

• **La definizione formale di spazio vettoriale.**

Uno spazio vettoriale su un campo K è un insieme V su cui sono definite due operazioni: la somma fra due elementi di V , il cui risultato è ancora un elemento di V , e il prodotto di un elemento del campo K (ossia di uno “scalare”) per un elemento di V , il cui risultato è un elemento di V . Tali operazioni soddisfano le seguenti proprietà:

- $\forall u, v, w \in V$ vale $(u+v)+w = u+(v+w)$ (proprietà associativa dell’addizione).
- $\forall v, w \in V$ vale $v+w = w+v$ (proprietà commutativa dell’addizione).
- esiste $O \in V$ tale che $\forall v \in V$ vale $v+O = v$ (O è l’elemento identico per l’addizione. ATTENTI! L’elemento O appartiene a V , non va confuso con lo scalare 0 che appartiene al campo K).
- $\forall v \in V$ esiste un elemento “ $-v$ ” in V tale che $v+(-v) = 0$ (esistenza dell’opposto per l’addizione).
- $\forall \lambda, \mu \in K, \forall v, w \in V$ vale $\lambda(v+w) = \lambda v + \lambda w$ e anche $(\lambda+\mu)v = \lambda v + \mu v$ (proprietà distributive della moltiplicazione per scalare).
- $\forall \lambda, \mu \in K, \forall v \in V$ vale $(\lambda\mu)v = \lambda(\mu v)$ (proprietà associativa della moltiplicazione per scalare).
- $\forall v \in V$ vale $1v = v$ e $0v = O$ (proprietà di 0 e 1 rispetto alla moltiplicazione).

Di solito chiameremo “vettori” gli elementi di uno spazio vettoriale V , mentre, come si è visto, gli elementi del campo K saranno gli “scalari”.

L’esempio che studieremo di più è quello di \mathbb{R}^n . Si tratta, come insieme, del prodotto cartesiano $\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$ (dove \mathbb{R} compare n volte); per vederlo come spazio vettoriale sul campo \mathbb{R} dei numeri reali usiamo le seguenti notazioni:

(1) Gli elementi, ossia i vettori, sono le liste ordinate formate da n numeri reali:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \\ \dots \\ a_{n-1} \\ a_n \end{pmatrix}$$

(2) La somma fra vettori è definita da:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \\ \dots \\ a_{n-1} \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \dots \\ \dots \\ b_{n-1} \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \\ \dots \\ \dots \\ a_{n-1} + b_{n-1} \\ a_n + b_n \end{pmatrix}$$

(3) Il prodotto per scalari è definito da:

$$\lambda \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \\ \dots \\ a_{n-1} \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \lambda a_3 \\ \dots \\ \dots \\ \lambda a_{n-1} \\ \lambda a_n \end{pmatrix}$$

Consideriamo ora uno spazio vettoriale V sul campo K .

Definizione 0.1. Un SOTTOSPAZIO vettoriale di V è un sottoinsieme $N \subseteq V$ che soddisfa le seguenti tre proprietà:

- (1) Per ogni $u, v \in N$ vale $u + v \in N$ (in altre parole, N è chiuso rispetto alla somma).
- (2) Per ogni $k \in K$ e per ogni $u \in N$ vale $ku \in N$ (N è chiuso rispetto alla moltiplicazione per scalare).
- (3) Il vettore O appartiene a N .

L'ultima proprietà in realtà è ridondante, perché si ricava dalle altre, ma abbiamo voluto evidenziarla lo stesso: se vi capita di studiare un insieme che non contiene lo O , potete subito concludere che non è un sottospazio vettoriale !

Esempio 0.2. La circonferenza $x^2 + y^2 = 1$ in \mathbb{R}^2 non è un sottospazio vettoriale. Tutti e soli i sottospazi vettoriali di \mathbb{R}^2 sono \mathbb{R}^2 , $\{O\}$ e tutte le rette passanti per O (dimostratelo per esercizio).

Esempio 0.3. L'anello dei polinomi $K[x]$ può essere visto come uno spazio vettoriale su K (basta "dimenticarsi" che esiste il prodotto fra polinomi, ci serve solo la somma

e il prodotto per le costanti $k \in K$). Consideriamo il sottoinsieme dato dai polinomi che hanno 1 come radice:

$$L = \{p(x) \in K[x] \mid p(1) = 0\}$$

e verifichiamo che è un sottospazio vettoriale.

- (1) Se $p, q \in L$ allora $p + q$ appartiene a L , infatti $p(1) + q(1) = 0 + 0 = 0$.
- (2) Se $p \in L$ e $k \in K$ allora $kp \in L$, infatti $kp(1) = k0 = 0$.
- (3) Il polinomio 0, che è lo O di $K[x]$, appartiene a L , infatti ha 1 come radice (addirittura ogni elemento di K è una radice di 0).

• **Qualche osservazione sulle basi di spazi vettoriali.**

Sia V uno spazio vettoriale sul campo K .

Definizione 0.4. Si dice che un insieme di vettori $\{v_1, v_2, \dots, v_k\}$ è un “insieme di generatori” di V se ogni elemento $v \in V$ può essere espresso come combinazione lineare dei vettori v_1, v_2, \dots, v_k , ossia se per ogni $v \in V$ si trovano degli scalari a_1, a_2, \dots, a_k tali che

$$a_1v_1 + a_2v_2 + \dots + a_kv_k = v$$

Talvolta si dice anche, più brevemente, che i vettori v_1, v_2, \dots, v_k “generano” V .

Se abbiamo un insieme di generatori, potrebbe essere possibile trovare più di una combinazione lineare che esprime lo stesso vettore $v \in V$: l’unicità non è garantita.

Esempio 0.5. Si verifica (esercizio) che i vettori

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix}$$

generano \mathbb{R}^3 . Si possono facilmente trovare due distinte combinazioni lineari di tali vettori che esprimono il vettore

$$\begin{pmatrix} 2 \\ 2 \\ 5 \end{pmatrix}$$

Per esempio:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Ci chiediamo allora quale ulteriore condizione bisogna imporre ad un insieme di generatori per poter essere sicuri che ogni elemento dello spazio venga espresso in maniera unica come combinazione lineare.

Il concetto chiave è quello di “indipendenza lineare”:

Definizione 0.6. Si dice che un insieme di vettori $\{v_1, v_2, \dots, v_r\}$ è un “insieme di vettori linearmente indipendenti” se l’unico modo di scrivere O come combinazione lineare di questi vettori è quando tutti i coefficienti sono nulli, ossia se

$$a_1v_1 + a_2v_2 + \dots + a_rv_r = O \iff a_1 = a_2 = \dots = a_r = 0$$

Talvolta si dice anche, più brevemente, che i vettori v_1, v_2, \dots, v_r sono “linearmente indipendenti”.

Se invece i vettori v_1, v_2, \dots, v_r non sono linearmente indipendenti, si dice che sono “linearmente dipendenti” (o che l’insieme $\{v_1, v_2, \dots, v_r\}$ è un “insieme di vettori linearmente dipendenti”).

Sottolineiamo che la caratteristica di essere “linearmente indipendenti” riguarda tutto l’insieme di vettori che stiamo considerando: se per esempio abbiamo un insieme di 5 vettori che non sono linearmente indipendenti, potrebbe accadere che, levandone uno, l’insieme residuo sia costituito da 4 vettori che stavolta sono linearmente indipendenti.

Esempio 0.7. Verificate che i vettori di \mathbb{R}^3 dell’esempio precedente

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix}$$

non sono linearmente indipendenti.

Però se considerate solo i primi tre, o gli ultimi tre, allora avete degli insiemi di vettori linearmente indipendenti (esercizio). Ma è sempre vero che levandone uno avete garantita la lineare indipendenza? (NO, provate a levare il terzo vettore....)

Definizione 0.8. Si dice che un insieme di vettori $\{v_1, v_2, \dots, v_n\}$ che generano lo spazio e sono anche linearmente indipendenti, è una “base” di V .

Fissiamo uno spazio vettoriale V (sul campo K) e supponiamo che ammetta una base finita $\{v_1, v_2, \dots, v_n\}$.

Proposizione 0.9. *Ogni vettore $v \in V$ si scrive IN MODO UNICO come combinazione lineare degli elementi della base.*

Dimostrazione. Il vettore v si può scrivere come combinazione lineare degli elementi della base perché gli elementi della base generano V . L’unicità di una tale combinazione lineare è conseguenza della lineare indipendenza degli elementi della base. Infatti, supponiamo che si possa scrivere:

$$v = a_1v_1 + a_2v_2 + \dots + a_nv_n = b_1v_1 + b_2v_2 + \dots + b_nv_n$$

dove gli a_i e b_j sono elementi del campo K . Sottraendo abbiamo:

$$O = (a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \dots + (a_n - b_n)v_n$$

Ma sappiamo che i vettori v_1, v_2, \dots, v_n sono linearmente indipendenti. Dunque la combinazione lineare che abbiamo scritto sopra, e che ha come risultato O , deve avere tutti i coefficienti nulli. Così possiamo concludere che $a_i = b_i$ per ogni i , ossia che

esiste un solo modo di scrivere v come combinazione lineare degli elementi della base data. ■

Come annunciato, in questo corso considereremo quasi sempre spazi vettoriali che ammettono una base finita. Abbiamo detto in classe che risulterà che tutte le basi di uno spazio che ammette una base finita hanno la stessa cardinalità e che tale cardinalità si chiama la “dimensione” dello spazio vettoriale; discuteremo più avanti nel corso questa affermazione. Per il momento ci limitiamo ad osservare, mediante il seguente teorema, che se uno spazio vettoriale ammette un insieme finito di generatori, allora ammette anche una base finita. In altre parole i generatori dati potrebbero essere molti, e sovrabbondanti, ma è sempre possibile estrarre dall’insieme dei generatori un sottoinsieme che è una base.

Teorema 0.10. *Sia V uno spazio vettoriale (sul campo K) e supponiamo che V sia generato da un insieme finito di vettori $\{w_1, w_2, \dots, w_s\}$. Allora è possibile estrarre da $\{w_1, w_2, \dots, w_s\}$ un sottoinsieme $\{w_{i_1}, w_{i_2}, \dots, w_{i_n}\}$ che è una base di V .*

Dimostrazione. Consideriamo l’insieme:

$$\mathcal{M} = \{A \subset \{w_1, w_2, \dots, w_s\} \mid A \text{ è un insieme di vettori lin. indep.}\}$$

e notiamo che \mathcal{M} non è vuoto, in quanto contiene certamente i sottoinsiemi di $\{w_1, w_2, \dots, w_s\}$ di cardinalità 1, tipo $\{w_1\}$ o $\{w_2\}$. Fra tutti gli elementi di \mathcal{M} , ossia fra tutti i sottoinsiemi che appartengono a \mathcal{M} , ne prendo uno di cardinalità massima: $\{w_{i_1}, w_{i_2}, \dots, w_{i_n}\}$; per la osservazione precedente sarà $n \geq 1$ ossia tale insieme sarà non vuoto.

Questo $\{w_{i_1}, w_{i_2}, \dots, w_{i_n}\}$ è proprio il nostro candidato ad essere una base di V .

Parte bene perché per come lo abbiamo costruito è un insieme di vettori linearmente indipendenti. Resta da dimostrare che genera V . Per questo basterà mostrare che con combinazioni lineari dei vettori di $\{w_{i_1}, w_{i_2}, \dots, w_{i_n}\}$ posso ottenere uno qualunque dei vettori di $\{w_1, w_2, \dots, w_s\}$, visto che sappiamo che questi generano V (verificare di aver capito bene questo passaggio!).

