Appunti provvisori del secondo semestre del corso di Matematica Discreta

Pietro Di Martino e Giovanni Gaiffi, 29 maggio 2012
(rispetto alla versione dell'11 maggio sono state corrette alcune sviste da voi segnalate).

Indice

CAPITOLO 1

Spazi vettoriali e applicazioni lineari

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 a partire da un insieme V su cui si possa introdurre due operazioni: una tra due elementi di V sarà detta **somma vettoriale** o più semplicemente **somma**, e l'altra tra un elemento di V e uno di $\mathbb{K}$ sarà detta **moltiplicazione per scalare** o **prodotto esterno**[1]. 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[2]:

- $\forall\ u, v, w \in V$ 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).
- $\forall\ v \in V$ 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).

[1] 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 $\mathbb{K}$.

[2] 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.

- $\forall \lambda, \mu \in \mathbb{K}$, $\forall\ v \in V$ 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).

Osservazione 1.2. Osserviamo che affinché V sia uno spazio vettoriale su un campo $\mathbb{K}$ si richiede, in particolare, che V sia un gruppo commutativo rispetto alla somma.

Esercizio 1.3. 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:

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

Osservazione 1.4. È molto importante sottolineare la differenza tra l'elemento neutro della somma vettoriale O e lo 0 elemento neutro della somma di $\mathbb{K}$: in particolare 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 O.

Esercizio 1.5. 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 $-v$.

Esercizio 1.6. 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 1.3 segue che $0v = O$.

Esercizio 1.7. 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:

$$O \underbrace{=}_{esercizio\ 1.6} 0v = (-1 + 1)v \underbrace{=}_{\text{prop.dist.}} -1v + 1v \underbrace{=}_{\text{invariante moltiplicativo}} -1v + v$$

Ovvero $-1v$ è l'opposto per l'addizione di v. Dall'esercizio 1.5 segue che $-1v = -v$.

Cominciamo adesso a vedere alcuni esempi di spazio vettoriale:

Esempio 1.8. 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 $\mathbb{Q}$.

Da questo segue ogni campo $\mathbb{K}$ è uno spazio vettoriale su qualsiasi sottocampo $\mathbb{F}$ di $\mathbb{K}$ stesso (con le operazioni definite come sopra). Ad esempio $\mathbb{R}$ è uno spazio vettoriale su $\mathbb{Q}$.

Sarà vero anche il viceversa, ovvero che se $\mathbb{F}$ è un sottocampo proprio di $\mathbb{K}$ allora $\mathbb{F}$ è uno spazio vettoriale su $\mathbb{K}$ con somma e prodotto definite come quelle che rendono $\mathbb{F}$ e $\mathbb{K}$ campi? Cominciamo ad analizzare un caso particolare: è vero che $\mathbb{Q}$ è uno spazio vettoriale su R con le operazioni di somma e prodotto usuali? La risposta è NO, infatti abbiamo che la somma di due elementi di $\mathbb{Q}$ è ancora un elemento di $\mathbb{Q}$, ma non è vero che il prodotto di un elemento di $\mathbb{Q}$ per uno di $\mathbb{R}$ sia sempre un elemento di $\mathbb{Q}$. Per esempio:

$$\underbrace{2}_{\in\mathbb{Q}} \cdot \underbrace{\sqrt{2}}_{\in\mathbb{R}} = \underbrace{2\sqrt{2}}_{\notin\mathbb{Q}}$$

.
In generale se $\mathbb{F}$ è un sottocampo proprio di $\mathbb{K}$, $\mathbb{F}$ non è uno spazio vettoriale su $\mathbb{K}$ con le operazioni di somma e prodotto che definiscono i due campi. Infatti possiamo ripercorrere lo stesso ragionamento fatto per $\mathbb{Q}$ e $\mathbb{R}$: consideriamo un elemento $a \in \mathbb{K} \setminus \mathbb{F}$ e il prodotto scalare tra a e l'elemento neutro del prodotto 1 che, per definizione di campo, appartiene a $\mathbb{F}$.
Questa osservazione è particolarmente importante perché ricorda che la chiusura dello spazio vettoriale V rispetto alle due operazioni, ovvero che la somma di due elementi di V sia ancora un elemento di V e che il prodotto scalare tra un elemento di V e un elemento di $\mathbb{K}$ sia ancora un elemento di V, fa parte della definizione di spazio vettoriale. È facile infatti provare che nel caso $V = \mathbb{Q}$ e $\mathbb{K} = \mathbb{R}$ tutte le altre proprietà di spazio vettoriale varrebbero!

Esercizio 1.9. Dimostrare che $\mathbb{R}^2 = \{(a, b) | a, b \in \mathbb{R}\}$, l'insieme delle coppie di numeri naturali (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 ci aspettiamo, come segue:

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

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

Osservazione 1.10. Osserviamo, nel caso di $\mathbb{R}^2$ come spazio vettoriale su $\mathbb{R}$, la differenza tra il vettore O elemento neutro della somma vettoriale e lo 0 scalare. Lo O elemento neutro della somma è l'elemento $(0, 0)$ di $\mathbb{R}^2$.

Esercizio 1.11. 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:

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

(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 1.12. $\mathbb{R}^n$ 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 $\mathbb{K}$.

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 costituito dal solo vettore 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 contenente il solo vettore O.

Esempio 1.16. $\mathbb{Q}$ è un sottospazio vettoriale proprio (sul campo $\mathbb{Q}$) di $\mathbb{R}$.

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 1.17. *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 per forza. Quello che ci interessa è però 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. Anche in questo caso, per tutte le proprietà che devono valere per ogni elemento di V, si può concludere che a maggior ragione varranno per tutti gli elementi di W che è un sottoinsieme di V. Dunque sicuramente valgono 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 1.5) 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. □

Osservazione 1.18. 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 1.19. 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.
Osserviamo, ad esempio, che la circonferenza $x^2+y^2 = 1$ in $\mathbb{R}^2$ non è un sottospazio

vettoriale, così come tutte le circonferenze di raggio $r > 0$. Tali insiemi infatti non contengono lo O. Ma anche considerando l'unione tra la circonferenza e lo O non avremmo uno spazio vettoriale, infatti tali insieme non sarebbero chiusi né per somma vettoriale, né per prodotto scalare (provarlo per esercizio). Dimostriamo che:

Esercizio 1.20. Tutti e soli i sottospazi vettoriali propri di $\mathbb{R}^2$ sono le rette passanti per l'origine O.

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:

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

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:

$$(x_1, kx_1) + (x_2, kx_2) = (x_1 + x_2, kx_1 + kx_2) = (x_1 + x_2, k(x_1 + x_2)) \in r$$
$$h(x_1, kx_1) = (hx_1, hkx_1) = (hx_1, khx_1) \in r$$

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 $\mathbb{R}^2$. Lasciamo, per ora, questo punto all'intuizione geometrica: se esistesse w, 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: si tratta di *vedere* 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 $z = w + v$.

Esercizio 1.21. 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[3].

Esercizio 1.22. 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 $n \in \mathbb{N}$.[4]
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 1.17 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 $\mathbb{K}^{\leq n}[x]$.

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

[4]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.

- Sia $p(x) \in \mathbb{K}^{\leq n}[x]$ se $\lambda \in \mathbb{K}$ è diverso da zero si ha che:

$$deg(\lambda \cdot p(x)) = deg(p(x)) \leq n$$

e quindi per ogni $\lambda \in \mathbb{K}$ si ha che $\lambda \cdot p(x) \in \mathbb{K}^{\leq n}[x]$. Altrimenti, se $\lambda = 0$, sappiamo, per ipotesi, che $\lambda \cdot p(x) = 0 \in \mathbb{K}^{\leq n}[x]$.

Ricordiamo che un polinomio $p(x) \in \mathbb{K}[x]$ dà origine ad una funzione da $\mathbb{K}$ in $\mathbb{K}$, che indicheremo sempre con p. Per esempio, $p(x) = 3x^4 + 2x + 1 \in \mathbb{R}[x]$ dà origine alla funzione $p : \mathbb{R} \to \mathbb{R}$ che assegna ad ogni valore $r \in \mathbb{R}$ il valore della valutazione del polinomio $p(x)$ in r, ovvero $p(r) = 3r^4 + 2r + 1$, per cui $p(2) = 53$.

Osservazione 1.23. Il principio d'identità dei polinomi afferma che, se il campo $\mathbb{K}$ dei coefficienti è infinito, allora due polinomi a coefficienti in $\mathbb{K}$ definiscono la stessa funzione da $\mathbb{K}$ in $\mathbb{K}$ (ovvero assegnano ad ogni k in $\mathbb{K}$ lo stesso valore) se e solo se sono lo stesso polinomio.

Bisogna fare attenzione che tale principio non vale più nel caso di $\mathbb{K}$ finiti. Infatti è ovvio che due polinomi uguali continuano a definire la stessa funzione, ma il viceversa può *saltare*. Consideriamo per esempio $\mathbb{K} = \mathbb{Z}_2[x]$ e i due polinomi $f(x) = x + 1$ e $g(x) = x^{134} + 1$; essi sono diversi (due polinomi si dicono uguali se hanno lo stesso grado e gli stessi coefficienti) ma danno luogo alla stessa funzione da $\mathbb{Z}_2$ a $\mathbb{Z}_2$, infatti:

$$f(0) = g(0) = 1 \text{ e } f(1) = g(1) = 0$$

Approfondiremo più avanti questa osservazione.

Esempio 1.24. 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:

$$kp(1) \underbrace{=}_{p(x)\in L} k \cdot 0 = 0$$

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 1.24. 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:

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

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

□

2. 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[5] di V.

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

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

Definizione 1.34. Un insieme di vettori $\{v_1, v_2, \ldots, v_k\}$ 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

$$a_1v_1 + a_2v_2 + \cdots + a_tv_t = v$$

, 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.

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 $\{v_1, \ldots, v_t\}$ di V sappiamo dunque che ogni v in V si può scrivere, con una opportuna scelta dei coefficienti $\{k_1, \ldots, k_t\}$, come:

$$v = \sum_{i=1}^{t} k_i v_i \tag{2.1}$$

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 2.1). In generale la risposta a questa domanda è no, come possiamo vedere dal seguente esempio:

Esempio 1.35. Si verifica (esercizio) che i vettori

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

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:

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

Abbiamo quindi mostrato che, senza qualche ipotesi aggiuntiva sull'insieme dei generatori di uno spazio vettoriale, non possiamo essere sicuri che ogni vettore si scriva in maniera **unica** come combinazione lineare dei generatori. Ci chiediamo allora quale ulteriore condizione bisogna imporre ad un insieme di generatori per poter essere sicuri che ogni elemento dello spazio venga espresso in maniera unica come combinazione lineare. Il concetto chiave è quello di *indipendenza lineare*:

Definizione 1.36. Si dice che un insieme finito di vettori $\{v_1, v_2, \ldots, v_r\}$ 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

$$a_1v_1 + a_2v_2 + \cdots + a_rv_r = O <=> a_1 = a_2 = \cdots = a_r = 0$$

Talvolta si dice anche, più brevemente, che i vettori $v_1, v_2, \ldots, v_r$ sono **linearmente indipendenti**.
Se invece i vettori $v_1, v_2, \ldots, v_r$ 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**).

È bene sottolineare, come si evince dalla definizione, che la caratteristica di essere un insieme di vettori linearmente indipendenti è per l'appunto una caratteristica che riguarda tutto l'insieme di vettori che stiamo considerando. Può accadere che se da un insieme di 5 vettori linearmente dipendenti ne togliamo uno, l'insieme residuo, costituito da 4 vettori, sia un insieme di vettori linearmente indipendenti.

Esempio 1.37. Verificate che i vettori dell'esempio precedente

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

non sono linearmente indipendenti.
Considerando solo l'insieme costituito dai primi tre vettori (o quello costituito dagli ultimi tre) dimostrate che si tratta questa volta di un insieme di vettori linearmente indipendenti. Ma è sempre vero che levando un vettore da un insieme di vettori linearmente dipendenti è garantita la lineare indipendenza? La risposta abbastanza ovvia è NO: provate per esempio a considerare l'insieme di tre vettori ottenuto considerando solo il primo, il secondo e il quarto vettore.

Definizione 1.38. Un insieme di vettori $\{v_1, v_2, \ldots, v_n\}$ che generano lo spazio V e sono anche linearmente indipendenti si dice una **base** (finita) di V.

Non sempre uno spazio vettoriale ammette un numero finito di generatori, un caso che abbiamo incontrato di spazio vettoriale che non ammette una base finita è lo spazio vettoriale sul campo $\mathbb{K}$ dei polinomi a coefficienti in $\mathbb{K}$. Potete facilmente dimostrare questo fatto. Ma ammette una base infinita oppure non ha una base? Per rispondere a questa domanda dobbiamo ampliare la definizione di indipendenza lineare a insiemi infiniti e questo si fa chiedendo che le combinazioni lineari che appaiono nelle formule siano comunque composte da un numero finito di addendi; dopodiché è facile verificare che l'insieme delle potenze di x con esponente $n \in \mathbb{N}$ è una base di $\mathbb{K}[x]$.
Ma in questo corso considereremo quasi sempre spazi vettoriali che ammettono una base finita. Fissiamo dunque uno spazio vettoriale V (sul campo $\mathbb{K}$) e supponiamo che ammetta una base finita $\{v_1, v_2, \ldots, v_n\}$. Mostriamo che 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 1.39. *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 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 $\mathbb{K}$. Sottraendo abbiamo:

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

Ma sappiamo che i vettori $v_1, v_2, \ldots, v_n$ sono linearmente indipendenti. Dunque la combinazione lineare che abbiamo scritto sopra, e che ha come risultato O, deve avere tutti i coefficienti nulli. Così possiamo concludere che $a_i = b_i$ per ogni i, ossia che esiste un solo modo di scrivere v come combinazione lineare degli elementi della base data. □

Il seguente risultato è particolarmente importante perché mostra che se uno spazio vettoriale ammette un insieme finito di generatori, allora ammette anche una base finita. In altre parole i generatori dati potrebbero essere molti, e sovrabbondanti, ma è sempre possibile estrarre dall'insieme dei generatori un sottoinsieme che è una base. Il teorema non si limita al risultato in sé, ma nella dimostrazione si caratterizza la base come un sottoinsieme massimale di vettori linearmente indipendenti dell'insieme dei generatori (massimale rispetto alla proprietà di essere linearmente indipendenti):

Teorema 1.40. *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:

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

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 $\{w_2\}$. Fra tutti gli elementi di $\mathcal{M}$ consideriamone uno di cardinalità massima: $\{w_{i_1}, w_{i_2}, \ldots, w_{i_n}\}$ (sicuramente è $n \geq 1$ ossia tale insieme non è vuoto).

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

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

Se $\{w_{i_1}, w_{i_2}, \ldots, w_{i_n}\} = \{w_1, w_2, \ldots, w_s\}$ abbiamo già finito. 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, se no 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:

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

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. □

Osservazione 1.41. Dalla dimostrazione del teorema 1.40 segue che ogni sottoinsieme massimale di $\mathcal{M}$ è una base di V; questo ci suggerisce che la base di uno spazio vettoriale non è unica. Osserviamo inoltre 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, qualsiasi sia $\lambda \in \mathbb{K} \backslash \{0\}$.

L'osservazione 1.41 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 nel Paragrafo 2 del Capitolo 2), 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 1.42. Sia V uno spazio vettoriale con basi di cardinalità n. Allora chiamiamo n la **dimensione** di V.

Esercizio 1.43. Dimostrare che $\mathbb{K}^{\leq n}[x]$ (che sappiamo dall'Esercizio 1.22 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 1.44. 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 una base di W.

(1) 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).
(2) Un polinomio $p(x)$ di W è per definizione del tipo:

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

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. È facile dimostrare 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 abbiamo scoperto avere dimensione 3.

Esercizio 1.45. Dimostrare che la dimensione di un sottospazio intersezione $U \cap V$ è minore o uguale del minimo fra la dimensione di U e quella di V. Quando è esattamente uguale?

Esercizio 1.46. Sia W il sottospazio di $\mathbb{R}^3$ generato dai vettori: $(1, -1, 0)$ e $(b - 1, b + 1, -b)$ (con $b \in \mathbb{R}$) e U il sottoinsieme di $\mathbb{R}^3$ definito da:

$$U = \{(x, y, z) \in \mathbb{R}^3 | x + y + 2z = 0\}$$

(1) Dimostrare che U è sottospazio di $\mathbb{R}^3$.
(2) Calcolare la dimensione di U e, al variare di $b \in \mathbb{R}$, la dimensione di W.
(3) Calcolare, al variare di $b \in \mathbb{R}$, la dimensione di $U \cap W$.

(1) È ovvio che il vettore $(0, 0, 0)$ appartiene ad U infatti: $0 + 0 + 2 \cdot 0 = 0$. Siano (x_1, y_1, z_1) e (x_2, y_2, z_2) due vettori di U e consideriamo:

$$(x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$$

e

$$\lambda \cdot (x_1, y_1, z_1) = (\lambda x_1, \lambda y_1, \lambda z_1)$$

al variare di λ in $\mathbb{R}$. Dobbiamo mostrare che questi due elementi appartengono ancora ad U, ma la verifica è del tutto banale infatti:

$$(x_1 + x_2) + (y_1 + y_2) + 2(z_1 + z_2) = \underbrace{x_1 + y_1 + 2z_1}_{=0} + \underbrace{x_2 + y_2 + 2z_2}_{=0} = 0$$

e

$$\lambda x_1 + \lambda y_1 + 2\lambda z_1 = \lambda \underbrace{(x_1 + y_1 + 2z_1)}_{=0} = 0$$

(2) Per determinare la dimensione di U cerchiamo una base. Il sottospazio U è determinato dalla condizione $x + y + 2z$ che ha due gradi di libertà, ovvero se fissiamo un valore s alla variabile z e un valore t alla variabile y abbiamo che il valore di x è univocamente determinato ed è uguale a $-t - 2s$, si ha cioè:

$$\begin{cases} y = s \\ z = t \\ x = -t - 2s \end{cases}$$

e quindi ogni vettore (x, y, z) di U è del tipo:

$$(-t - 2s, t, s) = (-1, 1, 0) \cdot t + (-2, 0, 1) \cdot s$$

cioè $\{(-1, 1, 0), (-2, 0, 1)\}$ è un insieme di generatori di U (quindi la dimensione di U è minore o uguale a 2). È facile mostrare che i due vettori sono anche linearmente indipendenti (esercizio!) e quindi $\{(-1, 1, 0), (-2, 0, 1)\}$ è una base di U e $\dim(U) = 2$.
Per quanto riguarda W bisogna capire, al variare di $b \in \mathbb{R}$, se i due vettori $(1, -1, 0)$ e $(b - 1, b + 1, -b)$ siano linearmente indipendenti (e quindi costituiscano una base di W di cardinalità 2) oppure linearmente dipendenti (e quindi $\dim(W) = 1$ e una sua base è composta da uno dei due vettori):

$$h \cdot (1, -1, 0) + l \cdot (b - 1, b + 1, -b) = (0, 0, 0)$$

Otteniamo il sistema:

$$\begin{cases} h + l \cdot (b - 1) = 0 \\ -h + l \cdot (b + 1) = 0 \\ -bl = 0 \end{cases}$$

Se $b \neq 0$ allora l deve essere uguale a zero e di conseguenza anche $h = 0$. Quindi se $b \neq 0$ i due vettori $(1, -1, 0)$ e $(b-1, b+1, -b)$ sono linearmente indipendenti. Se $b = 0$ allora il sistema diventa:

$$\begin{cases} h - l = 0 \\ -h + l = 0 \\ 0 = 0 \end{cases}$$

Che è risolto per ogni scelta di $h = l$ e quindi $(1, -1, 0)$ e $(b-1, b+1, -b)$ sono linearmente dipendenti e W ha dimensione uno.

(3) Osserviamo che i generatori di W stanno in U infatti:

$$1 - 1 + 2 \cdot 0 = 0 \quad \text{e} \quad (b-1) + (b+1) + 2 \cdot (-b) = 2b - 2b = 0$$

Quindi $W \subseteq Q$ e $W \cap Q = W$. Perciò abbiamo già trovato la dimensione di $W \cap U$ nel punto precedente.

3. Applicazioni lineari e matrici.

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

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

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

È naturale chiedersi quali funzioni "rispettano" le strutture vettoriali che abbiamo introdotto; per esempio domandiamoci: quali funzioni mandano sottospazi in sottospazi? Gli oggetti giusti da prendere in considerazione sono le applicazioni lineari:

Definizione 1.47. Siano V e W spazi vettoriali di dimensione finita sul campo $\mathbb{K}$. Una applicazione lineare L da V a W è una funzione

$$L : V \to W$$

che soddisfa le seguenti due proprietà:

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

Osservazione 1.48. Le due proprietà della definizione possono essere espresse in maniera equivalente dalla seguente richiesta: per ogni $v_1, v_2 \in V$ e per ogni $a, b \in \mathbb{K}$ vale

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

Definizione 1.49. Siano V e W spazi vettoriali sul campo $\mathbb{K}$ e consideriamo una applicazione lineare L da V a W. Chiameremo *nucleo*[6] di L il seguente sottoinsieme di V:

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

Il nucleo di una applicazione lineare è dunque costituito da tutti i vettori che vengono mandati in O. I seguenti due esercizi individuano importanti proprietà del nucleo e dell'immagine di una applicazione lineare.

[6]La parola inglese per nucleo è kernel, questo spiega il simbolo $Ker\ L$ che sta verrà utilizzato.

Esercizio 1.50. Siano V e W spazi vettoriali sul campo $\mathbb{K}$ e consideriamo una applicazione lineare L da V a W. Dimostrare che $Ker\ L$ è un sottospazio vettoriale di V e che $Imm\ L$ è un sottospazio vettoriale di W.

Esercizio 1.51. Siano V e W spazi vettoriali sul campo $\mathbb{K}$ e consideriamo una applicazione lineare L da V a W. Dimostrare che L è iniettiva se e solo se $Ker\ L = \{O\}$.

Per poter fare calcoli con le applicazioni lineari spesso conviene rappresentarle tramite le *matrici*, come spiegheremo nei prossimi due paragrafi.

3.1. Matrici e prodotto fra matrici.

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 $\mathbb{K}$:

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

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}. Chiameremo gli elementi a_{ij} i *coefficienti* della matrice.

Definizione 1.52. Dati due interi positivi m, n, chiamiamo $Mat_{m\times n}(\mathbb{K})$ l'insieme costituito dalle matrici $m \times n$ a coefficienti in $\mathbb{K}$.

Sull'insieme $Mat_{m\times n}(\mathbb{K})$ è possibile definire una somma e una moltiplicazione per scalare. Mostriamo queste due operazioni con due esempi, da cui il lettore potrà facilmente intuire qual è la definizione generale. Esempio per la somma:

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

Esempio per la moltiplicazione per scalare:

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

Esercizio 1.53. Dimostrare che, con le operazioni somma e prodotto per scalare illustrate sopra, $Mat_{m\times n}(\mathbb{K})$ è uno spazio vettoriale su $\mathbb{K}$.

C'è un'altra operazione fra matrici su cui vogliamo soffermarci: il *prodotto righe per colonne*. Innanzitutto introduciamo una notazione più compatta per scrivere le matrici. Indicheremo una matrice $m \times n$ nel seguente modo: $A = (a_{ij})$. Talvolta, per ricordare meglio quali sono le dimensioni della matrice scriveremo:

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

Ora, sia $A = (a_{ij})$ una matrice $m \times n$ e sia $B = (b_{st})$ una matrice $n \times k$; il prodotto righe per colonne AB è una matrice $m \times k$ che chiameremo $C = (c_{rh})$ tale che per ogni r, h:

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

Facciamo un esempio:

$$\begin{pmatrix} 1 & 2 & 4 \\ 0 & 6 & 3 \end{pmatrix} \begin{pmatrix} 2 & 2 & 2 \\ 5 & 6 & -8 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 12 & 18 & -14 \\ 30 & 39 & -48 \end{pmatrix}$$

Controlliamo nel dettaglio come è stato ottenuto il coefficiente $c_{21} = 30$. Abbiamo utilizzato la seconda riga di A e la prima colonna di B:

$$c_{21} = 0 \cdot 2 + 6 \cdot 5 + 3 \cdot 0 = 30$$

Osservazione 1.54. Come si può notare, c_{rh} si ottiene utilizzando i coefficienti della riga r-esima di A e quelli della colonna h-esima di B; per come è definita questa operazione è fondamentale che il numero di colonne di A sia uguale al numero di righe di B. Se questo non accade il prodotto righe per colonne fra due matrici A e B non è definito. In particolare questo prodotto ci dà una nuova operazione sull'insieme $Mat_{n\times n}(\mathbb{K})$. Mostreremo più avanti che con la somma introdotta in precedenza e col prodotto righe per colonne l'insieme $Mat_{n\times n}(\mathbb{K})$ è un anello non commutativo.

Esercizio 1.55. Mostrare con un esempio che, date due matrici $n \times n$ A, B, non è detto che $AB = BA$ (si trovano esempi anche con $n = 2$).

3.2. Rappresentare gli elementi di uno spazio vettoriale con vettori in colonna e le applicazioni lineari tramite matrici.

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

$$L : V \to W$$

Scegliamo in V una base $\{e_1, e_2, \ldots, e_n\}$ e in W una base $\{\epsilon_1, \epsilon_2, \ldots, \epsilon_m\}$.
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:

$$v = b_1e_1 + b_2e_2 + \cdots + b_ne_n$$

Per la linearità di L allora:

$$L(v) = b_1L(e_1) + b_2L(e_2) + \cdots + b_nL(e_n)$$

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

Esercizio 1.56 (Importante!)**.** Dimostrare, in base alle osservazioni precedenti, che $ImmL =< L(e_1), L(e_2), \ldots, L(e_n) >$.

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

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

In conclusione 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, permette di esprimere $L(v)$ come combinazione lineare degli elementi della base $\{\epsilon_1, \epsilon_2, \ldots, \epsilon_m\}$:

(3.1)

$$L(v) = (b_1a_{11}+b_2a_{12}+\cdots+b_na_{1n})\epsilon_1+(b_1a_{21}+b_2a_{22}+\cdots+b_na_{2n})\epsilon_2+\cdots+(b_1a_{m1}+b_2a_{m2}+\cdots+b_na_{mn})\epsilon_m$$

Utilizzeremo la seguente comoda notazione. Possiamo esprimere il vettore $L(e_i)$ in colonna:

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

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

Osservazione 1.57. Useremo spesso, per uno spazio vettoriale di cui sia stata fissata una base, questa notazione con i vettori messi 'in colonna'.

Il conto appena svolto si può dunque tradurre in questa nuova notazione nel seguente modo:

$$L(v) = b_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \cdots \\ \cdots \\ a_{m1} \end{pmatrix} + b_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \\ \cdots \\ \cdots \\ a_{m2} \end{pmatrix} + \cdots + b_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ \cdots \\ \cdots \\ a_{mn} \end{pmatrix}$$

E infine la relazione (3.1) diventa:

$$L(v) = \begin{pmatrix} b_1a_{11} + b_2a_{12} + \cdots + b_na_{1n} \\ b_1a_{21} + b_2a_{22} + \cdots + b_na_{2n} \\ b_1a_{31} + b_2a_{32} + \cdots + b_na_{3n} \\ \cdots \\ \cdots \\ b_1a_{m1} + b_2a_{m2} + \cdots + b_na_{mn} \end{pmatrix} \tag{3.2}$$

Ecco che entrano in scena le matrici.

Definizione 1.58. La matrice associata all'applicazione lineare L nelle basi $\{e_1, e_2, \ldots, e_n\}$ e $\{\epsilon_1, \epsilon_2, \ldots, \epsilon_m\}$ è data dalla seguente griglia di m righe per n colonne:

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

Notiamo che la matrice $[L]$ (da ora in poi, per semplificare la notazione, ometteremo il riferimento alle basi tutte le volte che potremo farlo senza creare ambiguità) si ottiene ponendo uno accanto all'altro i vettori $L(e_1), L(e_2), \ldots, L(e_n)$, scritti come colonne nella base scelta di W.

Torniamo al problema di calcolare $L(v)$ in maniera conveniente.
Scriviamo v come colonna:

$$v = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \ldots \\ \ldots \\ b_n \end{pmatrix}$$

dove abbiamo messo uno sotto l'altro i coefficienti di v rispetto alla base $\{e_1, e_2, \ldots, e_n\}$.
Una semplice verifica ci permette di osservare che con queste notazioni vale:

$$L(v) = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \ldots & a_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \ldots \\ \ldots \\ b_n \end{pmatrix}$$

dove l'operazione che entra in gioco è il prodotto righe per colonne (per verificarlo basta svolgere il prodotto righe per colonne e confrontare con (3.1)). Dunque la matrice $[L]$, mediante il prodotto righe per colonne, ci permette di calcolare come agisce l'applicazione L sui vettori di V.

Esempio 1.59. Facciamo ora un esempio che mostra come la matrice associata ad una applicazione lineare dipenda dalle basi scelte. Consideriamo gli spazi vettoriali $\mathbb{R}^4$ e $\mathbb{R}^3$ con le loro basi standard, rispettivamente

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

e

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

Consideriamo poi la applicazione lineare

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

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

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

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

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

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

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

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

Dunque, preso per esempio il vettore

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

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

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

Il risultato è

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

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

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

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

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

Proviamo a scrivere la matrice

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

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

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

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

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

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

e dunque

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

Allora il vettore da inserire come prima colonna della matrice

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

è

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

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

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

Osservazione 1.60. Dati due spazi vettoriali V, W, c'è una applicazione lineare da V a W la cui matrice associata non cambia mai qualunque siano le 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.

Un altro caso speciale che vorremmo segnalare è quello dell'applicazione *identità* $I : V \to V$, quella che lascia fisso ogni elemento di v: $I(v) = v \quad \forall v \in V$. Se scegliamo in V *la stessa base* sia in arrivo sia in partenza, si verifica immediatamente che la matrice $[I] = (a_{ij})$ 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 $i = 1, 2, \ldots, n$. Tale matrice è l'elemento neutro rispetto alla moltiplicazione in $Mat_{n\times n}(K)$. 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.

4. Altri esercizi

Esercizio 1.61. 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. Proviamo che effettivamente L è un'applicazione *lineare*[7], 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 andiamo a provare 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:

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

E ora passiamo alle altre richieste dell'esercizio:

(2) Per scrivere una matrice associata a L bisogna fissare una base (ordinata[8]) per ognuno dei due spazi. In questo caso le basi sono state fissate:

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

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

Per scrivere la matrice associata a L rispetto a queste basi bisogna calcolare l'immagine degli elementi di $ß_{\mathbb{R}^3}$ tramite L. Le coordinate di questi elementi rispetto alla base $ß_{\mathbb{R}^2}$ ci forniscono le colonne della matrice

[7] In effetti è facile provare che lo sono tutte le applicazioni da $\mathbb{R}^n$ a $\mathbb{R}^m$ che *agiscono* sulle coordinate in maniera che il risultato sia una combinazione lineare delle stesse.

