Dispense del Corso di Algebra Lineare

Parte 1

Ultimo aggiornamento: 7 maggio 2025

Dr. Michele Ginesi

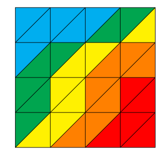
N.B. Queste note sono in corso d'opera. È caldamente **sconsigliata** la stampa, almeno fino a fine corso.

Introduzione e Informazioni sul Corso

La prima parte del corso si focalizzerà su vari aspetti dell'algebra lineare, focalizzandosi in particolare sui concetti di *spazio vettoriale*, *sistema lineare*, e *calcolo matriciale e vettoriale*. Durante il corso, verranno fornite anche delle interpretazioni "geometriche" dei concetti presentati.

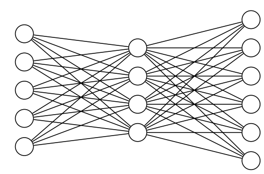
L'algebra lineare è un fondamento necessario per comprendere, e sviluppare, gran parte (se non quasi tutta) la ricerca scientifica moderna. Vediamo alcuni esempi.

La risoluzione numerica di equazioni differenziali (come l'equazione del calore o l'equazione d'onda) viene ottenuta tramite calcolo matriciale. Per fare ciò, si suddivide la superficie in triangolini e si scrive una matrice che rappresenta in qualche modo il problema differenziale da risolvere.



Nel controllo della poszione di un robot, la relazione tra gli *angoli di giunto* (ovvero la configurazione del robot) e la *posa* (i.e. posizione ed orientamento) dell'end-effector può essere espressa tramite l'uso di matrici e vettori.

Nell'ambito del *machine learning*, molte delle operazioni sono definite da operazioni tra matrici e prodotti matrici-vettori. In particolare, nel caso delle reti neurali le operazioni interne ai *layers* della rete sono prodotti matrice-vettore.



L'**esame** consisterà in una prova scritta che copre entrambe le parti del corso e conterrà sia domande di teoria che esercizi. Per chiunque ottenga un voto maggiore o uguale a 27 ci sarà la possibilità di fare una prova orale facoltativa.
Tutti gli argomenti di questa dispensa, esclusi quelli col simbolo "$\diamondsuit$" a lato, sono argomento d'esame.

I **testi** consigliati per questa prima parte di corso sono:

- Algebra Lineare e Geometria, di Francesco Bottacin (su cui gran parte di queste note sono basate).
- Algebra Lineare, di Enrico Gregorio e Luigi Salce;
- Esercizi di algebra lineare e geometria, di Francesco Bottacin.

Durante entrambe le parti del corso, ci sarà un tutor che svolgerà delle ore extra di esercizi.

Diario delle Lezioni

07 Mar 2022 [2h] Definizione di gruppo 1.1, anello 1.3, campo 1.4, e spazio vettoriale 1.5 con esempi. Definizione di combinazione lineare 1.6, (in)dipendenza lineare 1.7. Enunciata e dimostrata proposizione 1.1.
08 Mar 2022 [3h] Enunciata e dimostrata proposizione 1.2. Definizioni di sottospazio vettoriale 1.8. Enunciata proposizione 1.3. Enunciata e deimostrata proposizione 1.4. Definizione di sottospazio generato 1.9. Enunciata proposizione 1.5. Definizioni di somma e somma diretta di spazzi vettoriali 1.10, 1.11. Enunciata proposizione 1.6. Definizioni di insieme di generatori 1.12, base 1.13, e spazio finitamente generato 1.14. Enunciata e dimostrata proposizione dell'estrazione della base 1.8
10 Mar 2022 [3h] Esercizi su spazi e sottospazi vettoriali; insieme di generatori e basi; intersezione, somma, e somma diretta di spazi vettoriali.
14 Mar 2022 [2h] Enunciato e dimostrato il lemma dello scambio 1.1. Enunciati, senza dimostrazione, i corollari 1.1, 1.3, 1.5, e la proposizione 1.9. Enunciato e dimostrato corollario 1.2, completamento alla base 1.4, e formula di Grassmann 1.1. Definita la base canonica 1.16.
15 Mar 2022 [3h] Iniziata la teoria delle funzioni lineari. Date le definizioni 2.1, 2.2, 2.3, 2.4. Enunciati, senza dimostrazione, la proposizione 2.2, corollario 2.1, proposizione 2.6, corollario 2.2. Enunciati e dimostrati i seguenti risultati: proposzioni 2.1, 2.3, 2.4, 2.5, teorema di nullità + rango 2.1, proposizione 2.7
17 Mar 2022 [3h] Esercizi su applicazioni lineari. Verificare se una funzione è lineare, calcolare nucleo e immagine.
21 Mar 2022 [2h] Definizione 2.5 di matrice. Esempio di calcolo di una matrice dato un endomorfismo e due basi. Matrici simili (definizione 2.6) con esempi. Somma tra matrici (definizione 2.7) e prodotto matrice per scalare (definizione 2.8).
22 Mar 2022 [3h] Definiti il prodotto tra matrici 2.9, la matrice identica 2.10, la matrice trasposta 2.12 e simmetrica 2.13 e la matrice inversa 2.15. Enunciate, senza dimostrazione, diverse proprietà del prodotto tra matrici 2.8, delle matrici trasposte 2.9 e delle matrici inverse 2.10. Definito il prodotto scalare 2.14 tra due vettori. Definiti i concetti di nucleo, immagine, rango e nullità di una matrice 2.16 2.17. Teorema nullità + rango per le matrici 2.2.
24 Mar 2022 [3h] Esercizi sul calcolo di una matrice dato un omomorfismo e due basi; calcolo del prodotto tra due matrici; calcolo di nucleo e immagine di una matrice; determinare su un omomorfismo è iniettivo, suriettivo, o biiettivo.
28 Mar 2022 [2h] Enunciati, senza dimostrazione, la proprietà 2.11, il teorema 2.3, ed il corollario 2.3. Definiti i concetti di rango per righe e rango per colonne. Visto il procedimento del cambio di base (sezione 2.2.2).
29 Mar 2022 [3h] Sistemi lineari. Forma matriciale di un sistema, matrice aumentata del sistema, sistema omogeneo. Enunciate e dimostrate le proposzioni 3.1, 3.2, 3.3, ed il teorema di Rouché-Capelli 3.1. Risoluzione di un sistema lineare tramite Eliminazione di Gauss.
31 Mar 2022 [3h] Esercizi su calcolo della matrice del cambio di base, risoluzione di sistemi lineari tramite eliminazione di Gauss.
04 Apr 2022 [2h] Enunciata proposizione 3.4. Uso di EG per calcolare il rango di una matrice. Matrici associate alle operazioni elementari di EG. Uso di EG per calcolare l'inversa di una matrice (non finito).
05 Apr 2022 [3h] Algoritmo per calcolare la matrice inversa tramite Eliminazione di Gauss. Esercizi su calcolo del rango di una matrice e sul calcolo della matrice inversa.

07 Apr 2022 [3h] Esercizi: simulazione della prima prova parziale.

11 Apr 2022 [2h] Determinante. Definizione 4.1 di una funzione determinante. Enunciati, senza dimostrazione, proposzioni 4.1, 4.2, teorema 4.1 sull'univocità del determinante a meno di costanti, e teorema di Laplace 4.3 sull'esistenza del determinante. Enunciato e dimostrato il teorema di Binet 4.2, la proposizione 4.3 sul determinante di una matrice inversa. Definizione "computativa" di determinante con esempi.

12 Apr 2022 [3h]. Uso di EG per il calcolo del determinante. Definizioni di complemento algebrico 4.2 e matrice aggiunta 4.3 con esempi. Enunciati e dimostrati la proposizion 4.4 sulla relazione tra inversa e aggiunta ed il teorema di Cramér 4.4.

Introdotto il problema della diagonalizzazione di endomorfismi. Definiti i concetti di endomorfismo diagonalizzabile 5.1 e matrice diagonalizzabile 5.1.

14 Apr 2022 [3h]. Esercizi sul calcolo del determinante (tramite formula di Laplace, regola di Sarrus, e eliminazione di Gauss). Risoluzione di sistemi lineari tramite regola di Cramér. Presentato un metodo alternativo per calcolare la matrice del cambio di base.

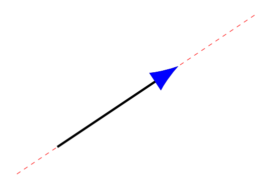
21 Apr 2022 [3h] Definizione 5.3 di autovalore, autovettore e spettro. Enunciata proposizione 5.1. Definizione 5.4. Enunciata e dimostrata proposizone 5.2 Definizione 5.5 di molteplicità algebrica e geometrica di autovalori. Enunciata proposizione 5.3. Enunciati e dimostrati proposizione 5.4 e teorema 5.1 della diagonalizzazione.

26 Apr 2022 [3h] Esempi di diagonalizzazione di matrici. Esercizi in preparazione alla prima prova parziale.

02 Mag 2022 [2h] Esercizi su diagonalizzazione di matrici.

1 Spazi Vettoriali

In questa prima parte, introdurremo il concetto di *spazio vettoriale*. Come prima cosa, introduciamo il concetto 'geometrico' di vettore, traendo ispirazione dal loro uso in fisica. Infatti, molti concetti in fisica vengono misurati usando quantià *scalari*, ovvero grandezze che possono essere espresse con un singolo numero. Ad esempio, la temperatura, la carica elettrica, ed il tempo sono grandezze scalari. D'altro canto, invece, ci sono quantità fisiche che non possono essere descritte da un singolo numero. Ad esempio, quando vogliamo descrivere il movimento di un corpo, non basta conoscerne la velocità, ma è necessario anche conoscere in che *direzione* e *verso*[1] il corpo si sta muovendo (verso destra, sinistra, in avanti, indietro, ...).
Più dettagliatamente, per definire un vettore dobbiamo specificare tre componenti:



- un numero che misura la *lunghezza* del vettore;
- una retta, chiamata *direzione*;
- un *verso* che indica in quale dei due possibili modi il vettore è orientato sulla retta.

Poichè questi oggetti vengono usati per descrivere, ad esempio, degli spostamenti, delle velocità, o le forze esercitate su un corpo; abbiamo bisogno di poter svolgere alcune operazioni. Ad esempio, la moltiplicazione per uno scalare (andiamo al *doppio* della velocità); la somma tra vettori (quando due, o più, forze vengono esercitate); ed altre che vedremo durante il corso.

1.1 Dal Concetto di Gruppo a quello di Spazio Vettoriale

Introdurremo ora una serie di definizioni che ci porteranno, alla fine, di definire in modo formale il concetto di *spazio vettoriale*.

La prima definizione è quella di *gruppo*.

Definizione 1.1 (Gruppo)**.** Un *gruppo* $(G,+)$ è un insieme G non vuoto, $G \neq \emptyset$ munito di un'operazione binaria $+ : G \times G \to G$ che ad ogni coppia di elementi a, b di G associa un elemento di G per cui valgono le seguenti proprietà:

- *proprietà associativa*: dati tre elementi $a, b, c \in G$, vale la seguente identità $(a+b)+c = a+(b+c)$;
- *esistenza dell'elemento neutro*: esiste un elemento $e \in G$ tale per cui $a+e = e+a = a$ per ogni elemento a di G;
- *esistenza dell'elemento inverso*: per ogni elemento a di G, esiste un elemento $a' \in G$, detto *inverso* di a, tale per cui $a+a' = a'+a = e$.

Vediamo un esempio.

Esempio 1.1. L'insieme dei numeri interi $\mathbb{Z} \doteq \{0,1,-1,2,-2,\ldots\}$ con la somma come operazione, $(\mathbb{Z},+)$, è un gruppo. L'elemento neutro è lo zero $e = 0$, mentre l'elemento inverso è l'opposto $a' = -a$.
Al contrario, l'insieme dei numeri naturali $\mathbb{N} \doteq \{0,1,2,\ldots\}$ non forma un gruppo con l'operazione di somma in quanto non esiste l'elemento inverso.

[1] concetti che chiariremo nel dettaglio tra poco

Definizione 1.2 (Gruppo abeliano)**.** Un gruppo $(G,+)$ si dice *abeliano* quando l'operazione $+$ soddisfa la *proprietà commutativa*, ovvero per ogni coppia di elementi a,b di G, $a+b=b+a$.

Una generalizzazione del concetto di gruppo è quello di *anello*.

Definizione 1.3 (Anello)**.** Un anello $(R,+,\cdot)$ è un insieme R non vuoto $R\neq\emptyset$ munito di due operazioni binarie, chiamate *somma* e *moltiplicazione*, $+$ e $\cdot$ tale per cui:

- $(R,+)$ è un gruppo abeliano con elemento neutro $e_+=0$;
- $(R,\cdot)$ è un *monoide*, ovvero:
  - l'operazione $\cdot$ è associativa: $a\cdot(b\cdot c)=(a\cdot b)\cdot c$, $\forall a,b,c\in R$;
  - esiste l'elemento neutro $e_\cdot$ per l'operazione $\cdot$: $a\cdot e_\cdot=e_\cdot\cdot a=a$, $\forall a\in R$.
- la moltiplicazione è *distributiva* rispetto all'addizione, ovvero per ogni terna di elementi di R, $a,b,c,\in R$ valgono le seguenti identità:

$$
\begin{aligned}
&a\cdot(b+c)=(a\cdot b)+(a\cdot c),\\
&(b+c)\cdot a=(b\cdot a)+(c\cdot a.)
\end{aligned}
$$

Un anello si dice *commutativo* quando la moltiplicazione $\cdot$ soddisfa la proprietà commutativa: $a\cdot b=b\cdot a$, $\forall a,b,\in R$.

