Appunti delle lezioni del modulo di Algebra Lineare per il corso MDAL

Pietro Di Martino e Giovanni Gaiffi

Indice

Capitolo 1. Spazi vettoriali: introduzione 5
1. Definizione di spazio vettoriale e primi esempi 5
2. Sottospazi vettoriali 8
3. Intersezione e somma di sottospazi vettoriali 12
4. Base di uno spazio vettoriale 16
5. Applicazioni lineari 23
6. Matrici e vettori in colonna per rappresentare applicazioni lineari e elementi di uno spazio vettoriale 25
7. Esercizi di fine capitolo 32

Capitolo 2. Il rango delle applicazioni lineari e la riduzione a scalini delle matrici 37
1. Le operazioni elementari sulle colonne 37
2. La riduzione a scalini per colonne applicata allo studio delle basi 43
3. Il teorema della dimensione del nucleo e dell'immagine di una applicazione lineare 47
4. La riduzione a scalini per colonna vista come moltiplicazione per matrici invertibili 49
5. La riduzione a scalini per righe 52
6. Ancora sulla riduzione a scala di una matrice e lo studio delle applicazioni lineari 56
7. Altri esercizi 59

Capitolo 3. Sistemi lineari 65
1. Risolvere un sistema usando le operazioni elementari di riga 65
2. Altri esercizi 70
3. Come funziona Google 77

Capitolo 4. La formula di Grassmann 83
1. La formula di Grassmann per le intersezioni e le somme di sottospazi. 83
2. Un metodo per calcolare l'intersezione di due sottospazi vettoriali 85
3. Somma diretta di sottospazi 87
4. Altri esercizi 89

Capitolo 5. Applicazioni lineari e matrici invertibili 91
1. Endomorfismi lineari invertibili 91
2. Il metodo per trovare l'inversa (se esiste) di una matrice quadrata 92
3. Cambiamento di base nel caso degli endomorfismi lineari 95
4. Altri esercizi 98

Capitolo 6. Informazioni sul determinante 101

1. Definizione del determinante di una matrice quadrata 101
2. Il determinante e il calcolo del rango di una matrice 102
3. Il teorema di Binet 104
4. Proprietà del determinante rispetto alle mosse di riga e di colonna 105
5. Altri esercizi 105

Capitolo 7. Diagonalizzazione di endomorfismi lineari 107
1. Autovalori e autovettori di un endomorfismo lineare 107
2. Il polinomio caratteristico di un endomorfismo 109
3. Una strategia per scoprire se un endomorfismo è diagonalizzabile 112
4. Il criterio della molteplicità algebrica e della molteplicità geometrica 115
5. Esempi 118
6. Altri esercizi 121

Capitolo 8. Prodotti scalari e spazi euclidei 125
1. Prodotto scalare 125
2. Ortogonalità 126
3. La disuguaglianza di Cauchy-Schwarz 129
4. Sottospazi ortogonali 130
5. Esercizi 131

Capitolo 9. Qualche appunto sul teorema spettrale 133
1. Introduzione al teorema spettrale 133
2. Endomorfismi simmetrici definiti positivi o negativi 137
3. Esercizi 138

Bibliografia 141

CAPITOLO 1

Spazi vettoriali: introduzione

1. Definizione di spazio vettoriale e primi esempi

In questo e nei prossimi capitoli concentreremo la nostra attenzione sulla struttura matematica di spazio vettoriale V su un campo $\mathbb{K}$. Tale struttura sarà definita avendo in mente il modello dei vettori applicati del piano[1] e le proprietà delle operazioni tra vettori applicati che abbiamo discusso.

La definizione di spazio vettoriale dunque sarà a partire da un insieme V non vuoto, su cui si possa introdurre due operazioni: una tra due elementi di V, che sarà detta **somma vettoriale** o più semplicemente **somma**; e l'altra tra un elemento di V e uno di $\mathbb{K}$, che sarà detta **moltiplicazione per scalare** o **prodotto esterno**[2]. Chiameremo **vettori** gli elementi di uno spazio vettoriale V e **scalari** gli elementi del campo $\mathbb{K}$.

Definizione 1.1. Uno **spazio vettoriale su un campo** $\mathbb{K}$ è un insieme V su cui sono definite la somma (o addizione) fra due elementi di V (il cui risultato è ancora un elemento di V, si dice che V è chiuso per la somma), e il prodotto di un elemento del campo $\mathbb{K}$ per un elemento di V (il cui risultato è un elemento di V si dice che V è chiuso per il prodotto con elementi di $\mathbb{K}$) che verificano le seguenti proprietà[3]:

- $\forall\ u, v, w \in V$ 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 neutro per l'addizione).

---

[1]In quel caso il campo numerico $\mathbb{K}$ che abbiamo considerato è il campo $\mathbb{R}$, che sarà il campo che utilizzeremo maggiormente anche nel seguito. Visto che però dal punto di vista delle definizioni il riferirsi ad un generico campo $\mathbb{K}$ non complica le cose, e che comunque talvolta vedremo anche spazi vettoriali su campi diversi da $\mathbb{R}$, preferiamo dare la definizione più generale di spazio vettoriale su un campo $\mathbb{K}$ e non limitarsi agli spazi vettoriali sul campo $\mathbb{R}$.

[2]Il nome prodotto esterno ricorda il fatto che, a differenza della somma vettoriale che è una operazione tra elementi di V, la moltiplicazione è tra un elemento di V ed un elemento di $\mathbb{K}$.

[3]Indicheremo il risultato della somma tra due vettori v, w con $v+w$ e il risultato del prodotto scalare tra lo scalare λ di $\mathbb{K}$ e il vettore v di V con $\lambda \cdot v$ o più frequentemente, omettendo il simbolo $\cdot$, con λv.

- $\forall\ v \in V$ esiste un elemento w in V tale che $v + w = O$ (esistenza dell'opposto per l'addizione).

- $\forall \lambda, \mu \in \mathbb{K},\ \forall\ v, w \in V$ vale $\lambda(v+w) = \lambda v + \lambda w$ e anche $(\lambda+\mu)v = \lambda v + \mu v$ (proprietà distributive della moltiplicazione per scalare).

- $\forall \lambda, \mu \in \mathbb{K},\ \forall\ v \in V$ vale $(\lambda\mu)v = \lambda(\mu v)$ (proprietà associativa della moltiplicazione per scalare).

- $\forall\ v \in V$ vale $1v = v$ (proprietà di esistenza dell'invariante moltiplicativo).

Esercizio 1.2. L'elemento neutro O della somma vettoriale è unico.

Supponiamo che in uno spazio vettoriale V su un campo $\mathbb{K}$ esistano due elementi neutri per la somma vettoriale O e O', considerando la somma tra O e O' e sfruttando da una parte il fatto che O è elemento neutro e dall'altra che anche O' lo è, si dimostra $O = O'$, infatti:

$$O \underbrace{=}_{\text{O}' \text{ è el.neutro somma}} O + O' \underbrace{=}_{\text{O è el.neutro somma}} O'$$

Esercizio 1.3. Dimostrare che dato uno spazio vettoriale V sul campo $\mathbb{K}$ e un elemento v di V, l'opposto per l'addizione di v è unico. Indicheremo l'opposto di v con $-v$.

Osservazione 1.4. È molto importante sottolineare la differenza tra l'elemento neutro della somma vettoriale $\mathbf{O}$ e lo 0 elemento neutro della somma di $\mathbb{K}$: in particolare $\mathbf{O}$ è un vettore, appartiene a V, mentre 0 è uno scalare di $\mathbb{K}$. Vedremo meglio più avanti, quando faremo degli esempi di spazi vettoriali, la distinzione tra questi due elementi. Il seguente esercizio fornisce una relazione tra i due elementi: il prodotto scalare tra lo scalare 0 e qualsiasi vettore di V restituisce l'elemento neutro della somma vettoriale $\mathbf{O}$.

Esercizio 1.5. Dalla proprietà di spazio vettoriale, dimostrare che dato V spazio vettoriale su un campo $\mathbb{K}$ vale la seguente legge di annullamento del prodotto scalare:

$$\forall v \in V \quad 0v = O$$

Dato $v \in V$ consideriamo $(0+0)v$:

$$0v \underbrace{=}_{\text{0 è el.neutro somma in } \mathbb{K}} (0+0)v \underbrace{=}_{\text{prop.distr.}} 0v + 0v$$

Da questa uguaglianza (sommando da entrambe le parti per l'opposto di $0v$) segue che $0v$ è l'elemento neutro della somma vettoriale. Dall'esercizio 1.2 segue che $0v = O$.

Esercizio 1.6. Dimostrare che, dato V spazio vettoriale su un campo $\mathbb{K}$, il prodotto scalare di un qualsiasi vettore v per lo scalare -1 è uguale all'opposto di v.

Basta osservare che:

$$O \underbrace{=}_{esercizio\ 1.5} 0v = (-1+1)v \underbrace{=}_{\text{prop.dist.}} -1v + 1v \underbrace{=}_{\text{invariante moltiplicativo}} -1v + v$$

Ovvero $-1v$ è l'opposto per l'addizione di v. Dall'esercizio 1.3 segue che $-1v = -v$.

Cominciamo adesso a vedere alcuni esempi di spazio vettoriale:

Esempio 1.7. Ogni campo $\mathbb{K}$ è uno spazio vettoriale su $\mathbb{K}$ stesso con le operazioni di somma vettoriale e prodotto per scalare che sono definite identiche alle operazioni di somma e prodotto del campo. In particolare dunque $\mathbb{R}$ è uno spazio vettoriale su $\mathbb{R}$, così come $\mathbb{Q}$ è uno spazio vettoriale su $\mathbb{Q}$.

Esempio 1.8. Abbiamo praticamente[4] visto che $\mathbb{R}^2 = \{(a,b)|a,b \in \mathbb{R}\}$, l'insieme delle coppie di numeri reali (che possiamo *vedere* geometricamente come il piano) è uno spazio vettoriale su $\mathbb{R}$ con le operazioni di somma vettoriale e prodotto scalare definite come segue[5]:

$$(a,b) +_{\mathbb{R}^2} (c,d) \underbrace{=}_{def} (a+c, b+d)$$

$$\lambda \cdot_{\mathbb{R}^2} (a,b) \underbrace{=}_{def} (\lambda \cdot a, \lambda \cdot b)$$

Osservazione 1.9. Osserviamo, nel caso di $\mathbb{R}^2$ come spazio vettoriale su $\mathbb{R}$, la differenza tra il vettore $\mathbf{O}$ elemento neutro della somma vettoriale e lo 0 scalare. Lo $\mathbf{O}$ elemento neutro della somma è l'elemento $(0,0)$ di $\mathbb{R}^2$.

Esercizio 1.10. Dimostrare in generale che, dato un campo $\mathbb{K}$, l'insieme $\mathbb{K}^n$ delle n-uple ordinate $(a_1, a_2, \ldots, a_n)$ di elementi di $\mathbb{K}$, che molto spesso rappresenteremo in colonna

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

anziché in riga, è uno spazio vettoriale su $\mathbb{K}$ con le operazioni di somma vettoriale definite come segue:

(1) La somma fra vettori di $\mathbb{K}^n$ è definita da:

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

[4]Scriviamo praticamente perché in realtà non abbiamo introdotto le coordinate, ma basta pensare i vettori applicati in un riferimento cartesiano, ed identificare ogni vettore con le coordinate cartesiane del suo secondo estremo.

[5]Quanto scritto nelle formule è la traduzione in termini di coordinate delle definizioni date nella sezione precedente in termini di vettori geometrici.

(2) Il prodotto tra un vettore di $\mathbb{K}^n$ e uno scalare di $\mathbb{K}$ è definito da:

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

Nel nostro corso l'esempio di spazio vettoriale che studieremo di più è quello di $\mathbb{R}^n$ come spazio vettoriale su $\mathbb{R}$ con le operazioni definite come sopra. Si tratta, come insieme, del prodotto cartesiano $\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ (dove $\mathbb{R}$ compare n volte), i cui elementi, ossia i vettori, sono le liste ordinate formate da n numeri reali.

Esercizio 1.11. $\mathbb{R}^n$ è anche uno spazio vettoriale su $\mathbb{Q}$?

Si può rispondere all'esercizio mostrando, più in generale, che se V è uno spazio vettoriale su un campo $\mathbb{K}$ e $\mathbb{F}$ è un sottocampo di $\mathbb{K}$ allora V è uno spazio vettoriale su $\mathbb{F}$. L'osservazione chiave è che se le proprietà del prodotto scalare valgono per tutti gli elementi di $\mathbb{K}$, *a maggior ragione* varranno per tutti gli elementi di un sottoinsieme ($\mathbb{F}$) degli elementi di $\mathbb{K}$.

Esempio 1.12. Un altro esempio di spazio vettoriale su un campo $\mathbb{K}$ è l' insieme dei polinomi $\mathbb{K}[x]$ con la somma tra polinomi e il prodotto tra polinomi e costanti di $\mathbb{K}$ definiti nel modo usuale. Ovvero il polinomio somma di $p(x)$ e $q(x)$ è quello il cui coefficiente di grado n (per ogni $n \in \mathbb{N}$) è la somma dei coefficienti di grado n dei polinomi $p(x)$ e $q(x)$. E il polinomio prodotto di $k \in \mathbb{K}$ e $p(x)$ è il polinomio che ha come coefficiente di grado n (per ogni $n \in \mathbb{N}$) k volte il coefficiente di grado n di $p(x)$.

2. Sottospazi vettoriali

Dato uno spazio vettoriale V sul campo $\mathbb{K}$ viene abbastanza naturale definire un sottospazio vettoriale di V come segue:

Definizione 1.13. Un **sottospazio vettoriale** W di V è un sottoinsieme $W \subseteq V$ che (rispetto alle operazioni $+$ e $\cdot$ che rendono V uno spazio vettoriale su $\mathbb{K}$) è uno spazio vettoriale su $\mathbb{K}$.

Esempio 1.14. Dato uno spazio vettoriale V su un campo $\mathbb{K}$, V e l'insieme $\{O\}$ sono sempre sottospazi di V (qualunque sia V).

Definizione 1.15. Chiameremo **sottospazio proprio** (o non banale) di V un qualsiasi sottospazio vettoriale di V che sia diverso da V e dal sottospazio $\{O\}$.

Una domanda che sorge abbastanza spontanea è la seguente: dato uno spazio vettoriale V su $\mathbb{K}$ e un suo sottoinsieme W, per provare che W è un sottospazio vettoriale su V dobbiamo verificare per W tutte le proprietà di spazio vettoriale o non tutte le verifiche sono necessarie? La risposta è che in realtà basta verificare che W contiene lo O e che sia chiuso per somma vettoriale e prodotto scalare:

Proposizione 1.16. *Dato V spazio vettoriale su $\mathbb{K}$ e W sottoinsieme di V, W è sottospazio vettoriale di V (rispetto alle operazioni $+$ e $\cdot$ che rendono V uno spazio vettoriale su $\mathbb{K}$) se e solo se:*

(1) *Il vettore O appartiene a W.*
(2) *Per ogni $u, v \in W$ vale $u + v \in W$.*
(3) *Per ogni $k \in \mathbb{K}$ e per ogni $u \in W$ vale $ku \in W$.*

Dim. Se W è un sottospazio vettoriale di V deve verificare tutte le proprietà di spazio vettoriale su $\mathbb{K}$ e dunque le tre proprietà elencate devono valere.

Verifichiamo il viceversa, ovvero che se valgono le tre proprietà esplicitate nella proposizione (esistenza dell'elemento neutro della somma e chiusura per somma vettoriale e moltiplicazione per scalare) allora W è sottospazio.
Dobbiamo provare che valgono tutte le altre proprietà che definiscono uno spazio vettoriale. Dal fatto che valgono in V, segue immediatamente che valgono in W la proprietà associativa e commutativa dell'addizione, la proprietà distributiva e associativa della moltiplicazione per scalare, e l'esistenza dell'invariante moltiplicativo (infatti questo appartiene al campo $\mathbb{K}$ ed ha la proprietà che moltiplicato per ogni elemento v di V, e dunque *a maggior ragione* di W, restituisce v).
Rimane dunque solo da provare l'esistenza in W, per ogni elemento w di W, dell'opposto per l'addizione. Ma noi sappiamo (esercizio 1.3) che l'opposto di w è il risultato del prodotto scalare tra -1 e w, ed essendo W chiuso per prodotto scalare, $-1w$ è un elemento di W. $\square$

Osservazione 1.17. Su alcuni libri di testo si può trovare al posto della proprietà che il vettore O appartenga a W il fatto che W sia diverso dal vuoto, cioè contenga almeno un elemento. In effetti se v è un elemento di W, allora dal fatto che W è chiuso per prodotto con scalari (terza proprietà) e dal fatto che $0 \cdot v = O$ si ha che $O \in W$. Viceversa, per definizione, se $O \in W$ allora W non è vuoto. Dunque richiedere che O appartenga a W equivale a richiedere che W non sia vuoto.

Esempio 1.18. Consideriamo $\mathbb{R}^2$ come spazio vettoriale su $\mathbb{R}$ e cerchiamo di capire quali sono i sottospazi vettoriali di $\mathbb{R}^2$. Sicuramente ci sono quelli banali ovvero $\mathbb{R}^2$ stesso e il sottospazio costituito dall'elemento neutro per la somma O, che sappiamo essere l'origine $(0, 0)$ del piano.
La Figura 1 mostra un esempio di un sottoinsieme di $\mathbb{R}^2$ che non è uno spazio vettoriale Come ulteriore esempio, osserviamo che la circonferenza $x^2 + y^2 = 1$ in $\mathbb{R}^2$ non è un sottospazio vettoriale. La circonferenza infatti non contiene lo O. Ma anche considerando l'unione tra la circonferenza e lo O non avremmo uno spazio vettoriale, infatti tale insieme non sarebbe chiuso né per somma vettoriale, né per prodotto scalare.

Esercizio 1.19. Dimostrare che una circonferenza nel piano (qualunque sia il raggio e qualunque sia il centro) non è un sottospazio vettoriale del piano.

Si può dimostrare che tutti e soli i sottospazi vettoriali propri di $\mathbb{R}^2$ sono le rette passanti per l'origine O (vedi Figura 2).

Infatti una retta r passante per l'origine del piano è caratterizzata dal numero reale k che identifica la *pendenza* della retta stessa (la retta avrà l'equazione $y = kx$). L'insieme r dei punti appartenenti alla retta è dunque il seguente:

$$r = \{(x, kx) | x \in \mathbb{R}\}$$

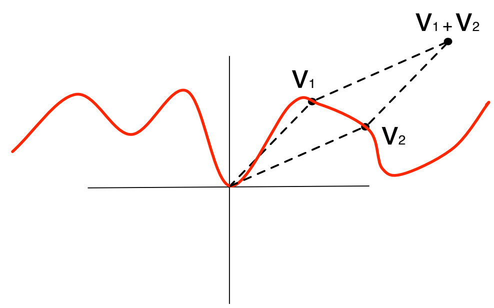


FIGURA 1. I due vettori V_1 e V_2 appartengono alla curva disegnata in rosso, ma la loro somma $V_1 + V_2$ non appartiene alla curva: dunque la curva non è un sottospazio vettoriale di $\mathbb{R}^2$. Notare che il vettore somma si può ottenere tramite la ben nota *regola del parallelogramma*.

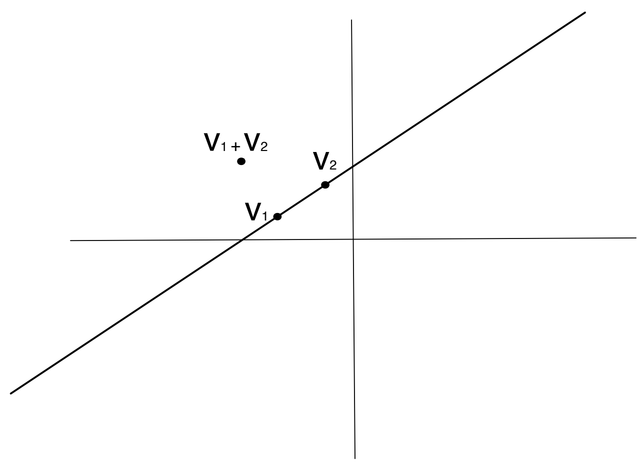


FIGURA 2. La somma di due vettori che giacciono su una retta non passante per l'origine non appartiene alla retta.

Ora osserviamo che $(0, 0) \in r$ e che per ogni coppia $(x_1, kx_1), (x_2, kx_2)$ di punti di r e per ogni $h \in \mathbb{R}$ si ha:

$$(x_1, kx_1) + (x_2, kx_2) = (x_1 + x_2, kx_1 + kx_2) = (x_1 + x_2, k(x_1 + x_2)) \in r$$
$$h(x_1, kx_1) = (hx_1, hkx_1) = (hx_1, k(hx_1)) \in r$$

Abbiamo dunque mostrato che tutte le rette passanti per l'origine sono sottospazi propri di $\mathbb{R}^2$. Viceversa dobbiamo mostrare che se V è un sottospazio proprio di $\mathbb{R}^2$, allora V è una retta passante per l'origine. Osserviamo che se V è sottospazio proprio, esiste un vettore $v \neq O$ in V. Ora, la chiusura di V per prodotto scalare ci dice che tutta la retta r che unisce O con v sta in V. Ci resta da dimostrare che se esistesse w in V non appartenente alla retta r, allora V sarebbe tutto $\mathbb{R}^2$. Infatti se esistesse w in $V - r$, allora (per lo stesso ragionamento seguito per v) in V sarebbe contenuta tutta la retta s passante per O e w. Ora V deve essere chiuso anche per somma vettoriale: ci resta dunque da dimostrare che, dato un qualsiasi z vettore di $\mathbb{R}^2$, si possono scegliere opportunamente un vettore w' su s e un vettore v' su r in modo tale che $z = w' + v'$. Lasciamo, per ora, questo punto al lettore (e alla sua intuizione geometrica); forniremo in seguito ulteriori strumenti per dimostrarlo. Si tratta, di trovare le coordinate di z rispetto alle rette s e r (prendendo come unit di misura su s e r le lunghezze rispettivamente di w' e v').

Esercizio 1.20. Dimostrare che l'anello dei polinomi $\mathbb{K}[x]$ è uno spazio vettoriale su $\mathbb{K}$ con la somma tra polinomi e il prodotto tra polinomi e costanti di $\mathbb{K}$ definiti nel modo usuale[6].

Esercizio 1.21. Dimostrare che l'insieme $\mathbb{K}^{\leq n}[x]$, i cui elementi sono il polinomio 0 e i polinomi a coefficienti in $\mathbb{K}$ di grado minore o uguale ad n, è un sottospazio vettoriale di $\mathbb{K}[x]$ qualsiasi sia $n \in \mathbb{N}$.[7]

Svolgimento. Qualsiasi sia n, abbiamo per definizione di $\mathbb{K}^{\leq n}[x]$ che il polinomio nullo appartiene a $\mathbb{K}^{\leq n}[x]$. Dalla Proposizione 1.16 segue che rimane da verificare che $\mathbb{K}^{\leq n}[x]$ sia chiuso per somma e per prodotto per scalare, ma questo segue banalmente dalle proprietà del grado, infatti:

- $deg(p(x) + q(x)) \leq max(deg(p(x)), deg(q(x)))$ quindi se $p(x)$ e $q(x)$ hanno grado minore o uguale di n, in quanto appartenenti a $\mathbb{K}^{\leq n}[x]$, allora $p(x)+q(x)$ ha grado minore o uguale a n e quindi anch'esso appartiene a $\mathbb{K}^{\leq n}[x]$.
- Sia $p(x) \in \mathbb{K}^{\leq n}[x]$ un polinomio non nullo. Se $\lambda \in \mathbb{K}$ è diverso da zero si ha che:
$$deg(\lambda \cdot p(x)) = deg(p(x)) \leq n$$
e quindi $\lambda \cdot p(x) \in \mathbb{K}^{\leq n}[x]$. Se $\lambda = 0$ vale $\lambda \cdot p(x) = 0$ e dunque anche in questo caso $\lambda \cdot p(x) \in \mathbb{K}^{\leq n}[x]$. Se $p(x) = 0$ vale $\lambda \cdot p(x) = 0$ per ogni $\lambda \in \mathbb{K}$, dunque è sempre vero che $\lambda \cdot p(x) \in \mathbb{K}^{\leq n}[x]$.

Esempio 1.22. Consideriamo il sottoinsieme L di $\mathbb{K}[x]$ che contiene tutti e soli i polinomi che hanno 1 come radice, ovvero:
$$L = \{p(x) \in \mathbb{K}[x] \mid p(1) = 0\}$$
Verifichiamo che L è un sottospazio vettoriale di $\mathbb{K}[x]$.

(1) Il polinomio 0, che è il vettore O di $\mathbb{K}[x]$, appartiene a L, infatti ha 1 come radice (addirittura ogni elemento di $\mathbb{K}$ è una radice di 0).
(2) Se $p(x), q(x) \in L$ allora $(p + q)(x)$ appartiene a L, infatti:
$$(p+q)(1) \underbrace{=}_{definizione\ di\ somma\ tra\ polinomi} p(1) + q(1) \underbrace{=}_{p(x)\in L,\ q(x)\in L} 0 + 0 = 0$$
(3) Se $p(x) \in L$ e $k \in \mathbb{K}$ allora $kp(x) \in L$, infatti:
$$(kp)(1) = k\ p(1) \underbrace{=}_{p(x)\in L} k \cdot 0 = 0$$

Esercizio 1.23. Dire quali dei seguenti sottoinsiemi di $\mathbb{R}^{\leq n}[x]$ sono sottospazi vettoriali di $\mathbb{R}^{\leq n}[x]$:

(1) $V_1 = \{p(x) \in \mathbb{R}^{\leq n}[x] | p(2) = 0\}$

---

[6]Dedicheremo più avanti un capitolo allo studio degli anelli di polinomi. In questo capitolo si utilizza solo la struttura di spazio vettoriale.

[7]Abbiamo dovuto aggiungere il polinomio 0 perché per tale polinomio non si definisce un grado, come vedremo più avanti, dunque non rientra fra i polinomi di grado minore o uguale a n.

(2) $V_2 = \{p(x) \in \mathbb{R}^{\leq n}[x] | p(1) = 1\}$
(3) $V_3 = \{p(x) \in \mathbb{R}^{\leq n}[x] | \sum_{i=0}^{n} a_i x^i, \quad a_i \in \mathbb{Z}\}$
(4) $V_4 = \{p(x) \in \mathbb{R}^{\leq n}[x] | p(1) = -p(2)\}$
(5) $V_5 = \{p(x) \in \mathbb{R}^{\leq n}[x] | \sum_{i=0}^{[n/2]} a_{2i} x^{2i}\}$ Dove con $[n/2]$ indichiamo la parte intera di $n/2$.

Svolgimento. Analizziamo punto per punto le richieste dell'esercizio

(1) La dimostrazione che V_1 è un sottospazio vettoriale di $\mathbb{R}^{\leq n}[x]$ ricalca quella vista nell'esempio 1.22. In generale l'insieme dei polinomi di $\mathbb{K}[x]$ che hanno una radice k in $\mathbb{K}$ è dunque uno spazio vettoriale; come vedremo questo insieme coincide con l'insieme dei polinomi di $\mathbb{K}[x]$ che sono divisibili per $x - k$.
(2) V_2 è un insieme che non verifica nessuna delle tre proprietà che definiscono un sottospazio vettoriale, ma per dimostrare che non è un sottospazio vettoriale basta osservare che una di esse non vale, per esempio basta osservare che il polinomio 0 non appartiene a V_2. Infatti tale polinomio valutato in qualsiasi elemento vale sempre 0 e non 1.
(3) Il polinomio 0 appartiene a V_3 e la somma di due polinomi a coefficienti interi è un polinomio a coefficienti interi: quindi V_3 è chiuso per la somma. *Sfortunatamente* però V_3 non è chiuso per prodotto per scalare; infatti sia $p(x) \in V_3$ non zero, se scegliamo un qualsiasi numero reale a che non sia intero e nemmeno razionale (per essere sicuri che non ci siano semplificazioni), per esempio $\sqrt{2}$, allora $a \cdot p(x)$ è un polinomio non a coefficienti interi. Quindi V_3 non è sottospazio vettoriale di $\mathbb{R}^{\leq n}[x]$.
(4) Si osserva subito che 0 appartiene a V_4. Siano $p(x)$ e $g(x)$ polinomi di V_4 allora valutiamo la loro somma e la moltiplicazione di uno dei due per uno scalare $r \in \mathbb{R}$ e verifichiamo se continua a valere la proprietà che definisce V_4:

$$(p+g)(1) = p(1) + g(1) \underbrace{=}_{p(x)\in V_4,\ g(x)\in V_4} -p(2) - g(2) = -(p+g)(2)$$

$$(r \cdot p)(1) = r \cdot p(1) \underbrace{=}_{p(x)\in V_4} r \cdot (-p(2)) = -(r \cdot p)(2)$$

Quindi V_4 è un sottospazio vettoriale di $\mathbb{R}^{\leq n}[x]$.
(5) V_5 è il sottoinsieme di $\mathbb{R}^{\leq n}[x]$ dei polinomi che hanno tutti i coefficienti dei termini di grado dispari uguali a zero. Dunque il polinomio 0 appartiene a V_5. Ora osserviamo che la somma tra due polinomi è definita facendo le somme tra monomi dello stesso grado, quindi se sommiamo due polinomi con solo monomi di grado pari otteniamo un polinomio formato solo da monomi di grado pari. È banale osservare che V_5 è chiuso anche per prodotto scalare e quindi è un sottospazio vettoriale di $\mathbb{R}^{\leq n}[x]$.

3. Intersezione e somma di sottospazi vettoriali

Dati due sottospazi vettoriali U e W di uno spazio vettoriale V, siamo interessati a cercare di caratterizzare, se esistono, il più piccolo sottospazio vettoriale di V che contenga sia U che W, e il più grande sottospazio vettoriale di V contenuto sia in U che in W.

Per questo secondo caso la prima idea che viene in mente è quella di considerare l'intersezione insiemistica tra U e W. Infatti se $U \cap W$ è un sottospazio vettoriale di V sicuramente è contenuto in U e in W (per definizione di intersezione) e inoltre non ci può essere nessun sottospazio H di U e di W che contiene $U \cap W$ (altrimenti esisterebbe $h \in H$ che non appartiene a $U \cap W$, ma questo significa che h non appartiene ad almeno uno tra U e W e di conseguenza che H non è contenuto in almeno uno dei due spazi vettoriali).
La seguente proposizione ci assicura che dati due sottospazi vettoriali U e W di uno spazio vettoriale V, $U \cap W$ è un sottospazio vettoriale di V:

Proposizione 1.24. *Sia V uno spazio vettoriale su un campo $\mathbb{K}$, U e W due sottospazi di V, allora $U \cap W$ è un sottospazio vettoriale di V.*

Dim. Dobbiamo mostrare che $U \cap W$ verifica le proprietà della definizione 1.13:

(1) $O \in U \cap W$, infatti essendo U e W due sottospazi, certamente $O \in U$ e $O \in W$.
(2) Siano $v_1, v_2 \in U \cap W$ allora:

$$\begin{cases} \underbrace{v_1 + v_2 \in U}_{U \text{ è sottospazio}} \\ \underbrace{v_1 + v_2 \in W}_{W \text{ è sottospazio}} \end{cases} \Rightarrow \quad v_1 + v_2 \in U \cap W$$

(3) Sia $v \in U \cap W$ allora per ogni $\lambda \in \mathbb{K}$ si ha:

$$\begin{cases} \underbrace{\lambda \cdot v \in U}_{U \text{ è sottospazio}} \\ \underbrace{\lambda \cdot v \in W}_{W \text{ è sottospazio}} \end{cases} \Rightarrow \quad \lambda \cdot v \in U \cap W \qquad \square$$

Esercizio 1.25. Dire quali dei seguenti sottoinsiemi di $\mathbb{R}^{\leq n}[x]$ sono sottospazi vettoriali di $\mathbb{R}^{\leq n}[x]$:

(1) $V_1 = \{p(x) \in \mathbb{R}^{\leq n}[x] | p(2) = 0\}$
(2) $V_2 = \{p(x) \in \mathbb{R}^{\leq n}[x] | p(1) = 1\}$
(3) $V_3 = \{p(x) \in \mathbb{R}^{\leq n}[x] | \sum_{i=0}^{n} a_i x^i, \quad a_i \in \mathbb{Z}\}$
(4) $V_4 = \{p(x) \in \mathbb{R}^{\leq n}[x] | p(1) = -p(2)\}$
(5) $V_5 = \{p(x) \in \mathbb{R}^{\leq n}[x] | \sum_{i=0}^{[n/2]} a_{2i} x^{2i}\}$ Dove con $[n/2]$ indichiamo la parte intera di $n/2$.

Analizziamo punto per punto le richieste dell'esercizio

(1) La dimostrazione che V_1 è un sottospazio vettoriale di $\mathbb{R}^{\leq n}[x]$ ricalca quella vista nell'esempio 1.22. È ovvio che le proprietà dimostrate non dipendono dalla radice scelta. In generale l'insieme dei polinomi di $\mathbb{K}[x]$ che hanno una radice k in $\mathbb{K}$ è dunque uno spazio vettoriale; come vedremo questo insieme equivale all'insieme dei polinomi di $\mathbb{K}[x]$ che sono divisibili per $x - k$.
(2) V_2 è un insieme che non verifica nessuna delle tre proprietà che definiscono un sottospazio vettoriale, ma per dimostrare che non è un sottospazio vettoriale basta osservare che una di esse non vale, per esempio basta osservare che il polinomio 0 non appartiene a V_2. Infatti tale polinomio

valutato in qualsiasi elemento vale sempre 0 e non potrà mai essere uguale a 1.

(3) Il polinomio 0 appartiene a V_3 e la somma di due polinomi a coefficienti interi è un polinomio a coefficienti interi: quindi V_3 è chiuso per la somma. *Sfortunatamente* però V_3 non è chiuso per prodotto per scalare; infatti sia $p(x) \in V_3$ non zero, se scegliamo un qualsiasi numero reale a che non sia intero e nemmeno razionale (per essere sicuri che non ci siano semplificazioni), per esempio $\sqrt{2}$, allora $a \cdot p(x)$ è un polinomio non a coefficienti interi. Quindi V_3 non è sottospazio vettoriale di $\mathbb{R}^{\leq n}[x]$.

(4) Si osserva subito che 0 appartiene a V_4. Siano $p(x)$ e $g(x)$ polinomi di V_4 allora valutiamo la loro somma e la moltiplicazione di uno dei due per uno scalare $r \in \mathbb{R}$ e verifichiamo se continua a valere la proprietà che definisce V_4:

$$(p+g)(1) = p(1) + g(1) \underbrace{=}_{p(x)\in V_4,\ g(x)\in V_4} -p(2) - g(2) = -(p+g)(2)$$

$$(r \cdot p)(1) = r \cdot p(1) \underbrace{=}_{p(x)\in V_4} r \cdot (-p(2)) = -(r \cdot p)(2)$$

Quindi V_4 è un sottospazio vettoriale di $\mathbb{R}^{\leq n}[x]$.

(5) V_5 è il sottoinsieme di $\mathbb{R}^{\leq n}[x]$ dei polinomi che hanno tutti i coefficienti dei termini di grado dispari uguali a zero. Dunque il polinomio 0 appartiene a V_5. Ora osserviamo che la somma tra due polinomi è definita facendo le somme tra monomi dello stesso grado, quindi se sommiamo due polinomi con solo monomi di grado pari otteniamo un polinomio formato solo da monomi di grado pari. È banale osservare che V_5 è chiuso anche per prodotto scalare e quindi è un sottospazio vettoriale di $\mathbb{R}^{\leq n}[x]$.

Dati due sottospazi vettoriali U e W di uno spazio vettoriale V, siamo interessati a cercare di caratterizzare, se esistono, il più piccolo sottospazio vettoriale di V che contenga sia U che W, e il più grande sottospazio vettoriale di V contenuto sia in U che in W.
Per questo secondo caso la prima idea che viene in mente è quella di considerare l'intersezione insiemistica tra U e W. Infatti se $U \cap W$ è un sottospazio vettoriale di V sicuramente è contenuto in U e in W (per definizione di intersezione) e inoltre non ci può essere nessun sottospazio H di U e di W che contiene $U \cap W$ (altrimenti esisterebbe $h \in H$ che non appartiene a $U \cap W$, ma questo significa che h non appartiene ad almeno uno tra U e W e di conseguenza che H non è contenuto in almeno uno dei due spazi vettoriali).
Effettivamente, la seguente proposizione, ci assicura che dati due sottospazi vettoriali U e W di uno spazio vettoriale V, $U \cap W$ è un sottospazio vettoriale di V:

Proposizione 1.26. *Sia V uno spazio vettoriale su un campo $\mathbb{K}$, U e W due sottospazi di V, allora $U \cap W$ è un sottospazio vettoriale di V.*

Dim. Dobbiamo mostrare che $U \cap W$ verifica le proprietà della definizione 1.13:

(1) $O \in U \cap W$, infatti essendo U e W due sottospazi, certamente $O \in U$ e $O \in W$.

(2) Siano $v_1, v_2 \in U \cap W$ allora:

$$\begin{cases} \underbrace{v_1 + v_2 \in U}_{U \text{ è sottospazio}} \\ \underbrace{v_1 + v_2 \in W}_{W \text{ è sottospazio}} \end{cases} \Rightarrow \quad v_1 + v_2 \in U \cap W$$

(3) Sia $v \in U \cap W$ allora per ogni $\lambda \in \mathbb{K}$ si ha:

$$\begin{cases} \underbrace{\lambda \cdot v \in U}_{U \text{ è sottospazio}} \\ \underbrace{\lambda \cdot v \in W}_{W \text{ è sottospazio}} \end{cases} \Rightarrow \quad \lambda \cdot v \in U \cap W \qquad \square$$

A questo punto andiamo alla caccia del più piccolo sottospazio contenente i sottospazi U e W di uno spazio vettoriale V. Anche qui verrebbe naturale considerare l'unione insiemistica: se infatti $U \cup W$ fosse sempre un sottospazio di V, sarebbe sicuramente il più piccolo sottospazio contenente sia U che W (provarlo per esercizio).
Sfortunatamente in generale non è vero che $U \cup W$ è un sottospazio vettoriale di V.

Esempio 1.27. Provare ad esempio che se $V = \mathbb{R}^2$ e U e W sono due rette distinte passanti per O, allora $U \cup W$ non è un sottospazio di V.

Basta mostrare che presi $u \in U$ e $w \in W$, entrambi diversi dall'origine, $v + w$ non appartiene alla unione $U \cup W$. Perché è vero?

Quanto sopra ci comincia a suggerire la strada: dovrebbe infatti essere chiaro che il più piccolo sottospazio vettoriale di V che contiene sia U sia W deve necessariamente (per essere chiuso per la somma) contenere tutti gli elementi della forma $u + w$ dove $u \in U$ e $w \in W$. Consideriamo dunque il seguente insieme:

Definizione 1.28. Dati due sottospazi vettoriali U e W di uno spazio vettoriale V su $\mathbb{K}$, chiamo *somma* di U e W l'insieme

$$U + W = \{u + w \mid u \in U, w \in W\}$$

La seguente proposizione fornisce la risposta alla nostra ricerca del più piccolo sottospazio contenente sia U che W: è $U + W$ appena definito.

Proposizione 1.29. *Dati due sottospazi vettoriali U e W di uno spazio vettoriale V su $\mathbb{K}$, $U + W$ è un sottospazio vettoriale di V (ed è il più piccolo contenente U e W).*

Dim. O appartiene ad $U + W$, infatti appartiene sia ad U che a W dunque:

$$O = \underbrace{O}_{\in U} + \underbrace{O}_{\in W}$$

Ora dati $a \in \mathbb{K}$ e $x, y \in U + W$, per definizione di $U + W$ esistono u_1, u_2 in U e w_1, w_2 in W tali che: $x = u_1 + w_1$ e $y = u_2 + w_2$. Dunque:

$$x + y = (u_1 + w_1) + (u_2 + w_2) = \underbrace{(u_1 + u_2)}_{\in U} + \underbrace{(w_1 + w_2)}_{\in W} \in U + W$$

$$ax = a(u_1 + w_1) = \underbrace{au_1}_{\in U} + \underbrace{aw_1}_{\in W} \in U + W \qquad \square$$

4. Base di uno spazio vettoriale

Sia V uno spazio vettoriale su un campo $\mathbb{K}$. Per definizione di V, se $v_1, v_2, \ldots, v_n$ sono n vettori di V, allora per qualsiasi scelta di n elementi $k_1, k_2, \ldots, k_n$ (non necessariamente distinti) di $\mathbb{K}$ il vettore:

$$v = k_1 v_1 + \ldots + k_n v_n = \sum_{i=1}^{n} k_i v_i$$

appartiene a V, in quanto V è chiuso per somma vettoriale e prodotto per scalare.

Definizione 1.30. Dato un insieme di vettori $\{v_1, v_2, \ldots, v_k\}$ di V, spazio vettoriale sul campo $\mathbb{K}$, il vettore:

$$v = k_1 \cdot v_1 + \ldots + k_n \cdot v_n$$

con $\{k_1, k_2, \ldots, k_n\}$ scalari di $\mathbb{K}$, si dice una **combinazione lineare** dei vettori $\{v_1, v_2, \ldots, v_k\}$. I k_i sono detti **coefficienti** della combinazione lineare.

Esempio 1.31. Consideriamo lo spazio vettoriale $\mathbb{R}^3$ su $\mathbb{R}$ e i due vettori seguenti:

$$v_1 = \begin{pmatrix} 3 \\ -1 \\ 3 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

Allora il vettore v_3 seguente:

$$v_3 = \begin{pmatrix} 5 \\ -1 \\ 7 \end{pmatrix} = 1 \cdot \begin{pmatrix} 3 \\ -1 \\ 3 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

È una combinazione lineare dell'insieme dei vettori $\{v_1, v_2\}$ di coefficienti 1 e 2.

Definizione 1.32. Dati $\{v_1, v_2, \ldots, v_t\}$ vettori di uno spazio vettoriale V sul campo $\mathbb{K}$, si definisce **span** dei vettori $v_1, \ldots, v_t$ (e si indica con $Span(v_1, v_2, \ldots, v_t)$ o anche con $< v_1, v_2, \ldots, v_t >$) l'insieme di tutte le possibili combinazioni lineari dell'insieme di vettori $\{v_1, v_2, \ldots, v_t\}$.

Esercizio 1.33. Dimostrare che dato uno spazio vettoriale V sul campo $\mathbb{K}$, per ogni $t > 0$ e per ogni scelta di vettori $\{v_1, v_2, \ldots, v_t\}$ di V si ha che $Span(v_1, v_2, \ldots, v_t)$ è un sottospazio vettoriale[8] di V.

Particolarmente interessante è il caso in cui, scelti t vettori di V spazio vettoriale sul campo $\mathbb{K}$, si ha che $V = Span(v_1, \ldots v_t)$:

Definizione 1.34. Un insieme di vettori $\{v_1, v_2, \ldots, v_t\}$ di V per cui $V = Span(v_1, \ldots v_t)$ (ovvero per ogni $v \in V$, esistono degli scalari $a_1, a_2, \ldots, a_t$ tali che

$$a_1 v_1 + a_2 v_2 + \cdots + a_t v_t = v$$

, si dice **un insieme di generatori** di V. In tal caso si dice anche che i vettori $v_1, v_2, \ldots, v_k$ **generano** V.