[8] Ordinata, perché per la costruzione della matrice è importante anche l'ordine in cui si considerano gli elementi delle due basi.

associata a 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 $[A]_{\substack{ß_{\mathbb{R}^3}\\ß_{\mathbb{R}^2}}}$ associata a L nelle basi suddette è la seguente:

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

(3) Sappiamo (vedi Esercizio 1.56) 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 sappiamo anche che questi non sono una base (infatti $\mathbb{R}^2$ ha dimensione 2 e quindi $Imm(L)$ ha al massimo dimensione 2). Lasciamo al lettore la facile 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 $z\in\mathbb{R}$. A questo punto sappiamo che i vettori che stanno in $Ker(L)$ sono 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 1.62. 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 1.63. 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 1.64. 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 1.65. Sia $F : \mathbb{C}^3 \to \mathbb{C}^3$ l'applicazione lineare definita, nella base standard di $\mathbb{C}^3$, dalla matrice:

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

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

Esercizio 1.66. 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 1.67. 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 $\mathbb{R}$. 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 1.68. 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 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 1.69. 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 1.70. 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 1.71. Sia $\mathcal{T} : Mat_{n\times n}(\mathbb{K}) \to \mathbb{K}$ la funzione *traccia* definita da

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

a) Dimostrare che $\mathcal{T}$ è una applicazione lineare (per la struttura di spazio vettoriale su $Mat_{n\times n}(\mathbb{K})$ vedi l'Esercizio 1.53).
b) 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. Studiare l'immagine di una applicazione lineare: le operazioni elementari sulle colonne e il concetto di rango

Consideriamo una matrice $m \times n$, a coefficienti in un campo $\mathbb{K}$. Numeriamo le colonne della matrice come sempre da sinistra verso destra e consideriamo i seguenti tre tipi di mosse sulle colonne (dette anche *operazioni elementari sulle colonne*):

- 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.

Vedremo fra poco che è sempre possibile, usando le mosse descritte sopra, ridurre la matrice in una forma detta *a scalini (per colonne)*. Per intenderci, ecco alcuni esempi di matrici in forma a scalini per colonne:

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

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

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

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

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

Per una definizione formale di matrice a scalini per colonne possiamo seguire questa strada: chiamiamo *profondità* di un vettore la posizione occupata, contata dal basso,

dal suo più alto coefficiente diverso da zero. Per esempio i vettori

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

hanno rispettivamente profondità 2,3,1. Allora una matrice a scalini per colonne è una matrice tale che

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

Negli esempi abbiamo aggiunto anche la condizione che in ogni vettore colonna il coefficiente più alto diverso da zero sia uguale a 1, ma questa richiesta non è essenziale.

Teorema 2.1. *Data una matrice $m \times n$, a coefficienti in un campo $\mathbb{K}$, è sempre possibile, usando le operazioni elementari sulle colonne, ridurre la matrice in forma a scalini per colonne.*

Dimostrazione. Per induzione su m. Il caso $m = 1$ è semplice: se la matrice è composta da una riga nulla $(0\ 0\ 0\ 0\ 0\ 0\ \ldots 0)$ è già in forma a scalini; se invece ha qualche coefficiente non 0, con le operazioni di colonna si può facilmente portare nella forma $(1\ 0\ 0\ 0\ 0\ 0\ \ldots 0)$.

Supponiamo ora che l'enunciato sia vero per matrici con $m-1$ righe, e consideriamo una matrice A di formato $m \times n$. Se una riga di A è nulla, possiamo considerare la matrice A' ottenuta da A togliendo tale riga. Per ipotesi induttiva sappiamo ridurre a scalini per colonne A', visto che ha $m-1$ righe; osserviamo a questo punto che la stessa esatta sequenza di operazioni per colonne riduce a scalini anche A.

Se invece A non ha righe nulle, si sceglie una colonna di profondità massima, cioè di profondità m. Possiamo supporre che sia la prima colonna da sinistra (se non lo è possiamo sempre fare uno scambio di colonne). A questo punto, sottraendo tale colonna, moltiplicata per opportuni scalari, alle altre colonne, si giunge ad avere una matrice A' del tipo:

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

Per ipotesi induttiva sappiamo che possiamo ridurre a scalini la sottomatrice

$$\begin{pmatrix} 0 & 0 & \ldots & 0 \\ a_{22} & a_{23} & \ldots & a_{2n} \\ a_{32} & a_{33} & \ldots & a_{3n} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix}$$

che ha una riga (e una colonna) in meno. Le stesse mosse, operate sulla matrice A', la riducono a scalini.

□

Osserviamo che, sempre con le mosse di colonna (in particolare con le operazioni del primo tipo) si può sempre trasformare ulteriormente una matrice a scalini ponendola in forma *a scalini per colonna ridotta*.

Ecco le forme a scalini ridotte negli esempi appena visti:

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

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

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

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

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

Esercizio 2.2. In base agli esempi qui sopra, il lettore provi a dare una definizione formale di "matrice a scalini per colonne ridotta".

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

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

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

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

Dunque M_{ij}^{-1} è quasi uguale alla M_{ij}: differisce solo per un coefficiente, perché al posto di λ compare $-\lambda$. Il lettore può verificare con un semplice calcolo che le due matrici indicate sono l'una l'inversa dell'altra.

Osservazione 2.3. Quanto detto sopra mette in evidenza che la mossa in questione è "reversibile", ossia, una volta fatta, possiamo fare la sua inversa e tornare esattamente alla matrice di partenza: lo stesso vale per gli altri 2 tipi di mosse (trovate per esercizio le matrici invertibili che le realizzano).

Consideriamo allora una applicazione lineare:

$$L : V \to W$$

dove V e W sono due spazi vettoriali sul campo $\mathbb{K}$, di dimensione n e m rispettivamente.
Prendiamo in V una base $\{e_1, e_2, \ldots, e_n\}$ e in W una base $\{\epsilon_1, \epsilon_2, \ldots, \epsilon_m\}$.
A questo punto alla L si può associare una matrice $[L]$, di forma $m \times n$, nelle basi scelte. Ora facciamo operazioni elementari sulle colonne di $[L]$ fino a ridurla in forma a scalini ridotta: questo equivale a dire che moltiplichiamo $[L]$ a destra per tante matrici invertibili $[M_1], [M_2], \ldots, [M_k]$ fino a che $[L][M_1][M_2]\cdots[M_k]$ è in forma a scalini ridotta.

Per semplificare la notazione, chiamiamo $[M] = [M_1][M_2]\cdots[M_k]$: sappiamo che $[M]$ è invertibile, visto che è il prodotto di matrici invertibili (la sua inversa è $[M]^{-1} = [M_k]^{-1}[M_{k-1}]^{-1}\cdots[M_1]^{-1}$).

Ma la matrice $[M]$ è associata ad una applicazione lineare $M : V \to V$. Infatti, fissata la base di V, in questo caso $\{e_1, e_2, \ldots, e_n\}$, osserviamo che la corrispondenza fra applicazioni lineari da V in V e matrici $n \times n$ a coefficienti in $\mathbb{K}$ è bigettiva, ossia data una applicazione lineare costruiamo la matrice, data una matrice troviamo la applicazione lineare da essa rappresentata.

E chi è la applicazione lineare associata a $[L][M]$? 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 2.4. *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 vale, rispetto alle basi fissate:*

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

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 $\begin{pmatrix} 1 \\ 0 \\ \ldots \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ \ldots \\ 0 \end{pmatrix}$ etc... e controllare che diano lo stesso risultato).

□

Torniamo alla applicazione lineare L, e studiamo la sua immagine $Imm\ L$.

Proposizione 2.5. *Vale che $Imm\ (L \circ M) = Imm\ L$, ossia, scritto con un'altra notazione, $(L \circ M)\ (V) = L(V)$.*

Dimostriamo questa proposizione dimostrando più in generale che

Proposizione 2.6. *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

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

□

Ricordiamo (vedi Esercizio 1.56) 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 colonna di $[L]$. Ma la proposizione garantisce che $Imm\ L = Imm\ (L \circ M)$, dunque $L(V)$ è anche generato dai vettori colonna 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 $[L][M]$ formano un insieme di vettori linearmente indipendenti, e dunque che tali colonne, siccome appunto sono linearmente indipendenti e inoltre generano $Imm\ L$, sono una base di $Imm\ L$.

Osservazione 2.7. Se riducessimo la matrice $[L]$ in forma a scalini con altre mosse potremmo trovare una forma a scalini ridotta diversa, rappresentata dalla matrice $[L][M']$. Ma lo stesso ragionamento ci permetterebbe di affermare che anche le colonne non nulle di $[L][M']$ danno una base di $Imm\ L$.

Introduciamo il concetto di *rango* di una applicazione lineare.

Definizione 2.8. 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 2.9. *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 ridotta (o anche solo a scalini, visto che il numero di colonne non nulle è lo stesso).*

Osservazione 2.10. Se avessimo fissato altre basi avremmo avuto una matrice $[L]$ diversa, ma, trasformandola in forma a scalini ridotta, avremmo ancora ovviamente trovato lo stesso numero di colonne non nulle, giacché tale numero è $dim\ Imm\ L$, ossia dipende dalla applicazione (è la dimensione della sua immagine) e non dalle basi scelte.

Osservazione 2.11. 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 1.40, 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 $[L]$.

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 che, dato uno spazio vettoriale che ammette una base finita, tutte le basi di tale spazio hanno la stessa cardinalità (risultato annunciato nel Paragrafo 2 del Capitolo 1).

Teorema 2.12. *Se uno spazio vettoriale V ha una base di cardinalità $n \in \mathbb{Z}^+$, allora tutte le altre basi di V hanno la stessa cardinalità.*

DIMOSTRAZIONE. Prendiamo due basi di V: $\{e_1, e_2, \ldots, e_n\}$ 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": per esempio, se M' fosse

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

si vedrebbe subito che il vettore

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

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

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

Corollario 2.13. *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. Consideriamo una base di V $e_1, e_2, \ldots, e_n$ e siano $v_1, v_2, \ldots, v_n$ i vettori in questione.

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 $n \times n$. In entrambi i casi previsti dal teorema la matrice M' in forma a scalini ridotta ottenuta a partire da M è la matrice identità[1]: $M' = I$. 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 $M' = I$.

D'altra parte, se i vettori $v_1, v_2, \ldots, v_n$ generano V allora le n colonne di M' devono generare V (sempre perché facendo le mosse di colonna lo spazio generato dalle colonne non cambia) e dunque di nuovo non ci possono essere scalini lunghi. □

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, e dati n vettori $v_1, v_2, \ldots, v_n$, tali vettori costituiscono una base di V o no. Il criterio è il seguente: esprimiamo i vettori $v_1, v_2, \ldots, v_n$ in termini di una base nota $e_1, e_2, \ldots, e_n$ 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, altrimenti no. Perché ?

Nella dimostrazione del Teorema 2.12 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 sono linearmente indipendenti e questo implica che le n colonne di M sono linearmente indipendenti, ossia che $v_1, v_2, \ldots, v_n$ sono linearmente indipendenti. Il corollario precedente ci assicura che questo basta per dire che $v_1, v_2, \ldots, v_n$ sono una base.

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

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

Si noti che si tratta dei vettori utilizzati in un esempio alla fine del Paragrafo 3 del Capitolo 1; il lettore avrà dunque già verificato che si tratta di una base. Ma ora possiamo farlo col nuovo metodo.

Scriviamo dunque la matrice

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

[1]Usiamo il simbolo I per indicare la matrice identità, come convenuto nella Osservazione 1.60.

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 seguente:

Teorema 2.14 (Teorema del Completamento)**.** *Dato uno spazio vettoriale V di dimensione n, ogni sottoinsieme $B = \{v_1, \ldots, v_k\} \subset V$ di vettori linearmente indipendenti di cardinalità $k \leqslant n$, 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 $v_1, v_2, \ldots, v_k$ come vettori colonna rispetto a una base data e si forma una matrice M. Poi si riduce M in forma a scalini per colonne. Tutte le volte che troviamo uno scalino lungo (diciamo di altezza $i \geq 2$) possiamo facilmente trovare $i-1$ vettori $w_1, w_2, \ldots, w_{i-1}$ tali che $\{v_1, v_2, \ldots, v_k, w_1, w_2, \ldots, w_{i-1}\}$ è ancora un insieme di vettori linearmente indipendenti. Ripetendo questa costruzione per ogni scalino lungo, troviamo alla fine n vettori linearmente indipendenti, dunque una base di V come richiesto. □

Illustriamo il metodo descritto in quest'ultima dimostrazione con un esempio. Supponiamo che $V = \mathbb{R}^7$ e di avere 4 vettori linearmente indipendenti che scriviamo rispetto alla base standard di $\mathbb{R}^7$:

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

La matrice M in questo caso è:

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

e una sua riduzione a scalini per colonne è :

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

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

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

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

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

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

3. Le operazioni elementari di riga e un approfondimento sul concetto di rango

Nei paragrafi precedenti abbiamo studiato le operazioni elementari di colonna su una matrice. Ripetiamo tutto per le righe invece che per le colonne; data una matrice $m \times n$ a coefficienti in un campo $\mathbb{K}$, ordiniamo le righe contandole dall'alto

in basso e consideriamo sulle righe le mosse "simmetriche" alle mosse di colonna che conosciamo già:

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

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

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

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

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

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

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

Osservazione 2.15. Per una definizione formale di matrice a scalini per righe possiamo procedere come nel caso delle matrici a scalini per colonne, chiamando stavolta *profondità di una riga* la posizione occupata, contata da destra, dal suo coefficiente diverso da zero che sta più a sinistra... (insomma la riga $(0, 0, 7, \sqrt{2}, 3, 0, 0, 9)$ ha profondità 6, visto che il coefficiente 7 occupa la sesta casella contando da destra).

Anche in questo caso possiamo sempre ottenere anche la forma *a scalini (per righe) ridotta*.

Ecco le forme a scalini ridotte negli esempi appena visti:

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

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

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

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

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

Esercizio 2.16. Dare una definizione formale di matrice a scalini per righe e matrice a scalini per righe ridotta.

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

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

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

Consideriamo allora 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 2.4.

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

Proposizione 2.17. *Sia*

$$B : W \to W$$

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

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

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

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

e verifichiamo che deve valere $a_1 = a_2 = \cdots = a_r = 0$.
Per la linearità di B possiamo riscrivere

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

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

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

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

Si osserva subito che $B(v_1), \dots, B(v_r)$ generano $Imm\,(B \circ L)$. Verifichiamolo per esercizio: gli elementi di $Imm\,(B \circ L)$ sono tutti i vettori della forma $B(u)$ dove u è un vettore che appartiene all'immagine di L. Ma allora u può essere scritto in termini della base $\{v_1, \dots, v_r\}$ di $Imm\, L$:

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

e di conseguenza

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

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

□

Dunque, siccome il rango di L e quello di $U \circ L$ sono uguali, allora il massimo numero di colonne linearmente indipendenti di $[L]$ deve essere uguale al massimo numero di colonne linearmente indipendenti della sua forma a scalini per righe $[U][L]$ (questo infatti è un modo di contare il rango, come sappiamo dalla Osservazione 2.11). 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 2.18. *Data l'applicazione lineare* L *come sopra, il suo rango è uguale al numero di righe non nulle che si trovano quando si riduce una matrice associata* $[L]$ *in forma a scalini per righe.*

Nel caso delle colonne avevamo osservato che il numero di colonne non zero della forma a scalini era uguale al massimo numero di colonne linearmente indipendenti della matrice iniziale: per ragioni puramente di simmetria 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:

Teorema 2.19. *Data l'applicazione lineare L come sopra, il suo rango è uguale al numero di righe non nulle che si trovano quando si riduce una matrice associata $[L]$ in forma a scalini per righe. Tale numero è anche uguale al massimo numero di righe linearmente indipendenti di $[L]$.*

E, in sintesi:

Teorema 2.20. *Data l'applicazione lineare $L : V \to W$, fissiamo una base in V e una in W e consideriamo la matrice $[L]$ associata a L rispetto a tali basi.*
1) Il massimo numero di righe linearmente indipendenti di questa matrice è uguale al massimo numero di colonne linearmente indipendenti, ossia al rango di L.
2) Se si riduce la matrice in forma a scalini, sia che lo si faccia per righe, sia che lo si faccia per colonne, il numero di scalini che otterremo sarà sempre uguale al rango di L.

Osservazione 2.21. Potremmo chiamare *rango di una matrice M* il massimo numero di colonne (o righe, abbiamo visto che è lo stesso) linearmente indipendenti. Con tale definizione il rango di una applicazione lineare L coincide con quello di una sua matrice associata $[L]$, dunque possiamo usare la parola rango senza tante attenzioni, applicandola sia alle matrici sia alle applicazioni. Ora, sappiamo che se componiamo L a destra o a sinistra per una applicazione invertibile, il rango non cambia. Dunque, se moltiplichiamo $[L]$ a destra o a sinistra per matrici invertibili, anche il rango delle matrici non cambia. Allora se abbiamo una applicazione lineare L e ci interessa calcolarne il rango, prendiamo una matrice associata $[L]$ (va bene una qualunque, il discorso è indipendente dalle basi) e possiamo ridurla a scalini usando sia mosse di riga sia mosse di colonna nell'ordine che ci torna più comodo.

4. Il teorema che lega dimensione dell'immagine e dimensione del nucleo

Usando il teorema del completamento (Teorema 2.14) possiamo dimostrare:

Teorema 2.22. *Considerata una applicazione lineare $L : V \to W$, dove V e W sono spazi vettoriali su di un campo $\mathbb{K}$ e V ha dimensione finita, vale*

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

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

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

Sappiamo che che $Imm\ L$ è il sottospazio generato da

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

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

$$Imm\ L =< L(w_1), L(w_2), \dots, L(w_{n-k}) > .$$

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

Per dimostrare la indipendenza lineare scriviamo:

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

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

Per linearità l'equazione equivale a:

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

ossia

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

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

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

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

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

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

Definizione 2.23. Una applicazione lineare bigettiva $L : V \to W$, dove V e W sono spazi vettoriali su di un campo $\mathbb{K}$, si dice un *isomorfismo lineare*.

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

5. Altri esercizi

Esercizio 2.24. Sia $a \in \mathbb{R}$ e $f_a : \mathbb{R}^4 \to \mathbb{R}^2$ l'applicazione lineare data da

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

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

(2) Sia W_a il sottospazio di $\mathbb{R}^4$ generato dai vettori $(7, a^2, -1, a), (-a, -4, 1, 8)$. Calcola la dimensione di $Ker\ f_a \cap W_a$ al variare di $a \in \mathbb{R}$.

Esercizio 2.25. Consideriamo in $\mathbb{R}^4$ il sottospazio V generato dai vettori

$$\begin{pmatrix} 1 \\ -6 \\ 3 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

e il sottospazio W generato dai vettori

$$\begin{pmatrix} 2 \\ 1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -6 \\ 2 \\ 3 \end{pmatrix}$$

a) Calcolare $dim\ V \cap W$.

b) Trovare un vettore $v \in \mathbb{R}^4$ che non appartiene a $V + W$.

Esercizio 2.26. 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 2.27. 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 $a_{11} = a_{22} = \cdots = a_{rr} = 1$.

Esercizio 2.28. Sia A la matrice a coefficienti in $\mathbb{Z}$ seguente:

$$A = \begin{pmatrix} 1 & 5 & -6 & 0 \\ 2 & -8 & -14 & 12 \\ -1 & 7 & 10 & 0 \end{pmatrix}$$

Per ogni p primo, denotiamo con A_p la matrice A pensata con i coefficienti in $\mathbb{Z}_p$. Determinare, al variare di p tra i primi, il rango della matrice A_p.

Svolgimento. Il rango di una matrice non varia se si fanno operazioni di riga o di colonna. Con alcune operazioni di riga si ottiene

$$A \overset{\substack{[2]=2[1]-[2] \\ [3]=[1]+[3]}}{\longrightarrow} B = \begin{pmatrix} 1 & 5 & -6 & 0 \\ 0 & 18 & 2 & -12 \\ 0 & 12 & 4 & 0 \end{pmatrix} \overset{[3]=3[3]-2[3]}{\longrightarrow} C = \begin{pmatrix} 1 & 5 & -6 & 0 \\ 0 & 18 & 2 & -12 \\ 0 & 0 & 8 & 24 \end{pmatrix}$$

Dunque A ha rango 3. Osserviamo che l'operazione di riga $[3] = 3[3] - 2[3]$ (che è in realtà la composizione di due operazioni elementari di riga, moltiplicare la terza riga per 3 e poi sostituirla con se stessa meno la seconda riga moltiplicata per 2) non può essere fatta in $\mathbb{Z}_3$) perché una delle due operazioni coinvolte è "sostituire la terza riga con se stessa moltiplicata per 3", che in $\mathbb{Z}_3$ non è una operazione ammessa (si può sostituire una riga con se stessa moltiplicata per λ ma λ deve

essere diverso da 0). A parte il caso $\mathbb{Z}_3$, le mosse che abbiamo fatto sono valide per tutti gli altri $\mathbb{Z}_p$. Dunque il rango di A_p sarà 3 per tutti i primi che non dividono 8 e 18, ovvero per tutti i primi diversi da 2 e da 3, (del resto il caso $p = 3$ doveva già essere studiato a parte). Ci rimane dunque da studiare questi due casi:

- Se $p = 2$, A_2 è la seguente matrice (si ottiene riducendo modulo 2 la matrice C):
$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
e dunque A_2 ha rango 1.
- Se $p = 3$, A_3 è la seguente matrice (si ottiene riducendo modulo 3 la matrice B, ovvero la matrice che avevamo prima di fare la mossa 'proibita' in $\mathbb{Z}_3$):
$$\begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$
e dunque A_3 ha rango 2.

Esercizio 2.29. Sia $f_h : (\mathbb{Z}_5)^3 \longrightarrow (\mathbb{Z}_5)^3$, al variare di h in $\mathbb{Z}_5$, l'applicazione lineare la cui matrice associata rispetto alle basi standard è la seguente:
$$[f_h] = \begin{pmatrix} 1 & h & 4 \\ 0 & 0 & 1 \\ 0 & h & 0 \end{pmatrix}$$
Sia V il sottospazio vettoriale $V = \{(x, y, z) \in (\mathbb{Z}_5)^3 | y - z = 0\}$.

Studiare per quali valori di $h \in \mathbb{Z}_5$ si ha che $f_h(V) \subseteq V$.

Svolgimento.

Troviamo una base di V. Gli elementi di V sono le terne di elementi di $\mathbb{Z}_5$ che risolvono il sistema lineare omogeneo $y - z = 0$. Perciò i vettori di V sono del tipo:
$$\begin{pmatrix} s \\ t \\ t \end{pmatrix}$$
al variare di s e t in $\mathbb{Z}_5$ (in particolare li possiamo anche contare, sono 25). Dunque il generico vettore di V può essere scritto, al variare di s e t, come segue:
$$\begin{pmatrix} s \\ t \\ t \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} s + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} t$$
Questo ci dice che i vettori:
$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; v_2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$
generano V, ed essendo anche linearmente indipendenti (perché?), sono una base di V (che dunque ha dimensione 2).

Cerchiamo ora una base di $f_h(V)$. Sappiamo che $f_h(V)$ è generata dall'immagine, tramite f_h, di una base di V, dunque calcoliamoci $f_h(v_1)$ e $f_h(v_2)$:

$$f_h(v_1) = \begin{pmatrix} 1 & h & 4 \\ 0 & 0 & 1 \\ 0 & h & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$f_h(v_2) = \begin{pmatrix} 1 & h & 4 \\ 0 & 0 & 1 \\ 0 & h & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} h+4 \\ 1 \\ h \end{pmatrix}$$

Per essere $f_h(V) \subseteq V$ deve essere che $f_h(v_1)$ e $f_h(v_2)$ appartengono a V. Osserviamo subito che $f_h(v_1)$ appartiene a V (è uno dei due elementi della base di V trovata), studiamo quando $f_h(v_2)$ appartiene a $V = Span(v_1, v_2)$. Ovvero studiamo quando esistono $\alpha, \beta \in \mathbb{Z}_5$ tali che:

$$\begin{pmatrix} h+4 \\ 1 \\ h \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Questo equivale a risolvere in $\mathbb{Z}_5$ il sistema lineare in α e β seguente:

$$\begin{pmatrix} \alpha \\ \beta \\ \beta \end{pmatrix} = \begin{pmatrix} h+4 \\ 1 \\ h \end{pmatrix}$$

Il sistema ammette soluzione se e solo se $h = 1$.

Esercizio 2.30. 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 può osservare che la matrice ha rango 2 e 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 2.22 che $Ker(f)$ ha dimensione 1 in quanto:

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

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 l $Ker(f)$, 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, dato un valore t a questa variabile si ottiene che x e y devono essere:

$$x = -\tfrac{1}{3}t \quad y = -\tfrac{2}{3}t$$

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

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

E dunque il vettore:

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

è una base di $Ker(f)$.

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, noto come *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 3 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':

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

Come sappiamo, $M' = RM$ 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

$$\begin{cases} x + 2y + 2z + 2t &= 1 \\ y + \frac{4}{3}z - \frac{4}{3}t &= -2 \\ 2z - \frac{12}{7}t &= -18 \\ 2t &= \frac{259}{139} \end{cases}$$

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, come suggerito dall'Esercizio 1.62, 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 $\mathbb{R}^n$. 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 proprio dalla matrice M'. Ricordiamo (vedi Teorema 2.20) che il rango di ϕ, cioè la dimensione di $Imm\ \phi$, è uguale a k. Dunque, per il Teorema 2.22 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 $p = \begin{pmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{pmatrix} \in S$. Vogliamo dimostrare che $p \in v + S_o$. Osserviamo che $c_1, c_2, \dots, c_n$ 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).*

Osservazione 3.6. Invitiamo il lettore a considerare la "somiglianza" delle osservazioni su S qui sopra con quanto abbiamo visto nella prima parte del corso a riguardo dell'insieme delle soluzioni di una equazione diofantea lineare: anche in quel caso, se l'equazione diofantea non è omogenea, l'insieme delle soluzioni è dato da un "traslato" dell'insieme delle soluzioni della equazione omogenea associata.

Gli esercizi del prossimo paragrafo permetteranno di mettere in pratica il metodo di Gauss. Il lettore potrà inoltre trovare alla pagina web [**AlgGauss**] un 'risolutore' di sistemi lineari, basato sulla riduzione di Gauss, che mostra, a fini didattici, le mosse utilizzate.

2. Altri esercizi

Esercizio 3.7. Discutere la risolubilità, ed eventualmente trovare tutte le soluzioni, del seguente sistema a coefficienti in $\mathbb{Z}_5$:

$$\begin{cases} x_1 - x_2 + x_3 + x_4 = 1 \\ x_2 - x_4 = 0 \\ x_3 + x_4 = 1 \end{cases} \tag{2.1}$$

Svolgimento. La matrice completa associata al sistema 2.1 è:

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

che è già in forma a scalini per righe. Gli scalini sono 3, e il sistema è risolubile (il rango della matrice incompleta è 3 e dunque è uguale al rango della matrice completa). L'unica variabile libera è x_4, quindi troveremo le soluzioni del sistema in funzione di x_4 e avremo 5 soluzioni distinte, una per ogni scelta possibile di x_4 in $\mathbb{Z}_5$:

$$\begin{cases} x_1 = x_4 - (1 - x_4) - x_4 + 1 \\ x_2 = x_4 \\ x_3 = 1 - x_4 \end{cases} \longrightarrow \begin{cases} x_1 = x_4 \\ x_2 = x_4 \\ x_3 = 1 - x_4 \end{cases}$$

Le soluzioni del sistema sono dunque del tipo i vettori $(x_4, x_4, 1-x_4, x_4)$ (appartenenti a $(\mathbb{Z}_5)^4$) per ogni scelta di x_4 in $\mathbb{Z}_5$.

Esercizio 3.8. Discutere la risolubilità, ed eventualmente trovare tutte le soluzioni, del seguente sistema a coefficienti in $\mathbb{Z}_7$:

$$\begin{cases} x_1 + x_2 - x_3 = 1 \\ x_1 + x_3 = 0 \\ x_1 + x_2 + 6x_3 = 0 \end{cases} \tag{2.2}$$

Svolgimento. La matrice completa associata al sistema 2.2 è:

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

Portiamola in forma a scalini con operazioni elementari di riga (ribadiamo che la scelta delle operazioni da effettuare non è univoca):

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

Il sistema associato alla matrice A_2 non ha soluzioni, come si vede osservando l'ultima riga, quindi il sistema 2.2 non ha soluzioni.

Esercizio 3.9. Discutere la risolubilità, ed eventualmente trovare tutte le soluzioni, del seguente sistema a coefficienti in $\mathbb{Q}$:

$$\begin{cases} x_1 - 3x_2 + x_3 + 2x_4 = 0 \\ 2x_1 - 6x_2 + x_3 + 5x_4 = 1 \\ 3x_1 - 9x_2 + 2x_3 + 10x_4 = 4 \end{cases} \tag{2.3}$$

Svolgimento. La matrice completa associata al sistema 2.3 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 sistema associato ad A_3 è risolubile, e 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 2.3):

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

Perciò le soluzioni del sistema 2.3 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 3.10. Trovare tutte le soluzioni del seguente sistema lineare omogeneo a coefficienti in $\mathbb{Z}_{11}$:

$$\begin{cases} 6x + y + 4z = 0 \\ 7x + 8y + 8z = 0 \\ 10x + y + z = 0 \\ 2x + y + 7z = 0 \end{cases} \tag{2.4}$$

Svolgimento. Possiamo limitarci a ridurre a scalini la matrice incompleta corrispondente a 2.4 in quanto l'ultima colonna, che ha tutti i coefficienti uguali a 0, rimarrà immutata dopo ogni operazione di riga:

$$B = \begin{pmatrix} 6 & 1 & 4 \\ 7 & 8 & 8 \\ 10 & 1 & 1 \\ 2 & 1 & 7 \end{pmatrix}$$

Per portare la matrice a scalini in questo caso risolviamo alcune congruenze. Per prima cosa troviamo l'inverso di 7 in $\mathbb{Z}_{11}$: tale inverso è 8 perché $7 \cdot 8 = 56 \equiv 1\ (11)$. Allora moltiplichiamo la seconda riga per $8 \cdot 6$, di modo che il suo primo coefficiente a sinistra sia uguale a 6, dopodiché sottraiamo la prima riga alla seconda:

$$B \to B_1 = \begin{pmatrix} 6 & 1 & 4 \\ 0 & 9 & 6 \\ 10 & 1 & 1 \\ 2 & 1 & 7 \end{pmatrix}$$

L'inverso di 10 in $\mathbb{Z}_{11}$ è ovviamente 10 stesso (ossia -1) quindi moltiplichiamo la terza riga per $-1 \cdot 6$ e poi sottraiamo la prima riga alla terza:

$$B_1 \to B_2 = \begin{pmatrix} 6 & 1 & 4 \\ 0 & 9 & 6 \\ 0 & 4 & 1 \\ 2 & 1 & 7 \end{pmatrix}$$

Altri simili passaggi portano alla seguente forma a scalini:

$$B_3 = \begin{pmatrix} 6 & 1 & 4 \\ 0 & 9 & 6 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

La matrice B_3 ha tre scalini, e il sistema (omogeneo) ad essa associato ha un'unica soluzione che è la soluzione identicamente nulla: $(0, 0, 0)$.