Se $\{w_{i_1}, w_{i_2}, \dots, w_{i_n}\} = \{w_1, w_2, \dots, w_s\}$ abbiamo già finito. Se invece

$$\{w_{i_1}, w_{i_2}, \dots, w_{i_n}\} \subsetneq \{w_1, w_2, \dots, w_s\}$$

allora prendiamo un vettore, diciamo w_r , che non appartiene a $\{w_{i_1}, w_{i_2}, \dots, w_{i_n}\}$. Dobbiamo dimostrare che w_r si può scrivere come combinazione lineare dei vettori $\{w_{i_1}, w_{i_2}, \dots, w_{i_n}\}$.

Se consideriamo l’insieme $\{w_r, w_{i_1}, w_{i_2}, \dots, w_{i_n}\}$ notiamo che certamente questo non è un insieme di vettori linearmente indipendenti, se no apparterebbe a \mathcal{M} e non sarebbe più vero che $\{w_{i_1}, w_{i_2}, \dots, w_{i_n}\}$ aveva cardinalità massima fra gli elementi di \mathcal{M} .

Dunque esiste una combinazione lineare:

$$a_r w_r + a_{i_1} w_{i_1} + a_{i_2} w_{i_2} + \dots + a_{i_n} w_{i_n} = 0$$

che è non banale, ossia i coefficienti non sono tutti zero. In particolare risulta che non può essere $a_r = 0$, altrimenti resterebbe una combinazione lineare non banale:

$$a_{i_1}w_{i_1} + a_{i_2}w_{i_2} + \cdots + a_{i_n}w_{i_n} = 0$$

che contraddirebbe la lineare indipendenza di $\{w_{i_1}, w_{i_2}, \dots, w_{i_n}\}$.

Visto dunque che $a_r \neq 0$, si può dividere tutto per a_r ottenendo:

$$w_r = -\frac{a_{i_1}}{a_r}w_{i_1} - \frac{a_{i_2}}{a_r}w_{i_2} - \cdots - \frac{a_{i_n}}{a_r}w_{i_n}$$

che è la combinazione lineare cercata. ■

Esempio 0.11. Potete facilmente dimostrare che l'anello dei polinomi $K[x]$ non è uno spazio vettoriale di dimensione finita, ossia non può avere una base finita. Ma ammette una base infinita? Per rispondere osserviamo per prima cosa che le definizioni di insieme di generatori e insieme di vettori linearmente indipendenti (e dunque anche quella di base) vanno adattate al caso di insiemi infiniti (si "aggiustano" chiedendo che le combinazioni lineari che appaiono nelle formule siano comunque composte da un numero finito di addendi); dopodiché potete verificare che l'insieme $\{1, x, x^2, x^3, \dots, x^n, \dots\}$ è una base.

Invece se ci limitiamo a considerare l'insieme $K[x]^{\leq n}$ costituito dai polinomi di grado minore o uguale ad un certo intero $n \geq 0$ fissato (e uniamo anche il polinomio 0 che per noi non ha grado), allora abbiamo uno spazio vettoriale, sottospazio del precedente, che ha dimensione $n + 1$ (verificare esibendo una base).

• **Applicazioni lineari e matrici.**

Consideriamo la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definita da

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \\ x^2 \end{pmatrix}$$

Si osserva subito che la f manda la retta di equazione $x = y$, ossia la retta data dai punti di coordinate $\begin{pmatrix} x \\ x \end{pmatrix}$, nella parabola di equazione $y = x^2$. Ma la retta è un sottospazio di \mathbb{R}^2 , mentre la parabola no..

Ci piacerebbe lavorare con funzioni che mandano sottospazi in sottospazi, ossia che "rispettano" le strutture vettoriali che abbiamo introdotto. Gli oggetti giusti da prendere in considerazione sono le applicazioni lineari:

Definizione 0.12. Siano V e W spazi vettoriali sul campo K . Una applicazione lineare L da V a W è una funzione

$$L : V \rightarrow W$$

che soddisfa le seguenti due proprietà:

- (1) per ogni $v_1, v_2 \in V$ vale $L(v_1 + v_2) = L(v_1) + L(v_2)$;
- (2) per ogni $\lambda \in K$ e per ogni $v \in V$ vale $L(\lambda v) = \lambda L(v)$.

Osservazione 0.13. Le due proprietà della definizione possono essere espresse in maniera equivalente dalla seguente richiesta: per ogni $v_1, v_2 \in V$ e per ogni $a, b \in K$ vale

$$L(av_1 + bv_2) = aL(v_1) + bL(v_2)$$

Consideriamo ora una applicazione lineare:

$$L : V \rightarrow W$$

e prendiamo in V una base $\{e_1, e_2, \dots, e_n\}$ e in W una base $\{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}$. Dato un elemento $v \in V$, proviamo a scrivere la sua immagine $L(v)$. Sappiamo che v si può scrivere in modo unico come combinazione lineare degli elementi della base scelta:

$$v = a_1e_1 + a_2e_2 + \dots + a_n e_n$$

Per la linearità di L allora:

$$L(v) = a_1L(e_1) + a_2L(e_2) + \dots + a_nL(e_n)$$

Dunque, per conoscere L , ossia per saper dire qual è l'immagine di un qualsiasi elemento $v \in V$, basta conoscere $L(e_1), L(e_2), \dots, L(e_n)$.

Per poter descrivere $L(e_1), L(e_2), \dots, L(e_n)$, che sono vettori di W , possiamo servirci della base $\{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}$: per ogni i , $L(e_i)$ si può scrivere in modo unico come

$$L(e_i) = a_{1i}\epsilon_1 + a_{2i}\epsilon_2 + \dots + a_{mi}\epsilon_m$$

Ecco che entrano in scena le matrici.

Definizione 0.14. La matrice associata all'applicazione lineare L nelle basi $\{e_1, e_2, \dots, e_n\}$ e $\{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}$ è data dalla seguente griglia di m righe per n colonne:

$$[L] \begin{matrix} e_1, e_2, \dots, e_n \\ \epsilon_1, \epsilon_2, \dots, \epsilon_m \end{matrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}$$

che si può anche scrivere sinteticamente come:

$$[L] \begin{matrix} e_1, e_2, \dots, e_n \\ \epsilon_1, \epsilon_2, \dots, \epsilon_m \end{matrix} = (a_{ij}) \quad \begin{matrix} i = 1, 2, 3, \dots, n \\ j = 1, 2, 3, \dots, m \end{matrix}$$

Notiamo che la matrice $[L]$ (da ora in poi, per semplificare la notazione, ometteremo il riferimento alle basi tutte le volte che questo non creerà ambiguità) si ottiene ponendo uno accanto all'altro i vettori $L(e_1), L(e_2), \dots, L(e_n)$, scritti come vettori colonna nella base scelta di W . Dunque guardando la matrice possiamo sapere tutto quello che ci serve sull'applicazione lineare L .

Applichiamo ciò che abbiamo detto fin qui al calcolo di $L(v)$ dove, come prima, $v \in V$ e

$$v = a_1e_1 + a_2e_2 + \dots + a_n e_n$$

Allora

$$L(v) = a_1L(e_1) + a_2L(e_2) + \cdots + a_nL(e_n)$$

ossia

$$L(v) = a_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \dots \\ \dots \\ a_{m1} \end{pmatrix} + a_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \\ \dots \\ \dots \\ a_{m2} \end{pmatrix} + \cdots + a_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ \dots \\ \dots \\ a_{mn} \end{pmatrix}$$

Ribadiamo che i vettori che compaiono nella espressione scritta qui sopra sono i vettori colonna della matrice $[L]$, e che stiamo considerando lo spazio W munito della sua base $\{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}$. Dunque per esempio il vettore colonna

$$\begin{pmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ \dots \\ \dots \\ a_{mn} \end{pmatrix}$$

rappresenta l'elemento

$$L(e_n) = a_{1n}\epsilon_1 + a_{2n}\epsilon_2 + \cdots + a_{mn}\epsilon_m$$

Possiamo allo stesso modo rappresentare i vettori di V come vettori colonna relativamente alla base $\{e_1, e_2, \dots, e_n\}$; in tal modo allora per esempio il nostro v si scrive:

$$v = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \\ \dots \\ a_n \end{pmatrix}$$

Ricordando il prodotto “righe per colonne” di cui abbiamo parlato in classe (e che qui non ridefiniamo), possiamo anche scrivere così:

$$L(v) = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \\ \dots \\ a_n \end{pmatrix}$$

Dunque la matrice $[L]$, mediante il prodotto righe per colonne, ci permette di calcolare come agisce la applicazione L sui vettori di V .

Esempio 0.15. Facciamo ora un esempio che mostra come la matrice associata ad una applicazione lineare dipenda dalle basi scelte. Consideriamo gli spazi vettoriali \mathbb{R}^4 e \mathbb{R}^3 con le loro basi standard, rispettivamente

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, e_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

e

$$\epsilon_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \epsilon_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \epsilon_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Consideriamo poi la applicazione lineare

$$L : \mathbb{R}^4 \rightarrow \mathbb{R}^3$$

così definita (sappiamo, per quanto osservato sopra, che per definire una applicazione lineare basta dare il suo valore sugli elementi di una base):

$$L(e_1) = 2\epsilon_1 + \sqrt{3}\epsilon_2$$

$$L(e_2) = 3\epsilon_1 + \epsilon_2 + \epsilon_3$$

$$L(e_3) = \epsilon_1 + 7\epsilon_2 + 8\epsilon_3$$

$$L(e_4) = 2\epsilon_2 + 4\epsilon_3$$

Come sappiamo, a questa applicazione corrisponde la seguente matrice relativamente alle basi standard:

$$[L]_{\substack{e_1, e_2, e_3, e_4 \\ \epsilon_1, \epsilon_2, \epsilon_3}} = \begin{pmatrix} 2 & 3 & 1 & 0 \\ \sqrt{3} & 1 & 7 & 2 \\ 0 & 1 & 8 & 4 \end{pmatrix}$$

Dunque, preso per esempio il vettore

$$v = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

(scritto rispetto alla base standard) per calcolare $L(v)$ basta fare il prodotto:

$$\begin{pmatrix} 2 & 3 & 1 & 0 \\ \sqrt{3} & 1 & 7 & 2 \\ 0 & 1 & 8 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

che dà come risultato

$$L(v) = \begin{pmatrix} 11 \\ \sqrt{3} + 31 \\ 42 \end{pmatrix}$$

che è un vettore di \mathbb{R}^3 scritto nella base standard.

Supponiamo ora di voler cambiare le basi. Prendiamo allora in \mathbb{R}^4 la nuova base (verificare che si tratta davvero di una base !):

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, v_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

e in \mathbb{R}^3 la nuova base (anche qui verificare !):

$$w_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, w_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, w_3 = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

Proviamo a scrivere la matrice

$$[L] \begin{matrix} v_1, v_2, v_3, v_4 \\ w_1, w_2, w_3 \end{matrix}$$

che rappresenterà la solita applicazione lineare L (ma sarà diversa dalla matrice trovata prima, relativa alle basi standard).