[8] Non è detto che sia un sottospazio proprio.

L'esistenza di un sistema finito di generatori per uno spazio vettoriale V su un campo $\mathbb{K}$ è un fatto, come si può intuire, molto importante: si riduce infatti la descrizione di uno spazio vettoriale che potrà avere cardinalità infinita, alla lista di un numero finito di vettori (i generatori) dalle cui combinazioni lineari si possono ottenere tutti i vettori di V.
Dato un sistema di generatori $\{v_1, \dots, v_t\}$ di V sappiamo dunque che ogni v in V si può scrivere, con una opportuna scelta dei coefficienti $\{k_1, \dots, k_t\}$, come:

$$v = \sum_{i=1}^{t} k_i v_i \tag{4.1}$$

Ci chiediamo se tale scrittura è, in generale, unica (il che ci direbbe che, fissato il sistema di generatori, ogni vettore v di V è univocamente identificato e dunque determinato, dall'unica scelta di coefficienti per cui vale l'uguaglianza 4.1). In generale la risposta a questa domanda è no, come possiamo vedere dal seguente esempio:

Esempio 1.35. 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:

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

Esercizio 1.36. Dimostrare che i 4 vettori dell'esempio precedente generano $\mathbb{R}^3$, quindi che ogni vettore di $\mathbb{R}^3$ può essere scritto come combinazione lineare di questi 4 vettori.

Abbiamo dunque che, in generale, la scrittura di un vettore che appartiene allo Span di n vettori non è unica.

Ci chiediamo allora quale ulteriore condizione bisogna imporre sugli n vettori di cui consideriamo lo Span, 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 1.37. Si dice che un insieme finito di vettori $\{v_1, v_2, \dots, v_r\}$ è un **insieme di vettori linearmente indipendenti** se l'unico modo di scrivere il vettore O come combinazione lineare di questi vettori è con tutti i coefficienti nulli, ossia se

$$a_1 v_1 + a_2 v_2 + \cdots + a_r v_r = O <=> a_1 = a_2 = \cdots = a_r = 0$$

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 $\{v_1, v_2, \ldots, v_r\}$ è un **insieme di vettori linearmente dipendenti**).

Il concetto di lineare indipendenza di un insieme di vettori $A = \{v_1, \ldots, v_n\}$ in particolare equivale al fatto che nessuno dei vettori v_i si possa scrivere come combinazione lineare dell'insieme ottenuto da A togliendo il vettore v_i, ovvero che nessun v_i appartenga allo *Span* dell'insieme di vettori formato dai vettori di A tranne v_i:

Proposizione 1.38. *Un insieme $A = \{v_1, \ldots, v_n\}$ di vettori di uno spazio vettoriale V su $\mathbb{K}$ è un insieme di vettori linearmente indipendenti se e solo se nessun v_i appartenente ad A si può scrivere come combinazione lineare dell'insieme $B = A \backslash \{v_i\}$ (ovvero v_i non appartiene a Span(B)).*

Dimostrazione. $\Rightarrow$) Sia A un insieme di vettori linearmente indipendenti e supponiamo per assurdo che esista un v_i che si scriva come combinazione lineare dei vettori di A diversi da lui. Possiamo senza perdere di generalità supporre che tale vettore sia quello che abbiamo indicato con v_1, e dunque stiamo supponendo che $v_1 \in Span(v_2, \ldots, v_n)$, ovvero che esistono $n-1$ scalari $a_2, \ldots, a_n$ di $\mathbb{K}$ tali che:

$$\underbrace{\sum_{i=2}^{n} a_i \cdot v_i}_{a_2 \cdot v_2 + \ldots + a_n \cdot v_n} = v_1$$

Ma da questa, aggiungendo ad entrambi i membri dell'uguaglianza l'opposto di v_1, seguirebbe che:

$$a_2 \cdot v_2 + \ldots + a_n \cdot v_n - v_1 = O$$

Ovvero avremmo trovato una combinazione lineare non nulla (perché il coefficiente di v_1 è uguale a -1 (vedi Proposizione **??**) dei vettori di A, contro l'ipotesi che A sia un insieme di vettori linearmente indipendenti.

$\Leftarrow$) Siano i vettori di A tali che nessuno di loro si può scrivere come combinazione lineare degli altri vettori di A. Supponiamo per assurdo che A sia un insieme di vettori linearmente dipendenti. Allora devono esistere n scalari a_i, non tutti nulli, tali che:

$$\underbrace{\sum_{i=1}^{n} a_i \cdot v_i}_{a_1 \cdot v_1 + \ldots + a_n \cdot v_n} = O$$

Non essendo, per ipotesi, i coefficienti tutti nulli, almeno uno dei coefficienti deve essere diverso da zero. Eventualmente riordinando l'ordine dei vettori (dovrebbe essere abbastanza ovvio che la dipendenza lineare di un insieme di vettori non dipende dall'ordine in cui si elencano) possiamo supporre che sia a_1 diverso da 0, e scrivere:

$$\sum_{i=2}^{n} -a_i \cdot v_i = a_1 \cdot v_1$$

Ed essendo $a_1 \in \mathbb{K}$ diverso da 0, esiste l'inverso a_1^{-1} di a_1 in $\mathbb{K}$. Moltiplicando per questo inverso otteniamo:

$$\sum_{i=2}^{n} \frac{-a_i}{a_1} \cdot v_i = v_1$$

Ottenendo, contro l'ipotesi iniziale, che v_1 è combinazione lineare di $v_2, \ldots, v_n$. $\square$

Osservazione 1.39. La caratteristica di essere un insieme di vettori linearmente dipendenti, come ben espresso dalla definizione, è, per l'appunto, una caratteristica che riguarda l'intero insieme dei vettori. Ad esempio, può accadere (ma non in generale) che se da un insieme di n vettori linearmente dipendenti ne togliamo uno, l'insieme residuo, costituito da $n-1$ vettori, risulta un insieme di vettori linearmente indipendenti. Daremo un esempio fra poco: il lettore provi intanto a pensare ad un esempio nello spazio vettoriale $\mathbb{R}^2$.

Esercizio 1.40. Dimostrare che un insieme composto da un unico vettore v, diverso dal vettore nullo, è sempre un insieme di vettori linearmente indipendenti (Suggerimento: vedi Esercizio 1.3).

Esercizio 1.41. Sia V uno spazio vettoriale su $\mathbb{K}$. Dimostrare che se $A = \{v_1, \dots, v_n\}$ è un sottoinsieme di V di vettori linearmente indipendenti, allora qualsiasi sottoinsieme non vuoto di A è un insieme di vettori linearmente indipendenti.

Esercizio 1.42. Considerati i vettori dell'Esempio 1.35:

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

Dimostrare che:

(1) l'insieme contenente questi vettori è un insieme di vettori linearmente dipendenti, mostrando esplicitamente una combinazione lineare di questi vettori che sia uguale a 0, ma non abbia tutti i coefficienti nulli,
(2) gli insiemi costituiti rispettivamente dai primi tre vettori, e dagli ultimi tre, sono entrambi linearmente indipendenti,
(3) in riferimento a quanto detto nell'Osservazione 1.39, mostrare che se invece si considera l'insieme costituito dal primo, secondo e quarto vettore, tale insieme continua ad essere un insieme di vettori linearmente dipendenti.

Abbiamo definito cosa è un insieme di generatori di uno spazio vettoriale, e cosa significa insieme di vettori linearmente indipendenti, consideriamo gli insiemi di vettori che verificano entrambe queste due proprietà:

Definizione 1.43. Sia V uno spazio vettoriale su $\mathbb{K}$, un insieme di vettori $\{v_1, v_2, \dots, v_n\}$ di V, che genera lo spazio V e che è un insieme di vettori linearmente indipendenti si dice una **base** (finita) di V.

Osservazione 1.44. Nella definizione 1.43 è specificato *finita*. Non sempre uno spazio vettoriale ammette un numero finito di generatori, e di conseguenza nemmeno una base finita. Un caso che abbiamo incontrato di spazio vettoriale che non ammette una base finita è lo spazio vettoriale $\mathbb{K}[x]$ dei polinomi a coefficienti in $\mathbb{K}$. Basta osservare che i termini di grado n di un polinomio, si ottengono solo inserendo nell'insieme dei generatori un polinomio che contenga $a_n x^n$ con a_n diverso da 0 (perché?). A questo punto sorge spontanea la domanda: ma tale spazio vettoriale ammetterà una base infinita oppure non ha una base? Per rispondere però dovremmo ampliare la definizione di combinazione lineare (e poi di indipendenza lineare) a insiemi infiniti[9]. Ma in questo corso considereremo quasi esclusivamente

[9] Questo si fa chiedendo che le combinazioni lineari che appaiono nelle formule siano comunque composte da un numero finito di addendi, ovvero che tranne un numero finito di coefficienti, tutti gli altri siano uguali a 0. Dopodiché non è difficile verificare che l'insieme delle potenze di x con esponente $n \in \mathbb{N}$ è una base di $\mathbb{K}[x]$.

spazi vettoriali che ammettono una base finita, e ad ogni modo, cercheremo base solo di spazi vettoriali che ne hanno una finita.

Fissata la definizione di base (finita[10]) di uno spazio vettoriale, siamo interessati a capire:

(1) Se la scelta di una base, in luogo di un generico sistema di generatori (che potrebbero anche formare un insieme di vettori linearmente dipendenti), garantisce l'unicità di scrittura di un vettori in termini di combinazione lineare degli elementi della base.
(2) Quando uno spazio vettoriale ammette una base finita, ed in particolare se il fatto che uno spazio vettoriale V abbia un insieme finito di generatori, garantisce o no che V abbia una base finita.

Cominciamo analizzando il primo aspetto. Fissiamo dunque uno spazio vettoriale V (sul campo $\mathbb{K}$) e supponiamo che ammetta una base finita $\{v_1, v_2, \dots, v_n\}$. Dimostriamo che effettivamente la definizione di base è funzionale allo scopo di avere un'unica rappresentazione di ogni vettore di V come combinazione lineare dei vettori della base:

Proposizione 1.45. *Ogni vettore $v \in V$ si scrive IN MODO UNICO come combinazione lineare degli elementi della base.*

Dim. Il vettore v si può scrivere come combinazione lineare degli elementi della base perché, per definizione, 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:

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

dove gli a_i e i b_j sono elementi del campo $\mathbb{K}$. Allora:

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

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.

Questo implica che $a_i - b_i = 0$ per ogni i, ovvero che $a_i = b_i$ per ogni i. Abbiamo dunque provato che esiste un solo modo di scrivere v come combinazione lineare degli elementi della base data. $\square$

Osservazione 1.46. Una precisazione importante sul significato di *IN MODO UNICO*: si intende che per ottenere lo specifico vettore v, i coefficienti di ciascun vettore della base sono fissati, ma ovviamente (vale la proprietà commutativa) possiamo cambiare l'ordine degli addendi della combinazione lineare. Ad esempio se $\{v_1, v_2\}$ è una base dello spazio vettoriale V e $v = 3v_1 + 7v_2$, possiamo scrivere anche $v = 7v_2 + 3v_1$. Quello che rimane fisso (ed è quello che ci dice la proposizione 1.45) è la corrispondenza vettore-coefficiente: ovvero devo scegliere 3 per il vettore v_1 e 7 per il vettore v_2.

Se fissiamo una volta per tutte l'ordine degli elementi della base, ed è quello che solitamente si fa, allora fissiamo definitivamente la sequenza dei coefficienti.

[10]D'ora innanzi non ricorderemo più questo aspetto, sottointendendo che tutti i risultati che proveremo si riferiranno al caso finito.

Visti i risultati provati, dovrebbe cominciare ad essere chiara l'importanza di essere in grado di determinare, se possibile, la base di uno spazio vettoriale. Un primo risultato molto significativo da questo punto di vista, è riassunto nel prossimo teorema, che risponde alla seconda questione sulle basi che avevamo sollevato, ovvero se il fatto che uno spazio vettoriale V abbia un insieme finito di generatori, garantisce o no che V abbia una base finita. Il teorema è importante non solo dal punto di vista teorico (si riassume dicendo che si può sempre **estrarre** una base da un insieme finito di generatori), ma anche *pratico*. Infatti dalla dimostrazione del teorema segue una caratterizzazione della base che sarà utile per ricavare, appunto concretamente, una base, una volta conosciuto un insieme finito di generatori dello spazio vettoriale.

Teorema 1.47. *Sia V uno spazio vettoriale (sul campo $\mathbb{K}$) diverso da $\{O\}$ e generato dall'insieme finito di vettori non nulli $\{w_1, w_2, \ldots, w_s\}$. Allora è possibile estrarre da $\{w_1, w_2, \ldots, w_s\}$ un sottoinsieme $\{w_{i_1}, w_{i_2}, \ldots, w_{i_n}\}$ (con $n \leq s$) che è una base di V.*

Dim. Consideriamo l'insieme:

$$\mathcal{M} = \{A \subseteq \{w_1, w_2, \ldots, w_s\} \mid A \text{ è un insieme di vettori lin. ind.}\}$$

e notiamo che $\mathcal{M}$ non è vuoto, in quanto contiene certamente i sottoinsiemi di $\{w_1, w_2, \ldots, w_s\}$ di cardinalità 1, tipo $\{w_1\}$ o $\{w_2\}$ (vedi esercizio 1.40). Fra tutti gli elementi di $\mathcal{M}$ consideriamone uno di cardinalità massima: $W = \{w_{i_1}, w_{i_2}, \ldots, w_{i_n}\}$ (sicuramente è $n \geq 1$ ossia tale insieme non è vuoto).

Questo W è proprio il nostro candidato ad essere una base di V.

Osserviamo per prima cosa che W, per come lo abbiamo costruito (vedi esercizio 1.41), è un insieme di vettori linearmente indipendenti: resta da dimostrare che genera V.

Per questo bisogna mostrare che con combinazioni lineari dei vettori di W possiamo ottenere uno qualunque dei vettori $\{w_1, w_2, \ldots, w_s\}$ che sappiamo, per ipotesi, generare V[11].

Se $\{w_{i_1}, w_{i_2}, \ldots, w_{i_n}\} = \{w_1, w_2, \ldots, w_s\}$ abbiamo già finito: i vettori di partenza costituiscono di già un insieme di vettori linearmente indipendenti, e dunque sono una base finita di V.

Se invece

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

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

Se consideriamo l'insieme $\{w_r, w_{i_1}, w_{i_2}, \ldots, w_{i_n}\}$ notiamo che certamente questo non è un insieme di vettori linearmente indipendenti, altrimenti apparterrebbe a $\mathcal{M}$ e non sarebbe più vero che $\{w_{i_1}, w_{i_2}, \ldots, w_{i_n}\}$ ha cardinalità massima fra gli elementi di $\mathcal{M}$.

Dunque esiste una combinazione lineare:

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

[11]Verificare di aver capito bene questo passaggio! In pratica stiamo dicendo che se lo *Span* di un insieme di vettori di uno spazio vettoriale V, contiene un insieme di generatori di V, allora quello *Span* è uguale a V.

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 contraddirrebbe la lineare indipendenza di $\{w_{i_1}, w_{i_2}, \ldots, w_{i_n}\}$.

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. $\square$

Osservazione 1.48. Dalla dimostrazione del teorema 1.47 segue la caratterizzazione di cui parlavamo: ogni sottoinsieme massimale di vettori linearmente indipendenti dell'insieme dei generatori di V è una base di V; questo ci suggerisce che la base di uno spazio vettoriale non è unica. Osserviamo infatti che se uno spazio vettoriale V ammette una base finita $v_1, \ldots, v_n$ allora anche $\lambda \cdot v_1, v_2, \ldots, v_n$ è una base di V, diversa dalla precedente, qualsiasi sia $\lambda \in \mathbb{K} \backslash \{0\}$.

L'osservazione 1.48 sottolinea il fatto che uno spazio vettoriale che ammette una base ne ammette anche altre (se il campo $\mathbb{K}$ è infinito, ne ammette infinite altre). Può dunque sorgere il dubbio che il numero di elementi di una base di uno spazio vettoriale V dipenda dalla base scelta e non da V. Il seguente risultato (che dimostreremo nella seconda parte del corso), risponde a questo legittimo dubbio:

Sia V uno spazio vettoriale sul campo $\mathbb{K}$ che ammette una base finita. Allora tutte le basi di V hanno la stessa cardinalità.

Possiamo dunque dare la seguente definizione:

Definizione 1.49. Sia V uno spazio vettoriale con basi di cardinalità n. Tale cardinalità n è detta la **dimensione** di V.

Esercizio 1.50. Dimostrare che $\mathbb{K}^{\leq n}[x]$ (che sappiamo dall'Esercizio 1.21 essere un sottospazio vettoriale di $\mathbb{K}[x]$) ha dimensione $n+1$. (Suggerimento: mostrare che $\{1, x, x^2, \ldots, x^n\}$ è una base).

Esercizio 1.51. Consideriamo il seguente sottoinsieme di $\mathbb{Q}[x]$:

$$W = \{p(x) \in \mathbb{Q}[x] | deg(p(x)) \leq 3 \text{ e } p(x) \text{ è divisibile per } (x-4)\}$$

(1) Dimostrare che W è sottospazio vettoriale di $\mathbb{Q}[x]$.
(2) Trovare la dimensione di W.

Svolgimento. Per dimostrare che W è un sottospazio di $\mathbb{Q}[x]$ basta osservare che W è l'intersezione di due sottospazi di $\mathbb{Q}[x]$, ovvero U, l'insieme dei polinomi di grado minore o uguale a 3, e V, l'insieme dei polinomi divisibili per $x - 4$ (o equivalentemente che si annullano in 4).

Per trovare la dimensione del sottospazio W, ne determineremo una base. Osserviamo che un polinomio $p(x)$ di W è per definizione del tipo:

$$\underbrace{(x-4)}_{divisibile\ per\ x-4} \cdot \underbrace{(ax^2+bx+c)}_{di\ grado\ \leq 3} \quad a, b, c \in \mathbb{Q}$$

Ovvero:

$$p(x) = ax^2(x-4) + bx(x-4) + c(x-4)$$

Quindi $\{(x-4),\ x(x-4),\ x^2(x-4)\}$ è un insieme di generatori di W.

Verificare che questi tre vettori sono anche linearmente indipendenti (come del resto tutti gli insiemi di polinomi composti da polinomi che a due a due hanno grado diverso). Quindi $\{(x-4),\ x(x-4),\ x^2(x-4)\}$ è una base di W, che dunque abbiamo scoperto avere dimensione 3.

Esercizio 1.52. Dati due sottospazi vettoriali U e W di uno spazio vettoriale V su $\mathbb{K}$:

(1) dimostrare che la dimensione di un sottospazio intersezione $U \cap W$ è minore o uguale del minimo tra la dimensione di U e quella di W,
(2) discutere quando la dimensione dell'intersezione è esattamente uguale al minimo tra la dimensione di U e quella di W.

5. Applicazioni lineari

Introdotta la struttura di spazio vettoriale, e definito sottospazio vettoriale, è naturale chiedersi quali funzioni "rispettano" le strutture vettoriali introdotte; ci domandiamo dunque: quali funzioni mandano sottospazi in sottospazi[12]?

Mostriamo con un esempio che in generale, la proprietà di mandare sottospazi in sottospazi non è garantita, e dunque dovremo considerare funzioni "particolari" affinché la proprietà voluta sia garantita:

Esempio 1.53. Consideriamo la funzione $f : \mathbb{R}^2 \to \mathbb{R}^2$ definita da

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

La funzione f manda i punti $\begin{pmatrix} x \\ x \end{pmatrix}$, con la prima e seconda coordinata uguali, ovvero i punti della retta di equazione $x = y$, nella parabola di equazione $y = x^2$. Ma, come sappiamo dalla proposizione **??**, la retta $y = x$, passando dall'origine, è un sottospazio di $\mathbb{R}^2$, mentre la parabola non lo è (la proposizione **??** esplicita che tutti e solo i sottospazi propri di $\mathbb{R}^2$ sono le rette passanti per l'origine).

Dobbiamo dunque considerare funzioni (applicazioni) con proprietà *particolari*:

Definizione 1.54. Siano V e W spazi vettoriali di dimensione finita sul campo $\mathbb{K}$. Una applicazione L da V a W è è detta **lineare** se soddisfa le seguenti due proprietà:

$$\forall v_1, v_2 \in V \quad L(v_1 + v_2) = L(v_1) + L(v_2) \tag{5.1}$$

$$\forall \lambda \in \mathbb{K}\ \forall v \in V \quad L(\lambda v) = \lambda L(v) \tag{5.2}$$

Osservazione 1.55. È molto importante sottolineare come la definizione di applicazione lineare, vista la proprietà 5.2 che deve essere soddisfatta, e che mette in gioco lo scalare λ di $\mathbb{K}$ sia a sinistra che a destra dell'uguaglianza, abbia senso per applicazioni tra spazi vettoriali V e W sullo STESSO campo $\mathbb{K}$.

Osservazione 1.56. Il soddisfare le due proprietà 5.1 e 5.2 della definizione 1.54 da parte di una applicazione L, è equivalente a soddisfare la seguente proprietà:

$$\forall\ v_1, v_2 \in V\ forall\ a, b \in \mathbb{K} \quad L(av_1 + bv_2) = aL(v_1) + bL(v_2) \tag{5.3}$$

[12]Ovvero se f è una funzione da uno spazio vettoriale V ad uno spazio vettoriale W, vorremmo che per ogni sottospazio U di V, l'immagine $f(U)$ di U tramite f, fosse un sottospazio di W.

Esercizio 1.57. Dimostrare l'equivalenza tra il soddisfare le proprietà 5.1 e 5.2 e la proprietà 5.3 (ovvero provare che una applicazione L che soddisfa le due proprietà 5.1 e 5.2 necessariamente verifica anche la proprietà 5.3, e viceversa se L verifica la proprietà 5.3 allora necessariamente verifica anche le proprietà 5.1 e 5.2).

Nel caso di applicazioni lineari da uno spazio vettoriale V ad uno spazio vettoriale W (entrambi sul campo $\mathbb{K}$), ha proprietà molto interessanti l'insieme degli elementi (vettori) di V che hanno come immagine lo O di W. Per questo diamo un nome a questo insieme:

Definizione 1.58. Siano V e W spazi vettoriali sul campo $\mathbb{K}$ e consideriamo una applicazione lineare L da V a W. Chiameremo **nucleo** di L, e lo indicheremo[13] con $Ker(L)$, il seguente sottoinsieme di V:

$$Ker\ L = \{v \in V \mid L(v) = O\}$$

Cominciamo a delineare le prime proprietà del nucleo e dell'immagine di una applicazione lineare. Siano dunque V e W spazi vettoriali sul campo $\mathbb{K}$ e L una applicazione lineare da V a W.

(1) $Ker\ L$ è un sottospazio vettoriale di V.
(2) $Imm\ L$ è un sottospazio vettoriale di W.
(3) L è iniettiva[14] se e solo se $Ker\ L = \{O\}$.

Esercizio 1.59. Dimostrare le tre proprietà appena enunciate.

Svolgimento. Lasciamo le prime due verifiche al lettore e discutiamo la terza proprietà che caratterizza le applicazioni lineari iniettive.

Cominciamo supponendo che L sia iniettiva e dimostriamo che $Ker(L) = O$. Facciamo vedere che l'unico elemento di V che viene mandato nello O di W[15] è O.

Primo passo O di V va nello O di W:

$$L(O) \underbrace{=}_{def.O} L(O+O) \underbrace{=}_{L\ lineare} L(O) + L(O) \tag{5.4}$$

Essendo W uno spazio vettoriale, esiste l'opposto w di $L(O)$, ed aggiungendo w ad entrambi i membri dell'equazione 5.4, si ottiene proprio $L(O) = O$ (dove il primo O è l'elemento neutro in V, e il secondo, l'elemento neutro in W).

A questo punto, essendo L iniettiva, non ci possono essere altri elementi di V, diversi da O, che hanno la stessa immagine. Dunque $Ker(L) = \{O\}$.

A questo punto dobbiamo provare il viceversa, ovvero che se $Ker(L) = \{O\}$ allora L è iniettiva. Consideriamo dunque due vettori v, u che hanno la stessa immagine tramite L, e mostriamo che deve essere necessariamente $u = v$. Da $L(u) = L(v)$ sommando da entrambe le parti per l'opposto $-L(v)$ di $L(v)$ si ottiene:

$$L(u) - L(v) = O$$

[13] La parola inglese per nucleo è kernel, questo spiega il simbolo $Ker\ L$ che viene utilizzato per indicare il nucleo di una applicazione.

[14] Ricordiamo che una applicazione f tra due insiemi A e T è detta iniettiva se per ogni a, b di A, $f(a) = f(b)$ implica $a = b$. La stessa proprietà si può enunciare dicendo che per ogni a, b in A, se $a \neq b$ allora $f(a) \neq f(b)$ (ovvero l'iniettività significa che non ci sono elementi diversi di A che hanno la stessa immagine tramite f).

[15] Dovremmo indicare diversamente lo O di V e lo O di W, ad esempio usando O_V e O_W, in quanto in generale non saranno lo stesso elemento. Per non appesantire, quando non ci sia confusione sul fatto che stiamo parlando di un elemento di V o di W, useremo sempre la notazione O.

E per linearità di L, $L(u-v) = O$. Per ipotesi (il $Ker(L)$ contiene solo lo O), l'elemento $u-v$ deve essere O, ovvero $u-v=O$, cioè $u=v$.

6. Matrici e vettori in colonna per rappresentare applicazioni lineari e elementi di uno spazio vettoriale

Spesso in matematica, il modo di rappresentare gli oggetti può essere molto importante. Nel caso delle applicazioni lineari, è molto utile la rappresentazione tramite oggetti noti come *matrici*.

Definizione 1.60. Dati due interi positivi m, n, una **matrice** $m \times n$ a coefficienti in $\mathbb{K}$ è una griglia composta da m righe e n colonne in cui in ogni posizione c'è un elemento di $\mathbb{K}$ (tali elementi vengono chiamati coefficienti della matrice):

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \tag{6.1}$$

Anche la scelta degli indici dei coefficienti della matrice è strategica: come si può notare, l'elemento che si trova nella riga i-esima dall'alto e nella colonna j-esima da sinistra viene indicato con a_{ij}. Spesso per indicare la matrice A in 6.1 useremo la notazione sintetica $A = (a_{ij})$ e talvolta, per ricordare quali sono le dimensioni della matrice, scriveremo:

$$A = (a_{ij})_{\substack{i=1,2,\dots,m \\ j=1,2,\dots,n}}$$

Definizione 1.61. Dati due interi positivi m, n, chiamiamo $Mat_{m\times n}(\mathbb{K})$ l'insieme di tutte le matrici $m \times n$ a coefficienti in $\mathbb{K}$.

Definizione 1.62. Sull'insieme $Mat_{m\times n}(\mathbb{K})$ possiamo in maniera *naturale* definire la addizione e la moltiplicazione per scalare. Date due matrici $A = (a_{ij})$ e $B = (b_{ij})$ in $Mat_{m\times n}(\mathbb{K})$ e dato uno scalare k in $\mathbb{K}$, definiamo:

- la matrice somma $A + B = C = (c_{ij})$, il cui generico coefficiente nella i-esima riga e j-esima colonna è ottenuto sommando i coefficienti nella stessa posizione (cioè alla i-esima riga e j-esima colonna) di A e di B. Ovvero per ogni $i \leq m$ e per ogni $j \leq n$ $c_{ij} = a_{ij} + bij$.

  Vediamo un esempio di somma tra matrici 2×3:

  $$\begin{pmatrix} 1 & 2 & 4 \\ 0 & 6 & 3 \end{pmatrix} + \begin{pmatrix} 2 & 2 & 2 \\ 5 & 6 & -8 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 6 \\ 5 & 12 & -5 \end{pmatrix}$$

- la matrice moltiplicazione per scalare $k \cdot A = D = (d_{ij})$, il cui generico coefficiente nella i-esima riga e j-esima colonna è ottenuto moltiplicando lo scalare k per il coefficiente di A nella i-esima riga e j-esima colonna. Ovvero per ogni $i \leq m$ e per ogni $j \leq n$ $d_{ij} = k \cdot a_{ij}$.

  Vediamo un esempio di moltiplicazione per scalare di una matrice 2×3:

  $$3 \begin{pmatrix} 1 & 2 & 4 \\ 0 & 6 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 12 \\ 0 & 18 & 9 \end{pmatrix}$$

Esercizio 1.63. Dimostrare che, con le operazioni di addizione e moltiplicazione per scalare introdotte nella definizione 1.62, $Mat_{m\times n}(\mathbb{K})$ è uno spazio vettoriale su $\mathbb{K}$.

Esercizio 1.64. Dimostrare che la dimensione dello spazio vettoriale $Mat_{m\times n}(\mathbb{K})$ su $\mathbb{K}$ è $m \times n$.

Oltre alle operazioni introdotte nella definizione 1.62, si può introdurre un'altra operazione tra matrici: il cosidetto *prodotto righe per colonne*. Per definire tale prodotto è importante anche l'ordine in cui si considerano le due matrici in quanto sarà definito solo quando il numero di colonne di A è uguale al numero di righe di B. Cioè il prodotto righe per colonne tra una matrice A di $Mat_{m\times n}(\mathbb{K})$ e una matrice B di $Mat_{h\times k}(\mathbb{K})$ è definito solo se $n = h$.

Definizione 1.65. Data una matrice $A = (a_{ij})$ di $Mat_{m\times n}(\mathbb{K})$ e una matrice $B = (b_{st})$ di $Mat_{n\times k}(\mathbb{K})$, il **prodotto riga per colonna** AB, è la matrice $C = (c_{rh})$ di $Mat_{m\times k}(\mathbb{K})$, i cui coefficienti, per ogni r, h, sono definiti come segue:

$$c_{rh} = a_{r1}b_{1h} + a_{r2}b_{2h} + a_{r3}b_{3h} + \cdots + a_{rn}b_{nh}$$

Ovvero per ottenere l'elemento c_{rh} dobbiamo moltiplicare progressivamente (ovvero il primo con il primo, il secondo con il secondo, e così via) gli elementi della r-esima riga di A, con gli elementi della h-esima riga di B (da qui il nome prodotto riga per colonna) e sommare i risultati ottenuti.

Esempio 1.66. Consideriamo la matrice A di $Mat_{2\times 3}(\mathbb{K})$:

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 6 & 3 \end{pmatrix}$$

e la matrice B di $Mat_{3\times 3}(\mathbb{K})$:

$$\begin{pmatrix} 2 & 2 & 2 \\ 5 & 6 & -8 \\ 0 & 1 & 0 \end{pmatrix}$$

La definizione ci dice che possiamo definire $C = AB$ e che C è la matrice di

$Mat_{2\times 3}(\mathbb{K})$ i cui coefficienti sono ottenuti come segue:

$$c_{11} = \underbrace{1\cdot 2}_{a_{11}\cdot b_{11}} + \underbrace{2\cdot 5}_{a_{12}\cdot b_{21}} + \underbrace{4\cdot 0}_{a_{13}\cdot b_{31}} = 12$$

$$c_{12} = \underbrace{1\cdot 2}_{a_{11}\cdot b_{12}} + \underbrace{2\cdot 6}_{a_{12}\cdot b_{22}} + \underbrace{4\cdot 1}_{a_{13}\cdot b_{32}} = 18$$

$$c_{13} = \underbrace{1\cdot 2}_{a_{11}\cdot b_{13}} + \underbrace{2\cdot(-8)}_{a_{12}\cdot b_{23}} + \underbrace{4\cdot 0}_{a_{13}\cdot b_{33}} = -14$$

$$c_{21} = \underbrace{0\cdot 2}_{a_{21}\cdot b_{11}} + \underbrace{6\cdot 5}_{a_{22}\cdot b_{21}} + \underbrace{3\cdot 0}_{a_{23}\cdot b_{31}} = 30$$

$$c_{22} = \underbrace{0\cdot 2}_{a_{21}\cdot b_{12}} + \underbrace{6\cdot 6}_{a_{22}\cdot b_{22}} + \underbrace{3\cdot 1}_{a_{23}\cdot b_{32}} = 39$$

$$c_{23} = \underbrace{0\cdot 2}_{a_{21}\cdot b_{13}} + \underbrace{6\cdot(-8)}_{a_{22}\cdot b_{23}} + \underbrace{3\cdot 0}_{a_{23}\cdot b_{33}} = -48$$

E dunque si ha:

$$AB = C = \begin{pmatrix} 12 & 18 & -14 \\ 30 & 39 & -48 \end{pmatrix}$$

Esercizio 1.67. Mostrare, trovando un opportuno esempio, che date due matrici $n\times n$ A, B, in generale non vale la proprietà commutativa del prodotto riga per colonna (si trovano esempi anche con $n = 2$). E dunque che $Mat_{n\times n}(\mathbb{K})$ è un esempio di anello non commutativo.

Consideriamo due spazi vettoriali V, W di dimensione finita (rispettivamente n e m) su $\mathbb{K}$, e una applicazione lineare:

$$L\,:\,V\to W$$

Scegliamo in V una base $\{e_1,\dots,e_n\}$ e in W una base $\{\epsilon_1,\dots,\epsilon_m\}$. Il problema che affronteremo in questo paragrafo è il seguente: dato un elemento $v\in V$, calcolare la sua immagine $L(v)$ utilizzando le basi date e una notazione conveniente.

Sappiamo che v si può scrivere in modo unico come combinazione lineare degli elementi della base scelta, ovvero che esistono n scalari b_i tali che:

$$v = b_1e_1+\cdots+b_ne_n = \sum_{i=1}^n b_ie_i$$

Per la linearità di L si ha che:

$$L(v) = b_1L(e_1)+\cdots+b_nL(e_n) = \sum_{i=1}^n b_iL(e_i)$$

Dunque, per conoscere L, ossia per saper dire qual è l'immagine di un qualsiasi elemento $v\in V$, basta conoscere le immagini tramite L degli elementi di una base scelta $(L(e_1),\dots,L(e_n))$, e sapersi ricavare le coordinate $(b_1,\dots,b_n)$ di ogni v rispetto alla stessa base.

Esercizio 1.68 (Importante!)**.** Dimostrare, in base alle osservazioni precedenti, che $ImmL =< L(e_1),\dots,L(e_n) >$.

Esercizio 1.69. Dimostrare che dati due spazi vettoriali V e W, e fissata una base $\{e_1, \dots, e_n\}$ di V, e n elementi $w_1, \dots w_n$ di W, esiste una e una sola applicazione lineare L da V in W tale che, per ogni i, $L(e_i) = w_i$.

Per poter descrivere $L(e_1), \dots, L(e_n)$, che sono vettori di W, possiamo servirci della base $\{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}$. Per ogni i, sappiamo che esistono (unici) m scalari $a_j i$ (con j che varia da 1 a m) tali che:

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

Possiamo esprimere il vettore $L(e_i)$ in colonna[16]:

$$L(e_i) = \begin{pmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \\ \dots \\ \dots \\ a_{mi} \end{pmatrix}$$

dove abbiamo messo uno sotto l'altro i coefficienti di $L(e_i)$ rispetto alla base $\{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}$.

Da

$$L(v) = b_1 L(e_1) + \cdots + b_n L(e_n) = \sum_{i=1}^{n} b_i L(e_i)$$

sostituendo otteniamo:

$$L(v) = b_1(a_{11}\epsilon_1 + a_{21}\epsilon_2 + \cdots + a_{m1}\epsilon_m) + b_2(a_{12}\epsilon_1 + a_{22}\epsilon_2 + \cdots + a_{m2}\epsilon_m) + \cdots +$$

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

che, riordinando i termini, e raccogliendo a fattore gli ϵ_i, permette di esprimere $L(v)$ come combinazione lineare degli elementi della base $\{\epsilon_1, \dots, \epsilon_m\}$:
(6.2)

$$L(v) = (b_1 a_{11} + \cdots + b_n a_{1n})\epsilon_1 + \cdots + (b_1 a_{m1} + b_2 a_{m2} + \cdots + b_n a_{mn})\epsilon_m = \sum_{i=1}^{m} (\sum_{j=1}^{n} b_j a_{ij})\epsilon_i$$

Usando la notazione in colonna introdotta, il conto appena svolto si può tradurre nel seguente modo:

$$L(v) = b_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \dots \\ \dots \\ a_{m1} \end{pmatrix} + \cdots + b_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ \dots \\ \dots \\ a_{mn} \end{pmatrix} = \sum_{i=1}^{n} b_i \begin{pmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \\ \dots \\ \dots \\ a_{mi} \end{pmatrix}$$

[16] Useremo spesso, per uno spazio vettoriale di cui sia stata fissata una base, questa notazione con i vettori messi "in colonna".

Che ci permette di scrivere la relazione (6.2), usando lo strumento matrice appena introdotto, come segue:

$$L(v) = \begin{pmatrix} b_1a_{11} + b_2a_{12} + \cdots + b_na_{1n} \\ b_1a_{21} + b_2a_{22} + \cdots + b_na_{2n} \\ b_1a_{31} + b_2a_{32} + \cdots + b_na_{3n} \\ \cdots \\ \cdots \\ b_1a_{m1} + b_2a_{m2} + \cdots + b_na_{mn} \end{pmatrix} \tag{6.3}$$

Definizione 1.70. Data una applicazione lineare L da uno spazio vettoriale V di dimensione n ad uno spazio vettoriale W di dimensione m, si dice **matrice associata** all'applicazione lineare L nelle basi $\{e_1, e_2, \ldots, e_n\}$ di V e $\{\epsilon_1, \epsilon_2, \ldots, \epsilon_m\}$ di W la seguente matrice di m righe per n colonne:

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

Da ora in poi, per semplificare la notazione, ometteremo il riferimento alle basi tutte le volte che potremo farlo senza creare ambiguità, ma notiamo che la matrice $[L]$ associata all'applicazione L, non dipende solo da L stessa, ma anche dalle base scelte per V e W e che si ottiene ponendo uno accanto all'altro i vettori dei coefficienti di $L(e_1), L(e_2), \ldots, L(e_n)$ nella base scelta di W, scritti in colonna. Vediamo un esempio di determinazione della matrice di una applicazione lineare rispetto a basi diverse (in modo da impratichirsi nella determinazione della matrice associata ad una applicazione lineare, e anche di notare il fatto che cambiando le basi, cambia la matrice).

Esempio 1.71. 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 l'applicazione lineare

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

così definita (sappiamo, dall'esercizio 1.69, che, per determinare l'immagine di una applicazione lineare da V a W, e dunque per definirla, basta definire l'immagine degli elementi di una base di V):

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

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

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

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

A questa applicazione corrisponde la seguente matrice relativamente alle basi standard:

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

Consideriamo ora una base diversa di $\mathbb{R}^4$ (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 una base diversa anche di $\mathbb{R}^3$ (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

$$[L]_{\substack{v_1, v_2, v_3, v_4 \\ w_1, w_2, w_3}}$$

che rappresenta la solita applicazione lineare L ma rispetto a due basi diverse da quelle standard (osserveremo, come già anticipato, che troveremo una matrice diversa da $[L]_{\substack{e_1, e_2, e_3, e_4 \\ \epsilon_1, \epsilon_2, \epsilon_3}}$).

Procediamo esattamente come nel caso precedente (in cui i conti però erano immediati): 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 dobbiamo esprimerlo in termini della base $\{w_1, w_2, w_3\}$, ovvero dobbiamo trovare a, b, c scalari tali che:

$$\begin{pmatrix} 5 \\ \sqrt{3} + 1 \\ 1 \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + b \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} c \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

Impareremo tra breve a risolvere questi sistemi in maniera algoritmica, per ora accontentiamoci di verificare 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

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

è

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

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

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

Osservazione 1.72. Dati due spazi vettoriali V, W, esiste una sola applicazione lineare da V a W la cui matrice associata è indipendente dalle basi scelte. Si tratta della *applicazione nulla* $\mathcal{O} : V \to W$ che manda ogni $v \in V$ in $O \in W$. Qualunque siano le basi scelte, la matrice associata a tale applicazione avrà tutti i coefficienti uguali a 0.

Esempio 1.73. Consideriamo l'applicazione lineare L definita nell'Esempio 1.71. Abbiamo già trovato la matrice $[L]$ rispetto alle basi canoniche di $\mathbb{R}^4$ e $\mathbb{R}^3$:

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

Supponiamo di voler calcolare l'immagine del vettore v di $\mathbb{R}^4$ seguente:

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

Per quello che abbiamo appena osservato, si ha che i coefficienti di $L(v)$ nella base canonica di $\mathbb{R}^3$, sono dati dal prodotto riga per colonna seguente:

$$L(v) = [L] \begin{smallmatrix} e_1, e_2, e_3, e_4 \\ \epsilon_1, \epsilon_2, \epsilon_3 \end{smallmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 11 \\ \sqrt{3}+31 \\ 42 \end{pmatrix}$$

Osservazione 1.74. Consideriamo l'applicazione *identità* $I : V \to V$, che lascia fisso ogni elemento di v: $I(v) = v \quad \forall v \in V$, e fissiamo la base $\mathcal{B}$ di V. Si verifica che la matrice $[I] = (a_{ij})$ associata ad I rispetto a $\mathcal{B}$ sia in arrivo che in partenza (ovvero $[I]_{\substack{\mathcal{B}\\ \mathcal{B}}} = (a_{ij})$), è la matrice quadrata di formato $n \times n$ che ha tutti i coefficienti uguali a 0 eccetto quelli sulla diagonale, che sono invece uguali a 1: $a_{ij} = 0$ se $i \neq j$ e $a_{ii} = 1$ per ogni $i = 1, 2, \dots, n$.

Tale matrice è l'elemento neutro rispetto alla moltiplicazione riga per colonna in $Mat_{n\times n}(K)$ (verificare per esercizio).

Nel seguito useremo il simbolo I per indicare sia la applicazione lineare I sia la matrice identità, anche senza specificare di che formato sia $(2 \times 2, 3 \times 3, n \times n ...)$, visto che il contesto renderà sempre chiaro il significato.

7. Esercizi di fine capitolo

Esercizio 1.75. Si consideri la funzione $L : \mathbb{R}^3 \to \mathbb{R}^2$ definita sulle coordinate rispetto alle basi standard di $\mathbb{R}^3$ e $\mathbb{R}^2$ da:

$$L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - 2y - z \\ x + y + z \end{pmatrix}$$

(1) Verificare che L è lineare.
(2) Scrivere la matrice associata ad L rispetto alle basi standard di $\mathbb{R}^3$ e $\mathbb{R}^2$.
(3) Determinare una base di $Ker\ L$ e $Imm\ L$.

