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2, \mathbf{w}, \mathbf{w}_1, \mathbf{w}_2 \in V$ e per ogni $a_1, a_2, b_1, b_2 \in \mathbf{R}$:

1. $\langle a_1\mathbf{v}_1 + a_2\mathbf{v}_2, \mathbf{w} \rangle = a_1\langle \mathbf{v}_1, \mathbf{w} \rangle + a_2\langle \mathbf{v}_2, \mathbf{w} \rangle$
2. $\langle \mathbf{v}, b_1\mathbf{w}_1 + b_2\mathbf{w}_2 \rangle = b_1\langle \mathbf{v}, \mathbf{w}_1 \rangle + b_2\langle \mathbf{v}, \mathbf{w}_2 \rangle$
3. $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$
4. $\langle \mathbf{v}, \mathbf{v} \rangle \geqslant 0$, con $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ se e solo se $\mathbf{v} = \mathbf{0}$.

Le condizioni 1 e 2 dicono che $\langle\ ,\ \rangle$ è bilineare. La condizione 3 dice che $\langle\ ,\ \rangle$ è simmetrica. Una funzione bilineare simmetrica che soddisfa la condizione 4 è detta **definita positiva**.

La bilinearità implica che, per ogni $a \in \mathbf{R}$, $\mathbf{v}, \mathbf{w} \in V$:

5. $\langle \mathbf{v}, a\mathbf{w} \rangle = \langle a\mathbf{v}, \mathbf{w} \rangle = a\langle \mathbf{v}, \mathbf{w} \rangle$

Inoltre abbiamo, per ogni $\mathbf{v} \in V$:

6. $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$

Infatti, $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{0} + \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{0}, \mathbf{v} \rangle + \langle \mathbf{0}, \mathbf{v} \rangle$ quindi $\langle \mathbf{0}, \mathbf{v} \rangle = 0$.

Definizione. Uno spazio vettoriale reale dotato di un prodotto scalare è detto **spazio euclideo**.

Esempi

1. $\mathbf{R}^n$ standard. Consideriamo lo spazio $\mathbf{R}^n$ col seguente prodotto scalare:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i, \qquad \text{dove } \mathbf{x} = (x_1, \ldots, x_n),\ \mathbf{y} = (y_1, \ldots, y_n)$$

Lo spazio $\mathbf{R}^n$, con questo prodotto scalare, è detto **spazio euclideo $\mathbf{R}^n$ standard**. Notiamo che:

$$\text{posto } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \text{allora} \quad \langle \mathbf{x}, \mathbf{y} \rangle = {}^t X Y \text{ (come prodotto di matrici)}$$

2. Sia V lo spazio delle funzioni continue a valori reali definite sull'intervallo $[a, b]$ della retta reale. Allora
$$\langle f, g\rangle = \int_a^b f(x)g(x)\,\mathrm{d}t$$
è un prodotto scalare.

3. Consideriamo lo spazio euclideo $\mathbf{R}^3$ standard. Dati $\mathbf{x} = (x_1, x_2, x_3)$ ed $\mathbf{y} = (y_1, y_2, y_3)$; siano O l'origine di $\mathbf{R}^3$, P il punto di coordinate (x_1, x_2, x_3) e Q il punto di coordinate (y_1, y_2, y_3). Sia infine α l'angolo interno (cioè $0 \leqslant \alpha < \pi$) formato dai segmenti OP e OQ. Allora
$$\langle \mathbf{x}, \mathbf{y}\rangle = \|\mathbf{x}\| \cdot \|\mathbf{y}\| \cos\alpha$$
dove $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}$. In particolare, $\cos\alpha = \dfrac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$

4. Sia V uno spazio vettoriale reale di dimensione n. Se fissiamo una base $\mathsf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ possiamo definire il **prodotto scalare standard** $\langle ;, \rangle_{\mathsf{E}}$ rispetto alla base E nel modo seguente: dati $\mathbf{v} = \sum_i a_i\mathbf{e}_i$, e $\mathbf{w} = \sum_j b_j\mathbf{e}_i$, allora
$$\langle \mathbf{v}, \mathbf{w}\rangle_{\mathsf{E}} = \sum_i a_i b_i$$