Nella prima colonna della matrice che stiamo per costruire, dovremo mettere il vettore $L(v_1)$ scritto in termini della base $\{w_1, w_2, w_3\}$. Calcoliamolo, facendo in un primo tempo riferimento alle basi standard (d'altra parte la nostra L la abbiamo definita tramite le basi standard, dunque non possiamo far altro che ripartire da quella definizione).

$$L(v_1) = L\left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}\right) = L\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}\right) + L\left(\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 5 \\ \sqrt{3} + 1 \\ 1 \end{pmatrix}$$

Fin qui questo vettore è scritto ancora in termini della base standard di \mathbb{R}^3 . Ora lo esprimiamo in termini della base $\{w_1, w_2, w_3\}$. Si verifica che risulta

$$\begin{pmatrix} 5 \\ \sqrt{3} + 1 \\ 1 \end{pmatrix} = (4 - \sqrt{3}) \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + (\sqrt{3} + 1) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 2 \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

e dunque

$$L(v_1) = (4 - \sqrt{3})w_1 + (\sqrt{3} + 1)w_2 - 2w_3$$

Allora il vettore da inserire come prima colonna della matrice

$$[L] \begin{matrix} v_1, v_2, v_3, v_4 \\ w_1, w_2, w_3 \end{matrix}$$

è

$$\begin{pmatrix} 4 - \sqrt{3} \\ \sqrt{3} + 1 \\ -2 \end{pmatrix}$$

Procedendo allo stesso modo per le altre colonne si ottiene (verificare!):

$$[L] \begin{matrix} v_1, v_2, v_3, v_4 \\ w_1, w_2, w_3 \end{matrix} = \begin{pmatrix} 4 - \sqrt{3} & -4 & -8 & -2 \\ \sqrt{3} + 1 & 8 & 9 & 2 \\ -2 & \frac{5}{2} & \frac{11}{2} & 2 \end{pmatrix}$$

- **Le operazioni elementari di colonna, le matrici a scalini, il rango di una matrice.**

In classe abbiamo descritto un gioco: si prende una matrice $m \times n$, a coefficienti in un campo K , e abbiamo a disposizione tre tipi di mosse sulle colonne. Ecco le mosse:

- si somma alla colonna i la colonna j moltiplicata per uno scalare λ ;
- si moltiplica la colonna s per uno scalare $k \neq 0$;
- si permutano fra di loro due colonne, diciamo la i e la j .

Si è visto che è sempre possibile, usando le mosse descritte sopra, ridurre la matrice in una forma detta “a scalini (per colonne)”. Per intenderci, ecco alcuni esempi di matrici in forma a scalini (sempre per colonne, vedremo presto le matrici in forma a scalini per righe):

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ \sqrt{3} + 1 & 1 & 0 & 0 \\ -2 & \frac{5}{2} & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ \sqrt{3} + 1 & 1 & 0 & 0 \\ -2 & 7 & 0 & 0 \\ \sqrt{3} & 4 & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & \frac{5}{2} & 0 & 0 \end{pmatrix}$$

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

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 \\ -8 & 4 & 1 & 0 & 0 \\ -2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 9 & 1 \end{pmatrix}$$

Per una definizione formale di “matrice a scalini per colonne” abbiamo concordato che si potrebbe seguire questa strada: chiamiamo “profondità” di un vettore la posizione occupata, contata dal basso, del suo più alto coefficiente diverso da zero. Per esempio i vettori

$$\begin{pmatrix} 0 \\ \sqrt{3} + 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 4 - \sqrt{3} \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix}$$

hanno rispettivamente profondità 2,3,1. Allora una matrice a scalini per colonne è una matrice tale che le profondità dei suoi vettori colonna, guardati da sinistra a destra, risultano strettamente decrescenti (negli esempi abbiamo aggiunto anche la condizione che in ogni vettore colonna il coefficiente più alto diverso da zero sia uguale a 1, ma questa richiesta non è essenziale).

Operando poi sempre con le solite operazioni di colonna (in particolare con le operazioni del primo tipo) si può trasformare ancora la matrice ponendola in forma “a scalini ridotta”.

Ecco le forme a scalini ridotte negli esempi appena visti (vedendo gli esempi, date voi una definizione formale di “ridotta”...):

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ \sqrt{3}+1 & 1 & 0 & 0 \\ -2 & \frac{5}{2} & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ \sqrt{3}+1 & 1 & 0 & 0 \\ -2 & 7 & 0 & 0 \\ \sqrt{3} & 4 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2-7-7\sqrt{3} & 7 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & \frac{5}{2} & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & \frac{5}{2} & 0 & 0 \end{pmatrix}$$

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

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 \\ -8 & 4 & 1 & 0 & 0 \\ -2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 9 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Ogni singola mossa sulle colonne corrisponde a moltiplicare la nostra matrice $m \times n$, a destra, per una matrice $n \times n$ invertibile. Per esempio la mossa:

– si somma alla colonna i la colonna j moltiplicata per lo scalare λ ;

corrisponde a moltiplicare la nostra matrice per la matrice $n \times n$ (chiamiamola E_{ij}) che ha tutti 1 sulla diagonale, e 0 in tutte le altre caselle eccetto che nella casella identificata da “riga j , colonna i ”, dove troviamo λ . Verifichiamo che questa matrice è invertibile esibendo la sua inversa; ricordiamo che l’inversa E_{ij}^{-1} deve soddisfare $E_{ij}E_{ij}^{-1} = E_{ij}^{-1}E_{ij} = I$. Possiamo subito trovare E_{ij}^{-1} pensando che ci deve dare la mossa inversa della precedente, che sarebbe

– si sottrae alla colonna i la colonna j moltiplicata per lo scalare $-\lambda$;

Dunque E_{ij}^{-1} è quasi uguale alla E_{ij} : differisce solo per un coefficiente, perché al posto di λ compare $-\lambda$.

Osservazione 0.16. Quanto detto sopra mette in evidenza che la mossa in questione è “reversibile”, ossia, una volta fatta, possiamo fare la sua inversa e tornare esattamente alla matrice di partenza: lo stesso vale per gli altri 2 tipi di mosse (trovate per esercizio le matrici invertibili che le realizzano).

Consideriamo allora una applicazione lineare:

$$L : V \rightarrow W$$

dove V e W sono due spazi vettoriali sul campo K , di dimensione n e m rispettivamente.

Prendiamo in V una base $\{e_1, e_2, \dots, e_n\}$ e in W una base $\{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}$.

A questo punto alla L si può associare una matrice $[L]$, di forma $m \times n$, nelle basi scelte. Ora agiamo sulle colonne di $[L]$ fino a ridurla in forma a scalini ridotta: questo equivale a dire che moltiplichiamo $[L]$ a destra per tante matrici invertibili $[E_1], [E_2], \dots, [E_k]$ fino a che $[L][E_1][E_2] \cdots [E_k]$ è in forma a scalini ridotta.

Per semplificare la notazione, chiamiamo $[E] = [E_1][E_2] \cdots [E_k]$: sappiamo che $[E]$ è invertibile, visto che è il prodotto di matrici invertibili (la sua inversa è $[E]^{-1} = [E_k]^{-1}[E_{k-1}]^{-1} \cdots [E_1]^{-1}$).

Ma la matrice $[E]$ è associata ad una applicazione lineare - che chiameremo E , come c’era da aspettarsi - da V in V : infatti, fissata la base di V , in questo caso $\{e_1, e_2, \dots, e_n\}$, sappiamo che la corrispondenza fra applicazioni lineari da V in V e matrici $n \times n$ a coefficienti in K è bigettiva, ossia data una applicazione lineare costruiamo la matrice, data una matrice troviamo la applicazione lineare da essa rappresentata.

E chi è la applicazione lineare associata a $[L][E]$? È proprio la

$$L \circ E : V \rightarrow W$$

dato che il prodotto fra matrici è stato definito in modo da rispettare la composizione fra applicazioni.

Proposizione 0.17. *Vale che $\text{Imm } L \circ E = \text{Imm } L$, ossia, scritto con un’altra notazione, $L \circ E(V) = L(V)$.*

Dimostriamo questa proposizione dimostrando più in generale che

Proposizione 0.18. *Sia*

$$B : V \rightarrow V$$

una applicazione lineare invertibile. Allora vale $\text{Imm } L = \text{Imm } L \circ B$

Dimostrazione. Dato che B è una funzione invertibile, è bigettiva, ossia $B(V) = V$. Dunque

$$\text{Imm } L \circ B = L(B(V)) = L(V)$$

□

Sappiamo che l’immagine di una applicazione lineare è un sottospazio vettoriale: più esattamente sappiamo che è il sottospazio vettoriale generato dalle immagini degli elementi di una base dello spazio di partenza, ossia, una volta fissate le basi, dai vettori colonna della matrice che rappresenta l’applicazione.

Tradotto nel nostro caso: $\text{Imm } L$ è il sottospazio vettoriale di W generato da $L(e_1), L(e_2), \dots, L(e_n)$, che sono i vettori colonna di $[L]$. Ma la proposizione ci ha

appena garantito che $\text{Imm } L = \text{Imm } L \circ E$, dunque $L(V)$ è anche generato dai vettori colonna di $[L][E]$.

Grazie a queste osservazioni, possiamo a questo punto dimostrare il seguente importante teorema, che era stato annunciato fin dalla terza lezione:

Teorema 0.19. *Se uno spazio vettoriale V ha una base di cardinalità $n \in \mathbb{Z}^+$, allora tutte le altre basi di V hanno la stessa cardinalità.*

Dimostrazione. Prendiamo due basi di V : $\{e_1, e_2, \dots, e_n\}$ e $\{v_1, v_2, \dots, v_r\}$. Dobbiamo dimostrare che $n = r$. Scegliamo per il momento di usare $\{e_1, e_2, \dots, e_n\}$ come base di V ; ogni vettore v_j potrà essere espresso in maniera unica come combinazione lineare dei vettori e_1, e_2, \dots, e_n e possiamo quindi pensarlo come un vettore colonna

$$v_j = \begin{pmatrix} a_{j1} \\ a_{j2} \\ \dots \\ a_{jn} \end{pmatrix}$$

Possiamo formare una matrice M ponendo uno accanto all'altro i vettori v_1, v_2, \dots, v_r : M sarà del tipo $n(\text{righe}) \times r(\text{colonne})$.

Sappiamo che possiamo ridurre a scalini la M con le mosse di colonna ottenendo una nuova matrice M' . Ma che tipo di scalini avrà M' ? Lo spazio generato dai vettori colonna di M' è uguale allo spazio generato dai vettori colonna di M , dunque a V , visto che i vettori colonna di M sono i v_j che sono una base di V per ipotesi. Allora i vettori colonna di M' non possono formare degli scalini “lunghi”: per esempio, se M' fosse

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & \frac{5}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

si vedrebbe subito che il vettore

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

che ha tutti i coefficienti uguali a 0 eccetto un 1 in corrispondenza dello scalino lungo, non viene generato dalle colonne di M' . Il fatto che in M' non ci siano scalini lunghi si può esprimere anche dicendo che la profondità dei vettori colonna deve scendere ad ogni passo di 1 da sinistra a destra, e con gli scalini si deve “toccare il fondo”. Questo è possibile solo se ci sono abbastanza colonne, ossia se $r \geq n$.

Possiamo ripetere tutto il discorso invertendo il ruolo delle basi $\{e_1, e_2, \dots, e_n\}$ e $\{v_1, v_2, \dots, v_r\}$: in tal modo otterremo che deve valere $n \geq r$. Dunque $n = r$ come volevamo dimostrare. \square

Come conseguenza, possiamo “ufficialmente” dare la seguente definizione:

Definizione 0.20. Se uno spazio vettoriale V ha una base di cardinalità $n \in \mathbb{Z}^+$ si dice che V ha dimensione n e si scrive anche $\dim V = n$.

Torniamo ora alla applicazione lineare:

$$L : V \rightarrow W$$

alla sua matrice associata $[L]$ (rispetto alle basi fissate in precedenza). Perché ci piace la matrice $[L][E]$?

Perché ha delle colonne “semplici” con cui lavorare, visto che è in forma a scalini ridotta. Si vede subito che le colonne non nulle di $[L][E]$ formano un insieme di vettori linearmente indipendenti, e dunque che tali colonne, siccome appunto sono linearmente indipendenti e inoltre generano $\text{Imm } L$, sono UNA BASE di $\text{Imm } L$.

Se anche avessimo ridotto la matrice $[L]$ in forma a scalini con altre mosse, e dunque magari avessimo trovato una forma a scalini diversa, rappresentata dalla matrice $[L][E']$, di nuovo lo stesso ragionamento ci porterebbe a dire che le colonne non nulle di $[L][E']$ danno UNA BASE di $\text{Imm } L$.