Osservazione 1.1. Nella definizione di anello, solo la somma $+$ deve soddisfare la commutatività, mentre la moltiplicazione $\cdot$ potrebbe essere non commutativa. Inoltre, l'elemento inverso per la moltiplicazione $\cdot$ potrebbe non esistere.

Vediamo un esempio di anello.

Esempio 1.2. Gli insiemi dei numeri interi $\mathbb{Z}$, razionali $\mathbb{Q}$, e reali $\mathbb{R}$ formano degli anelli (commutativi) rispetto alla classica somma e moltiplicazione tra numeri. L'elemento neutro della somma è lo zero $e_+=0$, mentre l'elemento neutro della moltiplicazione è l'unità $e_\cdot=1$.

La prossima definizione riguarda il concetto di campo.

Definizione 1.4 (Campo)**.** Un anello commutativo $(R,+,\cdot)$ viene detto *campo* quando la moltiplicazione ammette un elemento inverso per ogni elemento di R diverso dall'elemento neutro della somma: $\forall a\in R\setminus\{e_+\},\exists a':a\cdot a'=a'\cdot a=e_\cdot$.

Esempio 1.3. Gli insiemi dei numeri razionali, reali, e complessi formano un campo rispetto alla somma e alla moltiplicazione. Al contrario, l'insieme degli interi **non è** un campo in quanto l'elemento inverso della moltiplicazione non esiste.

In questo corso, ci concentreremo in particolare sul campo dei reali $\mathbb{R}$ e dei complessi $\mathbb{C}$. Inoltre, da qui in poi, semplificheremo la notazione considerando l'elemento neutro della somma sempre uguale a 0, $e_+=0$; e l'elemento neutro della moltiplicazione sempre uguale a 1, $e_\cdot=1$.

Arriviamo finalmente alla definizione di *spazio vettoriale*.

Definizione 1.5 (Spazio vettoriale)**.** Uno *spazio vettoriale* su un campo $\mathbb{K}$ (con elementi neutri 0 ed 1 rispettivamente per la somma e la moltiplicazione) è un insieme non vuoto V, $V \neq \emptyset$, dotato di due operazioni binarie: $(V, +, \cdot)$ chiamate *somma* e *moltiplicazione per uno scalare*

$$
\begin{aligned}
+ &: V \times V \to V, \qquad (\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} + \mathbf{w}, \\
\cdot &: \mathbb{K} \times V \to V, \qquad (\lambda, \mathbf{v}) \mapsto \lambda \cdot \mathbf{v},
\end{aligned}
$$

che soddisfano le seguenti proprietà:

- $(V, +)$ è un gruppo abeliano con elemento neutro $\mathbf{0}$;
- vale la proprietà distributiva della moltiplicazione per scalare rispetto alla somma:

$$\lambda \cdot (\mathbf{v} + \mathbf{w}) = (\lambda \cdot \mathbf{v}) + (\lambda \cdot \mathbf{w}), \quad \forall \lambda \in \mathbb{K}, \quad \forall \mathbf{v}, \mathbf{w} \in V;$$

- vale la pseudo-distributività della moltiplicazione per scalare rispetto alla somma tra scalari:

$$(\lambda_1 + \lambda_2) \cdot \mathbf{v} = (\lambda_1 \cdot \mathbf{v}) + (\lambda_2 \cdot \mathbf{v}), \quad \forall \lambda_1, \lambda_2 \in \mathbb{K}, \quad \forall \mathbf{v} \in V;$$

- vale la compatibilità del prodotto tra scalari e del prodotto per scalari:

$$(\lambda_1 \lambda_2) \cdot \mathbf{v} = \lambda_1 \cdot (\lambda_2 \cdot \mathbf{v}), \quad \forall \lambda_1, \lambda_2 \in \mathbb{K}, \quad \forall \mathbf{v} \in V$$

- l'elemento neutro della moltiplicazione tra scalari è anche l'elemento neutro della moltiplicazione per scalari:

$$1 \cdot \mathbf{v} = \mathbf{v}.$$

Gli elementi del campo $\mathbb{K}$ vengono chiamati *scalari*, mentre gli elementi dell'insieme V vengono chiamati *vettori*.

Nonostante usiamo lo stesso simbolo per indicare le somme tra scalari e tra vettori, così come usiamo lo stesso simbolo per indicare la moltiplicazione scalare-vettore e la moltiplicazione scalara-scalare, queste sono operazioni diverse. Tuttavia non ci preoccupiamo di definire nuovi simboli in quanto le proprietà che soddisfano sono le stesse in entrambi i casi, e quale delle due si stia considerando è chiaro dal contesto.
In modo simile, in quanto segue semplificheremo la notazione in quando spesso ignoreremo il simbolo $\cdot$ per indicare la moltiplicazione (sia *tra* scalari che *per* scalare): $\lambda \cdot \mathbf{v} = \lambda \mathbf{v}$, $\lambda_1 \cdot \lambda_2 = \lambda_1 \lambda_2$.

Osservazione 1.2. Dalle proprietà di uno spazio vettoriale segue che $0 \cdot \mathbf{v} = \mathbf{0}$. Infatti, consideriamo le seguenti identità:

$$\mathbf{v} + 0\mathbf{v} = 1\mathbf{v} + 0\mathbf{v} = (1+0)\mathbf{v} = 1\mathbf{v} = \mathbf{v}.$$

Sommiamo l'opposto di $\mathbf{v}$ ad entrambi i lati:

$$
\begin{aligned}
-\mathbf{v} + \mathbf{v} + 0\mathbf{v} = -\mathbf{v} + \mathbf{v} &\Leftrightarrow \cancel{-\mathbf{v}} + \cancel{\mathbf{v}} + 0\mathbf{v} = \cancel{-\mathbf{v}} + \cancel{\mathbf{v}} \\
&\Leftrightarrow 0\mathbf{v} = \mathbf{0}.
\end{aligned}
$$

In modo simile, si ottiene che l'opposto di un vettore $\mathbf{v}$, $-\mathbf{v}$ si può scrivere come $-1 \cdot \mathbf{v}$.

Vediamo ora qualche esempio di spazio vettoriale.

Esempio 1.4. L'insieme $\mathbb{R}^n$ dotato della somma

$$(v_1, v_2, \ldots, v_n) + (w_1, w_2, \ldots, w_n) = (v_1 + w_1, v_2 + w_2, \ldots, wv_n + w_n)$$

e del prodotto per scalare

$$\lambda \cdot (v_1, v_2, \ldots, v_n) = (\lambda v_1, \lambda v_2, \ldots, \lambda v_n)$$

è uno spazio vettoriale su $\mathbb{R}$.

Esempio 1.5. L'insieme dei polinomi a coefficienti in $\mathbb{R}$ con variabile denotata con x, $\mathbb{R}[x]$, con le classiche operazioni di somma e moltiplicazione per una costante è uno spazio vettoriale su $\mathbb{R}$. Siano infatti $\mathscr{P}_1(x), \mathscr{P}_2(x) \in \mathbb{R}[x]$ due polinomi in x:

$$\begin{aligned}
\mathscr{P}_1(x) &= a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n, \\
\mathscr{P}_2(x) &= b_0 + b_1 x + b_2 x^2 + \ldots + b_n x^n.
\end{aligned}$$

La somma tra polinomi è definita come

$$\mathscr{P}_1(x) + \mathscr{P}_2(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \ldots (a_n + b_n)x^n,$$

mentre la moltiplicazione per uno scalare è definita come

$$\lambda \cdot \mathscr{P}_1(x) = \lambda\, a_0 + \lambda\, a_1 x + \lambda\, a_2 x^2 + \ldots \lambda\, a_n x^n.$$

Chiaramente, $(\mathbb{R}[x], +)$ forma un gruppo abeliano, il cui elemento nullo è lo zero $\mathbf{0} = 0 + 0x + 0x^2 + \ldots + 0x^n = 0$ e l'elemento inverso è il polinomio i cui coefficienti sono l'opposto dei coefficienti

$$-\mathscr{P}_1(x) = -a_0 + (-a_1)x + (-a_2)x^2 + \ldots + (-a_n)x^n.$$

Vale la proprietà distributiva:

$$\begin{aligned}
\lambda(\mathscr{P}_1(x) + \mathscr{P}_2(x)) &= \lambda(a_0 + a_1 x + \ldots + a_n x^n + b_0 + b_1 x + \ldots + b_n x^n) \\
&= (\lambda a_0 + \lambda a_1 x + \ldots + \lambda a_n x^n + \lambda b_0 + \lambda b_1 x + \ldots + \lambda b_n x^n) \\
&= (\lambda a_0 + \lambda a_1 x + \ldots + \lambda a_n x^n) + (\lambda b_0 + \lambda b_1 x + \ldots + \lambda b_n x^n) \\
&= \lambda(a_0 + a_1 x + \ldots + a_n x^n) + \lambda(b_0 + b_1 x + \ldots + b_n x^n) \\
&= \lambda \mathscr{P}_1(x) + \lambda \mathscr{P}_2(x)
\end{aligned}$$

Vale la pseudo-distributività:

$$\begin{aligned}
(\lambda_1 + \lambda_2)\mathscr{P}_1(x) &= (\lambda_1 + \lambda_2) \cdot (a_0 + a_1 x + \ldots a_n x^n) \\
&= (\lambda_1 + \lambda_2) \cdot a_0 + (\lambda_1 + \lambda_2) \cdot a_1 x + \ldots + (\lambda_1 + \lambda_2) \cdot a_n x^n) \\
&= (\lambda_1 \cdot a_0 + \lambda_2 \cdot a_0) + (\lambda_1 \cdot a_1 + \lambda_2 \cdot a_1)x + \ldots + (\lambda_1 \cdot a_n + \lambda_2 \cdot a_n)x^n) \\
&= (\lambda_1 a_0 + \lambda_1 a_1 x + \ldots + \lambda_1 a_n x^n) + (\lambda_2 a_0 + \lambda_2 a_1 x + \ldots + \lambda_2 a_n x^n) \\
&= \lambda_1(a_0 + a_1 x + \ldots + a_n x^n) + \lambda_2(a_0 + a_1 x + \ldots + a_n x^n) \\
&= \lambda_1 \mathscr{P}_1(x) + \lambda_2 \mathscr{P}_1(x)
\end{aligned}$$

Vale la compatibilità del prodotto tra scalari e del prodotto per scalari:

$$
\begin{aligned}
(\lambda_1\lambda_2)\mathscr{P}_1(x) &= (\lambda_1\lambda_2)(a_0 + a_1x + \ldots + a_nx^n) \\
&= (\lambda_1\lambda_2)a_0 + (\lambda_1\lambda_2)a_1x + \ldots + (\lambda_1\lambda_2)a_nx^n) \\
&= \lambda_1(\lambda_2a_0) + \lambda_1(\lambda_2a_1)x + \ldots + \lambda_1(\lambda_2a_n)x^n) \\
&= \lambda_1(\lambda_2a_0 + \lambda_2a_1x + \ldots + \lambda_2a_nx^n) \\
&= \lambda_1(\lambda_2\mathscr{P}_1(x))
\end{aligned}
$$

E l'unità è l'lemento neutro della moltiplicazione per scalare

$$1 \cdot \mathscr{P}_1(x) = \mathscr{P}_1(x).$$

Un concetto di fodamentale importanza nell'algebra lineare è quello di *combinazione lineare di vettori*.

Definizione 1.6 (Combinazione lineare)**.** Sia V uno spazio vettoriale su un campo $\mathbb{K}$. Una *combinazione lineare* di elementi di V è una somma **finita** del tipo

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_n\mathbf{v}_n,$$

con $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ e $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n \in V$.

A seguito di questa definizione, possiamo introdurre quelle di dipendenza ed indipendenza lineare di insiemi di vettori.

Definizione 1.7 (Dipendenza lineare)**.** Sia $A = \{\mathbf{v}_1, \mathbf{v}2, \ldots, \mathbf{v}_n\}$ un sottoinsieme dello spazio vettoriale V , $A \subset V$, sul campo $\mathbb{K}$. I vettori di A si dicono *linearmente dipendenti* se esistono degli scalari $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ non tutti nulli per cui

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots \lambda_n\mathbf{v}_n = \mathbf{0}.$$

Se un insieme di vettori non è linearmente dipendente, allora è detto *linearmente indipendente*.