Norma

Sia V uno spazio euclideo con prodotto scalare $\langle\, , \rangle$; sia $\mathbf{v} \in V$. Definiamo la **norma** di $\mathbf{v}$ come $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$.
Nello spazio euclideo $\mathbf{R}^n$ standard, se $\mathbf{x} = (x_1, \dots, x_n)$, allora $\|\mathbf{x}\| = \sqrt{\sum_i x_i^2}$.
La norma gode delle seguenti proprietà, per ogni $\mathbf{v} \in V$ ed $a \in \mathbf{R}$:

1. $\|\mathbf{v}\| \geqslant 0$, con $\|\mathbf{v}\| = 0$ se e solo se $\mathbf{v} = \mathbf{0}$
2. $\|a\mathbf{v}\| = |a| \cdot \|\mathbf{v}\|$

La prima proprietà segue immediatamente dalla definizione di prodotto scalare; per la seconda abbiamo che
$$\|a\mathbf{v}\|^2 = \langle a\mathbf{v}, a\mathbf{v}\rangle = a^2\|\mathbf{v}\|^2$$

Teorema 6.1 (Disuguaglianza triangolare, o, di Cauchy-Schwartz)
Sia V uno spazio euclideo. Allora, per ogni $\mathbf{v}, \mathbf{w} \in V$,
$$\|\mathbf{v} + \mathbf{w}\| \leqslant \|\mathbf{v}\| + \|\mathbf{w}\|$$

Dimostrazione. Dato che i numeri coinvolti nella disuguaglianza sono numeri reali non negativi, basta dimostrare la disuguaglianza sui loro quadrati. Abbiamo che
$$\|\mathbf{v} + \mathbf{w}\|^2 = \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{v}\rangle + \langle \mathbf{w}, \mathbf{w}\rangle + 2\langle \mathbf{v}, \mathbf{w}\rangle$$

mentre
$$(||\mathbf{v}|| + ||\mathbf{w}||)^2 = ||\mathbf{v}||^2 + ||\mathbf{w}||^2 + 2||\mathbf{v}||\,||\mathbf{w}|| = \langle \mathbf{v}, \mathbf{v}\rangle + \langle \mathbf{w}, \mathbf{w}\rangle + 2||\mathbf{v}|| \cdot ||\mathbf{w}||$$
Dobbiamo quindi dimostrare che
$$\langle \mathbf{v}, \mathbf{w}\rangle \leqslant ||\mathbf{v}|| \cdot ||\mathbf{w}||$$
Dimostriamo la disuguaglianza
$$|\langle \mathbf{v}, \mathbf{w}\rangle| \leqslant ||\mathbf{v}|| \cdot ||\mathbf{w}|| \tag{6.1}$$
che implica la precedente. Fissati $\mathbf{v}$ e $\mathbf{w}$, consideriamo la funzione $P(t) = ||t\mathbf{v} + \mathbf{w}||^2$ di variabile reale t. Essendo un quadrato, $P(t) \geqslant 0$ per ogni t. Abbiamo
$$P(t) = ||\mathbf{v}||^2 t^2 + 2t\langle \mathbf{v}, \mathbf{w}\rangle + ||\mathbf{w}||^2$$
cioè $P(t)$ è un polinomio di grado due in t. Dato che $P(t) \geqslant 0$, il discriminante Δ di $P(t)$ soddisfa $\Delta \leqslant 0$: in caso contrario, $P(t)$ avrebbe due radici reali distinte ed assumerebbe quindi valori negativi. Ora,
$$0 \leqslant \frac{\Delta}{4} = \langle \mathbf{v}, \mathbf{w}\rangle^2 - ||\mathbf{v}||^2||\mathbf{w}||^2$$
il che dimostra la (6.1). □

Definizione. Sia V uno spazio euclideo. Un vettore $\mathbf{v} \in V$ si dice **normale** se $||\mathbf{v}|| = 1$.

Si osservi che, se $\mathbf{v} \neq \mathbf{0}$, allora $\mathbf{v}/||\mathbf{v}||$ è un vettore normale, detto il **normalizzato** di $\mathbf{v}$.

6.2 Ortogonalità

Definizione. Sia V uno spazio euclideo con prodotto scalare $\langle\ ,\ \rangle$. Dati due vettori $\mathbf{v}, \mathbf{w}$ di V diciamo che $\mathbf{v}$ è **ortogonale** a $\mathbf{w}$ se $\langle \mathbf{v}, \mathbf{w}\rangle = 0$. Scriveremo $\mathbf{v} \perp \mathbf{w}$.

Ovviamente, $\mathbf{v} \perp \mathbf{w}$ se e solo se $\mathbf{w} \perp \mathbf{v}$. Si noti che l'unico vettore ortogonale a se stesso è il vettore nullo.