Esercizio 3.11. Discutere la risolubilità del seguente sistema a coefficienti in $\mathbb{Z}_{11}$ in dipendenza del parametro $\lambda \in \mathbb{Z}_{11}$:

$$\begin{cases} 3x + 2y - 5z = 6 \\ 5x + (2 + \lambda)y - 2z = 4 \\ 9x + 5y - 3z = 3\lambda \end{cases} \tag{2.5}$$

Svolgimento. La matrice completa associata al sistema 2.5 è:

$$A = \begin{pmatrix} 3 & 2 & -5 & 6 \\ 5 & 2+\lambda & -2 & 4 \\ 9 & 5 & -3 & 3\lambda \end{pmatrix}.$$

Portiamola in forma a scalini; potremmo per esempio fare le seguenti mosse:

$$A \to A_1 = \begin{pmatrix} 3 & 2 & -5 & 6 \\ 0 & 8+5\lambda & -5 & 3 \\ 9 & 5 & -3 & 3\lambda \end{pmatrix}$$

$$A_1 \to A_2 = \begin{pmatrix} 3 & 2 & -5 & 6 \\ 0 & 8+5\lambda & -5 & 3 \\ 0 & -1 & 1 & 3\lambda-7 \end{pmatrix}$$

A questo punto conviene scambiare la seconda e la terza riga,

$$A_2 \to A_3 = \begin{pmatrix} 3 & 2 & -5 & 6 \\ 0 & -1 & 1 & 3\lambda-7 \\ 0 & 8+5\lambda & -5 & 3 \end{pmatrix}$$

e sommare alla quarta riga la terza moltiplicata per $(8+5\lambda)$:

$$A_3 \to A_4 = \begin{pmatrix} 3 & 2 & -5 & 6 \\ 0 & -1 & 1 & 3\lambda-7 \\ 0 & 0 & 3+5\lambda & (3\lambda-7)\cdot(8+5\lambda)+3 \end{pmatrix}.$$

La matrice trovata è a scalini, ma dobbiamo capire quali sono i suoi coefficienti diversi da 0. Se $3+5\lambda \neq 0$ allora il sistema ha una e una sola soluzione.

Se invece $3+5\lambda = 0$ (ossia se λ è uguale a 6), abbiamo:

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

Quindi se $\lambda = 6$ il sistema 2.5 non ha soluzioni.

Esercizio 3.12. Consideriamo il seguente sistema a coefficienti in $\mathbb{Z}_p$:

$$\begin{cases} x+5y-6z=0 \\ 2x-8y-14z=12 \\ -x+7y+10z=0\lambda \end{cases} \tag{2.6}$$

Discutere al variare di p tra i numeri primi la risolubilità del sistema 2.6

Svolgimento. La matrice completa associata al sistema 2.6 è la seguente:

$$A = \begin{pmatrix} 1 & 5 & -6 & 0 \\ 2 & -8 & -14 & 12 \\ -1 & 7 & 10 & 0 \end{pmatrix}$$

Cerchiamo di portarla in forma a scalini:

$$A \to A_1 = \begin{pmatrix} 1 & 5 & -6 & 0 \\ 0 & -18 & -2 & 12 \\ -1 & 7 & 10 & 0 \end{pmatrix}$$

$$A_1 \to A_2 = \begin{pmatrix} 1 & 5 & -6 & 0 \\ 0 & -18 & -2 & 12 \\ 0 & 12 & 4 & 0 \end{pmatrix}.$$

Ora vorremmo moltiplicare la terza riga per 3 e sommarle la seconda riga moltiplicata per 2. Possiamo farlo, purché il campo in questione non sia $\mathbb{Z}_3$ (altrimenti staremmo moltiplicando la terza riga per 0, mossa non ammissibile). Quindi tratteremo a parte il caso $\mathbb{Z}_3$. Per tutti i $p \neq 3$ possiamo tranquillamente fare la mossa descritta e si ottiene:

$$A_2 \to A_3 = \begin{pmatrix} 1 & 5 & -6 & 0 \\ 0 & -18 & -2 & 12 \\ 0 & 0 & 8 & 24 \end{pmatrix}$$

Ora, se p non divide coefficienti -18 o 8, la matrice è in forma a scalini e ha tre scalini. Si può quindi dire che se $p \neq 2$ e $p \neq 3$ allora il sistema 2.6 ha una e una sola soluzione.

Se $p = 2$ la matrice A_3 ridotta modulo 2 diventa:

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

Quindi il sistema 2.6 ha due variabili libere e dunque 4 soluzioni in $(\mathbb{Z}_2)^3$, una per ogni scelta possibile della coppia di variabili libere (y, z). Ci resta da trattare il caso $p = 3$; torniamo alla matrice A_2, ovvero prima della sostituzione lineare che in $\mathbb{Z}_3$ non potevamo effettuare. Riducendo tale matrice modulo 3, si trova

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

che può essere portata nella seguente forma a scalini:

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

Quindi il sistema 2.6 in $(\mathbb{Z}_3)^3$ ha una variabile libera e dunque 3 soluzioni, una per ogni possibile scelta della variabile libera z.

Esercizio 3.13. Determinare per quale valori del parametro reale t il sistema lineare 2.7 nelle variabili x, y, z a coefficienti in $\mathbb{R}$ è risolubile e trovarne le soluzioni:

$$\begin{cases} x + y + tz = 1 \\ x + z = 0 \\ x + y + t^3 z = 3 \\ x + y + z = 0 \end{cases} \tag{2.7}$$

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.

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

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

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. 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 $[4] = [3] - (t^2+t+1)[4]$:

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

Il sistema 2.7 è dunque equivalente al sistema 2.8:

$$\begin{cases} x + y + tz = 1 \\ y = 0 \\ (t^3-1)z = 3 \\ 0 = 2 - t^2 - t \end{cases} \tag{2.8}$$

Per essere risolubile deve essere dunque $2 - t^2 - t = 0$, ovvero:

$$t = \frac{-1 \pm \sqrt{1+8}}{2} = \begin{matrix} \nearrow & 1 \\ \searrow & -2 \end{matrix}$$

Il sistema 2.7 può avere soluzioni solo per $t = 1$ o $t = -2$. Nel caso $t = 1$ però, sostituendo nel sistema 2.8, 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 2.8 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 2.7 ammette soluzioni solo nel caso $t = 2$. Per questo valore di t la soluzione del sistema è unica.

Esercizio 3.14. Trovare tutte le soluzioni in $\mathbb{Q}^4$ del seguente sistema lineare:

$$\begin{cases} x_1 - 3x_2 + x_3 + 2x_4 = 0 \\ 2x_1 - 6x_2 + x_3 + 5x_4 = 1 \\ 3x_1 - 9x_2 + 2x_3 + 10x_4 = 4 \end{cases} \tag{2.9}$$

Svolgimento La matrice dei coefficienti associata al sistema 2.10 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 2.10:

$$B \overset{[2]=[2]-2[1]}{\longrightarrow} B_1 = \begin{pmatrix} 1 & -3 & 1 & 2 & 0 \\ 0 & 0 & -1 & 1 & 1 \\ 3 & -9 & 2 & 10 & 4 \end{pmatrix} \overset{[3]=[3]-3[1]}{\longrightarrow} B_2 = \begin{pmatrix} 1 & -3 & 1 & 2 & 0 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & -1 & 4 & 4 \end{pmatrix}$$

$$B_2 \overset{[3]=[3]-[2]}{\longrightarrow} B_3 = \begin{pmatrix} 1 & -3 & 1 & 2 & 0 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 3 & 3 \end{pmatrix}$$

La matrice B_3 è a scalini e il sistema corrispondente ad essa, equivalente al sistema 2.9, è il seguente:

$$\begin{cases} x_1 - 3x_2 + x_3 + 2x_4 = 0 \\ -x_3 + x_4 = 1 \\ 3x_4 = 3 \end{cases} \tag{2.10}$$

Il sistema 2.10 ha una variabile libera (x_2). Dunque al variare del valore h di x_2 in $\mathbb{Q}$, si ha che le soluzioni del sistema 2.10sono gli elementi di $\mathbb{Q}^4$ del tipo: $(3h-2, h, 0, 1)$.

Esercizio 3.15. 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:

$$A \xrightarrow[[3]=2[1]+[3]]{[2]=[1]+[2]} A_1 = \begin{pmatrix} 1 & 1 & 10 \\ 0 & 2 & 12 \\ 0 & 3 & 21 \end{pmatrix} \xrightarrow{[3]=[3]-\frac{3}{2}[2]} A_2 = \begin{pmatrix} 1 & 1 & 10 \\ 0 & 2 & 12 \\ 0 & 0 & 3 \end{pmatrix}$$

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:

$$A \xrightarrow[[3]=4[1]-[3]]{[2]=[1]-[2]} A_1 = \begin{pmatrix} 1 & 1 & 1 & 10 \\ 0 & 2 & 0 & 8 \\ 0 & 6 & 3 & 39 \end{pmatrix} \xrightarrow{[3]=[3]-3[2]} A_2 = \begin{pmatrix} 1 & 1 & 1 & 10 \\ 0 & 2 & 0 & 8 \\ 0 & 0 & 3 & 15 \end{pmatrix}$$

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 potrebbe 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 3.16. 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 3.17. 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 3.18. 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 3.19. 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 3.20. 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 3.21. 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 2.22 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 3.22. 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 2.22 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 $Imm\ g$.

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 2.22.

Premettiamo una osservazine 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 $\{v_1, v_2, \ldots, v_n\}$ è una base di V e $\{w_1, \ldots, w_m\}$ è una base di W, allora $\{(v_1, O), (v_2, O), \ldots, (v_n, O), (O, w_1), \ldots, (O, w_m)\}$ 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 2.22.

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 2.23). 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 2.22 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 2.14) 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. 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 $dim\ U + dim\ W$ e vale anche il viceversa, ossia due sottospazi sono in somma diretta se e solo se vale $dim\ (U + W) = dim\ U + dim\ W$. 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 $dim\ U + dim\ W$).

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 4.5. 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.

Esercizio 4.6. Dimostrare che, dato un sottospazio vettoriale U di V che non sia V stesso, esiste sempre un complementare di U.

Suggerimento: prendere una base di U e completarla ad una base di V (vedi Teorema 2.14). I vettori che abbiamo aggiunto sono la base di uno spazio vettoriale complementare a U.

Osservazione 4.7. 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 $U_1, U_2, \ldots, U_k$ di uno spazio vettoriale V, si dice che tali sottospazi sono in somma diretta se, per ogni $i = 1, 2, \ldots, k$, 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 4.8. Dimostrare che, se $U_1, U_2, \ldots, U_k$ 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.

3. Altri esercizi

Esercizio 4.9. 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 4.10. 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 permette, data la matrice associata ad una applicazione lineare invertibile, 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. Chiameremo $End(V)$ 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 2.8), 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 2.22 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, basta 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. Visto che L è bigettiva, in particolare è iniettiva, dunque

$$L^{-1}(v + w) = L^{-1}(v) + L^{-1}(w)$$

vale se e solo se vale

$$L(L^{-1}(v + w)) = L\left(L^{-1}(v) + L^{-1}(w)\right).$$

Quest'ultima relazione si verifica facilmente. 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, $L\left(L^{-1}(v) + L^{-1}(w)\right) = L(L^{-1}(v)) + L(L^{-1}(w)) = v + w$.

□

Osservazione 5.2. Un endomorfismo invertibile è un caso particolare di *isomorfismo* (vedi Definizione 2.23). 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 (ossia è un isomorfismo) se e solo se ha rango n; l'applicazione inversa L^{-1} è anch'essa lineare.

Chiameremo $GL(V)$ il sottoinsieme di $End(V)$ costituito dagli endomorfismi invertibili. Come si verifica immediatamente, $GL(V)$ è un gruppo con la operazione di composizione fra funzioni $\circ$, chiamato il *gruppo generale lineare* su V.

Esercizio 5.3. Verificare che la composizione di due applicazioni $L, T \in GL(V)$ è ancora in $GL(V)$. Completare poi tutte le altre verifiche del fatto che $GL(V)$ è un gruppo.

Se fissiamo una base di V, ad ogni endomorfismo $L \in End(V)$ viene associata una matrice $[L] \in Mat_{n\times n}(\mathbb{K})$. 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 2.4 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 Osservazione1.60).

Dunque la matrice $[L]$ è invertibile e ha per inversa $[L^{-1}]$. Sempre applicando il Teorema 2.4 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 5.4. *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 2.20 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 $A \in Mat_{n\times n}(\mathbb{K})$. 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 5.4 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 deve avere la forma:

$$(I\ B)\,.$$

Affermiamo che la matrice $n \times n$ B 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 3 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 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 5.5. La relazione $UA = I$, ossia $BA = I$, ci dice solo che B è l'inversa sinistra di A. Ma possiamo mostrare facilmente che B coincide con l'inversa di A e dunque vale anche $AB = I$. Infatti l'inversa di A deve esistere (ed è unica[1]), visto che A è invertibile. Chiamiamola K; tale matrice K deve soddisfare per definizione $AK = KA = I$. Ora moltiplichiamo entrambi i membri della relazione $BA = I$, a destra, per K: $BAK = IK$. Usando la proprietà associativa del prodotto in $Mat_{n\times n}(\mathbb{K})$ otteniamo, visto che $AK = I$:

$$B = K,$$

ossia B è l'inversa di A.

Osservazione 5.6. 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 avremmo trovato meno di n scalini (cioè delle righe nulle).

Esercizio 5.7. 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.

[1] L'insieme delle matrici $n \times n$ invertibili è un gruppo rispetto alla moltiplicazione, come possiamo facilmente ricavare, per esempio, dal fatto che $GL(V)$ è un gruppo (vedi Esercizio 5.3). Ricordiamo che in un gruppo l'inverso di un elemento è unico: se b e c sono inversi di a e se indichiamo con e l'elemento neutro, dalla relazione $ab = e$ si ricava, moltiplicando a sinistra per c, $cab = ce$. A questo punto, utilizzando la proprietà associativa e il fatto che $ca = e$, si ottiene $b = c$.

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 2.4 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 5.8. *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 2.4:

$$\underset{v_i}{V} \xrightarrow{I} \underset{e_i}{V} \xrightarrow{L} \underset{e_i}{V} \xrightarrow{I} \underset{v_i}{V}$$

(sotto ogni copia di V il simbolo e_i oppure v_i indica quale base abbiamo scelto). Otteniamo[2]:

$$[I \circ L \circ I]_{\substack{v_1, v_2, \ldots, v_n \\ v_1, v_2, \ldots, v_n}} = [I]_{\substack{e_1, e_2, \ldots, e_n \\ v_1, v_2, \ldots, v_n}} [L]_{\substack{e_1, e_2, \ldots, e_n \\ e_1, e_2, \ldots, e_n}} [I]_{\substack{v_1, v_2, \ldots, v_n \\ e_1, e_2, \ldots, e_n}}$$

Visto che $I \circ L \circ I = L$ la formula appena ottenuta si può riscrivere come:

$$[L]_{\substack{v_1, v_2, \ldots, v_n \\ v_1, v_2, \ldots, v_n}} = M^{-1}[L]_{\substack{e_1, e_2, \ldots, e_n \\ e_1, e_2, \ldots, e_n}} M$$

□

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 1.59, ma il teorema appena dimostrato ha una grande importanza dal punto di vista teorico, come vedremo nei prossimi capitoli.

Per esempio, nell'Esercizio 1.71 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:

$$\mathcal{T}\left([L]_{\substack{v_1, v_2, \ldots, v_n \\ v_1, v_2, \ldots, v_n}}\right) = \mathcal{T}\left([L]_{\substack{e_1, e_2, \ldots, e_n \\ e_1, e_2, \ldots, e_n}}\right)$$

per ogni scelta delle basi $e_1, e_2, \ldots e_n$ e $v_1, v_2, \ldots v_n$.

[2]Il Teorema 2.4 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.

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 5.8 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 1.71), dunque:

$$\mathcal{T}\left(\left(M^{-1}[L]_{\substack{e_1,e_2,\dots,e_n\\ e_1,e_2,\dots,e_n}}\right)M\right) = \mathcal{T}\left(M\left(M^{-1}[L]_{\substack{e_1,e_2,\dots,e_n\\ e_1,e_2,\dots,e_n}}\right)\right) =$$

$$= \mathcal{T}\left(MM^{-1}[L]_{\substack{e_1,e_2,\dots,e_n\\ e_1,e_2,\dots,e_n}}\right) = \mathcal{T}\left([L]_{\substack{e_1,e_2,\dots,e_n\\ e_1,e_2,\dots,e_n}}\right)$$

Questa catena di uguaglianze conduce, come avevamo annunciato, a:

$$\mathcal{T}\left([L]_{\substack{v_1,v_2,\dots,v_n\\ v_1,v_2,\dots,v_n}}\right) = \mathcal{T}\left([L]_{\substack{e_1,e_2,\dots,e_n\\ e_1,e_2,\dots,e_n}}\right)$$

4. Altri esercizi

Esercizio 5.9. Trovare l'inversa della seguente matrice a coefficienti in $\mathbb{R}$:

$$\begin{pmatrix} 1 & -2 & 3 \\ 2 & -5 & 10 \\ 0 & 0 & 1 \end{pmatrix}$$

Esercizio 5.10. Trovare l'inversa della seguente matrice a coefficienti in $\mathbb{R}$:

$$\begin{pmatrix} -11 & 2 & 2 \\ -1 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix}$$

Esercizio 5.11. 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 5.12. 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.

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:

$$Det\ A = (-1)^{1+i}a_{i1}Det\ A_{i1} + (-1)^{2+i}a_{i2}Det\ A_{i2} + \cdots + (-1)^{i+n}a_{in}Det\ A_{in} \tag{1.1}$$

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:

$$Det\ A = (-1)^{1+j}a_{1j}Det\ A_{1j} + (-1)^{2+j}a_{2j}Det\ A_{2j} + \cdots + (-1)^{j+n}a_{nj}Det\ A_{nj} \tag{1.2}$$

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 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 $i_1, i_2, \ldots, i_k$) e k colonne fra le m colonne (diciamo che scegliamo le colonne $j_1, j_2, \ldots, j_k$).

[1]La regola prende nome dal matematico francese Pierre Frederic Sarrus (1798-1861).

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 $i \in \{i_1, i_2, \ldots, i_k\}$ e $j \in \{j_1, j_2, \ldots, j_k\}$.

Osservazione 6.7. Da una matrice $n \times m$ è possibile ricavare $\binom{n}{k}\binom{m}{k}$ minori di formato $k \times k$.

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 5.7) 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 sapevamo che 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, 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:

$$Det\left([L]_{\substack{v_1,v_2,\dots,v_n\\ v_1,v_2,\dots,v_n}}\right)=Det\left([L]_{\substack{e_1,e_2,\dots,e_n\\ e_1,e_2,\dots,e_n}}\right)$$

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 5.8 possiamo scrivere:

$$Det\left([L]_{\substack{v_1,v_2,\dots,v_n\\ v_1,v_2,\dots,v_n}}\right)=Det\left(M^{-1}[L]_{\substack{e_1,e_2,\dots,e_n\\ e_1,e_2,\dots,e_n}}M\right)$$

A questo punto, per il Teorema di Binet e per il Corollario 6.13, possiamo concludere che

[2]Il nome deriva dal matematico francese Jacques Philippe Marie Binet (1786-1856).

$$Det\begin{pmatrix} [L] & \\ & v_1, v_2, \ldots, v_n \\ & v_1, v_2, \ldots, v_n \end{pmatrix} = Det\left(M^{-1}\right) Det\begin{pmatrix} [L] & \\ & e_1, e_2, \ldots, e_n \\ & e_1, e_2, \ldots, e_n \end{pmatrix} Det\,(M) = Det\begin{pmatrix} [L] & \\ & e_1, e_2, \ldots, e_n \\ & e_1, e_2, \ldots, e_n \end{pmatrix}$$

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} nel Paragrafo 1 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. Dimostrare che il determinante della seguente *matrice di Vandermonde*[3] di formato $n \times n$ (con $n \geq 2$)

$$\begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ b_1 & b_2 & b_3 & \ldots & b_n \\ b_1^2 & b_2^2 & b_3^2 & \ldots & b_n^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ b_1^{n-1} & b_2^{n-1} & b_3^{n-1} & \ldots & b_n^{n-1} \end{pmatrix}$$

è uguale a $\prod_{i<j}(b_j - b_i)$, ossia al prodotto di tutte le possibili differenze fra i b_i (col segno opportuno, come risulta dalla formula). In particolare, se i b_i sono a due a due distinti, il determinante è diverso da 0 e la matrice è invertibile.

[3]Il nome deriva dal matematico francese Alexandre-Thophile Vandermonde (1735-1796)

Esercizio 6.15. Ripensare al fatto che una matrice di Vandermonde interviene nel metodo, illustrato nella prima parte del corso, per trovare una formula per le successioni definite per ricorrenza lineare e a coefficienti costanti.

Esercizio 6.16. 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 6.17. Sia L un'endomorfismo lineare dello spazio vettoriale V, sia A la matrice corrispondente all'endomorfismo A in una base fissata di A 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 zero** 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 $Ker\ T \supsetneq \{0\}$, 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, $\{v_1, v_2, \ldots, v_n\}$, composta solo da autovettori di T.

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

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

per certi autovalori λ_i (sui quali non abbiamo informazioni, per esempio potrebbero anche essere tutti uguali $\lambda_1 = \cdots = \lambda_n$).

Come sarebbe fatta la matrice

$$[T]_{\substack{\{v_1, v_2, \ldots, v_n\} \\ \{v_1, v_2, \ldots, v_n\}}}$$

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 $\{v_1, v_2, \ldots, v_n\}$, ossia

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

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

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

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, 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 $[T] - \lambda[I]$ 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: $P_T(t) = t^n + \ldots\ldots\ldots$ Queste dimostrazioni sono facoltative e consigliate!

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 5.8, le due matrici $[T]_{\substack{b\\b}}$ e $[T]_{\substack{b'\\b'}}$ sono legate dalla seguente relazione: esiste una matrice $[B]$ invertibile tale che

$$[B]^{-1}[T]_{\substack{b'\\b'}}\ [B] = [T]_{\substack{b\\b}}$$

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. In base 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 lo spazio V ha dimensione 3, ossia in cui la matrice $[T]$ è 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. □

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 4 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 $\lambda_1, \lambda_2, \ldots, \lambda_k$: a questo punto scopriamo chi sono i relativi autospazi $V_{\lambda_1}, V_{\lambda_2}, \ldots, V_{\lambda_k}$.
- PASSO 3. Un teorema (vedi Teorema 7.19) ci assicurerà che gli autospazi $V_{\lambda_1}, V_{\lambda_2}, \ldots, V_{\lambda_k}$ sono in somma diretta (vedi Paragrafo 2 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: $\lambda_1, \lambda_2, \ldots, \lambda_k$, tutti distinti fra loro. Vogliamo individuare gli autospazi $V_{\lambda_1}, V_{\lambda_2}, \ldots V_{\lambda_k}$.

Per questo basterà risolvere dei sistemi lineari: per ogni $i = 1, 2, \ldots, k$, l'autospazio V_{λ_i} è costituito per definizione dai vettori $v \in V$ tali che $T(v) = \lambda_i v$, ossia, scelta una base di V e dunque trovata la matrice $[T]$, dalle soluzioni del sistema lineare
$$([T] - \lambda_i I) \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ \cdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ \cdots \\ 0 \\ 0 \end{pmatrix}$$

3.3. Passo 3. Cominciamo col dimostrare il seguente teorema.

Teorema 7.17. *Dato un endomorfismo lineare $T : V \to V$, siano $\lambda_1, \lambda_2, \ldots, \lambda_k$ degli autovalori di T distinti fra loro. Consideriamo ora degli autovettori $v_1 \in V_{\lambda_1}, v_2 \in V_{\lambda_2}, \ldots, v_k \in V_{\lambda_k}$. Allora $\{v_1, v_2, \ldots, v_k\}$ 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..). Supponiamo di aver dimostrato che l'enunciato è vero fino a $k-1$ e cerchiamo di dimostrarlo per k. Supponiamo allora che valga:

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = O \tag{3.1}$$

Per mostrare che $\{v_1, v_2, \ldots, v_k\}$ 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 $\{v_1, v_2, \ldots, v_{k-1}\}$ 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 $\lambda_1, \lambda_2, \ldots, \lambda_k$ degli autovalori di T distinti fra loro. Gli autospazi $V_{\lambda_1}, V_{\lambda_2}, \ldots, V_{\lambda_k}$ sono in somma diretta.*

DIMOSTRAZIONE. Ricordiamo (vedi Paragrafo 2 del Capitolo 4) che dire che gli autospazi $V_{\lambda_1}, V_{\lambda_2}, \ldots, V_{\lambda_k}$ sono in somma diretta vuol dire che se ne prendo uno qualunque, diciamo V_{λ_1} tanto per fissare la notazione, la sua intersezione con la somma di tutti gli altri è banale, ossia

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

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

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

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

Dunque vale:

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

ossia

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

Questa è una combinazione lineare di autovettori relativi ad autovalori distinti, ma non è la combinazione lineare banale (infatti il coefficiente di v_1 è $1 \neq 0$). Dunque tali autovettori sarebbero linearmente dipendenti, assurdo perché contraddice il Teorema 7.17. □

Nelle ipotesi del teorema precedente sappiamo allora (vedi Esercizio 4.8), che la dimensione della somma degli autospazi è "la massima possibile", ossia

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

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

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

allora

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

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

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

e quindi è possibile trovare una base "buona", fatta da autovettori di T, insomma T è diagonalizzabile.
Per scrivere una simile base "buona", come sappiamo dal Paragrafo 2 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, ...,$dim\ V_{\lambda_k}$ coefficienti uguali a λ_k.

Il rango di T sarà uguale al numero dei coefficienti non zero che troviamo sulla diagonale, la dimensione del nucleo sarà uguale al numero dei coefficienti uguali a zero che troviamo sulla diagonale. In altre parole, se 0 non è un autovalore di T, allora $Ker\ T = \{O\}$; se invece 0 è un autovalore di T allora troveremo sulla diagonale $dim\ V_0$ coefficienti uguali a 0 - d'altronde avevamo già notato che $V_0 = Ker\ T$.

4. Il criterio della molteplicità algebrica e della molteplicità geometrica

Nel paragrafo precedente abbiamo trovato un criterio per decidere se un endomorfismo è diagonalizzabile o no. In questo paragrafo faremo una osservazione che ci permetterà di riformularlo in maniera più "operativa". 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

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

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 2.14), 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:

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

ossia ha una parte diagonale, dove troviamo λ_1 ripetuto $dim\ V_{\lambda_1}$ volte, λ_2 ripetuto $dim\ V_{\lambda_2}$ volte... λ_k ripetuto $dim\ V_{\lambda_k}$ volte, e poi sulle ultime r colonne, che corrispondono a $T(s_1), T(s_2), \ldots, T(s_r)$ 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 $i = 1, 2, \ldots, k$, $dim\ V_{\lambda_i} \leq a_i$. □

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 $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$.

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}.

[1]La molteplicità geometrica di una autovalore è sempre ≥ 1, visto che, per definizione, l'autospazio relativo a tale autovalore non è banale.

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

$$[T]_{\substack{v_1,v_2,v_3 \\ v_1,v_2,v_3}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix}$$

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 infatti 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 verificate, 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. 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 7.31. 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 7.32. 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 7.33. 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 7.34. 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:

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

Per quali valori di $a \in \mathbb{R}$ l'endomorfismo T_a è diagonalizzabile ?

Esercizio 7.35. 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 7.36. 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 7.37. 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 7.38 (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 7.39. 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 7.40. 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: $\lambda = 0$.

[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 7.41. 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 7.42. 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 7.43. 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.

Esercizio 7.44. 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 7.45 (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 7.46 (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, allota $T(V_\lambda) \subseteq V_\lambda$.

Esercizio 7.47. 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

Polinomi

Abbiamo già incontrato nei precedenti capitoli lo spazio vettoriale $\mathbb{K}[x]$ dato dai polinomi a coefficienti in un campo $\mathbb{K}$. Quando si considera anche la moltiplicazione fra polinomi, tale spazio vettoriale acquista una struttura di anello commutativo con identità. In questo capitolo ci proponiamo di studiare questa struttura, e ripartiamo 'dall'inizio', ossia dalla definizione di polinomio: i risultati ottenuti avranno utili applicazioni anche nell'algebra lineare.

1. Definizione, notazioni e uguaglianza tra polinomi

Definizione 8.1. Un polinomio nella variabile x, a coefficienti nel campo $\mathbb{K}$, è una espressione del tipo $p = p(x) = \sum_{i=0}^{\infty} a_i x^i$, con $n \in \mathbb{N}$ e gli $a_i \in \mathbb{K}$ tutti nulli eccetto un numero finito. Questo in particolare significa che esiste un $m \in \mathbb{N}$ tale che $a_n = 0$ per ogni $n > m$. Gli a_i si chiamano *coefficienti* del polinomio $p(x)$.

L'insieme dei polinomi a coefficienti nel campo $\mathbb{K}$ si indica con $\mathbb{K}[x]$.

Definizione 8.2 (Uguaglianza tra polinomi)**.** Diciamo che due polinomi

$$p(x) = \sum_{i=0}^{\infty} a_i x^i$$

e

$$q(x) = \sum_{i=0}^{\infty} b_i x^i$$

sono uguali se per ogni $i \in \mathbb{N}$ vale $a_i = b_i$.

Osservazione 8.3. Due osservazioni notazionali importanti per passare dalla definizione formale di polinomio a come poi si rappresentano concretamente i polinomi:
1) Solitamente nella scrittura di $p(x)$ si omettono tutti i termini con coefficiente uguale a zero (e questo permette di scrivere un polinomio per esteso: altrimenti non potremmo scrivere gli infiniti termini di un polinomio senza ricorrere alla scrittura *compatta* con la sommatoria).
A volte però (e lo vedremo già nella definizione delle operazioni tra polinomi) può essere utile ricordarsi che possiamo scrivere un polinomio, come ad esempio $f(x) = 3x^3+2x+1$, considerando anche *qualche* termine con coefficiente 0 (e dunque potremo scrivere $f(x)$ come $3x^3+0x^2+2x+1$, ma anche come $0x^4+3x^3+0x^2+2x+1$, eccetera...).
2) Nonostante nella definizione formale si scriva $p(x) = \sum_{i=0}^{\infty} a_i x^i$, di solito i polinomi si scrivono elencando in ordine decrescente di grado i suoi termini con coefficiente diverso da zero. Ovvero, il polinomio $f(x)$ dell'esempio sopra si scrive, di solito, $3x^3+2x+1$ e non $1+2x+3x^3$. È importante sottolineare che è una scelta puramente convenzionale (vedi Definizione 8.2 di uguaglianza tra polinomi).

Esempio 8.4. $f(x) = 3x^2 + \sqrt{2}x - 1$ è un esempio di polinomio a coefficienti in $\mathbb{R}$. Anche il polinomio $g(x) = 3x^3 - 5$ è un polinomio a coefficienti in $\mathbb{R}$. Il polinomio $g(x)$ appartiene anche a $\mathbb{Q}[x]$ (l'insieme dei polinomi a coefficienti nel campo $\mathbb{Q}$), il polinomio $f(x)$ non appartiene a $\mathbb{Q}[x]$.

Osservazione 8.5. Dalla definizione di polinomio (Definizione 8.1), e in particolare dal fatto che i coefficienti diversi da 0 sono in numero finito, segue che esiste $m \in \mathbb{N}$ tale che $a_n = 0$ se $n > m$. Ora, considerando che ogni insieme finito non vuoto ha massimo, ci chiediamo se esiste un indice m per cui $a_m \neq 0$ e $a_n = 0$ per ogni $n > m$ (cioè appunto m è il massimo indice per cui il coefficiente del polinomio è diverso da 0). La risposta è quasi sempre *sì, questo indice massimo esiste*. Diciamo quasi sempre, perché in un caso non riusciamo a trovarlo: se infatti consideriamo il polinomio con tutti i coefficienti uguali a 0, che rientra a pieno titolo nella definizione di polinomio e viene detto **polinomio nullo** e indicato con 0, non esiste nessun indice m per cui $a_m \neq 0$ e dunque l'esistenza del massimo *non si applica* perché l'insieme su cui dovremmo trovare il massimo è un insieme vuoto.