Poiché tutte le basi hanno la stessa cardinalità, con le osservazioni appena fatte siamo riusciti a dimostrare:

Teorema 0.21. *Data una applicazione lineare L come sopra e fissate le basi, vale che $\dim \text{Imm } L$ è uguale al numero di colonne non nulle che si trovano quando si trasforma $[L]$ in forma a scalini ridotta (o anche solo a scalini, visto che il numero di colonne non nulle è lo stesso).*

Osservazione 0.22. Se avessimo fissato altre basi avremmo avuto una matrice $[L]$ diversa, ma, trasformandola in forma a scalini ridotta, avremmo ancora ovviamente trovato lo stesso numero di colonne non nulle, giacché tale numero è $\dim \text{Imm } L$, ossia è invariante, dipende dalla applicazione (è la dimensione della sua immagine) e non dalle basi scelte.

Vogliamo insistere sul teorema appena dimostrato, perché è importantissimo. Enunciamolo di nuovo in un altro modo.

Fissiamo le basi come sempre; sappiamo che $\text{Imm } L$ è il sottospazio vettoriale di W generato dai vettori colonna di $[L]$. Da questi vettori, come abbiamo visto in una sezione precedente (“Qualche osservazione sulle basi di spazi vettoriali”), è possibile estrarre una base di $\text{Imm } L$.

Dunque, facendo il conto da questo punto di vista, quanto è $\dim \text{Imm } L$? È uguale al numero di elementi di una base di $\text{Imm } L$, ossia (rivedete la costruzione nella sezione) è uguale al MASSIMO NUMERO DI COLONNE LINEARMENTE INDIPENDENTI di $[L]$.

Se consideriamo una applicazione lineare invertibile

$$B : V \rightarrow V$$

vale, per la proposizione precedente, che $\dim \text{Imm } L$, ossia il massimo numero di colonne linearmente indipendenti di $[L]$, è uguale a $\dim \text{Imm } L \circ B$, ossia al massimo numero di colonne linearmente indipendenti di $[L][B]$.

Allora quando facciamo una mossa sulle colonne di una matrice che ha m come numero massimo di colonne linearmente indipendenti, la nuova matrice avrà ancora

m come numero massimo di colonne linearmente indipendenti. In altre parole, fare una mossa di colonne non cambia tale numero.

Enunciamo allora di nuovo il teorema in una forma “più lunga”, introducendo prima il concetto di “rango” di una applicazione lineare:

Definizione 0.23. Data una applicazione lineare $L : V \rightarrow W$, dove due spazi vettoriali V e W di dimensione finita sul campo K , il “rango” di L è il numero $\dim \text{Imm } L$.

Teorema 0.24. . Data una applicazione lineare L come sopra e fissate le basi, vale che $\dim \text{Imm } L$ è uguale al massimo numero di colonne linearmente indipendenti di $[L]$. Visto che fare mosse di colonna non cambia tale numero, vale che $\dim \text{Imm } L$ è uguale anche al massimo numero di colonne linearmente indipendenti che si trovano quando si trasforma $[L]$ in forma a scalini ridotta (o anche solo a scalini). Data la particolare forma delle matrici a scalini (le cui colonne non nulle costituiscono ovviamente un insieme di vettori linearmente indipendenti), $\dim \text{Imm } L$ non è altro che il numero di colonne non nulle che si trovano quando si trasforma $[L]$ in forma a scalini.

- **Quando un insieme di vettori è una base ?**

Le considerazioni riassunte nella sezione precedente ci permettono di discutere una importante applicazione del metodo di riduzione di una matrice in forma a scalini per colonne.

Applicazione. Domanda: dato uno spazio vettoriale V di dimensione n con una base nota e_1, e_2, \dots, e_n , come si fa a stabilire se n vettori v_1, v_2, \dots, v_n costituiscono anch’essi una base di V ?

Risposta: Mettiamo i vettori v_1, v_2, \dots, v_n , che avremo espresso in termini della base e_1, e_2, \dots, e_n , in colonna uno accanto all’altro. Così facendo otteniamo una matrice $[M]$ che è $n \times n$. Ora possiamo ridurre $[M]$ in forma a scalini ridotta $[M']$: se M' è l’identità allora $\{v_1, v_2, \dots, v_n\}$ è una base, altrimenti no. Perché ?

Nella dimostrazione del Teorema 0.19 abbiamo visto che se $\{v_1, v_2, \dots, v_n\}$ è una base allora è necessario che la forma a scalini ridotta $[M']$ abbia scalini tutti corti. Ma le matrici M e M' di cui stiamo parlando sono di forma $n \times n$, quindi $[M']$ è la matrice identità.

Viceversa, se M' è l’identità, le sue n colonne sono linearmente indipendenti e per la sezione precedente sappiamo che questo implica che le n colonne di $[M]$ sono linearmente indipendenti ossia che v_1, v_2, \dots, v_n sono linearmente indipendenti.

Basta per dire che sono una base ? Resta da vedere che generano V . Pensiamo alla applicazione M' che è associata a $[M']$: visto che $[M']$ è la matrice identità si osserva subito che $\text{Imm } M' = V$. Ma per le proposizioni discusse nella sezione precedente $\text{Imm } M' = \text{Imm } M$, dunque $\text{Imm } M = V$, cioè le colonne di $[M]$ generano V , quindi v_1, v_2, \dots, v_n generano V , come volevamo.

Corollario 0.25. *In uno spazio vettoriale V di dimensione n , se ho n vettori linearmente indipendenti questi sono anche una base di V . Allo stesso modo, se ho n vettori che generano V questi sono anche una base di V .*

Dimostrazione. In entrambi i casi, la matrice $n \times n$ M ottenuta con i vettori dati deve avere come forma a scalini $M' = I$. Il perché si basa su osservazioni già fatte, ma ripetiamole per esercizio: se i vettori erano linearmente indipendenti allora le n colonne di M' devono essere indipendenti (facendo le mosse di colonna il massimo numero di colonne linearmente indipendenti non cambia), e dunque non ci possono essere scalini lunghi. Se i vettori generavano V allora le n colonne di M' devono generare V (facendo le mosse di colonna lo spazio generato dalle colonne non cambia) e dunque di nuovo non ci possono essere scalini lunghi. \square

Facciamo ora un semplice esempio concreto di “riconoscimento” di una base. Consideriamo \mathbb{R}^4 con la sua base standard e poi i vettori

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, v_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Si noti che si tratta dei vettori utilizzati in un esempio alla fine della sezione “Applicazioni lineari e matrici”; avrete dunque già verificato che si tratta di una base. Ma ora lo possiamo fare col nuovo metodo.

Scriviamo dunque la matrice

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

e cerchiamo di portarla in forma a scalini ridotta. Sottraendo la quarta colonna alla terza otteniamo

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Sottraendo la terza colonna alla seconda otteniamo

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

e infine sottraendo la seconda colonna alla prima troviamo la matrice identità come volevamo:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Dunque $\{v_1, v_2, v_3, v_4\}$ è una base di \mathbb{R}^4 . In questo esempio i calcoli erano particolarmente semplici, ma è già possibile notare la “convenienza” di questo metodo.

• **Le operazioni di riga.**

Nel paragrafo precedente abbiamo visto le operazioni di colonna su una matrice. Ripetiamo tutto per le righe invece che per le colonne; data una matrice $m \times n$, a coefficienti in un campo K , consideriamo sulle righe le mosse “simmetriche” alle mosse di colonna che conosciamo già:

- si somma alla riga i la riga j moltiplicata per uno scalare λ ;
- si moltiplica la riga s per uno scalare $k \neq 0$;
- si permutano fra di loro due righe, diciamo la i e la j .

Per simmetria col caso delle colonne sappiamo che è sempre possibile, usando le mosse descritte sopra, ridurre la matrice in una forma detta “a scalini (per righe)”. Ecco alcuni esempi di matrici in forma a scalini per righe:

$$\begin{pmatrix} 1 & 5 & 4 & 0 \\ 0 & 1 & \sqrt{7} + 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3 & 0 & 9 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & \sqrt{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 5 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Osservazione 0.26. Per una definizione formale di “matrice a scalini per righe” possiamo procedere come nel paragrafo precedente, chiamando stavolta “profondità” di una riga la posizione occupata, contata da destra, dal suo coefficiente diverso da zero che sta più a sinistra. . . (insomma la riga $(0, 0, 7, \sqrt{2}, 3, 0, 0, 9)$ ha profondità 6, visto che il coefficiente 7 occupa la sesta casella contando da destra).

Anche in questo caso possiamo sempre ottenere anche la forma “a scalini (per riga) ridotta”.

Ecco le forme a scalini ridotte negli esempi appena visti:

$$\begin{pmatrix} 1 & 0 & -1 - 5\sqrt{7} & 0 \\ 0 & 1 & \sqrt{7} + 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 9 + 9\sqrt{3} \\ 0 & 1 & 0 & -3\sqrt{3} \\ 0 & 0 & 1 & \sqrt{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & -20 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 5 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Ogni singola mossa sulle righe corrisponde stavolta a moltiplicare la nostra matrice $m \times n$, a SINISTRA, per una matrice $n \times n$ invertibile. Per esempio la mossa:

– si somma alla riga i la riga j moltiplicata per uno scalare λ ;

corrisponde a moltiplicare la nostra matrice per la matrice $n \times n$ (chiamiamola U_{ij}) che ha tutti 1 sulla diagonale, e 0 in tutte le altre caselle eccetto che nella casella identificata da “riga i , colonna j ”, dove troviamo λ (questa matrice U_{ij} è la simmetrica rispetto alla diagonale della matrice E_{ij} analoga usata nel caso delle mosse di colonna: è dunque facile verificare che è invertibile..anche U_{ij}^{-1} sarà la simmetrica di E_{ij}^{-1} ..).

Consideriamo allora una applicazione lineare:

$$L : V \rightarrow W$$

dove V e W sono due spazi vettoriali sul campo K , di dimensione n e m rispettivamente.

Fissiamo come al solito una base in V e una in W e consideriamo la matrice $[L]$, associata a L .

Agire sulle righe di $[L]$ fino a ridurla in forma a scalini ridotta equivale a dire che moltiplichiamo $[L]$ a sinistra per delle matrici invertibili $[U_1], [U_2], \dots, [U_s]$ fino a che $[U_s][U_{s-1}] \cdots [U_1][L]$ è in forma a scalini (per righe) ridotta.

Per semplificare la notazione, chiamiamo $[U] = [U_s][U_{s-1}] \cdots [U_1]$: sappiamo che $[U]$ è una matrice invertibile, visto che è il prodotto di matrici invertibili, e chiamiamo U l'applicazione lineare da W in W che, rispetto alla base fissata di W , è associata a $[U]$.

L' applicazione lineare associata a $[U][L]$ è proprio la

$$U \circ L : V \rightarrow W$$

In questo caso non è vero che $\text{Imm } L = \text{Imm } U \circ L$; è vero però che queste due immagini hanno la stessa dimensione, come ci viene garantito dalla seguente

Proposizione 0.27. *Sia*

$$B : W \rightarrow W$$

una applicazione lineare invertibile. Allora vale $\dim \text{Imm } L = \dim \text{Imm } B \circ L$, ossia L e $B \circ L$ hanno lo stesso rango.

Dimostrazione. Sia $\{v_1, \dots, v_r\}$ una base di $\text{Imm } L$ (che dunque ha dimensione r). Se dimostriamo che $\{B(v_1), \dots, B(v_r)\}$ è una base di $\text{Imm } B \circ L$ abbiamo finito, perché allora anche $\text{Imm } B \circ L$ ha dimensione r .

Controlliamo per prima cosa che $B(v_1), \dots, B(v_r)$ sono linearmente indipendenti: consideriamo una combinazione lineare che si annulla

$$a_1 B(v_1) + \dots + a_r B(v_r) = O$$

e verifichiamo che deve valere $a_1 = a_2 = \dots = a_r = 0$.