Esempio 1.6. L'insieme dei polinomi $\{x^2, x, 1, 3x^2 - 2x\}$ è linearmente dipendente in quanto

$$(-3)x^2 + 2x + 0 \cdot 1 + (3x^2 + 2x) = 0 = \mathbf{0}.$$

Proposizione 1.1. *I vettori* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ *sono linearmente dipendenti se e solo se uno di essi può essere espresso come combinazione lineare degli altri:*

$$\mathbf{v}_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} \lambda_j\mathbf{v}_j.$$

Dimostrazione. Poiché i vettori $\mathbf{v}_j$ sono linearmente diepndenti, esiste un insieme di scalari λ_j non tutti nulli tale per cui

$$\lambda_1\mathbf{v}1 + \lambda_2\mathbf{v}_2 + \ldots \lambda_n\mathbf{v}_n = \mathbf{0}.$$

Si scelga allora un indice i per cui $\lambda_i \neq 0$. Dalla condizione di dipendenza segue

$$-\lambda_i \mathbf{v}_i = \lambda_1 \mathbf{v}_1 + \ldots + \lambda_{i-1} \mathbf{v}_{i-1} + \lambda_{i+1} \mathbf{v}_{i+1} + \ldots \lambda_n \mathbf{v}_n,$$

da cui il vettore $\mathbf{v}_i$ si può scrivere come

$$\mathbf{v}_i = \frac{-\lambda_1}{\lambda_i} \mathbf{v}_1 + \ldots + \frac{-\lambda_{i-1}}{\lambda_i} \mathbf{v}_{i-1} + \frac{-\lambda_{i+1}}{\lambda_i} \mathbf{v}_{i+1} + \ldots \frac{-\lambda_n}{\lambda_i} \mathbf{v}_n.$$

Viceversa, supponiamo che esista un vettore $\mathbf{v}_1$ che si possa scrivere come combinazione lineare degli altri:

$$\mathbf{v}_i = \lambda_1 \mathbf{v}_1 + \ldots + \lambda_{i-1} \mathbf{v}_{i-1} + \lambda_{i+1} \mathbf{v}_{i+1} + \ldots \lambda_n \mathbf{v}_n.$$

Allora si ha

$$\lambda_1 \mathbf{v}_1 + \ldots + \lambda_{i-1} \mathbf{v}_{i-1} + (-\mathbf{v}_i) + \lambda_{i+1} \mathbf{v}_{i+1} + \ldots \lambda_n \mathbf{v}_n = \mathbf{0}.$$

E quindi l'insieme è linearmente dipendente poiché il coefficiente λ_i è non nullo. □

Proposizione 1.2. *Siano $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in V$ dei vettori linearmente indipendenti. Se $\mathbf{v} \in V$ può essere scritto come combinazione lineare dei vettori $\mathbf{v}_i$, $i = 1, 2, \ldots, n$, allora i coefficienti sono determinati in modo unico.*

Dimostrazione. Supponiamo esistano due modi per scrivere il vettore $\mathbf{v}$ come combinazione lineare dei vettori $\mathbf{v}_i$:

$$\begin{aligned}
\mathbf{v} &= \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_n \mathbf{v}_n, \\
\mathbf{v} &= \mu_1 \mathbf{v}_1 + \mu_2 \mathbf{v}_2 + \ldots + \mu_n \mathbf{v}_n.
\end{aligned}$$

Allora abbiamo

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_n \mathbf{v}_n = \mu_1 \mathbf{v}_1 + \mu_2 \mathbf{v}_2 + \ldots + \mu_n \mathbf{v}_n,$$

che implica

$$(\lambda_1 - \mu_1) \mathbf{v}_1 + (\lambda_2 - \mu_2) \mathbf{v}_2 + \ldots + (\lambda_n - \mu_n) \mathbf{v}_n = \mathbf{0}.$$

Poiché i vettori sono linearmente indipendenti, questo implica $\lambda_0 - \mu_i = 0$ per ogni indice i. Quindi $\lambda_i = \mu_i$. □

1.2 Sottospazi Vettoriali

In questa parte, introdurremo le definizioni di sottospazi, insieme a quelle di somma e intersezione di spazi vettoriali.

Da ora in poi, salvo diversamente identicato, denoteremo con V un generico spazio vettoriale su un campo $\mathbb{K}$.

Definizione 1.8 (Sottospazio vettoriale)**.** Un sottoinsieme non vuoto $W \subset V$ è detto un *sottospazio vettoriale* se W è uno spazio vettoriale su $\mathbb{K}$ rispetto alle operazioni di somma e moltiplicazione per scalare definite su V.

Esempio 1.7. Ogni spazio vettoriale V è sottospazio di se stesso.

Esempio 1.8. L'insieme contenente l'elemento nullo di uno spazio vettoriale, $\{\mathbf{0}\} \subset V$ è un sottospazio, detto spazio nullo.

Esempio 1.9. Sia $V = \mathbb{R}^2$ lo spazio di tutti i vettori con due componenti. Allora, il sottoinsieme $W = \{[a, 2a] : a \in \mathbb{R}\}$ di tutti i vettori che giaciono sulla retta $y = 2x$ è un sottospazio vettoriale.

Esempio 1.10. Sia $\mathbb{R}^3$ lo spazio vettoriale di tutti i vettori a tre componenti con coefficienti reali, $\mathbb{R}^3 = \{[a, b, c] : a, b, c \in \mathbb{R}\}$. Allora l'insieme di tutti i vettori in cui la terza componente è nulla, $\{[a, b, 0] : a, b \in \mathbb{R}^3\}$, forma un sottospazio. Geometricamente, questo significa che un piano nello spazio tridimensionale è a sua volta uno spazio vettoriale.

Questa definizione ci dice quindi che se un sottoinsieme W di V soddisfa tutte le proprietà che definiscono uno spazio vettoriale, relativamente alle operazioni che rendono V uno spazio vettoriale, allora è un sottospazio.
Fortunatamente, per dimostrare che W sia un sottospazio, non è necessario dimostrare tutte le proprietà (essere un gruppo rispetto alla somma, proprietà associativa, ...), ma possiamo limitarci a verificare che la combinazione lineare di due elementi qualsiasi di W sia anch'essa elemento di W. Abbia,mo infatti il seguente risultato.

Proposizione 1.3. *Un sottoinsieme non vuoto W di uno spazio vettoriale V su un campo $\mathbb{K}$ è un sottospazio vettoriale di V se e solo se*

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 \in W,$$

per ogni coppia di scalari $\lambda_1, \lambda_2 \in \mathbb{K}$ e ogni coppia di vettori $\mathbf{v}_1, \mathbf{v}_2 \in W$.

Esempio 1.11. Usando gli insiemi V e W dell'esempio 1.9, abbiamo che dobbiamo dimostrare che l'elemento $\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2$ sia in W per ogni scelta dei coefficienti λ_1, λ_2 e dei vettori $\mathbf{w}_1 = [a, 2a], \mathbf{w}_2 = [b, 2b]$ con $a, b \in \mathbb{R}$. Consideriamo quindi la combinazione lineare

$$\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 = [\lambda_1 a + \lambda_2 b, 2\lambda_1 a + 2\lambda_2 b] = [\ell, 2\ell], \quad \ell \in \mathbb{R}.$$

Quindi $\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 \in W$.

Si ha che l'intersezione di sottospazi vettoriali è a sua volta un sottospazio vettoriale.

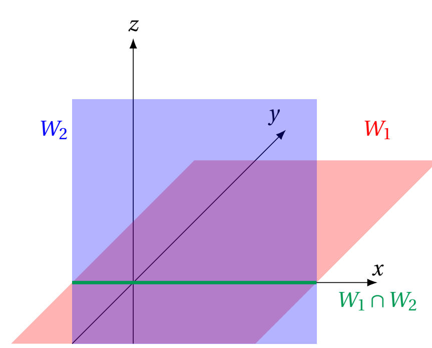
Proposizione 1.4. *Sia $\{W_i\}_{i=1,2,\ldots,N}$ un insieme di sottospazi vettoriali di V. Allora anche la loro intersezione $W = \bigcap_{i=1,2,\ldots,N} W_i$ è uno spazio vettoriale.*

Dimostrazione. Siano $\lambda_1, \lambda_2 \in \mathbb{K}$ due scalari, e $\mathbf{v}_1, \mathbf{v}_2 \in W$ due vettori. Allora, $\mathbf{v}_1, \mathbf{v}_2 \in W_i$ per ogni $i = 1, 2, \ldots, N$. Poiché W_i è uno spazio vettoriale per ogni i, allora anche $\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 \in W_i$ per ogni indice i. Quindi, $\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 \in W$. Di conseguenza, grazie alla Proposizione 1.3 anche W è un sottospazio di V. □

Esempio 1.12. Sia $V = \mathbb{R}^3$, e siano $W_1 = \{[x, y, 0] : x, y \in \mathbb{R}\}$ e $W_2 = \{[x, 0, z] : x, z \in \mathbb{R}\}$. L'intersezione è l'insieme

$$W_1 \cap W_2 = \{[x, 0, 0] : x \in \mathbb{R}\},$$

che forma uno spazio vettoriale. Geometricamente, W_1 è il piano 'orizzontale' nello spazio in cui $z = 0$, mentre W_2 è il piano 'verticale' in cui $y = 0$. L'intersezione è l'asse orizzontale x, che forma uno spazio vettoriale.



Al contrario, l'unione di due spazi vettoriali potrebbe non essere uno spazio vettoriale. Si consideri, ad esempio, gli spazi $V = \{[a,0] : a \in \mathbb{R}\}$ e $W = \{[0,b] : b \in \mathbb{R}\}$. La loro unione è l'insieme dei vettori con due componenti per cui almeno una componente è nulla. Tuttavia, il vettore $[1,1]$ può essere scritto come combinazione lineare degli elementi di $V \cup W$ ma non appartiene a $V \cup W$. Di conseguenza, l'unione $V \cup W$ non è uno spazio vettoriale.
Tuttavia c'è un modo per "generare" uno spazio vettoriale partendo da un insieme che non è uno spazio a sua volta.

Definizione 1.9 (Sottospazio generato)**.** Sia S un sottoinsieme di V. Il *sottospazio vettoriale generato dall'insieme* S, indicato con $\mathscr{L}(S)$, è il più piccolo sottospazio vettoriale di V che contenga S.

Esempio 1.13. Il sottospazio generato dall'insieme vuoto è lo spazio nullo $\mathscr{L}(\emptyset) = \{\mathbf{0}\}$.

Esempio 1.14. Il sottospazio generato dall'unione degli insiemi $U = \{[a,0] : a \in \mathbb{R}\}$ e $W = \{[0,b] : b \in \mathbb{R}\}$ è lo spazio $V = \{[a,b] : a,b,\in \mathbb{R}\}$.

Si può construire lo spazio vettoriale generato da un insieme tramite combinazioni lineari di elementi dell'insieme stesso.

Proposizione 1.5. *Il sottospazio vettoriale $\mathscr{L}(S)$ generato da un insieme non vuoto S è l'insieme di tutte le combinazioni lineari finite di elementi di S:*

$$\mathscr{L}(S) = \left\{ \sum_{i=1}^{n} \lambda_i \mathbf{v}_i : n \in \mathbb{N}, \lambda_i \in \mathbb{K}, \mathbf{v}_i \in S \right\}.$$

◊ *Dimostrazione.* Definiamo

$$\Lambda(S) = \left\{ \sum_{i=1}^{n} \lambda_i \mathbf{v}_i : n \in \mathbb{N}, \lambda_i \in \mathbb{K}, \mathbf{v}_i \in S \right\}.$$

Chiaramente $S \subset \Lambda(S)$. Inoltre, ogni sottospazio vettoriale di V contentente S contiene anche tutte le combinazioni lineari finite di elementi di S, e quindi contiene $\Lambda(S)$. Di conseguenza, l'insieme $\Lambda(S)$ è contenuto nell'intersezione di tutti i sottospazi vettoriali di V che contengono S, quindi $\Lambda(S) \subset \mathscr{L}(S)$.

D'altra parte, l'insieme $\Lambda(S)$ è chiaramente un sottospazio di V. E, poichè $\mathscr{L}(S)$ è il più piccolo sottospazio vettoriale contenente S, si ha $\mathscr{L}(S) \subset \Lambda(S)$.
Segue quindi che $\mathscr{L}(S) = \Lambda(S)$. □

Dato un insieme finito di vettori $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, il sottospazio vettoriale generato $\mathscr{L}(S)$ viene denotato anche come $\langle \mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n \rangle$.

Definizione 1.10 (Somma di spazi vettoriali)**.** Siano U e W due sottospazi vettoriali di V. La loro *somma* $U + W$ è definito come il sottospazio generato dall'unione di U e W: $U + W = \mathscr{L}(U \cup W)$.
Tale definizione si generalizza facilmente alla somma di più di due sottospazi.

Proposizione 1.6. *La somma di due sottospazi vettoriali U e W di V è lo spazio vettoriale i cui elementi sono somme degli elementi di U e W:*

$$U + W = \{\mathbf{u} + \mathbf{w} : \mathbf{u} \in U, \mathbf{w} \in W\}.$$

◊ *Dimostrazione.* È immediato verificare che l'insieme $\{\mathbf{u}+\mathbf{w} : \mathbf{u} \in U, \mathbf{w} \in W\}$ è un sottospazio vettoriale di V che contiene sia U che W e quindi contiene anche la loro unione $U \cup W$. Poiché, per definizione, $U+W$ è il più piccolo sottospazio vettoriale di V contenente $U \cup W$ si ha l'inclusione

$$U+W \subset \{\mathbf{u}+\mathbf{w} : \mathbf{u} \in U, \mathbf{w} \in W\}.$$

D'altra parte, ogni vettore nella forma $\mathbf{u}+\mathbf{w}$ appartiene sicuramente a $U+W$, quindi anche l'inclusione opposta vale:

$$U+W \supset \{\mathbf{u}+\mathbf{w} : \mathbf{u} \in U, \mathbf{w} \in W\}.$$

Di conseguenza, i due insiemi coincidono:

$$U+W = \{\mathbf{u}+\mathbf{w} : \mathbf{u} \in U, \mathbf{w} \in W\}.$$

□

Un caso particolare di somma di spazi vettoriali si ha nel caso della somma diretta.

Definizione 1.11 (Somma diretta di spazi vettoriali)**.** La somma di due spazi vettoriali U e W si dice *diretta*, e si indica con $U \dotplus W$, quando $U \cap W = \{\mathbf{0}\}$.

Anche in questo caso, la somma diretta si generalizza a più di due sottospazi. Infatti, siano W_i, $i = 1,2,\ldots,N$ dei sottospazi vettoriali. La somma degli spazi W_i si dice diretta, e si indica con

$$\bigoplus_{i=1,2,\ldots,N} W_i = W_1 \oplus W_2 \oplus \ldots \oplus W_N,$$

quando $W_i \cap W_j = \{\mathbf{0}\}$ se $i \neq j$.

Esempio 1.15. Sia $V = \mathbb{R}^3$ e siano $U = \{[a,b,0] : a,b \in \mathbb{R}\}$ e $W = \{[0,b,c] : b,c \in \mathbb{R}\}$. La somma di U e W coincide con V: $U+W = V$. Infatti, ogni vettore $\mathbf{v} = [\alpha,\beta,\gamma] \in V$ può essere scritto come $[\alpha,\beta,0]+[0,0,\gamma]$. In questo modo, il primo vettore appartiene ad U, mentre il secondo appartiene a W. Tuttavia, la somma **non** è direta in quanto $U \cap W = \{[0,b,0] : b \in \mathbb{R}\}$.