Definizione 8.6. Dato il polinomio **non nullo** $f(x) = \sum_{i=0}^{\infty} a_i x^i \in \mathbb{K}[x]$ consideriamo, se esiste, il più grande indice m per cui il coefficiente corrispondente a_m è diverso da zero. Tale m è detto **grado** del polinomio $f(x)$ ed è indicato con $deg(f(x))$:

$$deg(f(x)) = max\{m \in \mathbb{N} | a_m \neq 0\}$$

NOTAZIONE: Dopo aver definito il grado e visto che i coefficienti di un polinomio sono tutti nulli a parte un numero finito, per semplicità e per identificare immediatamente un generico polinomio di grado n useremo spesso la scrittura $p(x) = \sum_{i=0}^{n} a_i x^i$ in luogo di $p(x) = \sum_{i=0}^{\infty} a_i x^i$.
NOTAZIONE: Ricordiamo che, fin dal Capitolo 1, abbiamo indicato con $\mathbb{K}^{\leq m}[x]$ il sottoinsieme di $\mathbb{K}[x]$ contenente il polinomio 0 e tutti e soli i polinomi di grado minore o uguale di un fissato $m \in \mathbb{N}$:

$$\mathbb{K}^{\leq m}[x] = \{p(x) \in \mathbb{K}[x] | deg(p(x)) \leq m\} \cup \{0\}$$

Esempio 8.7. Consideriamo i polinomi dell'Esempio 8.4:

$$f(x) = 3x^2 + \sqrt{2}x - 1 \quad g(x) = 3x^3 - 5$$

Il grado di $f(x)$ è 2 e il grado di $g(x)$ è 3.

Osservazione 8.8. Il caso di $g(x)$, che può essere considerato un polinomio in $\mathbb{R}[x]$ ma anche in $\mathbb{Q}[x]$, può far sorgere la domanda se il grado dipenda dal campo $\mathbb{K}$ su cui è considerato il polinomio. In questo caso sembrerebbe di no, ma che cosa direste del grado di $g(x)$ considerato come polinomio in $\mathbb{Z}_3[x]$?

Osservazione 8.9. Dalla definizione di grado e di uguaglianza tra polinomi segue immediatamente che condizione necessaria affinché due polinomi siano uguali è che abbiano lo stesso grado. Ovviamente la condizione suddetta non è sufficiente.

Abbiamo introdotto i polinomi come scritture formali: somme di prodotti tra coefficienti di un campo $\mathbb{K}$ e potenze di una variabile x, ma uno degli aspetti che più ci interessa di questi *oggetti matematici* è il fatto che ad essi sia associata una funzione da $\mathbb{K}$ in $\mathbb{K}$. Per questo introduciamo il concetto di funzione associata ad un polinomio e allo stesso tempo sottolineiamo anticipatamente come i due concetti -

polinomio e *funzione polinomiale associata* - strettamente correlati, sono però distinti. In particolare vedremo che, per certi campi $\mathbb{K}$, due polinomi diversi possono essere associati alla stessa funzione polinomiale.

Definizione 8.10. Ad ogni polinomio $f(x) = \sum_{i=0}^{n} a_i x^i$ a coefficienti in un campo $\mathbb{K}$ può essere associata una funzione, da $\mathbb{K}$ in $\mathbb{K}$, che che indicheremo ancora con f e chiameremo **funzione associata al polinomio** $f(x)$; tale funzione associa ad ogni c in $\mathbb{K}$ il valore dell'espressione ottenuta sostituendo alla variabile x, l'elemento c di $\mathbb{K}$, ovvero:

$$\forall c \in \mathbb{K} \quad f(c) = \sum_{i=0}^{n} a_i c^i$$

Chiameremo il valore $f(c)$ di $\mathbb{K}$ **valutazione del polinomio** $f(x)$ **in** c.

Esempio 8.11. Il polinomio $3x^2 - 5x + 1$ di $\mathbb{R}[x]$ valutato nell'elemento 2 è uguale a $3 \cdot 2^2 - 5 \cdot 2 + 1 = 3$.

Osservazione 8.12. Se due polinomi $f(x) = \sum_{i=0}^{n} a_i x^i$ e $g(x) = \sum_{j=0}^{m} b_j x^j$ di $\mathbb{K}[x]$ sono uguali allora la funzione associata ad essi è uguale. Infatti i due polinomi hanno lo stesso grado $d = n = m$ e per ogni $i \leq d$ si ha che $a_i = b_i$, dunque per ogni $c \in \mathbb{K}$:

$$f(c) = \sum_{i=0}^{d} a_i c^i = \sum_{i=0}^{d} b_i c^i = g(c)$$

Si potrebbe pensare che sia vero anche il viceversa, ovvero che se le funzioni associate a due polinomi $f(x)$ e $g(x)$ sono uguali allora i polinomi $f(x)$ e $g(x)$ sono uguali, ma vedremo che questo non sempre è vero.

2. Somma, prodotto e divisione euclidea tra polinomi

Definiti *gli oggetti* polinomi come scritture formali, vogliamo definire anche come operare tra essi, in particolare come sommare e moltiplicare tra loro due polinomi. Definiremo anche, alla stregua di quel che abbiamo fatto in $\mathbb{Z}$, la divisione euclidea.

Definiamo preliminarmente la somma e il prodotto di un polinomio qualsiasi $p(x)$ in $\mathbb{K}[x]$ con il polinomio nullo nella maniera più *naturale*:

$$p(x) + 0 = 0 + p(x) = p(x)$$
$$p(x) \cdot 0 = 0 \cdot p(x) = 0$$

A questo punto possiamo definire la somma e il prodotto tra due generici polinomi non nulli $p(x)$ e $q(x)$ di grado rispettivamente n e m (non necessariamente diversi: ovvero vogliamo definire operazioni anche tra polinomi dello stesso grado). Ma come farlo? La scelta come si può intuire non è univoca: abbiamo *degli oggetti nuovi* (i polinomi) e possiamo definire le operazioni come vogliamo, ad esempio fissando che la somma di due polinomi di grado n è sempre uguale a x^n. È importante sottolineare come molte delle scelte fatte dai matematici (ad esempio nell'introduzione di qualche operazione e nelle *regole* che seguono tale introduzione) siano *strategiche* e finalizzate a qualche scopo. Nel caso dei polinomi, ad esempio, vorremmo riuscire a definire le operazioni in modo che continuino a valere tutte quelle proprietà a cui siamo abituati (commutatività, associatività, distributività,...). Inoltre, essendo $\mathbb{K}$ un sottoinsieme di $\mathbb{K}[x]$ (contenente il polinomio 0 e tutti i polinomi di grado 0), vorremmo che quando dobbiamo sommare (o moltiplicare) due elementi di $\mathbb{K}$ non

ci sia bisogno di chiedersi se stiamo facendo la somma (o la moltiplicazione) in $\mathbb{K}$ o in $\mathbb{K}[x]$, ma che le due operazioni: somma (moltiplicazione) in $\mathbb{K}$ e somma (moltiplicazione) in $\mathbb{K}[x]$ coincidano sugli elementi di $\mathbb{K}$. Per esempio la definizione di somma accennata sopra non risponde a questo criterio.

Definizione 8.13 (Somma di polinomi)**.** Dati i due polinomi di $\mathbb{K}[x]$:

$$\begin{aligned} p(x) &= \textstyle\sum_{i=0}^{n} a_i x^i \\ q(x) &= \textstyle\sum_{j=0}^{m} b_j x^j \end{aligned}$$

si definisce il **polinomio somma** di $p(x)$ e $q(x)$, che indicheremo con $(p+q)(x)$ come il polinomio il cui termine di grado i ha coefficiente uguale alla somma dei coefficienti di grado i di $p(x)$ e $q(x)$, ovvero[1]:

$$(p+q)(x) \underbrace{=}_{definizione} \sum_{i=0}^{max(m,n)} (a_i + b_i)x^i$$

Esercizio 8.14. Dimostrare che la somma tra polinomi è una funzione da $\mathbb{K}[x]$ in se stesso. Ovvero che dati $f(x), g(x)$ in $\mathbb{K}[x]$ il polinomio somma $(f+g)(x)$ appartiene a $\mathbb{K}[x]$.

La definizione di moltiplicazione tra polinomi appare a prima vista molto più cervellotica di quella data per la somma, e sicuramente è molto più complicato capire dalla scrittura formale del polinomio prodotto qual è quel polinomio. In realtà però anche la definizione di prodotto di polinomi è piuttosto *naturale* e nasce, come anticipato, dalla volontà di voler conservare le proprietà note delle operazioni.

Volendo moltiplicare due polinomi ci troveremo a moltiplicare tra loro anche polinomi di un solo termine (chiamati anche **monomi**) per esempio $4x$ per $2x^3$ e definiamo questo prodotto come $8x^4$ (ovvero mantenendo le note proprietà delle potenze anche per la moltiplicazione della variabile: $x^3 \cdot x = x^{3+1} = x^4$). Deciso questo abbiamo in realtà già quasi definito la moltiplicazione tra polinomi, il resto segue dalla volontà di conservare la proprietà distributiva del prodotto rispetto alla somma di polinomi. Quindi se vogliamo moltiplicare tra loro i polinomi $f(x) = 3x+1$ e $g(x) = x^3 + 3x + 1$, vogliamo che il risultato sia uguale a sommare tra loro questi prodotti di singoli termini: $x^3 \cdot 3x$, $x^3 \cdot 1$, $3x \cdot 3x$, $3x \cdot 1$, $1 \cdot 3x$, $1 \cdot 1$, che sappiamo calcolare.

Possiamo dunque dare la definizione di prodotto tra polinomi, sperando che, a questo punto, sia un po' meno oscura.

Definizione 8.15 (Prodotto di polinomi)**.** Dati i due polinomi di $\mathbb{K}[x]$:

$$\begin{aligned} p(x) &= \textstyle\sum_{i=0}^{n} a_i x^i \\ q(x) &= \textstyle\sum_{j=0}^{m} b_j x^j \end{aligned}$$

si definisce il **polinomio prodotto** di $p(x)$ e $q(x)$, che indicheremo con $(p \cdot q)(x)$, come segue:

$$(p \cdot q)(x) \underbrace{=}_{definizione} \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j x^{i+j}$$

[1]Osserviamo come dalla scrittura formale del polinomio somma, emerga che facciamo la somma fino al massimo tra m e n. Questo perché da quel punto in poi tutte le somme sono uguali a 0 e quindi ininfluenti per la scrittura del polinomio. Osserviamo altresì che per fare queste somme fino al massimo tra m e n stiamo pensando i due polinomi scritti con tutti i coefficienti - anche quelli nulli - dall'indice 0 all'indice uguale al massimo tra m e n.

Esercizio 8.16. Dati due polinomi $p(x) = \sum_{i=0}^{n} a_i x^i$ e $q(x) = \sum_{j=0}^{m} b_j x^j$ a coefficienti in $\mathbb{K}$, verificare che la valutazione del polinomio somma (prodotto) è la somma (prodotto) delle valutazioni di $p(x)$ e $q(x)$. Ovvero che per ogni c di $\mathbb{K}$ si ha che:

$$(p+q)(c) = p(c) + q(c)$$
$$(p \cdot q)(c) = p(c) \cdot q(c)$$

Ovvero che la valutazione del polinomio somma (prodotto) è uguale alla somma (prodotto) delle valutazioni dei due singoli polinomi.

Definite somma e prodotto su $\mathbb{K}[x]$ possiamo valutare il grado della somma e del prodotto di due polinomi. Come vedremo, mentre il grado del prodotto è determinato dal grado dei due polinomi, sul grado della somma possiamo solo dire che non supererà un certo grado.

Proposizione 8.17 (proprietà del grado)**.** *Consideriamo $f(x)$ e $g(x)$ diversi dal polinomio nullo*[2]*, allora:*

$$deg(\underbrace{f(x) + g(x)}_{(f+g)(x)}) \leq \max(deg(f(x)), deg(g(x)))$$
$$deg(\underbrace{f(x) \cdot g(x)}_{(f \cdot g)(x)}) = deg(f(x)) + deg(g(x))$$

Dimostrazione. Supponiamo $f(x)$ e $g(x)$ di grado rispettivamente m e n:

$$f(x) = \sum_{i=0}^{m} a_i x^i \quad a_m \neq 0$$
$$g(x) = \sum_{j=0}^{n} b_j x^j \quad b_n \neq 0$$

Osserviamo che:

(1) Se $k > max(m,n)$ allora $a_k + b_k = 0$ (ovvero tutti i coefficienti di $(f+g)(x)$ di indice maggiore del massimo tra m e n sono nulli) per cui

$$deg((f+g)(x)) \leq max(m,n)$$

(2) Applicando la definizione di polinomio prodotto si ha

$$(f \cdot g)(x) = a_m b_n x^{n+m} + \text{termini di grado inferiore}$$

Infatti non ci possono essere termini di grado superiore a $m+n$ in quanto $a_i \cdot b_j = 0$ se $i > m$ o $j > n$. A questo punto basta osservare che essendo $a_m \neq 0$ e $b_n \neq 0$ per ipotesi sul grado di $f(x)$ e $g(x)$, allora $a_m b_n \neq 0$, dunque:

$$deg(f \cdot g)(x) = m + n = deg(f(x)) + deg(g(x))$$

□

Dopo aver introdotto somma e prodotto per l'insieme $\mathbb{K}[x]$ dei polinomi a coefficienti nel campo $\mathbb{K}$, è bene verificare che effettivamente con le definizioni **scelte** valgono le proprietà che ci interessavano. Dimostrare dunque per esercizio il seguente teorema, valido qualsiasi sia il campo $\mathbb{K}$:

[2]Nel caso di operazioni con il polinomio nullo, la determinazione del grado è immediata. Il prodotto di un polinomio per il polinomio nullo è uguale al polinomio nullo e quindi il grado non è definito. La somma di un polinomio $f(x)$ con il polinomio nullo è uguale ad $f(x)$ e quindi il grado della somma è quello di $f(x)$.

Teorema 8.18. $(\mathbb{K}[x], +, \cdot, 0, 1)$ *è un anello commutativo con identità* [3].

In questo paragrafo studieremo le (molte) analogie tra $\mathbb{K}[x]$ e l'anello degli interi $\mathbb{Z}$.

Innanzitutto stabiliamo, tramite le proprietà del grado sul prodotto tra polinomi evidenziate dalla Proposizione 8.17, quali polinomi $f(x) \in \mathbb{K}[x]$ hanno l'inverso per la moltiplicazione[4], ovvero per quali $f(x)$ esiste $g(x)$ tale che $f(x) \cdot g(x)$ è uguale al polinomio identità 1:

Proposizione 8.19. *In* $\mathbb{K}[x]$ *gli unici polinomi invertibili sono quelli di grado* 0*, ovvero tutte e sole le costanti diverse da* 0*.*

Dimostrazione. Il fatto che le costanti diverse da 0 siano invertibili in $\mathbb{K}[x]$ è conseguenza del fatto che $\mathbb{K}$ è un campo[5]: dunque tutte le costanti diverse da 0 sono invertibili. Viceversa vogliamo mostrare che se $f(x)$ è invertibile allora è una costante diversa da 0. Sia dunque $f(x)$ invertibile, allora esiste $g(x)$ tale che $f(x) \cdot g(x) = 1$. Dalla Proposizione 8.17 segue che:

$$0 = deg(1) = deg(f(x) \cdot g(x)) = deg(f(x)) + deg(g(x))$$

Ovvero $deg(f(x)) + deg(g(x)) = 0$, da cui necessariamente segue che (essendo il grado un numero naturale):

$$deg(f(x)) = deg(g(x)) = 0$$

cioè $f = a_0 \neq 0$ e $g = a_0^{-1}$. □

Osservazione 8.20. Osserviamo che la scelta di usare come insieme dei coefficienti un campo ha anche la conseguenza (oltre a quella già discussa del fatto che tutte le costanti diverse da 0 hanno inverso in $\mathbb{K}[x]$) che in $\mathbb{K}[x]$ non ci sono divisori dello zero.

L'analogia tra l'anello degli interi $\mathbb{Z}$ e l'anello $\mathbb{K}[x]$ dei polinomi a coefficienti nel campo $\mathbb{K}$ diventa molto evidente con la definizione di una divisione euclidea (divisione con resto) anche tra polinomi. Come vedremo, nella divisione euclidea tra polinomi gioca un ruolo cruciale il concetto di grado di un polinomio.

Definizione 8.21 (Divisione euclidea)**.** Siano $p(x), s(x) \in \mathbb{K}[x]$, con $\mathbb{K}$ un campo e $s(x) \neq 0$. Diciamo che $q(x), r(x) \in \mathbb{K}[x]$ sono **quoziente** e **resto** della divisione di $p(x)$ per $s(x)$ se $p(x) = q(x)s(x) + r(x)$ e inoltre $r(x) = 0$ oppure $deg(r(x)) < deg(s(x))$.

È naturale chiedersi se l'analogia con $\mathbb{Z}$ continua anche a riguardo dell'esistenza e unicità dei polinomi quoziente e resto appena definiti. Ovvero, dati due polinomi $p(x)$, $s(x)$ qualsiasi (con $s(x) \neq 0$) esistono sempre il quoziente e il resto del polinomio $p(x)$ diviso $s(x)$? Se esistono, sono unici? Ci risponde il seguente teorema:

---

[3] Sottolineiamo che 1 rappresenta il polinomio $f(x) = \sum_{i=0}^{\infty} a_i x^i$ con $a_0 = 1$ e $a_i = 0$ per ogni $i \neq 0$.

[4] In $\mathbb{Z}$, come sappiamo, gli unici elementi ad avere inverso moltiplicativo sono 1 e -1.

[5] Se considerassimo $A[x]$ con A anello ma non campo, questo non sarebbe più vero. Infatti A anello e non campo significa che esiste almeno un $a \in A$ diverso da 0 non invertibile. Dunque il polinomio a di $A[x]$ di grado 0, non è invertibile in $A[x]$ in quanto moltiplicandolo per qualsiasi altro polinomio di grado 0 non sarà mai uguale 1 per ipotesi, e moltiplicandolo per polinomi $g(x)$ di grado $m > 0$ non sarà mai 1: per la Proposizione 8.17 si ha che $deg(a \cdot g(x)) = 0 + m = m$, mentre il polinomio 1 ha grado 0.

Teorema 8.22 (Teorema sulla divisione euclidea tra polinomi)**.** *Dati $p(x), s(x) \in \mathbb{K}[x]$, con $\mathbb{K}$ un campo e $s(x) \neq 0$, esistono e sono unici $q(x), r(x) \in \mathbb{K}[x]$* **quoziente** *e* **resto** *della divisione di $p(x)$ per $s(x)$.*

DIMOSTRAZIONE. Per prima cosa dimostriamo che, se esistono i polinomi quoziente e resto della divisione tra $p(x)$ e $q(x)$, questi polinomi sono unici. Poi dimostreremo che tali polinomi (quoziente e resto) esistono sempre.
Unicità: supponiamo che esistano $q_1(x), q_2(x), r_1(x), r_2(x)$ tali che

$$p(x) = q_1(x)s(x) + r_1(x), \quad p(x) = q_2(x)s(x) + r_2(x)$$

con $r_1(x) = 0$ oppure $deg(r_1(x)) < deg(s(x))$ e $r_2(x) = 0$ oppure $deg(r_2(x)) < deg(s(x))$. Facendo la differenza fra le due equazioni precedenti otteniamo:

$$(q_1(x) - q_2(x))s(x) = -(r_1(x) - r_2(x))$$

Se $q_1(x) \neq q_2(x)$, allora $q_1(x) - q_2(x) \neq 0$ e quindi il polinomio a primo membro ha grado pari a $deg(s(x)) + deg(q_1(x) - q_2(x)) \geq deg(s(x))$. A secondo membro però abbiamo la somma di due polinomi ($-r_1(x)$ e $r_2(x)$) ciascuno dei quali è 0 oppure ha grado strettamente più piccolo del grado di $s(x)$: dunque il polinomio somma è 0 oppure ha grado strettamente più piccolo di quello di $s(x)$ (vedi Proposizione 8.17). Dall'Osservazione 8.9 segue dunque che se $q_1(x) \neq q_2(x)$ allora $(q_1(x) - q_2(x))s(x)$ e $-(r_1(x) - r_2(x))$ non sono uguali.

Deve quindi valere $q_1(x) = q_2(x)$, da cui:

$$-(r_1(x) - r_2(x)) = (q_1(x) - q_2(x))s(x) = 0 \cdot s(x) = 0 \rightarrow r_1(x) = r_2(x).$$

Esistenza: procediamo per induzione su $deg(p(x))$.
Passo base: $deg(p(x)) = 0$ (ovvero $p(x)$ è una costante c di $\mathbb{K}$ diversa da 0). Se anche $s(x)$ è una costante a di $\mathbb{K}$ diversa da zero allora si ha:

$$c = \underbrace{\frac{c}{a}}_{q(x)} \cdot a + \underbrace{0}_{r(x)}$$

Se invece $s(x)$ ha grado maggiore di 0, allora $deg(p(x)) < deg(s(x))$. In questo caso per avere polinomio quoziente e resto con le proprietà cercate, basta considerare $q(x) = 0$ e $r(x) = p(x)$:

$$p(x) = \underbrace{0}_{q(x)} \cdot s(x) + \underbrace{p(x)}_{r(x)} \text{ e } deg(r(x)) = deg(p(x)) < deg(s(x))$$

Passo induttivo: supponiamo che per ogni polinomio $p(x)$ di $\mathbb{K}[x]$ di grado minore o uguale a n esistano il polinomio quoziente e resto della divisione per qualsiasi polinomio $s(x)$ diverso da 0 di $\mathbb{K}[x]$ e mostriamo che da questo segue che anche per i polinomi di grado $n+1$ esistono quoziente e resto. Sia dunque $deg(p(x)) = n+1$. Sia $a_{n+1} \in \mathbb{K}$ il coefficiente di grado massimo di $p(x)$ e b il coefficiente di grado massimo di $s(x)$ e consideriamo il seguente polinomio $p_1(x)$ di $\mathbb{K}[x]$:

$$p_1(x) = p(x) - \underbrace{(\frac{a_{n+1}}{b} x^{deg(p(x)) - deg(s(x))})}_{t(x)} s(x)$$

Il coefficiente di grado $n+1$ di $p_1(x)$ è dato dalla somma del coefficiente di grado $n+1$ di $p(x)$ (ovvero a_{n+1}) con il coefficiente di grado $n+1$ di $t(x) \cdot s(x)$ che è:

$$-\frac{a_{n+1}}{b} \cdot b = -a_{n+1}$$

Dunque il coefficiente di grado $n+1$ di $p_1(x)$ è 0, cioè $p_1(x)$ ha grado strettamente più basso del grado di $p(x)$. Da questo segue, per ipotesi induttiva, che esistono $q_1(x), r_1(x)$ tali che

$$p_1(x) = q_1(x)s(x) + r_1(x)$$

con $r_1(x) = 0$ oppure $deg(r_1(x)) < deg(s(x))$. Si ha quindi che

$$p(x) = \underbrace{(q_1(x)s(x) + r_1(x))}_{p_1(x)} + t(x)s(x) = (q_1(x) + t(x))s(x) + r_1(x)$$

e $r_1(x)$ ha la proprietà richiesta per essere il resto della divisione euclidea. □

La dimostrazione del teorema fornisce anche un algoritmo per calcolare quoziente e resto in una divisione fra polinomi; si tratta in sostanza della "divisione fra polinomi" che probabilmente avete già visto alle scuole superiori. Il punto di partenza è confrontare i termini di grado maggiore dei due polinomi; assume dunque particolare rilevanza il coefficiente di grado massimo di un polinomio:

Definizione 8.23. Dato un polinomio $f(x) = \sum_{i=0}^{n} a_i x^i$ di grado n in $\mathbb{K}[x]$ chiameremo **coefficiente direttivo di** $f(x)$ (o principale) il coefficiente a_n di $f(x)$. Indicheremo il coefficiente direttivo di un polinomio $f(x)$ con la notazione $L.C.(f)$. Chiameremo il termine $a_n x^n$ **termine principale** del polinomio $f(x)$ e lo indicheremo con la notazione[6] $L.T.(f)$.

Esempio 8.24. Dato il polinomio $f(x) = -4x^4 + 3x^2 - 1$, il coefficiente direttivo di $f(x)$ è $L.C.(f) = -4$.

Definizione 8.25. Un polinomio con coefficiente direttivo uguale a 1 di dice **monico**.

A questo punto vediamo con qualche esempio come funziona l'algoritmo di divisione tra polinomi.

Esempio 8.26. In $\mathbb{Q}[x]$, dividere il polinomio $p(x) = 2x^4 + x^3 + -x^2 + 1$ per $s(x) = 3x^2 + 1$. L'algoritmo di divisione tra polinomi è simile all'algoritmo di divisione tra numeri interi: definiamo $q(x)$ e $r(x)$ per approssimazioni successive, seguendo il metodo di dimostrazione del teorema.

$$p_1(x) = \underbrace{2x^4 + x^3 + -x^2 + 1}_{p(x)} - \underbrace{\frac{2}{3}x^2}_{\frac{L.C.(p)}{L.C.(s)} x^{deg(p(x)) - deg(s(x))}} \cdot \underbrace{(3x^2 + 1)}_{s(x)} = x^3 - \frac{5}{3}x^2 + 1$$

In pratica il polinomio $p_1(x)$ è un *resto* (ma attenzione non *il polinomio resto* della divisione tra $p(x)$ e $s(x)$, che deve avere grado strettamente minore di $s(x)$) di un primo passaggio di divisione tra $p(x)$ e $s(x)$.
Iteriamo il procedimento, sostituendo $p_1(x)$ a $p(x)$ (questo passaggio dovrebbe ricordare qualcosa di analogo nella divisione euclidea tra interi), e definendo dunque un polinomio $p_2(x)$ come segue:

$$p_2(x) = p_1(x) - \underbrace{\frac{1}{3}x}_{\frac{L.C.(p_1)}{L.C.(s)} x^{deg(p_1(x)) - deg(s(x))}} \cdot (3x^2 + 1) = -\frac{5}{3}x^2 - \frac{1}{3}x + 1$$

[6] I due acronimi usati per le notazioni seguono dai termini inglesi *leading coefficient* e *leading term*.

Quando ci fermeremo in questa iterazione? Quando uno dei polinomi $p_i(x)$ è 0 oppure ha grado strettamente minore di $s(x)$: tale polinomio sarà il polinomio resto della divisione tra $q(x)$ e $s(x)$ e troveremo il polinomio quoziente *risalendo* le divisioni fatte.
Torniamo al nostro esempio: $p_2(x)$ non ha grado strettamente minore di $s(x)$, dunque ripetiamo una terza volta il procedimento con $p_2(x)$ al posto di $p(x)$

$$p_3(x) = p_2(x) + \frac{5}{9}(3x^2+1) = -\frac{1}{3}x + \frac{14}{9}$$

Essendo $p_3(x)$ di grado minore a $s(x)$, $p_3(x)$ è il polinomio resto $r(x)$ che cercavamo, e abbiamo

$$\begin{aligned} &p(x) = p_1(x) + \left(\tfrac{2}{3}x^2\right)\cdot s(x) = \\ &= p_2(x) + \left(\tfrac{1}{3}x + \tfrac{2}{3}x^2\right)\cdot s(x) = \\ &= p_3(x) + \left(-\tfrac{5}{9} + \tfrac{1}{3}x + \tfrac{2}{3}x^2\right)\cdot s(x) \end{aligned}$$

Riassumendo, abbiamo calcolato

$$p(x) = \underbrace{\left(-\frac{5}{9} + \frac{1}{3}x + \frac{2}{3}x^2\right)}_{q(x)}\underbrace{(3x^2+1)}_{s(x)} + \underbrace{\left(-\frac{1}{3}x + \frac{14}{9}\right)}_{r(x)}$$

Osservazione 8.27. Osserviamo un aspetto importante dal punto di vista *operativo*: il metodo delle divisioni successive mostrato nell'esempio precedente termina sempre in un numero finito di passaggi (al più uguale al grado del polinomio $p(x)$ che si vuole dividere). Infatti all'i-esimo passaggio, il polinomio $p_i(x)$ o è 0, oppure il suo grado è strettamente minore di quello di $p_{i-1}(x)$. Dunque la successione dei gradi dei polinomi $p_i(x)$ è una successione strettamente decrescente di numeri naturali che parte da $deg(p(x))$.

Esempio 8.28. Calcoliamo quoziente e resto della divisione tra i due seguenti polinomi di $\mathbb{Q}[x]$:

$$f(x) = x^6 - 1 \quad g(x) = x^4 + x^3 + x^2 - 4x + 1$$

Come abbiamo visto nell'Esempio 8.26 bisogna confrontare i termini principali (ovvero quelli di grado massimo) dei due polinomi:

$$L.T.(x^6-1) = x^6 \quad L.T.(x^4+x^3+x^2-4x+1) = x^4$$

A questo punto *per cosa dobbiamo moltiplicare x^4 per arrivare ad x^6?* La risposta è x^2: dunque moltiplichiamo $g(x)$ per x^2 e il risultato lo sottraiamo da $f(x)$. Quello che otterremo sarà un polinomio di grado minore di 6. Continueremo fino a che non otteniamo 0 o un polinomio di grado minore a $g(x)$ (cioè fino a che non otteniamo il resto della divisione tra $f(x)$ e $g(x)$). Questo è quello che abbiamo fatto nell'Esempio 8.26, vediamo come farlo con una forma grafica appropriata:

$$\begin{array}{lllllll|l} x^6 & & & & & -1 & | & x^4+x^3+x^2-4x+1 \\ x^6 & +x^5 & +x^4 & -4x^3 & +x^2 & & | & x^2 \\ & -x^5 & -x^4 & +4x^3 & -x^2 & -1 & & \end{array}$$

A questo punto dobbiamo *confrontare* x^4 (il termine principale di $g(x)$) con $-x^5$ (il termine principale del polinomio $-x^5 - x^4 + 4x^3 - x^2 - 1$). Il secondo passaggio sarà quindi quello di moltiplicare $g(x)$ per $-x$:

$$\begin{array}{lllllll|l}
x^6 & & & & & & -1 & | & x^4+x^3+x^2-4x+1 \\
x^6 & +x^5 & +x^4 & -4x^3 & +x^2 & & & | & x^2-x \\
 & -x^5 & -x^4 & +4x^3 & -x^2 & & -1 & & \\
 & -x^5 & -x^4 & -x^3 & +4x^2 & -x & & & \\
 & & & 5x^3 & -5x^2 & +x & -1 & &
\end{array}$$

Il polinomio ottenuto è di grado minore di $g(x)$ quindi abbiamo terminato l'algoritmo di divisione tra $f(x)$ e $g(x)$ trovando il polinomio quoziente $q(x)$ e il polinomio resto $r(x)$:

$$f(x) = g(x) \cdot \underbrace{(x^2 - x)}_{q(x)} + \underbrace{(5x^3 - 5x^2 + x - 1)}_{r(x)}$$

3. Divisori e radici

Dimostrata la possibilità di eseguire la divisione euclidea tra polinomi, possiamo introdurre il concetto di divisore tra i polinomi e continuare così l'analogia con i concetti e le terminologie usate per l'anello $\mathbb{Z}$.