Per la linearità di B possiamo riscrivere

$$B(a_1 v_1 + \dots + a_r v_r) = O$$

Siccome B è invertibile, allora è bigettiva, in particolare è iniettiva. Visto che $B(O) = O$ (B è lineare), per la iniettività non possono esserci altri elementi di W mandati in O da B , dunque

$$a_1 v_1 + \dots + a_r v_r = O$$

Ma v_1, \dots, v_r sono linearmente indipendenti (sono una base di $\text{Imm } L$) e quindi deve valere $a_1 = a_2 = \dots = a_r = 0$ come volevamo.

Ci resta da verificare che $B(v_1), \dots, B(v_r)$ generano $\text{Imm } B \circ L$. Gli elementi di $\text{Imm } B \circ L$ sono tutti i vettori della forma $B(u)$ dove u è un vettore che appartiene all'immagine di L . Ma allora u può essere scritto in termini della base $\{v_1, \dots, v_r\}$ di $\text{Imm } L$:

$$u = b_1 v_1 + \dots + b_r v_r$$

e di conseguenza

$$B(u) = B(b_1 v_1 + \dots + b_r v_r) = b_1 B(v_1) + \dots + b_r B(v_r)$$

Questo dimostra che $B(v_1), \dots, B(v_r)$ generano $\text{Imm } B \circ L$. □

Dunque, siccome il rango di L e quello di $U \circ L$ sono uguali, allora il massimo numero di colonne linearmente indipendenti di $[L]$ deve essere uguale al massimo numero di colonne linearmente indipendenti della sua forma a scalini per righe $[U][L]$ (questo infatti è un modo di contare il rango). Ma si vede subito che una matrice in forma a scalini per righe ha tante colonne linearmente indipendenti quanti sono i suoi scalini ossia quante sono le righe non nulle. Dunque abbiamo dimostrato:

Proposizione 0.28. *Data l'applicazione lineare L come sopra, il suo rango è uguale al numero di righe non nulle che si trovano quando si riduce una matrice associata $[L]$ in forma a scalini per righe.*

Nel caso delle colonne avevamo osservato che il numero di colonne non zero della forma a scalini era uguale al massimo numero di colonne linearmente indipendenti della matrice iniziale: per ragioni puramente di simmetria, questo deve valere anche per le righe (qui consideriamo le righe come dei vettori di uno spazio vettoriale, scritti per riga invece che per colonna come facciamo di solito). Completiamo allora la proposizione:

Proposizione 0.29. *Data l'applicazione lineare L come sopra, il suo rango è uguale al numero di righe non nulle che si trovano quando si riduce una matrice associata $[L]$ in forma a scalini per righe. Tale numero è anche uguale al massimo numero di righe linearmente indipendenti di L .*

E, in sintesi:

Teorema 0.30. *Data l'applicazione lineare $L : V \rightarrow W$, fissiamo una base in V e una in W e consideriamo la matrice $[L]$ associata a L rispetto a tali basi.*

1) *Il massimo numero di righe linearmente indipendenti di questa matrice è uguale al massimo numero di colonne linearmente indipendenti, ossia al rango di L .*

2) *Se si riduce la matrice in forma a scalini, sia che lo si faccia per righe, sia che lo si faccia per colonne, il numero di scalini che otterremo sarà sempre uguale al rango di L .*

Osservazione 0.31. Potremmo chiamare “rango” di una matrice M il massimo numero di colonne (o righe, abbiamo visto che è lo stesso) linearmente indipendenti. Con tale definizione il rango di una applicazione lineare L coincide con quello di una sua matrice associata $[L]$, dunque siamo contenti e possiamo usare la parola “rango” senza tante attenzioni, applicandola sia alle matrici sia alle applicazioni. Ora, sappiamo che se componiamo L a destra o a sinistra per una applicazione invertibile, il rango non cambia. Dunque, se moltiplichiamo $[L]$ a destra o a sinistra per matrici invertibili, anche il rango delle matrici non cambia. Allora se abbiamo una applicazione lineare L e ci interessa calcolarne il rango, prendiamo una matrice associata $[L]$ (va bene una qualunque, il discorso è indipendente dalle basi) e possiamo ridurla a scalini usando sia mosse di riga sia mosse di colonna nell'ordine che ci torna più comodo.

- **Teorema del completamento e teorema sulla dimensione del nucleo e della immagine di una applicazione lineare.**

Nel paragrafo precedente abbiamo dimostrato questo teorema:

Teorema. *In uno spazio vettoriale V di dimensione n , se ho n vettori linearmente indipendenti questi sono anche automaticamente una base di V .*

Utilizziamolo per dimostrare il teorema di completamento ad una base. La dimostrazione qui di seguito mette bene in risalto l'uso del principio di induzione.

Teorema 0.32 (Teorema del Completamento). *Dato uno spazio vettoriale V di dimensione n , ogni sottoinsieme $B = \{v_1, \dots, v_k\} \subset V$ di vettori linearmente indipendenti di cardinalità $k \leq n$, può essere completato ad una base di V aggiungendo a B $n - k$ vettori di $V \setminus \text{Span}(B)$.*

Dimostrazione.

Procediamo per induzione su $i = n - k$.

Passo base. Per $i = 1$, ossia per $k = n - 1$, aggiungiamo a B un vettore $s \in (V \setminus \text{Span}(B))$, cioè costruiamo $B' = B \cup \{s\} = \{v_1, \dots, v_{n-1}, s\}$. I vettori di B' sono linearmente indipendenti. Infatti

$$a_1v_1 + \dots + a_{n-1}v_{n-1} + a_ns = 0$$

è una combinazione lineare di coefficienti tutti nulli, dal momento che se fosse $a_n \neq 0$ si avrebbe $s = \frac{1}{a_n}(-a_1v_1 - \dots - a_{n-1}v_{n-1})$ che è impossibile per la scelta di s , viceversa con $a_n = 0$ si avrebbe $a_1v_1 + \dots + a_{n-1}v_{n-1} = 0$ che per l'indipendenza dei vettori di B implica $a_1 = \dots = a_{n-1} = 0$. Per il teorema ricordato sopra sappiamo che ogni sottoinsieme di vettori linearmente indipendenti di cardinalità n è una base se lo spazio che li contiene ha dimensione n , dunque abbiamo concluso che B' è una base.

Passo induttivo. Adesso supponiamo vera la tesi per $i = r$ e mostriamo che vale per $i = r + 1$. Sia $B = \{v_1, \dots, v_k\} \subset V$ un insieme di vettori linearmente indipendenti con $k = n - (r + 1)$. In modo del tutto analogo a quanto visto per il passo base posso estendere B con un vettore $s_1 \in (V \setminus \text{Span}(B))$ ed avere $B' = B \cup \{s_1\} = \{v_1, \dots, v_k, s_1\}$ in modo tale che i vettori di B' risultino linearmente indipendenti. Ora è possibile applicare l'ipotesi induttiva e completare B' ad una base di V , provando il teorema per induzione. ■

In classe abbiamo dato una dimostrazione di questo stesso teorema che si basa sulla riduzione a scalini e fornisce un metodo concreto per scegliere i vettori che estendono v_1, v_2, \dots, v_k ad una base. Supponiamo di conoscere una base di V ; per prima cosa si scrivono i vettori v_1, v_2, \dots, v_k come vettori colonna e si forma una matrice M . Poi si riduce M in forma a scalini per colonne. Tutte le volte che troviamo uno scalino lungo (diciamo di altezza $i \geq 2$) possiamo facilmente trovare $i - 1$ vettori w_1, w_2, \dots, w_{i-1} tali che $\{v_1, v_2, \dots, v_k, w_1, w_2, \dots, w_{i-1}\}$ è ancora un insieme di vettori linearmente indipendenti. Ripetendo questa costruzione per ogni scalino lungo, troviamo alla fine n vettori linearmente indipendenti, dunque una base di V come richiesto. Illustriamo questo metodo con un esempio: supponiamo che $V = \mathbb{R}^7$ e di avere 4 vettori linearmente indipendenti che scriviamo rispetto alla base

standard di \mathbb{R}^7 :

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \\ 1 \\ 0 \\ 3 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}$$

La matrice M in questo caso è:

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 3 & 1 & 2 \end{pmatrix}$$

e una sua riduzione a scalini per colonne è :

$$M' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 \end{pmatrix}$$

Il primo scalino lungo ha altezza 3, e osserviamo subito che i vettori

$$w_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

non appartengono al sottospazio generato dalle colonne di M' (che coincide col sottospazio generato da v_1, v_2, v_3, v_4). Similmente, prendendo in considerazione il secondo scalino lungo (che ha altezza 2), notiamo che il vettore

$$w_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

non appartiene al sottospazio generato da v_1, v_2, v_3, v_4 . E' facile ora vedere che i vettori $v_1, v_2, v_3, v_4, w_1, w_2, w_3$ formano una base: quando si scrive la matrice 7×7 M'' formata da tali vettori si nota subito che la sua forma a scalini ridotta è l'identità (abbiamo proprio aggiunto a M tre vettori che "accorciano" i suoi scalini lunghi..).

Usando il teorema del completamento possiamo dimostrare:

Teorema 0.33. *Considerata una applicazione lineare $L : V \rightarrow W$, dove V e W sono spazi vettoriali su di un campo K e V ha dimensione finita, vale*

$$\dim Ker L + \dim Imm L = \dim V$$

Dimostrazione. Sia $n = \dim V$ e sia $\{z_1, z_2, \dots, z_k\}$ una base di $Ker L$ (che dunque ha dimensione $k \leq n$). Se $k = n$ abbiamo finito (l'applicazione L in tal caso è l'applicazione nulla e la formula è banalmente verificata).

Altrimenti, se $k < n$, per il teorema del completamento posso trovare w_1, w_2, \dots, w_{n-k} tali che $\{z_1, z_2, \dots, z_k, w_1, w_2, \dots, w_{n-k}\}$ sia una base di V .

Sappiamo che $Imm L$ è il sottospazio generato da

$$L(z_1), L(z_2), \dots, L(z_k), L(w_1), L(w_2), \dots, L(w_{n-k})$$

ma siccome gli z_j sono nel nucleo vale che, per ogni $j = 1, \dots, k$, $L(z_j) = 0$ e allora

$$Imm L = \langle L(w_1), L(w_2), \dots, L(w_{n-k}) \rangle$$

Se dimostriamo che $L(w_1), L(w_2), \dots, L(w_{n-k})$ sono linearmente indipendenti, abbiamo finito (perché allora sono una base di $Imm L$ che dunque ha dimensione $n - k$, e la formula è verificata).

Per dimostrare la indipendenza lineare scriviamo:

$$a_1 L(w_1) + a_2 L(w_2) + \dots + a_{n-k} L(w_{n-k}) = O$$

dove, per ogni i , $a_i \in K$. Se questo è vero solo quando $a_1 = \dots = a_{n-k} = 0$ allora $L(w_1), L(w_2), \dots, L(w_{n-k})$ sono linearmente indipendenti.

Per linearità l'equazione equivale a:

$$L(a_1 w_1 + a_2 w_2 + \dots + a_{n-k} w_{n-k}) = O$$

ossia

$$a_1 w_1 + a_2 w_2 + \dots + a_{n-k} w_{n-k} \in Ker L$$

Ma allora possiamo esprimere $a_1 w_1 + a_2 w_2 + \dots + a_{n-k} w_{n-k}$ come combinazione lineare di z_1, \dots, z_k visto che questi sono una base di $Ker L$:

$$a_1 w_1 + a_2 w_2 + \dots + a_{n-k} w_{n-k} = b_1 z_1 + \dots + b_k z_k$$

dove i $b_j \in K$, che diventa

$$a_1 w_1 + a_2 w_2 + \dots + a_{n-k} w_{n-k} - b_1 z_1 + \dots - b_k z_k = O$$

Siccome $\{z_1, z_2, \dots, z_k, w_1, w_2, \dots, w_{n-k}\}$ è una base di V allora tutti i coefficienti nella equazione sopra devono essere uguali a 0. In particolare $a_1 = \dots = a_{n-k} = 0$ come volevamo. \square

Definizione 0.34. Una applicazione lineare bigettiva $L : V \rightarrow W$, dove V e W sono spazi vettoriali su di un campo K , si dice un "isomorfismo lineare".