◊ **Proposizione 1.7.** *Ogni vettore* $\mathbf{v} \in U \oplus W$ *si può scrivene in modo unico nella forma* $\mathbf{v} = \mathbf{u}+\mathbf{w}$ *con* $\mathbf{u} \in U$ *e* $\mathbf{w} \in W$.

◊ *Dimostrazione.* Grazie alla Proposizione 1.6, $\mathbf{v}$ si può scrivere come somma di elementi di U e W. Resta da dimostrare che tale scrittura sia unica. Consideriamo due modi diversi per scrivere il vettore $\mathbf{v}$:

$$\mathbf{v} = \mathbf{u}_1 + \mathbf{w}_1 = \mathbf{u}_2 + \mathbf{w}_2, \quad \mathbf{u}_1, \mathbf{u}_2 \in U \, \mathbf{w}_1, \mathbf{w}_2 \in W, \quad \mathbf{u}_1 \neq \mathbf{u}_2, \mathbf{w}_1 \neq \mathbf{w}_2.$$

Allora si ha che $\mathbf{w}_1 - \mathbf{w}_2 = \mathbf{u}_2 - \mathbf{u}_1$. L'elemento a sinistra appartiene a W, mentre l'elemento a destra appartiene a U (infatti, sono combinazioni lineari di elementi di W ed U rispettivamente). Quindi $\mathbf{w}_1 - \mathbf{w}_2, \mathbf{u}_2 - \mathbf{u}_1 \in W \cap U$. Ma, poiché la somma tra U e W è diretta, abbiamo che l'intersezione è il sottospazio nullo: $W \cap U = \{\mathbf{0}\}$. Di conseguenza

$$\mathbf{w}_1 - \mathbf{w}_2 = \mathbf{0} \quad \Rightarrow \mathbf{w}_1 = \mathbf{w}_2,$$
$$\mathbf{u}_1 - \mathbf{u}_2 = \mathbf{0} \quad \Rightarrow \mathbf{u}_1 = \mathbf{u}_2.$$

Segue che $\mathbf{w}_1 = \mathbf{w}_2$ e $\mathbf{u}_1 = \mathbf{u}_2$. □

1.3 Insiemi di Generatori, Basi, e Spazi Finitamente Generati

Ora che abbiamo chiaro il concetto di *spazio generato*, possiamo passare al prossimo argomento, ovvero i concetti di *insieme generatori* e di *base*.

Definizione 1.12 (Insieme di generatori)**.** Un sottoinsieme $S \subset V$ di uno spazio vettoriale V è detto un *insieme di generatori* di V se $\mathscr{L}(S) = V$. In questo caso, si dice che S *genera* V.

Notiamo subito che *ogni* spazio vettoriale ammette un insieme di generatori. Banalmente, l'intero spazio V genera V.

Si ha, dalla Proposizione 1.5 che se S è iun insieme di generatori di V, allora ogni vettore di V, $\mathbf{v} \in V$ può essere scritto (in genere, non univocamente) come una combinazione lineare finita di elementi di S.

Definizione 1.13 (Base di uno spazio)**.** Un insieme S di generatori di uno spazio vettoriale V è chiamato *base* se gli elementi di S sono linearmente indipendenti.

Poichè una base è un insieme di generatori, abbiamo che ogni elemento di V può essere scritto come combinazione lineare degli elementi della base. Inoltre, grazie alla Proposizione 1.2, abbiamo che ogni elemento di V può essere scritto *in modo unico* come combinazione lineare degli elementi della base.

Esempio 1.16. L'insieme

$$S = \left\{\mathbf{v}_1 = \begin{bmatrix}1 & 0\end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix}0 & 1\end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix}1 & 1\end{bmatrix}\right\}$$

è un insieme di generatori per $V = \mathbb{R}^2$ ma non è una base in quanto $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2$. Tuttavia, qualsiasi sottoinsieme di S composto da due elementi è una base per V. Infatti, un generico elemento $\mathbf{v} = [a, b] \in V$ può essere scritto come

$$\begin{aligned}
\mathbf{v} &= a\mathbf{v}_1 + b\mathbf{v}_2,\\
\mathbf{v} &= (a-b)\mathbf{v}_1 + b\mathbf{v}_3,\\
\mathbf{v} &= (b-a)\mathbf{v}_2 + a\mathbf{v}_3.
\end{aligned}$$

Definizione 1.14 (Spazio finitamente generato)**.** Uno spazio vettoriale V viene detto *finitamente generato* se esiste un insieme finito di generatori di V.

Da questa definizione seguono due importanti proprietà: ogni spazio vettoriale finitamente generato ammette una base, e da ogni insieme di generatori di uno spazio si può estrarre una base. Iniziamo dimostrando il secondo risultato.

Proposizione 1.8 (Estrazione della base)**.** *Sia* $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ *un insieme di generatori di* V. *Allora* S *contiene dei vettori* $\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \ldots, \mathbf{v}_{i_m}$ *per un* $m \leq n$ *che formano una base di* V

Dimostrazione. Se i vettori di S sono linearmente indipendenti, allora S è una base di V. Altrimenti, uno dei vettori di S può essere scritto come combinazione lineare degli altri. Senza perdita di generalità, assumiamo che il vettore $\mathbf{v}_n$ si possa scrivere come combinazione lineare degli altri $n-1$ vettori di S:

$$\mathbf{v}_n = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_{n-1}\mathbf{v}_{n-1}.$$

Di conseguenza, l'insieme

$$S' = S \setminus \{\mathbf{v}_n\} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\}$$

è un insieme di generatori, concludendo la dimostrazione. Altrimenti, in modo analogo, si scrive uno degli elementi di S' come combinazione lineare degli altri e si considera l'insieme di generatori S'' definito come S' privato dell'elemento linearmente dipendente dagli altri.
Proseguendo in questo modo si arriverà prima o poi ad avere un insieme di generatori linearmente indipendenti, ovvero una base di V. □

Vediamo ora una serie di risultati che metteno in relazione i concetti di *insieme linearmente indipendenti*, *insieme di generatori*, e *base*.

Lemma 1.1 (Steinitz / dello scambio)**.** *Sia V uno spazio vettoriale finitamente generato. Sia inoltre $A = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ un insieme di generatori di V, e $B = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m\}$ un insieme di vettori linearmente indipendenti. Allora $m \leq n$.*

Dimostrazione. Poiché i vettori di A generano lo spazio V. Allora il vettore $\mathbf{b}_1$ può essere scritto come combinazione lineare di elementi di A:

$$\mathbf{b}_1 = \lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2 + \ldots, \lambda_n \mathbf{a}_n.$$

Poichè gli elementi di B sono linearmente indipendenti, allora $\mathbf{b}_1 \neq \mathbf{0}$, e gli scalari λ_i non possono essere tutti nulli. Quindi, esiste un coefficiente $\lambda_i \neq 0$. Di conseguenza, il vettore $\mathbf{a}_i$ si può scrivere come combinazione lineare degli altri elementi di A e $\mathbf{b}_1$:

$$\mathbf{a}_i = \frac{1}{\lambda_i}\mathbf{b}_1 + \frac{-\lambda_1}{\lambda_i}\mathbf{a}_1 + \ldots + \frac{-\lambda_{i-1}}{\lambda_i}\mathbf{a}_{i-1} + \frac{-\lambda_{i+1}}{\lambda_i}\mathbf{a}_{i+1} + \ldots + \frac{-\lambda_n}{\lambda_i}\mathbf{a}_n.$$

Senza perdita di generalità, poniamo $i = n$. Abbiamo quindi che l'insieme $\{\mathbf{b}_1, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-1}\}$ è un insieme di generatori di V. In modo analogo, il vettore $\mathbf{b}_2$ si può scrivere come combinazione lineare di $\{\mathbf{b}_1, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-1}\}$ e, senza perdita di generalità, anche l'insieme $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-2}\}$ è un insieme di generatori di V. Continuando in questo modo, si dimostra che tutti gli insiemi $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_h, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-h}\}$ sono insiemi di generatori di V.
Ora, assumiamo per assurdo $n < m$. Allora ponendo $h = n$, avremmo che i vettori $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ generano V e, di conseguenza, che il vettore $\mathbf{b}_{n+1}$ si può scrivere come combinazione lineare di $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$, contraddicendo il fatto che i vettori di B formano una base. Quindi, $m \leq n$. □

Dal lemma di Steinitz seguono immediatamente i seguenti corollari

Corollario 1.1. *Sia V uno spazio vettoriale finitamente generato, e sia $\mathscr{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base di V. Allora*

1. *per ogni insieme di vettori linearmente indipendenti $A = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ si ha $m \leq n$;*
2. *per ogni insieme $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ di generatori di V, si ha $p \geq n$.*

♢ *Dimostrazione.* Per dimostrare il punto 1. si ha che, poiché l'insieme $\mathscr{B}$ è una base, allora è anche un insieme di generatori. Poiché per il lemma dello scambio 1.1 un insieme di generatori ha un numero di elementi maggiore o uguali ad un insieme linearmente indiependente, segue $n \geq m$.

Per dimostrare il punto 2., invece, notiamo che la base $\mathscr{B}$ contiene, per definizione di base, elementi linearmente indipendenti. Per il lemma dello scambio 1.1, un insieme linearmente indipendente ha un numero di elementi minore o uguale ad un insieme di generatori, segue $n \leq p$. □

Corollario 1.2. *Due basi qualunque di uno stesso spazio vettoriale finitamente generato V hanno lo stesso numero di elementi.*

♢ *Dimostrazione.* Si considerino due basi $\mathscr{B}_1 = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ e $\mathscr{B}_2 = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ di V. Poichè gli elementi di $\mathscr{B}_1$ sono linearmente indipendenti e quelli di $\mathscr{B}_2$ generano lo spazio, $n \leq m$. In modo identico, poichè gli elementi di $\mathscr{B}_2$ sono linearmente indipendenti e quelli di $\mathscr{B}_1$ generano lo spazio, $m \leq n$. Di conseguenza, si ha $m = n$. □

Questo ultimo risultato è estremamente importante in quanto ci dice che la cardinalità (i.e. il numero di elementi) di una base di uno spazio vettoriale V non dipende dalla base scelta. Questo ci permette di dare la prossima definizione.

Definizione 1.15 (Dimensione di uno spazio)**.** La *dimensione* di uno spazio vettoriale finitamente generato V, indicata con $\dim(V)$, è il numero di elementi di una base di V.

Un importante esempio di base dello spazio vettoriale $\mathbb{K}^n$ è la *base canonica.*

Definizione 1.16 (Base canonica)**.** Consideriamo lo spazio vettoriale dei vettori ad n componenti su un campo $\mathbb{K}$ (e.g. $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$), $V = \mathbb{K}^n = \{[v_1, v_2, \ldots, v_n] : v_i \in \mathbb{K}\}$. La *base canonica* è la base $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ in cui $\mathbf{e}_i$ ha come unica componente non nulla la $i-$sima, che è pari ad uno

$$\mathbf{e}_i = [0, 0, \ldots, 0, \underset{i}{1}, 0, \ldots, 0, 0].$$

Per questa base, un generico elemento $\mathbf{v} = [\lambda_1, \lambda_2, \ldots, \lambda_n]$ può essere scritto come

$$\mathbf{v} = \lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \ldots + \lambda_n \mathbf{e}_n = \sum_{i=1}^{n} \lambda_i \mathbf{e}_i.$$

Dal lemma dello scambio 1.1 seguono i seguenti due risultati.

Corollario 1.3. *Sia V uno spazio vttoriale finitamente generato. Allora ogni sottospazio W di V è finitamente generato e* $\dim W \leq \dim V$.

♢ *Dimostrazione.* Poniamo $n = \dim V$. Consideriamo un insieme di vettori linearmente indipendenti $A = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ di W. Chiaramente, essi sono linearmente indipendenti anche per V. Di conseguenza, dal lemma di Steinitz 1.1 segue $m \leq n$.

Assumiamo ora che l'insieme A non sia un insieme di generatori di W. Quindi esiste un vettore $\mathbf{w}_{m+1} \in W$ che non può essere scritto come combinazione lineare di elementi di A. Da ciò segue che l'insieme

$$A' \doteq A \cup \{\mathbf{w}_{m+1}\} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m, \mathbf{w}_{m+1}\}$$

è linearmente indipendente. Se i vettori di A' non generano W, possiamo aggiungere un nuovo elemento $\mathbf{w}_{m+2}$ che non può essere scritto come combinazione lineare di A'. Continuando in questo modo, arriveremo ad un insieme

$$\tilde{A} \doteq \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m, \mathbf{w}_{m+1}, \dots, \mathbf{w}_{m+d}\},$$

per un certo d, che genera W.