Nello spazio euclideo $\mathbf{R}^3$ standard, la nozione di ortogonalità è quella standard della geometria euclidea. Infatti $\mathbf{x} \perp \mathbf{y}$ se e solo se $\cos\alpha = 0$, cioè $\alpha = \pi/2$, dove α è l'angolo interno formato da $\mathbf{x}$ ed $\mathbf{y}$.

Definizione. Un insieme di vettori $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ di V è detto **ortogonale** se $\langle \mathbf{v}_i, \mathbf{v}_j\rangle = 0$ per ogni $i \neq j$.

Un insieme di vettori $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ di V è detto **ortonormale** se è $\langle \mathbf{v}_i, \mathbf{v}_j\rangle = \delta_{ij}$. Cioè i vettori sono normali e ortogonali fra loro. In particolare, i vettori di un insieme ortonormale sono tutti non nulli.

Proposizione 6.2

Sia $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ un insieme ortogonale in cui $\mathbf{v}_i \neq \mathbf{0}$ per ogni i. Allora i vettori $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ sono linearmente indipendenti (questo in particolare è vero se l'insieme è ortonormale).

Dimostrazione. Supponiamo che $\sum_{j=1}^{m} a_j \mathbf{v}_j = \mathbf{o}$, allora, fissato i, avremo

$$0 = \langle \mathbf{v}_i, \sum_j a_j \mathbf{v}_j \rangle = a_i \|\mathbf{v}_i\|^2$$

essendo $\mathbf{v}_i \neq \mathbf{o}$, la sua norma è strettamente positiva; segue che $a_i = 0$. □

Proposizione 6.3

Sia $\mathsf{E} = \{\mathbf{e}_1, \dots \mathbf{e}_n\}$ *una base ortonormale dello spazio euclideo* V*; sia* $\mathbf{v} = \sum_i a_i \mathbf{e}_i$ *un vettore di* V*. Allora, per ogni i, abbiamo*

$$a_i = \langle \mathbf{v}, \mathbf{e}_i \rangle$$

Notazione. Scriveremo spesso *b.o.n.* anziché base ortonormale.

Proposizione 6.4

Sia $\mathsf{E} = \{\mathbf{e}_1, \dots \mathbf{e}_n\}$ *una base ortonormale dello spazio euclideo* V*. Dati due vettori* $\mathbf{v} = \sum_i a_i \mathbf{e}_i$*,* $\mathbf{w} = \sum_i b_i \mathbf{e}_i$*; allora*

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{n} a_i b_i$$

Dimostrazione. La dimostrazione segue subito dalla bilinearità del prodotto scalare, ricordando che $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$. □

Osservazione 6.5. Dato uno spazio euclideo V con prodotto scalare $\langle\, , \rangle$, se $\mathsf{E} = \{\mathbf{e}_1, \dots \mathbf{e}_n\}$ è una base ortonormale allora il prodotto scalare dato coincide col prodotto scalare standard $\langle\, , \rangle_E$ rispetto alla base E.

Teorema 6.6 (Gram-Schmidt)

Sia V *uno spazio euclideo di dimensione finita, allora esiste una base ortonormale.*

Dimostrazione. Sia $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base di V: vogliamo costruire una base ortonormale $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$. Sia $\mathbf{e}_1 = \mathbf{v}_1 / \|\mathbf{v}_1\|$. Il vettore $\mathbf{e}_2$ deve essere ortogonale ad $\mathbf{e}_1$, e lo cerchiamo tra i vettori in $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2)$. Precisamente, per ogni $\alpha \in \mathbf{R}$, avremo

$$\mathbf{v}_2 = \alpha \mathbf{e}_1 + (\mathbf{v}_2 - \alpha \mathbf{e}_1)$$

cerchiamo α in modo tale che $\mathbf{v}_2 - \alpha \mathbf{e}_1$ sia ortogonale ad $\mathbf{e}_1$. Imponiamo quindi che $\langle \mathbf{v}_2 - \alpha \mathbf{e}_1, \mathbf{e}_1 \rangle = 0$ cioè deve essere

$$\langle \mathbf{v}_2, \mathbf{e}_1 \rangle = \alpha \langle \mathbf{e}_1, \mathbf{e}_1 \rangle = \alpha$$