Svolgimento. Cominciamo, provando che effettivamente L è un'applicazione lineare[17], ovvero:

- $\forall\ v, w \in \mathbb{R}^3$ si ha che $L(v + w) = L(v) + L(w)$. Controlliamo che sussista questa uguaglianza; siano (x_1, y_1, z_1) e (x_2, y_2, z_2) le coordinate di v e w rispettivamente allora: $(v + w) = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$ quindi:

$$L(v+w) = L\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix} = \begin{pmatrix} (x_1 + x_2) - 2(y_1 + y_2) - (z_1 + z_2) \\ (x_1 + x_2) + (y_1 + y_2) + (z_1 + z_2) \end{pmatrix}$$

Mentre:

$$L(v)+L(w) = \begin{pmatrix} x_1 - 2y_1 - z_1 \\ x_1 + y_1 + z_1 \end{pmatrix} + \begin{pmatrix} x_2 - 2y_2 - z_2 \\ x_2 + y_2 + z_2 \end{pmatrix} = \begin{pmatrix} (x_1 + x_2) - 2(y_1 + y_2) - (z_1 + z_2) \\ (x_1 + x_2) + (y_1 + y_2) + (z_1 + z_2) \end{pmatrix}$$

- $\forall\ v \in V$ e $\forall\ k \in \mathbb{K}$ si ha che $L(k \cdot v) = k \cdot L(v)$. Anche in questo caso proviamo questa uguaglianza:

$$L(k \cdot v) = L\begin{pmatrix} k \cdot x \\ k \cdot y \\ k \cdot z \end{pmatrix} = \begin{pmatrix} k \cdot x - 2k \cdot y - k \cdot x \\ k \cdot x + k \cdot y + k \cdot z \end{pmatrix}$$

Mentre:

$$k \cdot L(v) = k \cdot \begin{pmatrix} x - 2y - z \\ x + y + z \end{pmatrix} = \begin{pmatrix} k \cdot x - 2k \cdot y - k \cdot x \\ k \cdot x + k \cdot y + k \cdot z \end{pmatrix}$$

[17]In effetti allo stesso modo in cui proviamo questo risultato, si può provare che sono applicazioni lineari da $\mathbb{R}^n$ a $\mathbb{R}^m$ tutte le applicazioni che agiscono sulle coordinate in maniera che il risultato sia una combinazione lineare delle stesse.

A questo punto, per scrivere la matrice associata a L rispetto alle basi standard:

$$ß_{\mathbb{R}^2} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

$$ß_{\mathbb{R}^3} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

bisogna calcolare, le coordinate rispetto alla base $ß_{\mathbb{R}^2}$, dell'immagine degli elementi di $ß_{\mathbb{R}^3}$ tramite L:

$$L\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \quad L\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}; \quad L\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Perciò la matrice associata a L nelle basi $ß_{\mathbb{R}^3}$ di $\mathbb{R}^3$ e $ß_{\mathbb{R}^2}$ di $\mathbb{R}^2$ è la seguente:

$$[A]_{\substack{ß_{\mathbb{R}^3} \\ ß_{\mathbb{R}^2}}} = \begin{pmatrix} 1 & -2 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

Sappiamo (vedi Esercizio 1.68) che $Imm\ L$ è generata dai vettori

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

ma, in questo caso, sappiamo anche *immediatamente* che questi non sono una base (se fossero linearmente indipendenti, sarebbe facile mostrare, per esempio usando il Teorema 1.47, che $\mathbb{R}^2$ ha dimensione ≥ 3, mentre sappiamo che la dimensione di $\mathbb{R}^2$ è 2). Lasciamo al lettore la verifica che, presi due qualunque vettori fra i tre scritti sopra, tali vettori costituiscono una base di $Imm\ L$.

Per trovare una base di $Ker\ L$, cerchiamo di capire come sono fatti i suoi elementi. Per definizione un vettore v sta in $Ker\ L$ se $L(v) = 0$. Questo si traduce, se poniamo $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, nella relazione:

$$\begin{pmatrix} 1 & -2 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Ovvero abbiamo il sistema:

$$\begin{aligned} x - 2y - z &= 0 \\ x + y + z &= 0 \end{aligned}$$

Risolviamo il sistema e troviamo:

$$\begin{aligned} &x = 2 \cdot (-\tfrac{2}{3}z) + z = -\tfrac{1}{3}z \\ &y = -\tfrac{2}{3}z \end{aligned}$$

Osserviamo che ci sono infinite soluzioni, una per ogni scelta di $z \in \mathbb{R}$. I vettori che stanno in $Ker\ L$ sono dunque della forma:

$$\begin{pmatrix} -\frac{1}{3}z \\ -\frac{2}{3}z \\ z \end{pmatrix} = z \begin{pmatrix} -\frac{1}{3} \\ -\frac{2}{3} \\ 1 \end{pmatrix}$$

Quindi il vettore:

$$\begin{pmatrix} -\frac{1}{3} \\ -\frac{2}{3} \\ 1 \end{pmatrix}$$

genera $Ker\ L$ e costituisce anche una base. In particolare L non è iniettiva perché $Ker\ L$ non è composto dal solo vettore nullo.

Esercizio 1.76. Dimostrare che l'insieme $S \subseteq \mathbb{R}^4$ delle soluzioni del seguente sistema lineare:

$$2x + y + t + 2z = 0$$

$$x + 3t + z = 0$$

$$x + y - 2t + z = 0$$

è un sottospazio vettoriale di $\mathbb{R}^4$. Secondo voi questo risultato vale anche in generale per qualunque sistema di equazioni lineari?

Esercizio 1.77. Sia $a \in \mathbb{R}$. Consideriamo in $\mathbb{R}^4$ il sottospazio V_a dato dalle soluzioni del seguente sistema lineare:

$$\begin{cases} x + 2y + z & = 0 \\ ay + z + 3t & = 0 \end{cases}$$

e il sottospazio W_a generato dai vettori

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

Calcolare, per $a = 3$, $dim\ V_a \cap W_a$ e $dim\ (V_a + W_a)$.
Calcolare, al variare di $a \in \mathbb{R}$, $dim\ V_a \cap W_a$ e $dim\ (V_a + W_a)$.

Esercizio 1.78. Sia $a \in \mathbb{R}$ e siano $f_a : \mathbb{R}^4 \to \mathbb{R}^3$, $g_a : \mathbb{R}^4 \to \mathbb{R}^3$ le funzioni date da

$$f_a(x, y, z, t) = (x + 2y + z, y + (a + 1)z, t + 1)$$

$$g_a(x, y, z, t) = ((a + 1)x + 2y + z, ay + (a + 1)z, az + (a + 1)t)$$

(1) Perché g_a è una applicazione lineare mentre f_a non lo è ?

(2) Scrivere una base per $Ker\ g_a$, quando $a = 5$.

(3) Scrivere una base per $Ker\ g_a$, al variare di $a \in \mathbb{R}$.

Esercizio 1.79. Sia $F : \mathbb{R}^2 \to \mathbb{R}^2$ l'applicazione lineare definita, nella base standard di $\mathbb{R}^2$, dalla matrice:

$$[F] = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}$$

Trovare $Ker\ F$ e $Imm\ F$.

Esercizio 1.80. Sia $F : \mathbb{C}^3 \to \mathbb{C}^3$ l'applicazione lineare definita, nella base standard di $\mathbb{C}^3$, dalla matrice:

$$[F] = \begin{pmatrix} -1 & 0 & -1 \\ 0 & 1 & i \\ 2-i & i & 1-i \end{pmatrix}$$

Trovare $Ker\ F$ e $Imm\ F$.

Esercizio 1.81. Sia $F : \mathbb{C}^3 \to \mathbb{C}^3$ l'applicazione lineare definita, nella base standard di $\mathbb{C}^3$, dalla matrice:

$$[F] = \begin{pmatrix} -1 & 0 & i \\ 2 & 1 & 2i \\ 2-i & 1 & 2i-1 \end{pmatrix}$$

Trovare $Ker\ F$ e $Imm\ F$.

Esercizio 1.82. Consideriamo i seguenti sottoinsiemi V e W dello spazio vettoriale $\mathbb{R}^{\leq 3}[x]$:

$$V = \{p(x) \in \mathbb{R}_3[x] |\ p(-1) = 0\}$$

e

$$W = \{p(x) \in \mathbb{R}_3[x] | p'(1) = 0\}$$

(Nota: con $p'(x)$ indichiamo la derivata del polinomio $p(x)$.)

(1) Dimostrare che V e W sono sottospazi vettoriali di $\mathbb{R}^{\leq 3}[x]$.
(2) Determinare una base di V, W, $W + V$ e $W \cap V$.

Esercizio 1.83. Consideriamo la matrice a coefficienti in $\mathbb{R}$

$$B = \begin{pmatrix} -1 & 1 \\ 2 & 2 \end{pmatrix}$$

Sia V lo spazio vettoriale delle matrici 2×2 a coefficienti in $\mathbb{R}$. Quale è la sua dimensione?

Dire se l'applicazione $L : V \to V$ tale che per ogni matrice X vale

$$L(X) = XB - BX$$

è lineare. Se è lineare, calcolare la dimensione del nucleo e dell'immagine.

Esercizio 1.84. Consideriamo i due seguenti sottoinsiemi di $\mathbb{R}^3$:

$$A = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}$$

$$\mathcal{B} = \left\{ \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

(1) Dimostrare che $\mathcal{B}$ è una base di $\mathbb{R}^3$ e che i vettori di A sono linearmente indipendenti. Completare poi A ad una base $\mathcal{C}$ di $\mathbb{R}^3$.
(2) Considerata l'applicazione lineare $L : \mathbb{R}^3 \to \mathbb{R}^3$ definita da $L(x, y, z) = (x+y, z, z)$ trovare una base di $Ker\ L$ e $Imm\ L$ e scrivere la matrice $[M]^{\mathcal{C}}_{\mathcal{B}}$ associata alla base $\mathcal{C}$ in partenza e alla base $\mathcal{B}$ in arrivo.

Esercizio 1.85. Siano

$$V_a = < \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} a \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} >$$

e

$$W = < \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} >$$

due sottospazi di $\mathbb{R}^4$.

(1) Al variare di a in $\mathbb{R}$, trovare la dimensione di $V_a + W$ e di $V_a \cap W$;
(2) Dire per quali valori dei parametri $a, b \in \mathbb{R}$ il vettore

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

appartiene al sottospazio $V_a \cap W$.

Esercizio 1.86. Consideriamo lo spazio vettoriale $\mathbb{R}[x]^{\leq 4}$. Sia

$$V = \{p(x) \in \mathbb{R}[x]^{\leq 4} \mid p(0) = p(1) = p(2)\}$$

- Si dimostri che V è un sottospazio vettoriale di $\mathbb{R}[x]^{\leq 4}$.
- Si calcoli la dimensione di V.

Esercizio 1.87. Sia $\mathcal{T} : Mat_{n\times n}(\mathbb{K}) \to \mathbb{K}$ la funzione *traccia* definita da

$$\mathcal{T}((a_{ij})) = a_{11} + a_{22} + \cdots + a_{nn}.$$

(1) Dimostrare che $\mathcal{T}$ è una applicazione lineare (per la struttura di spazio vettoriale su $Mat_{n\times n}(\mathbb{K})$ vedi l'Esercizio 1.63).
(2) Dimostrare che per ogni $A, B \in Mat_{n\times n}(\mathbb{K})$ vale $\mathcal{T}(AB) = \mathcal{T}(BA)$.

CAPITOLO 2

Il rango delle applicazioni lineari e la riduzione a scalini delle matrici

1. Le operazioni elementari sulle colonne

Consideriamo una generica matrice in $Mat_{m\times n}(\mathbb{K})$:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \tag{1.1}$$

e i tre seguenti tipi di *mossa* sulle colonne, detti anche **operazioni elementari sulle colonne** di una matrice:

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

Quello che vogliamo mostrare è che, attraverso l'utilizzo di queste 3 mosse, si può sempre trasformare una qualsiasi matrice $m \times n$ in una matrice a forma detta *a scalini (per colonne)*.

Esempio 2.1. Prima di dare una definizione formale di cosa sia una matrice a scalini per colonne, vediamo alcuni esempi, in quanto la visualizzazione rende piuttosto bene l'idea di cosa si intenda:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ \sqrt{3}+1 & 1 & 0 & 0 \\ -2 & \frac{5}{2} & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ \sqrt{3}+1 & 1 & 0 & 0 \\ -2 & 7 & 0 & 0 \\ \sqrt{3} & 4 & 1 & 0 \end{pmatrix}$$

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

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 \\ -8 & 4 & 3 & 0 & 0 \\ -2 & 1 & 0 & 4 & 0 \\ 0 & 0 & 0 & 9 & 1 \end{pmatrix}$$

Per una vera definizione di matrice a scalini per colonne possiamo seguire questa strada: chiamiamo *profondità* di una colonna la posizione occupata, contata dal basso, dal suo più alto coefficiente diverso da zero. Alla colonna nulla (con tutte i coefficienti uguali a 0) assegnamo per convenzione profondità 0.

Esempio 2.2. Consideriamo la seguente matrice:

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

Consideriamo la prima colonna della matrice A:

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

In questa colonna, il numero diverso da 0 che è più in alto è $\sqrt{3}+1$. Contando dal basso $\sqrt{3}+1$ è in seconda posizione, dunque la colonna ha profondità 2. Analogamente per la seconda colonna:

$$A = \begin{pmatrix} 4-\sqrt{3} \\ 0 \\ -2 \end{pmatrix}$$

il numero diverso da 0 più in alto è $4-\sqrt{3}$ che è in terza posizione contando dal basso, dunque la profondità della colonna è 3. E infine il numero diverso da 0 più in alto per la terza colonna è -2, che è in prima posizione contando dal basso. Dunque la terza colonna ha profondità 1.

A questo punto possiamo dare la definizione di matrice a scalini per colonne.

Definizione 2.3. Una matrice A in $Mat_{m\times n}(\mathbb{K})$, si dice **in forma a scalini per colonne** se rispetta le seguenti proprietà:

- leggendo la matrice da sinistra a destra, le colonne non nulle si incontrano tutte prima della colonne nulle;
- leggendo la matrice da sinistra a destra, le profondità delle sue colonne non nulle risultano strettamente decrescenti.

Definizione 2.4. In una matrice in forma a scalini per colonna, i coefficienti diversi da zero più alti di posizione di ogni colonna non nulla si chiamano **pivot**.

Esempio 2.5. Se consideriamo le matrici a scalini dell'Esempio 2.1, i pivot sono i coefficienti in neretto:

$$\begin{pmatrix} \mathbf{1} & 0 & 0 & 0 \\ \sqrt{3}+1 & \mathbf{1} & 0 & 0 \\ -2 & \frac{5}{2} & \mathbf{1} & 0 \end{pmatrix} \qquad \begin{pmatrix} \mathbf{1} & 0 & 0 & 0 \\ \sqrt{3}+1 & \mathbf{1} & 0 & 0 \\ -2 & 7 & 0 & 0 \\ \sqrt{3} & 4 & \mathbf{1} & 0 \end{pmatrix}$$

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

$$\begin{pmatrix} \mathbf{1} & 0 & 0 & 0 & 0 \\ -2 & \mathbf{1} & 0 & 0 & 0 \\ -8 & 4 & \mathbf{3} & 0 & 0 \\ -2 & 1 & 0 & \mathbf{4} & 0 \\ 0 & 0 & 0 & 9 & \mathbf{1} \end{pmatrix}$$

Teorema 2.6. *Data una matrice A in $Mat_{m\times n}(\mathbb{K})$ è sempre possibile, usando (un numero finito di*[1]*) operazioni elementari sulle colonne, ridurre la matrice in forma a scalini per colonne.*

Dimostrazione. Procediamo per induzione sul numero di righe m.

Il caso base, $m = 1$, è immediato. Se la matrice è composta da una riga nulla è già una matrice $1 \times n$ in forma a scalini; se invece ha qualche coefficiente non 0, con la terza delle operazioni elementari sulle colonne (ovvero la permutazione di colonne), si può portare un coefficiente non zero all'inizio della riga, dopodiché, utilizzando più volte la prima operazione per colonna, si possono ridurre a zero tutti gli altri coefficienti.

Supponiamo ora che l'enunciato sia vero per tutte le matrici con $m-1$ righe, e dimostriamolo per matrici con m righe. Sia A una matrice $m \times n$. Se una riga di A è nulla, possiamo considerare la matrice A' ottenuta da A togliendo tale riga. Per ipotesi induttiva, visto che A' ha $m-1$ righe, sappiamo che possiamo, usando solo le operazioni elementari sulle colonne, ridurre a scalini per colonne A'. Osserviamo a questo punto che la stessa sequenza di operazioni per colonne riduce a scalini anche A. Infatti, sommando tra loro due colonne (prima operazione elementare), o moltiplicando una colonna per uno scalare k diverso da 0 (seconda operazione elementare), o infine permutando due colonne (terza ed ultima operazione elementare ammessa sulle colonne), i valori su una riga composta da tutti 0, rimangono 0.

Se invece A non ha righe nulle, vuol dire che in A esiste una colonna di profondità m (prova a spiegare perché). Possiamo supporre che sia la prima colonna da sinistra (se non lo è possiamo sempre fare uno scambio di colonne), e dunque che $a_{1,1}$ sia diverso da zero. A questo punto, sottraendo tale colonna, moltiplicata per

[1]Questa precisazione è importante perché fa intuire come sia possibile costruire un algoritmo per ridurre una matrice in forma a scalini per colonne.

opportuni scalari, alle altre colonne, si giunge ad avere una matrice A' del tipo:

$$A' = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \tag{1.2}$$

Prima di proseguire con la dimostrazione del teorema vediamo, con un esempio, come effettivamente sia possibile trasformare una matrice A con $a_{1,1}$ diverso da 0, in una matrice della forma 1.2.

Consideriamo la matrice:

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

E vediamo come è possibile trasformarla nella forma di A', usando solo le operazioni elementari sulle colonne.

- Moltiplichiamo la prima colonna di A per $\frac{1}{3}$ (seconda operazione elementare). Scriveremo, come si può vedere qui sotto, $[1] = \frac{1}{3}[1]$ per indicare che sostituiamo alla prima colonna il risultato a destra dell'uguaglianza:

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 0 & 1 & 2 \\ -1 & 0 & 0 \\ 1 & 4 & 2 \end{pmatrix} \xrightarrow{[1]=\frac{1}{3}[1]} A_1 = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ -\frac{1}{3} & 0 & 0 \\ \frac{1}{3} & 4 & 2 \end{pmatrix}$$

- Sommiamo alla seconda colonna della matrice A_1 ottenuta, -2 volte la prima colonna di A_1 (scriveremo $[2] = [2] - 2[1]$):

$$A_1 \xrightarrow{[2]=[2]-2[1]} A_2 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ -\frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{3} & \frac{10}{3} & 2 \end{pmatrix}$$

- Sommiamo alla terza colonna della matrice A_2 ottenuta, la prima colonna di A_2 (scriveremo $[3] = [3] + [1]$):

$$A_2 \xrightarrow{[3]=[3]+[1]} A_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{10}{3} & \frac{7}{3} \end{pmatrix}$$

A questo punto possiamo tornare alla dimostrazione del nostro teorema.

Per ipotesi induttiva sappiamo che possiamo ridurre a scalini la sottomatrice di A' seguente:

$$\begin{pmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}$$

che ha una riga (e una colonna) in meno di A. Si osserva subito che le stesse mosse, operate sulla matrice A, la riducono a scalini. $\square$

Osservazione 2.7. Quando si riduce una matrice A in $Mat_{m\times n}(\mathbb{K})$ in forma a scalini, la forma a scalini ottenuta non è unica, basta osservare che se B è in forma scalini, allora anche $k \cdot B$ è a scalini (con $k \in \mathbb{K}$).

Osservazione 2.8. In quasi tutte le matrici dell'Esempio 2.1, in ogni colonna il coefficiente più alto diverso da zero è uguale a 1. Questa richiesta non rientra nella definizione di matrice a scalini per colonna. È vero o falso che quando si riduce una matrice A in $Mat_{m\times n}(\mathbb{K})$ in forma a scalini con questa ulteriore proprietà (ovvero che in ogni colonna il coefficiente più alto diverso da zero valga 1), la forma a scalini ottenuta è unica?

Esercizio 2.9. Vista la dimostrazione del teorema, provare a scrivere un algoritmo che data in input una matrice A in $Mat_{m\times n}(\mathbb{K})$, restituisce una forma a scalini per colonne di A.

Nel seguito, ci interesserà anche una forma ancora più *particolare* di matrice a scalini per colonne:

Definizione 2.10. Una matrice A in $Mat_{m\times n}(\mathbb{K})$, si dice **in forma a scalini per colonne ridotta** se:

- A è a scalini per colonne,
- Tutte le entrate nella stessa riga di un pivot, precedenti (leggendo la riga da sinistra a destra) al pivot, sono nulle.

Esempio 2.11. Ecco le forme a scalini ridotte per alcune delle matrici dell'Esempio 2.1:

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

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

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

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

Proposizione 2.12. *Data una matrice A in forma a scalini per colonne, è sempre possibile, usando solo la prima delle operazioni elementari sulle colonne, portare A in forma a scalini per colonna ridotta.*

Esercizio 2.13. Dimostrare la Proposizione 2.12.

Corollario 2.14. *Ogni matrice A può essere trasformata, attraverso le operazioni elementari sulle colonne, in una matrice in forma a scalini per colonna ridotta.*

Dimostrazione. Basta applicare il Teorema 2.6, per trasformare A in A' matrice in forma a scalini, e poi applicare la Proposizione 2.12 per trasformare A' in A'' in forma a scalini ridotta. □

Prima di proseguire verso lo studio delle applicazioni lineari tramite la *manipolazione* delle matrici corrispondenti, soffermiamoci ad analizzare alcune proprietà di quelle che abbiamo chiamato *operazioni elementari sulle colonne*.

Data una matrice A $m \times n$, a coefficienti in un campo $\mathbb{K}$, le n colonne possono essere considerati n vettori di $\mathbb{K}^m$. Ad esempio se abbiamo la matrice:

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

possiamo considerare i 4 vettori di $\mathbb{R}^5$ seguenti:

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 3 \\ 4 \\ 7 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ -1 \end{pmatrix}, v_3 = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, v_4 = \begin{pmatrix} 1 \\ 0 \\ 3 \\ 1 \\ 0 \end{pmatrix}$$

Data una matrice A in $Mat_{m\times n}(\mathbb{K})$, indicate con $v_1, \dots, v_n$ i vettori di $\mathbb{K}^m$ formati dalle colonne di A, si può in particolare considerare il sottospazio di $\mathbb{K}^m$ generato da $v_1, \dots, v_n$, ovvero $Span(v_1, \dots, v_n)$.

Proposizione 2.15. *Operando attraverso le operazioni elementari sulle colonne di una matrice, lo Span dei* vettori colonna *rimane invariato. Ovvero se indichiamo con $v_1, \dots, v_n$ i vettori colonna di una matrice A in $Mat_{m\times n}(\mathbb{K})$, per ogni matrice A' ottenuta da A attraverso le operazioni elementari sulle colonne, si ha, indicando con $w_1, \dots, w_n$ i vettori colonna di A', che:*

$$Span(v_1, \dots, v_n) = Span(w_1, \dots, w_n)$$

Dimostrazione. Basta osservare che:

- $Span(v_1, \dots, v_i, \dots, v_n) = Span(v_1, \dots, v_i + k \cdot v_j, \dots, v_n)$. Infatti se un vettore v appartiene a $Span(v_1, \dots, v_i, \dots, v_n)$ allora può essere scritto come

  $$v = \lambda_1 v_1 + \dots + \lambda_i v_i + \dots + \lambda_j v_j + \dots + \lambda_n v_n$$

  Ma questo vettore può essere scritto anche come

  $$v = \lambda_1 v_1 + \dots + \lambda_i (v_i + k \cdot v_j) + \dots + (\lambda_j - k\lambda_i) v_j + \dots + \lambda_n v_n$$

  dunque appartiene anche a $Span(v_1, \dots, v_i + k \cdot v_j, \dots, v_n)$.
  Viceversa, se un vettore w appartiene a $Span(v_1, \dots, v_i + k \cdot v_j, \dots, v_n)$ allora può essere scritto come

  $$w = \mu_1 v_1 + \dots + \mu_i (v_i + k \cdot v_j) + \dots + \mu_j v_j + \dots + \mu_n v_n$$

Si nota che può essere allora scritto anche come

$$w = \mu_1 v_1 + \cdots + \mu_i v_i + \cdots + (\mu_j + k\mu_i) v_j + \cdots + \lambda_n v_n$$

dunque appartiene anche a $Span(v_1, \ldots, v_i, \ldots, v_n)$.

- Se $k \neq 0$, $Span(v_1, \ldots, v_i, \ldots, v_n) = Span(v_1, \ldots, k \cdot v_i, \ldots, v_n)$ (dimostrazione simile alla precedente, anzi ancora più immediata).
- $Span(v_1, \ldots, v_i, \ldots, v_j, \ldots v_n) = Span(v_1, \ldots, v_j, \ldots, v_i, \ldots v_n)$ (dimostrazione simile alla precedente, anzi ancora più immediata).

□

Un'altra osservazione importante, riguardo alle operazioni elementari sulle colonne, è che ogni singola operazione è *reversibile*. Ossia, una volta fatta, possiamo fare la sua inversa e tornare esattamente alla matrice di partenza.

2. La riduzione a scalini per colonne applicata allo studio delle basi

Possiamo utilizzare le osservazioni sul metodo di riduzione a scalini per colonne per dimostrare finalmente il teorema che afferma che tutte le basi di uno spazio vettoriale hanno la stessa cardinalità.

Teorema 2.16. *Sia V uno spazio vettoriale sul campo $\mathbb{K}$ che ammette una base finita. Allora tutte le basi di V hanno la stessa cardinalità.*

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

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

Possiamo formare una matrice M ponendo uno accanto all'altro i vettori $v_1, v_2, \ldots, v_r$: M sarà del tipo n(righe) $\times$ r(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*, ovvero la differenza di profondità tra colonne adiacenti non può essere più di 1. 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 potrebbe venire 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 $\{e_1, e_2, \ldots, e_n\}$ e $\{v_1, v_2, \ldots, v_r\}$: in tal modo otterremo che deve valere $n \geq r$. Dunque $n = r$ come volevamo dimostrare $\square$

Corollario 2.17. *In uno spazio vettoriale V di dimensione n, dati n vettori linearmente indipendenti questi sono anche una base di V. Allo stesso modo, dati n vettori che generano V questi sono anche una base di V.*

Dimostrazione. Fissiamo una base di V $e_1, e_2, \ldots, e_n$ e siano $v_1, v_2, \ldots, v_n$ vettori linearmente indipendenti.

Esprimiamo i vettori $v_1, v_2, \ldots, v_n$ in termini della base $e_1, e_2, \ldots, e_n$ e poniamoli in colonna uno accanto all'altro. Così facendo otteniamo una matrice M che è $n \times n$. La matrice M' in forma a scalini ridotta ottenuta a partire da M è la matrice identità: $M' = I$. Il perché si basa su osservazioni già fatte, ma ripetiamole per esercizio: se i vettori $v_1, v_2, \ldots, v_n$ sono linearmente indipendenti allora lo *Span* delle colonne di M ha dimensione n. Ma tale *Span* coincide con lo *Span* delle colonne di M': le n colonne di M' devono dunque essere indipendenti. Questo può accadere solo se sono non nulle e di profondità diverse. L'unico modo è che $M' = I$. Riassumendo, lo *Span* delle colonne di M, ovvero $Span(v_1, v_2, \ldots, v_n)$, è uguale allo *Span* delle colonne di $M' = I$, che è tutto V. Abbiamo dimostrato che i vettori $v_1, v_2, \ldots, v_n$ generano V, e dunque sono una base.

Per quel che riguarda l'altra parte dell'enunciato, ossia quella in cui si considerano dei vettori $v_1, v_2, \ldots, v_n$ che generano V, per dimostrare che sono linearmente indipendenti basta applicare il Teorema 1.47: se non fossero linearmente indipendenti sarebbe possibile estrarre un sottoinsieme di cardinalità minore di n che è una base, dunque lo spazio avrebbe due basi di cardinalità diversa, assurdo.

$\square$

Considerazioni simili a quelle esposte fin qui ci permettono di descrivere un criterio concreto per decidere se, dato uno spazio vettoriale V di dimensione n ed una base $e_1, e_2, \ldots, e_n$ di V, e dati n vettori $v_1, v_2, \ldots, v_n$ di V, tali vettori costituiscono una base di V o no. Il criterio è il seguente: esprimiamo i vettori $v_1, v_2, \ldots, v_n$ trovandone i coefficienti rispetto alla base $e_1, e_2, \ldots, e_n$ e poniamoli 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 $\{v_1, v_2, \ldots, v_n\}$ è una base di V, altrimenti no. Proviamo a capire il perché di quest'ultima affermazione.

Nella dimostrazione del Teorema 2.16 abbiamo visto che se $\{v_1, v_2, \ldots, v_n\}$ è una base allora la forma a scalini ridotta M' non ha scalini *lunghi*. 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 generano V. Questo implica che le n colonne di M generano V, ossia che $v_1, v_2, \ldots, v_n$ generano V. Per il Corollario 2.17 si conclude che $v_1, v_2, \ldots, v_n$ sono una base.

Esempio 2.18. Facciamo ora un semplice esempio concreto di *riconoscimento* di una base. Consideriamo $\mathbb{R}^4$ con la sua base standard e poi i vettori

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

Verifichiamo col nuovo metodo che si tratta di una base.
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.

Un'altra importante applicazione delle osservazioni sulla riduzione a scalini per colonne è data dal prossimo teorema, il quale - in un certo senso - è il *duale* del Teorema 1.47. In quel caso avevamo dimostrato che da un insieme A di generatori di uno spazio V si può sempre estrarre una base di V (ovvero trovare un sottoinsieme di A che è base di V). Ora dimostriamo che un insieme B di vettori linearmente indipendenti di V si può sempre completare ad una base di V (ovvero esiste un insieme C che contiene B e che è base di V). La dimostrazione del teorema, non solo garantisce teoricamente che questo completamento si può fare, ma descrive anche un algoritmo per trovare un completamento.

Teorema 2.19 (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, con $1 \leq k \leqslant n$, può essere completato ad una base di V aggiungendo a B $n-k$ vettori di $V \setminus Span(B)$.*

Dimostrazione. Per prima cosa si scrivono i vettori $v_1, v_2, \dots, v_k$ come vettori colonna rispetto a una base data, e si considera la matrice M che ha questi vettori colonna come colonne. Poi si riduce M in forma a scalini per colonne. Tutte le volte che troviamo uno scalino lungo (diciamo di altezza $i \geq 2$, dove l'altezza è la differenza di profondità tra due colonne adiacenti che formano lo scalino lungo) 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. Supponiamo infatti che M abbia uno scalino di lunghezza i, ovvero in una certa colonna abbia il pivot alla riga t e nella colonna successiva alla riga $t+i$, e sia $i > 1$. Allora per j che varia tra 1 ed $i-1$, basta scegliere come vettore w_j il vettore definito come segue: w_j ha tutti 0 tranne un 1 in corrispondenza della riga $t+j$-esima. È facile osservare infatti che ogni w_j così costruito non appartiene allo $Span(v_1, \dots, v_k, w_1, \dots, w_{j-1})$. Dunque, per la Proposizione 1.38, ad ogni aggiunta di w_j, l'insieme $\{v_1, \dots, v_k, w_1, \dots, w_{j-1}, w_j\}$ rimane un insieme di vettori linearmente indipendenti di V.

Concludiamo la dimostrazione, e poi esemplificheremo quanto abbiamo spiegato qui in simboli.

Ripetendo questa *aggiunta* di vettori w_k per ogni scalino lungo che troviamo in M, troviamo alla fine n vettori linearmente indipendenti, dunque una base di V come richiesto. $\square$

Esempio 2.20. Illustriamo il metodo descritto con un esempio. Supponiamo che $V = \mathbb{R}^7$ e siano dati i 4 vettori linearmente indipendenti che, scritti rispetto alla base standard di $\mathbb{R}^7$, sono rappresentati come segue:

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

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 :

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

Il primo scalino lungo ha altezza 3 (tra i due pivot della seconda e terza colonna, indicati in neretto, c'è infatti una differenza di profondità di 3). Come osservato nella dimostrazione del Teorema 2.19, 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), e si osserva immediatamente che $\{v_1, v_2, v_3, v_4, w_1, w_2\}$ è un insieme di vettori linearmente indipendenti di $\mathbb{R}^7$.

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 $v_1, v_2, v_3, v_4, w_1, w_2$. A questo punto $\{v_1, v_2, v_3, v_4, w_1, w_2, w_3\}$ sono un insieme di vettori linearmente indipendenti di $\mathbb{R}^7$, e si può concludere che formano una base di $\mathbb{R}^7$, sia osservando che quando si scrive la matrice 7×7 M'' formata da tali vettori la sua forma a scalini ridotta è l'identità (abbiamo proprio aggiunto a M tre vettori che *accorciano* i suoi scalini lunghi...), sia avendo in mente che $\mathbb{R}^7$ ha dimensione 7, perciò ogni insieme di vettori linearmente indipendenti di cardinalità 7 è una sua base.

3. Il teorema della dimensione del nucleo e dell'immagine di una applicazione lineare

Il teorema del completamento, ha come importante corollario un risultato che stabilisce una relazione tra la dimensione del nucleo e quella dell'immagine di una applicazione lineare.

Teorema 2.21. *Considerata una applicazione lineare $L : V \to W$, dove V e W sono spazi vettoriali su un campo $\mathbb{K}$, vale*

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

Dimostrazione. Sia $n = dim\ V$ e consideriamo il $Ker\ L$ che sappiamo essere un sottospazio di V.
Osserviamo che se $Ker\ L = V$ (ovvero ha dimensione n) o $Ker\ L = \{O\}$ (ovvero ha dimensione 0) è facile dimostrare la tesi del teorema. Infatti, nel primo caso L è l'applicazione nulla, e dunque $Imm\ L$, che contiene solo il vettore O, è il sottospazio di dimensione $\{0\}$ di W.

Nel secondo caso, sappiamo che $(L(e_1), \dots, L(e_n))$ (dove con $(e_1, \dots, e_n)$ abbiamo indicato una base di V) genera $Imm\ L$, e, visto che $Ker\ L = \{O\}$, ovvero

L è iniettiva, sappiamo che $(L(e_1), \dots, L(e_n))$ sono linearmente indipendenti (controllate di sapere spiegare bene il perché). Dunque $(L(e_1), \dots, L(e_n))$ costituiscono una base di $Imm\ L$, che ha dimensione n.

Mettiamoci dunque nel caso in cui $dim\ Ker\ L = k$, con $0 < k < n$, e sia $\{z_1, \dots, z_k\}$ una base di $Ker\ L$. Per il teorema del completamento possiamo trovare $w_1, \dots, w_{n-k}$ tali che $\{z_1, \dots, z_k, w_1, \dots, w_{n-k}\}$ sia una base di V. Sappiamo che $Imm\ L$ è il sottospazio generato da

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

ma, essendo gli z_j nel nucleo di L, per ogni $j = 1, \dots, k$, $L(z_j) = O$ e allora

$$Imm\ L =< L(w_1), \dots, L(w_{n-k}) > .$$

Abbiamo per ora dimostrato che $dim\ Imm\ L \leq n - k$. Vorremmo far vedere, per concludere la dimostrazione del teorema, che $dim\ Imm\ L = n - k$. Questo è vero se $L(w_1), \dots, L(w_{n-k})$ sono linearmente indipendenti.

Consideriamo dunque una combinazione lineare nulla degli $L(w_i)$:

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

e facciamo vedere che i coefficienti di tale combinazione lineare devono essere tutti nulli.

Per linearità l'equazione sopra equivale a:

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

ossia

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

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

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

dove i $b_j \in \mathbb{K}$, che diventa

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

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

Definizione 2.22. Una applicazione lineare bigettiva $L : V \to W$, tra due spazi vettoriali V e W sul campo $\mathbb{K}$, si dice un **isomorfismo lineare**.

Dal Teorema 2.21 segue che:

- Se $L : V \to W$ è una applicazione lineare iniettiva allora $dim\ Imm\ L = dim\ V$. Infatti sappiamo che $Ker\ L = \{O\}$, dunque $dim\ Ker\ L = 0$, e quindi dal teorema si ha $dim\ Imm\ L = dim\ V$.
- Se $L : V \to W$ è un isomorfismo lineare allora $dim\ V = dim\ W$. Infatti se L è bigettiva, in particolare è iniettiva (dunque $dim\ Imm\ L = dim\ V$), e surgettiva (dunque $Imm\ L = W$ e quindi $dim\ Imm\ L = dim\ W$).
- Se $L : V \to W$ è una applicazione lineare iniettiva, allora L pensata come applicazione da V ad $Imm\ L$ è un isomorfismo lineare.

4. La riduzione a scalini per colonna vista come moltiplicazione per matrici invertibili

La riduzione a scalini per colonna risulta molto utile anche per lo studio delle applicazioni lineari, ed in particolare per la determinazione di dimensione e base dell'immagine di una applicazione lineare.

Dati due spazi vettoriali V e W sul campo $\mathbb{K}$ di dimensione n e m rispettivamente, consideriamo una applicazione lineare:

$$L : V \to W$$

Fissiamo una base $\{e_1, e_2, \ldots, e_n\}$ di V e una base $\{\epsilon_1, \epsilon_2, \ldots, \epsilon_m\}$ di W. Indichiamo con $[L]$ la matrice, di forma $m \times n$, associata a L nelle basi scelte.

Per quanto abbiamo fin qui detto, possiamo, tramite un numero finito k di operazioni elementari sulle colonne di $[L]$, portarla in forma a scalini ridotta. Ma c'è di più. Ogni operazione elementare sulle colonne corrisponde a moltiplicare la matrice iniziale $[L] \in Mat_{m\times n}(\mathbb{K})$, a destra, per una matrice B $n \times n$ 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 quadrata $n \times n$ (chiamiamola $M_{ij\lambda}$) che ha tutti 1 sulla diagonale, e 0 in tutte le altre caselle eccetto che nella casella alla "riga j, colonna i", dove troviamo λ.

Esercizio 2.23. Dimostrare che moltiplicando $[L]$ in $Mat_{m\times n}(\mathbb{K})$ per la matrice $M_{ij\lambda}$ definita sopra, si ottiene proprio la matrice $[L]'$ uguale ad $[L]$ tranne che per la colonna i-esima, che in $[L]'$ è ottenuta sommando alla colonna i-esima di $[L]$, λ volte la colonna j-esima di $[L]$.

Verifichiamo che $M_{ij\lambda}$ è effettivamente invertibile nell'anello $Mat_{n\times n}(\mathbb{K})$ esibendo la sua inversa; ricordiamo che l'inversa $M_{ij\lambda}^{-1}$ deve soddisfare $M_{ij\lambda}M_{ij\lambda}^{-1} = M_{ij\lambda}^{-1}M_{ij\lambda} = I$. Senza perderci in calcoli, possiamo trovare $M_{ij\lambda}^{-1}$ in maniera più astuta, pensando che ci deve dare la mossa inversa di quella compiuta, ovvero

- si sottrae alla colonna i la colonna j moltiplicata per lo scalare λ, che corrisponde a sommare alla colonna i la colonna j moltiplicata per lo scalare $-\lambda$;

Dunque $M_{ij\lambda}^{-1}$ sarà fatta come $M_{ij\lambda}$, differendo solo per il coefficiente alla riga i e colonna j, in cui, al posto di λ, comparirà $-\lambda$, pertanto secondo la nostra notazione potremo indicarla con $M_{ij(-\lambda)}$.

Esercizio 2.24. Provare che effettivamente

$$M_{ij\lambda} \cdot M_{ij(-\lambda)} = I = M_{ij(-\lambda)} \cdot M_{ij\lambda}$$

Esercizio 2.25. Provare che anche per le mosse corrispondenti alle altre due operazioni elementari sulle colonne si trovano delle matrici invertibili che le realizzano.

Dunque, ridurre a scalini per colonne la matrice $[L]$ equivale a moltiplicare $[L]$ a destra per tante matrici invertibili $[M_1], [M_2], \ldots, [M_k]$ ed avere che $[L][M_1][M_2]\cdots[M_k]$ è in forma a scalini ridotta.

Per semplificare la notazione, chiamiamo $[M]$ la matrice $n \times n$ ottenuta moltiplicando tra loro a destra le matrici $[M_i]$, ovvero:

$$[M] = [M_1] \cdot \ldots \cdot [M_k]$$

$[M]$ è invertibile, visto che è il prodotto di matrici invertibili, e la sua inversa è

$$[M]^{-1} = [M_k]^{-1} \cdot \ldots \cdot [M_1]^{-1}$$

Sappiamo che ad $[M]$ è associata ad una applicazione lineare $M : V \to V$. Ci chiediamo: quale è la applicazione lineare associata a $[L][M]$? È proprio la

$$L \circ M : V \to W$$

dato che il prodotto fra matrici è stato definito in modo da rispettare la composizione fra applicazioni. Vale infatti il teorema, la cui dimostrazione lasciamo come esercizio (con suggerimento):

Teorema 2.26. *Siano V, W, U spazi vettoriali su $\mathbb{K}$, e fissiamo per ciascuno una base. Siano $T : V \to W$, $S : W \to U$ applicazioni lineari. Allora, rispetto alle basi fissate, vale:*

$$[S \circ T] = [S][T]$$

dove nel membro di destra stiamo considerando il prodotto righe per colonne fra matrici.

Dimostrazione. Lasciata per esercizio (suggerimento: applicare le matrici $[S \circ T]$ e $[S][T]$ ai vettori colonna $\begin{pmatrix} 1 \\ 0 \\ ... \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ ... \\ 0 \end{pmatrix}$ etc... e controllare che diano lo stesso risultato). □

Torniamo all'applicazione L: sappiamo già che lo span delle colonne di $[L]$ coincide con lo span delle colonne della matrice ottenuta portando in forma a scalini $[L]$, ovvero che $Imm\ L$ coincide con l'immagine dell'applicazione lineare associata della matrice ottenuta portando in forma a scalini $[L]$ (tale applicazione lineare è $L \circ M$)
.

Riteniamo però istruttivo fornire una nuova dimostrazione di questo fatto.

Proposizione 2.27. *$Imm\ (L \circ M) = Imm\ L$, ossia, scritto con un'altra notazione, $(L \circ M)\ (V) = L(V)$.*

Dimostriamo questa proposizione dimostrando più in generale che

Proposizione 2.28. *Sia*

$$B : V \to V$$

una applicazione lineare invertibile. Allora vale $Imm\ L = Imm\ (L \circ B)$

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

$$Imm\ (L \circ B) = L(B(V)) = L(V)$$

□

Osservazione 2.29. Dopo quanto abbiamo dimostrato, abbiamo un algoritmo in 3 passi che, data una applicazione lineare tra due spazi vettoriali V e W, e fissata una base $\{e_1, e_2, \ldots, e_n\}$ di V e una base $\{\epsilon_1, \epsilon_2, \ldots, \epsilon_m\}$ di W, ci permette di determinare la dimensione ed una base di $Imm\ L$:

(1) Scrivere la matrice $[L]$ associata ad L rispetto alle basi fissate,

(2) ridurre $[L]$ in forma a scalini (che indichiamo con $[S]$) attraverso operazioni elementari su colonna,
(3) Sapendo che $[L]$ e $[S]$ sono associate a due applicazioni lineari distinte, ma con la stessa immagine (ovvero $Imm\ L$), leggiamo le informazioni relative ad $Imm\ L$ attraverso la forma a scalini $[S]$ di $[L]$.

Ma come si *leggono*, tramite $[S]$, le informazioni cercate su $Imm\ L$? Ovvero come si applica il punto 3) sopra?