Poiché W è un sottospazio di V, l'insieme $\tilde{A}$ non può avere più di n elementi. Quindi, l'insieme $\tilde{A}$ è una base di W con un numero non maggiore di n elementi. Quindi, $\dim W \leq \dim V$. □

Corollario 1.4 (Completamento alla base)**.** *Sia V uno spazio vettoriale finitamente generato. Allora ogni insieme di vettori linearmente indipendenti $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ può essere completato ad una base di V. Questo significa che esistono dei vettori $\mathbf{v}_{m+1}, \mathbf{v}_{m+2}, \dots, \mathbf{v}_n$ tali per cui l'insieme $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m, \mathbf{v}_{m+1}, \mathbf{v}_{m+2}, \dots, \mathbf{v}_n\}$ è una base di V.*

Dimostrazione. Se l'insieme $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ genera V, la dimostrazione è conclusa. Altrimenti, esiste un vettore $\mathbf{v}_{m+1}$ che non è una combinazione lineare dei vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$. Quindi, l'insieme $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m, \mathbf{v}_{m+1}\}$ è un insieme linearmente indipendente. Se questi vettori generano V, la dimostrazione è conclusa, altrimenti si può considerare un vettore $\mathbf{v}_{m+2}$ ch non puoò essere scritto come combinazione lineare di $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m, \mathbf{v}_{m+1}$. Quindi, l'insieme $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m, \mathbf{v}_{m+1}, \mathbf{v}_{m+2}\}$ è linearmente indipendente.
Proseguendo in questo modo, si ha che esiste un valore n per cui l'insieme linearmente indiependente

$$\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m, \mathbf{v}_{m+1}, \mathbf{v}_{m+2}, \dots, \mathbf{v}_{n-1}, \mathbf{v}_n\}$$

genera lo spazio V.
Se così non fosse, significherebbe che non esiste un insieme finito che genera V, contro l'ipotesi per cui V sia finitamente generato. □

Se la dimensione dello spazio vettoriale V è nota, verificare che un insieme di vettori formi una base è un esercizio molto semplice, grazie al prossimo risultato.

Proposizione 1.9. *Sia V uno spazio vettoriale di dimensione n, $\dim V = n$, e siano $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in V$.*

1. *se i vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ sono linearmente indipendenti, allora sono anche un insieme di generatori di V, e quindi formano una base di V;*

2. *se i vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ generano V, allora sono anche linearmente indipendenti, e quindi formano una base dello spazio.*

◇ *Dimostrazione.* Iniziamo dal punto 1. Dal corollario 1.4 segue che l'insieme linearmente indiependente $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ è sottoinsieme di una base

$$\mathscr{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n, \mathbf{v}_{n+1}, \dots, \mathbf{v}_{n+d}\}$$

di V. Ma poiché la dimensione di V è n, abbaimo che $d =$ e quindi l'insieme $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ forma una base di V.

Per il punto 2. supponiamo che l'insieme $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ generi V. Se, per assurdo, fosse un insieme linearmente dipendente, uno dei vettori sarebbe combinazione lineare degli altri. Quindi esiste un indice i per cui l'insieme $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n\}$ genera V. Ma allora avremmo che $\dim V \leq n-1$, in contraddizione con l'ipotesi per cui $\dim V = n$. Quindi, i vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ sono linearmente indipendenti. □

Corollario 1.5. *Sia V uno spazio vettoriale finitamente generato e $W \subsetneq V$ uno sottospazio proprio. Allora* $\dim W < \dim V$.

◇ *Dimostrazione.* Dal corollario 1.3 sappiamo che $\dim W \leq \dim V$. Supponiamo, per assurdo, che $\dim W = \dim V = n$. Consideriamo quindi una base $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n\}$ di W. Questo insieme è un insieme di vettori linearmente indipendente per V e, poiché n è la dimensione di V, per la proposizione 1.9 essi sono una base di V. Di conseguenze, $V = W$. Ma questa è una contraddizione in quanto W, per ipotesi, è sottospazio proprio di V. □

Il prossimo risultato ci permette di mettere in relazione le dimensioni di due sottospazi U, W di V con la loro somma $U \oplus W$ e la loro intersezione $U \cap W$.

Teorema 1.1 (Formula di Grassmann)**.** *Siano U e W due sottospazi vettoriali di uno spazio vettoriale finitamente generato V. Si ha*

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W).$$

Dimostrazione. Chiamiamo

$$u = \dim U, \qquad w = \dim W, \qquad \ell = \dim(U \cap W).$$

Consideriamo ora una base $\mathscr{B}_{U \cap W} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_\ell\}$ di $U \cap W$. Questo insieme è un insieme linearmente indipendente sia per U che per W. Grazie al corollario del completamento alla base 1.4, possiamo quindi completare questo insieme ad una base $\mathscr{B}_U = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_\ell, \mathbf{u}_1, \dots, \mathbf{u}_{u-\ell}\}$ di U ed a una base $\mathscr{B}_W = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_\ell, \mathbf{w}_1, \dots, \mathbf{w}_{w-\ell}\}$ di W.
Poiché ogni vettore di $U+W$ può essere scritto come una somma di un vettore di U e di un vettore di W, esiste una combinazione lineare dell'insieme $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_\ell, \mathbf{u}_1, \dots, \mathbf{u}_{u-\ell}, \mathbf{w}_1, \dots, \mathbf{w}_{w-\ell}\}$ che genera qualsiasi elemento di $U+W$.

Dimostriamo ora che questi vettori sono linearmente indipendenti. Supponiamo quindi che esista una combinazione lineare a coefficienti non tutti nulli che risulti nel vettore nullo:

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots \alpha_\ell \mathbf{v}_\ell + \beta_1 \mathbf{u}_1 + \dots + \beta_{u-\ell} \mathbf{u}_{u-\ell} + \gamma_1 \mathbf{w}_1 + \dots + \gamma_{w-\ell} \mathbf{w}_{w-\ell} = \mathbf{0}.$$

Dobbiamo dimostrare che tutti i coefficienti $\alpha_i, \beta_j, \gamma_k$ sono nulli. Questa identià può essere riscritta come segue

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots \alpha_\ell \mathbf{v}_\ell + \beta_1 \mathbf{u}_1 + \dots + \beta_{u-\ell} \mathbf{u}_{u-\ell} = -(\gamma_1 \mathbf{w}_1 + \dots + \gamma_{w-\ell} \mathbf{w}_{w-\ell}).$$

Chiamiamo ora $\mathbf{x}$ il vettore rappresentato dalle queste espressioni. Il vettore $\mathbf{x}$ appartiene sia ad U che a W, in quanto può essere scritto sia come combinazione lineare di elementi di U (i termini a sinistra sono elementi di $\mathscr{B}_U$) e di W (i termini a destra sono elementi di $\mathscr{B}_W$): $\mathbf{x} \in U \cap W$. Poiché l'insieme $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_\ell\}$ forma una base di $U \cap W$, il vettore $\mathbf{x}$ può essere scritto, in modo unico, come

$$\mathbf{x} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \dots + \lambda_\ell \mathbf{v}_\ell.$$

Inoltre, il vettore **x** si può scrivere anche nei seguenti due modi:

$$\underbrace{\lambda_1\mathbf{v}_1+\ldots\lambda_\ell\mathbf{v}_\ell}_{\mathbf{x}} = \underbrace{\alpha_1\mathbf{v}_1+\alpha_2\mathbf{v}_2+\ldots\alpha_\ell\mathbf{v}_\ell+\beta_1\mathbf{u}_1+\ldots+\beta_{u-\ell}\mathbf{u}_{u-\ell}}_{\in U}, \tag{1a}$$

$$\underbrace{\lambda_1\mathbf{v}_1+\ldots\lambda_\ell\mathbf{v}_\ell}_{\mathbf{x}} = \underbrace{-(\gamma_1\mathbf{w}_1+\ldots+\gamma_{w-\ell}\mathbf{w}_{w-\ell})}_{\in W}. \tag{1b}$$

La seconda equazione, (1b), si può riscrivere come

$$\lambda_1\mathbf{v}_1+\ldots\lambda_\ell\mathbf{v}_\ell+\gamma_1\mathbf{w}_1+\ldots+\gamma_{w-\ell}\mathbf{w}_{w-l}=\mathbf{0}.$$

Questo significa che una combinazione lineaare di elementi in $\mathscr{B}_W$ risulta nel vettore nullo. Ma poiché $\mathscr{B}_W$ forma una base, allora tutti i coefficienti λ_i, γ_j sono nulli.

Da questo segue anche che il termine a sinistra di (1a) coincide col vettore nullo, $\mathbf{x}=\mathbf{0}$ e, poiché i termini a sinistra in (1b) costituiscono una combinazione lineare di elementi di $\mathscr{B}_U$, che è una base, allora anche i coefficienti α_i e β_j sono tutti nulli.

Abbiamo quindi dimostrato che

$$\lambda_1=\ldots=\lambda_\ell=0, \qquad \alpha_1=\ldots=\alpha_\ell=0, \qquad \beta_1=\ldots=\beta_{u-\ell}=0, \qquad \gamma_1=\ldots=\gamma_{w-\ell}=0.$$

Di conseguenza, l'insieme

$$\{\mathbf{v}_1,\ldots,\mathbf{v}_\ell,\mathbf{u}_1,\ldots,\mathbf{u}_{u-\ell},\mathbf{w}_1,\ldots,\mathbf{w}_{w-\ell}\}$$

forma una base di $U+W$. Da questo segue l'identità

$$\dim(U+W)=\ell+(u-\ell)+(w-\ell)=u+w-\ell,$$

che, rifacendosi alle definizioni di u, w, ed ℓ restituisce la tesi. □

Osservazione 1.3. Nel caso di somma diretta, si ha

$$\dim(U\oplus W)=\dim U+\dim W,$$

poiché $U\cap W=\{\mathbf{0}\}$ e $\dim(\{\mathbf{0}\})=0$.

2 Applicazioni Lineari e Matrici

In questa parte del corso, ci concentreremo sulla teoria delle funzioni lineari tra spazi vettoriali, introducendo il concetto di matrice. Successivamente, ci concentreremo sulla teoria delle matrici, studiando, ad esempio, le operazioni tra matrici. Questi risultati ci permetteranno di affrontare la teoria dei sistemi di equazioni lineari.

2.1 Applicazioni Lineari

Come prima cosa dobbiamo definire cosa sia una funzione lineare.

Definizione 2.1 (Funzione lineare / omomorfismo)**.** Siano V e W due spazi vettoriali su uno stesso campo $\mathbb{K}$. Una funzione $f : V \to W$ è detta *lineare* (rispetto al campo $\mathbb{K}$) se soddisfano le due seguenti proprietà:

$$f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2), \qquad \forall \mathbf{v}_1, \mathbf{v}_2 \in V \tag{2a}$$

$$f(\lambda \cdot \mathbf{v}) = \lambda \cdot f(\mathbf{v}), \qquad \forall \mathbf{v} \in V, \quad \forall \lambda \in \mathbb{K}. \tag{2b}$$

Una funzione lineare tra spazi vettoriali è detta anche *omomorfismo* di spazi vettoriali.

Una funzione che soddisfa (2a) è detta *additiva*.

Il prossimo risultato ci fornisce una definizione equivalente alla definizione 2.1 di omomorfismo. Infatti, le condizioni (2) possono essere compresse in una sola condizione come segue.

Proposizione 2.1. *Una funzione $f : V \to W$ è lineare se e solo se soddisfa la seguente proprietà:*

$$f(\alpha \mathbf{v}_1 + \beta \mathbf{v}_2) = \alpha f(\mathbf{v}_1) + \beta f(\mathbf{v}_2). \tag{3}$$

Dimostrazione. Come prima cosa, mostriamo che l'equazione (3) implica le equazioni (2). Infatti, l'identità (2a) è un caso particolare di (3) con $\alpha = \beta = 1$, mentre (2b) è un caso particolare di (3) nel caso in cui $\beta \mathbf{v}_2 = \mathbf{0}$.

Per dimostrare l'implicazione opposta, calcoliamo il valore $f(\alpha \mathbf{v}_1 + \beta \mathbf{v}_2)$ usando le equazioni in (2). Utilizzando (2a) abbiamo

$$f(\alpha \mathbf{v}_1 + \beta \mathbf{v}_2) = f(\alpha \mathbf{v}_1) + f(\beta \mathbf{v}_2).$$

Ora, utilizzando invece (2b), otteniamo che il termine a destra di quest'ultima equazione si riscrive come

$$f(\alpha \mathbf{v}_1) + f(\beta \mathbf{v}_2) = \alpha f(\mathbf{v}_1) + \beta f(\mathbf{v}_2).$$

□

Una funzione lineare da uno spazio vettoriale in se stesso, $f : V \to V$ è detto *endomorfismo*.
Un caso particolare di omomorfismo è l'isomorfismo.

Definizione 2.2 (Isomorfismo)**.** Sia $f : V \to W$ un omomorfismo tra spazi vettoriali. La funzione f è detta *isomorfismo* se esiste un omomorfismo $g : W \to V$ tale per cui

$$f \circ g : W \to W = \mathrm{id}_W, \qquad g \circ f : V \to V = \mathrm{id}_V,$$

dove id denota la mappa identità.
Quindi, essere un isomorfismo significa essere un omomorfismo invertibile, in cui la mappa inversa è lineare.

Due spazi vettoriali V e W si dicono *isomorfi*, e si indica $V \cong W$, quando esiste un isomorfismo tra V e W.
Un isomorfismo da uno spazio vettoriale in se stesso è chiamato *automorfismo*.