Definizione 8.29. Dati due polinomi $s(x), p(x) \in \mathbb{K}[x]$, diciamo che il polinomio $s(x)$ **divide** il polinomio $p(x)$ se esiste un polinomio $q(x) \in \mathbb{K}[x]$ tale che $p(x) = q(x)s(x)$. In modo equivalente, possiamo dire che $s(x)$ divide $p(x)$ se il resto della divisione di $p(x)$ per $s(x)$ è uguale a zero. Per indicare che $s(x)$ divide $p(x)$ useremo la notazione $s(x)|p(x)$. Se $s(x)$ divide $p(x)$ si dice che $s(x)$ è un **fattore** del polinomio $p(x)$ o anche che $p(x)$ è un **multiplo** di $s(x)$.

Esempio 8.30. Il polinomio $g(x) = x^2 - 3x + 2$ è un divisore del polinomio $f(x) = 2x^3 - 4x^2 - 2x + 4$ infatti:

$$2x^3 - 4x^2 - 2x + 4 = (x^2 - 3x + 2) \cdot (2x + 2)$$

Esercizio 8.31. Se $f(x)$ divide $g(x)$ allora $deg(f(x)) \leq deg(g(x))$.

Definizione 8.32. Dato un polinomio $p(x) \in \mathbb{K}[x]$, diciamo che $\lambda \in \mathbb{K}$ è una **radice** del polinomio $p(x)$ se $p(\lambda) = 0$ (ricordando la Definizione 8.10, le radici di un polinomio sono gli zeri della funzione da $\mathbb{K}$ in $\mathbb{K}$ ad esso associata).

Osservazione 8.33. È importante osservare come un polinomio $p(x)$ di $\mathbb{K}[x]$ possa non avere radici in $\mathbb{K}$ ed averle in un campo $\mathbb{L}$ contenente $\mathbb{K}$ (d'altra parte, essendo i coefficienti di $p(x)$ elementi di $\mathbb{K} \subset \mathbb{L}$, essi sono anche elementi di $\mathbb{L}$ e $p(x)$ è anche un polinomio di $\mathbb{L}[x]$). Ad esempio consideriamo in $\mathbb{R}[x]$ il polinomio $p(x) = x^2 + 1$. Tale polinomio non ha radici in $\mathbb{R}$: se $\lambda \in \mathbb{R}$, si ha che $p(\lambda) = 1 + \lambda^2 \geq 1 > 0$. Invece $p(x)$ ha due radici in $\mathbb{C}$: $\imath$ e $-\imath$.

La proprietà di avere una radice è collegata alla divisibilità di un polinomio:

Teorema 8.34 (Teorema di Ruffini)**.** *Un elemento $\lambda \in \mathbb{K}$ è radice del polinomio $p(x) \in \mathbb{K}[x]$ se e solo se il polinomio $x - \lambda$ divide $p(x)$.*

Dimostrazione. Iniziamo dimostrando che se il polinomio $x - \lambda$ divide $p(x)$ allora λ è radice di $p(x)$. Per definizione di divisibilità tra polinomi (Definizione 8.29), esiste $q(x)$ tale che $p(x) = q(x)(x - \lambda)$. Essendo uguali i due polinomi sappiamo, dall'Osservazione 8.12, che le funzioni associate ad essi sono uguali. Dunque

valutando i polinomi a primo e secondo membro in λ in questa espressione otteniamo che $p(\lambda) = 0$, ovvero che λ è radice di $p(x)$.

Mostriamo adesso che se λ è una radice di $p(x)$ allora $x - \lambda$ divide $p(x)$. Questo equivale a mostrare che, se indichiamo con $q(x), r(x)$ i polinomi quoziente e resto della divisione di $p(x)$ per il polinomio $x - \lambda$, allora $r(x) = 0$. Per le proprietà del polinomio resto, o vale $r(x) = 0$ e in tal caso abbiamo $(x - \lambda)|p(x)$, oppure il grado di $r(x)$ è minore del grado di $x - \lambda$, e dunque $r(x)$ è una costante $c \neq 0$ di $\mathbb{K}$. Mostriamo che questo non può accadere:

$$p(x) = q(x)(x - \lambda) + c$$

Dunque

$$p(\lambda) = q(\lambda) \cdot (\lambda - \lambda) + c \tag{3.1}$$

Ma λ per ipotesi è una radice di $p(x)$ e dunque $p(\lambda) = 0$ e l'equazione (3.1) equivale a $c = 0$ che contraddice la nostra assunzione $c \neq 0$. □

Possiamo, a questo punto, definire la molteplicità di una radice di un polinomio.

Definizione 8.35. Data una radice $a \in \mathbb{K}$ di un polinomio $f(x) \in \mathbb{K}[x]$, si chiama **molteplicità** il numero intero positivo m tale che $(x-a)^m$ divide $f(x)$ e $(x-a)^{m+1}$ non divide $f(x)$.

Il legame tra radici e fattori di grado 1 di un polinomio evidenziato dal teorema di Ruffini pone un limite al numero di radici che un polinomio $p(x)$ di $\mathbb{K}[x]$ di grado n può avere in $\mathbb{K}$:

Corollario 8.36. *Dato un campo $\mathbb{K}$, un polinomio $p(x) \in \mathbb{K}[x]$, diverso dal polinomio 0, di grado $n \in \mathbb{N}$, ha al più n radici distinte in $\mathbb{K}$.*

Dimostrazione. Procediamo per induzione su $deg(p(x))$.
Passo base: Se $deg(p(x)) = 0$ l'enunciato è vero perché $p(x)$ è una costante diversa da zero e quindi il polinomio ha 0 radici.
Passo induttivo: Supponiamo che i polinomi di $\mathbb{K}[x]$ di grado d minore di $n + 1$ abbiamo al più d radici in $\mathbb{K}$. Mostriamo che da questo segue che i polinomi di grado $n + 1$ in $\mathbb{K}[x]$ hanno al più $n + 1$ radici in $\mathbb{K}$. Sia dunque $p(x)$ un generico polinomio in $\mathbb{K}[x]$ di grado $n + 1$. Possono presentarsi due casi:

(1) $p(x)$ non ha radici, e allora verifica banalmente l'enunciato: il numero di radici (cioè 0) è minore (in questo caso strettamente) del grado di $p(x)$ che è $n + 1$;
(2) $p(x)$ ha radici, e allora sia $\alpha \in \mathbb{K}$ una radice di $p(x)$. Per il teorema di Ruffini, $p(x) = p_1(x)(x - \alpha)$. Per ipotesi induttiva, $p_1(x)$, che ha grado n, ha al più n radici distinte in $\mathbb{K}$, chiamiamole $\alpha_1, \ldots, \alpha_r$ con $r \leq n$. Possiamo concludere che non ci sono radici di $p(x)$ diverse da $\alpha, \alpha_1, \ldots, \alpha_r$[7] e dunque ce ne sono al più $r + 1$ distinte e $r + 1 \leq n + 1$ come volevamo.

[7] Infatti se per assurdo β fosse una radice diversa da $\alpha, \alpha_1, \ldots, \alpha_r$, allora si potrebbe scrivere

$$p(\beta) = p_1(\beta)(\beta - \alpha)$$

ossia

$$0 = p_1(\beta)(\beta - \alpha)$$

Poiché $\beta - \alpha \neq 0$, deve valere $p_1(\beta) = 0$ e dunque β sarebbe una radice di $p_1(x)$ ma questo è assurdo perché β è diversa da $\alpha_1, \ldots, \alpha_r$.

□

Osservazione 8.37. Osserviamo che dal Corollario 8.36 segue che un polinomio $p(x)$ di $\mathbb{K}[x]$ diverso dal polinomio nullo e di grado n, ha al più n radici distinte anche se le cerchiamo in qualsiasi campo $\mathbb{L}$ che contenga $\mathbb{K}$.

Infatti se il campo $\mathbb{L}$ contiene $\mathbb{K}$, allora possiamo pensare il polinomio $p(x)$ come un polinomio a coefficienti in $\mathbb{L}$, e dunque come un elemento di $\mathbb{L}[x]$ e applicare il Corollario 8.36 a $p(x) \in \mathbb{L}[x]$.

Il Corollario 8.36 è molto importante perché, oltre a dare un limite al numero di radici di un polinomio di grado n, permette di dimostrare quello che è noto come principio d'identità dei polinomi per i campi infiniti. Il principio di identità dei polinomi riguarda la domanda che ci eravamo posti nell'Osservazione 8.12: se due funzioni associate ad un polinomio sono uguali allora sono uguali anche i polinomi a cui esse sono associate? Detto in altre parole, la funzione che fa corrispondere ad ogni polinomio la funzione polinomiale associata è iniettiva? La risposta del seguente teorema è SÌ nel caso in cui il campo $\mathbb{K}$ sia infinito, subito dopo vedremo che questo non è vero nel caso finito:

Teorema 8.38 (Principio d'identità dei polinomi)**.** *Se $\mathbb{K}$ è un campo infinito, allora due polinomi $f(x), g(x)$ che sono uguali come funzioni da $\mathbb{K}$ in $\mathbb{K}$ sono anche uguali come polinomi (ovvero, vedi definizione di uguaglianza tra polinomi 8.2, hanno lo stesso grado e tutti i coefficienti di grado corrispondente uguale).*

Dimostrazione. Consideriamo il polinomio $h(x) = f(x) - g(x)$. Per ipotesi la funzione associata ad esso è la funzione nulla che, ad ogni $c \in \mathbb{K}$, associa l'elemento 0 di $\mathbb{K}$, infatti:

$$h(c) = f(c) - g(c) \underbrace{=}_{\substack{\forall c \in \mathbb{K} \\ f(c)=g(c)}} 0$$

Dunque il polinomio $h(x)$ ha infinite radici in $\mathbb{K}$ (ogni valore di $\mathbb{K}$ è radice). Dal Corollario 8.36 segue che $h(x)$ non può essere un polinomio di grado n per nessun $n \in \mathbb{N}$. Dunque $h(x) = f(x) - g(x)$ può essere solo il polinomio 0 per cui non è definito il grado, ovvero il polinomio con tutti i coefficienti uguali a zero. Da questo segue che i polinomi $f(x)$ e $g(x)$ sono dello stesso grado e hanno tutti i coefficienti uguali. □

Osservazione 8.39. Nella dimostrazione precedente del principio d'identità dei polinomi entra in maniera decisiva l'ipotesi di considerare polinomi a coefficienti in un campo infinito: tutto segue infatti dall'osservazione che se $h(c) = 0$ per ogni c in $\mathbb{K}$ allora $h(x)$ *ha infinite radici in* $\mathbb{K}$.

Come già anticipato facciamo vedere che il principio d'identità dei polinomi, che sembra molto intuitivo (ma solo a causa del fatto che siamo molto più abituati a lavorare con campi infiniti quali $\mathbb{Q}$ e $\mathbb{R}$), non vale nel caso di campo dei coefficienti finito.
In $\mathbb{Z}_p$ consideriamo il polinomio nullo $f(x) = 0$ ed il polinomio $g(x) = x^p - x$. I due polinomi, alla luce della Definizione 8.2 di uguaglianza tra polinomi, sono evidentemente diversi, ma, dal piccolo teorema di Fermat sappiamo che, per ogni $c \in \mathbb{Z}_p$ si ha:

$$c^p \equiv c \ (p)$$

Ovvero per ogni c in $\mathbb{Z}_p$, si ha $g(c) = 0$. Dunque le funzioni associate ai polinomi $f(x)$ e $g(x)$ sono identiche (la funzione nulla), ma i polinomi sono diversi.

4. Massimo comun divisore tra polinomi e lemma di Bezout

Una volta definito il concetto di divisibilità e di divisore tra polinomi (Definizione 8.29) non c'è nessun ostacolo a proseguire l'analogia tra $\mathbb{Z}$ e $\mathbb{K}[x]$, introducendo la definizione di massimo comun divisore tra polinomi.

Definizione 8.40. Dati due polinomi $p_1(x), p_2(x)$ in $\mathbb{K}[x]$, non entrambi nulli, un **massimo comun divisore** di $p_1(x), p_2(x)$ è un polinomio $d(x)$ che divide sia $p_1(x)$ che $p_2(x)$, e tale che ogni altro polinomio che divide sia $p_1(x)$ che $p_2(x)$ ha grado minore o uguale a quello di $d(x)$.

Dimostriamo innanzitutto, in maniera del tutto analoga a quanto fatto in $\mathbb{Z}$, che tra due polinomi non entrambi nulli esiste sempre un massimo comun divisore e che vale l'analogo del lemma di Bezout.

Teorema 8.41 (Esistenza del massimo comun divisore e lemma di Bezout tra polinomi)**.** *Dati due polinomi $f(x)$,$g(x)$ in $\mathbb{K}[x]$ non entrambi nulli, esiste un polinomio $d(x)$ che è un massimo comun divisore tra $f(x)$ e $g(x)$.*
Esistono inoltre due polinomi $t(x)$ e $h(x)$ in $\mathbb{K}[x]$ tali che:

$$d(x) = f(x) \cdot h(x) + g(x) \cdot t(x)$$

Dimostrazione. Consideriamo il seguente sottoinsieme A di $\mathbb{K}[x]$:

$$A = \{z(x) \in \mathbb{K}[x] \backslash \{0\} | z(x) = t(x) \cdot f(x) + h(x) \cdot g(x) \text{ con } t(x), h(x) \in \mathbb{K}[x]\}$$

Scegliamo un elemento $d(x) = t(x) \cdot f(x) + h(x) \cdot g(x) \in A$ di grado minimo, ossia tale che $\forall z(x) \in A$ valga $deg(d(x)) \leq deg(z(x))$. Dividiamo $f(x)$ per $d(x)$ indicando con $q(x)$ e $r(x)$ rispettivamente il polinomio quoziente e resto (ovvero $r(x) = 0$ oppure $deg(r(x)) < deg(d(x))$):

$$f(x) = q(x)d(x) + r(x) \tag{4.1}$$

Osserviamo che, se $r(x)$ fosse diverso da 0, allora apparterrebbe ad A, infatti dall'equazione 4.1 segue che:

$$\begin{aligned} r(x) &= f(x) - q(x) \cdot \underbrace{(t(x) \cdot f(x) + h(x) \cdot g(x))}_{d(x)} = \\ &= (1 - q(x) \cdot t(x)) \cdot f(x) + (-q(x) \cdot h(x)) \cdot g(x) \end{aligned}$$

D'altra parte, avendo indicato con $d(x)$ l'elemento di grado minimo di A e sapendo che $deg(r(x)) < deg(d(x))$ abbiamo che $r(x)$ non può appartenere ad A, dunque $r(x)$ deve essere il polinomio nullo. Questo significa che $d(x)$ divide $f(x)$. Ripetendo lo stesso ragionamento per $g(x)$ troviamo che $d(x)$ divide anche $g(x)$.

Resta da mostrare che se un polinomio $c(x)$ divide sia $f(x)$ che $g(x)$ allora $c(x)$ ha grado minore o uguale di $d(x)$. In realtà dimostreremo *qualcosa di più*[8], ovvero che $c(x)$ divide $d(x)$. Per ipotesi esistono due polinomi $k(x)$ e $l(x)$ tali che:

$$f(x) = c(x) \cdot k(x) \quad g(x) = c(x) \cdot l(x)$$

Dunque:

$$d(x) = t(x) \cdot \underbrace{(c(x) \cdot k(x))}_{f(x)} + h(x) \cdot \underbrace{(c(x) \cdot l(x))}_{g(x)} = c(x) \cdot (t(x) \cdot k(x) + b(x) \cdot l(x))$$

[8]In realtà, in analogia con quel che accade in $\mathbb{Z}$, si può facilmente dimostrare che nella definizione di massimo comun divisore è equivalente la condizione divisore comune di grado massimo o divisore comune diviso da tutti gli altri divisori comuni.

Ovvero $c(x)$ divide $d(x)$, da cui segue che il grado di $c(x)$ è minore o uguale di quello di $d(x)$. □

La dimostrazione del Teorema 8.41 non è *costruttiva*, ovvero non fornisce un metodo per calcolare un massimo comun divisore tra due polinomi, né per calcolare i due polinomi $h(x)$ e $t(x)$ dell'enunciato. Ma proprio l'analogia tra $\mathbb{Z}$ e $\mathbb{K}[x]$ permette di dimostrare un risultato analogo a quello noto come algoritmo di Euclide per gli interi:

Teorema 8.42 (Algoritmo di Euclide)**.** *Per calcolare un massimo comun divisore $d(x)$ tra due polinomi $f(x)$ e $g(x)$ di $\mathbb{K}[x]$ si può usare l'algoritmo di Euclide.*

DIMOSTRAZIONE. La dimostrazione del fatto che alla fine l'algoritmo di Euclide restituisce un massimo comun divisore tra $f(x)$ e $g(x)$ è del tutto analoga a quella fatta per gli interi. □

Osservazione 8.43. Come nel caso di $\mathbb{Z}$, se $d(x)$ è un massimo comun divisore tra $f(x)$ e $g(x)$, i polinomi $h(x)$ e $t(x)$ del lemma di Bezout[9] si possono calcolare *risalendo* l'algoritmo di Euclide.

Prima di mostrare un esempio concreto di calcolo di massimo comun divisore tra due polinomi, osserviamo che fin dalla definizione abbiamo scritto **un** massimo comun divisore e non **il** massimo comun divisore: abbiamo infatti stabilito che esiste per ogni coppia di polinomi non entrambi nulli, ma non abbiamo discusso se è unico o meno.

È abbastanza immediato accorgersi che il massimo comun divisore non è unico. Infatti se $d(x)$ è un massimo comun divisore tra due polinomi $f(x)$ e $g(x)$ di $\mathbb{K}[x]$, allora qualsiasi polinomio ottenuto moltiplicando $d(x)$ per un elemento $k \neq 0$ di $\mathbb{K}$ continua ad esserlo. Infatti $k \cdot d(x)$ ha lo stesso grado di $d(x)$ (k è diverso da zero) inoltre è un divisore comune di $f(x)$ e $g(x)$. Infatti sappiamo che:

$$\underbrace{f(x) = d(x) \cdot h(x)}_{ipotesi\ d(x)\ divide\ f(x)} \qquad \underbrace{g(x) = d(x) \cdot t(x)}_{ipotesi\ d(x)\ divide\ g(x)}$$

e di conseguenza:

$$\underbrace{f(x) = (k \cdot d(x)) \cdot k^{-1} \cdot h(x)}_{quindi\ anche\ k \cdot d(x)\ divide\ f(x)} \qquad \underbrace{g(x) = (k \cdot d(x)) \cdot k^{-1} \cdot t(x)}_{quindi\ anche\ k \cdot d(x)\ divide\ g(x)}$$

Definizione 8.44. Due polinomi $f(x), g(x)$ in $\mathbb{K}[x]$ si dicono **associati** se differiscono moltiplicativamente per una costante non nulla, ossia se esiste $k \in \mathbb{K}$, $k \neq 0$ tale che $f(x) = k \cdot g(x)$.

Abbiamo appena osservato che:

Proposizione 8.45. *Se $d(x)$ è un massimo comun divisore tra due polinomi non entrambi nulli $f(x), g(x)$ in $\mathbb{K}[x]$ allora tutti i polinomi associati a $d(x)$ sono massimi comun divisori.*

Vogliamo dimostrare il viceversa. Osserviamo che, nella definizione di massimo comun divisore $d(x)$ tra due polinomi $p_1(x), p_2(x)$, la condizione *tale che ogni altro polinomio che divide sia $p_1(x)$ che $p_2(x)$ ha grado minore o uguale a quello di $d(x)$*

[9] Ovvero tali che $d(x) = f(x) \cdot h(x) + g(x) \cdot t(x)$.

è equivalente a *tale che ogni altro polinomio che divide sia $p_1(x)$ che $p_2(x)$ divide anche $d(x)$*.

Da questo segue che, dato un massimo comun divisore $d(x)$ tra $f(x)$ e $g(x)$, tutti gli altri massimi comun divisori si ottengono moltiplicando $d(x)$ per una costante diversa da zero, ovvero (vedi Proposizione 8.19) **per gli invertibili** dell'anello $\mathbb{K}[x]$. Infatti, se $d(x)$ e $t(x)$ sono due massimi comun divisori, allora entrambi dividono sia $f(x)$ che $g(x)$ quindi, essendo $t(x)$ massimo comun divisore, $d(x)|t(x)$, ovvero esiste $t_1(x)$ tale che $t(x) = d(x) \cdot t_1(x)$. Analogamente, essendo $d(x)$ massimo comun divisore, $t(x)|d(x)$, ovvero esiste $d_1(x)$ tale che $d(x) = t(x) \cdot d_1(x)$. Da questo segue che:

$$d(x) = t(x) \cdot d_1(x) = d(x) \cdot t_1(x) \cdot d_1(x)$$

ovvero

$$d(x)(1 - t_1(x) \cdot d_1(x)) = 0$$

che implica

$$d_1(x) \cdot t_1(x) = 1.$$

Dalla Proposizione 8.19 segue a questo punto che $d_1(x)$ e $t_1(x)$ sono costanti e una l'inversa dell'altra. Abbiamo dunque dimostrato quello che ci eravamo prefissi, ovvero:

Proposizione 8.46. *Dati due polinomi non entrambi nulli $f(x)$ e $g(x)$ in $\mathbb{K}[x]$, se $d_1(x)$ e $d_2(x)$ sono due massimi comun divisori tra $f(x)$ e $g(x)$, allora $d_1(x)$ e $d_2(x)$ sono polinomi associati.*

Nel seguito, dati due polinomi $f(x)$ e $g(x)$ di $\mathbb{K}[x]$ non entrambi nulli, chiameremo talvolta **il** massimo comun divisore di $f(x)$ e $g(x)$ (e lo indicheremo con $M.C.D.(f(x), g(x))$) l'unico polinomio monico nell'insieme dei massimi comun divisori di $f(x)$ e $g(x)$.
Illustriamo a questo punto con un esempio come si applica l'algoritmo di Euclide tra polinomi:

Esempio 8.47. Calcolare il M.C.D. in $\mathbb{Z}_3[x]$ tra $f(x) = x^2 - x + 4$ e $g(x) = x^3 + 2x^2 + 3x + 2$

Osserviamo che essendo il campo dei coefficienti $\mathbb{Z}_3$, i coefficienti stessi sono classi di equivalenza modulo 3. In particolare abbiamo (usando come rappresentanti delle classi i numeri compresi tra 0 e 2):

$$x^2 - x + 4 = x^2 + 2x + 1 \quad \text{e} \quad x^3 + 2x^2 + 3x + 2 = x^3 + 2x^2 + 2$$

Iniziamo l'algoritmo di divisione; confrontando i due termini principali osserviamo che bisogna moltiplicare x^2 per x per ottenere x^3:

$$\begin{array}{cccc|l} x^3 & +2x^2 & +3x & +2 & x^2 + 2x + 1 \\ x^3 & +2x^2 & +x & & x \\ & & 2x & +2 & \end{array}$$

Abbiamo quindi terminato il primo passo dell'algoritmo di Euclide:

$$g(x) = f(x) \cdot \underbrace{x}_{q_1(x)} + \underbrace{(2x+2)}_{r_1 x}$$

Il secondo passo prevede di dividere $f(x)$ per $r_1(x)$. I coefficienti direttivi dei due polinomi sono rispettivamente x^2 e $2x$, ma **attenzione** non si può moltiplicare per $\frac{1}{2} \cdot x$ perchè questa scrittura in $\mathbb{Z}_3$ non ha alcun senso. Il ragionamento però è analogo a quello che si farebbe in $\mathbb{Q}[x]$ (dove appunto moltiplicheremmo per $\frac{1}{2}x$): si tratta di moltiplicare per x moltiplicato per l'inverso di 2 in $\mathbb{Z}_3$ (che sappiamo esistere, essendo $\mathbb{Z}_3$ un campo!). Trovare questo inverso non è difficile visto che in $\mathbb{Z}_3$ ci sono tre elementi di cui uno è l'elemento nullo. Basta osservare che $2 \cdot 2 = 4$ ovvero 1 in $\mathbb{Z}_3$. Perciò l'inverso di 2 in $\mathbb{Z}_3$ è 2 e quindi nella divisione tra $f(x)$ e $r_1(x)$ dobbiamo moltiplicare per $2x$ il polinomio $r_1(x)$ e togliere il risultato da $f(x)$:

$$\begin{array}{lll|l} x^2 & +2x & +1 & 2x+2 \\ 4x^2 & +4x & & 2x \\ -3x^2 & -2x & +1 & \end{array}$$

Osserviamo che in $\mathbb{Z}_3$ il polinomio trovato è uguale al polinomio $x+1$, possiamo perciò continuare nell'algoritmo di divisione perché questo polinomio ha lo stesso grado di $2x+2$, che dovremo moltiplicare per 2:

$$\begin{array}{lll|l} x^2 & +2x & +1 & 2x+2 \\ 4x^2 & +4x & & 2x+2 \\ & x & +1 & \\ & x & +1 & \\ & & 0 & \end{array}$$

Perciò il secondo passo dell'algoritmo di Euclide è anche quello conclusivo:

$$f(x) = r_1(x) \cdot \underbrace{(2x+2)}_{q_2(x)} + \underbrace{0}_{r_2(x)}$$

E un M.C.D. tra $f(x)$ e $g(x)$ è l'ultimo resto non zero dell'algoritmo di Euclide, ovvero $r_1(x)$.

Vista la semplicità dei conti non è difficile trovare i due polinomi $t(x)$ e $h(x)$ dell'identità di Bezout, ovvero tali che:

$$f(x) \cdot t(x) + g(x) \cdot h(x) = r_1(x)$$

Infatti dal primo passo dell'algoritmo di Euclide segue che:

$$g(x) - f(x) \cdot x = r_1(x)$$

E perciò i due polinomi cercati sono: $t(x) = -x$ (o, se si vuole mantenere la convenzione di usare come rappresentanti delle classi in $\mathbb{Z}_3$ i numeri 0, 1 e 2, $t(x) = 2x$) e $h(x) = 1$.

Esercizio 8.48. Trovare il M.C.D. in $\mathbb{Q}[x]$ tra le seguenti coppie di polinomi:

(1) $x^5 - x^4 + 3x^2 - 2x^2 + 2x - 1$ e $x^6 + x^5 - 2x^4 + 6x^3 + 5x + 3$
(2) $x^9 - 1$ e $x^{11} - 1$
(3) $x^5 - 2x^3 + x^2 - 3x - 3$ e $x^3 + x^2 - 3x - 3$
(4) $x^5 + x^4 - 3x^3 + 2x^2 - 1$ e $x^3 - 2x^2 - x + 2$
(5) $x^3 + 3x^2 - 4$ e $x^3 - x^2 - 3x + 6$

Risoluzione. Svolgiamo qui di seguito il calcolo del M.C.D. tra $f(x) = x^9 - 1$ e $g(x) = x^{11} - 1$, lasciando gli altri per esercizio. Ricordiamoci, prima di procedere, che il campo dei coefficienti è $\mathbb{Q}$.

$$\begin{array}{llll} x^{11} & & -1 & \mid \; x^9 - 1 \\ x^{11} & -x^2 & & \mid \; x^2 \\ & x^2 & -1 & \end{array}$$

Primo passo algoritmo di Euclide:

$$g(x) = f(x) \cdot \underbrace{x^2}_{q_1(x)} + \underbrace{(x^2 - 1)}_{r_1(x)}$$

Continuiamo dividendo $f(x)$ per $r_1(x)$:

$$\begin{array}{llllll} x^9 & & & & -1 & \mid \; x^2 - 1 \\ x^9 & -x^7 & & & & \mid \; x^7 + x^5 + x^3 + x \\ & x^7 & & & -1 & \\ & x^7 & -x^5 & & & \\ & & x^5 & & -1 & \\ & & x^5 & -x^3 & & \\ & & & x^3 & -1 & \\ & & & x^3 & -x & \\ & & & & x & -1 \end{array}$$

Secondo passo algoritmo di Euclide:

$$f(x) = r_1(x) \cdot \underbrace{(x^7 + x^5 + x^3 + x)}_{q_2(x)} + \underbrace{x - 1}_{r_2(x)}$$

L'algorimo continua dividendo $r_1(x)$ per $r_2(x)$, è evidente (prodotto notevole) senza fare la divisione che il terzo passo dell'algoritmo di Euclide sarà:

$$\underbrace{(x^2 - 1)}_{r_1(x)} = \underbrace{(x - 1)}_{r_2(x)} \cdot \underbrace{(x + 1)}_{q_2(x)} + \underbrace{0}_{r_3(x)}$$

Perciò l'algoritmo è terminato e un M.C.D. tra $f(x)$ e $g(x)$ è l'ultimo resto non zero, ovvero $r_2(x) = x - 1$.

Esercizio 8.49. Trovare, dati $f(x) = x^9 - 1$ e $g(x) = x^{11} - 1$, i due polinomi dell'algoritmo di Bezout, ovvero $t(x)$ e $h(x)$ tali che:

$$f(x) \cdot t(x) + g(x) \cdot h(x) = x - 1.$$

Esercizio 8.50. Trovare, dati $f(x) = x^9 - 1$ e $g(x) = x^{11} - 1$, un[10] m.c.m. tra $f(x)$ e $g(x)$ ovvero un polinomio $m(x)$ che è multiplo sia di $f(x)$ che di $g(x)$ e tale che ogni polinomio che è multiplo comune di $f(x)$ e $g(x)$ ha grado maggiore o uguale di $m(x)$.

[10] L'uso dell'articolo indeterminato ha la stessa spiegazione che nel caso del M.C.D., infatti se $m(x)$ è un minimo comun multiplo tra $f(x)$ e $g(x)$, allora anche ogni altro polinomio ottenuto dalla moltiplicazione di $m(x)$ per una costante diversa da zero è m.c.m. tra $f(x)$ e $g(x)$.

5. Polinomi irriducibili e teorema di fattorizzazione unica

In questo paragrafo, che tratterà della fattorizzazione di polinomi, considereremo (per motivi che diverranno chiari nel corso del paragrafo stesso) anche polinomi a coefficienti in $\mathbb{Z}$, ovvero in un anello che non è un campo. Cercheremo di sottolineare le differenze principali nei due casi, una per esempio è che in $\mathbb{Z}[x]$ non è più vero che tutti i polinomi di grado 0 (ovvero le costanti non nulle) sono invertibili: gli unici polinomi invertibili sono il polinomio 1 e il polinomio -1 (per le altre costanti a non esiste un polinomio di grado 0 b in $\mathbb{Z}[x]$ tale che $a \cdot b = 1$). Useremo la notazione $A[x]$ quando considereremo il caso allargato di polinomi a coefficienti in un anello A commutativo, con unità e privo di divisori di zero.[11] I casi che ci interesseranno saranno essenzialmente quelli dei polinomi a coefficienti in $\mathbb{Z}, \mathbb{Z}_p, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, dunque tutti del tipo descritto e, a parte $\mathbb{Z}$, tutti campi.

Cominciamo introducendo il concetto di polinomio irriducibile in $A[x]$, che avrà lo stesso ruolo del concetto di numero primo in $\mathbb{Z}$.

Si tratta dunque di considerare quei polinomi che non possono essere scritti come prodotti di due altri polinomi. In realtà messa così non avremmo speranze, infatti qualsiasi polinomio $p(x)$ di $A[x]$ può essere scritto come il polinomio 1 per $p(x)$ o anche come $a \cdot a^{-1} \cdot p(x)$ al variare di a tra gli invertibili di A[12]. Proprio da questa osservazione vogliamo partire per definire i polinomi irriducibili.

Definizione 8.51. Dato un polinomio $p(x)$ di $A[x]$ con A anello, se esistono due polinomi $f(x)$ e $g(x)$ in $A[x]$ entrambi non invertibili e tali che

$$p(x) = f(x) \cdot g(x)$$

il prodotto $f(x) \cdot g(x)$ si dice **una fattorizzazione** di $p(x)$ in $A[x]$.