Osservazione 2.30. Ricordiamo (vedi Esercizio 1.68) che $Imm\ L$ coincide con 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: $Imm\ L$ è il sottospazio vettoriale di W generato da $L(e_1)$, $L(e_2)$,..., $L(e_n)$, che sono i vettori corrispondenti alle colonne di $[L]$. Ma la proposizione garantisce che $Imm\ L = Imm\ (L \circ M)$, dunque $L(V)$ è anche generato dai vettori di W corrispondenti alle colonne di $[L][M]$. Ora, la matrice $[L][M]$ ha delle colonne *semplici* con cui lavorare, visto che è in forma a scalini ridotta. Si vede subito che le colonne non nulle di $[L][M]$ (dove sono presenti i pivot) formano un insieme di vettori linearmente indipendenti di $\mathbb{K}^m$, e dunque che i vettori di W corrispondenti a tali colonne, siccome appunto sono linearmente indipendenti e generano $Imm\ L$, sono una base di $Imm\ L$. In particolare $dim\ Imm\ L$ è uguale al numero di scalini (pivot) della matrice ridotta a scala.

Ricordando il Teorema 2.16, e osservando che il numero di colonne non nulle in una matrice in forma a scalini corrisponde al numero di pivot, possiamo riassumere quanto abbiamo appena detto:

Proposizione 2.31. *Data una matrice A, qualsiasi riduzione a scalini per colonne di A ottenuta attraverso le operazioni elementari su colonna, ha come invariante il numero n di pivot. In particolare, se vediamo A come la matrice associata ad una applicazione lineare L tra due spazi vettoriali V e W rispetto a due basi fissate di V e W, allora il numero n di pivot è la dimensione del sottospazio di W $Imm\ L$.*

A questo punto possiamo introdurre un concetto importante nello studio delle applicazioni lineari, quello di *rango* di una applicazione lineare.

Definizione 2.32. Data una applicazione lineare $L : V \to W$, dove V e W sono due spazi vettoriali di dimensione finita sul campo $\mathbb{K}$, il **rango** di L è il numero $dim\ Imm\ L$.

Le osservazioni fatte fin qui in particolare dimostrano il seguente:

Teorema 2.33. *Data una applicazione lineare L come sopra e fissate le basi, vale che il rango di L è uguale al numero di colonne non nulle che si trovano quando si trasforma $[L]$ in forma a scalini, ovvero al numero di pivot.*

Osservazione 2.34. Se avessimo fissato altre basi avremmo avuto una matrice $[L]$ diversa, ma, trasformandola in forma a scalini, avremmo ancora ovviamente trovato lo stesso numero di colonne non nulle, giacché tale numero è $dim\ Imm\ L$, ossia dipende dalla applicazione (è la dimensione della sua immagine) e non dalle basi scelte.

Osservazione 2.35. Osserviamo che il rango di una applicazione lineare L è anche uguale al **massimo numero di colonne linearmente indipendenti** di $[L]$. Infatti sappiamo che $Imm\ L$ è il sottospazio vettoriale di W generato dai vettori colonna di $[L]$. Da questi vettori, come risulta dal Teorema 1.47, è possibile estrarre una base di $Imm\ L$ e, ricordando la dimostrazione di quel teorema, possiamo dire che $dim\ Imm\ L$ è uguale al massimo numero di colonne linearmente indipendenti di $[L]$.

5. La riduzione a scalini per righe

Nei paragrafi precedenti di questo capitolo abbiamo studiato le operazioni elementari di colonna su una matrice. Possiamo ripetere molte delle cose dette operando sulle righe anziché sulle colonne.

Innanzitutto, data una matrice in $Mat_{m\times n}(\mathbb{K})$, è possibile definire le *mosse* sulle righe, dette anche **operazioni elementari sulle righe** di una matrice, in totale simmetria (ovvero sostituendo alla parola colonna, la parola riga) con quanto fatto per le colonne:

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

Procedendo sempre in maniera simmetrica a quanto fatto nel caso delle colonne, si può definire la forma a scalini per righe di una matrice. Chiamiamo stavolta *profondità* di una riga la posizione occupata, contata da destra, dal suo coefficiente diverso da zero che sta più a sinistra sulla riga. Alla riga nulla (con tutte i coefficienti uguali a 0) assegnamo per convenzione profondità 0.

Esempio 2.36. Ad esempio la riga:

$$(0, 0, 7, \sqrt{2}, 3, 0, 0, 9)$$

ha profondità 6, in quanto, il termine diverso da zero più a sinistra della riga è il 7, e se si comincia a contare da destra della riga, il 7 è alla posizione 6.

A questo punto possiamo dare la definizione di matrice a scalini per righe.

Definizione 2.37. Una matrice A in $Mat_{m\times n}(\mathbb{K})$, si dice **in forma a scalini per righe** se rispetta le seguenti proprietà:

- leggendo la matrice dall'alto verso il basso, le righe non nulle si incontrano tutte prima della righe nulle;
- leggendo la matrice dall'alto verso il basso, le profondità delle sue righe non nulle risultano strettamente decrescenti.

Esercizio 2.38. Dire quale, tra le seguenti matrici, è in forma a scalini per righe:

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

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

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

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

In totale simmetria con quanto dimostrato nel caso delle colonne è possibile dimostrare che:

Teorema 2.39. *Data una matrice A in $Mat_{m\times n}(\mathbb{K})$ è sempre possibile, usando (un numero finito di) operazioni elementari sulle righe, ridurre la matrice in forma a scalini per righe.*

In particolare, anche quando abbiamo una matrice in forma a scalini per riga, si possono definire i **pivot** della matrice, come i coefficienti più a sinistra delle righe non nulle.
Inoltre, anche nel caso delle righe, è possibile definire una forma a scalini *particolare*: la forma a **scalini per righe ridotta**. In questo caso, lasciamo come esercizio al lettore, di dare una definizione formale, presentando prima alcuni esempi di forma a scalini per righe ridotta.

Esempio 2.40. Ecco alcune matrici in forma a scalini per righe ridotta:

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

$$\begin{pmatrix} 1 & 0 & 0 & -20 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 5 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Corollario 2.41. *Ogni matrice A può essere trasformata, attraverso le operazioni elementari sulle righe, in una matrice in forma a scalini per righe ridotta.*

Esercizio 2.42. Dimostrare il Corollario 2.41.

Proseguiamo lo studio delle analogie tra mosse per colonna e mosse per riga cercando di capire a cosa equivale una operazione elementare su riga in termini di operazioni tra matrici.
Ogni singola mossa (operazione elementare) sulle righe di una generica matrice A di $Mat_{m\times n}(\mathbb{K})$, equivale stavolta a moltiplicare A **a sinistra**, per una matrice $m \times m$ invertibile.[2] Per esempio la mossa:

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

corrisponde a moltiplicare la nostra matrice a sinistra per la matrice $m \times m$ 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, che indichiamo con $U_{ij\lambda}$, è la simmetrica rispetto alla diagonale della matrice $M_{ij\lambda}$ analoga usata nel caso delle mosse di colonna. È immediato verificare che è invertibile e $U_{ij\lambda}^{-1}$ è la simmetrica di $M_{ij\lambda}^{-1}$.

Consideriamo a questo punto una applicazione lineare:

$$L \,:\, V \to W$$

dove V e W sono due spazi vettoriali sul campo $\mathbb{K}$, 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 per righe ridotta equivale a dire che moltiplichiamo $[L]$ a sinistra per delle matrici invertibili $[U_1], [U_2], \ldots, [U_s]$ 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, ha per matrice $[U]$.

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

$$U \circ L \,:\, V \to W$$

come sappiamo per il Teorema 2.26.

Però in generale non è vero che $Imm\, L = Imm\,(U \circ L)$. In questo caso infatti, prima si applica L e poi una applicazione U bigettiva su W. Dunque l'immagine di $U \circ L$ sarà uguale all'immagine di U applicata a $Imm(L)$, che non è detto sia uguale ad $Imm(L)$. Vale però che la dimensione rimane invariata.

Teorema 2.43. *Sia*

$$L \,:\, V \to W$$

una applicazione lineare fra due spazi vettoriali V e W sul campo $\mathbb{K}$. Sia

$$B \,:\, W \to W$$

una applicazione lineare invertibile. Allora vale $Ker\ L = Ker\ B \circ L$ e inoltre vale $dim\, Imm\, L = dim\, Imm\,(B \circ L)$, ossia L e $B \circ L$ hanno lo stesso rango.

[2] Teniamo a mente questo cambiamento (da moltiplicazione a destra nel caso delle colonne, a moltiplicazione a sinistra nel caso delle righe, perché spiegherà alcune differenze che ci sono nelle risultato della riduzione per riga e per colonna.

DIMOSTRAZIONE. Dimostriamo per prima cosa che $Ker\ L = Ker\ B \circ L$. Verifichiamo innanzitutto che $Ker\ L \subseteq Ker\ B \circ L$.

Sia v un vettore di $Ker\ L$, allora $L(v) = O$. Ma $(B \circ L)(v) = B(L(v)) = B(O) = O$, per cui $v \in Ker\ B \circ L$.

Verifichiamo ora che $Ker\ B \circ L \subseteq Ker\ L$. Sia v' un vettore che appartiene a $Ker\ B \circ L$. Dobbiamo mostrare che $v' \in Ker\ L$, ossia che $L(v') = O$. Sappiamo che $(B \circ L)(v') = O$, ovvero che $B(L(v')) = O$. Ora, l'applicazione B è bigettiva, dunque iniettiva, e l'unico vettore che manda in O è O per cui deve valere $L(v') = O$.

A questo punto, possiamo applicare il Teorema 2.21 alle due applicazioni L e $B \circ L$. Otteniamo

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

e

$$dim\ Imm\ B \circ L + dim\ Ker\ B \circ L = dimV$$

Da queste due equazioni, sapendo che $Ker\ L = Ker\ B \circ L$, si ricava immediatamente che $dim\, Imm\, L = dim\, Imm\,(B \circ L)$. $\square$

Torniamo al nostro caso della applicazione lineare L tale che la sua matrice $[L]$ ha come forma a scalini per righe ridotta la matrice $[U][L]$. Come sappiamo, per calcolare il nucleo di L in concreto si risolve il sistema lineare

$$[L] \begin{pmatrix} x_1 \\ x_2 \\ ... \\ ... \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ ... \\ 0 \end{pmatrix}$$

Una prima conseguenza del teorema appena dimostrato è che per calcolare il nucleo di L basterà risolvere il sistema lineare

$$[U][L] \begin{pmatrix} x_1 \\ x_2 \\ ... \\ ... \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ ... \\ 0 \end{pmatrix}$$

in cui la matrice $[U][L]$ è molto più semplice perché è in forma a scalini ridotta per riga. Di questo aspetto ci occuperemo più estesamente nel prossimo capitolo, dedicato ai sistemi lineari.

Un'altra conseguenza del Teorema 2.43 è la seguente: 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, come sappiamo dalla Osservazione 2.35). 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:

Teorema 2.44. *Sia*

$$L\ :\ V \to W$$

una applicazione lineare tra gli spazi vettoriali V e W rispettivamente di dimensione n e m sul campo $\mathbb{K}$. Il rango di L (ovvero la dimensione della sua immagine) 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 lo stesso argomento vale 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 il Teorema 2.44 come nel caso della riduzione per colonne:

Teorema 2.45. *Sia*

$$L : V \to W$$

una applicazione lineare tra gli spazi vettoriali V e W rispettivamente di dimensione n e m sul campo $\mathbb{K}$. Il rango di L (ovvero la dimensione della sua immagine) 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 della matrice $[L]$.

6. Ancora sulla riduzione a scala di una matrice e lo studio delle applicazioni lineari

Riassumendo e combinando tra loro i risultati degli ultimi due paragrafi si ha:

Teorema 2.46. *Data una applicazione lineare $L : V \to W$ tra due spazi vettoriali V e W sul campo $\mathbb{K}$, e data la matrice $[L]$ associata a L rispetto a due basi fissate di V e W:*

1. *Il massimo numero di righe linearmente indipendenti di $[L]$ è uguale al massimo numero di colonne linearmente indipendenti di $[L]$, ossia al rango di L.*
2. *Se si riduce la matrice $[L]$ in forma a scalini, sia che lo si faccia per righe, sia che lo si faccia per colonne, il numero di scalini che otterremo (ovvero il numero di pivot della matrice) sarà sempre uguale al rango di L.*

Definizione 2.47. Data una matrice M, chiamiamo **rango della matrice** il massimo numero di colonne (o righe, abbiamo visto che è lo stesso) linearmente indipendenti di M.

Osservazione 2.48. Con la definizione appena introdotta, e dal teorema 2.46, si ha che il rango di una applicazione lineare L coincide con quello di una sua matrice associata $[L]$ rispetto a due basi fissate. Dunque possiamo usare la parola rango senza fare troppa attenzione, applicandola sia alle matrici sia alle applicazioni.

Sappiamo inoltre 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.

Dall'Osservazione 2.48 segue un algoritmo piuttosto semplice per calcolare il rango (ovvero la dimensione dell'immagine) di una applicazione lineare L:

1. Scriviamo la matrice associata $[L]$ rispetto ad una qualunque coppia di basi.
2. Riduciamo $[L]$ a scalini usando sia mosse di riga sia mosse di colonna nell'ordine che ci torna più comodo.
3. Contiamo il numero di scalini (pivot) della matrice ridotta a scala.

Lo schema in Figura 1 riassume graficamente le relazioni tra le riduzioni a scala (per riga e per colonna) di una matrice $[L]$ associata ad una applicazione lineare L tra due spazi vettoriali V e W discusse fin qui.

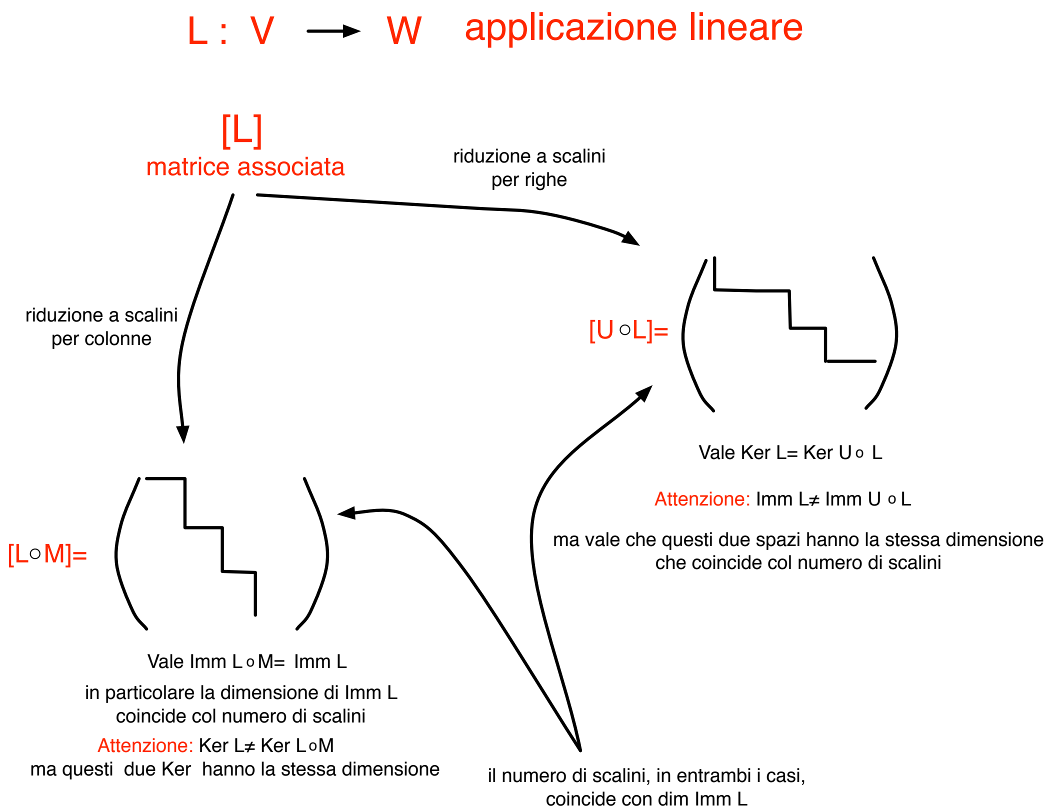


FIGURA 1. Schema riassuntivo sulla riduzione per righe e per colonne

Dunque, data L applicazione lineare tra due spazi vettoriali V e W sul campo $\mathbb{K}$, attraverso la riduzione a scala per colonne di una matrice associata a L, si può determinare una base di $Imm\ L$ considerando i vettori corrispondenti alle colonne non nulle della matrice così ottenuta. La base che troviamo con questo metodo ha il pregio di essere *a scalini*, e dunque si presta bene per alcuni scopi: ad esempio è quasi immediato, diremmo quasi *a colpo d'occhio*, stabilire se un certo vettore v appartiene ad $Imm\ L$, ovvero è combinazione lineare dei vettori di questa base oppure no.

Ma come abbiamo già accennato, e come vedremo anche nel prossimo capitolo, per determinare il $Ker\ L$, può essere comodo portare la matrice a scalini per righe: in questo caso dobbiamo necessariamente calcolare anche la forma a scalini per colonne per determinare una base di $Imm\ L$? La risposta a questa domanda è no, come spiegheremo tra poco: dalla forma a scalini per righe di una matrice associata ad L, si ricava una informazione che risulta utile per trovare una base di $Imm\ L$, che in generale però non sarà a scalini. Vediamo come partendo da un esempio.

Supponiamo che una applicazione lineare L, tra due spazi vettoriali V e W su $\mathbb{K}$, rispetto a certe basi fissate in V e W, sia rappresentata dalla matrice:

$$[L] = \begin{pmatrix} 1 & 2 & 5 & 7 & 18 \\ 4 & 8 & 14 & 20 & 54 \\ 2 & 4 & 7 & 10 & 28 \\ 3 & 6 & 9 & 13 & 36 \end{pmatrix}$$

Con operazioni di riga, si può portare $[L]$ nella forma a scalini per righe seguente (dove in neretto abbiamo indicato i pivot, e dunque la posizione degli scalini):

$$R = \begin{pmatrix} \mathbf{1} & 2 & 2 & 3 & 4 \\ 0 & 0 & \mathbf{3} & 4 & 6 \\ 0 & 0 & 0 & 0 & \mathbf{2} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Osserviamo che gli scalini sono nella prima, nella terza e nella quinta colonna. Quello che vogliamo dimostrare è che da questo segue che la prima, la terza e la quinta colonna **della matrice iniziale** $[L]$ (ATTENZIONE! Proprio della matrice iniziale $[L]$, non della sua forma ridotta a scalini per riga R) sono linearmente indipendenti, e quindi una base di $Imm\ L$ è data dai vettori corrispondenti alla prima (v_1), alla terza (v_2), e alla quinta colonna (v_3) di $[L]$:

$$\begin{pmatrix} 1 \\ 4 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 5 \\ 14 \\ 7 \\ 9 \end{pmatrix}, \begin{pmatrix} 18 \\ 54 \\ 28 \\ 36 \end{pmatrix}$$

Cerchiamo di capire come mai è proprio così. Supponiamo di estrarre dalla matrice iniziale la matrice più piccola in cui abbiamo tenuto solo la prima, la terza e la quinta colonna:

$$\begin{pmatrix} 1 & 5 & 18 \\ 4 & 14 & 54 \\ 2 & 7 & 28 \\ 3 & 9 & 36 \end{pmatrix}$$

Se su tale matrice facciamo le stesse identiche mosse di riga che avevamo fatto sulla $[L]$ per ottenere la sua forma a scalini per riga A, che matrice otterremo? Senza fare nessun conto, sappiamo già che otterremo la seguente matrice:

$$\begin{pmatrix} 1 & 2 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Infatti durante le operazioni di riga ciascuna colonna si trasforma senza interagire con le altre colonne, dunque nel nostro caso le tre colonne si trasformano esattamente come si erano trasformate prima.

A questo punto però sappiamo che i tre vettori

$$\begin{pmatrix} 1 \\ 4 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 5 \\ 14 \\ 7 \\ 9 \end{pmatrix}, \begin{pmatrix} 18 \\ 54 \\ 28 \\ 36 \end{pmatrix}$$

sono linearmente indipendenti, perché la matrice che essi formano ha rango 3 (questo lo sappiamo perché abbiamo mostrato che le operazioni elementari per riga - o colonna - non modificano il rango della matrice). Ora il rango di L, ovvero $dim\ Imm\ L$, è proprio 3 (la matrice ridotta a scalini per righe aveva appunto 3 scalini). Dunque tali vettori sono una base di $Imm\ L$.

Ricalcando questa dimostrazione, si può dimostrare in generale che:

Proposizione 2.49. *Se A è una matrice $m \times n$ a valori su un campo $\mathbb{K}$, ed indichiamo con $e_1, \ldots, e_n$ le sue colonne, e B è una riduzione a scalini per righe di A, allora le colonne di A in corrispondenza alla posizione dei pivot di B formano una base dello Span delle colonne di A ($Span(e_1, \ldots, e_n)$).*

Osservazione 2.50. Si può anche considerare il caso 'simmetrico' a quello discusso: se abbiamo una matrice ridotta a scalini per colonna e vediamo che gli scalini sono, per esempio, nella seconda, nella terza e nella quarta riga, possiamo dedurre che la seconda, la terza e la quarta riga *della matrice iniziale* sono una base dello spazio generato dalle righe. A cosa potrebbe essere utile questa affermazione? Rifletteteci per esercizio, partendo dalla considerazione che il sistema dato dalle righe della matrice iniziale è equivalente al sistema in cui si considerano solo la seconda, la terza e la quarta riga...

Osservazione 2.51. Quello che abbiamo dimostrato e riassunto nella Proposizione 2.49, permette di fornire un algoritmo per *estrarre* una base di uno spazio vettoriale, a partire da un insieme di generatori. Cosa che teoricamente sappiamo essere sempre possibile, a seguito del Teorema 1.47.
Sia infatti V uno spazio vettoriale, di dimensione n e di cui conosciamo una base $\{e_1, \ldots, e_n\}$. Consideriamo l vettori $v_1, \ldots, v_l$ di V, ed, in particolare, il sottospazio vettoriale di V $Span(v_1, \ldots, v_l)$. Se vogliamo estrarre una base di $Span(v_1, \ldots, v_l)$ da $v_1, \ldots, v_l$ (che sappiamo, per definizione, generare $Span(v_1, \ldots, v_l)$), scriviamo le coordinate dei v_i rispetto alla base $\{e_1, \ldots, e_n\}$ e le consideriamo (ogni n-upla di coordinate del singolo v_i) come colonne di una matrice M. A questo punto si porta M in forma a scalini per riga (chiamiamo la matrice a scalini per riga S), e, i vettori v_i con indice i in corrispondenza dei pivot di S formano una base di $Span(v_1, \ldots, v_l)$, estratta dall'insieme di generatori $\{v_1, \ldots, v_l\}$.

7. Altri esercizi

Esercizio 2.52. Consideriamo l'applicazione lineare $L_a : \mathbb{R}^3 \to \mathbb{R}^3$ dipendente dal parametro reale a e definita da:

$$L_a(x, y, z) = (ax + y + z, x + ay + z, -x + y + az)$$

(1) Trovare per quali valori di a l'applicazione L_a non è surgettiva.
(2) Fissato un valore $\bar{a}$ di a per cui L_a non è surgettiva, determinare la dimensione di $Ker(L_{\bar{a}})$ e di $Imm(L_{\bar{a}})$, ed una base di quest'ultimo spazio vettoriale.

Svolgimento Scriviamo la matrice associata ad L_a rispetto alla base *canonica* e_1, e_2, e_3 di $\mathbb{R}^3$ (sia "in partenza", che come spazio "di arrivo").

Per scrivere la matrice $[L_a]$ associata ad L_a rispetto a questa base, bisogna calcolare le coordinate di $L_a(e_1),\ L_a(e_2),\ L_a(e_3)$ rispetto alla base e_1, e_2 ed e_3:

$$\begin{cases} L_a(1,0,0) = (a, 1, -1) \\ L_a(0,1,0) = (1, a, 1) \\ L_a(0,0,1) = (1, 1, a) \end{cases}$$

Dunque:

$$[L_a] = \begin{pmatrix} a & 1 & 1 \\ 1 & a & 1 \\ -1 & 1 & a \end{pmatrix}$$

Riduciamo $[L_a]$ a scalini per righe, attraverso operazioni elementari di riga:

$$[L_a] \overbrace{\rightarrow}^{(1)\circlearrowleft(2)} A_1 = \begin{pmatrix} 1 & a & 1 \\ a & 1 & 1 \\ -1 & 1 & a \end{pmatrix} \underbrace{\overbrace{\rightarrow}^{(2)=(2)-a(1)}}_{(3)=(3)+(1)} A_2 = \begin{pmatrix} 1 & a & 1 \\ 0 & 1-a^2 & 1-a \\ 0 & a+1 & a+1 \end{pmatrix}$$

Se $1 - a^2$ è diverso da 0, ovvero se $a \neq \pm 1$, possiamo dividere la seconda riga per $1 - a^2$ e avere uno scalino sulla seconda riga, in corrispondenza della seconda colonna. Dunque, prima di andare avanti, studiamo a parte i casi *sfortunati*, ovvero $a = 1$ e $a = -1$, per cui non si può procedere così.

- Se $a = 1$ la matrice A_2 è:

  $$A_2 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 2 & 2 \end{pmatrix}$$

  Dunque non è ancora a scalini per righe, ma basta fare la mossa di riga di scambiare la seconda con la terza riga per avere la forma a scalini cercata:

  $$A_3 = \begin{pmatrix} \mathbf{1} & 1 & 1 \\ 0 & \mathbf{2} & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

  Si vede subito (si vedeva anche da A_2 a dire il vero...) che il rango di A_2, che sappiamo essere uguale a rango di $[L_1]$, che è a sua volta uguale alla $dim\ Imm\ L_1$, è 2 (tanto quanti sono i pivot, evidenziati in neretto nella matrice A_3). Dunque L_1 non è surgettiva (la dimensione di $\mathbb{R}^3$ è 3), e la dimensione del nucleo (dal Teorema 2.21) è 1.
- Se $a = -1$ la matrice A_2 è:

  $$A_2 = \begin{pmatrix} \mathbf{1} & -1 & 1 \\ 0 & 0 & \mathbf{2} \\ 0 & 0 & 0 \end{pmatrix}$$

  che è già in forma a scalini per righe. La lettura della matrice A_2, ci dice che anche L_{-1} non è surgettiva, e anche in questo caso $dim\ Imm\ L_{-1} = 2$ e $dim\ Ker\ L_{-1} = 1$.

Occupiamoci dei casi diversi dai due precedenti (stiamo suppondendo dunque che $a \neq \pm 1$):

$$A_2 \underbrace{\rightarrow} A_3 = \begin{pmatrix} 1 & a & 1 \\ 0 & 1 & \frac{1}{1+a} \\ 0 & 0 & a \end{pmatrix}$$

In questo caso A_3 è in forma a scalini per righe, ma il numero di pivot dipende dal valore di a. Se $a = 0$ allora A_3 ha 2 pivot, e dunque L_2 non è surgettiva, altrimenti L_a ha tre pivot e dunque è surgettiva (e il nucleo ha, in questo caso, dimensione 0). Concludendo il primo punto dell'esercizio, L_a è surgettiva per ogni valore di a diverso da $-1, 1, 0$. In questi 3 casi invece l'immagine ha dimensione 2 e il nucleo 1.

Per rispondere alla seconda domanda, fissiamo per esempio $a = 1$. Conosciamo già la riduzione a scalini per righe di L_1 che è:

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

Perciò una base di $Imm\ L_1$ è data dai vettori colonna della matrice $[L_1]$, associata ad L_1 rispetto alla base canonica, che si trovano in corrispondenza dei pivot (ovvero dalla prima e dalla seconda colonna). Scriviamo L_1:

$$[L_1] = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}$$

E dunque una base di $Imm\ L_1$ è:

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

Esercizio 2.53. Si consideri l'applicazione lineare $L : \mathbb{R}^5 \to \mathbb{R}^4$ che è data, nelle basi standard di $\mathbb{R}^5$ e $\mathbb{R}^4$, dalla seguente matrice:

$$\begin{pmatrix} 1 & 3 & 4 & 0 & 0 \\ 6 & 9 & 11 & 8 & 3 \\ 5 & 6 & 7 & 8 & 3 \\ 1 & 0 & 0 & 2 & 0 \end{pmatrix}$$

e consideriamo poi i 4 vettori di $\mathbb{R}^4$:

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 4 \\ 4 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 3 \\ 3 \\ 0 \end{pmatrix}, v_4 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 2 \end{pmatrix}$$

a) Verificare che $\{v_1, v_2, v_3, v_4\}$ è una base di $\mathbb{R}^4$.
b) Scrivere la matrice associata all'applicazione L rispetto alla base standard di $\mathbb{R}^5$ e alla base $\{v_1, v_2, v_3, v_4\}$ di $\mathbb{R}^4$.
c) L'applicazione L è surgettiva ?

Esercizio 2.54. Siano V, W spazi vettoriali di dimensione finita sul campo $\mathbb{K}$ e sia $L : V \to W$ una applicazione lineare di rango r. Dimostrare che esistono una base di V e una base di W tali che la matrice $[L]$ associata a L rispetto a tali basi abbia la forma:

$$[L] = (a_{ij}) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & ... \end{pmatrix}$$

in cui solo r coefficienti sono diversi da zero (e uguali a 1), ossia $a_{ij} = 0$ eccetto i coefficienti $a_{11} = a_{22} = \dots = a_{rr} = 1$.

Esercizio 2.55. Si consideri la trasformazione lineare $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definita da:

$$f(x, y, z) = (x - 2y - z, x + y + z)$$

Scrivere la matrice A di tale trasformazione rispetto alle basi canoniche di $\mathbb{R}^3$ e $\mathbb{R}^2$ e determinare una base dell'immagine e del nucleo di f.

Svolgimento. Per trovare la matrice A associata ad f rispetto alle basi canoniche di $\mathbb{R}^3$ e $\mathbb{R}^2$ è sufficiente calcolare i coefficienti, rispetto alla base canonica di $\mathbb{R}^2$, dell'immagine degli elementi della base canonica di $\mathbb{R}^3$. Tali coefficienti costituiscono le colonne ordinate della matrice cercata. Dunque la prima colonna di A è data dai coefficienti dell'immagine di $(1, 0, 0)$, la seconda colonna dai coefficienti dell'immagine di $(0, 1, 0)$ e la terza colonna dai coefficienti dell'immagine di $(0, 0, 1)$. Calcoliamoli:

$$f(1, 0, 0) = (1, 1) \quad f(0, 1, 0) = (-2, 1) \quad f(0, 0, 1) = (-1, 1)$$

Dunque la matrice A cercata è la seguente:

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

Per trovare una base dell'immagine di f (che è generata dalle colonne della matrice A) si osserva facilmente che la matrice ha rango 2 (perché?), dunque basta individuare due colonne linearmente indipendenti, per esempio le prime due. Dunque

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

costituiscono una base di $Im(f)$.

Sappiamo per il Teorema 2.21 che $Ker\ f$ ha dimensione 1 in quanto:

$$\underbrace{dim(\mathbb{R}^3)}_{=3} = \underbrace{dim\ Im(f)}_{=2} + dim\ Ker\ f$$

Per trovare una base di $Ker\ f$ con operazioni di riga trasformiamo A nella matrice a scalini per righe B:

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

I vettori $w = (x, y, z)$ di $\mathbb{R}^3$ appartenenti a $Ker\ f$, sono le soluzioni del sistema omogeneo $Aw = 0$, che è equivalente al sistema omogeneo $Bw = 0$, ovvero sono i vettori di $\mathbb{R}^3$ le cui coordinate risolvono il sistema:

$$\begin{cases} x - 2y - z = 0 \\ 3y + 2z = 0 \end{cases}$$

Il sistema ha una variabile libera z, e si ottiene che x e y devono essere:

$$x = -\frac{1}{3}z \quad y = -\frac{2}{3}z$$

Il generico vettore di $Ker\ f$ è dunque, al variare del valore t di z in $\mathbb{R}$, del tipo:

$$\begin{pmatrix} -\frac{1}{3}z \\ -\frac{2}{3}z \\ z \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ -\frac{2}{3} \\ 1 \end{pmatrix} z$$

E dunque il vettore:

$$\begin{pmatrix} -\frac{1}{3} \\ -\frac{2}{3} \\ 1 \end{pmatrix}$$

è una base di $Ker\ f$.

Esercizio 2.56 (Attenzione, questo esercizio presenta una tecnica interessante). Consideriamo lo spazio vettoriale $\mathbb{R}_{\leq 3}[x]$ su $\mathbb{R}$ dei polinomi a coefficienti reali, di grado minore o uguale a 3. Sia V il sottospazio di $\mathbb{R}_{\leq 3}[x]$, generato dai seguenti polinomi:

$$p_1(x) = x^3 - 2x + 3,\ p_2(x) = x^3 + 1,\ p_3(x) = x - 5,\ p_4(x) = 2$$

Ovvero $V = Span(p_1(x),\ p_2(x),\ p_3(x),\ p_4(x))$.

(1) Estrarre una base di V dall'insieme $\{p_1(x),\ p_2(x),\ p_3(x),\ p_4(x)\}$.
(2) Completare la base di V ad una base di $\mathbb{R}_{\leq 3}[x]$.

Svolgimento: $\mathbb{R}_{\leq 3}[x]$ è uno spazio vettoriale su $\mathbb{R}$ di dimensione 4, una cui base è $\{1, x, x^2, x^3\}$.

Facciamo vedere come *in un colpo solo* possiamo rispondere ad entrambe le domande dell'esercizio. Scriviamo le coordinate (le indichiamo in neretto) dei vettori che generano V rispetto alla base di $\mathbb{R}_{\leq 3}[x]$ riportata sopra:

$$\begin{aligned} p_1(x) &= \mathbf{3} \cdot 1 + (\mathbf{-2}) \cdot x + \mathbf{0} \cdot x^2 + \mathbf{1} \cdot x^3 \\ p_2(x) &= \mathbf{1} \cdot 1 + \mathbf{0} \cdot x + \mathbf{0} \cdot x^2 + \mathbf{1} \cdot x^3 \\ p_3(x) &= \mathbf{-5} \cdot 1 + \mathbf{1} \cdot x + \mathbf{0} \cdot x^2 + \mathbf{0} \cdot x^3 \\ p_4(x) &= \mathbf{2} \cdot 1 + \mathbf{0} \cdot x + \mathbf{0} \cdot x^2 + \mathbf{0} \cdot x^3 \end{aligned}$$

A questo punto scriviamo la matrice A, 4×8, che ha per colonne le coordinate dei vettori $(p_1(x),\ p_2(x),\ p_3(x),\ p_4(x),\ 1,\ x,\ x^2,\ x^3)$, rispettando questo ordine:

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

Portiamo A in forma a scalini per riga, innanzitutto scambiando l'ordine delle righe e portando la quarta riga come prima riga:

$$A_1 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 3 & 1 & -5 & 2 & 1 & 0 & 0 & 0 \\ -2 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$A_1 \xrightarrow{[2]=[2]-3[1],[3]=[3]+2[1]} A_2 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & -2 & -5 & 2 & 1 & 0 & 0 & -3 \\ 0 & 2 & 1 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$A_2 \xrightarrow{[3]=[3]+[2]} A_3 = \begin{pmatrix} \mathbf{1} & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & -\mathbf{2} & -5 & 2 & 1 & 0 & 0 & -3 \\ 0 & 0 & -\mathbf{4} & 2 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 \end{pmatrix}$$

A_3 è in forma a scalini per riga, ed ha i pivot (evidenziati in neretto) nella prima, seconda, terza e settima colonna. Questo ci dice che V ha dimensione 3, che una sua base è quella formata dai vettori $p_1(x) = x^3 - 2x + 3$, $p_2(x) = x^3 + 1$, $p_3(x) = x - 5$, e che questa base è completata ad una base di $\mathbb{R}_{\leq 3}[x]$ aggiungendo il vettore corrispondente ai coefficienti della settima colonna, ovvero x^2.
È importante osservare come se avessimo applicato lo stesso algoritmo dopo aver dato un diverso ordine alla base di V, di $\mathbb{R}_{\leq 3}[x]$, o di entrambe, avremmo ottenuto un altro risultato. Del resto, sappiamo fin dal primo capitolo che la base di uno spazio vettoriale non è unica.

CAPITOLO 3

Sistemi lineari

1. Risolvere un sistema usando le operazioni elementari di riga

Illustreremo un metodo molto conveniente per risolvere sistemi lineari di equazioni.

Sappiamo già risolvere sistemi *omogenei*, ossia quelli in cui tutte le equazioni hanno la forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$$

Sono i sistemi con cui abbiamo avuto a che fare quando abbiamo dovuto calcolare il nucleo di una applicazione lineare. Vogliamo mostrare che il metodo che abbiamo usato (ridurre la matrice associata a scalini per righe), si estende anche al caso di sistemi non omogenei. Questo metodo è noto anche col nome *metodo di eliminazione di Gauss*.

Cominciamo con un esempio. Consideriamo il sistema

$$\begin{cases} x + 2y + 2z + 2t &= 1 \\ x + 5y + 6z - 2t &= -5 \\ 8x - y - 2z - 2t &= 0 \\ 2y + 6z + 8t &= 3 \end{cases}$$

Per prima cosa, osserviamo che tutte le informazioni del sistema sono contenute nella seguente matrice di numeri (la *matrice completa associata al sistema*):

$$M = \begin{pmatrix} 1 & 2 & 2 & 2 & 1 \\ 1 & 5 & 6 & -2 & -5 \\ 8 & -1 & -2 & -2 & 0 \\ 0 & 2 & 6 & 8 & 3 \end{pmatrix}.$$

Ogni riga contiene i coefficienti di una delle equazioni (per esempio la terza equazione $8x - y - 2z - 2t = 0$ è 'codificata' dalla terza riga riga $(8 - 1 - 2 - 2\ 0)$).

Sia $S \subset \mathbb{R}^4$ l'insieme delle soluzioni del sistema, ovvero il sottoinsieme di $\mathbb{R}^4$ costituito dai vettori $\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$ tali che, se poniamo $a = x, b = y, c = z, d = t$, tutte le equazioni del sistema diventano delle uguaglianze vere.

Si osserva subito che $\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$ appartiene a S se e solo se il vettore $\begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix} \in \mathbb{R}^5$

soddisfa

$$\begin{pmatrix} 1 & 2 & 2 & 2 & 1 \\ 1 & 5 & 6 & -2 & -5 \\ 8 & -1 & -2 & -2 & 0 \\ 0 & 2 & 6 & 8 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

cioè

$$M \begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Come sappiamo dal Paragrafo 5 del Capitolo 2, possiamo agire sulle righe di M con le mosse elementari di riga fino a ridurla a scalini per righe. Ci sono vari modi per farlo. Uno di essi ci porta alla seguente matrice M':

$$M' = \begin{pmatrix} 1 & 2 & 2 & 2 & 1 \\ 0 & 1 & \frac{4}{3} & -\frac{4}{3} & -2 \\ 0 & 0 & 2 & -\frac{122}{7} & -18 \\ 0 & 0 & 0 & 2 & \frac{259}{139} \end{pmatrix}.$$

Sempre dal Paragrafo 5 del Capitolo 2 sappiamo anche che $M' = RM$ dove R è una matrice 4×4 invertibile.

Ora osserviamo che un vettore di $\mathbb{R}^5$ della forma $\begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix}$ soddisfa

$$M \begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{1.1}$$

se e solo se soddisfa

$$M' \begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{1.2}$$

Infatti se $\begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix}$ soddisfa la (1.1) allora

$$M' \begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix} = RM \begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix} = R \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Viceversa, se $\begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix}$ soddisfa la (1.2) allora vuol dire che

$$RM \begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

ma, poiché R è invertibile, possiamo moltiplicare entrambi i membri dell'uguaglianza per R^{-1} ottenendo

$$R^{-1}RM \begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix} = R^{-1} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

ossia

$$M \begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

dunque $\begin{pmatrix} a \\ b \\ c \\ d \\ -1 \end{pmatrix}$ soddisfa la (1.1).

Abbiamo dimostrato che l'insieme $S \subseteq \mathbb{R}^4$ delle soluzioni del sistema associato alla matrice M coincide con l'insieme delle soluzioni del sistema associato alla matrice ridotta a scalini M'. Per trovare le soluzioni del sistema iniziale, dunque, possiamo studiare le soluzioni del sistema

$$\begin{cases} x + 2y + 2z + 2t & = & 1 \\ y + \frac{4}{3}z - \frac{4}{3}t & = & -2 \\ 2z - \frac{122}{7}t & = & -18 \\ 2t & = & \frac{259}{139} \end{cases}$$

Ciò è un grande vantaggio, perché questo sistema (come tutti i sistemi associati a matrici ridotte a scalini per righe) si risolve immediatamente: si ricava dall'ultima equazione $t = \frac{259}{278}$, poi si sostituisce questo valore di t nella penultima equazione e si ricava un valore per z e così via...troveremo la soluzione del sistema (sottolineiamo che, in questo caso, c'è una sola la soluzione).

Quanto abbiamo illustrato per questo esempio vale in generale per qualunque sistema di equazioni lineari, con dimostrazione analoga. Abbiamo dunque il seguente:

Teorema 3.1. *L'insieme delle soluzioni di un sistema di equazioni lineari associato alla matrice M (a coefficienti nel campo $\mathbb{K}$) coincide con l'insieme delle soluzioni del sistema associato alla matrice M' ottenuta riducendo M, attraverso operazioni di riga, in forma a scalini per righe (o a scalini per righe ridotta).*

Osserviamo innanzitutto che questo metodo può talvolta portare a sistemi finali rappresentati da matrici a scalini del tipo:

$$M' = \begin{pmatrix} 1 & 0 & 2 & 2 & 1 \\ 0 & 0 & 2 & 15 & -18 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}.$$

Come si vede, c'è una riga che ha tutti i coefficienti uguali a 0 salvo l'ultimo: (0 0 0 0 4). Questo significa che il sistema non ammette soluzioni, poiché la corrispondente equazione $0x + 0y + 0z + 0t = 4$ non ha soluzioni. Dunque neppure il sistema iniziale, associato ad M, ammette soluzioni.

Esercizio 3.2. Esprimere in generale il contenuto dell'osservazione qui sopra, ossia dimostrare che, dato un sistema con matrice associata M, e chiamata $\overline{M}$ la sottomatrice di M ottenuta togliendo l'ultima colonna (chiamata talvolta la *matrice incompleta associata al sistema*), il sistema ammette soluzione se e solo se il rango di M è uguale al rango di $\overline{M}$.

Supponiamo ora che per un certo sistema l'insieme delle soluzioni S non sia vuoto: quali ulteriori caratteristiche possiede questo insieme?

Consideriamo un sistema lineare con m equazioni a coefficienti in $\mathbb{K}$ e n incognite $x_1, x_2, \ldots, x_n$, e sia M la matrice associata (tale matrice risulta di formato $m \times (n+1)$). Innanzitutto è facile dimostrare che, se il sistema lineare è *omogeneo*, ossia se l'ultima colonna della matrice M ha i coefficienti tutti uguali a 0, l'insieme S delle soluzioni non è vuoto (contiene infatti il vettore O) ed è un sottospazio vettoriale di $\mathbb{R}^n$ (vedi l'Esercizio 1.76; un modo conveniente per dimostrare che è un sottospazio è pensare all'insieme delle soluzioni come nucleo dell'applicazione lineare rappresentata dalla matrice $\overline{M}$). Invece, se il sistema non è omogeneo, si osserva che il vettore O non appartiene a S, dunque S non è un sottospazio vettoriale.