Il teorema precedente ci permette subito di notare che se $L : V \rightarrow W$ è un isomorfismo allora $\dim V = \dim W$. Se invece abbiamo una applicazione lineare $\Gamma : V \rightarrow W$ che sappiamo essere iniettiva (ossia $\text{Ker } \Gamma = \{O\}$) allora possiamo dire che $\dim \text{Imm } \Gamma = \dim V$ e Γ è un isomorfismo lineare quando viene pensata come funzione da V a $\text{Imm } \Gamma$ invece che come funzione da V a W .

• **La formula di Grassmann per le intersezioni e le somme di sottospazi.**

Dati due piani A e B (sottospazi vettoriali di dimensione 2, dunque contenenti O) in \mathbb{R}^3 , di che dimensione può essere la loro intersezione? Possono intersecarsi lungo una retta R , in tal caso si nota che il sottospazio generato dai vettori di $A \cup B$ è tutto \mathbb{R}^3 . Oppure possono essere coincidenti, ossia $A = B$, e allora la loro intersezione è A e ha dimensione 2, e il sottospazio generato dai vettori di $A \cup B$ è sempre A e ha dimensione 2.

In entrambi i casi, la somma della dimensione di $A \cap B$ e del sottospazio generato dai vettori di $A \cup B$ è sempre uguale a 4 (che a sua volta è uguale a $\dim A + \dim B$).

E se consideriamo un piano C e un sottospazio 3-dimensionale D in \mathbb{R}^4 ? Possono darsi tre casi per l'intersezione: $C \cap D = \{O\}$, $\dim(C \cap D) = 1$, $C \cap D = C$. Qualunque sia il caso, si verifica sempre che (esercizio!) la somma della dimensione di $C \cap D$ e del sottospazio generato dai vettori di $C \cup D$ è uguale a $5 = \dim C + \dim D$.

Sembra dunque che possa esserci una relazione fra le dimensioni in gioco: se due sottospazi A e B si intersecano “tanto”, allora generano “poco”, più precisamente:

$$\dim A + \dim B = \dim A \cap B + \dim \langle A \cup B \rangle$$

Questa formula ci dice, per esempio, che in \mathbb{R}^5 due sottospazi di dimensione 3 devono avere intersezione non banale: $\dim A = \dim B = 3$ e inoltre $\dim \langle A \cup B \rangle \leq 5$, dunque $\dim A \cap B \geq 1$.

Dimostreremo questa formula (detta “formula di Grassmann”) come applicazione dei teoremi dimostrati nella sezione precedente.

Per enunciarla bene introduciamo qualche notazione nuova. Per prima cosa ci occupiamo di $\langle A \cup B \rangle$.

Definizione 0.35. Dati due sottospazi vettoriali A e B di uno spazio vettoriale V , chiameremo $A + B$ l'insieme

$$A + B = \{a + b \mid a \in A, \quad b \in B\}$$

Osservazione 0.36. Si verifica facilmente (esercizio) che $A + B$ è un sottospazio vettoriale di V , e coincide con quello che finora abbiamo chiamato $\langle A \cup B \rangle$, ossia il sottospazio generato dai vettori di $A \cup B$.

Studiamo adesso il prodotto cartesiano di due spazi vettoriali: dati due spazi vettoriali V e W sul campo K , sul loro prodotto cartesiano $V \times W$ c'è una struttura “naturale” di spazio vettoriale, dove la somma è definita da:

$$(v, w) + (v_1, w_1) = (v + v_1, w + w_1)$$

e il prodotto per scalare da:

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

Si verifica che, se $\{v_1, v_2, \dots, v_n\}$ è una base di V e $\{w_1, \dots, w_m\}$ è una base di W , allora $\{(v_1, O), (v_2, O), \dots, (v_n, O), (O, w_1), \dots, (O, w_m)\}$ è una base di $V \times W$, che dunque ha dimensione $n + m = (\dim V) + (\dim W)$.

Possiamo finalmente enunciare e dimostrare la formula di Grassmann:

Teorema 0.37. *Dati due sottospazi A, B di uno spazio vettoriale V di dimensione finita, vale*

$$\dim A + \dim B = \dim A \cap B + \dim (A + B)$$

Dimostrazione. Consideriamo l'applicazione

$$\Phi : A \times B \rightarrow V$$

definita da $\Phi((a, b)) = a - b$. Si verifica (esercizio) che Φ è lineare.

Chi è il nucleo di Φ ? Per definizione

$$\text{Ker } \Phi = \{(a, b) \in A \times B \mid a - b = O\} = \{(a, b) \in A \times B \mid a = b\} = \{(z, z) \mid z \in A \cap B\}$$

Si nota subito che la mappa $\theta : A \cap B \rightarrow \text{Ker } \Phi$ data da $z \rightarrow (z, z)$ è un isomorfismo lineare. Dunque

$$\dim \text{Ker } \Phi = \dim A \cap B$$

Cosa sappiamo dire dell'immagine di Φ ? Per definizione

$$\text{Imm } \Phi = \{a - b \in V \mid a \in A, b \in B\} = \{a + b \in V \mid a \in A, b \in B\} = A + B$$

Dunque

$$\text{Imm } \Phi = A + B$$

Per il teorema che lega le dimensioni dell'immagine e del nucleo di una applicazione lineare sappiamo che:

$$\dim (A \times B) = \dim \text{Ker } \Phi + \dim \text{Imm } \Phi$$

che possiamo tradurre:

$$\dim A + \dim B = \dim A \cap B + \dim (A + B)$$

□

Esercizio. Dire se è possibile trovare in \mathbb{R}^4 tre sottospazi vettoriali di dimensione due A, B, C tali che $A \cap B = \{O\}$, $A \cap C = \{O\}$ e $B \cap C = \{O\}$.

Esercizio. Dati tre sottospazi vettoriali A, B, C di uno spazio vettoriale V , dare una buona definizione di $A + B + C$ e dire se è vera la formula:

$$\dim (A+B+C) = \dim A + \dim B + \dim C - \dim (A \cap B) - \dim (B \cap C) - \dim (A \cap C) + \dim (A \cap B \cap C)$$

• **Autovalori e autovettori di un endomorfismo lineare.**

Sia $T : V \rightarrow V$ una applicazione lineare da uno spazio vettoriale V (sul campo K) in sé. Una tale T è chiamata anche endomorfismo lineare.

Definizione 0.38. Un vettore $v \in V - \{O\}$ si dice un autovettore di T se

$$T(v) = \lambda v$$

per un certo $\lambda \in K$.

In altre parole un autovettore di T è un vettore NON NULLO dello spazio V che ha la seguente proprietà: la T lo manda in un multiplo di sé stesso.

Definizione 0.39. Se $v \in V - \{O\}$ è un autovettore di T tale che

$$T(v) = \lambda v$$

allora lo scalare $\lambda \in K$ si dice autovalore di T relativo a v (e viceversa si dirà che v è un autovettore relativo all'autovalore λ).

Si noti che è l'autovettore che non può essere $O \in V$, mentre l'autovalore può benissimo essere $0 \in K$: se per esempio T non è iniettiva, ossia $Ker T \supsetneq \{0\}$, tutti gli elementi $w \in (Ker T) - \{O\}$ soddisfano

$$T(w) = O = 0 w$$

ossia sono autovettori relativi all'autovalore 0.

Definizione 0.40. Dato $\lambda \in K$ chiamiamo l'insieme

$$V_\lambda = \{v \in V \mid T(v) = \lambda v\}$$

“autospatio relativo a λ ”.

Si verifica facilmente che un autospatio V_λ è un sottospazio vettoriale di V . Anche se abbiamo definito l' autospatio V_λ per qualunque $\lambda \in K$, in realtà V_λ è sempre uguale a $\{O\}$ a meno che λ non sia un autovalore. Questo è dunque il caso interessante: se λ è un autovalore di T allora V_λ è costituito da tutti gli autovettori relativi a λ più lo O . In particolare abbiamo notato poco fa che $V_0 = Ker T$.

Perché per noi sono importanti autovettori e autovalori ?

Supponiamo che V abbia dimensione n e pensiamo a cosa succederebbe se riuscissimo a trovare una base di V , $\{v_1, v_2, \dots, v_n\}$, composta solo da autovettori di T .

Avremmo, per ogni $i = 1, 2, \dots, n$,

$$T(v_i) = \lambda_i v_i$$

per certi autovalori λ_i (sui quali non sappiamo nulla, sono semplicemente elementi di K , addirittura potrebbero anche essere tutti uguali $\lambda_1 = \dots = \lambda_n$).

Come sarebbe fatta la matrice

$$[T] \begin{matrix} \{v_1, v_2, \dots, v_n\} \\ \{v_1, v_2, \dots, v_n\} \end{matrix}$$

associata a T rispetto a questa base ?

Ricordandosi come si costruiscono le matrici osserviamo che la prima colonna conterrebbe il vettore $T(v_1)$ scritto in termini della base $\{v_1, v_2, \dots, v_n\}$, ossia

$$T(v_1) = \lambda_1 v_1 + 0v_2 + 0v_3 + 0v_4 + \dots + 0v_n$$

la seconda il vettore $T(v_2) = 0v_1 + \lambda_2 v_2 + 0v_3 + \dots + 0v_n$ e così via. Quindi la matrice sarebbe diagonale:

$$[T] \begin{matrix} \{v_1, v_2, \dots, v_n\} \\ \{v_1, v_2, \dots, v_n\} \end{matrix} = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_{n-1} & 0 \\ 0 & 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

Ora, una matrice diagonale è per noi “leggibilissima”; solo guardandola possiamo sapere tutto di T : il suo rango (dunque anche la dimensione del nucleo), quali sono esattamente i vettori di $\text{Ker } T$, quali sono i sottospazi in cui T si comporta come l'identità, ossia i sottospazi costituiti dai vettori di V che la T lascia fissi...

Dunque studiamo gli autovalori e gli autovettori di T nella speranza di trovare “basi buone” che ci permettano di conoscere bene il comportamento di T .

Ma esistono sempre queste “basi buone”, ossia basi costituite solo da autovettori di T ? NO, non sempre. Se per una certa T esiste una base buona si dice che T è “diagonalizzabile”, altrimenti T è “non diagonalizzabile”.

Vorremmo trovare dei criteri semplici che ci permettono di decidere se una applicazione lineare è diagonalizzabile o no. Un buon primo passo è quello di avere un metodo che, dato un endomorfismo $T : V \rightarrow V$ e posto $n = \dim V$, ci permetta di decidere se uno scalare $\lambda \in K$ è o no un autovalore di T . Entrano qui in gioco i polinomi e le loro radici.

Innanzitutto osserviamo che, perché $\lambda \in K$ sia un autovalore, secondo la definizione bisogna che esista un $v \in V - \{O\}$ tale che

$$T(v) = \lambda v$$

Questo si può riscrivere anche come

$$T(v) - \lambda I(v) = O$$

dove $I : V \rightarrow V$ è l'identità. Riscriviamo ancora:

$$(T - \lambda I)(v) = O$$

Abbiamo “scoperto” che la applicazione lineare $T - \lambda I$ non è iniettiva: infatti manda il vettore v in O . Dunque, se scegliamo una base qualunque per V e costruiamo la matrice $[T]$ associata a T , la matrice $[T] - \lambda[I]$ dovrà avere determinante uguale a 0: $\det([T] - \lambda[I]) = 0 = \det(\lambda[I] - [T])$. A dire la verità scrivere $[I]$ è un formalismo inutile: in qualunque base la si scriva, la matrice associata all'identità è comunque

diagonale con tutti i coefficienti diagonali uguali a 1, quindi da ora in poi ometteremo le parentesi quadre e scriveremo semplicemente I .