A questo punto possiamo caratterizzare quelli che vogliamo chiamare polinomi irriducibili in $A[x]$:

Definizione 8.52. Sia $f(x)$ un polinomio di $A[x]$ non invertibile. Il polinomio $f(x)$ si dice **riducibile** (o fattorizzabile) in $A[x]$ se in $A[x]$ esiste almeno una fattorizzazione di $f(x)$. Altrimenti il polinomio $f(x)$ si dice **irriducibile**.

Osservazione 8.53. Un modo equivalente di dire che un polinomio $f(x)$ di $A[x]$ è irriducibile (ed è quello che solitamente viene richiamato negli esercizi e nelle dimostrazioni) è affermare che qualsiasi scrittura di $f(x)$ come prodotto di polinomi di $A[x]$:

$$f(x) = g(x)h(x)$$

implica che uno dei due polinomi sia invertibile in $A[x]$. Ovvero nel caso di polinomi a coefficienti in un campo $\mathbb{K}$, essendo gli invertibili tutti e soli i polinomi di grado 0 (le costanti), $f(x)$ è irriducibile in $\mathbb{K}[x]$ se e solo se $f(x)$ ha grado maggiore o uguale a 1 e non può essere scritto come prodotto di due polinomi (non necessariamente distinti) di grado maggiore di 0.

Cominciamo a discutere qualche proprietà sulla irriducibilità che vale nei $\mathbb{K}[x]$ (ma in generale, vedremo, non vale per gli $A[x]$). La seguente proposizione è di facile dimostrazione (esercizio!).

[11] Si dice in tal caso che A è un dominio. Per esempio l'anello Z è un dominio, mentre l'anello $\mathbb{Z}_{15}$ non lo è.

[12] Osserviamo che anche per un numero primo in $\mathbb{Z}$ è esattamente la stessa cosa: può essere scritto come $1 \cdot p$, oppure $(-1) \cdot (-1) \cdot p$.

Proposizione 8.54. *Se $f(x), g(x) \in \mathbb{K}[x]$ sono polinomi associati (vedi Definizione 8.44) allora $f(x)$ è irriducibile se e solo $g(x)$ lo è.*

Osservazione 8.55. Dato un polinomio $f(x)$ in $\mathbb{Q}[x]$ si può considerare il polinomio ad esso associato $g(x) = s \cdot f(x)$ dove s è il minimo comun denominatore dei coefficienti di $f(x)$. In particolare $g(x)$ è un polinomio a coefficienti interi. La Proposizione 8.54 ci dice che studiare in $\mathbb{Q}[x]$ l'irriducibilità di $f(x)$ e $g(x)$ è equivalente. Tra poco dimostreremo un risultato - noto come lemma di Gauss - che afferma che un polinomio a coefficienti interi è irriducibile in $\mathbb{Q}[x]$ se e solo se è irriducibile in $\mathbb{Z}[x]$. Tutto questo spiega perché consideriamo anche polinomi a coefficienti in $\mathbb{Z}$: quando dovremo studiare l'irriducibilità di un polinomio in $\mathbb{Q}[x]$, considereremo il polinomio associato a coefficienti in $\mathbb{Z}[x]$ e ne analizzeremo la riducibilità in $\mathbb{Z}[x]$ (cosa che risulterà conveniente).

Proposizione 8.56. *Negli anelli di polinomi $\mathbb{K}[x]$, con $\mathbb{K}$ campo, tutti i polinomi di grado 1 sono irriducibili.*

Dimostrazione. Supponiamo che il polinomio $f(x) \in \mathbb{K}[x]$ di grado 1 sia il prodotto di due polinomi $g(x)$ e $h(x)$ di $\mathbb{K}[x]$:

$$f(x) = g(x)h(x)$$

Per le proprietà del grado del prodotto di polinomi (Proposizione 8.17) abbiamo che:

$$1 = deg(f(x)) = deg(g(x)) + deg(h(x))$$

Ovvero uno dei due polinomi deve avere grado 0. E sappiamo, dalla Proposizione 8.19, che in $\mathbb{K}[x]$ tutti i polinomi di grado 0 sono invertibili. □

Osservazione 8.57. Le due proposizioni precedenti non sono vere, in generale, per polinomi a coefficienti in un anello. Mostriamo, ad esempio, che in $\mathbb{Z}[x]$ esistono polinomi di primo grado riducibili. Consideriamo $f(x) = 2x - 4$, possiamo scriverlo come $2 \cdot (x - 2)$ ed i polinomi 2 e $x - 2$ non sono invertibili in $\mathbb{Z}[x]$. Questo offre anche un esempio di due polinomi associati (appunto $f(x)$ e $x - 2$), il primo dei quali è riducibile, mentre il secondo è irriducibile (come dimostreremo dopo).

Essenzialmente quello che *salta* rispetto alla dimostrazione fatta nel caso di polinomi a coefficienti in un campo, è che negli anelli (che non sono campi), non è vero che tutti i polinomi di grado 0 sono invertibili. Dunque, se $g(x) \in \mathbb{Z}[x]$ è di grado 1 ed è possibile raccogliere un fattore a di grado 0 non invertibile dai coefficienti di $g(x)$, $g(x)$ si fattorizza come:

$$g(x) = a \cdot \frac{g(x)}{a}$$

Osserviamo che dividere $g(x)$ per a, pur essendo in $\mathbb{Z}[x]$, è una *operazione lecita* in quanto a è un fattore comune a tutti i coefficienti di $g(x)$.

La conclusione dell'Osservazione 8.57 fornisce lo spunto per introdurre la seguente definizione:

Definizione 8.58. Un polinomio $f(x) = \sum_{i=0}^{n} a_i x^i$ in $\mathbb{Z}[x]$ si dice **primitivo** se il massimo comun divisore tra i suoi coefficienti $a_0, a_1, \ldots, a_n$ è uguale a 1.

Proposizione 8.59. *Ogni $f(x) \in \mathbb{Q}[x]$ è associato ad un polinomio primitivo.*

DIMOSTRAZIONE. Sappiamo (Osservazione 8.55) che $f(x)$ può essere associato ad un polinomio $g(x)$ a coefficienti interi, moltiplicando per il minimo comun denominatore s dei coefficienti di $f(x)$. A sua volta $g(x)$ è associato al polinomio $h(x)$ primitivo ottenuto dividendo $g(x)$ per il massimo comun divisore d dei suoi coefficienti. Perciò:

$$h(x) = \frac{s}{d} f(x)$$

che è a coefficienti interi e primitivo per costruzione, è associato a $f(x)$. □

Esercizio 8.60. Dimostrare che se $f(x)$ e $t \cdot f(x)$, con $t \in \mathbb{Q}$, sono polinomi in $\mathbb{Z}[x]$ primitivi allora $t = 1$ o $t = -1$.

Risoluzione Sia $f(x) = \sum_{i=0}^{n} a_i x^i$ e $t = \frac{p}{q}$ ridotto ai minimi termini. Mostriamo che $q = \pm 1$ e $p = \pm 1$. Se q è - in valore assoluto - maggiore di 1 allora esiste almeno un coefficiente a_i di $f(x)$ che non è multiplo di q: questo perché, essendo $f(x)$ primitivo, i suoi coefficienti non hanno fattori in comune maggiori di 1. Ma allora $t \cdot a_i$, che è l'i-esimo coefficiente di $t \cdot f(x)$, non sarebbe intero contro l'ipotesi di $t \cdot f(x)$ primitivo. Analogamente se p fosse - in valore assoluto - maggiore di 1, allora tutti i coefficienti di $t \cdot f(x)$ sarebbero multipli di p e ancora una volta questo contraddirebbe la primitività di $t \cdot f(x)$.

La definizione di polinomio primitivo ci permette di individuare i polinomi irriducibili di primo grado in $\mathbb{Z}[x]$ (e dunque di mostrare che effettivamente il polinomio $x - 2$ è irriducibile in $\mathbb{Z}[x]$).

Proposizione 8.61. *In $\mathbb{Z}[x]$ i polinomi di primo grado sono irriducibili se e solo se sono primitivi.*

DIMOSTRAZIONE. Se $f(x) = ax + b \in \mathbb{Z}[x]$ di primo grado è il prodotto di due polinomi, allora, per la proprietà del grado[13], deve essere il prodotto di un polinomio di primo grado $h(x) = sx + t$, per un polinomio di grado 0, ovvero una costante $c \in \mathbb{Z}$. Questo, per la definizione di uguaglianza tra polinomi significa che $c \cdot s = a$ e $c \cdot t = b$, dunque che c è un divisore comune dei coefficienti di $f(x)$. Dunque esiste c non invertibile (ovvero diverso da 1 o -1), e quindi una fattorizzazione di $f(x)$ (ovvero $c \cdot h(x)$) se e solo se $f(x)$ non è primitivo. □

Abbiamo dunque discusso l'irriducibilità dei polinomi di grado 1 in $\mathbb{K}[x]$ e in $\mathbb{Z}[x]$. Per quanto riguarda i polinomi di grado maggiore di 1, una discussione importante è quella che lega la irriducibilità di un polinomio $f(x)$ in $\mathbb{K}[x]$ di grado $n > 1$ al fatto che esso abbia radici in $\mathbb{K}$. Dal teorema di Ruffini (Teorema 8.34) segue che se $f(x)$ ha una radice α in $\mathbb{K}$ allora è riducibile. Infatti si ha che il polinomio $(x - \alpha)$ divide $f(x)$:

$$\underbrace{f(x)}_{grado>1} = g(x) \cdot \underbrace{(x-1)}_{grado=1}$$

Inoltre, per le proprietà del grado, $g(x)$ ha grado maggiore di 0, ovvero non è invertibile.

[13]Le proprietà del grado continuano a valere in $A[x]$ con A dominio, come potete facilmente verificare.

Viceversa in generale **non è vero** che se un polinomio di grado maggiore di 1 non ha radici allora è irriducibile. Ad esempio il polinomio $x^4 + 2x^2 + 1$ di $\mathbb{R}[x]$ è riducibile in $\mathbb{R}[x]$:

$$x^4 + 2x^2 + 1 = (x^2 + 1)^2$$

ma non ha radici in $\mathbb{R}$ (non esiste nessun numero reale che elevato al quadrato è uguale a -1).

L'unica cosa certa è che un polinomio che non ha radici in $\mathbb{K}$ allora non ha fattori di grado 1 nella sua fattorizzazione in $\mathbb{K}[x]$. Da questo segue che:

Corollario 8.62. *Un polinomio $f(x) \in \mathbb{K}[x]$ di grado 2 e 3 è riducibile se e solo se ha una radice in $\mathbb{K}$.*

Dimostrazione. Abbiamo osservato che, in generale, un polinomio di grado $n > 1$ che ha una radice in $\mathbb{K}$ è riducibile in $\mathbb{K}[x]$. Viceversa se un polinomio di grado 2 o 3 è riducibile allora, sfruttando le proprietà del grado del prodotto di polinomi, necessariamente nel primo caso ($n = 2$) deve essere il prodotto di due fattori di grado 1, mentre nel secondo caso ($n = 3$) può essere il prodotto di un polinomio di grado 1 per un polinomio di grado 2 o il prodotto di tre polinomi di grado 1. Ovvero abbiamo stabilito che i polinomi di grado 2 o 3 riducibili hanno necessariamente un fattore di grado 1 e il teorema di Ruffini ci dice che avere un fattore di grado 1 in $\mathbb{K}[x]$ equivale ad avere una radice in $\mathbb{K}$. □

Gli elementi irriducibili di $\mathbb{K}[x]$ hanno molte analogie con i numeri primi di $\mathbb{Z}$. Un primo risultato importante è quello che ci dice che *se un polinomio irriducibile divide un prodotto di polinomi, allora divide uno dei due fattori.* Enunciamo questo risultato nel seguente teorema, la cui dimostrazione, lasciata come esercizio, coinvolge, analogamente a quello che accade in $\mathbb{Z}$, il lemma di Bezout.

Teorema 8.63 (Primalità di un polinomio irriducibile)**.** *Se $p(x)$ è un polinomio irriducibile in $\mathbb{K}[x]$ dove $\mathbb{K}$ è un campo, e $p(x) \mid f(x) \cdot g(x)$ (dove $f(x), g(x) \in \mathbb{K}[x]$), allora o vale $p(x) \mid f(x)$ o vale $p(x) \mid g(x)$.*

Vale anche l'analogo del teorema di fattorizzazione unica (la dimostrazione è un esercizio caldamente consigliato; è una applicazione del teorema di primalità: si procede in maniera del tutto simile alla dimostrazione della fattorizzazione unica in $\mathbb{Z}$).

Teorema 8.64 (Teorema di fattorizzazione unica per polinomi)**.** *Ogni polinomio di grado ≥ 1 in $\mathbb{K}[x]$ (dove $\mathbb{K}$ è un campo) è irriducibile o si fattorizza come prodotto di polinomi irriducibili. Inoltre, se*

$$f(x) = p_1(x) \cdot p_2(x) \cdot \ldots \cdot p_s(x) = q_1(x) \cdot q_2(x) \cdot \ldots \cdot q_t(x)$$

sono due fattorizzazioni del polinomio $f(x)$ come prodotto di irriducibili, allora vale che $s = t$ e che i polinomi $p_i(x)$ e i polinomi $q_j(x)$ sono a due a due associati.

Nel teorema di fattorizzazione unica per polinomi i $p_i(x)$ non sono necessariamente distinti. Proprio come nel caso della fattorizzazione tra gli interi, possiamo scrivere la fattorizzazione di un polinomio *accorpando* i fattori uguali e usando le potenze. Si scriverà dunque

$$h(x) = a \cdot q_1^{r_1}(x) \cdot q_2^{r_2}(x) \cdots q_t^{r_t}(x)$$

dove a è il coefficiente direttivo di $h(x)$, i $q_j(x)$ sono i polinomi irriducibili distinti monici della fattorizzazione di $h(x)$, gli r_i sono i numeri naturali positivi che evidenziano quante volte ricorre il polinomio $q_i(x)$ nella fattorizzazione di $h(x)$.[14].

Avendo questa fattorizzazione è molto facile individuare, proprio come avveniva in $\mathbb{Z}$, il *M.C.D.* di due polinomi. Se infatti consideriamo un polinomio $g(x)$ e la sua fattorizzazione in irriducibili:

$$g(x) = b \cdot p_1^{s_1}(x)p_2^{s_2}(x) \cdot \ldots \cdot p_j^{r_j}(x)$$

allora il *M.C.D.* $(h(x), g(x))$ si otterrà facendo il prodotto degli irriducibili che compaiono sia fra i $p_m(x)$ che fra i $q_n(x)$, ciascuno preso col minimo esponente fra i due esponenti che troviamo nelle due fattorizzazioni.

Esempio 8.65. Consideriamo in $\mathbb{Q}[x]$,

$$h(x) = (x-1)^2(x^2-5)^3(x^4-7x+7)$$

e

$$g(x) = (x-1)^7(x^2-5)(x^5+11x^2+11)^2$$

e supponiamo di sapere che i fattori che compaiono nelle fattorizzazioni sono irriducibili (presto discuteremo un criterio che permette di verificarlo facilmente); allora il *M.C.D.* $(h(x), g(x))$ è

$$(x-1)^2(x^2-5)$$

Gli altri *M.C.D.* $(h(x), g(x))$, come sappiamo, sono tutti i polinomi associati a $(x-1)^2(x^2-5)$.

Osservazione 8.66. L'unicità della fattorizzazione in $\mathbb{K}[x]$ è a meno dell'ordine dei fattori e di moltiplicazione per invertibili, cioè le costanti. Ovvero la fattorizzazione $(x-1)\cdot(x-2)$ del polinomio x^2-3x+2 potrebbe essere scritta anche $(x-2)\cdot(x-1)$, ma questa fattorizzazione la consideriamo identica alla precedente, abbiamo cambiato solo l'ordine dei fattori. Così come consideriamo identica la fattorizzazione $\frac{1}{2}\cdot(x-1)\cdot 2\cdot(x-2)$, in quanto abbiamo solo moltiplicato per invertibili (il cui prodotto è 1) i due fattori irriducibili.

Anche in questo caso osserviamo l'analogia con l'unicità della fattorizzazione in primi dei numeri in $\mathbb{Z}$. Il numero 21 è uguale a $7\cdot 3$; noi consideriamo identica (perchè cambiamo solo l'ordine) la fattorizzazione $3\cdot 7$, ma anche la fattorizzazione che si può ottenere moltiplicando per invertibili il cui prodotto totale sia 1. Gli invertibili in $\mathbb{Z}$ sono 1 e -1. Dunque 21 lo possiamo fattorizzare anche come $-1\cdot 3\cdot(-1)\cdot 7$ ovvero come $-3\cdot(-7)$.

Osservazione 8.67. Il teorema di fattorizzazione unica vale per ogni $\mathbb{K}[x]$ con $\mathbb{K}$ campo. Per la dimostrazione usiamo il teorema di primalità che a sua volta si dimostra tramite il teorema di Bezout che vale in $\mathbb{K}[x]$ con K campo. Cosa succede se l'insieme dei coefficienti A è un anello ma non un campo? Vale la fattorizzazione unica? La risposta è *"dipende"*... Si può infatti dimostrare che il teorema di fattorizzazione unica vale anche in $\mathbb{Z}[x]$, ma anche mostrare esempi di anelli (che non sono campi) per cui il il teorema di fattorizzazione unica non vale. Consideriamo ad esempio l'insieme $\mathbb{Z}_{30}[x]$ ed il polinomio x^2-1. Facendo i conti si può verificare che:

$$x^2-1 = (x-1)(x-29) = (x-19)(x-11)$$

[14]Detto in formule r_i è quel numero naturale tale che $q_i(x)^{r_i}$ divide $h(x)$ e $q_i(x)^{r_i+1}$ non divide $h(x)$.

Queste sono due distinte fattorizzazioni in irriducibili.

6. Fattorizzazione in $\mathbb{C}[x]$, $\mathbb{R}[x]$, $\mathbb{Q}[x]$ e $\mathbb{Z}_p[x]$

Affrontiamo ora il problema della fattorizzazione nell'anello dei polinomi $\mathbb{K}[x]$, variando $\mathbb{K}$ tra uno dei seguenti campi: $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}_p$ (con p primo).

6.1. Fattorizzazione in $\mathbb{C}[x]$. Il campo $\mathbb{C}$ dei numeri complessi ha una proprietà molto importante per quanto riguarda le radici di polinomi a coefficienti in $\mathbb{C}$, proprietà che non a caso si chiama **teorema fondamentale dell'algebra** e di cui noi riportiamo solo l'enunciato (la dimostrazione di questo risultato esula dagli obiettivi di questo testo).

Teorema 8.68 (Teorema fondamentale dell'algebra)**.** *Ogni polinomio $f(x)$ a coefficienti in $\mathbb{C}$ di grado maggiore di zero ammette almeno una radice in $\mathbb{C}$.*

Usando il teorema fondamentale dell'algebra e il teorema di Ruffini abbiamo una caratterizzazione completa degli irriducibili in $\mathbb{C}$. Infatti una immediata conseguenza è che:

Corollario 8.69. *Ogni polinomio $f \in \mathbb{C}[x]$ di grado $n > 0$ è il prodotto di n fattori di primo grado in $\mathbb{C}[x]$.*

Dimostrazione. Procediamo per induzione sul grado n di f. Se f è di primo grado la tesi segue immediatamente. Sia ora $f(x) = \sum_{i=0}^{n} a_i x^i$ con $a_i \in \mathbb{C}$ e $a_n \neq 0$, $n > 1$. Possiamo scrivere $f(x) = a_n g(x)$ con $g(x)$ monico. Sia α radice di $g(x)$, la cui esistenza è assicurata dal Teorema 8.68 allora:

$$f(x) = a_n(x - \alpha)g_1(x) \quad \text{con} \quad deg(g_1) = n - 1$$

quindi g_1 e di conseguenza f si scrivono come prodotto di fattori di grado 1. □

Dal Corollario 8.69 segue che:

In $\mathbb{C}[x]$ un polinomio è irriducibile se e solo se è di primo grado

In $\mathbb{C}[x]$ quindi fattorizzare un polinomio equivale a trovarne le radici perchè tutti i suoi fattori irriducibili sono di grado 1. Dobbiamo cioè essere in grado di risolvere equazioni polinomiali a coefficienti complessi, cosa che può essere anche molto complicata. Prima di vedere un esempio, sottolineiamo il fatto che la ricerca di radici complesse è importante, come vedremo, anche per la fattorizzazione in $\mathbb{R}[x]$.

Esempio 8.70. Fattorizzare il polinomio $x^2 + 4x + 5 \in \mathbb{C}[x]$ come prodotto di irriducibili.

Dobbiamo trovare le radici complesse del polinomio $x^2 + 4x + 5$, ovvero le soluzioni complesse dell'equazione

$$x^2 + 4x + 5 = 0 \tag{6.1}$$

La formula risolutiva dell'equazione di secondo grado ci permette di trovare le soluzioni complesse (anche se il delta è negativo!):

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Nel nostro caso:

$$x_{1,2} = \frac{-4 \pm 2i}{2} = -2 \pm i$$

Quindi il polinomio $x^2 + 4x + 5 \in \mathbb{C}[x]$ si fattorizza in irriducibili come:

$$(x - (-2 + i)) \cdot (x - (-2 - i))$$

Per riprova possiamo calcolarci questo prodotto osservando che:

$$(x - (-2 + i)) \cdot (x - (-2 - i)) = ((x + 2) + i) \cdot ((x + 2) - i)$$

E questo sappiamo essere un prodotto notevole (ovvero la differenza di quadrati):

$$((x + 2) + i) \cdot ((x + 2) - i) = (x + 2)^2 - i^2 = x^2 + 4x + 5$$

Per la ricerca di radici complesse in polinomi a coefficienti reali (e dunque utile sia per la fattorizzazione in $\mathbb{C}[x]$ che in $\mathbb{R}[x]$) è importante ricordare la funzione coniugio da $\mathbb{C}$ in $\mathbb{C}$:

Definizione 8.71. Chiamiamo funzione **coniugio** la funzione da $\mathbb{C}$ in $\mathbb{C}$ che al numero complesso $a + ib$ associa $\overline{a + ib} = a - ib$.

Esercizio 8.72. Usando la definizione dimostrare le seguenti proprietà della funzione coniugio:

(1) I suoi punti fissi, ovvero gli $z \in \mathbb{C}$ tali che $\overline{z} = z$, sono tutti e soli i numeri reali.
(2) Il coniugio della somma è la somma dei coniugi, ovvero per ogni $z, w \in \mathbb{C}$ $\overline{z + w} = \overline{z} + \overline{w}$.
(3) Il coniugio del prodotto è il prodotto dei coniugi, ovvero per ogni $z, w \in \mathbb{C}$ $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$.
(4) Il prodotto di un numero complesso per il suo coniugato è un numero reale, ovvero per ogni $z \in \mathbb{C}$ si ha che $z \cdot \overline{z} \in \mathbb{R}$.

Il coniugio permette di dimostrare una interessante proprietà delle radici complesse di un polinomio a coefficienti reali (**ATTENZIONE:** sottolineiamo il fatto che tra le ipotesi che stiamo considerando c'è che i coefficienti del polinomio siano reali), ovvero che se z è una radice di un polinomio $p(x)$ a coefficienti reali, allora $\overline{z}$ è una radice di $p(x)$. Questo è ovvio, ma non è di nessuna utilità, se z è reale in quanto $\overline{z} = z$, ma è invece importante nel caso in cui $z \in \mathbb{C} - \mathbb{R}$:

Proposizione 8.73. *Sia $f(x) \in \mathbb{R}[x] \subset \mathbb{C}[x]$ e sia $\alpha \in \mathbb{C}$ una radice di f. Allora anche $\overline{\alpha}$ è una radice di f.*

Dimostrazione. sia $f(x) = \sum_{i=0}^{n} a_i x^i$ con $a_i \in \mathbb{R}$. Per ipotesi:

$$0 = f(\alpha) = \sum_{i=0}^{n} a_i \alpha^i$$

quindi, dall'enunciato dell'Esercizio 8.72, segue che:

$$\overline{0} = \overline{\sum_{i=0}^{n} a_i \alpha^i} = \sum_{i=0}^{n} \overline{a_i \alpha^i} = \sum_{i=0}^{n} a_i \overline{\alpha^i} = \sum_{i=0}^{n} a_i \overline{\alpha}^i$$

Cioè $f(\overline{\alpha}) = \overline{0} = 0$. □

Nel prossimo esercizio useremo il risultato della Proposizione 8.73 per fattorizzare un polinomio a coefficienti reali in $\mathbb{C}[x]$.

Esercizio 8.74. Sapendo che $f(x) = x^4 - 4x^3 + 3x^2 + 14x + 26$ ha radice $3 + 2i$, fattorizzare il polinomio in $\mathbb{C}[x]$.

Risoluzione. Il polinomio considerato è a coefficienti interi, quindi in particolare reali. Allora possiamo applicare la Proposizione 8.73 e concludere che anche $3-2i$ è radice del polinomio; da questo segue che $(x-(3+2i))\cdot(x-(3-2i)) = x^2-6x+13$ divide $f(x)$:

$$\begin{array}{cccccc|l} x^4 & -4x^3 & +3x^2 & +14x & +26 & | & x^2-6x+13 \\ x^4 & -6x^3 & +13x^2 & & & | & x^2+2x+2 \\ & 2x^3 & -10x^2 & +14x & +26 & & \\ & 2x^3 & -12x^2 & +26x & & & \\ & & 2x^2 & -12x & +26 & & \\ & & 2x^2 & -12x & +26 & & \\ & & & & 0 & & \end{array}$$

Quindi:

$$f(x) = \underbrace{(x-(3+2i))\cdot(x-(3-2i))}_{x^2-6x+13}\cdot(x^2+2x+2)$$

E per completare la fattorizzazione in $\mathbb{C}[x]$ resta da fattorizzare il polinomio $x^2 + 2x + 2$.
Calcoliamo le radici del polinomio attraverso la formula risolutiva delle equazioni di secondo grado:

$$x_{1,2} = \frac{-2 \pm \sqrt{-4}}{2} = \frac{-2 \pm 2i}{2} = \frac{2\cdot(-1 \pm i)}{2} = -1 \pm i$$

Per cui la fattorizzazione di $f(x)$ è data da:

$$(x-(3+2i))\cdot(x-(3-2i))\cdot(x+(1+i))\cdot(x+(1-i))$$

Osservazione 8.75. Osserviamo, senza ancora aver parlato di fattorizzazione in $\mathbb{R}[x]$, che la fattorizzazione in $\mathbb{C}[x]$ del polinomio $f(x) = x^4 - 4x^3 + 3x^2 + 14x + 26$ dell'Esercizio 8.74 fornisce indicazioni importanti sulla fattorizzazione dello stesso polinomio in $\mathbb{R}[x]$.

6.2. Fattorizzazione in $\mathbb{R}[x]$. Anche in $\mathbb{R}[x]$ si possono caratterizzare i polinomi irriducibili attraverso il grado, utilizzando quello che sappiamo della fattorizzazione in $\mathbb{C}[x]$.

Consideriamo un generico polinomio $f(x) \in \mathbb{R}[x]$ di grado n. In particolare $f(x)$ può essere visto come elemento di $\mathbb{C}[x]$ e indichiamo con $z_1, \ldots, z_r$ le sue radici complesse e con $m_1, \ldots, m_r$ le loro rispettive molteplicità[15]. La fattorizzazione di $f(x)$ in $\mathbb{C}[x]$ è dunque la seguente:

$$\prod_{i=1}^{r}(x-z_i)^{m_i} \tag{6.2}$$

Come si passa dalla fattorizzazione in $\mathbb{C}[x]$ a quella in $\mathbb{R}[x]$? Si osserva che se $z_i \in \mathbb{R}$ allora $(x-z_i)^{m_i}$ è un fattore di $f(x)$ in $\mathbb{R}[x]$, mentre se $z_i \in \mathbb{C}-\mathbb{R}$, allora il fattore $(x-z_i)^{m_i}$ non appartiene a $\mathbb{R}[x]$, ma sappiamo che esiste un'altra radice z_j di $f(x)$

[15]Sappiamo, dal Corollario 8.69, che $\sum_{i=1}^{r} m_i = n$, ma in generale $r \leq n$. È $r = n$ solo se $f(x)$ ha tutte radici distinte in $\mathbb{C}[x]$.

tale che $z_j = \overline{z_i}$ e $m_i = m_j$.
Dunque, nella fattorizzazione 6.2, è presente il fattore

$$((x - z_i) \cdot (x - \overline{z_i}))^{m_i}$$

L'osservazione chiave è che il fattore di secondo grado $(x-z_i)\cdot(x-\overline{z_i})$ è un polinomio reale. Infatti sia $z = a + ib$, $a, b \in \mathbb{R}$ e $b \neq 0$, allora:

$$(x - (\underbrace{a + ib}_{z})) \cdot (x - (\underbrace{a - ib}_{\overline{z}})) = x^2 - 2ax + a^2 + b^2$$

Come anticipato, i coefficienti del polinomio (1, $-2a$ e $a^2 + b^2$) sono reali.
Riassumendo, date le radici complesse $z_1, \ldots, z_r$ di $f(x)$, se z_i è un numero reale allora $x - z_i$ è un fattore irriducibile di primo grado di $f(x)$ (ripetuto m_i volte) della fattorizzazione in $\mathbb{R}[x]$, se z_i non è un numero reale (ovvero $z_i = a + ib$ con $b \neq 0$) allora $(x - z_i) \cdot (x - \overline{z_i})$ è un fattore di secondo grado della fattorizzazione in $\mathbb{R}[x]$ (ripetuto m_i volte) ed è irriducibile. Quest'ultima proprietà deriva dal fatto che, essendo di secondo grado, o è irriducibile o è il prodotto di due fattori di primo grado. Ma questa seconda opzione possiamo escluderla in quanto, dal teorema di Ruffini sappiamo che i fattori di primo grado sono associati ad una radice nel campo, e sappiamo, per ipotesi, che le radici del polinomio (che sono z e $\overline{z}$) non sono reali ($b \neq 0$).[16].
Dunque la fattorizzazione 6.2 di $f(x)$ in $\mathbb{C}[x]$ fatta di tutti fattori di grado 1, si *trasforma* in una fattorizzazione in $\mathbb{R}[x]$ di $f(x)$ tenendo inalterati i fattori con radici reali e *accorpando* in fattori irriducibili di secondo grado quelli corrispondenti a radici non reali (moltiplicando $x - z$ per $x - \overline{z}$).
Abbiamo scoperto che:

Proposizione 8.76. *Ogni polinomio di grado maggiore di* 2 *in* $\mathbb{R}[x]$ *è riducibile.*

DIMOSTRAZIONE. Infatti in $\mathbb{C}[x]$ il polinomio $f(x)$ ha $n = deg(f(x))$ radici (non necessariamente distinte)[17] $z_1, \ldots, z_n$. Se una di queste n radici è reale, allora $f(x)$ ha un fattore di grado 1 e dunque è riducibile, altrimenti se sono tutte radici complesse non reali, $f(x)$ è divisibile per il polinomio reale di secondo grado $(x - z_1) \cdot (x - \overline{z_1})$:

$$f(x) = (x - z_1) \cdot (x - \overline{z_1}) \cdot h(x)$$

E per la proprietà del grado del prodotto di polinomi, $h(x)$ ha grado maggiore di 1 e dunque non è invertibile. □

Per concludere la piena caratterizzazione degli irriducibili in $\mathbb{R}[x]$, sapendo che (Proposizione 8.56) in ogni campo i polinomi di grado 1 sono irriducibili, ci resta da approfondire il caso dei polinomi di grado 2. Ma questo è molto semplice, infatti dal Corollario 8.62, sappiamo che $f(x) \in \mathbb{K}[x]$ di grado 2 è riducibile se e solo se ha una radice in $\mathbb{K}$. Nel caso di $\mathbb{K} = \mathbb{R}$ è noto dalla scuola superiore che, se $f(x) = ax^2 + bx + c$ è un generico polinomio reale di grado 2, allora $f(x)$ ha radici in $\mathbb{R}$ se e solo se:

$$b^2 - 4ac \geq 0$$

Abbiamo dunque la completa caratterizzazione degli irriducibili in $\mathbb{R}[x]$:
In $\mathbb{R}[x]$ un polinomio è irriducibile se e solo è di primo grado oppure di

[16]Si poteva anche esprimere questa osservazione utilizzando il Corollario 8.62: un polinomio di grado 2 è irriducibile se e solo se non ha radici nel campo.