Quindi α è univocamente determinato, e il vettore $\mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{e}_1 \rangle \mathbf{e}_1$ è ortogonale a $\mathbf{e}_1$. Prendiamo quindi come secondo vettore $\mathbf{e}_2$ il suo normalizzato:

$$\mathbf{e}_2 = \frac{\mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{e}_1 \rangle \mathbf{e}_1}{\|\mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{e}_1 \rangle \mathbf{e}_1\|}$$

I vettori $\mathbf{e}_1, \mathbf{e}_2$ sono quindi una b.o.n. di $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2)$. Analogamente, se cerchiamo un vettore della forma $\mathbf{v}_3 - \alpha_1\mathbf{e}_1 - \alpha_2\mathbf{e}_2$ che sia ortogonale sia ad $\mathbf{e}_1$ che ad $\mathbf{e}_2$ troveremo

$$\alpha_1 = \langle \mathbf{v}_3, \mathbf{e}_1 \rangle, \quad \alpha_2 = \langle \mathbf{v}_3, \mathbf{e}_2 \rangle$$

Prendiamo quindi come vettore $\mathbf{e}_3$ il normalizzato del vettore così determinato:

$$\mathbf{e}_3 = \frac{\mathbf{v}_2 - \langle \mathbf{v}_3, \mathbf{e}_1 \rangle \mathbf{e}_1 - \langle \mathbf{v}_3, \mathbf{e}_2 \rangle \mathbf{e}_2}{\|\mathbf{v}_2 - \langle \mathbf{v}_3, \mathbf{e}_1 \rangle \mathbf{e}_1 - \langle \mathbf{v}_3, \mathbf{e}_2 \rangle \mathbf{e}_2\|}$$

Procedendo in maniera analoga, determiniamo un insieme $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ di vettori ortonormali, che quindi costituiscono una base ortonormale di V. □

Osservazione 6.7. Il processo utilizzato nel teorema precedente per passare da una base $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ ad una b.o.n. $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ è detto **processo di ortonormalizzazione di Gram-Schmidt**.

Si osservi che, in tale processo, i vettori $\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ appartengono allo spazio vettoriale $\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ e ne costituiscono una b.o.n.

Corollario 6.8

Sia V uno spazio euclideo di dimensione n. Allora un insieme ortonormale $\{\mathbf{e}_1, \ldots, \mathbf{e}_r\}$ può essere completato ad una b.o.n. $\{\mathbf{e}_1, \ldots, \mathbf{e}_r, \ldots, \mathbf{e}_n\}$

Dimostrazione. I vettori $\{\mathbf{e}_1, \ldots, \mathbf{e}_r\}$ sono linearmente indipendenti per la proposizione 6.2. Possiamo completare l'insieme $\{\mathbf{e}_1, \ldots, \mathbf{e}_r\}$ ad una base $\{\mathbf{e}_1, \ldots, \mathbf{e}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ Applicando il processo di ortonormalizzazione di Gram-Schmidt otteniamo una b.o.n.; durante tale processo i primi r vettori non vengono modificati, come si verifica immediatamente, in quanto già ortonormali. □

6.3 Endomorfismi ortogonali

Sia V uno spazio euclideo di dimensione n ed f un endomorfismo di V.

Definizione. Diciamo che l'endomorfismo f

1. conserva il prodotto scalare se $\langle \mathbf{v}, \mathbf{w} \rangle = \langle f(\mathbf{v}), f(\mathbf{w}) \rangle$ per ogni $\mathbf{v}, \mathbf{w} \in V$;
2. conserva la norma se $\|\mathbf{v}\| = \|f(\mathbf{v})\|$ per ogni $\mathbf{v} \in V$.

Teorema 6.9

Sia V uno spazio euclideo di dimensione n ed f un endomorfismo di V. Allora le condizioni seguenti sono equivalenti

a) f conserva il prodotto scalare
b) f conserva la norma
c) se $\{\mathbf{e}_1, \ldots, \mathbf{e}_r\}$ è un insieme ortonormale, anche $\{f(\mathbf{e}_1), \ldots, f(\mathbf{e}_r)\}$ è ortonormale
d) se $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ è una b.o.n., allora $\{f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)\}$ è una b.o.n.

e) esiste una base ortonormale $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ *tale che* $\{f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)\}$ *è una b.o.n.*