Esempio 2.1. Sia

$$V = \mathbb{R}^3 = \{[v_1, v_2, v_3] : v_1, v_2, v_3 \in \mathbb{R}\},$$

e sia la funzione f definita come segue:

$$f : V \to V, \qquad f : \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \mapsto \begin{bmatrix} v_3 \\ 2\,v_2 \\ 3\,v_1 \end{bmatrix}.$$

La funzione f è lineare, infatti

$$\begin{aligned} f(\mathbf{v}+\mathbf{w}) = f\left(\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}\right) = f\left(\begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{bmatrix}\right) = \begin{bmatrix} v_3 + w_3 \\ 2(v_2 + w_2) \\ 3(v_1 + w_1) \end{bmatrix} = \begin{bmatrix} v_3 + w_3 \\ 2\,v_2 + 2\,w_2 \\ 3\,v_1 + 3\,w_1 \end{bmatrix} \\ = \begin{bmatrix} v_3 \\ 2\,v_2 \\ 3\,v_1 \end{bmatrix} + \begin{bmatrix} w_3 \\ 2\,w_2 \\ 3\,w_1 \end{bmatrix} = f\left(\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}\right) + f\left(\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}\right) = f(\mathbf{v}) + f(\mathbf{w}) \end{aligned}$$

e

$$f(\lambda\,\mathbf{v}) = f\left(\lambda \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}\right) = f\left(\begin{bmatrix} \lambda\,v_1 \\ \lambda\,v_2 \\ \lambda\,v_3 \end{bmatrix}\right) = \begin{bmatrix} \lambda\,v_3 \\ 2(\lambda\,v_2) \\ 3(\lambda\,v_1) \end{bmatrix} = \begin{bmatrix} \lambda\,v_3 \\ \lambda\,2\,v_2 \\ \lambda\,3\,v_1 \end{bmatrix} = \lambda \begin{bmatrix} v_3 \\ 2\,v_2 \\ 3\,v_1 \end{bmatrix} = \lambda f\left(\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}\right) = \lambda f(\mathbf{v})$$

È inoltre un isomorfismo con funzione inversa

$$g : \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \mapsto \begin{bmatrix} \frac{1}{3} w_3 \\ \frac{1}{2} w_2 \\ w_1 \end{bmatrix}.$$

Esempio 2.2. Sia $V = \mathbb{R}^3$ ed f definita come

$$f : \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \mapsto \begin{bmatrix} v_1 \\ v_2 \\ 0 \end{bmatrix}.$$

Allora f è un omomorfismo ma non un isomorfismo.

Dalla definizione di isomorfismo 2.2 segue che un isomorfismo è una funzione biiettiva, in quanto deve ammettere una funzione inversa. È vero anche il viceversa.

Proposizione 2.2. *Sia $f : V \to W$ un omomorfismo di spazi vettoriali. Se la funzione f è biiettiva allora è anche un isomorfismo.*

$\diamondsuit$ *Dimostrazione.* Poiché f è biiettiva, allora è anche invertibile. Bisogna quindi dimostrare che anche la mappa $f^{-1} : W \to V$ è lineare.

Si considerino, a tale scopo, due elementi $\mathbf{w}_1, \mathbf{w}_2 \in W$, e siano

$$\mathbf{v}_1 = f^{-1}(\mathbf{w}_1), \qquad \mathbf{v}_2 = f^{-1}(\mathbf{w}_2).$$

Dall'additività di f (equazione (2a)) si ha

$$\mathbf{w}_1 + \mathbf{w}_2 = f(\mathbf{v}_1) + f(\mathbf{v}_2) = f(\mathbf{v}_1 + \mathbf{v}_2),$$

la quale si può scrivere in modo equivalente

$$f^{-1}(\mathbf{w}_1 + \mathbf{w}_2) = f^{-1}\left(f(\mathbf{v}_1 + \mathbf{v}_2)\right) = \mathbf{v}_1 + \mathbf{v}_2 = f^{-1}(\mathbf{w}_1) + f^{-1}(\mathbf{w}_2).$$

Quindi, anche f^{-1} soddisfa (2a).

Si consideri ora anche uno scalare $\lambda \in \mathbb{K}$. Da (2b) si ha

$$\lambda \mathbf{w}_1 = \lambda f(\mathbf{v}_1) = f(\lambda \mathbf{v}_1),$$

che, si può riscrivere come segue:

$$f^{-1}(\lambda \mathbf{w}_1) = f^{-1}\left(f(\lambda \mathbf{v}_1)\right) = \lambda \mathbf{v}_1 = \lambda f^{-1}(\mathbf{w}_1).$$

Quindi, f^{-1} soddisfa (2b). $\square$

L'importanza del concetto di isomorfismo è data dal fatto che consente di identificare, in qualche senso, spazi vettoriali diversi, a patto che siano isomorfi. Possiamo quindi usare il concetto di isomorfismo per "classificare" spazi vettoriali.

Proposizione 2.3. *Sia V uno spazio vettoriale di dimensione n,* $\dim V = n$*, sul campo $\mathbb{K}$. Allora V è isomorfo a $\mathbb{R}^n$.*

Dimostrazione. Consideriamo una base $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ di V e definiamo una funzione $f : V \to \mathbb{R}^n$ che, ad ogni vettore $\mathbf{v} \in V$ associa l'**unica** n−upla $(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n$ tale per cui

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_n \mathbf{v}_n.$$

Tale funzione è lineare. Infatti, siano $\mathbf{w} = \sum_{i=1}^n \mu_i \mathbf{v}_i$ e $\mathbf{u} = \sum_{i=1}^n \nu_i \mathbf{v}_i$. Dimostriamo prima che f è additiva, ovvero soddisfa (2a):

$$\begin{aligned}
f(\mathbf{w} + \mathbf{u}) &= f\left(\sum_{i=1}^n \mu_i \mathbf{v}_i + \mathbf{u} + \sum_{i=1}^n \nu_i \mathbf{v}_i\right) \\
&= f\left(\sum_{i=1}^n (\mu_i + \nu_i)\mathbf{v}_i\right) \\
&= (\mu_1 + \nu_1, \mu_2 + \nu_2, \ldots, \mu_n + \nu_n) \\
&= (\mu_1, \mu_2, \ldots, \mu_n) + (\nu_1, \nu_2, \ldots, \nu_n) \\
&= f(\mathbf{w}) + f(\mathbf{u}).
\end{aligned}$$

Dimostriamo ora che f soddisfa (2b). Consideriamo uno scalare α e calcoliamo

$$\begin{aligned} f(\alpha\mathbf{v}) &= f(\alpha(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_n\mathbf{v}_n)) \\ &= f(\alpha\lambda_1\mathbf{v}_1 + \alpha\lambda_2\mathbf{v}_2 + \ldots + \alpha\lambda_n\mathbf{v}_n) \\ &= (\alpha\lambda_1, \alpha\lambda_2, \ldots, \alpha\lambda_n) \\ &= \alpha(\lambda_1, \lambda_2, \ldots, \lambda_n) \\ &= \alpha f(\mathbf{v}). \end{aligned}$$

Inoltre, poiché l'insieme $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ forma una base di V, la mappa f è anche biiettiva. Questo perché per ogni elemento $\mathbf{v}$ esiste una e una sola $n-$upla di coordinate rispetto ad una base.

Quindi, f è un isomorfismo. □

Osservazione 2.1. Nella dimostrazione abbiamo usato il fatto che un vettore $\mathbf{v} \in V$ può essere scritto in modo univoco come combinazione lineare degli elementi di una base $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$:

$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_n\mathbf{v}_n.$$

Gli scalari λ_i vengono chiamati *componenti* del vettore rispetto alla base.

Dalla proposizione 2.3, segue il prossimo corollario.

Corollario 2.1. *Due spazi vettoriali V e W su uno stesso campo $\mathbb{K}$ e di dimensione finita sono isomorfi se e solo se hanno la stessa dimensione:* $\dim V = \dim W$.

Dimostrazione. Entrambi gli spazi sono isomorfi a $\mathbb{K}^n$:

$$V \cong \mathbb{K}^n \cong W.$$

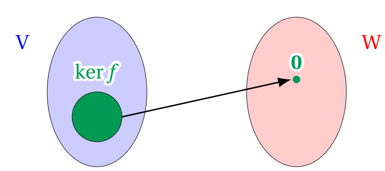
Quindi, $V \cong W$. □

2.1.1 Nucleo e Immagine

Introduciamo ora due particolari sottospazi vettoriali 'associati' ad una mappa lineare.

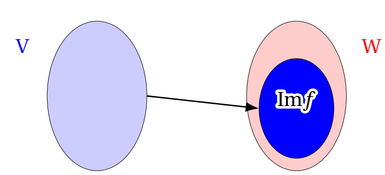
Definizione 2.3 (Nucleo)**.** Sia $f : V \to W$ una funzione lineare tra due spazi vettoriali. Il *nucleo* di f è l'insieme dei vettori che vengono mappati, tramite f, nel vettore nullo

$$\ker(f) \doteq \{\mathbf{v} \in V : f(\mathbf{v}) = \mathbf{0}\}. \tag{4}$$



Definizione 2.4 (Immagine)**.** Sia $f : V \to W$ una funzione lineare tra due spazi vettoriali. L' *immagine* di f è l'insieme dei vettori che possono essere ottenuti applicando la mappa f ad un vettore $\mathbf{v} \in V$

$$\mathrm{Im}(f) \doteq \{\mathbf{w} \in W : \exists\mathbf{v}, f(\mathbf{v}) = \mathbf{w}\}. \tag{5}$$



Esempio 2.3. Prendendo l'esempio 2.2, il nucleo è l'insieme

$$\ker f = \left\{ \begin{bmatrix} 0 \\ 0 \\ v_3 \end{bmatrix} : v_3 \in \mathbb{K} \right\},$$

mentre l'immagine è l'isnieme

$$\operatorname{Im} f = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ 0 \end{bmatrix} : v_1, v_2 \in \mathbb{K} \right\}.$$

Si ha che gli insiemi nucleo e immagine sono a loro volta spazi vettoriali.

Proposizione 2.4. *Sia $f : V \to W$ una mappa lineare. Allora il nucleo* $\ker f$ *è un sottospazio vettoriale di V, mentre l'immagine* $\operatorname{Im} f$ *è un sottospazio di W.*

Dimostrazione. Siano $\mathbf{v}_1, \mathbf{v}_2 \in \ker f$, e consideriamo una combinazione lineare $\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2$. Poiché f è una funzione lineare, abbiamo

$$f(\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2) = \lambda_1 f(\mathbf{v}_1) + \lambda_2 f(\mathbf{v}_2) = \mathbf{0}.$$

Quindi, anche la combinazione lineare è nel nucleo: $\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 \in \ker f$. Quindi, il nucleo è un sottospazio di V.

Siano ora $\mathbf{w}_1, \mathbf{w}_2$ due elementi dell'immagine di f, $\mathbf{w}_1, \mathbf{w}_2 \in \operatorname{Im} f$, di modo che $\mathbf{w}_1 = f(\mathbf{v}_1)$ e $\mathbf{w}_2 = f(\mathbf{v}_2)$. Poiché f è lineare, segue che

$$f(\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2) = \lambda_1 f(\mathbf{v}_1) + \lambda_2 f(\mathbf{v}_2) = \lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2,$$

da cui segue che l'immagine è un sottospazio vettoriale di W. □

I prossimi risultati servono a mettere in relazione i concetti di nucleo e di funzione lineare. In particolare, il prossimo risultato dimostra che l'iniettività di una funzione lineare coincide con l'avere lo spazio nullo come nucleo.

Proposizione 2.5. *Sia $f : V \to W$ una funzione lineare. Allora f è iniettiva se e solo se il nucleo è lo spazio nullo:* $\ker f = \{\mathbf{0}\}$.

Dimostrazione. Iniziamo supponendo che f sia iniettiva e sia $\mathbf{v}$ un generico elemento del nucleo di f: $\mathbf{v} \in \ker f$. Per una mappa lineare, $f(\mathbf{0}) = \mathbf{0}$, e, poiché f è iniettiva, esiste un solo elemento che viene mappato nel vettore nullo. Quindi $\mathbf{v} = \mathbf{0}$.

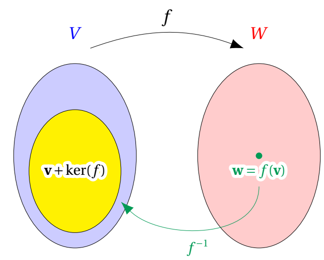
Viceversa, supponiamo che il nucleo si f sia lo spazio nullo, $\ker f = \{\mathbf{0}\}$. Consideriamo ora due vettori $\mathbf{v}, \mathbf{w} \in V$ tali per cui $f(\mathbf{v}) = f(\mathbf{w})$. Poiché f è lineare, abbiamo $f(\mathbf{v} - \mathbf{w}) = f(\mathbf{v}) - f(\mathbf{w}) = \mathbf{0}$. Quindi, il vettore $\mathbf{v} - \mathbf{w}$ è nel nucleo di f, $\mathbf{v} - \mathbf{w} \in \ker f$. Poiché per ipotesi il nucleo contiene solo il vettore nullo, $\ker f = \{\mathbf{0}\}$, si ha $\mathbf{v} - \mathbf{w} = \mathbf{0}$, ovvero $\mathbf{v} = \mathbf{w}$. Quindi, f è iniettiva. □

Il prossimo risultato mette in relazione la *contro-immagine* di una funzione f col nucleo della funzione stessa.

Proposizione 2.6. *Sia $f : V \to W$ una funzione lineare e sia $\mathbf{w} \in W$. Se $\mathbf{w} \in \operatorname{Im} f$, allora si ha*

$$f^{-1}(\mathbf{w}) = \mathbf{v} + \ker f = \left\{\mathbf{v} + \mathbf{u} : \mathbf{u} \in \ker f\right\},$$

ove $\mathbf{v}$ è un qualsiasi vettore tale che $f(\mathbf{v}) = \mathbf{w}$. Se invece $\mathbf{w} \notin \operatorname{Im} f$, si ha $f^{-1}(\mathbf{w}) = \emptyset$.



Dimostrazione. L'ultima affermazione è ovvia.

Per dimostrare il risultato principale, sia $\mathbf{v} \in V$ tale per cui $\mathbf{w} = f(\mathbf{v})$. Per ogni $\mathbf{u} \in \ker f$, si ha $f(\mathbf{v}+\mathbf{u}) = f(\mathbf{v}) + f(\mathbf{u}) = \mathbf{w} + \mathbf{0} = \mathbf{w}$. Quindi, $\mathbf{v} + \ker f \subset f^{-1}(\mathbf{w})$.