[17]Potrebbe essere anche tutte uguali e dunque una radice di molteplicità n.

secondo grado (del tipo ax^2+bx+c con $a \neq 0$) con $\Delta = b^2-4ac$ minore di zero.
Abbiamo dunque un *algoritmo* molto rapido per sapere se un polinomio $f(x)$ è riducibile in $\mathbb{R}[x]$ (basta guardare il grado ed eventualmente calcolare il delta nel caso il grado sia 2). Ma sapere che un polinomio $f(x)$ è riducibile non implica che la sua fattorizzazione in fattori irriducibili sia semplice da trovare.

Esercizio 8.77. Fattorizzare il polinomio $x^4-2x^2-3 \in \mathbb{R}[x]$.

Questo polinomio è di grado 4 ed è dunque riducibile in $\mathbb{R}[x]$: o è il prodotto di quattro polinomi di grado 1 (4 radici reali non necessariamente distinte), o il prodotto di un polinomio di grado 2 e due di grado 1 (2 radici reali non necessariamente distinte e 2 complesse coniugate) o il prodotto di due polinomi di grado 2 (4 radici complesse a due a due coniugate e non necessariamente distinte). Come si evince da questa prima analisi sarebbe fondamentale riuscire a determinarne le radici complesse. Esiste una formula risolutiva per le equazioni di quarto grado, ma non la conosciamo e dunque cerchiamo di agire diversamente, osservando che il polinomio considerato è, in un certo senso, *particolare*: non ha termini di grado dispari. Possiamo quindi, con la semplice sostituzione $x^2=t$, ottenere un polinomio di grado 2 associato a quello di partenza: t^2-2t-3. Cerchiamo di fattorizzare questo polinomio in $\mathbb{R}[t]$. Dalla formula risolutiva delle equazioni di secondo grado otteniamo:

$$t_{1,2} = \frac{2 \pm \sqrt{16}}{2}$$

Ovvero $t^2-2t-3 = (t-3)\cdot(t+1)$. Quindi:

$$x^4-2x^2-3 \underbrace{=}_{x^2=t} t^2-2t-3 = (t-3)\cdot(t+1) \underbrace{=}_{t=x^2} (x^2-3)\cdot(x^2+1)$$

In questo caso è facile vedere che x^2+1 è irriducibile in $\mathbb{R}[x]$ (ha radici complesse i e $-i$), mentre $x^2-3 = (x-\sqrt{3})\cdot(x+\sqrt{3})$. Concludendo si ha che la fattorizzazione in irriducibili di $x^4-2x^2-3 \in \mathbb{R}[x]$ è data da:

$$(x-\sqrt{3})\cdot(x+\sqrt{3})\cdot(x^2+1)$$

6.3. Fattorizzazione in $\mathbb{Q}[x]$. In $\mathbb{Q}[x]$, a differenza di quanto visto per $\mathbb{C}[x]$ e $\mathbb{R}[x]$, vedremo che per ogni naturale n esistono polinomi di grado n irriducibili. In $\mathbb{Q}[x]$ c'è però un *trucco* per facilitare lo studio della riducibilità o meno di un polinomio, e questo trucco è legato a quanto detto nella Osservazione 8.55, ovvero che il problema dell'irriducibilità e della fattorizzazione in $\mathbb{Q}[x]$ può essere ridotto allo studio di polinomi a coefficienti interi (considerando il polinomio a coefficienti interi primitivo $g(x)$ associato a $f(x)$).

Questo è molto importante perché dimostreremo ora che, nel caso di un polinomio primitivo a coefficienti interi, la sua irriducibilità in $\mathbb{Q}[x]$ è equivalente alla sua irriducibilità in $\mathbb{Z}[x]$. Questo è un risultato per niente banale e scontato: infatti, per esempio, se è vero che è ovvio che un polinomio $f(x) \in \mathbb{K}[x]$, riducibile in $\mathbb{K}[x]$, è riducibile in qualsiasi campo $\mathbb{L}$ che contenga strettamente $\mathbb{K}$ (basta considerare la stessa fattorizzazione, infatti i polinomi di $\mathbb{K}[x]$ sono in particolare polinomi di $\mathbb{L}[x]$), il viceversa non è in generale vero. Ad esempio qualsiasi polinomio di secondo grado irriducibile in $\mathbb{R}[x]$ (ad esempio x^2+1) è riducibile in $\mathbb{C}[x]$ (nel caso di x^2+1 è uguale a $(x-i)\cdot(x+i)$).

Lemma 8.78 (Lemma di Gauss)**.** *Sia $f(x) \in \mathbb{Z}[x]$. Se $f(x) = a(x)b(x)$ in $\mathbb{Q}[x]$ allora possiamo trovare due polinomi $a_1(x) \in \mathbb{Z}[x]$, associato a $a(x)$, e $b_1 \in \mathbb{Z}[x]$, associato a $b(x)$, tali che*

$$f(x) = a_1(x)b_1(x)$$

DIMOSTRAZIONE FACOLTATIVA !. Per dimostrare questo lemma ci serve un risultato intermedio che lasciamo come utile esercizio:

Esercizio 8.79. Il prodotto di due polinomi primitivi è un polinomio primitivo.

A questo punto se $f(x) \in \mathbb{Z}[x]$ è per ipotesi riducibile in $\mathbb{Q}[x]$ ($f(x) = g(x)h(x)$) consideriamo il suo associato primitivo $p(x)$ (vedi Proposizione 8.59). $p(x)$ è riducibile in $\mathbb{Q}[x]$, infatti $p(x)$ è il prodotto di $f(x)$ per una costante $t \in \mathbb{Q}$, dunque:

$$p(x) = t \cdot g(x)h(x)$$

Ora esistono (sempre per la Proposizione 8.59) due numeri razionali w, q tali che i polinomi $w \cdot t \cdot g(x)$ e $q \cdot h(x)$ sono primitivi. Per l'Esercizio 8.79 si ha che il seguente polinomio è primitivo:

$$w \cdot q \cdot f(x) = \underbrace{(w \cdot t \cdot g(x))}_{primitivo} \cdot \underbrace{q \cdot h(x)}_{primitivo} \tag{6.3}$$

Essendo $f(x)$ e $w \cdot q \cdot f(x)$ primitivi, dall'Esercizio 8.60 segue che $w \cdot q$ è uguale a 1 o a -1 e dunque sostituendo nell'equazione 6.3, si trova che:

$$\pm f(x) = \underbrace{(w \cdot t \cdot g(x))}_{primitivo} \cdot \underbrace{q \cdot h(x)}_{primitivo}$$

Ovvero $f(x)$ è il prodotto dei due polinomi a coefficienti interi $w \cdot t \cdot g(x)$ e $q \cdot h(x)$ □

Riassumendo, $g(x) \in \mathbb{Q}[x]$ è riducibile se e solo se il polinomio primitivo a coefficienti interi $f(x)$ ad esso associato è riducibile in $\mathbb{Z}[x]$. Abbiamo in definitiva ridotto la fattorizzazione in $\mathbb{Q}[x]$ a quella in $\mathbb{Z}[x]$ con notevoli vantaggi come vedremo da qui in avanti.

Cominciamo mostrando un primo criterio molto utile per riconoscere (e costruire) polinomi irriducibili in $\mathbb{Q}[x]$.

Teorema 8.80 (Criterio di Eisenstein)**.** *Sia*

$$f(x) = \sum_{i=0}^{n} a_i x^i$$

un polinomio primitivo di grado maggiore di 1 a coefficienti interi. Se esiste un numero primo p tale che:

1. *p NON divide il coefficiente direttivo a_n,*
2. *p divide tutti gli a_i con $i < n$,*
3. *p^2 non divide il termine noto a_0,*

allora $f(x)$ è irriducibile in $\mathbb{Z}[x]$, e dunque - per il lemma di Gauss - in $\mathbb{Q}[x]$.

DIMOSTRAZIONE. Supponiamo che $f(x)$ sia uguale al prodotto dei due polinomi $g(x) = \sum_{i=0}^{r} b_i x^i$ e $h(x) = \sum_{i=0}^{s} c_i x^i$ di $\mathbb{Z}[x]$, entrambi di grado maggiore o uguale a 1. Da $f(x) = g(x)h(x)$ e dalla definizione di uguaglianza tra polinomi,

segue che tutti i coefficienti del polinomio a destra sono uguali a tutti i coefficienti del polinomio a sinistra. Facendo i conti, otteniamo un sistema dove gli $n+1$ coefficienti a_i di $f(x)$ sono espressi tramite i coefficienti di $g(x)$ e $h(x)$ come segue[18]:

$$a_i = \sum_{j=0}^{i} b_j \cdot c_{i-j} \tag{6.4}$$

Partiamo *dal basso* del sistema 6.4: $a_0 = b_0 c_0$. Per ipotesi p divide a_0, ma p^2 non divide a_0: questo significa che p divide uno tra b_0 e c_0, ma non entrambi. Il ruolo dei b_i e dei c_i è simmetrico quindi possiamo, senza perdere di generalità, supporre che p divida b_0 e non c_0.

A questo punto la seconda equazione del sistema 6.4 è $a_1 = b_1 c_0 + b_0 c_1$, che diventa:

$$b_1 c_0 = a_1 - b_0 c_1$$

Ora sappiamo che p divide a_1 (ipotesi), p divide b_0 (appena stabilito) e dunque p divide $b_1 c_0$. Sappiamo anche che p non divide c_0 e di conseguenza divide b_1.

Iterando questo procedimento si ottiene che p divide ogni b_i e di conseguenza divide $a_n = b_r c_s$: ma questo è contro l'ipotesi. L'assurdo nasce dal fatto di aver supposto che $f(x)$, che verifica le tre condizioni del criterio di Eisenstein, possa essere scritto come prodotto di due polinomi di grado maggiore o uguale a 1. □

Come detto il criterio di Eisenstein permette di costruire polinomi irriducibili in $\mathbb{Q}[x]$ e addirittura permette di trovarne *infiniti* per ogni grado $n > 0$:

Corollario 8.81. *In $\mathbb{Q}[x]$ esistono polinomi irriducibili di grado $n > 0$ qualsiasi.*

DIMOSTRAZIONE. Basta considerare il polinomio $x^n - 2$ ed applicare Eisenstein con primo $p = 2$. Infatti 2 divide il termine noto (2), ma il quadrato di p (4) non divide il termine noto. E infine 2 non divide il coefficiente direttivo (1). Lo stesso ragionamento permette di dimostrare che $x^n - p$, per un qualsiasi primo p, è irriducibile. □

Un altro punto importante per fattorizzare in $\mathbb{Q}[x]$ un polinomio $f(x)$ a coefficienti interi è il fatto che la conoscenza del coefficiente direttivo e del termine noto di $f(x)$ permette di limitare la ricerca delle *possibili* radici razionali di $f(x)$ (e dunque, in termini di fattorizzabilità, dei possibili fattori di grado 1 di $f(x)$) ad un insieme finito di numeri razionali. Per la precisione:

Proposizione 8.82. *Se $f(x) \in \mathbb{Z}[x]$ e r/s (ridotto ai minimi termini, ovvero con $(r, s) = 1$) è una radice in $\mathbb{Q}$, allora r divide il termine noto e s divide il coefficiente direttivo di $f(x)$.*

DIMOSTRAZIONE. Sia $f(x) = \sum_{j=0}^{m} b_j x^j$ a coefficienti interi, l'ipotesi che r/s sia radice equivale a:

$$\sum_{i=0}^{n} b_i \left(\frac{r}{s}\right)^i = 0$$

[18] Esclusivamente per semplicità di notazione consideriamo anche i coefficienti nulli di $g(x)$ e $h(x)$ dei termini di grado maggiore rispettivamente di r e s. Ovvero $b_j = 0$ se $j > r$ e $c_t = 0$ se $t > s$.

Moltiplicando tutto per s^n si ottiene:

$$b_n r^n + \underbrace{b_{n-1} r^{n-1} s + ... + b_0 s^n}_{\text{è un multiplo di } s} = 0 \tag{6.5}$$

Per cui $s|b_n r^n$, ma essendo $(s, r) = 1$ questo implica $s|b_n$. Analogamente se raccogliamo in 6.5 r, otteniamo che r deve dividere $b_0 s^n$, ma essendo $(r, s) = 1$ questo implica che $r|b_0$. □

Esempio 8.83. Consideriamo il polinomio $f(x) = x^4 + 3x^3 + x^2 - 6x - 6$. Dalla Proposizione 8.82 segue che se r/s è una radice razionale, allora r divide -6 e s divide 1. Ovvero sappiamo che le uniche radici razionali possibili di $f(x)$ sono da ricercare nell'insieme finito:

$$A = \{\pm 1, \pm 2, \pm 3, \pm 6\}$$

Sostituendo in $f(x)$ non si trova 0 in nessuno di questi casi, dunque $f(x)$ non ha radici razionali.
ATTENZIONE: questo non significa che $f(x)$ sia irriducibile! Sappiamo solo che $f(x)$ non ha fattori di grado 1, ma potrebbe essere il prodotto di due fattori irriducibili di grado 2.

Esercizio 8.84. Il polinomio dell'esempio precedente è irriducibile in $\mathbb{Q}[x]$?
Suggerimento: se non vi riesce leggete più avanti...

Esercizio 8.85 (Divagazione aritmetica)**.** $\sqrt{2}$ è irrazionale.

Svolgimento Consideriamo il polinomio a coefficienti interi $x^2 - 2$. Per la Proposizione 8.82 non ha radici in $\mathbb{Q}$. Ora $\sqrt{2}$ è proprio una radice reale di $x^2 - 2$: dunque $\sqrt{2}$ non è razionale.

La Proposizione 8.82 è di fondamentale importanza in quanto limita ad un insieme finito e ristretto la ricerca di possibili radici razionali (e quindi fattori irriducibili di grado 1) di un polinomio a coefficienti interi. Questo permette per esempio di avere un algoritmo per discutere l'irriducibilità di polinomi di grado 2 e 3 in $\mathbb{Q}[x]$, infatti un polinomio di questo tipo o è irriducibile o ha una radice razionale.

Esercizio 8.86. Dire se $f(x) = x^3 - x^2 - 8x + 12$ è irriducibile in $\mathbb{Q}[x]$.

Risoluzione. I divisori del termine noto sono $\{\pm 1, \pm 2, \pm 3, \pm 4, \pm 6, \pm 12\}$, i divisori del coefficiente del termine di grado massimo sono $\{\pm 1\}$ quindi le possibili radici razionali sono: $\{\pm 1, \pm 2, \pm 3, \pm 4, \pm 6, \pm 12\}$. Proviamo a calcolare la funzione polinomiale $f(x)$ per questi valori fino a che non troviamo una radice; se non la troviamo vuol dire che $f(x)$ è irriducibile in $\mathbb{Q}[x]$:

$$f(1) = 4 \neq 0 \qquad f(-1) = 18 \neq 0 \qquad f(2) = 0$$

Dunque $f(x)$ è riducibile e ha $(x - 2)$ come fattore di grado 1.

A questo punto si potrebbe continuare a cercare altre radici razionali per vedere se ci sono altri fattori di $f(x)$ di grado 1 diversi da $(x - 2)$, ma forse nel caso di un polinomio di grado 3 conviene procedere dividendo $f(x)$ per $(x - 2)$ in modo da trovare un fattore di grado 2 che sappiamo dire se è riducibile o meno in $\mathbb{Q}[x]$

attraverso la formula risolutiva delle equazioni di secondo grado:

$$\begin{array}{cccc|l} x^3 & -x^2 & -8x & +12 & | & x-2 \\ x^3 & -2x^2 & & & | & x^2+x-6 \\ & x^2 & -8x & +12 & | & \\ & x^2 & -2x & & | & \\ & & -6x & +12 & | & \\ & & -6x & +12 & | & \\ & & & 0 & | & \end{array}$$

Quindi $f(x) = (x-2)\cdot(x^2+x-6)$. Si tratta di vedere se $x^2+x-6=0$ ha o meno due soluzioni razionali. Dalla formula risolutiva si ottiene:

$$x_{1,2} = \frac{-1 + \pm\sqrt{25}}{2} = \frac{-1 \pm 5}{2}$$

E quindi x^2+x-6 è riducibile in $\mathbb{Q}[x]$ e si fattorizza come $(x+3)\cdot(x-2)$. La fattorizzazione in irriducibili di $x^3-x^2-8x+12$ in $\mathbb{Q}[x]$ è dunque data da:

$$x^3 - x^2 - 8x + 12 = (x-2)^2 \cdot (x+3)$$

A questo punto cominciamo ad avere diversi strumenti per la fattorizzazione in $\mathbb{Q}[x]$: innanzitutto sappiamo che ci possiamo ridurre ad un polinomio, associato a quello di partenza, primitivo e a coefficienti interi. Sui polinomi primitivi a coefficienti interi conosciamo un criterio *diretto* di irriducibilità (Eisenstein). Inoltre, la fattorizzazione è molto più semplice in $\mathbb{Z}[x]$. Cerchiamo di capire perché riprendendo in mano il polinomio $f(x)$ dell'Esempio 8.83. Abbiamo già visto che non ha radici, dunque se è fattorizzabile è il prodotto di due polinomi di grado 2 (che per il lemma di Gauss possiamo supporre a coefficienti interi).

Consideriamo due generici polinomi di grado 2 in $\mathbb{Z}[x]$:

$$g(x) = ax^2 + bx + c$$
$$h(x) = dx^2 + ex + f$$

Per quanto osservato sopra, $f(x) = x^4 + 3x^3 + x^2 - 6x - 6$ è fattorizzabile se e solo se è il prodotto di due polinomi di grado 2, ovvero se e solo se esiste una soluzione del seguente sistema di 5 equazioni a coefficienti interi:

$$\begin{cases} 1 = a \cdot d \\ 3 = a \cdot e + b \cdot d \\ 1 = a \cdot f + b \cdot e + c \cdot d \\ -6 = b \cdot f + c \cdot e \\ -6 = c \cdot f \end{cases}$$

Sapere che, pur cercando la fattorizzazione in $\mathbb{Q}[x]$, possiamo risolvere in $\mathbb{Z}$ è di grande aiuto. Infatti risolvere *algoritmicamente* questo sistema in $\mathbb{Z}$ è possibile: ogni singola equazione infatti può avere solo un numero finito (anche uguale a 0) di soluzioni intere; studiando tutti i casi possibili e *risalendo* il sistema o si determina una soluzione intera o altrimenti si deduce che il sistema è irrisolvibile e dunque $f(x)$ è irriducibile in $\mathbb{Z}[x]$ e di conseguenza in $\mathbb{Q}[x]$. Questo procedimento di fattorizzazione in $\mathbb{Z}[x]$ risolvendo il sistema per casi è noto come **metodo della forza *bruta***. Applichiamo questo metodo al nostro sistema: vedremo così concretamente i vantaggi di sapere di potersi limitare a cercare soluzioni intere del

sistema. Da $1 = a \cdot d$ ad esempio, segue che o $a = d = 1$ oppure $a = d = -1$ (ma se $f(x) = g(x) \cdot h(x)$, allora $f(x) = -g(x) \cdot (-h(x))$ e dunque possiamo considerare $a = d = 1$). Andiamo dunque a riscriverci il nostro sistema:

$$\begin{cases} 1 = a \cdot d \\ 3 = a \cdot e + b \cdot d \\ 1 = a \cdot f + b \cdot e + c \cdot d \\ -6 = b \cdot f + c \cdot e \\ -6 = c \cdot f \end{cases} \leftrightarrow \begin{cases} a = d = 1 \\ 3 = e + b \\ 1 = f + b \cdot e + c \\ -6 = b \cdot f + c \cdot e \\ -6 = c \cdot f \end{cases}$$

Da $-6 = c \cdot f$ si ottiene che o $c = 1$ e $f = -6$, o $c = -1$ e $f = 6$, o $c = 2$ e $f = -3$ o infine $c = -2$ e $f = 3$ (essendo $g(x)$ e $h(x)$ dello stesso grado generici, il loro ruolo è completamente simmetrico e dunque non è necessario considerare anche i casi speculari tipo $c = 6$ e $f = -1$). Otteniamo dunque 4 sistemi con meno variabili. Bisogna studiarli tutti:

$$\begin{cases} a = d = 1 \\ 3 = e + b \\ 1 = -6 + b \cdot e + 1 \\ -6 = -6b + e \\ c = 1 \\ f = -6 \end{cases} \quad \begin{cases} a = d = 1 \\ 3 = e + b \\ 1 = 6 + b \cdot e - 1 \\ -6 = 6b - e \\ c = -1 \\ f = 6 \end{cases}$$

$$\begin{cases} a = d = 1 \\ 3 = e + b \\ 1 = -3 + b \cdot e + 2 \\ -6 = -3b + 2e \\ c = 2 \\ f = -3 \end{cases} \quad \begin{cases} a = d = 1 \\ 3 = e + b \\ 1 = 3 + b \cdot e - 2 \\ -6 = 3b - 2e \\ c = -2 \\ f = 3 \end{cases}$$

È facile verificare che i primi tre sistemi non hanno soluzioni intere (portano rispettivamente alle seguenti equazioni irrisolvibili in $\mathbb{Z}$: $5e = 12$, $7b = -3$, $5e = 3$), mentre l'ultimo ha soluzione (con $b = 0$ ed $e = 3$). Dunque esiste una fattorizzazione di $f(x)$ in $\mathbb{Q}[x]$ (nonostante $f(x)$ non abbia radici razionali):

$$\underbrace{x^4 + 3x^3 + x^2 - 6x - 6}_{f(x)} = (\underbrace{x^2 - 2}_{g(x)})(\underbrace{x^2 + 3x + 3}_{h(x)})$$

Osservazione 8.87. Osserviamo che applicando il metodo della forza bruta ad un polinomio $f(x)$ di $\mathbb{Q}[x]$ a coefficienti interi e monico, possiamo supporre che anche i due *polinomi-fattori* $g(x)$ e $q(x)$ siano monici. Infatti o i coefficienti direttivi sono entrambi uguali ad 1 o entrambi uguali a -1, ma in questo secondo caso possiamo considerare $-g(x)$ e $-q(x)$, che sono monici e il cui prodotto è sempre $f(x)$.

Per concludere la parte sulla fattorizzazione in $\mathbb{Q}[x]$ ed introdurre la fattorizzazione in $\mathbb{Z}_p[z]$ osserviamo che la irriducibilità in $\mathbb{Z}_p[x]$ e quella in $\mathbb{Q}[x]$ sono tra loro legate:

Proposizione 8.88 (Criterio della riduzione modulo p)**.** *Sia $f(x) = \sum_{i=0}^{n} a_i x^i$ primitivo a coefficienti interi e sia p un primo che non divide il coefficiente direttivo a_n di $f(x)$. Se il polinomio $\overline{f(x)}$ di $\mathbb{Z}_p[x]$ ottenuto riducendo i coefficienti modulo p (e considerando dunque le corrispondenti classi di resto) è irriducibile in $\mathbb{Z}_p[x]$ allora $f(x)$ è irriducibile in $\mathbb{Q}[x]$.*

DIMOSTRAZIONE. Se $f(x) = g(x) \cdot h(x)$ con $g(x)$ e $h(x)$ in $\mathbb{Z}[x]$ di grado maggiore o uguale a 1, allora $\overline{f(x)} = \overline{g(x)h(x)}$ e se p non divide a_n, ovvero $\overline{f(x)}$ ha lo stesso grado di $f(x)$, allora p non divide nemmeno i coefficienti direttivi di $g(x)$ e $h(x)$, ovvero anch'essi hanno il grado rispettivamente uguale a quello di $\overline{g(x)}$ e $\overline{h(x)}$. Dunque $\overline{f(x)} = \overline{g(x)h(x)}$ sarebbe una fattorizzazione di $\overline{f(x)}$ contro l'ipotesi di partenza. □

Esempio 8.89. Provare che il polinomio $f(x) = 5x^4 + 3x^3 + 5x + 1$ è irriducibile in $\mathbb{Q}[x]$.

Consideriamo il polinomio $\overline{f(x)}$ in $\mathbb{Z}_2[x]$:

$$\overline{f(x)} = x^4 + x^3 + x + 1$$

Questo polinomio è riducibile in $\mathbb{Z}_2[x]$ perché ha una radice, infatti:

$$\overline{f(1)} = 1 + 1 + 1 + 1 = 4 = 0$$

Dunque la Proposizione 8.88 non fornisce nessuna indicazione, infatti la proposizione NON PERMETTE DI CONCLUDERE NIENTE nel caso il polinomio $\overline{f(x)}$ ridotto modulo p sia riducibile in $\mathbb{Z}_p[x]$. Proviamo a considerare il polinomio $\overline{f(x)}$ in $\mathbb{Z}_3[x]$:

$$\overline{f(x)} = 2x^4 + 2x + 1$$

Questo polinomio non ha radici in $\mathbb{Z}_3[x]$ infatti:

$$\overline{f(0)} = 1 \quad \overline{f(1)} = 2 \quad \overline{f(2)} = 1$$

Dunque se $\overline{f(x)}$ è riducibile in $\mathbb{Z}_3[x]$ deve essere il prodotto di due polinomi irriducibili di grado 2. Impostiamo il sistema:

$$2x^4 + 2x + 1 = (ax^2 + bx + c)(dx^2 + ex + f)$$

Da cui, per la definizione di uguaglianza tra polinomi, si trova il seguente sistema di 5 equazioni in $\mathbb{Z}_3$:

$$\begin{cases} a \cdot d = 2 \\ a \cdot e + b \cdot d = 0 \\ a \cdot f + b \cdot e + c \cdot d = 0 \\ b \cdot f + c \cdot e = 2 \\ c \cdot f = 1 \end{cases}$$

Procediamo con il metodo della forza bruta[19]: da $a \cdot d = 2$ segue che $a = 2$ e $d = 1$[20], mentre da $c \cdot f = 1$ segue che $c = f = 1$ oppure $c = f = 2$. Studiamo i due casi per vedere se esiste una soluzione e quindi se $\overline{f(x)}$ è riducibile in $\mathbb{Z}_3[x]$:

[19] Sottolineiamo come in $\mathbb{Z}_p[x]$ esistano metodi algoritmici molto più efficienti per la fattorizzazione.

[20] Come già osservato ad ora i due fattori, entrambi di grado 2, sono completamente simmetrici e dunque possiamo fissare quale coefficiente tra a e d è uguale ad 1 e quale è uguale a 2, diverso sarebbe stato se i due fattori fossero stati di grado differente.

(1)
$$\begin{cases} a = 1 \\ d = 2 \\ c = f = 1 \\ e + 2b = 0 \\ 1 + b \cdot e + 2 = 0 \\ b + e = 2 \end{cases}$$
Questo sistema è impossibile, infatti da $e + 2b = 0$ ed $e + b = 2$ segue che $b = e = 1$, ma questo non è compatibile con l'altra equazione $1+b\cdot e+2 = 0$.

(2)
$$\begin{cases} a = 1 \\ d = 2 \\ c = f = 2 \\ e + 2b = 0 \\ 2 + b \cdot e + 1 = 0 \\ 2b + 2e = 2 \end{cases}$$
Questo sistema è impossibile, infatti da $e+2b = 0$ ed $2e+2b = 2$ segue che $b = e = 2$, ma questo non è compatibile con l'altra equazione $2+b\cdot e+1 = 0$.

Dunque $\overline{f(x)} = 2x^4 + 2x + 1$ è irriducibile in $\mathbb{Z}_3[x]$ e dalla Proposizione 8.88 segue che $f(x)$ è irriducibile in $\mathbb{Q}[x]$.

6.4. Fattorizzazione in $\mathbb{Z}_p[x]$. Anche in questo caso vale, come per $\mathbb{Q}[x]$, che per ogni n esiste un polinomio di grado n in $\mathbb{Z}_p[x]$ irriducibile. Rispetto a $\mathbb{Q}[x]$ abbiamo però un algoritmo finito elementare (anche se può essere molto dispendioso come tempo) per mostrare che un polinomio è irriducibile o trovarne una fattorizzazione in irriducibili. Sia infatti $f(x) \in \mathbb{Z}_p[x]$ di grado n; allora se è riducibile ha un fattore irriducibile che ha grado minore o uguale a $n/2$ se n è pari e a $(n-1)/2$ se n è dispari. Essendo $\mathbb{Z}_p$ finito i polinomi di grado minore o uguale di un fissato k sono finiti (sono p^{k+1}) e quindi un modo per trovare una fattorizzazione di $f(x)$ è provare a dividere per tutti i polinomi di grado minore o uguale di $n/2$ (o $(n-1)/2$ nel caso n dispari). Se nessuno di questi divide $f(x)$ allora $f(x)$ è irriducibile, altrimenti continuiamo con lo stesso procedimento fino a che non si scrive $f(x)$ come prodotto di irriducibili. Proviamo con questo metodo a risolvere il seguente esercizio:

Esercizio 8.90. Dimostrare che $f(x) = x^5 + x^2 + 1$ è irriducibile in $\mathbb{Z}_2[x]$

Risoluzione. Per prima cosa vediamo di escludere che ci siano fattori di grado 1 che fattorizzano $f(x)$. Per questo basta mostrare che $f(x)$ non ha radici in $\mathbb{Z}_2$, in effetti $f(0) = 1$ e $f(1) = 3 = 1$. Mostrato che non ci possono essere fattori di grado 1 l'unica possibile fattorizzazione di $f(x)$ può essere come un polinomio di grado 3 per un polinomio di grado 2. I polinomi di grado uguale a 2 in $\mathbb{Z}_2[x]$ sono x^2+x+1, $x^2 + x$ e $x^2 + 1$ e x^2...allora con pazienza facciamo la divisione di $f(x)$ per questi quattro polinomi. $f(x)$ risulterà irriducibile se da queste divisioni non verrà mai resto uguale a zero, cioè se nessun polinomio di grado 2 in $\mathbb{Z}_2[x]$ divide $f(x)$:

$$\begin{array}{ccccccc|l} x^5 & & & & +x^2 & +1 & | & x^2+x+1 \\ x^5 & +x^4 & +x^3 & & & & | & x^3+x^2 \\ & x^4 & +x^3 & +x^2 & & +1 & | & \\ & x^4 & +x^3 & +x^2 & & & | & \\ & & & & & +1 & | & \end{array}$$

Quindi $f(x) = (x^2+x+1)\cdot(x^3+x^2) + \underbrace{1}_{resto\ diverso\ da\ 0}$

$$\begin{array}{rrrrr|l}
x^5 & & & +x^2 & +1 & x^2+x \\
x^5 & +x^4 & & & & x^3+x^2+x \\
 & x^4 & & +x^2 & +1 & \\
 & x^4 & +x^3 & & & \\
 & & x^3 & +x^2 & +1 & \\
 & & x^3 & +x^2 & & \\
 & & & & +1 &
\end{array}$$

Quindi $f(x) = (x^2+x)\cdot(x^3+x^2+x) + \underbrace{1}_{resto\ diverso\ da\ 0}$

$$\begin{array}{rrrrr|l}
x^5 & & +x^2 & & +1 & x^2+1 \\
x^5 & +x^3 & & & & x^3+x+1 \\
 & x^3 & +x^2 & & +1 & \\
 & x^3 & & +x & & \\
 & & x^2 & +x & +1 & \\
 & & x^2 & & +1 & \\
 & & & x & &
\end{array}$$

Quindi $f(x) = (x^2+1)\cdot(x^3+x+1) + \underbrace{x}_{resto\ diverso\ da\ 0}$

$$\begin{array}{rrrr|l}
x^5 & +x^2 & +1 & & x^2 \\
x^5 & & & & x^3+1 \\
 & x^2 & +1 & & \\
 & x^2 & & & \\
 & & +1 &
\end{array}$$

Quindi $f(x) = (x^2)\cdot(x^3+1) + \underbrace{1}_{resto\ diverso\ da\ 0}$

Esercizio 8.91. Determinare e contare i polinomi irriducibili in $\mathbb{Z}_2[x]$ di grado minore o uguale a 4.