Questa osservazione ci avvicina alla seguente definizione:

Definizione 0.41. Dato un endomorfismo lineare $T : V \rightarrow V$ con $n = \dim V$, scegliamo una base per V e costruiamo la matrice $[T]$ associata a T rispetto a tale base. Ora definiamo il polinomio caratteristico di T $p_T(t) \in K[t]$ tramite la seguente equazione:

$$p_T(t) \equiv \det(tI - [T])$$

Ossevazioni importanti.

1) Perché la definizione data abbia senso innanzitutto bisogna verificare che $\det(tI - [T])$ sia veramente un polinomio. Questo si può dimostrare facilmente per induzione sulla dimensione n di V . Sempre per induzione si può dimostrare un po' di più, ossia che $p_T(t)$ è un polinomio di grado n con coefficiente direttivo 1: $p_T(t) = t^n + \dots$. Queste dimostrazioni sono facoltative e consigliate !

2) Inoltre ci interessa molto che la definizione appena data non dipenda dalla base scelta di V : non sarebbe una definizione tanto buona se con la scelta di due basi diverse ottenessimo due polinomi caratteristici diversi !

Questo problema per fortuna non si verifica. In generale vale il fatto seguente: dato l'endomorfismo $A : V \rightarrow V$, scegliamo una base b di V e costruiamo la matrice $[A]_b$. Poi scegliamo un'altra base b' e costruiamo $[A]_{b'}$. Le due matrici $[A]_b$ e $[A]_{b'}$ sono legate dalla seguente relazione: esiste una matrice $[B]$ invertibile tale che

$$[B][A]_b [B]^{-1} = [A]_{b'}$$

e quindi anche

$$[B]^{-1}[A]_{b'} [B] = [A]_b$$

Lasciamo facoltativa la dimostrazione di quanto appena detto. Usando il teorema di Binet a questo punto sappiamo verificare che

$$\det [A]_{b'} = \det [A]_b$$

Tornando al caso del nostro polinomio caratteristico, possiamo dunque stare sicuri che $p_T(t) = \det (tI - [T])$ non dipende dalla scelta della base.

Assicurateci che la definizione è ben data, enunciamo il teorema principale che spiega l'utilità del polinomio caratteristico.

Teorema 0.42. *Considerata T come sopra, vale che uno scalare $\lambda \in K$ è un autovalore di T se e solo se λ è una radice di $p_T(t)$, ossia se e solo se $p_T(\lambda) = 0$*

Dimostrazione. Abbiamo già visto (l'osservazione prima della definizione del polinomio caratteristico) che se λ è un autovalore di T allora $\det(\lambda I - [T]) = p_T(\lambda) = 0$. Resta dunque da dimostrare il viceversa. Supponiamo che $\det(\lambda I - [T]) = p_T(\lambda) = 0$: allora l'applicazione lineare $\lambda I - T$ non è iniettiva. dunque esiste $v \in V - \{O\}$ tale che $(\lambda I - T)(v) = 0$, che si può riscrivere anche come

$$T(v) = \lambda v$$

Abbiamo trovato un autovettore che ha autovalore λ e quindi abbiamo mostrato, come volevamo, che λ è un autovalore di T . ■

• **Una strategia per scoprire se un endomorfismo è diagonalizzabile.**

In questo paragrafo descriviamo una strategia in 4 passi che ci permette di scoprire se un endomorfismo $T : V \rightarrow V$, dove V è uno spazio vettoriale sul campo K con $n = \dim V$, è diagonalizzabile, e, in caso sia diagonalizzabile, di trovare una base che lo diagonalizza, ossia una base di V fatta tutta da autovettori di T . La nostra strategia sarà la seguente:

- PASSO 1. Data T , troviamo gli autovalori di T utilizzando il polinomio caratteristico.
- PASSO 2. Supponiamo di aver trovato gli autovalori $\lambda_1, \lambda_2, \dots, \lambda_k$: a questo punto scopriamo chi sono i relativi autospazi $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_k}$.
- PASSO 3. Un teorema ci assicurerà che gli autospazi $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_k}$ sono in somma diretta, ossia che vale

$$V_{\lambda_1} \oplus V_{\lambda_2} \oplus \dots \oplus V_{\lambda_k}$$

Quindi se

$$\dim V_{\lambda_1} + \dim V_{\lambda_2} + \dots + \dim V_{\lambda_k} = n = \dim V$$

allora è possibile trovare una base “buona”, fatta da autovettori di T e T è diagonalizzabile. Per scrivere una base “buona” basta scegliere una base per ogni V_{λ_i} e poi fare l'unione. Altrimenti, se

$$\dim V_{\lambda_1} + \dim V_{\lambda_2} + \dots + \dim V_{\lambda_k} < n = \dim V$$

T non è diagonalizzabile.

- PASSO 4. Se T è risultata diagonalizzabile, usando la base trovata si scrive la matrice diagonale $[T]$.

Vediamo i dettagli passo per passo.

PASSO 1.

Di questo ci siamo già occupati nel paragrafo precedente: per sapere quali sono gli autovalori di un endomorfismo T basta calcolare il polinomio caratteristico $p_T(t)$ e trovare le sue radici in K .

PASSO 2.

Supponiamo dunque di aver scoperto che T ha i seguenti autovalori: $\lambda_1, \lambda_2, \dots, \lambda_k$, tutti distinti fra loro. Vogliamo capire chi sono gli autospazi $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_k}$.

Per questo basterà risolvere dei sistemi lineari: per ogni $i = 1, 2, \dots, k$, l'autospazio V_{λ_i} è costituito per definizione dai vettori $v \in V$ tali che $T(v) = \lambda_i v$, ossia, scelta una base di V e dunque trovata la matrice $[T]$, dalle soluzioni del sistema lineare

$$([T] - \lambda_i I) \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ \cdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ \cdots \\ 0 \\ 0 \end{pmatrix}$$

PASSO 3.

Qui cominciamo col dimostrare il seguente teorema.

Teorema 0.43. *Dato un endomorfismo lineare $T : V \rightarrow V$, siano $\lambda_1, \lambda_2, \dots, \lambda_k$ degli autovalori di T distinti fra loro. Consideriamo ora degli autovettori $v_1 \in V_{\lambda_1}, v_2 \in V_{\lambda_2}, \dots, v_k \in V_{\lambda_k}$. Allora $\{v_1, v_2, \dots, v_k\}$ è un insieme di vettori linearmente indipendenti.*

Osservazione. Spesso ci si riferisce a questo teorema con la frase: “autovettori relativi ad autovalori distinti sono linearmente indipendenti”.

Dimostrazione. Procediamo per induzione su k . Per $k = 1$ l'enunciato è banale perché $\{v_1\}$ è banalmente un insieme di vettori linearmente indipendenti (c'è un vettore solo..). Supponiamo di aver dimostrato che l'enunciato è vero fino a $k - 1$ e cerchiamo di dimostrarlo per k . Supponiamo allora che valga:

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = O$$

Per mostrare che $\{v_1, v_2, \dots, v_k\}$ è un insieme di vettori linearmente indipendenti dobbiamo mostrare che questo può accadere solo quando $a_1 = a_2 = \dots = a_k = 0$.

Dalla equazione scritta ne ricaviamo due in due modi diversi. Prima applichiamo T ad entrambi i membri e per linearità ricaviamo

$$a_1T(v_1) + a_2T(v_2) + \dots + a_kT(v_k) = O$$

che svolgendo il calcolo diventa

$$a_1\lambda_1v_1 + a_2\lambda_2v_2 + \dots + a_k\lambda_kv_k = O$$

Poi invece moltiplichiamo l'equazione per λ_k ottenendo:

$$a_1\lambda_kv_1 + a_2\lambda_kv_2 + \dots + a_k\lambda_kv_k = O$$

Per sottrazione da queste due equazioni ricaviamo:

$$a_1(\lambda_k - \lambda_1)v_1 + a_2(\lambda_k - \lambda_2)v_2 + \dots + a_{k-1}(\lambda_k - \lambda_{k-1})v_{k-1} = O$$

Ma questa è una combinazione lineare dei $k - 1$ vettori $\{v_1, v_2, \dots, v_{k-1}\}$ uguale a O : per ipotesi induttiva tutti i coefficienti devono essere uguali a 0. Visto che gli scalari $\lambda_k - \lambda_i$ sono tutti diversi da zero (gli autovalori in questione sono distinti fra loro per ipotesi) questo implica che $a_1 = a_2 = \dots = a_{k-1} = 0$. Sostituendo nella equazione iniziale, notiamo che deve essere anche $a_k = 0$. ■

Il seguente teorema è un rafforzamento del precedente.

Teorema 0.44. *Teorema. Dato un endomorfismo lineare $T : V \rightarrow V$, siano $\lambda_1, \lambda_2, \dots, \lambda_k$ degli autovalori di T distinti fra loro. Gli autospazi $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_k}$ sono in somma diretta, ossia vale che*

$$V_{\lambda_1} \oplus V_{\lambda_2} \oplus \dots \oplus V_{\lambda_k}$$

Osservazione 0.45. Ricordiamo che dire che gli autospazi $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_k}$ sono in somma diretta vuol dire che se ne prendo uno qualunque, diciamo V_{λ_1} tanto per fissare la notazione, la sua intersezione con la somma di tutti gli altri è banale, ossia

$$V_{\lambda_1} \cap (V_{\lambda_2} + \dots + V_{\lambda_k}) = \{O\}$$

Questo garantisce che la dimensione di

$$V_{\lambda_1} + V_{\lambda_2} + \dots + V_{\lambda_k}$$

che possiamo scrivere anche

$$V_{\lambda_1} \oplus V_{\lambda_2} \oplus \dots \oplus V_{\lambda_k}$$

per far capire che sappiamo che la somma è diretta, è “la massima possibile”, ossia

$$\dim V_{\lambda_1} \oplus V_{\lambda_2} \oplus \dots \oplus V_{\lambda_k} = \dim V_{\lambda_1} + \dim V_{\lambda_2} + \dots + \dim V_{\lambda_k}$$

La dimostrazione è un facile esercizio (per induzione su k applicando la formula di Grassmann..).

Dimostrazione del Teorema. Supponiamo per assurdo che non sia così, e che ci sia un vettore $w \neq O$ tale che

$$w \in V_{\lambda_1} \cap (V_{\lambda_2} + \cdots + V_{\lambda_k})$$

Allora possiamo scrivere w in due modi:

$$w = v_1 \in V_{\lambda_1} - \{O\}$$

perché $w \in V_{\lambda_1}$, e

$$w = a_2v_2 + a_3v_3 + \cdots + a_kv_k$$

(dove gli a_j sono scalari e i $v_j \in V_{\lambda_j}$ per ogni j) visto che $w \in V_{\lambda_2} + \cdots + V_{\lambda_k}$.

Dunque vale:

$$v_1 = w = a_2v_2 + a_3v_3 + \cdots + a_kv_k$$

ossia

$$v_1 - a_2v_2 - a_3v_3 - \cdots - a_kv_k = O$$

Questa è una combinazione lineare di autovettori relativi ad autovalori distinti, ma non è la combinazione lineare banale (infatti almeno il coefficiente di v_1 è 1 dunque non nullo). Dunque tali autovettori sarebbero linearmente dipendenti, assurdo perché contraddice il teorema precedente. ■

Nelle ipotesi del teorema qui sopra, ci resta allora da osservare che, siccome

$$\dim (V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}) = \dim V_{\lambda_1} + \dim V_{\lambda_2} + \cdots + \dim V_{\lambda_k}$$

allora abbiamo un criterio per dire se T è diagonalizzabile o no. Infatti, se

$$\dim V_{\lambda_1} + \dim V_{\lambda_2} + \cdots + \dim V_{\lambda_k} = n = \dim V$$

allora

$$V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}$$

è un sottospazio di V che ha la stessa dimensione di V . Questo dimostra che

$$V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k} = V$$

e quindi è possibile trovare una base “buona”, fatta da autovettori di T , insomma T è diagonalizzabile.