Dimostrazione. Dimostriamo dapprima che a) e b) sono equivalenti:
a) implica b) perché

$$\|f(\mathbf{v})\|^2 = \langle f(\mathbf{v}), f(\mathbf{v})\rangle = \langle \mathbf{v}, \mathbf{v}\rangle = \|\mathbf{v}\|^2$$

b) implica a) perché dall'uguaglianza $\|f(\mathbf{v}) + f(\mathbf{w})\|^2 = \|\mathbf{v} + \mathbf{w}\|^2$ otteniamo

$$\|f(\mathbf{v})\|^2 + \|f(\mathbf{w})\|^2 + 2\langle f(\mathbf{v}), f(\mathbf{w})\rangle = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\langle \mathbf{v}, \mathbf{w}\rangle$$

da cui $\langle f(\mathbf{v}), f(\mathbf{w})\rangle = \langle \mathbf{v}, \mathbf{w}\rangle$.
Dimostriamo ora le seguenti implicazione:

$$\text{a)} \Rightarrow \text{c)} \Rightarrow \text{d)} \Rightarrow \text{e)} \Rightarrow \text{a)}$$

a) implica c) perché

$$\langle f(\mathbf{e}_i), f(\mathbf{e}_j)\rangle = \langle \mathbf{e}_i, \mathbf{e}_j\rangle = \delta_{ij}$$

c) implica d) banalmente
d) implica e) perché il teorema di Gram-Schmidt ci assicura l'esistenza di una b.o.n.
Rimane da mostrare che e) implica a). Sia $\mathsf{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ una base ortonormale tale che $\mathsf{F} = \{f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)\}$ sia una b.o.n. Dati due vettori $\mathbf{v} = \sum_i a_i\mathbf{e}_i, \mathbf{w} = \sum_i b_i\mathbf{e}_i$ avremo, essendo E una b.o.n.,

$$\langle \mathbf{v}, \mathbf{w}\rangle = \sum_i a_i b_i$$

D'altra parte, essendo F una b.o.n,

$$f(\mathbf{v}) = \sum_i a_i f(\mathbf{e}_1), \quad f(\mathbf{w}) = \sum_i b_i f(\mathbf{e}_1) \quad \Longrightarrow \quad \langle f(\mathbf{v}), f(\mathbf{w})\rangle = \sum_i a_i b_i$$

Cioè f conserva il prodotto scalare. □

Definizione. Un endomorfismo che soddisfa le condizioni equivalenti del teorema è detto anche **isometria** oppure **endomorfismo ortogonale**.

Proposizione 6.10

Sia V uno spazio euclideo di dimensione n e sia f un'isometria di V. Allora f è un isomorfismo.

Dimostrazione. Infatti f è un endomorfismo di uno spazio vettoriale di dimensione finita il cui nucleo è il solo vettore $\mathbf{o}$: se $\mathbf{v} \neq \mathbf{o}$, allora $\|\mathbf{v}\| \neq 0$ e quindi $\|f(\mathbf{v})\| = \|\mathbf{v}\| \neq 0$. □

Proposizione 6.11

Sia V uno spazio euclideo di dimensione n ed f un'isometria di V. Se λ è un autovalore reale di f, allora $\lambda = \pm 1$.

Dimostrazione. Sia $\mathbf{v}$ un autovettore relativo a λ, cioè $f(\mathbf{v}) = \lambda\mathbf{v}$. Sappiamo che $||f(\mathbf{v})|| = ||\mathbf{v}||$ cioè

$$|\lambda| \cdot ||\mathbf{v}|| = ||\mathbf{v}||$$

□

Osservazione 6.12. In generale, un'isometria f non ha autovalori reali. Per esempio, nello spazio euclideo $\mathbf{R}^2$ standard, una rotazione di un angolo θ è rappresentata da una matrice

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

che non ha autovalori reali a meno che $\theta = k\pi$ con $k \in \mathbf{Z}$.

Matrici ortogonali

Definizione. Sia P una matrice in $\mathrm{M}_n(\mathbf{R})$. Diciamo che P è **ortogonale** se $P\,{}^tP = I_n$. In particolare, una matrice ortogonale è invertibile e le matrici ortogonali sono esattamente le matrici P tali che $P^{-1} = {}^tP$. L'uguaglianza $P\,{}^tP = I_n$ è equivalente a ${}^tP P = I_n$.
L'insieme delle matrici ortogonali in $\mathrm{M}_n(\mathbf{R})$ è indicato con $\mathrm{O}(n)$, ed è un sottogruppo di $\mathrm{GL}_n(\mathbf{R})$.