Viceversa, per ogni $\mathbf{v}' \in f^{-1}(\mathbf{w})$ definiamo $\mathbf{u} = \mathbf{v}' - \mathbf{v}$. Per la linearità di f abbiamo che $f(\mathbf{u}) = f(\mathbf{v}') - f(\mathbf{v}) = \mathbf{w} - \mathbf{w} = \mathbf{0}$. Quindi $\mathbf{u} \in \ker f$, da cui segue $f^{-1}(\mathbf{w}) \subset \mathbf{v} + \ker f$. Quindi, poiché valgono entrambe le inclusioni, abbiamo che i due insiemi coincidono. □

Abbiamo inoltre il prossimo risultato, che sarà di fondamentale importanza quando introdurremo il concetto di matrici e di sistema lineari.

Teorema 2.1 (Nullità + rango)**.** *Sia $f : V \to W$ una funzione lineare. Se V ha dimensione finita, allora*

$$\dim V = \dim(\ker f) + \dim(\operatorname{Im} f).$$

Dimostrazione. Denotiamo con n ed m rispettivamente le dimensioni di V e del nucleo di f:

$$n = \dim V, m = \dim(\ker f).$$

Dobbiamo dimostrare che $\dim(\operatorname{Im} f) = n - m$. Consideriamo una base

$$\mathscr{B}_k = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$$

del nucleo di f ed un completamento di questa ad una base

$$\mathscr{B}_V = \{\underbrace{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m}_{\in \mathscr{B}_k}, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_n\}$$

dell'intero spazio V. Le immagini di questi vettori formano un insieme di generatori di $\operatorname{Im} f$. Poiché le immagini dei vettori della base sono il vettore nullo, $f(\mathbf{v}_1) = f(\mathbf{v}_2) = \ldots = f(\mathbf{v}_m) = \mathbf{0}$, si ha che i vettori $\mathbf{w}_1 = f(\mathbf{v}_{m+1}), \mathbf{w}_2 = f(\mathbf{v}_{m+2}), \ldots, \mathbf{w}_{n-m} = f(\mathbf{v}_n)$ formano un insieme di generatori per $\operatorname{Im} f$.

Dimostriamo ora che questi vettori sono linearmente indipendenti. Per fare ciò, consideriamo una combinazione lineare

$$\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \ldots + \lambda_{n-m} \mathbf{w}_{n-m} = \mathbf{0}, \tag{6}$$

ed osserviamo che, poiché f è lineare, abbiamo

$$\begin{aligned}
\mathbf{0} &= \lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \ldots + \lambda_{n-m} \mathbf{w}_{n-m} \\
&= \lambda_1 f(\mathbf{v}_{m+1}) + \lambda_2 f(\mathbf{v}_{m+2}) + \ldots + \lambda_{n-m} f(\mathbf{v}_n) \\
&= f(\lambda_1 \mathbf{v}_{m+1} + \lambda_2 \mathbf{v}_{m+2} + \ldots + \lambda_{n-m} \mathbf{v}_n)
\end{aligned}$$

da cui l'argomento appartiene al nucleo:

$$\lambda_1\mathbf{v}_{m+1} + \lambda_2\mathbf{v}_{m+2} + \ldots + \lambda_{n-m}\mathbf{v}_n \in \ker f.$$

Di conseguenza, questa quantità può essere scritta come combinazione lineare degli elementi della base del nucle di f $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$:

$$\lambda_1\mathbf{v}_{m+1} + \lambda_2\mathbf{v}_{m+2} + \ldots + \lambda_{n-m}\mathbf{v}_n = \mu_1\mathbf{v}_1 + \mu_2\mathbf{v}_2 + \ldots + \mu_m\mathbf{v}_m.$$

L'ultima identità può essere scritta come

$$\mu_1\mathbf{v}_1 + \mu_2\mathbf{v}_2 + \ldots + \mu_m\mathbf{v}_m - \lambda_1\mathbf{v}_{m+1} - \lambda_2\mathbf{v}_{m+2} - \ldots - \lambda_{n-m}\mathbf{v}_n = \mathbf{0}.$$

Poiché per ipotesi i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ formano una base di V, si ha che tutti i coefficienti sono nulli: $\lambda_1 = \lambda_2 = \ldots = \lambda_{n-m} = 0$, $\mu_1 = \mu_2 = \ldots = \mu_m = 0$. Di conseguenza, abbiamo che da (6) segue il fatto che tutti i coefficienti sono nulli. Quindi, i vettori $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{n-m}$ sono linearmente indipendenti, e, essendo anche un insieme di generatori, formano una base dell'immagine di f. Quindi, la dimensione di $\mathrm{Im} f$ è pari ad $n - m$. □

Le seguenti sono due denominazioni che useremo spesso quando parleremo di matrici. La dimensione dell'immagine di una funzione lineare è chiamata *rango*,

$$\mathrm{rk} f = \dim(\mathrm{Im} f),$$

mentre la dimensione del nucleo è chiamata *nullità*,

$$\mathrm{null} f = \dim(\ker f).$$

Esempio 2.4. Consideriamo la funzione dell'esempio 2.2. Poiché $V = \mathbb{R}^3$, abbiamo che $\dim V = 3$. Lo spazio nullo di f è composto da tutti i vettori della forma

$$\ker f = \left\{ \begin{bmatrix} 0 \\ 0 \\ \alpha \end{bmatrix} : \alpha \in \mathbb{R} \right\}.$$

Una base di questo spazio è la seguente:

$$\mathscr{B} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Quindi $\dim(\ker f) = 1$, da cui segue che $\dim(\mathrm{Im} f) = 3 - 1 = 2$.

2.1.2 Applicazioni Lineari e Basi

In questa parte, studieremo le proprietà di una funzione lineare $f : V \to W$ in relazione alla scelta delle basi sia del dominio V che del codominio W.

Il primo risultato che dimostreremo spiega come una funzione lineare sia totalmente determinata dalle immagini dei vettori di una base del dominio.

Proposizione 2.7. *Siano V e W due spazi vettoriali e sia $\{\mathbf{v}_i\}_{i\in I}$ una base di V. Allora*

1. *un omomorfismo $f : V \to W$ è univocamente determinato dalle immagini dei vettori $\mathbf{v}_i$, $i \in I$;*

2. *dato un insieme di vettori $\{\mathbf{w}_i\}_{i\in I}$ in W, esiste un unico omomorfismo $f : V \to W$ tale per cui $f(\mathbf{v}_i) = \mathbf{w}_1$ per ogni $i \in I$.*

Dimostrazione. Per dimostrare il punto 1, sia $f : V \to W$ una funzione lineare e supponiamo di conoscere $f(\mathbf{v}_i)$ per ogni $i \in I$. Poiché i vettori $\mathbf{v}_i$ formano una base di V, ogni vettore di V può essere scritto come combinazione lineare di vettori $\mathbf{v}_i$: $\mathbf{v} = \sum_{i\in I} \lambda_i \mathbf{v}_i$. Poiché la funzione f è lineare, abbiamo

$$f(\mathbf{v}) = \sum_{i\in I} \lambda_i f(\mathbf{v}_i) = \lambda_1 f(\mathbf{v}_1) + \lambda_2 f(\mathbf{v}_2) + \dots.$$

Quindi, conoscere i valori $f(\mathbf{v}_i)$ è sufficiente per conoscere il valore della funzione applicata a qualsiasi vettore $\mathbf{v} \in V$. Inoltre, poiché i vettori $\mathbf{v}_i$ formano una base, esiste un unico modo per definire il generico vettore $\mathbf{v}$, e quindi un unico modo per definire $f(\mathbf{v})$, dimostrando la tesi.

Dimostriamo ora il punto 2. Per ogni $i \in I$, scegliamo arbitrariamente un vettore $\mathbf{w}_i$ con $i \in I$ e definiamo una funzione f che soddisfi $f(\mathbf{v}_i) = \mathbf{w}_1$. Ora, possiamo estendere questa funzione per linearità. Questo significa che, per ogni vettore $\mathbf{v} = \sum_{i\in I} \lambda_i \mathbf{v}_i$, definiamo

$$f(\mathbf{v}) = \sum_{i\in I} \lambda_i f(\mathbf{v}_i).$$

La funzione f è ben definita e lineare. Inoltre, grazie al punto 1, f è unica. □

Osservazione 2.2. In questo risultato **non** è necessario porre che V sia finitamente generato. Quindi, l'insieme degli indici I può essere un insieme a cardinalità infinita.

Corollario 2.2. *Siano V e W due spazi vettoriali su $\mathbb{K}$, sia $\{\mathbf{v}_i\}_{i\in I}$ una base di V, e sia $f : V \to W$ una mappa lineare. Allora valgono i seguenti:*

1. *f è iniettiva se e solo se l'insieme $\{f(\mathbf{v}_i)\}_{i\in I}$ è libero (ovvero ogni suo sottoinsieme finito è linearmente indipendente);*

2. *f è suriettiva se e solo se l'insieme $\{f(\mathbf{v}_i)\}_{i\in I}$ genera W;*

3. *f è biiettiva se e solo se l'insieme $\{f(\mathbf{v}_i)\}_{i\in I}$ forma una base di W.*

Si osservi che, nel caso di spazi vettoriali finitamente generati, il punto 1. afferma che un omomorfismo f è iniettivo se e solo se l'insieme $\{f(\mathbf{v}_i)\}_{i\in I}$ è linearmente indipendente.

2.2 Matrici

Siano V e W due spazi vettoriali finitamente generati su uno stesso campo $\mathbb{K}$. Siano n ed m le loro dimensioni, $n = \dim V$, $m = \dim W$, e siano gli insiemi $\mathscr{B}_V = \{\mathbf{v}_j\}_{j=1,2,\dots,n}$ e $\mathscr{B}_W = \{\mathbf{w}_i\}_{i=1,2,\dots,m}$ basi per V e W rispettivamente.
Grazie alla Proposizione 2.7, abbiamo che una funzione lineare $f : V \to W$ è determinata in modo unico dalla conoscenza delle immagini di $\mathbf{v}_j$ tramite f, $f(\mathbf{v}_j)$. Inoltre, poiché l'insieme $\{\mathbf{w}_i\}_{i=1,2,\dots,m}$ è una base di W, abbiamo che esistono dei coefficienti $a_{ij} \in \mathbb{K}$ per $i = 1,2,\dots,m$ e $j = 1,2,\dots,n$ per cui

$$f(\mathbf{v}_j) = \sum_{i=1}^{m} a_{ij} \mathbf{w}_i.$$

Di conseguenza, date una base del dominio ed una del codominio, una funzione lineare è determinata univocamente dagli elementi $a_{ij} \in \mathbb{K}$. Questi elementi costituiscono ciò che chiamiamo matrice.

Definizione 2.5 (Matrice)**.** Una *matrice* con m righe ed n colonne a coefficienti in un campo $\mathbb{K}$ è un insieme ordinato di $m \cdot n$ elementi in $\mathbb{K}$, solitamente scritto sotto forma della tabella

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

o, più semplicemente, con la scrittura

$$\mathbf{A} = \left[a_{ij}\right]_{i=1,2,\dots,m}^{j=1,2,\dots,n}.$$

Il valore i è detto *indice di riga*, mentre il valore j è detto *indice di colonna*.

L'insieme delle matrici da m righe ed n colonne su un campo $\mathbb{K}$ è denotato con $\mathscr{M}_{m,n}(\mathbb{K})$ o, equivalentemente, con $\mathbb{K}^{m\times n}$. Quindi, una matrice $\mathbf{A}$ di m righe ed n colonne a coefficienti in $\mathbb{K}$ viene indicata come $\mathbf{A} \in \mathbb{K}^{m\times n}$ o $\mathbf{A} \in \mathscr{M}_{m,n}(\mathbb{K})$. La quantità $m \times n$ è detta anche *dimensione* della matrice.
Una matrice si dice *quadrata* quando il numero di righe è uguale al numero di colonne.

Abbiamo quindi che ad ogni applicazione lineare $f : V \to W$ tra due spazi vettoriali V e W possiamo associare una matrice. Occorre però prestare attenzione che la scelta di tale matrice dipende dalla scelta delle basi di V e W.

Esempio 2.5. Siano $V = \mathbb{R}^2$ e $W = \mathbb{R}^3$, e definiamo le basi $\mathscr{B}_V$ e $\mathscr{B}_W$ come segue

$$\mathscr{B}_V = \left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}, \qquad \mathscr{B}_W = \left\{ \mathbf{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{w}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\};$$

e la mappa lineare $f : V \to W$ nel seguente modo:

$$f\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} a+b \\ a-b \\ a \end{bmatrix}.$$

Calcoliamo la matrice della mappa f rispetto alle basi $\mathscr{B}_V$ e $\mathscr{B}_W$. Per fare ciò, iniziamo scrivendo $f(\mathbf{v}_1)$ ed $f(\mathbf{v}_2)$:

$$f(\mathbf{v}_1) = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \qquad f(\mathbf{v}_2) = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$$

Abbiamo le seguenti identità:

$$f(\mathbf{v}_1) = 3\mathbf{w}_1 - \mathbf{w}_2 - 2\mathbf{w}_3,$$
$$f(\mathbf{v}_2) = 4\mathbf{w}_1 - \mathbf{w}_2 - 2\mathbf{w}_3.$$

Da cui segue che la matrice $\mathbf{A}$ associata ad f è

$$\mathbf{A} = \begin{bmatrix} 3 & 4 \\ -1 & -1 \\ -2 & -2 \end{bmatrix}.$$

Esempio 2.6. Consideriamo ora due spazi vettoriali $V = \mathbb{R}^3$ e $W = \mathbb{R}^2$, utilizzando, per entrambi gli spazi, la relativa base canonica. Consideriamo ora il seguente omomorfismo:

$$f : V \to W, \qquad \begin{bmatrix} a \\ b \\ c \end{bmatrix} \mapsto \begin{bmatrix} a+b+c \\ a+2b+3c \end{bmatrix}.$$

Applichiamo ora la mappa agli elementi della base canonica di $V = \mathbb{R}^3$:

$$f\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad f\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \qquad f\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Questi tre valori possono essere scritti nel seguente modo in termine della base canonica di W:

$$f\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad f\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad f\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Quindi, la matrice $\mathbf{A}$ associata alla funzione f è data da

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}.$$

Ricordiamo che dalla proposizione 2.7 segue che, dato un vettore $\mathbf{v} \in V$, questo può essere scritto in modo univoco come combinazione lineare degli elementi di una base $\mathscr{B}_V \doteq \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ di V: $\mathbf{v} = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i$. I coefficienti $\lambda_i \in \mathbb{K}$ sono chiamati *coordinate* di $\mathbf{v}$ rispetto alla base $\mathscr{B}_V$.