Prima di procedere a scrivere la lista dei polinomi irriducibili in $\mathbb{Z}_2[x]$, osserviamo che questa lista potrà essere usata negli esercizi di fattorizzazione in $\mathbb{Z}_2[x]$ ma anche per mostrare la irriducibilità di qualche polinomio in $\mathbb{Q}[x]$ (vedi Proposizione 8.88). Infatti se cerchiamo un fattore di grado $d \leq 4$ di un polinomio $f(x)$ in $\mathbb{Z}_2[x]$ potremo usare questa lista per limitare il numero di divisioni da eseguire.

(1) **Polinomi irriducibili di grado 1 in $\mathbb{Z}_2[x]$**
Essendo $\mathbb{Z}_2$ un campo, tutti i polinomi di primo grado di $\mathbb{Z}_2[x]$ sono irriducibili. I polinomi di primo grado in $\mathbb{Z}_2[x]$ sono due: x e $x+1$.

(2) **Polinomi irriducibili di grado 2 in $\mathbb{Z}_2[x]$**
I polinomi di grado 2 in $\mathbb{Z}_2[x]$ sono 4:

$$x^2 \quad x^2+x \quad x^2+x+1 \quad x^2+1$$

Sappiamo che un polinomio di grado 2 è riducibile se e solo se ha una radice. Ora x^2 ha radice 0 doppia e infatti è fattorizzato come x per x,

$x^2 + x$ ha radici 0 e 1 e infatti è fattorizzato come x per $x + 1$, $x^2 + 1$ ha radice 1 doppia e infatti (in $\mathbb{Z}_2[x]$!) è fattorizzato come $x + 1$ per $x + 1$. Il polinomio $x^2 + x + 1$ invece, valutato in 0 e in 1 vale 1, di conseguenza non ha radici e dunque è l'unico irriducibile di grado 2 in $\mathbb{Z}_2[x]$.

(3) **Polinomi irriducibili di grado 3 in $\mathbb{Z}_2[x]$**
I polinomi di grado 3 in $\mathbb{Z}_2[x]$ sono 8:

$$\begin{array}{cccc} x^3 + x^2 & x^3 + x^2 + x & x^3 + x^2 + x + 1 & x^3 + x^2 + 1 \\ x^3 & x^3 + x & x^3 + x + 1 & x^3 + 1 \end{array}$$

Anche i polinomi di grado 3 sono riducibili se e solo se hanno una radice. Cercando le radici (è molto facile in $\mathbb{Z}_2$ che ha solo due elementi: basta calcolare il valore del polinomio in 0 e 1) e applicando - nel caso che i polinomi abbiano effettivamente radici - Ruffini, abbiamo il seguente quadro degli 8 polinomi di grado 3 di $\mathbb{Z}_2[x]$:

$$\begin{array}{cc} x^3 + x^2 = x \cdot x \cdot (x + 1) & x^3 + x^2 + x = x \cdot (x^2 + x + 1) \\ x^3 + x^2 + x + 1 = (x + 1)^3 & x^3 + x^2 + 1 \\ x^3 = x \cdot x \cdot x & x^3 + x = x \cdot (x + 1)^2 \\ x^3 + x + 1 & x^3 + 1 = (x + 1) \cdot (x^2 + x + 1) \end{array}$$

Dunque esistono solo due polinomi irriducibili in $\mathbb{Z}_2[x]$: $x^3 + x^2 + 1$ e $x^3 + x + 1$.

Osservazione 8.92. Osserviamo che in $\mathbb{Z}_2[x]$ una condizione sufficiente (ma non necessaria) affinché un polinomio $f(x)$ di grado maggiore di 1 sia riducibile è che $f(x)$ abbia un numero pari $2k$ di termini, in quanto valutando $f(x)$ in 1 si ottiene $2k$ che è 0 in $\mathbb{Z}_2$. Dunque 1 è radice di tutti i polinomi con un numero pari di termini in $\mathbb{Z}_2[x]$ e per il teorema di Ruffini questo equivale al fatto che $x - 1$ (che in $\mathbb{Z}_2[x]$ è $x + 1$) divide tutti i polinomi con un numero pari di termini. Ad esempio $x^3 + x^2 + x + 1$ ha 4 termini e infatti valutato in 1 dá come risultato 4, ovvero 0 in $\mathbb{Z}_2[x]$.

(4) **Polinomi irriducibili di grado 4 in $\mathbb{Z}_2[x]$**
I polinomi di grado 4 in $\mathbb{Z}_2[x]$ sono 16 di cui 8 con un numero pari di termini e dunque sicuramente riducibili:

$$\begin{array}{cc} x^4 + x^3 & x^4 + x^3 + x^2 + x \\ x^4 + x^3 + x + 1 & x^4 + x^3 + x^2 + 1 \\ x^4 + x^2 + x + 1 & x^4 + x^2 \\ x^4 + x & x^4 + 1 \end{array}$$

Rimangono 8 polinomi con un numero dispari di termini, di cui dobbiamo discutere l'irrudicibilità. Come abbiamo osservato, il criterio del numero pari di termini è una condizione sufficiente, ma non necessaria, per essere irriducibili in $\mathbb{Z}_2[x]$. In particolare anche tutti i polinomi che si annullano in 0 sono riducibili; tra gli 8 che sono rimasti ce ne sono 4: x^4, $x^4 + x^3 + x^2$, $x^4 + x^3 + x$ e $x^4 + x^2 + x$.

A questo punto rimangono solo i 4 polinomi di quarto grado che non hanno radici in $\mathbb{Z}_2[x]$:

$$x^4 + x^3 + x^2 + x + 1, \quad x^4 + x^3 + 1, \quad x^4 + x^2 + 1, \quad x^4 + x + 1.$$

Non è detto però che tutti siano irriducibili: potrebbero essere il prodotto di polinomi di grado 2 irriducibili. Sappiamo che in $\mathbb{Z}_2[x]$ l'unico polinomio

irriducibile di grado 2 è x^2+x+1, dunque l'unico polinomio di grado 4 di $\mathbb{Z}_2[x]$ che non ha radici ma è riducibile è:

$$(x^2+x+1)^2 = x^4+x^2+1$$

Perciò ci sono 3 polinomi di grado 4 irriducibili in $\mathbb{Z}_2[x]$ e sono:

$$x^4+x^3+x^2+x+1, \quad x^4+x^3+1, \quad x^4+x+1.$$

7. Esercizi sulla fattorizzazione

Esercizio 8.93 (Compito d'esame 2005)**.** Sia $g(x) \in \mathbb{R}[x]$ il polinomio

$$g(x) = x^3 - 2x^2 + 2x - 1$$

(1) Fattorizzare $g(x)$ in prodotto di polinomi irriducibili.
(2) Considerato il polinomio

$$f_a(x) = x^4 - 2ax^2 + 2ax - 1$$

dimostrare che, per ogni $a \in \mathbb{R}$, un M.C.D. tra $g(x)$ e $f_a(x)$ è il polinomio $x-1$.

Risoluzione. Sappiamo che il polinomio $g(x)$ è riducibile in $\mathbb{R}[x]$, in quanto ha grado 3. Questo in particolare significa che $g(x)$ ha una radice reale. Osserviamo che non abbiamo studiato formule risolutive delle equazioni di terzo grado, quindi con i nostri strumenti possiamo trovare questa radice solo se è razionale (il polinomio che stiamo considerando in $\mathbb{R}[x]$ è a coefficienti interi): possiamo cioè provare tutte le possibili radici razionali che otteniamo dai divisori del coefficiente direttivo e del termine noto.
Però leggendo il testo dell'esercizio non abbiamo bisogno nemmeno di questo passaggio, infatti se dobbiamo mostrare che $x-1$ è un M.C.D. di $g(x)$ con un altro polinomio, allora $x-1$ dovrà essere un divisore di $g(x)$ (e quindi 1 una radice di $g(x)$). Andiamo a verificare che $x-1$ è un fattore irriducibile di $g(x)$: che sia irriducibile è certo, visto che è di grado 1; dobbiamo mostrare che effettivamente è un divisore di $g(x)$ (se così non fosse potremmo intanto concludere che l'affermazione della seconda parte dell'esercizio è falsa). In realtà si vede subito che $x-1$ è un divisore perchè $g(1) = 1-2+2-1 = 0$, ma a noi per la fattorizzazione interessa comunque dividere i due polinomi:

$$\begin{array}{cccc|l} x^3 & -2x^2 & +2x & -1 & x-1 \\ x^3 & -x^2 & & & x^2-x+1 \\ \cline{1-3} & -x^2 & +2x & -1 & \\ & -x^2 & +x & & \\ \cline{2-3} & & x & -1 & \\ & & & 0 & \end{array}$$

Abbiamo trovato che $g(x) = (x-1)\cdot(x^2-x+1)$, a questo punto verifichiamo se x^2-x+1 è riducibile o meno in $\mathbb{R}[x]$ attraverso il calcolo del delta: essendo negativo ($\Delta = 1-4 = -3$) il polinomio è irriducibile in $\mathbb{R}[x]$ e quindi la fattorizzazione cercata è proprio:

$$g(x) = (x-1)\cdot(x^2-x+1).$$

A questo punto per dimostrare che $x-1$ è un $M.C.D.(g(x), f_a(x))$ cominciamo mostrando che $x-1$ divide $f_a(x)$ per ogni $a \in \mathbb{R}$ (e quindi è un fattore comune). Basta osservare che $f_a(1) = 1-2a+2a-1 = 0$. Ora se mostriamo che x^2-x+1 non

è un divisore di $f_a(x)$ per qualsiasi scelta di a in $\mathbb{R}$, abbiamo la tesi. Procediamo dunque calcolando il resto della divisione di $f_a(x)$ per $x^2 - x + 1$, che sarà un polinomio $r_a(x)$ che dipenderà dal coefficiente a. Dovremo osservare che $r_a(x)$ non è uguale al polinomio nullo qualsiasi sia la scelta di a in $\mathbb{R}$:

$$\begin{array}{cccccc|l}
x^4 & & -2ax^2 & +2ax & -1 & | & x^2 - x + 1 \\
x^4 & -x^3 & +x^2 & & & | & x^2 + x - 2a \\
 & x^3 & +x^2 \cdot (-1 - 2a) & +2ax & -1 & | & \\
 & x^3 & -x^2 & +x & & | & \\
 & & -2a \cdot x^2 & +x \cdot (2a - 1) & -1 & | & \\
 & & -2a \cdot x^2 & +2a \cdot x & -2a & | & \\
 & & & -x & -1 + 2a & | &
\end{array}$$

Osserviamo che il polinomio resto $r_a(x)$ è sempre di grado 1 qualsiasi sia la scelta di a in $\mathbb{R}$: in particolare non sarà mai uguale al polinomio nullo.

Esercizio 8.94. Dato il polinomio $g(x) = 4x^3 + 5x^2 + 3x + 1$ fattorizzarlo in prodotto di irriducibili in $\mathbb{Q}[x]$ e in $\mathbb{Z}_{13}[x]$.

Risoluzione. Sappiamo che un polinomio di grado 3 è sicuramente riducibile in $\mathbb{R}[x]$ o in $\mathbb{C}[x]$, ma non conosciamo un algoritmo per trovare questa fattorizzazione. In $\mathbb{Q}[x]$ e in $\mathbb{Z}_p[x]$ un polinomio di grado 3 non sappiamo se è riducibile o no, ma abbiamo un algoritmo finito per rispondere a questa domanda e per trovare un'eventuale fattorizzazione in irriducibili del polinomio stesso. Questo perchè, come già osservato, la riducibilità di un polinomio di grado 3 è equivalente all'esistenza di una radice nel campo. Nel caso della riducibilità in $\mathbb{Q}[x]$ se il polinomio è a coefficienti interi (come $g(x)$) la Proposizione 8.82 permette di limitare le possibili radici razionali ad un insieme finito (tramite il calcolo dei divisori del termine noto e del coefficiente direttivo), mentre nel caso della riducibilità in $\mathbb{Z}_p[x]$ il numero delle possibili radici è ovviamente finito in quanto è finito il campo dei coefficienti. I divisori del coefficiente direttivo sono $\{\pm 1, \pm 2 \pm 4\}$ mentre quelli del termine noto sono $\{\pm 1\}$, quindi le possibili radici razionali di $g(x)$ sono i numeri: $\{\pm\frac{1}{2}, \pm\frac{1}{4}, \pm 1\}$. Proviamoli, ma prima osserviamo che il polinomio $g(x)$ ha tutti coefficienti positivi e quindi non potrà avere radici positive. Ci possiamo dunque limitare a provare, tra le possibili radici razionali, quelle negative:

$$\begin{array}{l}
g(-\frac{1}{4}) = -\frac{1}{16} + \frac{5}{16} - \frac{3}{4} + 1 = \frac{1}{2} \\
g(-1) = -4 + 5 - 3 + 1 = -1 \\
g(-\frac{1}{2}) = -\frac{1}{2} + \frac{5}{4} - \frac{3}{2} + 1 = \frac{1}{4}
\end{array}$$

$g(x)$ non ha dunque radici razionali e quindi è irriducibile in $\mathbb{Q}[x]$.
Per quanto riguarda $\mathbb{Z}_{13}[x]$ valutando $g(x)$ per tutti gli elementi del campo si può verificare se esistono una o più radici. In questo caso troviamo $g(1) = 13 = 0$, quindi $g(x)$ è riducibile in $\mathbb{Z}_{13}[x]$ perché ha una radice e dunque per Ruffini è divisibile per

$x-1$:

$$\begin{array}{llll|l} 4x^3 & +5x^2 & +3x & +1 & x-1 \\ 4x^3 & -4x^2 & & & 4x^2+9x+12 \\ \cline{1-4} & 9x^2 & +3x & +1 & \\ & 9x^2 & -9x & & \\ \cline{2-4} & & 12x & +1 & \\ & & 12x & -12 & \\ \cline{3-4} & & & +13 & \\ & & & 0 & \end{array}$$

Dunque $g(x) = (x-1)\cdot(4x^2+9x+12)$ in $\mathbb{Z}_{13}[x]$, si tratta di vedere se $4x^2+9x+12$ è irriducibile o meno in $\mathbb{Z}_{13}[x]$. Per questo si può procedere in due modi: o si provano tutti gli elementi di $\mathbb{Z}_{13}[x]$ alla ricerca di un'eventuale radice, oppure si usa la seguente osservazione:

Osservazione 8.95. La formula per la risoluzione delle equazioni di secondo grado vale in ogni campo $\mathbb{K}$ (e quindi in particolare per campi finiti).

Dimostrazione. Supponiamo di dover risolvere:

$$ax^2 + bx + c = 0 \tag{7.1}$$

con a, b, c appartenenti ad un qualsiasi campo $\mathbb{K}$ e $a \neq 0$ (questo per garantire che effettivamente stiamo risolvendo un'equazione di secondo grado). Ripercorriamo i passi che portano alla formula risolutiva delle equazioni reali di secondo grado per far vedere che le uniche cose che usiamo sono le proprietà di campo di $\mathbb{R}$:

(1) Sommiamo ad entrambi i membri di 7.1 l'opposto di c:

$$ax^2 + bx = -c \tag{7.2}$$

(2) Moltiplichiamo entrambi i membri per l'inverso di a che indichiamo con a^{-1} (sappiamo che esiste in $\mathbb{K}$ l'inverso di $a \neq 0$):

$$x^2 + a^{-1}\cdot bx = a^{-1}\cdot(-c) \tag{7.3}$$

(3) Aggiungiamo ad entrambi i membri di 7.3 $[(2a)^2]^{-1}\cdot b^2$:

$$x^2 + a^{-1}\cdot bx + [(2a)^2]^{-1}\cdot b^2 = a^{-1}\cdot(-c) + [(2a)^2]^{-1}\cdot b^2 \tag{7.4}$$

(4) È facile vedere (sfruttando la commutatività in $\mathbb{K}$) che il primo membro di 7.4 non è nient'altro che $(x + (2a)^{-1}\cdot b)^2$, si ha dunque:

$$(x + (2a)^{-1}\cdot b)^2 = a^{-1}\cdot(-c) + [(2a)^2]^{-1}\cdot b^2$$

che ha soluzione in $\mathbb{K}$ se e solo se:

$$a^{-1}\cdot(-c) + [(2a)^2]^{-1}\cdot b^2 = [(2a)^2]^{-1}\cdot(b^2 - 4a\cdot c)$$

è un quadrato in $\mathbb{K}$.

Per concludere, basta osservare che $a \neq 0$ è un quadrato in $\mathbb{K}$ se e solo se a^{-1} è un quadrato in $\mathbb{K}$ e quindi $[(2a)^2]^{-1}$ è sempre un quadrato. Perciò l'equazione 7.1 ha soluzione in $\mathbb{K}$ se e solo se $b^2 - 4a\cdot c$ (che solitamente indichiamo con Δ) è un quadrato in $\mathbb{K}$. Se in $\mathbb{K}$ esiste radice di Δ e Δ è diverso da zero, allora ne esistono

esattamente 2 distinte[21]. Le soluzioni dell'equazione 7.1 in questo caso sono allora due distinte e si ottengono sommando $-(2a)^{-1} \cdot b$ alle radici di Δ. □

Il Δ in questo caso è uguale a $81 - 192 = -111$ che in $\mathbb{Z}_{13}$ è equivalente a 6. Dobbiamo controllare se 6 è un quadrato in $\mathbb{Z}_{13}$:

$$0^2 = 0, 1^2 = 1, 2^2 = 4, 3^2 = 9, 4^2 = 16 = 3, 5^2 = 25 = 12, 6^2 = 10$$

E qui ci possiamo fermare perchè in $\mathbb{Z}_{13}$ $7 = -6$, $8 = -5$, $9 = -4$, $10 = -3$, $11 = -2$, $12 = -1$ e quindi i loro quadrati sono identici. Si può dunque concludere che 6 non è un quadrato in $\mathbb{Z}_{13}$ e quindi $4x^2 + 9x + 12$ è irriducibile in $\mathbb{Z}_{13}[x]$.

Esercizio 8.96. Fattorizzare il polinomio $f(x) = x^5 + x^2 + 1$ in $\mathbb{Q}[x]$.

Risoluzione. Il polinomio $f(x)$ (di cui abbiamo già studiato la riducibilità, ma in $\mathbb{Z}_2[x]$, nell'Esercizio 8.90) non ha radici in $\mathbb{Q}[x]$. Infatti dalla Proposizione 8.82 sappiamo che le uniche possibili radici razionali di $f(x)$ sono 1 e -1, ma valutando il polinomio in questi due valori si ottiene:

$$f(1) = 3 \quad f(-1) = 1$$

Il teorema di Ruffini ci dice dunque che $f(x)$ non ha fattori lineari in $\mathbb{Q}[x]$. A questo punto o $f(x)$ è irriducibile o è il prodotto di due polinomi irriducibili rispettivamente di secondo e terzo grado. Procediamo con il metodo della forza bruta (sappiamo che possiamo prendere i due eventuali polinomi fattore monici, vedi Osservazione 8.87):

$$\begin{aligned} &x^5 + x^2 + 1 = (x^3 + ax^2 + bx + c)(x^2 + dx + e) = \\ &= x^5 + (a + d)x^4 + (e + ad + b)x^3 + (ae + bd + c)x^2 + (be + cd)x + ce \end{aligned}$$

Abbiamo dunque il seguente sistema a coefficienti interi:

$$\begin{cases} a + d = 0 \\ e + ad + b = 0 \\ ae + bd + c = 1 \\ be + cd = 0 \\ ce = 1 \end{cases}$$

Da $ce = 1$ seguono due possibilità $c = e = 1$ oppure $c = e = -1$, in entrambi i casi si ha $b = a = -d$. Sostituendo in $e + ad + b = 0$ si ottiene, nel caso $e = 1$:

$$a^2 - a - 1 = 0$$

e nel caso $e = -1$:

$$a^2 - a + 1 = 0$$

In entrambi i casi non esistono soluzioni intere. Dunque il metodo della forza bruta ci dice che il polinomio $f(x)$ è irriducibile in $\mathbb{Q}[x]$.

Esercizio 8.97. Fattorizzare il polinomio $f(x) = x^4 - 1$ in $\mathbb{Z}_5[x]$.

---

[21] Supponiamo $\Delta \neq 0$ abbia radice in $\mathbb{K}$ allora l'equazione $x^2 = \Delta$ è equivalente a

$$(x - \sqrt{\Delta}) \cdot (x + \sqrt{\Delta}) = 0$$

che in un campo, dove non ci sono divisori di zero, ha esattamente due soluzioni distinte $\sqrt{\Delta}$ e $-\sqrt{\Delta}$.

Risoluzione. Il polinomio $f(x)$ ha 1 come radice, dunque per il teorema di Ruffini è divisibile per $x-1$. Osserviamo prima di proseguire che il risultato della divisione restituirà $f(x)$ come prodotto di $x-1$ per un polinomio $g(x)$ di terzo grado. Per completare la fattorizzazione di $f(x)$ dovremo dunque studiare la riducibilità di $g(x)$ che, essendo di terzo grado, è equivalente alla ricerca di radici in $\mathbb{Z}_5[x]$ del polinomio suddetto. Procediamo ora con la divisione di $f(x)$ per $x-1$:

$$\begin{array}{cccccc|l}
x^4 & & & & -1 & | & x-1 \\
x^4 & -x^3 & & & & | & x^3+x^2+x+1 \\
 & x^3 & & & -1 & | & \\
 & x^3 & -x^2 & & & | & \\
 & & x^2 & & -1 & | & \\
 & & x^2 & -x & & | & \\
 & & & x & -1 & | & \\
 & & & x & -1 & | & \\
 & & & & 0 & | &
\end{array}$$

Dunque:

$$x^4-1=(x-1)\underbrace{(x^3+x^2+x+1)}_{g(x)}$$

Valutiamo se $g(x)$ ha radici in $\mathbb{Z}_5$:

$$g(0)=1 \quad g(1)=4 \quad g(2)=15=0 \quad g(3)=40=0 \quad g(4)=85=0$$

Perciò da Ruffini segue che $g(x)$ è fattorizzabile come:

$$g(x)=(x-2)(x-3)(x-4)$$

Concludendo:

$$f(x)=(x-1)(x-2)(x-3)(x-4)$$

Osserviamo che potevamo arrivare alla conclusione in maniera molto più rapida sfruttando le proprietà degli $\mathbb{Z}_p$ ed in particolare il piccolo teorema di Fermat. Infatti sappiamo che il polinomio x^5-x si annulla per ogni valore di $\mathbb{Z}_5$ e basta osservare che:

$$x^5-x=x(x^4-1)$$

Ovvero x^4-1 si annulla in tutti gli elementi di $\mathbb{Z}_5$ tranne che in 0 e dunque è fattorizzabile proprio come:

$$f(x)=(x-1)(x-2)(x-3)(x-4)$$

Esercizio 8.98. Sia $p(x)=x^4-4x^3+6x^2-4x+5$. Sapendo che $2+i$ è una radice complessa del polinomio $p(x)$ fattorizzarlo in $\mathbb{R}[x]$ e in $\mathbb{C}[x]$.

Risoluzione. Se $\alpha=2+i$ è radice, allora (Proposizione 8.73) anche il suo complesso coniugato $\overline{\alpha}=2-i$ è radice di $p(x)$. Dunque il polinomio è divisibile per:

$$(x-(2+i))(x-(2-i))=(x-2)^2-i^2=x^2-4x+4+1=x^2-4x+5$$

Eseguiamo la divisione:

$$\begin{array}{ccccc|l}
x^4 & -4x^3 & +6x^2 & -4x & +5 & | & x^2-4x+5 \\
x^4 & -4x^3 & +5x^2 & & & | & x^2+1 \\
 & & x^2 & -4x & +5 & | & \\
 & & x^2 & -4x & +5 & | & \\
 & & & & 0 & | &
\end{array}$$

Abbiamo dunque trovato che:

$$p(x) = (x^2 - 4x + 5)(x^2 + 1)$$

che è la fattorizzazione in irriducibili in $\mathbb{R}[x]$, infatti entrambi i polinomi di secondo grado non hanno soluzioni reali. Visto che $x^2 + 1$ ha come radici complesse i e $-i$ la fattorizzazione in irriducibili di $p(x)$ in $\mathbb{C}[x]$ è:

$$p(x) = (x - (2 + i))(x - (2 - i))(x - i)(x + i)$$

Esercizio 8.99. Elencare tutti i polinomi irriducibili di grado minore o uguale a 3 in $\mathbb{Z}_3[x]$.

Esercizio 8.100. Trovare un massimo comune divisore fra i seguenti polinomi appartenenti a $\mathbb{Q}[x]$:

$$f(x) = x^3 + x^2 + 7x + 7$$

$$g(x) = x^4 + x^3 + 2x^2 + 4x + 2$$

Spiegare come mai è possibile trovare due polinomi $a(x)$ e $b(x)$ in $\mathbb{Q}[x]$ tali che

$$a(x)f(x) + b(x)g(x) = x^2 + x$$

Tali polinomi sono unici ?
Trovare esplicitamente due polinomi $a(x)$ e $b(x)$ in $\mathbb{Q}[x]$ che soddisfano l'equazione del punto precedente.

Esercizio 8.101. Fattorizzare il polinomio $x^4+4x^3-19x^2+8x-42$ come prodotto di irriducibili in $\mathbb{R}[x],\ \mathbb{Q}[x],\ \mathbb{Z}_3[x],\ \mathbb{Z}_{13}[x]$.

Esercizio 8.102. Fattorizzare il polinomio $x^4 - 4x^3 + x^2 + 8x - 6$ come prodotto di irriducibili in $\mathbb{R}[x],\ \mathbb{Q}[x],\ \mathbb{Z}_7[x],\ \mathbb{Z}_{11}[x]$.

Esercizio 8.103. Consideriamo il polinomio

$$p(x) = x^4 - x^3 - x^2 - x - 2$$

Trovare la scomposizione di $p(x)$ in fattori irriducibili in $\mathbb{R}[x]$, $\mathbb{C}[x]$, $\mathbb{Z}_3[x]$.

Esercizio 8.104. Consideriamo in $\mathbb{K}[x]$ ($\mathbb{K}$ campo), il seguente polinomio, dipendente dal parametro $a \in \mathbb{K}$:

$$4x^3 + (2a + 4)x^2 + (2a + 1)x + 1$$

a) Fattorizzare il polinomio quando $\mathbb{K} = \mathbb{R}$
b) Fattorizzare il polinomio quando $\mathbb{K} = \mathbb{C}$
c) Fattorizzare il polinomio quando $\mathbb{K} = \mathbb{Z}_3$.

Esercizio 8.105. Fattorizzare il polinomio $f(x) = x^6 - x^5 - 2x^4 - 2x^2 + 2x + 4$ in $\mathbb{Q}[x], \mathbb{R}[x], \mathbb{C}[x],\ \mathbb{Z}_2[x], \mathbb{Z}_3[x]$.

Esercizio 8.106. Dimostrare che per ogni $p \in \mathbb{N}$ primo, il polinomio:

$$\sum_{i=0}^{p-1} x^i$$

è irriducibile in $\mathbb{Q}[x]$.

Esercizio 8.107. Sia V uno spazio vettoriale di dimensione finita su un campo $\mathbb{K}$ e sia $T : V \to V$ un endomorfismo lineare. Siano $h(t), g(t)$ due polinomi in $\mathbb{K}[t]$ il cui massimo comun divisore è 1. Dimostrare che

$$V = Ker\ h(T) \oplus Ker\ g(T)$$

Esercizio 8.108. Sia V uno spazio vettoriale di dimensione finita su un campo $\mathbb{K}$ e sia $T : V \to V$ un endomorfismo lineare. Sia I_T l'insieme dei polinomi $f(t)$ in $\mathbb{K}[t]$ tali che $f(T)$ è l'endomorfismo nullo.
a) Dimostrare che I_T è un *ideale* dell'anello $\mathbb{K}[t]$, ovvero è un gruppo rispetto alla operazione $+$ e soddisfa inoltre la seguente proprietà: per ogni $f(t) \in I_T$ e per ogni $g(t) \in \mathbb{K}[t]$ vale $f(t)g(t) \in I_T$.
b) Dimostrare che esiste un polinomio monico $\mu_T(t) \in I_T$ che divide ogni altro polinomio in I_T. Tale polinomio si chiama *polinomio minimo di* T.

Esercizio 8.109. Sia V uno spazio vettoriale di dimensione finita su un campo $\mathbb{K}$ e sia $T : V \to V$ un endomorfismo lineare. Il polinomio minimo $\mu_T(t)$ divide il polinomio caratteristico $P_T(t)$, visto che $P_T(T)$ è l'endomorfismo nullo (vedi teorema di Cayley-Hamilton, Esercizio 7.38). Dimostrare che ogni radice di $P_T(t)$ è anche una radice di $\mu_T(t)$, ovvero che le radici di $\mu_T(t)$ in $\mathbb{K}$ sono tutti e soli gli autovalori di T in $\mathbb{K}$.

Esercizio 8.110. Sia V uno spazio vettoriale di dimensione finita su un campo $\mathbb{K}$ e sia $T : V \to V$ un endomorfismo lineare. Dimostrare che T è diagonalizzabile se e solo se il polinomio minimo $\mu_T(t)$ si fattorizza in $\mathbb{K}[t]$ come prodotto di fattori di grado 1 tutti distitni fra loro, ovvero:

$$\mu_T(t) = (t - \lambda_1) \cdots (t - \lambda_k)$$

con i λ_i a due a due distinti.
Suggerimento: usare l'Esercizio 8.107.

Indice analitico

Bibliografia

[Ab] M. Abate, Algebra Lineare, McGraw-Hill.
[C] L.Childs, *Algebra: un'introduzione concreta*, ETS, 1991.
[DM-D] P. Di Martino, (con la revisione di R. Dvornicich), *Algebra*, Edizioni Plus, 2003.
[H] I.N.Herstein, *Algebra*, Editori Riuniti, 1988.
[R] K. H. Rosen, *Discrete mathematics and its applications*, Mc Graw-Hill, 2003.
[AlgGauss] http://marekrychlik.com/cgi-bin/gauss.cgi
[WIMS] http://wims.unice.fr/wims/wims.cgi