Per scrivere una simile base “buona” basta scegliere una base per ogni V_{λ_i} e poi fare l’unione. Tale unione è una base di tutto V : infatti i vettori di tale base generano V perché $V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}$ e sono proprio n .

Altrimenti, se

$$\dim V_{\lambda_1} + \dim V_{\lambda_2} + \cdots + \dim V_{\lambda_k} < n = \dim V$$

T non è diagonalizzabile. Infatti non è possibile trovare una base di autovettori: se la trovassimo contraddiremmo

$$\dim V_{\lambda_1} + \dim V_{\lambda_2} + \cdots + \dim V_{\lambda_k} < n = \dim V$$

PASSO 4.

Se l'endomorfismo T è diagonalizzabile, scegliamo dunque una base di autovettori nel modo descritto al passo 3, e avremo una matrice associata $[T]$ che risulterà diagonale. Manteniamo le notazioni introdotte al passo 3: allora sulla diagonale troveremo $\dim V_{\lambda_1}$ coefficienti uguali a λ_1 , $\dim V_{\lambda_2}$ coefficienti uguali a λ_2 , \dots , $\dim V_{\lambda_k}$ coefficienti uguali a λ_k . Il rango di T sarà dunque uguale al numero dei coefficienti non zero che troviamo sulla diagonale, la dimensione del nucleo sarà uguale al numero dei coefficienti uguali a zero che troviamo sulla diagonale. In altre parole, se 0 non è un autovalore di T , allora $\text{Ker } T = \{O\}$, come avevamo già osservato in precedenza; se invece 0 è un autovalore di T allora troveremo sulla diagonale $\dim V_0$ coefficienti uguali a 0 - d'altronde avevamo già notato che $V_0 = \text{Ker } T$.

• **Il criterio della molteplicità algebrica e della molteplicità geometrica.**

Nella sezione precedente abbiamo trovato un criterio per decidere se un endomorfismo è diagonalizzabile o no. In questa sezione faremo una osservazione che ci permetterà di migliorare tale criterio. Consideriamo come al solito un endomorfismo $T : V \rightarrow V$, dove V è uno spazio vettoriale sul campo K con $n = \dim V$.

Calcoliamo il suo polinomio caratteristico e fattorizziamolo in $K[t]$. Otterremo una espressione del tipo:

$$P_T(t) = (t - \lambda_1)^{a_1} (t - \lambda_2)^{a_2} \dots (t - \lambda_k)^{a_k} f(t)$$

dove $\lambda_1, \lambda_2, \dots, \lambda_k$ sono gli autovalori di T in K e sono tutti distinti fra loro, e $f(t)$ o è 1 oppure è un polinomio irriducibile in $K[t]$ di grado > 1 .

Se la T è diagonalizzabile, allora esiste una base b di V in cui la matrice associata $[T]_b^b$ ha forma diagonale e sulla diagonale compaiono $\lambda_1, \lambda_2, \dots, \lambda_k$. Più esattamente, per ogni $i = 1, 2, \dots, k$, λ_i compare $\dim V_{\lambda_i}$ volte. Dunque in questo caso possiamo ricalcolare il polinomio caratteristico P_T usando $[T]_b^b$ e otteniamo

$$P_T(t) = \text{Det } (tI - [T]_b^b)$$

Si tratta di calcolare il determinante di una matrice diagonale e si osserva allora che P_T si spezza nel prodotto di fattori lineari:

$$P_T(t) = (t - \lambda_1)^{\dim V_{\lambda_1}} (t - \lambda_2)^{\dim V_{\lambda_2}} \dots (t - \lambda_k)^{\dim V_{\lambda_k}}$$

il che dimostra che il fattore $f(t)$ è 1.

In sintesi:

Proposizione 0.46. *Se l'endomorfismo T è diagonalizzabile sul campo K , allora il suo polinomio caratteristico $P_T(t)$ si fattorizza come prodotto di fattori lineari in $K[t]$:*

$$P_T(t) = (t - \lambda_1)^{\dim V_{\lambda_1}} (t - \lambda_2)^{\dim V_{\lambda_2}} \dots (t - \lambda_k)^{\dim V_{\lambda_k}}$$

Dunque se nella fattorizzazione di P_T rimane un fattore irriducibile $f(T)$ di grado > 1 possiamo concludere che T non è diagonalizzabile. Ma cosa possiamo dire del viceversa? Se p_T si fattorizza come prodotto di fattori lineari in $K[t]$ allora T è diagonalizzabile? NO, in generale non è vero. Basta considerare per esempio

l'applicazione lineare $L : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ che nelle basi standard è rappresentata dalla matrice

$$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

Il polinomio caratteristico è $p_L(t) = (t-2)^2$ ma l'applicazione non è diagonalizzabile: possiamo verificarlo applicando il criterio del paragrafo precedente, infatti L ha il solo autospazio V_2 e se ne calcoliamo la dimensione scopriamo che $\dim V_2 = 1 < 2 = \dim \mathbb{R}^2$.

Prima di enunciare il nuovo criterio diamo qualche definizione:

Definizione 0.47. Data T come sopra con polinomio caratteristico

$$P_T(t) = (t - \lambda_1)^{a_1} (t - \lambda_2)^{a_2} \cdots (t - \lambda_k)^{a_k} f(t)$$

diremo che, per ogni $i = 1, 2, \dots, k$, a_i è la “molteplicità algebrica” dell'autovalore λ_i . Chiameremo invece “molteplicità geometrica” dell'autovalore λ_i il numero intero positivo $\dim V_{\lambda_i}$.

Proposizione 0.48. Dati $T : V \rightarrow V$ e

$$P_T(t) = (t - \lambda_1)^{a_1} (t - \lambda_2)^{a_2} \cdots (t - \lambda_k)^{a_k} f(t)$$

come sopra, per ogni autovalore λ_i vale che la sua molteplicità geometrica è minore o uguale alla sua molteplicità algebrica:

$$\dim V_{\lambda_i} \leq a_i$$

Dimostrazione. Nella proposizione precedente abbiamo già visto che se l'applicazione T è diagonalizzabile, allora vale

$$\dim V_{\lambda_i} = a_i \quad \forall i = 1, 2, \dots, k$$

Se invece T non è diagonalizzabile, ricordando che gli autospazi sono in somma diretta

$$V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}$$

possiamo cominciare a costruire una base di V prendendo una base per ogni V_{λ_i} e facendo l'unione b' . Poiché in questo caso

$$\dim V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k} < \dim V$$

b' non è ancora una base di V , ma è solo un insieme di vettori linearmente indipendenti; possiamo allora, per il teorema di completamento, scegliere degli elementi s_1, \dots, s_r tali che $b = b' \cup \{s_1, \dots, s_r\}$ sia una base. Rispetto a questa base la matrice

di T ha la seguente forma:

$$[T]_b^b = \begin{pmatrix} \lambda_1 & 0 & & & & * & .. & * & * \\ 0 & \lambda_1 & & & & * & .. & * & * \\ 0 & 0 & .. & & & & & & \\ 0 & 0 & & .. & .. & & & & \\ .. & .. & & & & & & & \\ .. & .. & & & & & & & \\ .. & & & & & & & & \\ & & & & .. & 0 & & & \\ & & & & & \lambda_k & 0 & & \\ & & & & & & \lambda_k & * & .. & * & * \\ & & & & & & 0 & * & .. & * & * \\ .. & .. & & & & & 0 & * & .. & * & * \\ 0 & 0 & & & & & 0 & * & .. & * & * \\ 0 & 0 & & & & & 0 & * & .. & * & * \end{pmatrix}$$

ossia ha una parte diagonale, dove troviamo λ_1 ripetuto $\dim V_{\lambda_1}$ volte, λ_2 ripetuto $\dim V_{\lambda_2}$ volte... λ_k ripetuto $\dim V_{\lambda_k}$ volte, e poi sulle ultime r colonne, che corrispondono a $T(s_1), T(s_2), \dots, T(s_r)$ non sappiamo dire nulla.

Osserviamo però che, sviluppando il determinante di $tI - [T]_b^b$ a partire dalla prima colonna, poi dalla seconda, poi dalla terza, e così via otteniamo:

$$p_T(t) = \text{Det}(tI - [T]_b^b) = (t - \lambda_1)^{\dim V_{\lambda_1}} (t - \lambda_2)^{\dim V_{\lambda_2}} \dots (t - \lambda_k)^{\dim V_{\lambda_k}} \text{Det } M$$

dove M è la sottomatrice $r \times r$ che sta nell'angolo in fondo a destra di $tI - [T]_b^b$.

Ricordiamo ora la fattorizzazione in irriducibili per P_T

$$P_T(t) = (t - \lambda_1)^{a_1} (t - \lambda_2)^{a_2} \dots (t - \lambda_k)^{a_k} f(t)$$

L'unicità di tale fattorizzazione ci dice che la potenza massima di $(t - \lambda_1)$ che divide $P_T(t)$ è a_1 . Dunque, qualunque polinomio sia $\text{Det } M$, possiamo dire che, per ogni $i = 1, 2, \dots, k$, $\dim V_{\lambda_i} \leq a_i$. ■

Le disuguaglianze appena dimostrate implicano subito il risultato principale di questa sezione:

Teorema 0.49 (Criterio delle molteplicità algebrica e geometrica.). *Dato un endomorfismo $T : V \rightarrow V$ di uno spazio vettoriale V (di dimensione finita n) sul campo K , siano $\lambda_1, \lambda_2, \dots, \lambda_k$ gli autovalori (distinti fra loro) di T in K . Allora T è diagonalizzabile se e solo se P_T si fattorizza come prodotto di fattori lineari e, per ogni autovalore λ_i , la sua molteplicità algebrica e quella geometrica sono uguali.*

Dimostrazione. Abbiamo già visto, nelle dimostrazioni delle proposizioni precedenti, che se T è diagonalizzabile allora P_T si fattorizza come prodotto di fattori lineari e, per ogni i

$$\text{molteplicità algebrica di } \lambda_i = \dim V_{\lambda_i}$$

Viceversa, se P_T si fattorizza come prodotto di fattori lineari

$$P_T(t) = (t - \lambda_1)^{a_1} (t - \lambda_2)^{a_2} \dots (t - \lambda_k)^{a_k}$$

e, per ogni autovalore λ_i , la sua molteplicità algebrica a_i e quella geometrica sono uguali, allora calcoliamo

$$\dim V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}$$

Tale dimensione è uguale a

$$\sum_{i=1}^k \dim V_{\lambda_i}$$

ma per la nostra ipotesi

$$\sum_{i=1}^k \dim V_{\lambda_i} = \sum_{i=1}^k a_i$$

che è uguale al grado del polinomio caratteristico P_T , e dunque a $n = \dim V$. Allora

$$\dim V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k} = V$$

e T è diagonalizzabile come volevamo dimostrare. ■

• **Breve nota sull'ultima lezione e segnalazione di errore.**

Nell'ultima lezione abbiamo affrontato due argomenti:

- 1) Diagonalizzabilità di endomorfismi simmetrici.
- 2) Cenni su forme bilineari e prodotti scalari.

Vorrei correggere degli errori commessi nello spiegare il punto 1). Il punto 2) invece va bene.

Nel punto 1) abbiamo enunciato questo teorema (dimenticate l'enunciato in forma più generale):

Teorema 0.50. *Consideriamo un endomorfismo $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ la cui matrice associata rispetto alla base standard è simmetrica. Allora T è diagonalizzabile.*

Il teorema è VERO, la dimostrazione invece non va bene. L'algoritmo che abbiamo proposto diagonalizza la matrice ma non ci garantisce che il cambio di base in partenza e in arrivo sia lo stesso....

Dunque di questa lezione ricordate solo l'enunciato del teorema e il punto 2), ossia i cenni sulle forme bilineari e i prodotti scalari.

Come parte facoltativa, chi vuole può rimediare al mio errore e trovare la dimostrazione giusta (potete venire a chiedermela).