Studiamo prima il caso dei sistemi omogenei: come è possibile capire che dimensione ha il sottospazio S?

Basta guardare la forma della matrice a scalini M'. Se possiede k scalini (ossia se il rango di M', che del resto è uguale al rango di M, è uguale a k), allora S ha dimensione $n - k$. Infatti possiamo pensare ad S come al nucleo della applicazione lineare $\phi : \mathbb{K}^n \to \mathbb{K}^m$ che rispetto alle basi standard è rappresentata dalla matrice $\overline{M}'$, ossia dalla matrice M' meno l'ultima colonna. Ricordiamo (vedi Teorema 2.46) che il rango di ϕ, cioè la dimensione di $Imm\ \phi$, è uguale a k. Dunque, per il Teorema

2.21 sappiamo che $dim\ Ker\ \phi + dim\ Imm\ \phi = n$, ovvero $dim\ Ker\ \phi + k = n$ da cui, dato che $Ker\ \phi = S$, ricaviamo $dim\ S = n - k$.

Osservazione 3.3. In concreto questo significa che, nel risolvere il sistema, ogni scalino lungo lascerà “libere” alcune variabili, come vediamo nel seguente esempio. Supponiamo che un certo sistema omogeneo a coefficienti in $\mathbb{R}$ conduca alla matrice a scalini:

$$M' = \begin{pmatrix} 1 & 0 & 2 & 2 & 0 \\ 0 & 1 & \sqrt{3} & 12 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Allora il sistema finale associato è

$$\begin{cases} x + 2z + 2t & = & 0 \\ y + \sqrt{3}z + 12t & = & 0 \\ 6t & = & 0 \end{cases}$$

Risolvendolo, otteniamo dall'ultima equazione $t = 0$ e, sostituendo, $y = -\sqrt{3}z$ e $x = -2z$. La variabile z resta “libera” e l'insieme delle soluzioni è il seguente sottospazio di $\mathbb{R}^4$:

$$S = \{ \begin{pmatrix} -2z \\ -\sqrt{3}z \\ z \\ 0 \end{pmatrix} \mid z \in \mathbb{R}\} = \{z \begin{pmatrix} -2 \\ -\sqrt{3} \\ 1 \\ 0 \end{pmatrix} \mid z \in \mathbb{R}\} =< \begin{pmatrix} -2 \\ -\sqrt{3} \\ 1 \\ 0 \end{pmatrix} >$$

Cosa possiamo dire invece di S se il sistema non è omogeneo e ammette soluzioni? Sia M la matrice associata al sistema e sia M_o la matrice che si ricava da M ponendo uguali a 0 tutti i coefficienti dell'ultima colonna. Possiamo pensare M_o come la matrice associata al sistema omogeneo ottenuto dal sistema iniziale ponendo uguali a 0 tutti i membri di destra delle equazioni. Chiamiamo S_o le soluzioni di questo sistema omogeneo e sia $v = \begin{pmatrix} a_1 \\ a_2 \\ ... \\ a_n \end{pmatrix}$ un elemento di S.

Teorema 3.4. *Con le notazioni introdotte sopra, vale che*

$$S = v + S_o = \{v + w \mid w \in S_o\}$$

ossia le soluzioni del sistema iniziale si ottengono tutte sommando il vettore v alle soluzioni del sistema omogeneo.

Dimostrazione. Sia $w = \begin{pmatrix} b_1 \\ b_2 \\ ... \\ b_n \end{pmatrix} \in S_o$. Vogliamo mostrare che $v + w \in S$.

Sia $\gamma_1 x_1 + \gamma_2 x_2 + \cdots + \gamma_n x_n = \delta$ una equazione del sistema. Allora $b_1, b_2, \ldots, b_n$ verificano

$$\gamma_1 b_1 + \gamma_2 b_2 + \cdots + \gamma_n b_n = 0$$

mentre $a_1, a_2, \ldots, a_n$ verificano

$$\gamma_1 a_1 + \gamma_2 a_2 + \cdots + \gamma_n a_n = \delta$$

Dunque

$$\gamma_1(a_1 + b_1) + \gamma_2(a_2 + b_2) + \cdots + \gamma_n(a_n + b_n) =$$
$$= (\gamma_1 a_1 + \gamma_2 a_2 + \cdots + \gamma_n a_n) + (\gamma_1 b_1 + \gamma_2 b_2 + \cdots + \gamma_n b_n) = \delta + 0 = \delta$$

Ripetendo questa osservazione per tutte le equazioni del sistema, si verifica dunque che $v + w \in S$.

Viceversa, sia $p = \begin{pmatrix} c_1 \\ c_2 \\ \cdots \\ c_n \end{pmatrix} \in S$. Vogliamo dimostrare che $p \in v + S_o$.

Osserviamo che $c_1, c_2, \ldots, c_n$ verificano

$$\gamma_1 c_1 + \gamma_2 c_2 + \cdots + \gamma_n c_n = \delta$$

Dunque

$$\gamma_1(a_1 - c_1) + \gamma_2(a_2 - c_2) + \cdots + \gamma_n(a_n - c_n) =$$
$$= (\gamma_1 a_1 + \gamma_2 a_2 + \cdots + \gamma_n a_n) - (\gamma_1 c_1 + \gamma_2 c_2 + \cdots + \gamma_n c_n) = \delta - \delta = 0$$

Ripetendo questa osservazione per tutte le equazioni del sistema dimostriamo che $v - p \in S_o$, dunque possiamo scrivere $p - v = w_o$ dove w_o è un certo elemento di S_o. Allora $p = v + w_o$ ossia $p \in v + S_o$. □

Corollario 3.5. *L'insieme S delle soluzioni di un sistema lineare non omogeneo, a coefficienti in $\mathbb{K}$, con m equazioni e n incognite, o è vuoto oppure è il* traslato *di un sottospazio vettoriale di $\mathbb{K}^n$, ossia è della forma $v+S_o$, dove S_o (l'insieme delle soluzioni del sistema omogeneo associato) è un sottospazio vettoriale di dimensione uguale a $n -$ (rango di M_o).*

2. Altri esercizi

Esercizio 3.6. Discutere la risolubilità, ed eventualmente trovare tutte le soluzioni, del seguente sistema a coefficienti in $\mathbb{Q}$:

$$\left\{ \begin{array}{l} x_1 - 3x_2 + x_3 + 2x_4 = 0 \\ 2x_1 - 6x_2 + x_3 + 5x_4 = 1 \\ 3x_1 - 9x_2 + 2x_3 + 10x_4 = 4 \end{array} \right. \tag{2.1}$$

Svolgimento. La matrice completa associata al sistema 2.1 è:

$$A = \begin{pmatrix} 1 & -3 & 1 & 2 & 0 \\ 2 & -6 & 1 & 5 & 1 \\ 3 & -9 & 2 & 10 & 4 \end{pmatrix}$$

Portiamola in forma a scalini per righe:

$$A \to A_1 = \begin{pmatrix} 1 & -3 & 1 & 2 & 0 \\ 0 & 0 & -1 & 1 & 1 \\ 3 & -9 & 2 & 10 & 4 \end{pmatrix}$$

$$A_1 \to A_2 = \begin{pmatrix} 1 & -3 & 1 & 2 & 0 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & -1 & 4 & 4 \end{pmatrix}$$

$$A_2 \to A_3 = \begin{pmatrix} 1 & -3 & 1 & 2 & 0 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 3 & 3 \end{pmatrix}.$$

Il rango della matrice incompleta è tre ed è uguale al rango della matrice completa. Dunque sistema associato ad A_3 è risolubile; osserviamo che ha come unica variabile libera x_2. Troviamo l'espressione di queste soluzioni in funzione di x_2; scriviamo il sistema corrispondente alla matrice A_3 (che sappiamo essere equivalente a 2.1):

$$\begin{cases} x_1 & -3x_2 & +x_3 & +2x_4 & = 0 \\ & & -x_3 & +x_4 & = 1 \\ & & & +3x_4 & = 3 \end{cases} \longrightarrow \begin{cases} x_1 = 3x_2 - 2 \\ x_3 = 0 \\ x_4 = 1 \end{cases}$$

Perciò le soluzioni del sistema (2.1) sono tutti i vettori di $\mathbb{Q}^4$ del tipo $(3x_2 - 2, x_2, 0, 1)$ al variare di x_2 in $\mathbb{Q}$.

Esercizio 3.7. Determinare per quale valori del parametro reale t il sistema lineare (2.2) nelle variabili x, y, z a coefficienti in $\mathbb{R}$ è risolubile e trovarne le soluzioni:

$$\begin{cases} x + y + tz = 1 \\ x + z = 0 \\ x + y + t^3 z = 3 \\ x + y + z = 0 \end{cases} \tag{2.2}$$

Svolgimento La matrice completa associata al sistema è A:

$$A = \begin{pmatrix} 1 & 1 & t & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & t^3 & 3 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

Riduciamola in forma a scalini. Notazione: useremo la notazione di scrivere tra parentesi quadra le righe. Per esempio [2]=[1]-3[2] significherà che sostituiremo al posto della seconda riga, la prima riga meno tre volte la seconda.

$$A \overset{[2]=[4]-[2]}{\longrightarrow} A_1 = \begin{pmatrix} 1 & 1 & t & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & t^3 & 3 \\ 1 & 1 & 1 & 0 \end{pmatrix} \overset{[3]=[3]-[4]}{\longrightarrow} A_2 = \begin{pmatrix} 1 & 1 & t & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & t^3-1 & 3 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

$$A_2 \overset{[4]=[1]-[4]}{\longrightarrow} A_3 = \begin{pmatrix} 1 & 1 & t & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & t^3-1 & 3 \\ 0 & 0 & t-1 & 1 \end{pmatrix}$$

A questo punto osserviamo che $t^3 - 1 = (t-1)(t^2+t+1)$ e che t^2+t+1 è diverso da zero per qualsiasi valore di $t \in \mathbb{R}$. Dunque è la mossa che consiste nel moltiplicare la quarta riga per $-(t^2 + t + 1)$ è lecita e poi, come mossa successiva, possiamo sommare alla quarta riga la terza riga. Il risultato di queste due mosse può essere sintetizzato come $[4] = [3] - (t^2 + t + 1)[4]$:

$$A_3 \overset{[4]=[3]-(t^2+t+1)[4]}{\longrightarrow} A_4 = \begin{pmatrix} 1 & 1 & t & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & t^3-1 & 3 \\ 0 & 0 & 0 & 2-t^2-t \end{pmatrix}$$

Il sistema (2.2) è dunque equivalente al sistema (2.3):

$$(2.3)\qquad \begin{cases} x+y+tz=1 \\ y=0 \\ (t^3-1)z=3 \\ 0=2-t^2-t \end{cases}$$

Per essere risolubile deve essere dunque $2-t^2-t=0$, ovvero:

$$t=\frac{-1\pm\sqrt{1+8}}{2}= \begin{array}{l} \nearrow \;\; 1 \\ \searrow \;\; -2 \end{array}$$

Il sistema (2.2) può avere soluzioni solo per $t=1$ o $t=-2$. Nel caso $t=1$ però, sostituendo nel sistema (2.3), si ha che la terza equazione è $0=3$ e dunque anche per questo valore il sistema non ha soluzioni.
Rimane il caso $t=-2$. Sostituendo nel sistema (2.3) si ottiene:

$$\begin{cases} x-2z=1 \\ y=0 \\ -9z=3 \\ 0=0 \end{cases}$$

che ha una unica soluzione:

$$\begin{cases} x=\frac{1}{3} \\ y=0 \\ z=-\frac{1}{3} \end{cases}$$

Concludendo il sistema (2.2) ammette soluzioni solo nel caso $t=2$. Per questo valore di t la soluzione del sistema è unica.

Esercizio 3.8. Trovare tutte le soluzioni in $\mathbb{Q}^4$ del seguente sistema lineare:

$$(2.4)\qquad \begin{cases} x_1-3x_2+x_3+2x_4=0 \\ 2x_1-6x_2+x_3+5x_4=1 \\ 3x_1-9x_2+2x_3+10x_4=4 \end{cases}$$

Svolgimento La matrice dei coefficienti associata al sistema (2.5) è:

$$B=\begin{pmatrix} 1 & -3 & 1 & 2 & 0 \\ 2 & -6 & 1 & 5 & 1 \\ 3 & -9 & 2 & 10 & 4 \end{pmatrix}$$

Lavoriamo con sostituzioni di riga per trovare una matrice a scalini associata ad un sistema equivalente (ovvero con lo stesso insieme di soluzioni) al sistema 2.5:

$$B \overset{[2]=[2]-2[1]}{\longrightarrow} B_1=\begin{pmatrix} 1 & -3 & 1 & 2 & 0 \\ 0 & 0 & -1 & 1 & 1 \\ 3 & -9 & 2 & 10 & 4 \end{pmatrix} \overset{[3]=[3]-3[1]}{\longrightarrow} B_2=\begin{pmatrix} 1 & -3 & 1 & 2 & 0 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & -1 & 4 & 4 \end{pmatrix}$$

$$B_2 \overset{[3]=[3]-[2]}{\longrightarrow} B_3=\begin{pmatrix} 1 & -3 & 1 & 2 & 0 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 3 & 3 \end{pmatrix}$$

La matrice B_3 è a scalini e il sistema corrispondente ad essa, equivalente al sistema (2.4), è il seguente:

$$(2.5)\qquad \begin{cases} x_1-3x_2+x_3+2x_4=0 \\ -x_3+x_4=1 \\ 3x_4=3 \end{cases}$$

Il sistema (2.5) ha una variabile libera (che è x_2). Dunque al variare del valore di x_2 in $\mathbb{Q}$, si ha che le soluzioni del sistema (2.5) sono gli elementi di $\mathbb{Q}^4$ del tipo: $(3x_2 - 2, x_2, 0, 1)$.

Esercizio 3.9. Determinare un polinomio $g(x) \in \mathbb{Q}[x]$ tale che:

$$g(1) = 10 \quad g(-1) = 2 \quad g(-2) = 1$$

Svolgimento Scegliendo un grado per il polinomio $g(x)$ e imponendo le condizioni richieste, l'esercizio si traduce nel risolvere un sistema lineare per determinare i valori dei coefficienti di $g(x)$.

- Un polinomio di grado 0 è una costante e dunque non c'è speranza di trovare $g(x)$ di grado 0 che, valutato su tre valori diversi di x, assuma tre valori distinti.
- Proviamo a vedere se esiste un polinomio di primo grado con i valori richiesti. Poniamo dunque $g(x) = ax + b$ con $a, b \in \mathbb{Q}$. Le condizioni richieste equivalgono al seguente sistema in $\mathbb{Q}^2$:

$$\begin{cases} a + b = 10 \\ -a + b = 2 \\ -2a + b = 1 \end{cases}$$

La matrice dei coefficienti associata al sistema è:

$$A = \begin{pmatrix} 1 & 1 & 10 \\ -1 & 1 & 2 \\ -2 & 1 & 1 \end{pmatrix}$$

Portiamola a scalini:

$$A \underset{[3]=2[1]+[3]}{\overset{[2]=[1]+[2]}{\longrightarrow}} A_1 = \begin{pmatrix} 1 & 1 & 10 \\ 0 & 2 & 12 \\ 0 & 3 & 21 \end{pmatrix} \overset{[3]=[3]-\frac{3}{2}[2]}{\longrightarrow} A_2 = \begin{pmatrix} 1 & 1 & 10 \\ 0 & 2 & 12 \\ 0 & 0 & 3 \end{pmatrix}$$

Il sistema dunque risulta non risolubile in quanto equivalente al seguente sistema:

$$\begin{cases} a + b = 10 \\ 2b = 12 \\ 0 = 3 \end{cases}$$

Questo significa che non esistono polinomi $g(x)$ di grado 1 con la proprietà richiesta di assumere i valori $10, 2$ e 1 rispettivamente in $1, -1$ e -2.
- Proviamo con $g(x)$ di secondo grado. Poniamo dunque $g(x) = ax^2 + bx + c$ e imponiamo le condizioni richieste ottenendo il sistema in $\mathbb{Q}^3$ seguente:

$$\begin{cases} a + b + c = 10 \\ a - b + c = 2 \\ 4a - 2b + c = 1 \end{cases}$$

La matrice dei coefficienti associata al sistema è:

$$A = \begin{pmatrix} 1 & 1 & 1 & 10 \\ 1 & -1 & 1 & 2 \\ 4 & -2 & 1 & 1 \end{pmatrix}$$

Portiamola a scalini:

$$A \underset{[3]=4[1]-[3]}{\overset{[2]=[1]-[2]}{\longrightarrow}} A_1 = \begin{pmatrix} 1 & 1 & 1 & 10 \\ 0 & 2 & 0 & 8 \\ 0 & 6 & 3 & 39 \end{pmatrix} \overset{[3]=[3]-3[2]}{\longrightarrow} A_2 = \begin{pmatrix} 1 & 1 & 1 & 10 \\ 0 & 2 & 0 & 8 \\ 0 & 0 & 3 & 15 \end{pmatrix}$$

Questa matrice ha rango massimo e uguale a 3 (come il numero delle variabili) e dunque il sistema corrispondente ha una unica soluzione:

$$\begin{cases} a+b+c=10 \\ 2b=8 \\ 3c=15 \end{cases} \longrightarrow \begin{cases} a=1 \\ b=4 \\ c=5 \end{cases}$$

L'unico polinomio di secondo grado con la proprietà richiesta è dunque $g(x)=x^2+4x+5$.

Generalizzando quanto visto finora (e pensando il tutto in $\mathbb{R}$ invece che in $\mathbb{Q}$) si potrebbero dimostrare (o comunque ripensare in termini di algebra lineare) alcuni risultati di geometria analitica: la condizione richiesta equivale al fatto che il grafico della funzione $g(x)$ passi per i tre punti del piano $(1,10),(-1,2),(-2,1)$.
Ora il grafico del polinomio di primo grado $ax+b$ corrisponde ad una generica retta del piano, dunque l'unica speranza che passi per tre punti è che questi siano allineati.
Il grafico del polinomio di secondo grado ax^2+bx+c corrisponde ad una generica parabola del piano. Abbiamo dimostrato che esiste una e una sola parabola del piano passante per i tre punti richiesti. Generalizzando si potrebbe dimostrare che, scelti tre punti non allineati, esiste una e una sola parabola del piano passante per i tre punti.

Esercizio 3.10. Trovare tutte le soluzioni del sistema a coefficienti in $\mathbb{R}$:

$$\begin{cases} 2x+2y+z+2t &= 0 \\ 2y+3z-t &= -5 \\ +y-z-t &= 0 \end{cases}$$

Esercizio 3.11. Trovare tutte le soluzioni del sistema a coefficienti in $\mathbb{R}$:

$$\begin{cases} 2x+y+z+t+w &= 1 \\ 2y+3z-t+2w &= 0 \\ 2x+y-z-t+w &= 0 \\ x+y+3z+t+w &= 0 \end{cases}$$

Esercizio 3.12. Trovare tutte le soluzioni del sistema a coefficienti in $\mathbb{R}$:

$$\begin{cases} x+2y+z+2t+w &= 0 \\ 2y+3z-t+2w &= 0 \\ y-z-t+w &= 0 \\ 4x+y+3z+t+w &= 0 \end{cases}$$

Esercizio 3.13. Consideriamo il sistema lineare a coefficienti in $\mathbb{R}$:

$$\begin{cases} 2x+y+mz &= 1 \\ 2y+mz &= 0 \\ x+my+2z &= 1 \end{cases}$$

Stabilire per quali valori del parametro reale m il seguente sistema ammette soluzioni e, per tali valori, calcolare le soluzioni.

Esercizio 3.14. Consideriamo il sistema lineare a coefficienti in $\mathbb{R}$:

$$\begin{cases} 2x+2y+(k-3)z &= -2 \\ x+(k-2)y-(k+1)z &= -3 \\ x+2y+kz &= 1 \end{cases}$$

Stabilire per quali valori del parametro reale k il seguente sistema ammette soluzioni e, per tali valori, calcolare le soluzioni.

Esercizio 3.15. Si consideri l'applicazione lineare $A_t : \mathbb{R}^4 \to \mathbb{R}^3$ a cui, rispetto alle basi standard, è associata la seguente matrice:

$$[A_t] = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 1 & 2 & 0 \\ t & t^3 & 1+t & 1 \end{pmatrix}.$$

Trovare, se esistono, valori del parametro t per i quali si ha che $dim\ Ker\ A_t = 2$ ed esibire, in tal caso, una base di $Ker\ A_t$.

Svolgimento. Dal Teorema 2.21 segue che il nucleo di A_t ha dimensione 2 se e solo se la dimensione dell'immagine di A_t è uguale a 2, in altre parole se e solo se il rango di A_t è 2. Come sappiamo, il rango si può calcolare riducendo la matrice $[A_t]$ in forma a scalini. Lo si può fare con operazioni elementari di riga, oppure con operazioni elementari di colonna, oppure, se ci interessa esclusivamente il rango, si possono usare sequenze "miste" di operazioni elementari per riga e per colonna.

In questo caso è vero che in prima battuta ci interessa il rango, ma l'esercizio chiede anche di esibire una base del nucleo di A_t per certi valori di t, dunque di risolvere un sistema lineare. In previsione di questo, ci conviene utilizzare le mosse di riga, le uniche che non cambiano le soluzioni del sistema lineare.

Facciamo una rapida analisi della matrice in questione: le due prime righe sono sicuramente linearmente indipendenti, perciò il numero di scalini che otterremo è almeno 2 e al massimo sarà 3 (ci sono solo tre righe).

Portiamo A_t in forma a scalini

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 1 & 2 & 0 \\ t & t^3 & 1+t & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & 0 & -1 \\ t & t^3 & 1+t & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & t^3-t & 1-t & 1-t \end{pmatrix}$$

A questo punto affinché la matrice abbia rango 2 è necessario che l'ultima riga non abbia coefficienti non nulli prima della quarta colonna, ovvero che:

$$t^3 - t = 1 - t = 0$$

e ciò accade solo per $t = 1$. Si ha quindi che A_1 è l'unica applicazione del tipo considerato che ha il nucleo di dimensione 2. Per individuare $Ker\ A_1$ dobbiamo risolvere il il sistema omogeneo:

$$[A_1] \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Tale sistema, come sappiamo, equivale a quello con matrice a scalini per righe:

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Quindi dobbiamo risolvere il sistema trovando le variabili x e t in funzione delle variabili libere y e z:

$$\begin{cases} x + y + 2z + t = 0 \\ -t = 0 \\ 0 = 0 \end{cases}$$

Troviamo $t = 0$ e $x = -y - 2z$, quindi un generico vettore di $Ker\ A_1$ è della forma:

$$\begin{pmatrix} -y-2z \\ y \\ z \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \cdot y + \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} \cdot z$$

Si osserva immediatamente che i due vettori:

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

sono un insieme di generatori linearmente indipendenti (quindi una base) di $Ker\ A_1$.

Esercizio 3.16. Sia $g : \mathbb{Q}^3 \to \mathbb{Q}^2$ definita da:

$$g(x, y, z) = (2x + y + 2z, x + y + 3z).$$

Trovare una base di $Imm\ g$ e di $Ker\ g$.

Svolgimento. La matrice associata a g nelle basi canoniche di $\mathbb{Q}^3$ e $\mathbb{Q}^2$ è:

$$[g] = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 1 & 3 \end{pmatrix}$$

Portiamola in forma a scalini con operazioni di riga:

$$[g] \longrightarrow \begin{pmatrix} 2 & 1 & 2 \\ 0 & 1 & 4 \end{pmatrix}$$

Gli elementi di $Ker\ g$ sono le soluzioni del sistema:

$$\begin{cases} 2x + y + 2z = 0 \\ y + 4z = 0 \end{cases}$$

Lo risolviamo in funzione della variabile libera z, quindi: $y = -4z$ e $x = z$. Perciò un generico elemento di $Ker\ g$ è della forma:

$$\begin{pmatrix} z \\ -4z \\ z \end{pmatrix} = z \begin{pmatrix} 1 \\ -4 \\ 1 \end{pmatrix}$$

Dunque $Ker\ g$ ha dimensione 1 e una sua base è data dal vettore:

$$\begin{pmatrix} 1 \\ -4 \\ 1 \end{pmatrix}.$$

Per il Teorema 2.21 sappiamo a questo punto che $Imm\ g$ ha dimensione 2. Per esibire una base di $Imm\ g$ basta allora scegliere due colonne linearmente indipendenti nella matrice

$$[g] = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 1 & 3 \end{pmatrix}$$

Le prime due colonne, come si verifica immediatamente, sono linearmente indipendenti, dunque i vettori

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

costituiscono una base di $Imm\ g$.

3. Come funziona Google

NOTA: Questo paragrafo è ricavato da una delle schede di una serie intitolata "Modelli e Realtà", a cura di A. Abbondandolo e G. Gaiffi.

Le schede sono in forma di dialogo fra due personaggi, Ilaria e Orazio.

Orazio ha appena messo in rete la sua prima pagina web, che raccoglie gli appunti su alcuni dei problemi matematici che ha affrontato insieme a Ilaria. Però non è soddisfatto.

O - Niente da fare: Google non ci trova. La nostra pagina è invisibile a chi non conosca già l'indirizzo esatto.

I - Abbi un po' di pazienza. Google ci mette un po' ad accorgersi della presenza di nuove pagine. È vero che dispone di un intero esercito di computer, che giorno e notte navigano per la rete e indicizzano tutti i documenti che trovano, ma si tratta pur sempre di parecchi miliardi di documenti. Riproverei a fare una nuova ricerca tra qualche settimana.

O - E basterà cercare "Ilaria e Orazio" per far saltare fuori la nostra pagina?

I - Su questo non sarei troppo ottimista. In rete ci saranno migliaia di documenti che contengono i nostri nomi. Dipenderà da quanta rilevanza Google attribuirà alla nostra pagina: se sarà considerata rilevante, apparirà tra le prime risposte, altrimenti finirà in fondo ad una lista lunghissima, che nessuno avrà voglia di leggere.

O - E come fa Google a decidere quanto rilevante è la nostra pagina web? Si vanno a leggere tutto quello che scriviamo?

I - No, lo fa in modo automatico, basandosi sul numero e sulla rilevanza delle altre pagine che contengono un link alla nostra. Utilizza un algoritmo che si chiama "PageRank". Vuoi che ti spieghi come funziona?

O - Certo.

I - Immaginati un navigatore indeciso, che non sa bene quello che cerca. Parte da una pagina a caso, guarda tutti i link che questa contiene, ne sceglie uno a caso, lo segue e fa lo stesso con la nuova pagina raggiunta. E avanti così, all'infinito. Supponiamo di aver numerato tutte le pagine della rete con numeri da 1 a N e chiamiamo r_j la probabilità che ad un dato istante il navigatore indeciso si trovi sulla j-esima pagina. Il valore r_j può essere interpretato come una misura della rilevanza della pagina j-esima.

O - Sí, capisco perché: nel suo girovagare, il navigatore indeciso si troverà più spesso sulle pagine molto popolari, quelle a cui puntano numerosi link, mentre visiterà molto raramente le pagine poco segnalate. Ma la quantità r_j non dovrebbe essere funzione del tempo?

I - All'inizio della navigazione sì, ma è ragionevole supporre che dopo un po' queste probabilità si avvicinino a dei valori indipendenti dal tempo. Possiamo anche scrivere una relazione tra le varie r_j. Chiamiamo ℓ_j il numero dei link presenti nella j-esima pagina e A_i il sottoinsieme di $\{1, \ldots, N\}$ che corrisponde alle pagine che puntano verso la i-esima. Affinché ad un dato istante il navigatore si trovi sulla i-esima pagina, è necessario che all'istante precedente si trovasse in una delle pagine dell'insieme A_i. Se j è una delle pagine di questo insieme, la probabilità che all'istante precedente il navigatore fosse su j vale r_j e, in questi caso, la probabilità che segua il link che lo porta ad i è $1/\ell_j$. Quindi la probabilità r_i è data dalla somma su tutti gli elementi j di A_i di r_j moltiplicato per $1/\ell_j$. In una formula:

$$r_i = \sum_{j \in A_i} \frac{r_j}{\ell_j}, \qquad \forall i \in \{1, \dots, N\}. \tag{3.1}$$

O - Non sono sicuro di aver capito. Sai che la probabilità non è il mio forte...

I - Puoi anche prescindere dal modello probabilistico del navigatore indeciso e considerare la formula che ho scritto come una definizione della rilevanza di una pagina: la rilevanza r_i della pagina i-esima è pari alla somma dei quozienti tra la rilevanza r_j di ciascuna delle pagine che ad essa puntano per il numero dei link ℓ_j di tale pagina. È come se ciascuna pagina trasmettesse una porzione della propria rilevanza alle pagine verso le quali punta. Quindi la rilevanza di una pagina non dipende solamente dal numero dei link che ad essa puntano, ma anche e soprattutto dalla rilevanza delle pagine che contengono questi link. È molto ragionevole che essere citati da una pagina molto visitata sia considerato più rilevante che essere citati da dieci pagine che non legge nessuno.

O - Ma la formula che hai scritto non è una definizione! Per sapere quanto vale r_i hai bisogno di conoscere il valore delle altre r_j: è un gatto che si morde la coda.

I - Hai ragione, ma più precisamente la (3.1) è un sistema di N equazioni lineari (una per ciascuna i) in N incognite (le $r_1, \dots, r_N$).

O - È vero! Anzi, si tratta di un sistema lineare omogeneo. Come tale ha sempre la soluzione banale $r_1 = r_2 = \dots = r_N = 0$. Se le righe della matrice quadrata associata sono linearmente indipendenti, ossia se la matrice ha rango uguale a N, questa è l'unica soluzione. Altrimenti, se le righe non sono linearmente indipendenti, ve ne sono infinite. Infatti in questo caso riducendo la matrice a scalini per righe si trovano delle righe uguali a 0 e degli scalini "lunghi".

I - Proprio così. La soluzione banale ovviamente non ci interessa. Vorremmo che ci fosse una soluzione non banale che, fedeli al modello probabilistico, possiamo normalizzare in modo che la somma di tutte le r_i valga 1.

O - Ma non è detto che questa soluzione ci sia! Considera questo esempio semplicissimo: una rete composta da due sole pagine, in cui la prima contiene un link alla seconda, che invece non possiede link:

$$\boxed{1} \longrightarrow \boxed{2}$$

In questo caso, il tuo sistema lineare è

$$\begin{cases} r_1 = 0 \\ r_2 = r_1, \end{cases}$$

che ha soltanto la soluzione $r_1 = r_2 = 0$.

I - Hai di nuovo ragione, ho dimenticato di introdurre una piccola modifica. Nel tuo esempio il problema è causato dalla seconda pagina, che non possiede link. Ripensa al navigatore indeciso: come prosegue se va a cadere in una pagina senza link? Dobbiamo aggiungere una regola che dica cosa fare in questo caso. La più semplice è questa: da una pagina senza link si va in una qualunque delle N pagine della rete con probabilità $1/N$. Se V è il sottoinsieme di $\{1, \ldots, N\}$ che corrisponde alle pagine prive di link, la formula corretta è la seguente:

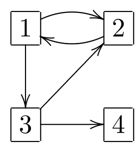
$$r_i = \sum_{j \in A_i} \frac{r_j}{\ell_j} + \frac{1}{N} \sum_{j \in V} r_j. \tag{3.2}$$

È come se si decidesse d'ufficio che le pagine senza link puntano a tutte le pagine della rete: una tale pagina trasmette una piccolissima parte della sua rilevanza ($1/N$ è un numero piccolissimo) a tutte le altre (inclusa sé stessa). Dopo questa correzione, il tuo esempio produce il sistema

$$\begin{cases} r_1 = r_2/2 \\ r_2 = r_1 + r_2/2, \end{cases}$$

che ha per soluzione generale $r_1 = s$, $r_2 = 2s$, al variare di s tra tutti i numeri reali. Imponendo la condizione di normalizzazione $r_1 + r_2 = 1$, si trova $r_1 = 1/3$ e $r_2 = 2/3$.

O - Sì, mi sembra ragionevole che la seconda pagina sia il doppio più rilevante della prima. Fammi provare con un altro esempio:



Questa volta il sistema è:

$$\begin{cases} r_1 = r_2 + r_4/4 \\ r_2 = r_1/2 + r_3/2 + r_4/4 \\ r_3 = r_1/2 + r_4/4 \\ r_4 = r_3/2 + r_4/4. \end{cases}$$

Provo a risolverlo... Ecco, la soluzione generale è

$$r_1 = 10s, \quad r_2 = 9s, \quad r_3 = 6s, \quad r_4 = 4s,$$

e normalizzando trovo

$$r_1 = 10/29, \quad r_2 = 9/29, \quad r_3 = 6/29, \quad r_4 = 4/29.$$

La quarta pagina è ovviamente la meno rilevante. Le prime due sono le più rilevanti, ma la prima lo è leggermente di più poiché beneficia per intero della rilevanza della seconda, la quale a sua volta riceve solamente metà della rilevanza della prima. In questo esempio ho trovato una soluzione non banale, ma sei sicura che il sistema (3.2) ne abbia sempre?

I - Sì, e te lo posso dimostrare facilmente. Cerchiamo di riscrivere il sistema (3.2) utilizzando il linguaggio dei vettori e delle matrici. Se chiamiamo r il vettore $(r_1, \ldots, r_N)$, possiamo riscrivere (3.2) come

$$r = Wr, \tag{3.3}$$

dove W è la matrice $N \times N$ in cui alla riga i-esima e alla colonna j-esima compare $1/\ell_j$ se la pagina j contiene un link verso la pagina i, $1/N$ se la pagina j non possiede link, e 0 se la pagina j contiene link che però non puntano verso la pagina i:

$$W_{i,j} = \begin{cases} 1/\ell_j & \text{se } j \in A_i, \\ 1/N & \text{se } j \in V, \\ 0 & \text{altrimenti.} \end{cases}$$

Nota che i coefficienti di questa matrice sono tutti numeri non negativi e che la somma dei coefficienti di ciascuna colonna vale 1. Le matrici con questa proprietà si chiamano "stocastiche".

O - Fammi capire perchè vale l'ultima proprietà che hai detto... Guardare una colonna significa fissare j e far variare i da 1 a N. Se la pagina j possiede link, allora sulla colonna j-esima trovo $1/\ell_j$ per ognuna delle ℓ_j pagine i a cui punta la pagina j, e tutti gli altri coefficienti sono zero: totale 1. Se invece la pagina j non possiede link, tutti i coefficienti della colonna j-esima valgono $1/N$: di nuovo, la somma è 1.

I - Esatto, mentre invece se avessi scritto allo stesso modo il sistema (3.1), avrei avuto delle colonne con coefficienti tutti uguali a 0 e la matrice W non sarebbe stata stocastica. Come hai giustamente osservato, l'equazione (3.3) rappresenta un sistema lineare omogeneo. Posso riscriverlo come

$$(I - W)r = 0,$$

dove I è la matrice identità, ossia la matrice $N \times N$ che ha 1 sulla diagonale e 0 fuori da essa. Scritto in questa forma, vediamo che (3.3) ha soluzioni non banali se e solamente se l'applicazione lineare rappresentata da $I - W$ ha nucleo non banale, ovvero se e solo se $I - W$ ha rango strettamente minore di N. Affermo che il fatto che W sia stocastica implica che il rango di $I - W$ sia $< N$.

O - Forse ci sono! La somma dei coefficienti di ciascuna colonna di $I-W$ vale 0 (c'è esattamente un 1 che viene dalla matrice identità, a cui devo togliere la somma dei coefficienti della corrispondente colonna della matrice stocastica W). Ma questo è come dire che la somma degli N vettori riga che compongono la matrice $I - W$ è il vettore nullo. In particolare, le righe di $I - W$ sono linearmente dipendenti e sappiamo che questo implica che il rango di $I - W$ è $< N$.

I - Proprio così. Quindi l'equazione (3.3), o equivalentemente il sistema (3.2), ha sempre soluzioni non banali.

O - Però se vogliamo che la (3.3) ci dia una buona definizione della rilevanza di tutte le pagine della rete, abbiamo anche bisogno che la soluzione non banale sia unica, una volta normalizzata in modo che la somma degli r_i valga 1...

I - Hai ragione. In effetti occorre apportare un'ultima modifica per avere l'unicità. Nota infatti che se la rete consiste di due sotto-reti non collegate tra loro da alcun link, non possiamo aspettarci che la soluzione sia unica.

O - È vero. In questo caso posso costruire due soluzioni linearmente indipendenti dell'equazione (3.3) in questo modo: risolvo l'equazione relativa ad una delle due sotto-reti e la estendo ad una soluzione globale ponendo la rilevanza di ciascuna pagina dell'altra sotto-rete uguale a zero.

I - Infatti l'idea della modifica che garantisce l'unicità è proprio quella di eliminare eventuali sotto-reti isolate. Si fa così: fissiamo un parametro ϵ compreso tra 0 e 1 e definiamo

$$G = (1 - \epsilon)W + \epsilon Q, \tag{3.4}$$

dove Q è la matrice $N \times N$ che ha tutti i coefficienti uguali a $1/N$. La matrice G è ancora stocastica, ma adesso tutti i suoi coefficienti sono positivi. Si può dimostrare che questo garantisce che l'equazione

$$r = Gr \tag{3.5}$$

possiede un'unica soluzione r tale che $\sum_i r_i = 1$. L'interpretazione probabilistica dietro a questa modifica è la seguente: a ciascun passo, il navigatore indeciso segue le regole che abbiamo descritto prima - e che sono riassunte dalla matrice W - con probabilità $1 - \epsilon$, mentre con probabilità ϵ si sposta su una pagina a caso.

O - E immagino che ϵ debba essere scelto piccolo, in modo da non modificare troppo la struttura effettiva delle rete.

I - Esatto, ma non troppo piccolo, perché il termine ϵQ risulta anche utile per calcolare effettivamente la soluzione r dell'equazione (3.5).

O - Già. Avevamo tralasciato completamente questo aspetto: i computer di Google devono risolvere un sistema lineare composto da qualche miliardo di equazioni. Come ci riescono?

I - In realtà si accontentano di una soluzione approssimata. Partono da un qualsiasi vettore con coefficienti positivi aventi somma 1, ad esempio $r(0) = (1/N, \ldots, 1/N)$. Gli applicano la matrice G, ottenendo il vettore $r(1) = Gr(0)$. Poi iterano il procedimento, determinando i vettori $r(2) = Gr(1)$, $r(3) = Gr(2)$, eccetera. Si può dimostrare che il fatto che la matrice stocastica G abbia tutti i coefficienti positivi implica che la successione di vettori $r(n)$ si avvicina sempre di più alla soluzione normalizzata dell'equazione (3.5). La convergenza è tanto più rapida quanto più grande è ϵ. Penso che Google scelga $\epsilon = 0,15$ ed ottenga una soluzione approssimata soddisfacente fermandosi dopo qualche decina di iterazioni. La rilevanza delle pagine della rete viene aggiornata a scadenze regolari ed il calcolo dura qualche giorno.

O - Fammi ricapitolare. I computer di Google esplorano in continuazione la rete. Le parole di ciascun documento vengono memorizzate in un database, mentre le informazioni sulla struttura della rete, ossia quali pagine contengano link a quali altre, vengono memorizzate nella matrice W. Il database serve per decidere quali documenti riportare in risposta alle varie ricerche. La matrice W è utiizzata invece per determinare una classifica assoluta di tutte le pagine, che vengono ordinate per rilevanza determinando una soluzione approssimata dell'equazione (3.5), dove la matrice G è ottenuta dalla W grazie alla formula (3.4). La rilevanza determina a sua volta l'ordine in cui Google fornisce le risposte ad una determinata ricerca.

I - Esatto. Quindi se vogliamo fare in modo che la nostra pagina abbia una rilevanza alta, dobbiamo convincere i gestori di qualche sito già rilevante - ad esempio qualche sito istituzionale sull'istruzione, o sulla divulgazione matematica - ad includerla tra i loro link.

CAPITOLO 4

La formula di Grassmann

1. La formula di Grassmann per le intersezioni e le somme di sottospazi.

Dati due sottospazi vettoriali A e B in $\mathbb{R}^3$ di dimensione 2 (dunque due piani contenenti O), di che dimensione può essere la loro intersezione?

Possono intersecarsi lungo una retta: in tal caso si nota che il sottospazio generato dai vettori di $A \cup B$, ossia $A + B$ (vedi Definizione 1.28), è tutto $\mathbb{R}^3$.

Oppure vale $A = B$: allora la loro intersezione è uguale ad A (e a B) e ha dimensione 2, e anche il sottospazio $A + B$ coincide con A.

In entrambi i casi, la somma delle dimensioni di $A \cap B$ e di $A + B$ è sempre uguale a 4 (che a sua volta è uguale a $dim\ A + dim\ B$).

E se in $\mathbb{R}^4$ consideriamo un piano C e un sottospazio D di dimensione 3?[1] 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 (esercizio!) che

$$dim\ C \cap D + dim\ (C + D) = 5 = dim\ C + dim\ D.$$

Sembra dunque che ci sia una relazione fra le dimensioni in gioco: se due sottospazi A e B di uno spazio vettoriale V si intersecano "tanto", allora generano "poco". Più precisamente:

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

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

Dimostreremo questa formula, detta *formula di Grassmann*, come applicazione del Teorema 2.21.

Premettiamo una osservazione sul prodotto cartesiano di due spazi vettoriali. Dati due spazi vettoriali V e W sul campo $\mathbb{K}$, 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:

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

[1] In generale, se V è uno spazio vettoriale di dimensione n e H è un sottospazio di dimensione $n - 1$ si dice che H è un *iperpiano* di V.

Si verifica immediatamente 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[2] di $V \times W$, che dunque ha dimensione $n + m = (dim\ V) + (dim\ W)$.

Teorema 4.1. *Dati due sottospazi A, B di uno spazio vettoriale V sul campo $\mathbb{K}$, vale*

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

Dimostrazione. Consideriamo l'applicazione

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

definita da $\Phi((a, b)) = a - b$. Si verifica (facile esercizio) che Φ è lineare. Dimostreremo il teorema studiando il nucleo e l'immagine di Φ e applicando il Teorema 2.21.

Cosa sappiamo dire del nucleo di Φ ? Per definizione

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

dunque

$$Ker\ \Phi = \{(a, b) \in A \times B \mid a = b\}$$

che equivale a scrivere:

$$Ker\ \Phi = \{(z, z) \in A \times B \mid z \in A \cap B\}.$$

Si nota subito che la applicazione lineare $\theta : A \cap B \to Ker\ \Phi$ data da $z \to (z, z)$ è iniettiva e surgettiva, dunque è un isomorfismo (vedi la Definizione 2.22). Allora il suo dominio e il suo codominio hanno la stessa dimensione, ovvero

$$dim\ Ker\ \Phi = dim\ A \cap B$$

Cosa sappiamo dire dell'immagine di Φ? Per definizione

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

Visto che B, come ogni spazio vettoriale, se contiene un elemento b contiene anche il suo opposto $-b$, possiamo scrivere la seguente uguaglianza fra insiemi:

$$\{a - b \mid a \in A,\ b \in B\} = \{a + b \in V \mid a \in A,\ b \in B\} = A + B.$$

Dunque

$$Imm\ \Phi = A + B$$

Per il Teorema 2.21 applicato a Φ sappiamo che:

$$dim\ (A \times B) = dim\ Ker\ \Phi + dim\ Imm\ \Phi.$$

Questa formula, viste le osservazioni fatte fin qui, si traduce come:

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

□

[2]Una precisazione: lo O che compare nelle coppie (v_i, O) è lo O dello spazio W, mentre lo O che compare in (O, w_j) è lo O di V. Qui e altrove nel testo abbiamo scelto, per semplicità, di non aggiungere indici al vettore O.