Infatti $I_n \in \mathrm{O}(n)$; se $P, Q \in \mathrm{O}(n)$, allora anche $PQ \in \mathrm{O}(n)$ perché

$$(PQ)\,{}^t(PQ) = PQ\,{}^tQ\,{}^tP = PI_n\,{}^tP = P\,{}^tP = I_n$$

Inoltre, se $P \in \mathrm{O}(n)$, anche $P^{-1} \in \mathrm{O}(n)$ perché

$${}^t(P^{-1}) = {}^t({}^tP) = P \quad \text{quindi} \quad P^{-1}\,{}^t(P^{-1}) = P^{-1}P = I_n$$

Lasciamo al lettore la facile verifica che, data $P \in \mathrm{GL}_n(\mathbf{R})$, allora $P \in \mathrm{O}(n)$ se e solo se $P^{-1} \in \mathrm{O}(n)$ se e solo se ${}^tP \in \mathrm{O}(n)$.

Osservazione 6.13. Date due matrici $A, B \in \mathrm{M}_n(\mathbf{R})$ indichiamo con $R_1, \dots, R_n$ le righe di A e con $C_1, \dots, C_n$ le colonne di B. Allora l'elemento di posto (i, j) della matrice prodotto AB non è altro che il prodotto scalare $\langle R_i, C_j \rangle$ dove R_i e C_j sono pensati come vettori dello spazio euclideo $\mathbf{R}^n$ standard.

Proposizione 6.14

Sia P una matrice in $\mathrm{M}_n(\mathbf{R})$. Allora $P \in \mathrm{O}(n)$ se e solo se le righe di P formano una base ortonormale dello spazio euclideo $\mathbf{R}^n$ standard se e solo se le colonne di P formano una base ortonormale dello stesso spazio.

Dimostrazione. Se $R_1, \dots, R_n$ sono le righe della matrice P, allora $R_1, \dots, R_n$ sono le colonne della matrice ${}^tP$, quindi l'elemento di posto (i, j) della matrice $P\,{}^tP$ è esattamente $\langle R_i, R_j \rangle$ per l'osservazione precedente. Segue che $P\,{}^tP = I_n$ se e solo se $\langle R_i, R_j \rangle = \delta_{ij}$, cioè $P \in \mathrm{O}(n)$ se e solo se $\{R_1, \dots, R_n\}$ è una b.o.n. Passando a ${}^tP$ si ottiene l'affermazione sulle colonne. □

Proposizione 6.15
Sia $P \in \mathrm{O}(n)$ una matrice ortogonale. Allora $\det(P) = \pm 1$.

Dimostrazione. Dalla $P\,{}^tP = I_n$ otteniamo, per il teorema di Binet,

$$1 = \det(P)\det({}^tP) = \det(P)^2$$

□

Il gruppo $\mathrm{SO}(n)$

L'insieme delle matrici $\{P \in \mathrm{O}(n) : \det(P) = 1\}$ formano un sottogruppo di $\mathrm{O}(n)$ denotato come $\mathrm{SO}(n)$ e detto **gruppo ortogonale speciale**.

Proposizione 6.16
Sia V uno spazio euclideo di dimensione n. Siano $\mathsf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ una b.o.n. ed $\mathsf{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ un insieme di n vettori. Se P è la matrice le cui colonne sono le componenti di $\mathbf{f}_1, \dots, \mathbf{f}_n$ nella base E, allora $P \in \mathrm{O}(n)$ se e solo se F è una b.o.n.

Dimostrazione. Sia $P = (p_{ij})$. Quindi $\mathbf{f}_k = \sum_j p_{jk}\mathbf{e}_j$. Scriviamo anche $P = (\mathbf{f}_1, \dots, \mathbf{f}_n)$ dove consideriamo $\mathbf{f}_j$ come la j colonna di P. Quindi

$${}^tP = \begin{pmatrix} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_n \end{pmatrix}$$

dove $\mathbf{f}_i$ è l'i-esima riga di ${}^tP$. L'elemento di posto (i, j) di ${}^tPP$ è $p_{1i}p_{1j} + \cdots + p_{ni}p_{nj}$ che coincide con $\langle \mathbf{f}_i, \mathbf{f}_j \rangle$ perché E è una base ortonormale. Segue che ${}^tPP = I_n$ se e solo se $\langle \mathbf{f}_i, \mathbf{f}_j \rangle = \delta_{ij}$, cioè $P \in \mathrm{O}(n)$ se e solo se F è una b.o.n. □

6.4 Spazi unitari

6.5 Endomorfismi unitari

6.6 Endomorfismi autoaggiunti

Indice analitico