Esercizio 4.2. Dare una dimostrazione della formula di Grassmann nel seguente modo: fissare una base $z_1, z_2, \ldots, z_k$ di $A \cap B$ e usare il teorema di completamento (Teorema 2.19) per completarla ad una base di A aggiungendo certi vettori $v_1, v_2, \ldots, v_r$. Poi usare di nuovo il teorema di completamento per completare la base di $A \cap B$ ad una base di B aggiungendo certi vettori $w_1, w_2, \ldots, w_s$. A questo punto dimostrare che i vettori $z_1, z_2, \ldots, z_k, v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_s$ sono una base di $A + B$.

Esercizio 4.3. Dire se è possibile trovare in $\mathbb{R}^4$ tre sottospazi vettoriali A, B, C di dimensione 2 tali che $A \cap B = \{O\}$, $A \cap C = \{O\}$ e $B \cap C = \{O\}$.

Esercizio 4.4. 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:

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

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

2. Un metodo per calcolare l'intersezione di due sottospazi vettoriali

Fin qui abbiamo visto essenzialmente due modi di presentare un sottospazio vettoriale di uno spazio vettoriale V: come span di certi vettori ($< v_1, ..., v_r >$) oppure come nucleo di una applicazione lineare (ovvero come insieme delle soluzioni di un sistema lineare omogeneo).

Consideriamo due sottospazi, U e W, di V. Se entrambi sono presentati come insieme delle soluzioni di un sistema è facile calcolare $U \cap W$: basta calcolare le soluzioni del sistema 'doppio', ottenuto considerando tutte le equazioni dei due sistemi.

Per esempio se U e W in $\mathbb{R}^4$ sono dati rispettivamente dalle soluzioni dei sistemi S_U:

$$\begin{cases} 3x + 2y + 4w & = 0 \\ 2x + y + z + w & = 0 \end{cases}$$

e S_W:

$$\begin{cases} x + 2y + z + w & = 0 \\ x + z + w & = 0 \end{cases}$$

allora $U \cap W$ è dato dalle soluzioni del sistema:

$$\begin{cases} 3x + 2y + 4w & = 0 \\ 2x + y + z + w & = 0 \\ x + 2y + z + w & = 0 \\ x + z + w & = 0 \end{cases}$$

Esercizio 4.5. Calcolare $U \cap W$ nell'esempio proposto.

Ma come possiamo calcolare $U \cap W$ se i due sottospazi sono presentati come span di certi vettori? Consideriamo per esempio U e W in $\mathbb{R}^5$ definiti così:

$$U = < \begin{pmatrix} 1 \\ 2 \\ 3 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 7 \\ 2 \\ -1 \end{pmatrix} >$$

$$W = < \begin{pmatrix} 1 \\ 2 \\ 0 \\ -2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -3 \\ -6 \\ 1 \end{pmatrix} >$$

Un metodo per calcolare $U \cap W$ è quello di esprimere U e W come soluzioni di un sistema lineare. Mostriamo come si può fare, cominciando da U.

Per prima cosa scriviamo la seguente matrice:

$$\begin{pmatrix} 1 & 2 & x_1 \\ 2 & 4 & x_2 \\ 3 & 7 & x_3 \\ -1 & 2 & x_4 \\ 2 & -1 & x_5 \end{pmatrix}$$

Le prime due colonne sono i vettori della base di U, la terza colonna è data dal vettore $v = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$, dove x_1, x_2, x_3, x_4, x_5 sono dei numeri reali. Cosa possiamo dire sul rango di questa matrice? Se il vettore v appartiene ad U significa che è combinazione lineare delle prime due colonne, dunque il rango della matrice è 2. Se il vettore non appartiene ad U allora il rango della matrice è 3, perchè la matrice è formata da tre colonne linearmente indipendenti.

Dunque il calcolo del rango della matrice può essere utilizzato per decidere se v appartiene o no ad U. Possiamo allora ridurre la matrice a scalini per righe, ottenendo:

$$\begin{pmatrix} 1 & 2 & x_1 \\ 0 & 1 & x_3 - 3x_1 \\ 0 & 0 & 2x_1 - x_2 \\ 0 & 0 & 13x_1 - 4x_3 + x_4 \\ 0 & 0 & -17x_1 + 5x_3 + x_5 \end{pmatrix}$$

Tale matrice ha lo stesso rango della matrice iniziale, e, studiando i suoi scalini, notiamo che ha rango due se e solo se i coefficienti x_1, x_2, x_3, x_4, x_5 soddisfano le equazioni

$$\begin{cases} 2x_1 - x_2 & = 0 \\ 13x_1 - 4x_3 + x_4 & = 0 \\ -17x_1 + 5x_3 + x_5 & = 0 \end{cases} \tag{2.1}$$

Dunque il vettore $v = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$ appartiene ad U se e solo se le sue coordinate soddisfano il sistema (2.1). Questo equivale a dire che le soluzioni del sistema coincidono con gli elementi di U. Abbiamo raggiunto il nostro scopo, ossia abbiamo presentato il sottospazio U come insieme di soluzioni di un sistema lineare.

Osservazione 4.6. Visto che U ha dimensione 2, un sistema le cui soluzioni coincidono con l'insieme U deve avere almeno 3 equazioni (infatti la dimensione dello spazio delle soluzioni di un sistema omogeneo, come sappiamo, è uguale al numero di variabili del sistema meno il rango della matrice; nel nostro caso le variabili sono cinque e dunque il rango deve essere tre, ovvero devono esserci almeno tre equazioni non nulle nel sistema). Questo vuol dire che nel nostro esempio abbiamo ottenuto una descrizione di U *col numero minimo possibile di equazioni* .

Esercizio 4.7. Dimostrare che l'osservazione precedente vale in generale: quando si parte da un sottospazio U di $\mathbb{K}^n$ di cui conosciamo una base $v_1, ...v_r$, il metodo descritto sopra permette di presentare U come lo spazio delle soluzioni di un sistema di $n - r$ equazioni, il numero minimo possibile di equazioni.

Per finire il calcolo di $U \cap W$ possiamo allo stesso modo ottenere un sistema le cui soluzioni sono tutti e soli i vettori di W:

$$\begin{cases} x_2 + x_3 + 2x_5 & = 0 \\ -10x_1 + 9x_2 - 5x_3 + 4x_4 & = 0 \end{cases}$$

A questo punto possiamo ottenere $U \cap W$ risolvendo il sistema:

$$\begin{cases} 2x_1 - x_2 & = 0 \\ 13x_1 - 4x_3 + 1x_4 & = 0 \\ -17x_1 + 5x_3 + 1x_5 & = 0 \\ x_2 + x_3 + 2x_5 & = 0 \\ -10x_1 + 9x_2 - 5x_3 + 4x_4 & = 0 \end{cases}$$

Esercizio 4.8. Verificare nel dettaglio tutti i calcoli di questo paragrafo e concludere l'esercizio, indicando una base di $U \cap V$. La dimensione risulta uguale a quella prevista dalla formula di Grassmann (4.1)?

3. Somma diretta di sottospazi

Si dice che due sottospazi U e W di uno spazio vettoriale V formano una *somma diretta* se vale che $U \cap W = \{O\}$. In questo caso, come sappiamo dalla formula di Grassmann, la dimensione di $U + W$ è 'la massima possibile', ovvero è uguale a $dim\ U + dim\ W$. Vale anche il viceversa, ossia due sottospazi sono in somma diretta se e solo se $dim\ (U + W) = dim\ U + dim\ W$. Quando siamo sicuri che $U + W$ è la somma di due sottospazi che sono in somma diretta, al posto di $U + W$ possiamo scrivere:

$$U \oplus W.$$

In particolare, per avere una base di $U \oplus W$ basta fare l'unione di una base di U con una base di W (si osserva immediatamente che i vettori di questa unione generano $U \oplus W$ e inoltre sono nel 'giusto numero', ossia il loro numero è $dim\ U + dim\ W$, vedi il Corollario 2.17).

Per esempio, in $\mathbb{R}^4$, il sottospazio

$$U =< \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \\ 2 \end{pmatrix} >$$

e il sottospazio

$$W = < \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} >$$

sono in somma diretta, e una base di $U \oplus W$ è data dai tre vettori

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

Esercizio 4.9. Motivare come mai è valido il seguente criterio per stabilire se, dato un sottospazio U, un certo vettore v vi appartiene o no: si controlla se U e $< v >$ sono in somma diretta, ovvero si calcola la dimensione di $U + < v >$ e se risulta uguale a $dim\ U + 1$ allora $v \notin U$, se invece è uguale a $dim\ U$ allora $v \in U$.

Dati due sottospazi U e W di uno spazio vettoriale V, può capitare che siano in somma diretta e che inoltre $U \oplus W = V$. Si dice in questo caso che i due sottospazi sono l'uno il *complementare* dell'altro. Un esempio banale di sottospazi complementari è fornito da $\{O\}$ e da V stesso, visto che vale $\{O\} \oplus V = V$. In generale è sempre possibile trovare lo spazio complementare ad un sottospazio proprio di V, come mostra il seguente esercizio.

Esercizio 4.10. Dimostrare che, dato un sottospazio vettoriale proprio U di V, esiste sempre un complementare di U.

Suggerimento: prendere una base di U e completarla ad una base di V (vedi Teorema 2.19). I vettori che abbiamo aggiunto sono la base di uno spazio vettoriale complementare a U.

Osservazione 4.11. Attenzione: un sottospazio vettoriale U di V che non è uguale a V possiede in generale molti complementari (infiniti, se il campo $\mathbb{K}$ ha infiniti elementi). Per esempio, in $\mathbb{R}^3$ un piano passante per l'origine ha per complementare una qualunque retta passante per l'origine e che non giace sul piano. Come ulteriore esempio, il lettore può facilmente verificare che, in $\mathbb{R}^4$, il sottospazio

$$U = < \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \\ 2 \end{pmatrix} >$$

ha come complementare

$$W_1 = < \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} >$$

ma anche

$$W_2 = < \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} >$$

In generale, dati k sottospazi $U_1, U_2, \dots, U_k$ di uno spazio vettoriale V, si dice che tali sottospazi sono in somma diretta se, per ogni $i = 1, 2, \dots, k$, vale che l'intersezione di U_i con la somma di tutti gli altri è uguale a $\{O\}$, ovvero

$$U_i \cap (U_1 + \cdots + \widehat{U_i} + \cdots + U_k) = \{O\}$$

dove il simbolo $\widehat{U_i}$ indica che nella somma si è saltato il termine U_i.

In tal caso per indicare $U_1 + U_2 + \cdots + U_k$ si può usare la notazione:

$$U_1 \oplus U_2 \oplus \cdots \oplus U_k$$

Esercizio 4.12. Dimostrare che, se $U_1, U_2, \dots, U_k$ sono in somma diretta, vale:

$$dim\ (U_1 \oplus U_2 \oplus \cdots \oplus U_k) = dim\ U_1 + dim\ U_2 + \cdots + \dim\ U_k$$

Suggerimento. Dimostrarlo per induzione su k.

In base all'esercizio precedente, osserviamo che per trovare una base di

$$U_1 \oplus U_2 \oplus \cdots \oplus U_k$$

basta scegliere una base per ognuno dei sottospazi U_i e poi fare l'unione (si vede immediatamente che questi elementi sono generatori e il loro numero è 'il numero giusto'). Se accade che

$$U_1 \oplus U_2 \oplus \cdots \oplus U_k = V$$

otterremo in tal modo una base dell'intero spazio.

4. Altri esercizi

Esercizio 4.13. Trovare un complementare in $\mathbb{R}^5$ del sottospazio

$$U = < \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \\ 2 \\ 1 \end{pmatrix} >$$

Esercizio 4.14. Stabilire se i due sottospazi di $\mathbb{R}^4$

$$U = < \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \\ 2 \end{pmatrix} >$$

e

$$W = < \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \end{pmatrix} >$$

sono in somma diretta.

CAPITOLO 5

Applicazioni lineari e matrici invertibili

1. Endomorfismi lineari invertibili

Abbiamo già incontrato nei capitoli precedenti applicazioni lineari invertibili. In questo paragrafo torniamo sull'argomento per alcuni approfondimenti; ci occuperemo in particolare di applicazioni lineari invertibili che hanno come dominio e codominio lo stesso spazio vettoriale V. Nel prossimo paragrafo descriveremo un algoritmo che, data la matrice associata ad una applicazione lineare invertibile, permette di trovare la matrice associata alla applicazione inversa.

Consideriamo uno spazio vettoriale V di dimensione n sul campo $\mathbb{K}$ e una applicazione lineare $L : V \to V$. Una tale applicazione si dice *endomorfismo lineare di V*. Indicheremo con $End(V)$ l'insieme di tutti gli endomorfismi lineari di V.

Proposizione 5.1. *Un endomorfismo L di V è invertibile se e solo se ha rango n. La funzione inversa $L^{-1} : V \to V$ è anch'essa una applicazione lineare.*

Dimostrazione. Supponiamo che L abbia rango n. Questo significa, per la definizione di rango (vedi Definizione 2.32), che $Imm\ L$ è un sottospazio di V di dimensione n; ma allora $Imm\ L = V$ e dunque L è surgettiva. Inoltre, per il Teorema 2.21 sappiamo che la dimensione di $Ker\ L$ è 0, dunque $Ker\ L = \{O\}$ e L è iniettiva. In conclusione, abbiamo mostrato che L è bigettiva e dunque invertibile.

Viceversa, se L è invertibile, allora in particolare è surgettiva, dunque $Imm\ L = V$ e il rango di L, che è uguale a $dim\ Imm\ L$, è uguale a n.

Quanto al fatto che l'inversa L^{-1} sia anch'essa lineare, dobbiamo verificare che, per ogni $v, w \in V$ e per ogni $\lambda \in \mathbb{K}$, valga $L^{-1}(v + w) = L^{-1}(v) + L^{-1}(w)$ e $L^{-1}(\lambda v) = \lambda L^{-1}(v)$. Facciamo a titolo di esempio la prima di queste due verifiche. Per dimostrare l'uguaglianza tra due elementi di V basta mostrare che hanno la stessa immagine tramite L, infatti, essendo L iniettiva, $a, b \in V$ sono uguali se e solo se $L(a) = L(b)$. Dunque nel caso particolare che ci interessa:

$$L^{-1}(v + w) = L^{-1}(v) + L^{-1}(w)$$

vale se e solo se vale

$$L(L^{-1}(v + w)) = L\left(L^{-1}(v) + L^{-1}(w)\right).$$

Si verifica facilmente, usando la linearit di L, che la seconda uguaglianza vera. Infatti per il membro di destra abbiamo $L(L^{-1}(v + w)) = v + w$ e per il membro di sinistra, utilizzando la linearità di L, $L\left(L^{-1}(v) + L^{-1}(w)\right) = L(L^{-1}(v)) + L(L^{-1}(w)) = v + w$. Dunque anche L^{-1} lineare, come volevamo dimostrare. $\square$

Osservazione 5.2. Il lettore può molto facilmente adattare (esercizio!) la dimostrazione precedente per ottenere un risultato che generalizza in questo senso quello precedente: dati due spazi vettoriali V e W entrambi di dimensione n e una

applicazione lineare $L : V \to W$, l'applicazione L è invertibile se e solo se ha rango n; l'applicazione inversa L^{-1} è anch'essa lineare.
Abbiamo invece gi osservato che se V e W hanno dimensioni diverse, rispettivamente m e n, nessuna applicazione lineare L da V a W pu essere invertibile. Infatti, avendo in mente la relazione che lega la dimensione del nucleo di L, dell'immagine di L e di V ($dim_{\mathbb{K}}(Imm(L)) + dim_{\mathbb{K}}(Ker(L)) = dim_{\mathbb{K}} V$), si ha:

- se $m > n$ allora la dimensione di $Imm(L)$ al massimo n. Quindi la dimensione di $Ker(L)$ almeno $m - n$, ovvero maggiore di 0, e L non iniettiva (in sintesi: non esistono applicazioni lineari iniettive da V a W, se la dimensione di V maggiore di quella di W);
- se $m < n$ allora la dimensione di $Imm(L)$ al massimo m. Quindi dimensione di $Imm(L)$ minore della dimensione di W, e L non surgettiva (in sintesi: non esistono applicazioni lineari surgettive da V a W, se la dimensione di V minore di quella di W);

Se fissiamo una base di V, ad ogni endomorfismo $L \in End(V)$ viene associata una matrice $[L] \in Mat_{n\times n}(\mathbb{K})$ (per costruire $[L]$ si è usata la base scelta come base sia "in partenza" sia "in arrivo"). Se L è invertibile, consideriamo l'inversa L^{-1} e la matrice ad essa associata $[L^{-1}]$. Visto che $L \circ L^{-1} = L^{-1} \circ L = I$, il Teorema 2.26 ci assicura che in $Mat_{n\times n}(\mathbb{K})$ vale

$$[L^{-1}][L] = [L][L^{-1}] = [I] = I$$

(ricordiamo che, quando la base scelta in partenza e in arrivo è la stessa, $[I]$ è la matrice identità, che avevamo convenuto di indicare sempre col simbolo I, vedi Osservazione 1.74).

Dunque la matrice $[L]$ è invertibile e ha per inversa $[L^{-1}]$. Sempre applicando il Teorema 2.26 otteniamo il viceversa: se la matrice $[L]$ associata ad un endomorfismo lineare è invertibile allora anche L è invertibile e la sua inversa è l'applicazione associata alla matrice $[L^{-1}]$.

Alla luce di questa osservazione, la proposizione precedente ha un immediato corollario.

Corollario 5.3. *Una matrice $A \in Mat_{n\times n}(\mathbb{K})$ è invertibile se e solo se il suo rango è n.*

Dimostrazione. Data una matrice $A \in Mat_{n\times n}(\mathbb{K})$ possiamo sempre supporre che sia la matrice associata ad un certo endomorfismo lineare L di uno spazio vettoriale V di dimensione n su cui è stata fissata una base. Dalla osservazione che precede il corollario sappiamo che A è invertibile se e solo se L è invertibile. Dalla Proposizione 5.1 sappiamo che L è invertibile se e solo se ha rango n. Dal Teorema 2.46 e dalla osservazione che lo segue sappiamo che il rango di A è uguale al rango di L. □

2. Il metodo per trovare l'inversa (se esiste) di una matrice quadrata

Come abbiamo visto nel paragrafo precedente, il problema di trovare una inversa di $L \in End(V)$ si può tradurre nel problema di trovare l'inversa in $Mat_{n\times n}(\mathbb{K})$ di una matrice data. Molto spesso questa traduzione è utile nelle applicazioni; dedichiamo questo paragrafo alla descrizione di un metodo concreto per trovare l'inversa di una matrice $A \in Mat_{n\times n}(\mathbb{K})$. Partiamo da una definizione (capiremo in seguito perché ci interessa questa forma ridotta): una matrice in forma a scalini per righe (o per colonne) ridotta, si dice *normalizzata*, se ha tutti i pivot uguali a 1.

Osservazione 5.4. Una matrice pu essere portata alla forma a scalini per righe (per colonne) ridotta e normalizzata, attraverso un numero finito di operazioni elementari sulle righe (colonne).
Basta infatti usare le proposizioni 2.41 e 2.12, che assicurano che con un numero finito di operazioni elementari, rispettivamente su righe e colonne, si pu portare qualsiasi matrice ad una forma a scalini per righe (colonne) ridotta, e osservare che per passare alla normalizzata basta moltiplicare le righe (colonne) con pivot a diversi da 1, per $\frac{1}{a}$.

Esercizio 5.5. La forma a scalini per righe (colonne) ridotta e normalizzata di una matrice unica. In particolare, una matrice nxn di rango n ha come forma a scalini per righe I_n, la matrice identità nxn (Provarlo!).

Cominciamo con un esempio.

2.1. Un esempio. Consideriamo la matrice $A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ che ha rango 3, dunque è invertibile, e calcoliamo la sua inversa. Per prima cosa formiamo la matrice

$$(A\ I) = \begin{pmatrix} 3 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Ora con delle operazioni elementari di riga portiamola in forma a scalini per righe ridotta, per esempio nel seguente modo: si sottrae alla prima riga la terza moltiplicata per 3

$$\begin{pmatrix} 3 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 0 & -1 & 1 & 1 & 0 & -3 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

poi si somma alla prima riga la seconda

$$\begin{pmatrix} 0 & -1 & 1 & 1 & 0 & -3 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 2 & 1 & 1 & -3 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

A questo punto si permutano le righe e si ottiene:

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

Per ottenere la forma a scalini ridotta, moltiplichiamo l'ultima riga per $\frac{1}{2}$

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

sottraiamo alla seconda riga la terza riga

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

infine sottraiamo alla prima riga la seconda:

$$\begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & -\frac{3}{2} \end{pmatrix}.$$

La matrice

$$B = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{3}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{3}{2} \end{pmatrix}$$

è l'inversa di A, come il lettore può immediatamente verificare.

2.2. Come mai il metodo funziona? Descriviamo di nuovo, da un punto di vista più generale, il metodo illustrato dall'esempio e spieghiamo come mai funziona. Consideriamo una matrice $A \in Mat_{n\times n}(\mathbb{K})$ e cerchiamo la sua inversa; dato il Corollario 5.3 possiamo supporre che A abbia rango n.

Per prima cosa creiamo una matrice $n \times 2n$ "ponendo accanto" le colonne di A e quelle di I. Indicheremo tale matrice col simbolo:

$$(A\ I)\,.$$

Adesso possiamo agire con operazioni elementari di riga in modo da ridurre la matrice in forma a scalini per righe ridotta. Poichè A ha rango n, anche $(A\ I)$ ha rango n. Un modo per rendersene conto è il seguente: il rango di $(A\ I)$ è minore o uguale a n visto che ha n righe, ed è maggiore o uguale a n, visto che si individuano facilmente n colonne linearmente indipendenti (quelle che provengono da A, oppure quelle che provengono da I).

Allora quando la matrice $(A\ I)$ viene ridotta in forma a scalini per righe ridotta, deve avere esattamente n scalini, dunque (vedi esercizio 5.5) deve avere la forma:

$$(I\ B)\,.$$

Affermiamo che la matrice B (quadrata, di formato $n\times n$) che si ricava dalla matrice precedente è proprio l'inversa di A che cercavamo.

Infatti agire con operazioni di riga equivale, come sappiamo dal Paragrafo 5 del Capitolo 2, a moltiplicare a sinistra la matrice $(A\ I)$ per una matrice invertibile U di formato $n \times n$, dunque:

$$U\,(A\ I) = (I\ B)\,.$$

Per come è definito il prodotto righe per colonne,

$$U\,(A\ I) = (UA\ \ UI)$$

(per rendersene conto può essere utile osservare che la colonna i-esima di $U\,(A\ I)$ è uguale a UC_i, dove C_i è la colonna i-esima di $(A\ I)$).

Dalle due uguaglianze precedenti ricaviamo

$$(UA\ \ UI) = (I\ B)$$

ossia le relazioni $UA = I$ e $UI = B$ che ci dicono che U è l'inversa di A e che $U = B$, come avevamo annunciato.

Osservazione 5.6. La relazione $UA = I$, ossia $BA = I$, ci dice solo che B è l'inversa sinistra di A. In generale sappiamo che il prodotto tra matrici (e anche la composizione tra funzioni) non commutativo. Dunque, il fatto che A sia invertibile, ci dice che esiste una matrice B tale che $BA = I$ (B appunto inversa sinistra di A), e che esiste una matrice C tale che $AC = I$ (C inversa destra di A). Ma possiamo mostrare facilmente che B coincide con l'inversa destra C di A. Come detto C deve soddisfare per definizione $AC = I$. Ora moltiplichiamo entrambi i membri della relazione $BA = I$, a destra, per C: $BAC = IC$. Usando la proprietà associativa del prodotto in $Mat_{n\times n}(\mathbb{K})$ otteniamo, visto che $AC = I$: $B = C$,.

Osservazione 5.7. Se la matrice A da cui siamo partiti non fosse stata invertibile, ossia se avesse avuto rango minore di n, il procedimento descritto per trovare l'inversa (ovviamente) non avrebbe funzionato: infatti nella riduzione a scalini non avremmo trovato n pivots nella parte di sinistra della matrice e non saremmo riusciti a portare $(A\ I)$ nella forma $(I\ B)$.

Esercizio 5.8. Dimostrare che una matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in Mat_{2\times 2}(\mathbb{K})$ è invertibile se e solo se $ad - bc \neq 0$. Calcolare, in tal caso, l'inversa.

Nota: il risultato sarà la matrice:

$$\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

.

3. Cambiamento di base nel caso degli endomorfismi lineari

Sia V uno spazio vettoriale di dimensione finita n sul campo $\mathbb{K}$ e sia $L \in End(V)$. Supponiamo di avere due basi di V, una data dai vettori $v_1, v_2, \ldots, v_n$ e l'altra dai vettori $e_1, e_2, \ldots e_n$. In questo paragrafo studieremo la relazione che lega le matrici associate a L rispetto a tali basi,

$$[L]_{\substack{v_1, v_2, \ldots, v_n \\ v_1, v_2, \ldots, v_n}} \quad \text{e} \quad [L]_{\substack{e_1, e_2, \ldots, e_n \\ e_1, e_2, \ldots, e_n}}$$

Per prima cosa scriviamo ogni vettore v_i come combinazione lineare dei vettori della base $e_1, e_2, \ldots e_n$:

$$v_1 = a_{11}e_1 + a_{21}e_2 + \cdots + a_{n1}e_n$$

$$v_2 = a_{12}e_1 + a_{22}e_2 + \cdots + a_{n2}e_n$$

$$\ldots\ldots\ldots\ldots$$

$$v_n = a_{1n}e_1 + a_{2n}e_2 + \cdots + a_{nn}e_n$$

Osserviamo a questo punto che la matrice associata all'endomorfismo identità $I \in End(V)$ prendendo come base in partenza $v_1, v_2, \ldots, v_n$ e come base in arrivo $e_1, e_2, \ldots e_n$ è la seguente:

$$[I]_{\substack{v_1,v_2,\ldots,v_n \\ e_1,e_2,\ldots,e_n}} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \ldots & a_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nn} \end{pmatrix}$$

Infatti nella prima colonna abbiamo scritto i coefficienti di $I(v_1)$ (che è uguale a v_1) rispetto alla base $e_1, e_2, \ldots e_n$, nella seconda colonna i coefficienti di $I(v_2) = v_2$ e così via... La matrice appena trovata è una *matrice di cambiamento di base* e la chiameremo M. Osserviamo subito che M è invertibile. Infatti pensiamo alla composizione di endomorfismi $I \circ I$ ovvero $V \xrightarrow{I} V \xrightarrow{I} V$ e consideriamo il primo spazio V e l'ultimo muniti della base $v_1, v_2, \ldots, v_n$, mentre lo spazio V al centro lo consideriamo con la base $e_1, e_2, \ldots, e_n$. Applicando il Teorema 2.26 otteniamo:

$$[I \circ I]_{\substack{v_1,v_2,\ldots,v_n \\ v_1,v_2,\ldots,v_n}} = [I]_{\substack{e_1,e_2,\ldots,e_n \\ v_1,v_2,\ldots,v_n}} [I]_{\substack{v_1,v_2,\ldots,v_n \\ e_1,e_2,\ldots,e_n}}$$

Visto che $I \circ I = I$ possiamo riscrivere

$$[I]_{\substack{v_1,v_2,\ldots,v_n \\ v_1,v_2,\ldots,v_n}} = [I]_{\substack{e_1,e_2,\ldots,e_n \\ v_1,v_2,\ldots,v_n}} [I]_{\substack{v_1,v_2,\ldots,v_n \\ e_1,e_2,\ldots,e_n}}$$

Ora la matrice al membro di sinistra è, come sappiamo, la matrice identità I, mentre la matrice più a destra è M, dunque:

$$I = [I]_{\substack{e_1,e_2,\ldots,e_n \\ v_1,v_2,\ldots,v_n}} M$$

Questo ci permette di concludere che M è invertibile e che $M^{-1} = [I]_{\substack{e_1,e_2,\ldots,e_n \\ v_1,v_2,\ldots,v_n}}$.

A questo punto possiamo enunciare il teorema che descrive la relazione fra le matrici associate a L rispetto alle due diverse basi:

Teorema 5.9. *Con le notazioni introdotte sopra, vale:*

$$[L]_{\substack{v_1,v_2,\ldots,v_n \\ v_1,v_2,\ldots,v_n}} = M^{-1}[L]_{\substack{e_1,e_2,\ldots,e_n \\ e_1,e_2,\ldots,e_n}} M$$

DIMOSTRAZIONE. Consideriamo la composizione di endomorfismi $I \circ L \circ I$ e applichiamo il Teorema 2.26:

$$\underset{v_i}{V} \xrightarrow{I} \underset{e_i}{V} \xrightarrow{L} \underset{e_i}{V} \xrightarrow{I} \underset{v_i}{V}$$

(sotto ogni copia di V il simbolo e_i oppure v_i indica quale base abbiamo scelto). Otteniamo[1]:

$$[I \circ L \circ I]_{\substack{v_1,v_2,\dots,v_n \\ v_1,v_2,\dots,v_n}} = [I]_{\substack{e_1,e_2,\dots,e_n \\ v_1,v_2,\dots,v_n}} [L]_{\substack{e_1,e_2,\dots,e_n \\ e_1,e_2,\dots,e_n}} [I]_{\substack{v_1,v_2,\dots,v_n \\ e_1,e_2,\dots,e_n}}$$

Visto che $I \circ L \circ I = L$ la formula appena ottenuta si può riscrivere come:

$$[L]_{\substack{v_1,v_2,\dots,v_n \\ v_1,v_2,\dots,v_n}} = M^{-1}[L]_{\substack{e_1,e_2,\dots,e_n \\ e_1,e_2,\dots,e_n}} M$$

□

Ricordiamo che il problema di trovare la matrice associata a L rispetto ad una base se si conosce la matrice associata rispetto ad un'altra base può essere affrontato anche senza scrivere le matrici M e M^{-1}, come mostra l'Esempio 1.71, ma il teorema appena dimostrato ha una grande importanza dal punto di vista teorico, come vedremo nei prossimi capitoli.

Per esempio, nell'Esercizio 1.87 abbiamo definito l'applicazione *traccia*

$$\mathcal{T} \,:\, Mat_{n\times n}(\mathbb{K}) \to \mathbb{K}$$

nel seguente modo:

$$\mathcal{T}\left((a_{ij})\right) = a_{11} + a_{22} + \cdots + a_{nn}.$$

È naturale chiedersi se, dato un endomorfismo $L \in End(V)$, la funzione traccia dia lo stesso valore su tutte le matrici che si possono associare a V, in altre parole se vale:

$$\mathcal{T}\left([L]_{\substack{v_1,v_2,\dots,v_n \\ v_1,v_2,\dots,v_n}}\right) = \mathcal{T}\left([L]_{\substack{e_1,e_2,\dots,e_n \\ e_1,e_2,\dots,e_n}}\right)$$

per ogni scelta delle basi $e_1, e_2, \dots e_n$ e $v_1, v_2, \dots v_n$.

La risposta è sì: la traccia non dipende dalla base scelta e dunque possiamo anche considerarla come applicazione lineare da $End(V)$ a $\mathbb{K}$. Per mostrarlo, innanzitutto utilizziamo il Teorema 5.9 e scriviamo:

$$\mathcal{T}\left([L]_{\substack{v_1,v_2,\dots,v_n \\ v_1,v_2,\dots,v_n}}\right) = \mathcal{T}\left(M^{-1}[L]_{\substack{e_1,e_2,\dots,e_n \\ e_1,e_2,\dots,e_n}} M\right)$$

A questo punto ricordiamo che per ogni $A, B \in Mat_{n\times n}(\mathbb{K})$ vale $\mathcal{T}(AB) = \mathcal{T}(BA)$ (vedi l'Esercizio 1.87), dunque:

$$\mathcal{T}\left(\left(M^{-1}[L]_{\substack{e_1,e_2,\dots,e_n \\ e_1,e_2,\dots,e_n}}\right) M\right) = \mathcal{T}\left(M\left(M^{-1}[L]_{\substack{e_1,e_2,\dots,e_n \\ e_1,e_2,\dots,e_n}}\right)\right) =$$

$$= \mathcal{T}\left(MM^{-1}[L]_{\substack{e_1,e_2,\dots,e_n \\ e_1,e_2,\dots,e_n}}\right) = \mathcal{T}\left([L]_{\substack{e_1,e_2,\dots,e_n \\ e_1,e_2,\dots,e_n}}\right)$$

[1]Il Teorema 2.26 si riferisce alla composizione di due applicazioni lineari, ma la versione con tre applicazioni lineari si ricava subito usando la proprietà associativa e applicando due volte il teorema.

Questa catena di uguaglianze conduce, come avevamo annunciato, a:

$$\mathcal{T}\left([L]\begin{array}{l} v_1, v_2, \ldots, v_n \\ v_1, v_2, \ldots, v_n \end{array}\right) = \mathcal{T}\left([L]\begin{array}{l} e_1, e_2, \ldots, e_n \\ e_1, e_2, \ldots, e_n \end{array}\right)$$

4. Altri esercizi

Esercizio 5.10. Trovare, se esiste, l'inversa della seguente matrice a coefficienti in $\mathbb{R}$:

$$\begin{pmatrix} 1 & -2 & 3 \\ 2 & 5 & 10 \\ 0 & 0 & 1 \end{pmatrix}$$

Esercizio 5.11. Trovare, se esiste, l'inversa della seguente matrice a coefficienti in $\mathbb{R}$:

$$\begin{pmatrix} -11 & 2 & 2 \\ -1 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix}$$

Esercizio 5.12. Trovare, se esistono, le inverse delle matrici degli Esercizi 5.10 e 5.11, lette stavolta come matrici a coefficienti in $\mathbb{Z}_2$.

Esercizio 5.13. Trovare, se esistono, le inverse delle matrici degli Esercizi 5.10 e 5.11, lette come matrici a coefficienti in $\mathbb{Z}_3$.

Esercizio 5.14. Dimostrare che se una matrice $A \in Mat_{n\times n}(\mathbb{R})$ ha tutti i coefficienti in $\mathbb{Z}$ ed è invertibile, allora la sua inversa ha coefficienti in $\mathbb{Q}$.

Esercizio 5.15. Sia $L : \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione lineare la cui matrice rispetto alla base standard è

$$[L] = \begin{pmatrix} 2 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{5}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{5}{2} \end{pmatrix}$$

Consideriamo adesso la base di $\mathbb{R}^3$ data dai vettori $v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Calcolare le matrici M e M^{-1} di cambiamento di base fra la base standard e la base v_1, v_2, v_3 e scrivere la matrice associata a L ripetto alla base v_1, v_2, v_3.

Esercizio 5.16. Sia $L : \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione lineare la cui matrice rispetto alla base standard è

$$[L] = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

Consideriamo adesso la base di $\mathbb{R}^3$ data dai vettori $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Calcolare la matrice associata a L rispetto alla base v_1, v_2, v_3 in due modi: direttamente (come nell'Esempio 1.71) e attraverso il calcolo delle matrici di cambiamento di base M e M^{-1}.

CAPITOLO 6

Informazioni sul determinante

In questo capitolo daremo alcune informazioni sulla funzione *determinante*. Alcuni teoremi non verranno dimostrati nel corso, ma utilizzeremo gli enunciati. Il lettore interessato può trovare le dimostrazioni per esempio in [**Ab**].

1. Definizione del determinante di una matrice quadrata

Il determinante è una funzione

$$Det\ :\ Mat_{n\times n}(\mathbb{K}) \to \mathbb{K}$$

Per alleggerire la notazione talvolta indicheremo con con $|a_{ij}|$ oppure con $Det\ (a_{ij})$ o con $Det\ A$ - invece che con $Det\,((a_{ij}))$ o $Det(A)$ - il determinante di una matrice $A = (a_{ij})$.

Il determinante è definito ricorsivamente, al crescere di n, nel seguente modo:

(1) il determinante di una matrice 1×1 è uguale all'unico coefficiente della matrice:
$$Det\ (a) = a$$
(2) dato $n \geq 2$, il deteminante di una matrice $A = (a_{ij})$ di formato $n \times n$ può essere ottenuto come combinazione lineare dei coefficienti di una qualunque riga, diciamo la i-esima, tramite la seguente formula:

$$Det\ A = (-1)^{1+i} a_{i1} Det\ A_{i1} + (-1)^{2+i} a_{i2} Det\ A_{i2} + \dots + (-1)^{i+n} a_{in} Det\ A_{in} \tag{1.1}$$

dove A_{ij} indica la matrice (quadrata) di formato $(n-1)\times(n-1)$ che si ottiene da A cancellando la riga i-esima e la colonna j-esima.

Osservazione 6.1. Dalla definizione possiamo immediatamente ricavare la seguente formula per il determinante di una matrice 2×2:

$$Det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

Osservazione 6.2. Il determinante si può ottenere anche come combinazione lineare dei coefficienti di una qualunque colonna, diciamo la j-esima, tramite la seguente formula:

$$Det\ A = (-1)^{1+j} a_{1j} Det\ A_{1j} + (-1)^{2+j} a_{2j} Det\ A_{2j} + \dots + (-1)^{j+n} a_{nj} Det\ A_{nj} \tag{1.2}$$

Ovviamente il fatto che sia equivalente calcolare il determinante a partire da una qualunque riga o da una qualunque colonna va dimostrato, ma noi abbiamo deciso di omettere questa dimostrazione.

Esempio 6.3. Data in $Mat_{3\times 3}(\mathbb{R})$ la matrice

$$A = \begin{pmatrix} 3 & 2 & 5 \\ 2 & 0 & 1 \\ 4 & 2 & 6 \end{pmatrix}$$

per calcolare il determinante si sceglie una riga e poi si applica la formula (1.1) (oppure si sceglie una colonna e poi si applica la (1.2)). Per esempio, scegliamo la seconda riga:

$$Det\ A = -2Det\begin{pmatrix} 2 & 5 \\ 2 & 6 \end{pmatrix} + 0Det\begin{pmatrix} 3 & 5 \\ 4 & 6 \end{pmatrix} - Det\begin{pmatrix} 3 & 2 \\ 4 & 2 \end{pmatrix} =$$
$$= -2(12-10) - (6-8) = -4 + 2 = -2$$

Osservazione 6.4. Nel caso delle matrici 3×3 il determinante si può calcolare anche mediante la seguente *regola di Sarrus*[1]. Data

$$B = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

si forma la seguente matrice 3×5

$$\begin{pmatrix} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{pmatrix}$$

dopodiché si sommano i tre prodotti dei coefficienti che si trovano sulle tre diagonali che scendono da sinistra a destra e si sottraggono i tre prodotti dei coefficienti che si trovano sulle tre diagonali che salgono da sinistra a destra:

$$Det\ B = aei + bfg + cdh - gec - hfa - idb$$

Verifichiamo che nel caso della matrice

$$A = \begin{pmatrix} 3 & 2 & 5 \\ 2 & 0 & 1 \\ 4 & 2 & 6 \end{pmatrix}$$

la regola di Sarrus dia lo stesso risultato -2 che abbiamo già calcolato:

$$3 \cdot 0 \cdot 6 + 2 \cdot 1 \cdot 4 + 5 \cdot 2 \cdot 2 - 4 \cdot 0 \cdot 5 - 2 \cdot 1 \cdot 3 - 6 \cdot 2 \cdot 2 = 8 + 20 - 6 - 24 = -2.$$

Esercizio 6.5. Dimostrare che il determinante di una matrice (a_{ij}) triangolare superiore (ossia tale che $a_{ij} = 0$ se $i > j$) è uguale al prodotto dei coefficienti che si trovano sulla diagonale. Lo stesso per una matrice triangolare inferiore.

2. Il determinante e il calcolo del rango di una matrice

Data una matrice $A = (a_{ij}) \in Mat_{n \times m}(\mathbb{K})$ e dato un numero intero positivo k minore o uguale al minimo fra m e n, possiamo scegliere k righe fra le n righe (diciamo che scegliamo le righe $i_1, i_2, \dots, i_k$) e k colonne fra le m colonne (diciamo che scegliamo le colonne $j_1, j_2, \dots, j_k$).

Definizione 6.6. Data la scelta di k righe e k colonne come sopra, chiamiamo *minore di A di formato $k \times k$* la matrice ottenuta da A cancellando tutti i coefficienti eccetto quelli che giacciono contemporaneamente su una delle righe e su una delle colonne scelte, in altre parole cancellando tutti i coefficienti eccetto gli a_{ij} per cui $i \in \{i_1, i_2, \dots, i_k\}$ e $j \in \{j_1, j_2, \dots, j_k\}$.

Osservazione 6.7. Da una matrice $n \times m$ è possibile ricavare $\binom{n}{k}\binom{m}{k}$ minori di formato $k \times k$.

[1]La regola prende nome dal matematico francese Pierre Frederic Sarrus (1798-1861).

I determinanti dei minori possono essere utilizzati per calcolare il rango di una matrice $n \times m$, come risulta dal seguente teorema.

Teorema 6.8. *Data una matrice $A = (a_{ij}) \in Mat_{n\times m}(\mathbb{K})$, supponiamo che esista un minore di formato $k \times k$ il cui determinante è diverso da 0. Allora il rango di A è maggiore o uguale a k. Se $k = n$ oppure $k = m$ allora il rango di A è uguale a k. Se $k < n$ e $k < m$ e tutti i determinanti dei minori di formato $(k+1) \times (k+1)$ sono uguali a 0 allora il rango di A è uguale a k.*

Osservazione 6.9. Se una matrice quadrata $n \times n$ ha determinante diverso da zero, allora per il teorema precedente ha rango n e dunque è invertibile. Sempre per il teorema vale anche il viceversa: se una matrice $n \times n$ ha determinante uguale a 0, allora il suo rango è strettamente minore di n e dunque la matrice non è invertibile. Nel caso 2×2, data

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

con determinante diverso da zero, ossia $ad - bc \neq 0$, l'inversa (vedi Esercizio 5.8) si scrive esplicitamente come:

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Esempio 6.10. La matrice

$$\begin{pmatrix} 3 & 9 & 4 & 7 & 12 \\ 1 & \mathbf{3} & \mathbf{2} & 0 & \mathbf{5} \\ 1 & \mathbf{2} & \mathbf{0} & 0 & \mathbf{1} \\ 1 & \mathbf{4} & \mathbf{2} & 7 & \mathbf{6} \end{pmatrix}$$

ha rango maggiore o uguale a 3 in quanto contiene un minore 3×3 (quello individuato dalle righe seconda, terza e quarta e dalle colonne seconda, terza e quinta, ovvero dai coefficienti in grassetto) che ha determinante diverso da 0 (è uguale a -2, come abbiamo calcolato nel paragrafo precedente). Inoltre, poiché tutti i minori 4×4 hanno determinante uguale a zero, la matrice ha rango esattamente 3. Quest'ultima verifica richiede il controllo del determinante di 5 minori. È più rapido osservare che la prima riga è uguale alla somma delle altre tre righe, dunque il rango è minore o uguale a 3: poiché dal calcolo del determinante del minore 3×3 sappiamo che il rango è maggiore o uguale a 3, allora è esattamente 3.

Il calcolo del rango attraverso il Teorema 6.8 può richiedere molte verifiche, ed è in generale meno conveniente della riduzione di Gauss. Il seguente teorema può comunque aiutare a ridurre i determinanti da calcolare:

Teorema 6.11 (Teorema degli orlati)**.** *Data una matrice $A = (a_{ij}) \in Mat_{n\times m}(\mathbb{K})$, supponiamo che esista un minore K di formato $k \times k$ (con $k < n$ e $k < m$) il cui determinante è diverso da 0. Se sono uguali a 0 tutti i determinanti dei minori di formato $(k+1) \times (k+1)$ che si ottengono aggiungendo una riga e una colonna a quelle scelte per formare il minore K, allora il rango di A è uguale a k, altrimenti è strettamente maggiore di k.*

Dunque se abbiamo un minore di formato $k \times k$ con determinante diverso da 0 e vogliamo decidere se la matrice ha rango k oppure ha rango strettamente maggiore di k basta controllare $(n-k)(m-k)$ minori di formato $(k+1) \times (k+1)$, non tutti i $\binom{n}{k+1}\binom{m}{k+1}$ minori di formato $(k+1) \times (k+1)$. Per rendersi conto del ‘risparmio’, anche con matrici piccole, consideriamo una matrice A di formato 5×6 e supponiamo di conoscere un minore 3×3 con determinante diverso da 0. Per controllare se la matrice ha rango 3 basta controllare i determinanti dei 6 minori 4×4 che si ottengono “orlando” il minore dato con una riga e una colonna in più; non occorre calcolare i determinanti di tutti i 75 minori 4×4 di A.

3. Il teorema di Binet

Il determinante non è una applicazione lineare (a parte il caso banale delle matrici 1×1). In particolare in generale $Det(A + B) \neq Det(A) + Det(B)$. Vale invece il seguente:

Teorema 6.12 (Teorema di Binet[2])**.** *Siano $A, B \in Mat_{n\times n}(\mathbb{K})$. Allora*

$$Det(AB) = Det(A)Det(B)$$

Come prima applicazione osserviamo

Corollario 6.13. *Se $M \in Mat_{n\times n}(\mathbb{K})$ è una matrice invertibile, allora*

$$Det(M^{-1}) = \frac{1}{Det(M)}$$

Dimostrazione. Calcoliamo $Det(M^{-1}M)$. Per il Teorema di Binet vale $Det(M^{-1}M) = Det(M^{-1})Det(M)$. D’altra parte $M^{-1}M = I$ e $Det(I) = 1$. □

Grazie al Teorema di Binet possiamo osservare che, dato un endomorfismo $L \in End(V)$, il determinante assume lo stesso valore su tutte le matrici che si associano a V al variare delle basi dello spazio, ossia vale:

$$Det\left([L]\,{}^{v_1,v_2,\dots,v_n}_{v_1,v_2,\dots,v_n}\right) = Det\left([L]\,{}^{e_1,e_2,\dots,e_n}_{e_1,e_2,\dots,e_n}\right)$$

per ogni scelta di due basi $e_1, e_2, \dots e_n$ e $v_1, v_2, \dots v_n$ di V.

Infatti, per il Teorema 5.9 possiamo scrivere:

$$Det\left([L]\,{}^{v_1,v_2,\dots,v_n}_{v_1,v_2,\dots,v_n}\right) = Det\left(M^{-1}[L]\,{}^{e_1,e_2,\dots,e_n}_{e_1,e_2,\dots,e_n}\,M\right)$$

A questo punto, per il Teorema di Binet e per il Corollario 6.13, possiamo concludere che

$$Det\left([L]\,{}^{v_1,v_2,\dots,v_n}_{v_1,v_2,\dots,v_n}\right) =$$

$$= Det\left(M^{-1}\right) Det\left([L]\,{}^{e_1,e_2,\dots,e_n}_{e_1,e_2,\dots,e_n}\right) Det\left(M\right) =$$

$$= Det\left([L]\,{}^{e_1,e_2,\dots,e_n}_{e_1,e_2,\dots,e_n}\right)$$

[2]Il nome deriva dal matematico francese Jacques Philippe Marie Binet (1786-1856).

4. Proprietà del determinante rispetto alle mosse di riga e di colonna

Studiamo come cambia il determinante se facciamo una operazione elementare di riga o di colonna su una matrice $A \in Mat_{n\times n}(\mathbb{K})$. Ricordiamo che le operazioni elementari di colonna sono di tre tipi:

- 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.

Una operazione del primo tipo corrisponde a moltiplicare A a destra per la matrice $n\times n$ (chiamata $M_{ij\lambda}$ nel Paragrafo 4 del Capitolo 2)) 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. La matrice M_{ij} è triangolare e il suo determinante è uguale a 1, dunque $Det(AM_{ij})$ è uguale a $Det(A)$ per il Teorema di Binet.

Analogamente, si osserva che una operazione di colonna del terzo tipo ha come effetto quello di cambiare il segno del determinante.

Quanto alle operazioni del secondo tipo, dalla definizione stessa di determinante si ricava che, se si moltiplica una colonna per uno scalare $k \neq 0$, anche il determinante della matrice risulterà moltiplicato per k (per convincersene basta calcolare il determinante proprio a partire da quella colonna).

Considerazioni analoghe valgono, ovviamente, per le operazioni elementari di riga.

Una conseguenza di queste osservazioni è che se vogliamo solo sapere se il determinante di una certa matrice è uguale a 0 oppure no (come capita quando ci interessa solo calcolare il rango, oppure in altre situazioni che incontreremo nei prossimi capitoli) possiamo, prima di calcolarlo, fare alcune operazioni di riga e/o di colonna. Di solito questo risulta utile se con tali operazioni otteniamo una riga, o una colonna, con molti coefficienti uguali a 0, facilitando il calcolo.

5. Altri esercizi

Esercizio 6.14. Sia A una matrice 2×2 a valori in $\mathbb{R}$.

(1) Dimostrare che esiste una matrice B, di formato 2×2, a valori in $\mathbb{R}$, diversa dalla matrice nulla, tale che $AB = 0$ se e solo se il determinante di A è uguale a 0.
(2) Il risultato precedente è vero anche per le matrici $n \times n$?

Esercizio 6.15. Sia L un endomorfismo lineare dello spazio vettoriale V, sia A la matrice corrispondente all'endomorfismo L in una base fissata di V e sia $Det(A) = 0$. Dire, giustificando la risposta, quali delle seguenti affermazioni sono vere e quali false:

(1) L'endomorfismo L non è surgettivo.
(2) $Ker\ L = \{O\}$.
(3) In una riduzione a scalini per righe di A almeno una riga è nulla.
(4) La matrice A ha almeno uno 0 sulla diagonale principale.
(5) Esiste una base $\mathcal{B}$ di V per cui la matrice associata ad L rispetto a $\mathcal{B}$ ha la prima colonna tutta di zeri.

CAPITOLO 7

Diagonalizzazione di endomorfismi lineari

1. Autovalori e autovettori di un endomorfismo lineare

Sia $T : V \to V$ un endomorfismo lineare dello spazio vettoriale V sul campo $\mathbb{K}$.

Definizione 7.1. Un vettore $v \in V - \{O\}$ si dice un autovettore di T se

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

per un certo $\lambda \in \mathbb{K}$.

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

Definizione 7.2. Se $v \in V - \{O\}$ è un autovettore di T tale che

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

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

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

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

ossia sono autovettori relativi all'autovalore 0.

Definizione 7.3. Dato $\lambda \in \mathbb{K}$ chiamiamo l'insieme

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

autospazio relativo a λ.

Esercizio 7.4. Verificare che un autospazio V_λ è un sottospazio vettoriale di V.

Osservazione 7.5. In particolare abbiamo notato poco fa che $V_0 = Ker\ T$.

Anche se abbiamo definito l'autospazio V_λ per qualunque $\lambda \in \mathbb{K}$, 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 O e da tutti gli autovettori relativi a λ.

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 abbiamo informazioni, per esempio 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"; a colpo d'occhio possiamo sapere tutto di T: il suo rango (dunque anche la dimensione del nucleo), quali sono esattamente i vettori di $Ker\ T$, quali sono (se esistono) i sottospazi in cui T si comporta come l'identità, ossia i sottospazi costituiti dai vettori di V che 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 un certo endomorfismo T esiste una base buona si dice che T è *diagonalizzabile*, altrimenti T è *non diagonalizzabile*.

Esempio 7.6. Consideriamo l'endomorfismo $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ dato da una *rotazione* di angolo θ con centro l'origine. Si verifica immediatamente che, rispetto alla base standard di $\mathbb{R}^2$, questo endomorfismo è rappresentato dalla matrice

$$[R_\theta =] \begin{pmatrix} \cos\theta & -\sin\theta \\ \\ \sin\theta & \cos\theta \end{pmatrix}$$

Per esempio, nel caso di una rotazione di 60^0 (ovvero $\frac{\pi}{3}$), abbiamo:

$$[R_{\frac{\pi}{3}}] = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

Nel caso in cui $0 < \theta < \pi$, non ci sono vettori $v \neq O$ che vengono mandati in un multiplo di se stessi, visto che tutti i vettori vengono ruotati di un angolo che non è nullo e non è di 180^0. Dunque non ci sono autovettori e autovalori.

Nel caso $\theta = 0$ la rotazione è l'identità, dunque tutti i vettori $v \neq O$ sono autovettori relativi all'autovalore 1, e $V_1 = \mathbb{R}^2$.

Nel caso $\theta = \pi$ la rotazione è uguale a $-I$, dunque tutti i vettori $v \neq O$ sono autovettori relativi all'autovalore -1, e $V_{-1} = \mathbb{R}^2$.

Esempio 7.7. Consideriamo l'endomorfismo $T : \mathbb{C}^2 \to \mathbb{C}^2$ che, rispetto alla base standard di $\mathbb{C}^2$, è rappresentato dalla matrice

$$[T] = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

Si tratta della stessa matrice che nell'esempio precedente era associata alla rotazione di $\frac{\pi}{3}$ nel piano reale ma stavolta, visto che stiamo considerando uno spazio vettoriale complesso, riusciamo a trovare autovettori e autovalori per T. Si verifica infatti (esercizio!) che il vettore $\begin{pmatrix} \imath \\ 1 \end{pmatrix}$ è un autovettore relativo all'autovalore $\frac{1+\imath\sqrt{3}}{2}$ e che il vettore $\begin{pmatrix} -\imath \\ 1 \end{pmatrix}$ è un autovettore relativo all'autovalore $\frac{1-\imath\sqrt{3}}{2}$. Poiché i due vettori sono linearmente indipendenti sul campo $\mathbb{C}$, costituiscono una base. Dunque T è diagonalizzabile e

$$V_{\frac{1+\imath\sqrt{3}}{2}} =< \begin{pmatrix} \imath \\ 1 \end{pmatrix} >$$

$$V_{\frac{1-\imath\sqrt{3}}{2}} =< \begin{pmatrix} -\imath \\ 1 \end{pmatrix} >$$

$$\mathbb{C}^2 = V_{\frac{1-\imath\sqrt{3}}{2}} \oplus V_{\frac{1+\imath\sqrt{3}}{2}}.$$

2. Il polinomio caratteristico di un endomorfismo

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

Innanzitutto osserviamo che, perché $\lambda \in \mathbb{K}$ sia una 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 $I : V \to V$ è l'identità. Riscriviamo ancora:

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

Abbiamo scoperto che, se T possiede un autovalore λ, allora l'endomorfismo $T - \lambda I$ non è iniettivo: 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 $[T - \lambda I] = [T] - \lambda I$ dovrà avere determinante uguale a 0 (vedi l'Osservazione 6.9):

$$det([T] - \lambda I) = 0 = det(\lambda I - [T])$$

dove come consuetudine abbiamo indicato con I anche la matrice identità.

Questa osservazione è la premessa per la seguente definizione:

Definizione 7.8. Dato un endomorfismo lineare $T : V \to V$ con $n = dim\ V$, scegliamo una base per V e costruiamo la matrice $[T]$ associata a T rispetto a tale base. Il polinomio caratteristico $P_T(t) \in \mathbb{K}[t]$ dell'endomorfismo T è definito da:

$$P_T(t) = det(t[I] - [T]).$$

Osservazione 7.9. 1) Perché la definizione data abbia senso innanzitutto bisogna verificare che $det(t[I] - [T])$ 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 direttore 1: $P_T(t) = t^n + \ldots\ldots\ldots$ Provate a fare questi esercizi!

2) È fondamentale inoltre che la definizione appena data non dipenda dalla base scelta di V: non sarebbe una definizione buona se con la scelta di due basi diverse ottenessimo due polinomi caratteristici diversi.

Questo problema per fortuna non si verifica. Infatti, se scegliamo due basi b e b' di V, come sappiamo dal Teorema 5.9, le due matrici $[T]_{\substack{b\\b}}$ e $[T]_{\substack{b'\\b'}}$ sono legate dalla seguente relazione: esiste una matrice $[B]$ invertibile tale che

$$[T]_{\substack{b\\b}} = [B]^{-1}[T]_{\substack{b'\\b'}}[B]$$

Usando il teorema di Binet (Teorema 6.12) a questo punto verifichiamo che

$$det\left(tI - [T]_{\substack{b\\b}}\right) = det\left(tI - [B]^{-1}[T]_{\substack{b'\\b'}}[B]\right) = det\left([B]^{-1}\left(tI - [T]_{\substack{b'\\b'}}\right)[B]\right) =$$

$$= det\left([B]^{-1}\right) det\left(tI - [T]_{\substack{b'\\b'}}\right) det\left([B]\right) = det\left(tI - [T]_{\substack{b'\\b'}}\right)$$

Abbiamo dunque mostrato che $P_T(t) = det\ (tI - [T])$ non dipende dalla scelta della base.

Esercizio 7.10. Grazie all'osservazione precedente, sappiamo in particolare che i coefficienti di $p_T(t)$ non dipendono dalla base scelta. Chiamiamo dunque

$$C_r : End(V) \to \mathbb{K}$$

la funzione che, ad ogni endomorfismo T associa il coefficiente di t^r in $P_T(t)$. Dimostrare che C_{n-1} è uguale a meno la traccia ossia $C_{n-1}(T) = -\mathcal{T}(T)$ (per la funzione traccia vedi il Paragrafo 3 del Capitolo 5) e che C_0 è uguale, a meno del segno, al determinante ossia $C_0(T) = \pm Det(T)$. Il polinomio caratteristico ci fornisce dunque l'esempio di altre funzioni che, come il determinante e la traccia, coinvolgono i coefficienti di una matrice $[T]$ ma in realtà non dipendono dalla base scelta.

Esercizio 7.11. Usando le stesse notazioni dell'esercizio precedente, calcolare la funzione C_1 nel caso in cui $dim\ V = 3$, ossia in cui la matrice $[T]$ sia 3×3.

Possiamo a questo punto enunciare il teorema principale che spiega l'utilità del polinomio caratteristico ai fini del problema della diagonalizzazione.

Teorema 7.12. *Considerato T come sopra, vale che uno scalare $\lambda \in \mathbb{K}$ 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 $(\lambda I - T)(v) = 0$. Questo 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. $\square$

Esempio 7.13. Consideriamo l'endomorfismo T dell'Esempio 7.7. Il suo polinomio caratteristico risulta $P_T(t) = t^2 - t + 1$ (verificare!). Questo polinomio ha due radici in $\mathbb{C}$, ovvero $\frac{1-\imath\sqrt{3}}{2}$ e $\frac{1+\imath\sqrt{3}}{2}$, che in effetti, come sappiamo, sono gli autovalori di T.

Notiamo che $t^2 - t + 1$ non ha invece radici in $\mathbb{R}$, coerentemente col fatto, osservato nell'Esempio 7.6, che la rotazione $R_{\frac{\pi}{3}}$ del piano reale non ammette autovettori.

Esercizio 7.14. Sia $F : \mathbb{R}^2 \to \mathbb{R}^2$ l'applicazione lineare definita, nella base standard di $\mathbb{R}^2$, dalla matrice:

$$[F] = \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix}$$

Calcolare $P_F(t)$.

Esercizio 7.15. Sia $F : \mathbb{C}^3 \to \mathbb{C}^3$ l'applicazione lineare definita, nella base standard di $\mathbb{C}^3$, dalla matrice:

$$[F] = \begin{pmatrix} -1 & 0 & 2 \\ 2i & 1 & 2i \\ 1 & 0 & 0 \end{pmatrix}$$

Calcolare $P_F(t)$.

Esercizio 7.16. Sia $F : \mathbb{C}^4 \to \mathbb{C}^4$ l'applicazione lineare definita, nella base standard di $\mathbb{C}^4$, dalla matrice:

$$[F] = \begin{pmatrix} -1 & 0 & 2 & 0 \\ 0 & 1 & 2i & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 3 & 0 & 0 \end{pmatrix}$$

Calcolare $P_F(t)$.

3. Una strategia per scoprire se un endomorfismo è diagonalizzabile

In questo paragrafo descriviamo una strategia in quattro passi che ci permette di scoprire se un endomorfismo $T : V \to V$, dove V è uno spazio vettoriale sul campo $\mathbb{K}$ di dimensione n, 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 (vedi Teorema 7.19) ci assicurerà che gli autospazi $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_k}$ sono in somma diretta (vedi Paragrafo 3 del Capitolo 4). Quindi se
$$V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k} = V$$
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
$$V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k} \subsetneq V$$
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.

3.1. Passo 1. Di questo ci siamo già occupati nel paragrafo precedente: per sapere quali sono gli autovalori di un endomorfismo T possiamo calcolare il polinomio caratteristico $P_T(t)$ e trovare le sue radici in $\mathbb{K}$.

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

3.3. Passo 3. Cominciamo col dimostrare il seguente teorema.

Teorema 7.17. *Dato un endomorfismo lineare* $T : V \to V$*, 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 7.18. 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 un insieme di vettori linearmente indipendenti (c'è un vettore solo... ed è diverso da O perché è un autovettore). Supponiamo di aver dimostrato che l'enunciato è vero fino al caso di $k-1$ vettori e cerchiamo di dimostrarlo per k. Supponiamo allora di avere un insieme di vettori $\{v_1, v_2, \dots, v_k\}$ come nelle ipotesi e che valga:

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = O \tag{3.1}$$

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 $a_1 = a_2 = \cdots = a_k = 0$.

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

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

che svolgendo il calcolo diventa

$$a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 + \cdots + a_k \lambda_k v_k = O$$

Poi invece moltiplichiamo l'equazione per λ_k ottenendo:

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

Per sottrazione da queste due equazioni ricaviamo:

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

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 $a_1 = a_2 = \cdots = a_{k-1} = 0$. Sostituendo nella equazione iniziale (3.1), notiamo che deve essere anche $a_k = 0$. □

Il seguente teorema è un rafforzamento del precedente.

Teorema 7.19. *Dato un endomorfismo lineare* $T : V \to V$*, siano* $\lambda_1, \lambda_2, \dots, \lambda_k$ *degli autovalori di* T *distinti fra loro. Allora gli autospazi* $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_k}$ *sono in somma diretta.*

Dimostrazione. Ricordiamo (vedi Paragrafo 3 del Capitolo 4) che dire che gli autospazi $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_k}$ sono in somma diretta vuol dire che se ne prendiamo uno qualunque, diciamo V_{λ_1} tanto per fissare la notazione, la sua intersezione con la somma di tutti gli altri è banale, ossia

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

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

$$w = v_2 + v_3 + \cdots + v_k$$

(dove i $v_j \in V_{\lambda_j}$ per ogni j) visto che $w \in V_{\lambda_2} + \cdots + V_{\lambda_k}$. Alcuni dei vettori v_j potrebbero essere uguali a O, ma non tutti (se fossero tutti uguali a O allora anche w sarebbe uguale a O). Diciamo che, a meno di riordinamenti, i vettori diversi da zero siano $v_2, ..., v_s$, con $s \leq k$.

Dunque vale:

$$v_1 = w = v_2 + v_3 + \cdots + v_s$$

Da questa uguaglianza ricaviamo la relazione

$$v_1 - v_2 - v_3 - \cdots - v_s = O$$

che rivela che gli autovettori $\{v_1, ...v_s\}$ sono linearmente dipendenti, assurdo perché contraddice il Teorema 7.17 (si tratta di autovettori relativi ad autovalori distinti). □

Nelle ipotesi del teorema precedente sappiamo allora (vedi Esercizio 4.12), che la dimensione della somma degli autospazi è "la massima possibile", ossia

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

Osserviamo che abbiamo dunque già ottenuto 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

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

e quindi è possibile trovare una base "buona", fatta da autovettori di T, insomma T è diagonalizzabile.
Per scrivere una simile base "buona", come sappiamo dal Paragrafo 3 del Capitolo 4, basta scegliere una base per ogni V_{λ_i} e poi fare l'unione.

Altrimenti, se

$$dim\ V_{\lambda_1} + dim\ V_{\lambda_2} + \cdots dim\ V_{\lambda_k} < n = dim\ V \tag{3.2}$$

T non è diagonalizzabile. Infatti non è possibile trovare una base di autovettori: se la trovassimo contraddiremmo la (3.2).

3.4. 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, $\ldots$,$dim\ V_{\lambda_k}$ coefficienti uguali a λ_k.

Il rango di T sarà uguale al numero dei coefficienti non zero che troviamo sulla diagonale di $[T]$, la dimensione del nucleo sarà uguale al numero dei coefficienti uguali a zero che troviamo sulla diagonale di $[T]$.

4. Il criterio della molteplicità algebrica e della molteplicità geometrica

Nel paragrafo precedente abbiamo descritto un algoritmo che ci permette di decidere se un endomorfismo è diagonalizzabile o no. In questo paragrafo faremo una osservazione che ci permetterà di riformulare questo algoritmo e di individuare alcune "scorciatoie".

Consideriamo come al solito un endomorfismo $T : V \to V$, dove V è uno spazio vettoriale sul campo $\mathbb{K}$ con $n = dim\ V$.

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

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

dove $\lambda_1, \lambda_2, \ldots, \lambda_k$ sono gli autovalori di T in $\mathbb{K}$ e sono tutti distinti fra loro, e $f(t)$ o è 1 oppure è un polinomio irriducibile in $\mathbb{K}[t]$ 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 $\lambda_1, \lambda_2, \ldots, \lambda_k$. Più esattamente, per ogni $i = 1, 2, \ldots, k$, λ_i compare $dim\ V_{\lambda_i}$ volte. Dunque in questo caso possiamo ricalcolare il polinomio caratteristico P_T usando $[T]_{b}^{b}$:

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

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

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

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

In sintesi:

Proposizione 7.20. *Se l'endomorfismo T è diagonalizzabile sul campo $\mathbb{K}$, allora il suo polinomio caratteristico $P_T(t)$ si fattorizza come prodotto di fattori lineari in $\mathbb{K}[t]$:*

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

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 $\mathbb{K}[t]$ allora T è diagonalizzabile ? NO, in generale non è vero. Basta considerare per esempio l'applicazione lineare $L : \mathbb{R}^2 \to \mathbb{R}^2$ 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 7.21. Data T come sopra con polinomio caratteristico

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

dove $f(t)$ o è 1 oppure è un polinomio irriducibile in $\mathbb{K}[t]$ di grado > 1, diremo che, per ogni $i = 1, 2, \ldots, k$, 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 7.22. *Dati* $T : V \to V$ *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

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

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 (Teorema 2.19), scegliere degli elementi $s_1, \ldots, s_r$ tali che $b = b' \cup \{s_1, \ldots, s_r\}$ sia una base. Rispetto a questa base la matrice di T ha la seguente forma:

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

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 $T(s_1), T(s_2), \ldots, T(s_r)$ 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:

$P_T(t) = Det\ (tI - [T]^{b}_{b}) = (t-\lambda_1)^{dim\ V_{\lambda_1}}(t-\lambda_2)^{dim\ V_{\lambda_2}} \cdots (t-\lambda_k)^{dim\ V_{\lambda_k}} Det\ M$

dove M è il minore $r \times r$ che sta nell'angolo in basso a destra di $tI - [T]^{b}_{b}$.

Ricordiamo ora la fattorizzazione in irriducibili per P_T

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

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 $Det\ M$, possiamo dire che, per ogni $i = 1, 2, \ldots, k$, $dim\ V_{\lambda_i} \le a_i$. □

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

Teorema 7.23 (Criterio delle molteplicità algebrica e geometrica.)**.** *Dato un endomorfismo lineare $T : V \to V$ di uno spazio vettoriale V (di dimensione finita n) sul campo $\mathbb{K}$, siano $\lambda_1, \lambda_2, \ldots, \lambda_k$ gli autovalori (distinti fra loro) di T in $\mathbb{K}$. 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

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

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. □

Esercizio 7.24. Dato un endomorfismo lineare $T : V \to V$ di uno spazio vettoriale V sul campo $\mathbb{K}$, dimostrare che se un autovalore di T ha molteplicità algebrica uguale a 1 allora anche la sua molteplicità geometrica è uguale a 1.

5. Esempi

Esempio 7.25. Sia $T : \mathbb{R}^3 \to \mathbb{R}^3$ un endomorfismo lineare la cui matrice rispetto alla base standard è:

$$[T] = \begin{pmatrix} 0 & 3 & 0 \\ 1 & -2 & 0 \\ 1 & -3 & 1 \end{pmatrix}$$

Vogliamo capire se è diagonalizzabile o no, e, se lo è, vogliamo trovare una base composta da autovettori. Innanzitutto calcoliamo il polinomio caratteristico:

$$P_T(t) = det \left(tI - \begin{pmatrix} 0 & 3 & 0 \\ 1 & -2 & 0 \\ 1 & -3 & 1 \end{pmatrix} \right) = det \begin{pmatrix} t & -3 & 0 \\ -1 & t+2 & 0 \\ -1 & 3 & t-1 \end{pmatrix} = (t-1)^2(t+3)$$

Gli autovalori sono dunque 1 e -3. La molteplicità algebrica di -3 è uguale a 1 e coincide con la sua molteplicità geometrica. Ripetiamo infatti in questo caso particolare il ragionamento che alcuni lettori avranno già utilizzato per risolvere l'Esercizio 7.24: infatti la molteplicità geometrica di -3 è ≥ 1 (visto che -3 è autovalore[1]), e per la Proposizione 7.22 deve essere minore o uguale alla molteplicità algebrica, quindi è esattamente 1 e coincide con la molteplicità algebrica.

Dunque, volendo applicare il criterio del Teorema 7.23, dobbiamo studiare l'autovalore 1, che ha molteplicità algebrica 2, e controllare se la sua molteplicità geometrica è uguale a 2 o no. La molteplicità geometrica di 1 è la dimensione dell'autospazio $V_1 = Ker\ (1I - T)$, dunque dobbiamo calcolare la dimensione di:

$$Ker \left(1I - \begin{pmatrix} 0 & 3 & 0 \\ 1 & -2 & 0 \\ 1 & -3 & 1 \end{pmatrix} \right) = Ker \begin{pmatrix} 1 & -3 & 0 \\ -1 & 3 & 0 \\ -1 & 3 & 0 \end{pmatrix}$$

Si osserva subito che la matrice ha rango uguale a 1, di conseguenza la dimensione del Ker è uguale a 2. Anche per quel che riguarda l'autovalore 1 la molteplicità geometrica risulta uguale alla molteplicità algebrica, dunque l'applicazione T è diagonalizzabile.

Per trovare una base formata da autovettori, dobbiamo scegliere una base di V_1 e una base di V_{-3} e fare l'unione. Cominciamo col trovare una base di V_1, ossia una base di

$$Ker\ (1I - T) = Ker \begin{pmatrix} 1 & -3 & 0 \\ -1 & 3 & 0 \\ -1 & 3 & 0 \end{pmatrix}$$

Risolvendo il sistema 'a occhio', si vede subito che una possibile base è data dai vettori $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$.

[1] La molteplicità geometrica di una autovalore è sempre ≥ 1, visto che, per definizione, l'autospazio relativo a tale autovalore non è banale.

Per trovare una base di V_{-3} che, come sappiamo, ha dimensione 1, basta individuare un vettore non nullo in

$$V_{-3} = Ker\left(-3I - \begin{pmatrix} 0 & 3 & 0 \\ 1 & -2 & 0 \\ 1 & -3 & 1 \end{pmatrix}\right) = Ker\begin{pmatrix} -3 & -3 & 0 \\ -1 & -1 & 0 \\ -1 & 3 & -4 \end{pmatrix}$$

Anche in questo caso il sistema associato si risolve immediatamente: è facile osservare che il vettore $\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$ costituisce una base di V_{-3}.

Dunque una base che diagonalizza l'endomorfismo T è data dai vettori

$$v_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

La matrice di T rispetto a tale base è data da

$$[T]\begin{smallmatrix} v_1, v_2, v_3 \\ v_1, v_2, v_3 \end{smallmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix}$$

Esempio 7.26. Si consideri l'applicazione lineare $F_a : \mathbb{R}^3 \to \mathbb{R}^3$ che, rispetto alla base standard, ha matrice:

$$[F_a] = \begin{pmatrix} a & 0 & 0 \\ 1 & a & 1 \\ 0 & -1 & 2 \end{pmatrix}$$

Vogliamo studiare, al variare del parametro $a \in \mathbb{R}$, la diagonalizzabilità di F_a. Per prima cosa calcoliamo il polinomio caratteristico $P_{F_a}(t)$:

$$P_{F_a}(t) = det\left(tI - \begin{pmatrix} a & 0 & 0 \\ 1 & a & 1 \\ 0 & -1 & 2 \end{pmatrix}\right) = det\begin{pmatrix} t-a & 0 & 0 \\ -1 & t-a & -1 \\ 0 & 1 & t-2 \end{pmatrix} =$$

$$= (t-a)(t^2 - (a+2)t + 2a + 1)$$

Ora osserviamo che il polinomio $t^2 - (a+2)t + 2a + 1$ ha radici

$$\frac{a + 2 \pm \sqrt{a^2 - 4a}}{2}$$

Tali radici sono reali se e solo se $a^2 \geq 4a$ ovvero se e solo se $a \geq 4$ oppure $a \leq 0$. Visto che il campo in cui stiamo cercando gli autovalori è $\mathbb{R}$, per il Teorema 7.23 possiamo intanto concludere che: *se $0 < a < 4$ l'endomorfismo F_a non è diagonalizzabile.*

Se invece $a \geq 4$ oppure $a \leq 0$, abbiamo tre autovalori reali:

$$\frac{a + 2 - \sqrt{a^2 - 4a}}{2}, \quad \frac{a + 2 + \sqrt{a^2 - 4a}}{2}, \quad a$$

e la prima cosa che ci conviene fare è calcolare le loro molteplicità algebriche, ossia capire se per qualche valore di a questi autovalori coincidono. Infatti, per i valori di a per cui questi tre autovalori sono a due a due distinti possiamo subito dire, in base al Teorema 7.23, che F_a è diagonalizzabile: gli autovalori hanno molteplicità algebrica uguale a 1, e dunque (vedi Esercizio 7.24), anche molteplicità geometrica uguale a 1.

Affrontiamo il problema della coincidenza studiando separatamente le tre possibili uguaglianze:

$$\frac{a+2-\sqrt{a^2-4a}}{2} = \frac{a+2+\sqrt{a^2-4a}}{2}$$

$$\frac{a+2-\sqrt{a^2-4a}}{2} = a$$

$$\frac{a+2+\sqrt{a^2-4a}}{2} = a$$

La prima di queste uguaglianze è vera se e solo se

$$-\sqrt{a^2-4a} = \sqrt{a^2-4a}$$

ovvero se e solo se $\sqrt{a^2-4a} = 0$ ovvero se e solo se $a = 0$ oppure $a = 4$. Invece si verifica subito che le uguaglianze

$$\frac{a+2-\sqrt{a^2-4a}}{2} = a$$

$$\frac{a+2+\sqrt{a^2-4a}}{2} = a$$

non sono mai vere, qualunque sia il valore di a.

Dunque i casi che richiedono attenzione sono solo $a = 0$ e $a = 4$; possiamo trarre una seconda conclusione: *se $a > 4$ oppure $a < 0$ l'endomorfismo F_a è diagonalizzabile.*

Studiamo infine i due casi rimasti: per quel che riguarda $a = 0$, gli autovalori di F_0 sono 0 e 1 e il polinomio caratteristico $P_{F_0}(t)$ è $t(t-1)^2$. Per capire se F_0 è diagonalizzabile bisogna calcolare la molteplicità geometrica di 1, ossia calcolare

$$dim\ V_1 = dim\ Ker\ \left(1I - \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix}\right) = dim\ Ker\ \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix}$$

La matrice ha rango 2, dunque il Ker ha dimensione 1. La molteplicità geometrica dell'autovalore 1 è uguale a 1, mentre la molteplicità algebrica è uguale a 2: l'endomorfismo F_0 non è diagonalizzabile.

Per quel che riguarda $a = 4$, gli autovalori di F_4 sono 4 e 3 e il polinomio caratteristico $P_{F_4}(t)$ è $(t-4)(t-3)^2$. Per capire se F_4 è diagonalizzabile bisogna calcolare la molteplicità geometrica di 3, ossia calcolare

$$dim\ V_3 = dim\ Ker\ \left(3I - \begin{pmatrix} 4 & 0 & 0 \\ 1 & 4 & 1 \\ 0 & -1 & 2 \end{pmatrix}\right) = det\ \begin{pmatrix} -1 & 0 & 0 \\ -1 & -1 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$

come nel caso precedente, il Ker ha dimensione 1 e risulta che l'endomorfismo F_4 non è diagonalizzabile.

6. Altri esercizi

Esercizio 7.27. Sia $A : \mathbb{R}^4 \to \mathbb{R}^4$ l'applicazione lineare che nella base standard è rappresentatata dalla matrice

$$[A] = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 1 & 0 & 2 & 0 \end{pmatrix}$$

Dire se A è diagonalizzabile. Descrivere gli autovalori e gli autospazi di A.

Esercizio 7.28. Sia $F : \mathbb{C}^3 \to \mathbb{C}^3$ l'applicazione lineare definita, nella base standard di $\mathbb{C}^3$, dalla matrice:

$$[F] = \begin{pmatrix} -1 & 0 & 1 \\ 2 & 1 & 2i \\ 1 & 0 & 0 \end{pmatrix}$$

Dire se F è diagonalizzabile e, se lo è, trovare un base di autovettori [nota: ricordiamo che stiamo lavorando sul campo $\mathbb{C}$].

Esercizio 7.29. Consideriamo l'endomorfismo lineare L_a di $\mathbb{R}^3$ dipendente dal parametro reale a e definito da:

$$L_a(x, y, z) = (ax + y + z, x + ay + z, -x + y + az)$$

(1) Discutere la diagonalizzabilità di L_a al variare del parametro reale a.
(2) Determinare, se esiste, una base di $\mathbb{R}^3$ di autovettori per L_0.

Esercizio 7.30. Sia $T : \mathbb{R}^3 \to \mathbb{R}^3$ l'endomorfismo la cui matrice rispetto alla base standard è

$$\begin{pmatrix} 9 & 6 & 9 \\ 3 & 12 & 9 \\ 3 & 6 & 15 \end{pmatrix}$$

Dire se T è diagonalizzabile e, se lo è, trovare una base fatta da autovettori.

Esercizio 7.31. a) Si consideri l'endomorfismo $A : \mathbb{C}^2 \to \mathbb{C}^2$ a cui nella base standard è associata la matrice

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

Trovare gli autovettori di A e dire se A è diagonalizzabile.
b) Si consideri l'endomorfismo $B : \mathbb{R}^2 \to \mathbb{R}^2$ a cui nella base standard è associata la matrice

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

i) Per quali valori di $b \in \mathbb{R}$ è diagonalizzabile ?
ii) Sia k un intero positivo. Si trovino, in funzione di k e del parametro b, gli autovalori dell'endomorfismo B^k.

Esercizio 7.32. Consideriamo l'applicazione lineare $A_a : \mathbb{R}^4 \to \mathbb{R}^4$ definita rispetto alla base standard dalla seguente matrice:

$$[A_a] = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & a & 0 \\ 1 & 0 & 1 & -1 \end{pmatrix}$$

Dire se esistono, e in caso affermativo trovare quali, valori del parametro $a \in \mathbb{R}$ per cui A_a è diagonalizzabile. Determinare inoltre gli autovettori di A_{-1}.

Esercizio 7.33. Sia $F_a : \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione lineare la cui matrice associata rispetto alla base canonica è la seguente:

$$[F_a] = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & 1 & a+1 \end{pmatrix}$$

(1) Determinare per quali valori del parametro a la matrice $[F_a$ è invertibile.
(2) Trovare i valori di a per i quali F_a è diagonalizzabile.
(3) Trovare, se esiste, una base di autovettori di F_a quando $a = 1/2$.

Esercizio 7.34. Sia $T_a : \mathbb{R}^3 \to \mathbb{R}^3$ l'endomorfismo lineare che, rispetto alla base standard di $\mathbb{R}^3$, è rappresentato dalla seguente matrice:

$$\begin{pmatrix} 1 & 2-a & 1 \\ 0 & a & 0 \\ -1 & 7a & a \end{pmatrix}$$

Per quali valori di $a \in \mathbb{R}$ l'endomorfismo T_a è diagonalizzabile ?

Esercizio 7.35. Sia $T_a : \mathbb{R}^3 \to \mathbb{R}^3$ l'endomorfismo lineare che, rispetto alla base standard di $\mathbb{R}^3$, è rappresentato dalla seguente matrice:

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

a) Per quali valori di $a \in \mathbb{R}$ l'endomorfismo T_a è diagonalizzabile ?

b) Trovare, per ogni a per cui T_a è diagonalizzabile, una base di $\mathbb{R}^3$ costituita da autovettori di T_a.

Esercizio 7.36. Sia $T_a : \mathbb{R}^3 \to \mathbb{R}^3$ l'endomorfismo lineare che, rispetto alla base standard di $\mathbb{R}^3$, è rappresentato dalla seguente matrice:

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

Per quali valori di $a \in \mathbb{R}$ l'endomorfismo T_a è diagonalizzabile ?

Esercizio 7.37. Sia $T_k : \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione lineare che, rispetto alla base standard di $\mathbb{R}^3$, è rappresentata dalla seguente matrice:

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

a) Per quali valori di $k \in \mathbb{R}$ T_k è diagonalizzabile ?

b) Nei casi in cui T_k è diagonalizzabile, trovare una base fatta da autovettori.

Esercizio 7.38. Sia $T_a : \mathbb{R}^4 \to \mathbb{R}^4$ l'endomorfismo lineare che, rispetto alla base standard di $\mathbb{R}^4$, è rappresentato dalla seguente matrice:

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

a) Per quali valori di $a \in \mathbb{R}$ l'endomorfismo T_a è diagonalizzabile ?

b) Trovare, per ogni a per cui T_a è diagonalizzabile, una base di $\mathbb{R}^4$ costituita da autovettori di T_a.

Esercizio 7.39. Sia V uno spazio vettoriale di dimensione finita n sul campo $\mathbb{K}$ e sia $T : V \to V$ un endomorfismo lineare. Dimostrare che esiste in $\mathbb{K}[t]$ un polinomio

$$f(t) = a_{n^2} t^{n^2} + \cdots + a_1 t + a_0$$

di grado minore o uguale a n^2 tale che

$$f(T) = a_{n^2} T^{n^2} + \cdots + a_1 T + a_0 I$$

è l'endomorfismo nullo.

Esercizio 7.40 (Teorema di Cayley-Hamilton)**.** Sia V uno spazio vettoriale di dimensione finita n sul campo $\mathbb{K}$ e sia $T : V \to V$ un endomorfismo lineare. Il teorema di Cayley-Hamilton afferma che l'endomorfismo $P_T(T)$ è l'endomorfismo nullo. Dimostrare questo teorema nel caso $n = 2$ e $n = 3$.

Esercizio 7.41. Dimostrare, come sopra, il teorema di Cayley-Hamilton nel caso $n = 2$ e $n = 3$ supponendo di sapere in più che l'endomorfismo T ammette un autovalore. La dimostrazione si semplifica?

Esercizio 7.42. Un endomorfismo lineare $T : V \to V$ si dice *nilpotente* se per un certo intero positivo n vale che $T^n = T \circ T \circ T \cdots \circ T$ è l'endomorfismo nullo.[2] Dimostrare che se T è nilpotente allora ha un unico autovalore: $\lambda = 0$.

Esercizio 7.43. Nel caso in cui V sia uno spazio vettoriale sul campo $\mathbb{C}$ dimostrare il viceversa dell'enunciato dell'esercizio precedente, ossia che se T ha un unico autovalore, uguale a 0, allora T è nilpotente. Se il campo è $\mathbb{R}$ e si sa che T ha un unico autovalore reale, uguale a 0, allora si può concludere che T è nilpotente?

Esercizio 7.44. Sia V uno spazio vettoriale di dimensione 4 sul campo $\mathbb{K}$ e sia $T : V \to V$ un endomorfismo lineare che, rispetto ad una certa base, è rappresentato dalla matrice:

$$[T] = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

L'endomorfismo T è nilpotente? È diagonalizzabile?

Esercizio 7.45. Sia $T : V \to V$ un endomorfismo lineare e sia λ un autovalore. Dimostrare che, per un ogni intero positivo n, λ^n è un autovalore di T^n.

[2] Analogamente, una matrice quadrata A si dice *nilpotente* se per un certo intero positivo n vale che A^n è la matrice nulla.

Esercizio 7.46. Sia V uno spazio vettoriale di dimensione finita sul campo $\mathbb{K}$ e sia $T : V \to V$ un endomorfismo lineare diverso da I e da $-I$. Supponiamo che valga $T^2 = I$. Individuare gli autovalori di T e dimostrare che T è diagonalizzabile.

Esercizio 7.47 (Proiezione lineare su un sottospazio)**.** Sia V uno spazio vettoriale di dimensione finita sul campo $\mathbb{K}$ e sia $T : V \to V$ un endomorfismo lineare diverso da I e dall'endomorfismo nullo. Supponiamo che valga $T^2 = T$. Dimostrare che T è diagonalizzabile e ha due autovalori, 1 e 0. Osservare che questo equivale a dire che T è una *proiezione lineare* di V su V_1: T manda V surgettivamente su V_1 e lascia fissi tutti i vettori di V_1. Sia $v_1, v_2, \ldots, v_n$ una base che diagonalizza T, con $V_1 =< v_1, v_2, \ldots, v_i >$ e $V_0 =< v_{i+1}, \ldots, v_n >$: scrivendo i vettori rispetto a questa base, la T è l'applicazione tale che

$$[T]\begin{pmatrix} a_1 \\ .. \\ a_i \\ a_{i+1} \\ .. \\ a_n \end{pmatrix} = \begin{pmatrix} a_1 \\ .. \\ a_i \\ 0 \\ .. \\ 0 \end{pmatrix}$$

Esercizio 7.48 (Diagonalizzazione simultanea di endomorfismi che commutano)**.** Sia V uno spazio vettoriale di dimensione finita n sul campo $\mathbb{K}$ e siano T ed S due endomorfismo lineari diagonalizzabili. Dimostrare che, se vale

$$T \circ S = S \circ T$$

allora esiste una base di V che diagonalizza S e T *simultaneamente*.

Suggerimento. Cominciare con l'osservare che, se λ è un autovalore per S e V_λ è il suo autospazio, allora $T(V_\lambda) \subseteq V_\lambda$.

Esercizio 7.49. Trovare, se possibile, una base di $\mathbb{R}^2$ che diagonalizza simultaneamante gli endomorfismi T ed S che, nella base standard, sono rappresentati rispettivamente dalle matrici:

$$[T] = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix} \qquad [S] = \begin{pmatrix} 5 & 2 \\ 6 & 1 \end{pmatrix}$$

CAPITOLO 8

Prodotti scalari e spazi euclidei

1. Prodotto scalare

Definizione 8.1. Sia V uno spazio vettoriale sul campo $\mathbb{K}$, dove $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$. Un *prodotto scalare* è una funzione che ad ogni coppia di vettori u, v appartenenti a V associa lo scalare $\langle u, v\rangle \in \mathbb{K}$, con le seguenti proprietà:

(1) $\langle au_1 + bu_2, v\rangle = a\langle u_1, v\rangle + b\langle u_2, v\rangle$, per ogni $u_1, u_2, v \in V$ e per ogni $a, b \in \mathbb{K}$;
(2) $\langle u, v\rangle = \overline{\langle v, u\rangle}$ per ogni $u, v \in V$ (dove $\overline{\langle v, u\rangle}$ indica il numero complesso coniugato di $\langle v, u\rangle$);
(3) per ogni $u \in V$ vale $\langle u, u\rangle \geq 0$ e $\langle u, u\rangle = 0$ vale se e solo se $u = O$.

Uno spazio vettoriale reale V munito di un prodotto scalare si dice *spazio euclideo*.

Osserviamo che, per la proprietà (2), il prodotto scalare $\langle u, u\rangle$ è sempre un numero reale, dunque ha senso la disuguaglianza che compare nella proprietà (3).

Per ogni vettore $u \in V$, scriveremo

$$||u|| = \sqrt{\langle u, u\rangle}$$

e diremo che $||u||$ è la *norma* di u.

Osservazione 8.2. Dalle proprietà (1) e (2) segue che

$$\langle u, av_1 + bv_2\rangle = \overline{a}\langle u, v_1\rangle + \overline{b}\langle u, v_2\rangle$$

Se il campo $\mathbb{K}$ è $\mathbb{R}$ ovviamente tutti i simboli di coniugazione complessa che compaiono nelle formule precedenti possono essere ignorati.

L'esempio principale di prodotto scalare è dato dal prodotto scalare standard in $\mathbb{R}^n$. Dati due vettori u, v, rappresentati, rispetto alla base standard, dalle colonne $u = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ \dots \\ a_n \end{pmatrix}$ e $v = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ \dots \\ b_n \end{pmatrix}$ il prodotto scalare standard è il seguente:

$$\langle u, v\rangle = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

Osservazione 8.3. Il prodotto scalare standard estende a tutti gli spazi $\mathbb{R}^n$ un concetto che in $\mathbb{R}^2$ e $\mathbb{R}^3$ ci è già familiare. Potete facilmente verificare che in questi casi per esempio $||u||$ coincide con quella che avete chiamato lunghezza del vettore u e che $||v-u||$ coincide con la distanza fra i vettori u e v. Inoltre $\langle u, v\rangle = 0$ se e solo se u e v sono ortogonali fra loro. Potete anche già osservare in $\mathbb{R}^2$ che vale la seguente relazione:

$$\langle u, v\rangle = ||u||||v|| \cos\theta$$

dove θ è l'angolo compreso fra i vettori u e v. Dimostreremo più in generale questa relazione fra qualche paragrafo.

Ecco altri due esempi di prodotto scalare:

Esempio 8.4. In $\mathbb{C}^n$, dati i vettori

$$u = \begin{pmatrix} a_1 \\ a_2 \\ ... \\ ... \\ a_n \end{pmatrix}, \quad v = \begin{pmatrix} b_1 \\ b_2 \\ ... \\ ... \\ b_n \end{pmatrix}$$

scritti rispetto alla base standard, abbiamo il prodotto scalare

$$\langle u, v \rangle = a_1 \overline{b_1} + a_2 \overline{b_2} + \cdots + a_n \overline{b_n}$$

Esempio 8.5. Nello spazio vettoriale delle funzioni continue reali nell'intervallo $[a, b]$,[1] abbiamo il seguente prodotto scalare:

$$\langle f, g \rangle = \int_a^b f(t)g(t)dt$$

2. Ortogonalità

Sia V uno spazio vettoriale munito di un prodotto scalare. Estendiamo il concetto di perpendicolarità che ci è noto in $\mathbb{R}^2$ e $\mathbb{R}^3$ dando la seguente definizione:

Definizione 8.6. Due vettori $u, v \in V$ sono detti *ortogonali* se

$$\langle u, v \rangle = 0$$

Notiamo che se due vettori verificano $\langle u, v \rangle = 0$ vale anche che $\langle v, u \rangle = \overline{\langle u, v \rangle} = 0$ dunque il concetto di ortogonalià non dipende dall'ordine con cui stiamo considerando i due vettori. Inoltre dalla definizione si ricava subito che il vettore O è ortogonale a tutti i vettori di V:

$$\langle O, v \rangle = \langle 0O, v \rangle = 0\langle O, v \rangle = 0$$

Sia U un sottospazio di V. L'insieme dato dai vettori di V che sono ortogonali ad ogni vettore di U si indica con $U^\perp$:

$$U^\perp = \{v \in V \mid \langle v, u \rangle = 0 \; \forall \, u \in U\}$$

Si verifica facilmente che $U^\perp$ è un sottospazio vettoriale di V, chiamato il *sottospazio ortogonale* a U.

Esercizio 8.7. Fare questa verifica.

Definizione 8.8. Un insieme $\{u_1, ..., u_r\}$ di vettori si dice *ortogonale* se i suoi elementi sono vettori a due a due ortogonali fra loro. L'insieme $\{u_1, ..., u_r\}$ si dice *ortonormale* se è ortogonale e se per ogni i vale $||u_i|| = 1$.

[1] Questo spazio vettoriale, a differenza di tutti gli altri spazi considerati in queste dispense, non ha una base finita; si tratta dunque di uno spazio vettoriale di dimensione infinita.

Esempio 8.9. Nello spazio euclideo $\mathbb{R}^n$ con il prodotto scalare standard, i vettori della base standard ($u_1 = \begin{pmatrix} 1 \\ 0 \\ ... \\ ... \\ 0 \end{pmatrix}$ etc...) costituiscono un insieme ortonormale.

Teorema 8.10. *Sia V uno spazio vettoriale munito di un prodotto scalare. Un insieme ortonormale di vettori $\{u_1, ..., u_r\}$ è linearmente indipendente e, per ogni $v \in V$, il vettore*

$$w = v - \langle v, u_1 \rangle u_1 - \langle v, u_2 \rangle u_2 - \cdots - \langle v, u_r \rangle u_r$$

è ortogonale ad ognuno degli u_i, dunque appartiene a $Span(u_1, ..., u_r)^\perp$.

Dimostrazione. Supponiamo di avere una combinazione lineare degli u_i uguale a O:

$$a_1 u_1 + \cdots + a_r u_r = O$$

Per dimostrare la lineare indipendenza dei vettori $u_1, ..., u_r$ dobbiamo dimostrare che per ogni i vale $a_i = 0$.

Fissato un i, consideriamo il prodotto scalare di entrambi i membri dell'uguaglianza con u_i:

$$\langle a_1 u_1 + \cdots + a_r u_r, u_i \rangle = \langle O, u_i \rangle$$

ovvero

$$a_1 \langle u_1, u_i \rangle + \cdots + a_r \langle u_r, u_i \rangle = 0$$

Vista la ortonormalità dell'insieme $\{u_1, ..., u_r\}$ questo implica $a_i = 0$.

Ripetendo questa considerazione per ogni $i = 1, ..., r$ si dimostra che i vettori $u_1, ..., u_r$ sono linearmente indipendenti.

Verifichiamo adesso che w è ortogonale a u_i per ogni i:

$$\langle w, u_i \rangle = \langle v, u_i \rangle - \langle v, u_1 \rangle \langle u_1, u_i \rangle - \cdots - \langle v, u_r \rangle \langle u_r, u_i \rangle =$$

$$= \langle v, u_i \rangle - \langle v, u_1 \rangle 0 - \cdots - \langle v, u_i \rangle 1 - \cdots - \langle v, u_r \rangle 0 = 0$$

□

Il teorema appena dimostrato è l'ingrediente principale del procedimento di ortogonalizzazione di Gram-Schmidt, che permette di ottenere una base ortonormale "modificando" una base data.

Teorema 8.11 (Procedimento di ortogonalizzazione di Gram-Schmidt)**.** *Sia V uno spazio vettoriale munito di un prodotto scalare, e sia $v_1, ... v_n$ una qualunque base di V. Allora esiste una base $u_1, ..., u_n$ di V che è ortonormale e tale che, per ogni $i = 1, ..., n$, $u_i \in Span(v_1, ... v_i)$.*

Dimostrazione. Per prima cosa poniamo $u_1 = \frac{v_1}{||v_1||}$. L'insieme $\{u_1\}$ è ortonormale.

Come secondo passo consideriamo il vettore $v_2 - \langle v_2, u_1 \rangle u_1$: questo vettore è diverso da O per la lineare indipendenza di v_1, v_2 e per il Teorema 8.10, applicato riferendosi all'insieme ortonormale $\{u_1\}$, è ortogonale a u_1. Inoltre appartiene al sottospazio $Span(v_1, v_2)$.

È dunque un buon candidato per essere il nostro u_2; l'unico problema è che non ha norma 1. Per rimediare a questo basta moltiplicarlo per un opportuno scalare. Poniamo dunque :

$$u_2 = \frac{v_2 - \langle v_2, u_1 \rangle u_1}{||v_2 - \langle v_2, u_1 \rangle u_1||}$$

Per le osservazioni precedenti, l'insieme $\{u_1, u_2\}$ è ortonormale.

Facciamo un ulteriore passo, per illustrare meglio il procedimento. A questo punto consideriamo il vettore $v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$. Questo vettore è diverso da O (altrimenti v_3 apparterrebbe a $Span(v_1, v_2)$ contro l'ipotesi che $v_1, v_2, ..., v_n$ è una base) e per il Teorema 8.10, applicato riferendosi all'insieme ortonormale $\{u_1, u_2\}$, è ortogonale a $Span(u_1, u_2)$.

Inoltre, visto che u_1 e u_2 appartengono a $Span(v_1, v_2)$, il vettore $v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$ appartiene a $Span(v_1, v_2, v_3)$ e ha dunque tutte le proprietà da noi richieste, eccetto quella di avere norma 1.

Come sopra, poniamo dunque:

$$u_3 = \frac{v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2}{||v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2||}$$

Una semplice dimostrazione per induzione ci mostra che possiamo proseguire questo procedimento ricorsivo fino a compiere n passi e definire u_n come

$$u_n = \frac{v_n - \langle v_n, u_1 \rangle u_1 - \langle v_n, u_2 \rangle u_2 - \cdots - \langle v_n, u_{n-1} \rangle u_{n-1}}{||v_n - \langle v_n, u_1 \rangle u_1 - \langle v_n, u_2 \rangle u_2 - \cdots - \langle v_n, u_{n-1} \rangle u_{n-1}||}$$

L'insieme $\{u_1, ..., u_n\}$ sarà ortonormale e tale che, per ogni i, $u_i \in Span(v_1, ..., v_i)$.

In particolare, i vettori $u_1, ..., u_n$ sono una base di V, visto che per il Teorema 8.10 sono linearmente indipendenti e $n = dim\ V$.

□

Il teorema appena dimostrato dimostra in particolare che ogni spazio V munito di un prodotto scalare ha una base ortonormale.

Esercizio 8.12. Applicare il procedimento di ortogonalizzazione alla seguente base di $\mathbb{R}^2$:

$$v_1 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

Esercizio 8.13. Trovare una base ortonormale di $\mathbb{R}^3$ il cui primo vettore sia $(\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$.

Osservazione 8.14 (Angolo fra due vettori in $\mathbb{R}^n$)**.** Consideriamo $V = \mathbb{R}^n$ munito del prodotto scalare standard. Siano u e v due vettori linearmente indipendenti di $\mathbb{R}^n$. Come sappiamo dal Teorema 8.10, applicato partendo dall'insieme ortonormale di vettori dato da un solo vettore, ossia $\{\frac{u}{||u||}\}$, il vettore

$$w = v - \langle v, \frac{u}{||u||} \rangle \frac{u}{||u||}$$

è ortogonale a $\frac{u}{||u||}$. Possiamo riscrivere questa relazione così

$$v = w + \langle v, \frac{u}{||u||} \rangle \frac{u}{||u||}$$

e ci accorgiamo che ci indica che nel piano $Span(u, v)$ i vettori $O, z = \langle v, \frac{u}{||u||}\rangle \frac{u}{||u||}, v$ determinano un triangolo rettangolo, con l'ipotenusa di lunghezza $||v||$ e i due cateti di lunghezza rispettivamente $||w||$ e $|\langle v, \frac{u}{||u||}\rangle| = \frac{|\langle v,u\rangle|}{||u||}$ (vedi Figura 1).

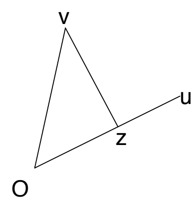


FIGURA 1. Dati due vettori $u, v \in \mathbb{R}^n$, i vettori $O, z = \langle v, \frac{u}{||u||}\rangle \frac{u}{||u||}, v$ determinano un triangolo rettangolo, visto che $w = v - \langle v, \frac{u}{||u||}\rangle \frac{u}{||u||}$ è ortogonale alla retta generata da u.

Chiamiamo θ l'angolo compreso fra l'ipotenusa e il cateto di lunghezza $\frac{|\langle v,u\rangle|}{||u||}$, ovvero l'angolo compreso fra u e v (andrebbe indicato quale dei due angoli stiamo considerando, ma siccome ne prenderemo il coseno non occorre in realtà precisare). Come sappiamo dalla geometria elementare vale

$$||v|| \cos\theta = \frac{|\langle v, u\rangle|}{||u||}$$

da cui ricaviamo la relazione

$$||u||||v|| \cos\theta = |\langle v, u\rangle|$$

a cui avevamo già accennato nella Osservazione 8.3. Come potete facilmente verificare, tale relazione vale anche se u e v non sono linearmente indipendenti, dunque vale per qualsiasi coppia di vettori in $\mathbb{R}^n$.

Introducendo il prodotto scalare standard, abbiamo dunque "recuperato" la trigonometria all'interno dell'algebra lineare.

3. La disuguaglianza di Cauchy-Schwarz

Dato un prodotto scalare $\langle\ ,\ \rangle$ su V, abbiamo il seguente teorema:

Teorema 8.15 (Disuguaglianza di Cauchy-Schwarz)**.** *Per ogni coppia di vettori $u, v \in V$ vale*

$$|\langle u, v\rangle| \leq ||u||\ ||v||$$

DIMOSTRAZIONE. Cominciamo osservando che se $v = O$, la disuguaglianza diventa $0 \leq 0$ dunque è vera.

Supponiamo dunque $v \neq O$, e calcoliamo, usando le proprietà del prodotto scalare, il quadrato della norma del vettore $u - \langle u, v\rangle tv$, dove $t \in \mathbb{R}$.

$$||u - \langle u, v\rangle tv||^2 = \langle u - \langle u, v\rangle tv, u - \langle u, v\rangle tv\rangle =$$
$$= \langle u, u\rangle - \overline{\langle u, v\rangle} t\langle u, v\rangle - \langle u, v\rangle t\langle v, u\rangle + \overline{\langle u, v\rangle}\langle u, v\rangle t^2\langle v, v\rangle =$$
$$= ||u||^2 - 2t(|\langle u, v\rangle|)^2 + (|\langle u, v\rangle|)^2 t^2 ||v||^2$$

Osserviamo che $||u - \langle u, v\rangle tv||^2 \geq 0$ dunque possiamo scrivere

$$||u||^2 - 2t(|\langle u, v\rangle|)^2 + (|\langle u, v\rangle|)^2 t^2 ||v||^2 \geq 0$$

Questo vale per qualsiasi $t \in \mathbb{R}$. Scegliendo adesso $t = \frac{1}{||v||^2}$ ricaviamo:

$$||u||^2 - \frac{(|\langle u, v \rangle|)^2}{||v||^2} \geq 0$$

da cui si ottiene la tesi. □

Esempio 8.16. Nel caso in cui V sia $\mathbb{R}^n$ col prodotto scalare standard, la disuguaglianza di Cauchy-Schwarz si traduce così. Dati

$$u = \begin{pmatrix} a_1 \\ a_2 \\ ... \\ ... \\ a_n \end{pmatrix}, \quad v = \begin{pmatrix} b_1 \\ b_2 \\ ... \\ ... \\ b_n \end{pmatrix}$$

vale

$$|a_1 b_1 + a_2 b_2 + \cdots + a_n b_n| \leq (a_1^2 + \cdots + a_n^2)(b_1^2 + \cdots + b_n^2)$$

che può essere pensata anche, svincolandosi dai vettori e dagli spazi vettoriali, come una disuguaglianza che riguarda due qualsiasi n-uple $a_1, ..., a_n$ e $b_1, ..., b_n$ di numeri reali.

Esempio 8.17. Nel caso in cui V sia lo spazio delle funzioni continue reali definite sull'intervallo $[0, 1]$, col prodotto scalare descritto nell'Esempio 8.5, la disuguaglianza di Cauchy-Schwarz si traduce nella seguente importante disuguaglianza fra integrali:

$$\left(\int_0^1 f(t)g(t)dt \right)^2 \leq \int_0^1 f(t)^2 dt \int_0^1 g(t)^2 dt$$

4. Sottospazi ortogonali

Sia V, come nei paragrafi precedenti, uno spazio vettoriale munito di prodotto scalare.

Proposizione 8.18. *Sia U un sottospazio vettoriale di V. Allora esiste una base ortonormale di U che è un sottoinsieme di una base ortonormale di V.*

Dimostrazione. Partiamo da una qualunque base $v_1, ..., v_r$ di U, ed estendiamola ad una base $v_1, ..., v_n$ di V (usando il Teorema 2.19). Ora utilizziamo il procedimento di ortogonalizzazione di Gram-Schmidt alla base $v_1, ..., v_n$: otteniamo una base ortonormale $u_1, ..., u_n$ di V, con la proprietà che i vettori $u_1, ..., u_r$ appartengono a $Span(v_1, ..., v_r)$, ossia a U. Dunque $u_1, ..., u_r$ è una base ortonormale di U, che è un sottoinsieme di una base ortonormale di V, come volevamo.

□

Teorema 8.19. *Sia U un sottospazio vettoriale di V. Allora V si decompone come somma diretta di U e di $U^\perp$:*

$$V = U \oplus U^\perp$$

In particolare vale che $dim\ U^\perp = dim\ V - dim\ U$.

Dimostrazione. Osserviamo innanzitutto che $U \cap U^{\perp} = \{O\}$, dunque i due sottospazi sono in somma diretta.

Infatti se $v \in U \cap U^{\perp}$ allora $\langle v, v \rangle = 0$ (in altre parole, visto che $v \in U^{\perp}$, v deve essere ortogonale a tutti i vettori di U, fra cui se stesso). Ma per la proprietà (3) del prodotto scalare da $\langle v, v \rangle = 0$ si ricava $v = O$.

Non ci resta che dimostrare che $U + U^{\perp} = V$, ossia che ogni vettore di V può essere scritto come somma di un vettore di U e di un vettore di $U^{\perp}$. Poniamo $r = dim\ U$. Per la Proposizione 8.18 sappiamo che esiste una base ortonormale $u_1, ..., u_n$ di V tale che i primi r vettori $u_1, ..., u_r$ sono una base ortonormale di U.

Un qualunque vettore $v \in V$ può essere scritto come combinazione lineare dei vettori di questa base:

$$v = a_1 u_1 + \cdots a_r u_r + a_{r+1} u_{r+1} + \cdots + a_n u_n$$

Osserviamo a questo punto che $a_1 u_1 + \cdots a_r u_r \in U$ come osservato sopra, mentre $a_{r+1} u_{r+1} + \cdots + a_n u_n \in U^{\perp}$ per la ortonormalità della base $u_1, ..., u_n$, dunque la scrittura

$$v = (a_1 u_1 + \cdots a_r u_r) + (a_{r+1} u_{r+1} + \cdots + a_n u_n)$$

esprime v come somma di un vettore di U e di un vettore di $U^{\perp}$.

□

Esercizio 8.20. Sia V uno spazio vettoriale di dimensione finita munito di un prodotto scalare e sia U un sottospazio vettoriale di V. Dimostrare che $(U^{\perp})^{\perp} = U$.

Osservazione 8.21 (Ortogonalità e sistemi lineari)**.** Consideriamo $V = \mathbb{R}^n$ munito del prodotto scalare standard. Dato un sottospazio W, con base $w_1, ..., w_r$, dove $w_1 = \begin{pmatrix} a_1 \\ a_2 \\ ... \\ ... \\ a_n \end{pmatrix}$, $w_2 = \begin{pmatrix} b_1 \\ b_2 \\ ... \\ ... \\ b_n \end{pmatrix}$, etc... se pensiamo a come si scrive il prodotto scalare ci rendiamo conto i vettori di $W^{\perp}$ sono esattamente le soluzioni del sistema

$$\begin{pmatrix} a_1 & a_2 & ... & a_n \\ b_1 & b_2 & ... & b_n \\ ... & ... & ... & ... \\ ... & ... & ... & ... \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ ... \\ ... \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ ... \\ ... \\ 0 \end{pmatrix}$$

Esercizio 8.22. Dimostrare che, dato un sottospazio U di uno spazio V munito di un prodotto scalare, e data una base ortonormale $u_1, ..., u_n$ di V tale che i primi r vettori $u_1, ..., u_r$ sono una base ortonormale di U, la mappa $T : V \to V$ data da $T(v) = v - \langle v, u_1 \rangle u_1 - \langle v, u_2 \rangle u_2 - \cdots - \langle v, u_r \rangle u_r$ è lineare, ha per immagine $U^{\perp}$ e ha nucleo uguale a U. Dimostrare inoltre che $T^2 = T$ (dunque, in base all'Esercizio 7.47 possiamo dire che T è una proiezione di V su $U^{\perp}$; in particolare T viene chiamata la *proiezione ortogonale* di V su U).

5. Esercizi

Esercizio 8.23. Sia V lo spazio vettoriale delle funzioni continue reali nell'intervallo $[-\pi, \pi]$, munito del prodotto scalare:

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(t) g(t) dt$$

Verificare che il seguente sottoinsieme è ortogonale:

$$1, \cos t, cos2t, ..., \sin t, \sin 2t, ...$$

[Attenzione, come vedete si tratta di un insieme infinito: del resto lo spazio vettoriale che stiamo considerando è, a differenza di tutti gli altri spazi considerati in queste dispense, di dimensione infinita. L'insieme proposto è molto importante in matematica, in particolare nella teoria delle *serie di Fourier*.]

Esercizio 8.24. Dato uno spazio vettoriale V munito di un prodotto scalare, dimostrare che la norma soddisfa la seguente *disuguaglianza triangolare*:

$$||u + v|| \leq ||u|| + ||v||$$

Esercizio 8.25. Si consideri $\mathbb{R}^4$ munito del prodotto scalare standard. Dato il sottospazio W generato dai vettori $(1, -2, 3, 3)$ e $(3, -5, 4, 4)$, trovare una base del sottospazio $W^\perp$.

Esercizio 8.26. Si consideri $\mathbb{R}^4$ munito del prodotto scalare standard. Dato il sottospazio W generato dai vettori $(1, -2, 3, 3)$ e $(3, -5, 4, 4)$, trovare una base ortonormale di W usando il procedimento di ortogonalizzazione di Gram-Schmidt.

Esercizio 8.27. Si consideri $\mathbb{R}^4$ munito del prodotto scalare standard. Dato il sottospazio W generato dai vettori $(1, 1, 0, 0)$ e $(0, 1, 1, 0)$, $(0, 0, 1, 1)$, trovare una base ortonormale di W usando il procedimento di ortogonalizzazione di Gram-Schmidt.

Esercizio 8.28. Si consideri $\mathbb{C}^3$ munito del prodotto scalare descritto nell'Esempio 8.4. Sia W il sottospazio generato dai vettori $(1, \imath, 1)$ e $(1+\imath, -0, 1)$. Trovare una base ortonormale per W.

CAPITOLO 9

Qualche appunto sul teorema spettrale

1. Introduzione al teorema spettrale

Cominciamo con l'enunciare, senza dimostrazione, il seguente teorema.

Teorema 9.1. *Sia $T : V \to V$ un endomorfismo di uno spazio vettoriale V di dimensione finita sul campo $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$) e munito di prodotto scalare. Esiste allora un endomorfismo T^* univocamente determinato tale che*

$$\langle T(u), v\rangle = \langle u, T^*(v)\rangle$$

per ogni $u, v \in V$. Se si fissa in V una base ortonormale vale che la matrice $[T^*]$ di T rispetto a tale base è la trasposta della coniugata di $[T]$.

Sottolineiamo che per l'esistenza dell'endomorfismo T^* è fondamentale l'ipotesi che lo spazio vettoriale V abbia dimensione finita.

Non dimostreremo questo teorema, ma vogliamo verificarlo in un caso particolare. Consideriamo lo spazio vettoriale $\mathbb{C}^3$ con la sua base standard (che è ortonormale). Sia $T : \mathbb{C}^3 \to \mathbb{C}^3$ l'endomorfismo dato, rispetto alla base standard, dalla matrice

$$[T] = \begin{pmatrix} 2 & \imath & 3 \\ 0 & \imath & 1 \\ 1 & 2\imath & 3 \end{pmatrix}$$

Consideriamo due vettori di $\mathbb{C}^3$:

$$u = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, v = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

e calcoliamo $\langle T(u), v\rangle$. Osserviamo che:

$$T(u) = \begin{pmatrix} 2x_1 + \imath x_2 + 3x_3 \\ \imath x_2 + x_3 \\ x_1 + 2\imath x_2 + 3x_3 \end{pmatrix}$$

dunque

$$\langle T(u), v\rangle = (2x_1 + \imath x_2 + 3x_3)\overline{y}_1 + (\imath x_2 + x_3)\overline{y}_2 + (x_1 + 2\imath x_2 + 3x_3)\overline{y}_3$$

Consideriamo ora la matrice associata a T^* nella base standard. Per il Teorema 9.1 sappiamo che è uguale alla trasposta coniugata:

$$[T^*] = \begin{pmatrix} 2 & 0 & 1 \\ -\imath & -\imath & -2\imath \\ 3 & 1 & 3 \end{pmatrix}$$

Calcoliamo $\langle u, T^*(v)\rangle$. Osserviamo che:

$$T^*(v) = \begin{pmatrix} 2y_1 + y_3 \\ -\imath y_1 - \imath y_2 - 2\imath y_3 \\ 3y_1 + y_2 + 3y_3 \end{pmatrix}$$

e quindi

$$\langle u, T^*(v)\rangle = x_1\overline{(2y_1 + y_3)} + x_2\overline{(-\imath y_1 - \imath y_2 - 2\imath y_3)} + x_3\overline{(3y_1 + y_2 + 3y_3)}$$

Svolgendo i calcoli si verifica facilmente che $\langle T(u), v\rangle = \langle u, T^*(v)\rangle$, in accordo col Teorema 9.1.

Definizione 9.2. L'endomorfismo T^* si dice *aggiunto* di T. Se vale che $T = T^*$ allora T si dice *autoaggiunto.*

Osserviamo che nel caso in cui $\mathbb{K} = \mathbb{R}$, se fissiamo una base ortonormale di V e consideriamo un endomorfismo T autoaggiunto, allora la matrice $[T]$ rispetto a tale base è simmetrica, ossia, indicando con $[T]^t$ la trasposta di $[T]$, vale

$$[T] = [T]^t$$

Viceversa, se un endomorfismo è rappresentato, rispetto ad una base ortonormale, da una matrice simmetrica, allora si mostra facilmente che è autoaggiunto. Per questo motivo nel caso in cui $\mathbb{K} = \mathbb{R}$ un endomorfismo autoaggiunto si dice anche *simmetrico.*

Teorema 9.3. *Sia $T : V \to V$ un endomorfismo autoaggiunto di uno spazio vettoriale V di dimensione finita sul campo $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$). Sia λ un autovalore di T. Allora $\lambda \in \mathbb{R}$.*

Dimostrazione. Sia λ un autovalore di T e sia $v \in V - \{O\}$ un autovettore relativo a λ. Dimostreremo che $\lambda\langle v, v\rangle = \overline{\lambda}\langle v, v\rangle$. Visto che $\langle v, v\rangle \neq 0$ questo implica che $\lambda = \overline{\lambda}$ ossia che $\lambda \in \mathbb{R}$.

Per dimostrare che $\lambda\langle v, v\rangle = \overline{\lambda}\langle v, v\rangle$ osserviamo che, per l'ipotesi che T è autoaggiunto, vale:

$$\langle T(v), v\rangle = \langle v, T(v)\rangle$$

Ora

$$\langle T(v), v\rangle = \langle \lambda v, v\rangle = \lambda\langle v, v\rangle$$

mentre

$$\langle v, T(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle$$

dunque $\lambda\langle v, v\rangle = \overline{\lambda}\langle v, v\rangle$ come volevamo. □

Studiamo adesso alcune proprietà degli endomorfismi simmetrici.

Teorema 9.4. *Sia $T : V \to V$ un endomorfismo simmetrico di uno spazio vettoriale V di dimensione finita su $\mathbb{K} = \mathbb{R}$. Allora sono vere le seguenti affermazioni:*

1. *Il polinomio caratteristico $p_T(t)$ ha tutte le radici reali e si fattorizza come prodotto di fattori lineari su $\mathbb{R}$.*
2. *L'endomorfismo T ha almeno un autovettore non nullo.*
3. *Siano $v_1, ..., v_r$ autovettori di T relativi ad autovalori distinti $\lambda_1, ..., \lambda_r$. Allora l'insieme di vettori $\{v_1, ..., v_r\}$ è ortogonale.*

Dimostrazione. Consideriamo il polinomio caratteristico $p_T(t)$. Per il teorema fondamentale dell'algebra tale polinomio si fattorizza in $\mathbb{C}[t]$ come prodotto di fattori lineari:

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

Le radici $\lambda_1, ..., \lambda_k$ del polinomio sono gli autovalori di T. Dato che T è simmetrico, per il Teorema 9.3 sappiamo che tali autovalori appartengono a $\mathbb{R}$. Dunque la fattorizzazione scritta sopra è in realtà una fattorizzazione di $p_T(t)$ in $\mathbb{R}[t]$. Questo dimostra il punto (1).

Inoltre dalla fattorizzazione di $p_T(t)$ deduciamo che T ha almeno un autovalore, dunque per dimostrare il punto (2) basta prendere un autovettore relativo a questo autovalore.

Per quel che riguarda il punto (3), sia v un autovettore relativo all'autovalore λ e sia w un autovettore relativo all'autovalore μ, con $\lambda, \mu \in \mathbb{R}$ e $\lambda \neq \mu$. Per l'ipotesi che T è simmetrico vale:

$$\langle T(v), w\rangle = \langle v, T(w)\rangle$$

Ora osserviamo che

$$\langle T(v), w\rangle = \langle \lambda v, w\rangle = \lambda\langle v, w\rangle$$

e anche

$$\langle v, T(w)\rangle = \langle v, \mu w\rangle = \mu\langle v, w\rangle$$

Dalle uguaglianze scritte sopra deduciamo che

$$\lambda\langle v, w\rangle = \mu\langle v, w\rangle$$

Se fosse $\langle v, w\rangle \neq 0$ questo sarebbe assurdo. Dunque vale $\langle v, w\rangle = 0$, ossia v e w sono ortogonali.

□

Vogliamo ora dimostrare il teorema spettrale, che afferma che un endomorfismo simmetrico è diagonalizzabile, e inoltre esiste una base ortonormale che lo diagonalizza.[1] Premettiamo il seguente lemma:

Lemma 9.5. *Sia V uno spazio vettoriale sul campo $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$, di dimensione finita e munito di prodotto scalare. Sia $T : V \to V$ un endomorfismo, e sia W un sottospazio di V invariante per T, ossia tale che $T(W) \subset W$. Allora $W^\perp$ è invariante per T^*.*

Dimostrazione. Consideriamo un elemento $v \in W^\perp$. Dobbiamo dimostrare che $T^*(v) \in W^\perp$. Ora, per ogni $w \in W$ vale che $T(w) \in W$, dunque possiamo scrivere:

$$0 = \langle T(w), v\rangle = \langle w, T^*(v)\rangle$$

dunque $T^*(v)$ è ortogonale a ogni $w \in W$. Questo dimostra che $T^*(v) \in W^\perp$.

□

Teorema 9.6 (Teorema spettrale, caso reale). *Sia $T : V \to V$ un endomorfismo simmetrico di uno spazio vettoriale V su $\mathbb{K} = \mathbb{R}$ di dimensione finita maggiore di 0. Allora esiste una base ortonormale di V i cui elementi sono autovettori di T.*

[1] L'aggettivo 'spettrale' proviene da alcune applicazioni, ed è collegato allo 'spettro' di una grandezza fisica. L'insieme degli autovalori è lo 'spettro' dell'endomorfismo.

Dimostrazione. Si dimostra per induzione sulla dimensione di V. Se la dimensione di V è uguale ad 1 si vede subito che l'enunciato è vero (basta prendere la base formata da un vettore non nullo di norma 1). Supponiamo dunque che $dim\ V = n > 1$ e che l'enunciato sia vero per ogni endomorfismo simmetrico di uno spazio vettoriale di dimensione $n-1$.

Per il Teorema 9.4 esiste un autovettore v_1 di T, che possiamo supporre di norma uguale a 1 (a meno di moltiplicarlo per scalare). Sia $W = Span(v_1)$ e consideriamo $W^\perp$. Come sappiamo dal Teorema 8.19 il sottospazio $W^\perp$ ha dimensione $n-1$. Inoltre sappiamo dal Lemma 9.5 che $W^\perp$ è invariante per T^*. Ma T^*, visto che T è simmetrico, coincide con T. Dunque $W^\perp$ è invariante per T e ha senso considerare l'endomorfismo T ristretto a $W^\perp$:

$$T_{|W^\perp} : W^\perp \to W^\perp$$

Osserviamo che $T_{|W^\perp}$ è simmetrico.[2] Per ipotesi induttiva sappiamo allora che esiste una base ortonormale $v_2, ..., v_n$ di $W^\perp$ composta di autovettori di $T_{|W^\perp}$. È immediato a questo punto verificare che $v_1, v_2, ..., v_n$ è una base ortonormale di V composta da autovettori di T.

□

Esempio 9.7. Consideriamo l'endomorfismo $T : \mathbb{R}^2 \to \mathbb{R}^2$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Osserviamo che l'endomorfismo T è simmetrico (infatti la sua matrice rispetto alla base standard, che è una base ortonormale, è simmetrica). Dunque per il Teorema Spettrale sappiamo che T è diagonalizzabile e ci aspettiamo di trovare un base ortonormale che lo diagonalizza. Verifichiamo. Il polinomio caratteristico è

$$p_T(t) = t^2 - 2t - 3 = (t-3)(t+1)$$

dunque gli autovalori sono 3 e -1, entrambi con molteplicità algebrica 1. Questo già ci conferma che T è diagonalizzabile. Troviamo l'autospazio V_3:

$$V_3 = Ker(T - 3I) = Ker \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix}$$

L'autospazio V_3 ha dunque dimensione 1 (come previsto, dato che la molteplicità algebrica dell'autovalore 3 è 1) ed è generato dal vettore $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Troviamo l'autospazio V_{-1}:

$$V_{-1} = Ker(T + I) = Ker \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$

[2] Per controllare in dettaglio questo fatto, osserviamo che $(T_{|W^\perp})^*$ è l'endomorfismo di $W^\perp$ che soddisfa

$$\langle T_{|W^\perp}(u_1), u_2 \rangle = \langle u_1, (T_{|W^\perp})^*(u_2) \rangle$$

per ogni $u_1, u_2 \in W^\perp$. Ma per le proprietà di T^* vale anche che

$$\langle T(u_1), u_2 \rangle = \langle u_1, T^*(u_2) \rangle$$

per ogni $u_1, u_2 \in W^\perp$, da cui, per l'unicità dell'aggiunto (vedi Teorema 9.1) segue che $(T^*)_{|W^\perp}$ e $(T_{|W^\perp})^*$ coincidono. Ma $T^* = T$ perché T è simmetrico, dunque $(T^*)_{|W^\perp} = T_{|W^\perp}$ e in conclusione $(T_{|W^\perp})^* = T_{|W^\perp}$ ossia $T_{|W^\perp}$ è simmetrico.

L'autospazio V_{-1} ha dunque dimensione 1 (come previsto) ed è generato dal vettore $w = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. I due autovettori v e w sono ortogonali (in accordo col Teorema 9.4, visto che sono due autovettori di un endomorfismo simmetrico relativi a due autovalori distinti). Concludiamo dunque che la base ortornormale data dai vettori

$$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

è una base di autovettori per T, in accordo col teorema spettrale.

Per completezza, diamo l'enunciato generale del teorema spettrale che vale anche nel caso $\mathbb{K} = \mathbb{C}$:

Teorema 9.8 (Teorema spettrale, enunciato valido anche su $\mathbb{C}$)**.** *Sia $T : V \to V$ un endomorfismo autoaggiunto di uno spazio vettoriale V di dimensione finita su $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$). Allora esiste una base ortonormale di V i cui elementi sono autovettori di T.*

2. Endomorfismi simmetrici definiti positivi o negativi

Come abbiamo visto nel precedente paragrafo, un endomorfismo autoaggiunto è diagonalizzabile e che tutti i suoi autovalori sono reali. Per alcune applicazioni che incontrerete nel secondo semestre del corso (Analisi II) è utile sapere che segno hanno gli autovalori.

Definizione 9.9. Sia $T : V \to V$ un endomorfismo autoaggiunto di uno spazio vettoriale V di dimensione finita su $\mathbb{K}$ (con $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$). Se tutti i suoi autovalori sono numeri reali ≥ 0 si dice che T è *semidefinito positivo*. Se tutti i suoi autovalori sono numeri reali ≤ 0 si dice che T è *semidefinito negativo*. In particolare, se tutti i suoi autovalori sono numeri reali > 0 si dice che T è *definito positivo*; se invece tutti i suoi autovalori sono numeri reali < 0 si dice che T è *definito negativo*.

La definizione appena data si applica in particolare nel caso reale agli endomorfismi simmetrici.

Consideriamo per esempio cosa accade nel caso in cui V ha dimensione 2. Ricordiamo innanzitutto che, dato un endomorfismo, sono ben definiti la sua traccia e il suo determinante (ossia non dipendono dalla base che si fissa e dalla particolare matrice che rappresenta l'endomorfismo in quella base: riguardate il Paragrafo 3 del Capitolo 5).

Proposizione 9.10. *Sia $T : \mathbb{R}^2 \to \mathbb{R}^2$ un endomorfismo simmetrico. Vale che:*

(1) *T è definito positivo se e solo se il suo determinante e la sua traccia sono entrambi > 0.*
(2) *T è definito negativo se e solo se il suo determinante è > 0 e la sua traccia è < 0.*

Dimostrazione. Come sappiamo dal teorema spettrale, T è diagonalizzabile. Siano λ_1, λ_2 i suoi autovalori (eventualmente coincidenti). Fissata una base di autovettori, la matrice che rappresenta T rispetto a tale base è:

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

da cui si ricava subito che il determinante di T è $\lambda_1\lambda_2$ e la traccia è $\lambda_1 + \lambda_2$. La dimostrazione a questo punto consiste in una analisi molto facile: per esempio il determinante è > 0 se solo se λ_1 e λ_2 sono diversi da 0 e hanno entrambi lo stesso segno. Se anche la traccia è > 0 allora λ_1 e λ_2 devono essere entrambi positivi, etc..(finite voi per esercizio).

$\square$

Esercizio 9.11. Sia $T : \mathbb{R}^2 \to \mathbb{R}^2$ un endomorfismo simmetrico. Dimostrate che se il suo determinante è $= 0$ e la sua traccia è ≥ 0 allora T è semidefinito positivo e che se il suo determinante è $= 0$ e la sua traccia è < 0 allora T è semidefinito negativo. Cosa si può dire di T se il suo determinante e la sua traccia sono entrambi uguali a 0?

3. Esercizi

Esercizio 9.12. Consideriamo l'endomorfismo $T : \mathbb{R}^2 \to \mathbb{R}^2$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$$

Trovare una base ortonormale che diagonalizza T.

Esercizio 9.13. Consideriamo l'endomorfismo $T : \mathbb{R}^2 \to \mathbb{R}^2$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Trovare una base ortonormale che diagonalizza T.

Esercizio 9.14. Consideriamo l'endomorfismo $T : \mathbb{R}^3 \to \mathbb{R}^3$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Trovare una base ortonormale che diagonalizza T.

Esercizio 9.15. Consideriamo l'endomorfismo $T : \mathbb{R}^3 \to \mathbb{R}^3$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Trovare una base ortonormale che diagonalizza T.

Esercizio 9.16. Consideriamo l'endomorfismo $T : \mathbb{R}^3 \to \mathbb{R}^3$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix}$$

Trovare una base ortonormale che diagonalizza T.

Esercizio 9.17. Dimostrare il Teorema 9.8.[Suggerimento: resta da dimostrare il caso $\mathbb{K} = \mathbb{C}$. Si può procedere per induzione. Si comincia prendendo un autovettore dell'endomorfismo T, e questa volta per dire che esiste basta osservare che un endomorfismo di uno spazio vettoriale complesso ha sempre un autovettore (infatti c'è sempre un autovalore perché il polinomio caratteristico ha sempre una radice complessa).]

Esercizio 9.18 (Esercizio di ripasso)**.** Verificate di saper dimostrare che, dato uno spazio vettoriale V di dimensione n su un qualunque campo $\mathbb{K}$, e dato un endomorfismo diagonalizzabile $T : V \to V$, il determinante di T è uguale al prodotto di tutti gli autovalori di T (ciascuno compare come fattore un numero di volte uguale alla sua molteplicità) e la traccia di T è uguale alla somma di tutti gli autovalori (ciascuno compare come addendo un numero di volte uguale alla sua molteplicità). [Suggerimento: prendete spunto dalla dimostrazione della Proposizione 9.10....]

Esercizio 9.19. Consideriamo l'endomorfismo $T : \mathbb{R}^2 \to \mathbb{R}^2$ la cui matrice, rispetto alla base standard è

$$[T] = \begin{pmatrix} 6 & 18 \\ -1 & 10 \end{pmatrix}$$

Scrivere il polinomio caratteristico di T senza calcolare $det(tI - [T])$ ma solo basandosi sulla traccia e sul determinante di T.

Bibliografia

[Ab] M. Abate, *Algebra Lineare*, McGraw-Hill.
[AlgGauss] http://marekrychlik.com/cgi-bin/gauss.cgi
[Pe] C. Petronio, *Geometria e Algebra Lineare*, Esculapio.