Osservazione 2.3. Dalla definizione della matrice $\mathbf{A}$ associata ad una mappa lineare $f : V \to W$ rispetto alle basi $\mathscr{B}_V = \{\mathbf{v}_i\}_{i=1}^{n}$ e $\mathscr{B}_W = \{\mathbf{w}_j\}_{j=1}^{m}$ di V e W rispettivamente, segue che le coordinate del vettore $f(\mathbf{v}_j)$ rispetto a $\mathscr{B}_W$ costituiscono la j−sima colonna di $\mathbf{A}$

Esempio 2.7. Usando i dati dell' Esempio 2.5, abbiamo che la prima colonna $(3, -1, -2)$ sono le coordinate di $f(\mathbf{v}_1)$ rispetto a $\mathscr{B}_W$, mentre la seconda colonna, $(4, -1, -2)$, sono le ccordinate di $f(\mathbf{v}_2)$ rispetto a $\mathscr{B}_W$.

Prima di studiare le operazioni tra matrici, diamo un'ultima definizione.

Definizione 2.6 (Matrici simili)**.** Due matrici quadrate si dicono *simili* se esse rappresentano lo stesso endomorfismo (ovvero un omomorfismo da uno spazio in se' stesso) $f : V \to V$ rispetto a basi diverse.

Esempio 2.8. Sia $V = \mathbb{R}^3$ e sia $f : \mathbb{R}^3 \to \mathbb{R}^3$ definito come segue:

$$f : \begin{bmatrix} x, y, z \end{bmatrix} \mapsto \begin{bmatrix} x+y \\ y+z \\ z-x \end{bmatrix}.$$

Siano inoltre $\mathscr{B}_1$ e $\mathscr{B}_2$ le seguenti basi:

$$\mathscr{B}_1 = \left\{ \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}, \qquad \mathscr{B}_2 = \left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Calcoliamo la matrice $\mathbf{A}_1$ associata ad f rispetto alla base $\mathscr{B}_1$ (sia il dominio che il codiminio sono equipaggiate con la stessa base). Calcoliamo $f(\mathrm{e}_i)$ e scriviamo questi vettori come combinazione lineare di $\mathbf{e}_i$:

$$f(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \mathbf{e}_1 - \mathbf{e}_3, \qquad f(\mathbf{e}_2) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \mathbf{e}_1 + \mathbf{e}_2, \qquad f(\mathbf{e}_3) = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \mathbf{e}_2 + \mathbf{e}_3.$$

Quindi, la matrice $\mathbf{A}_1$ associata ad f rispetto a $\mathscr{B}_1$ è

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix}.$$

In modo simile, per calcolare la matrice $\mathbf{A}_2$ associata ad f rispetto a $\mathscr{B}_2$ (base con cui sia dominio che codominio sono equipaggiati), calcoliamo $f(\mathbf{v}_i)$ e scriviamo i vettori così ottenuti come combinazione lineare degli stessi $\mathbf{v}_i$:

$$f(\mathbf{v}_1) = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} = 2\mathbf{v}_1 - \mathbf{v}_2, \qquad f(\mathbf{v}_2) = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \mathbf{v}_1 + \mathbf{v}_2, \qquad f(\mathbf{v}_3) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \mathbf{v}_1.$$

Quindi, la matrice $\mathbf{A}_2$ associata ad f rispetto a $\mathscr{B}_2$ è

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Quindi, le matrici $\mathbf{A}_1$ ed $\mathbf{A}_2$ sono simili.

2.2.1 Operazioni tra Matrici

In questa parte, introdurremo le operazioni tra matrici. Chiaramente, desideriamo che queste operazioni siano in qualche modo “compatibili” con le operazioni tra funzioni lineari (somma, moltiplicazione per scalare, composizione, ...), e che ad una operazione tra funzioni lineari corrisponda una operazione tra matrici.

Iniziamo con la *somma*. Siano $f, g : V \to W$ due applicazioni lineari. Siano inoltre $\mathbf{A} = [a_{ij}]$ ed $\mathbf{B} = [b_{ij}]$ le matrici associate ad f e g rispettivamente. La *somma* di f e g è l’omomorfismo[2] definito come

$$(f+g)(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v})$$

per ogni $\mathbf{v} \in V$. Per come è definita, abbiamo che la funzione $(f+g)$ può essere scritta utilizzando le matrici $\mathbf{A}$ e $\mathbf{B}$ come segue

$$\begin{aligned}(f+g)(\mathbf{v}_j) &= f(\mathbf{v}_j) + g(\mathbf{v}_j) \\ &= \sum_{i=1}^{m} a_{ij}\mathbf{w}_j + \sum_{i=1}^{m} b_{ij}\mathbf{w}_j \\ &= \sum_{i=1}^{m} (a_{ij} + b_{ij})\mathbf{w}_j.\end{aligned}$$

[2] è banale verificare che la funzione definita sia lineare

Quindi, la matrice associata alla somma di f e g ha come componenti le somme dei componenti di $\mathbf{A}$ ed $\mathbf{B}$. Diamo quindi la seguente definizione.

Definizione 2.7 (Somma di matrici)**.** Siano $\mathbf{A} = \left[a_{ij}\right]_{i=1,2\dots,m}^{j=1,2,\dots,n}$ e $\mathbf{B}\left[b_{ij}\right]_{i=1,2\dots,m}^{j=1,2,\dots,n}$ due matrici di dimensione $m \times n$ su un campo $\mathbb{K}$. Definiamo la *somma* tra $\mathbf{A}$ e $\mathbf{B}$ come la matrice le cui componenti sono le somme dei componenti di $\mathbf{A}$ con i componenti di $\mathbf{B}$:

$$\mathbf{A} + \mathbf{B} = \left[a_{ij} + b_{ij}\right]_{i=1,2\dots,m}^{j=1,2,\dots,n}.$$

Esempio 2.9. Siano $\mathbf{A}, \mathbf{B}$ due matrici 2×3 definite come

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 3 & 1 & 3 \\ 0 & 2 & 3 \end{bmatrix}.$$

La loro somma è

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 1 & 3 \\ 0 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1+3 & 2+1 & 3+3 \\ 3+0 & 2+2 & 1+3 \end{bmatrix} = \begin{bmatrix} 4 & 3 & 6 \\ 3 & 4 & 4 \end{bmatrix}.$$

Osservazione 2.4. Poiché la somma tra matrici è definita come la matrice con componenti pari alla somma delle componenti, abbiamo che tutte le proprietà della sooma tra scalari (commutativa, associativa) valgono anche per la somma tra matrici.

La prossima operazione che andremo a considerare è la *moltiplicazione per scalare*. Sia $f : V \to W$ una funzione lineare e sia $\lambda \in \mathbb{K}$ uno scalare. Definiamo la funzione (λf) come $(\lambda f)(\mathbf{v}) = \lambda \cdot f(\mathbf{v})$. Applicando la funzione ai vettori $\mathbf{v}_j$ della base $\mathscr{B}_V$ di V, otteniamo

$$(\lambda f)(\mathbf{v}_j) = \lambda \cdot f(\mathbf{v}_j) = \lambda \sum_{i=1}^{m} a_{ij}\mathbf{w}_i = \sum_{i=1}^{m} (\lambda a_{ij})\mathbf{w}_i.$$

Quindi, in modo simile a quanto fatto per la somma, abbiamo che la matrice associata alla funzione f è la matrice le cui componenti sono λ volte i componenti della matrice $\mathbf{A}$ associata ad f.

Definizione 2.8 (Prodotto tra matrice e scalare)**.** Siano $\lambda \in \mathbb{K}$ e $\mathbf{A} \in \mathbb{K}^{m\times n}$, il *prodotto* $\lambda\mathbf{A}$ è la matrice definita come segue:

$$\lambda\mathbf{A} = \left[\lambda a_{ij}\right]_{i=1,2,\dots,m}^{j=1,2,\dots,n}.$$

♢ *Osservazione* 2.5. Poiché ad ogni applicazione lineare $f \in Hom(\mathbb{K}^m, \mathbb{K}^n) \doteq \{f : \mathbb{K}^m \to \mathbb{K}^n | f \text{ lineare}\}$ corrisponde una e una sola matrice $\mathbf{A} \in \mathscr{M}_{m,n}(\mathbb{K}) \doteq \{\mathbf{A} \in \mathbb{K}^{m\times n}\}$, abbiamo che esiste una biiezione tra gli insiemi $Hom(\mathbb{K}^m, \mathbb{K}^n)$ e $\mathscr{M}_{m,n}$. Inoltre, dal fatto che $Hom(\mathbb{K}^m, \mathbb{K}^n)$ costituisca uno spazio vettoriale con le operazioni di somma e moltiplicazione per scalare, segue che l'insieme $\mathscr{M}_{m,n}(\mathbb{K})$ forma a sua volta, con le operazioni di somma e moltiplicazione per scalare definite sopra, uno spazio vettoriale.

♢ *Osservazione* 2.6. L'insieme delle matrici di dimesione $m \times n$, $\mathscr{M}_{m,n}(\mathbb{K})$, forma uno spazio vettoriale rispetto alla somma tra matrici ed al prodotto matrice-scalare.

Anche per lo spazio $\mathscr{M}_{m,n}(\mathbb{K})$ possiamo definire una base canonica considerando come elementi della base le matrici $\mathbf{E}_{i,j}$ in cui tutte le componenti eccetto una sono nulle. La componente non nulla si trova in riga i e colonna j ed assume valore unitario.

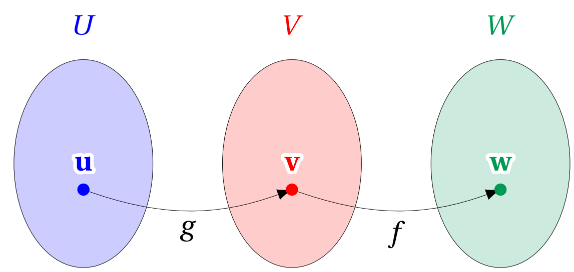
$$\mathbf{E}_{i,j} = \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

con questa definizione, qualsiasi matrice $\mathbf{A} = \left[a_{ij}\right]$ può essere scritta come

$$\mathbf{A} = \left[a_{ij}\right]_{i=1,2\dots,m}^{j=1,2,\dots,n} = \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij}\mathbf{E}_{i,j}.$$

La prossima operazione riguarda la *composizione* di omomorfismi, che ci consentirà di definire il *prodotto tra matrici*.

A tale scopo, definiamo quindi tre spazi vettoriali U, V, e W, ognuno equipaggiato con una base $\mathscr{B}_U = \{\mathbf{u}_j\}_{j=1}^{\ell}, \mathscr{B}_V = \{\mathbf{v}_h\}_{h=1}^{n}, \mathscr{B}_W = \{\mathbf{w}_i\}_{i=1}^{m}$. consideriamo inoltre due funzioni lineari $f : V \to W$, $g : U \to V$. Siano inoltre $\mathbf{A} \in \mathbb{K}^{m\times n}, \mathbf{B} \in \mathbb{K}^{n\times \ell}$, ed $\mathbf{C} \in \mathbb{K}^{m\times \ell}$ le matrici associate alle applicazioni lineari f, g, ed $f \circ g : U \to W$ rispetto alle basi associate ai tre spazi.



Per ogni vettore $\mathbf{u}_j$ della base $\mathscr{B}_U$ di U si ha

$$\begin{aligned}(f \circ g)(\mathbf{u}_j) &= f(g(\mathbf{u}_j)) \\ &= f\left(\sum_{h=1}^{n} b_{hj}\mathbf{v}_h\right)\end{aligned}$$

Ora, poiché f è lineare, possiamo riscrivere

$$f\left(\sum_{h=1}^{n} b_{hj}\mathbf{v}_h\right) = \sum_{h=1}^{n} b_{hj} f(\mathbf{v}_h).$$

Esprimendo ora f tramite i coefficienti della matrice, abbiamo

$$\begin{aligned}\sum_{h=1}^{n} b_{hj} f(\mathbf{v}_h) &= \sum_{h=1}^{n} b_{hj}\left(\sum_{i=1}^{m} a_{ih}\mathbf{w}_i\right) \\ &= \sum_{i=1}^{m}\left(\sum_{h=1}^{n} a_{ih}b_{hj}\right)\mathbf{w}_i \\ &= \sum_{i=1}^{m}\sum_{h=1}^{n} a_{ih}b_{hj}\mathbf{w}_i\end{aligned}$$

Dal fatto che $\mathbf{C}$ sia la matrice associata alla mappa $f \circ g$, abbiamo l'identità

$$(f \circ g)(\mathbf{u}_j) = \sum_{i=1}^{m} c_{ij}\mathbf{w}_i.$$

Abbiamo quindi che possiamo definire l'operazione di prodotto tra matrici.

Definizione 2.9 (Prodotto tra matrici)**.** Date due matrici $\mathbf{A} = [a_{ih}] \in \mathbb{K}^{m\times\ell}, \mathbf{B} = \left[b_{hj}\right] \in \mathbb{K}^{\ell\times n}$, il loro *prodotto* è la matrice $\mathbf{C} = \left[c_{ij}\right] \in \mathbb{K}^{m\times n}$ i cui coefficienti sono dati da

$$c_{ij} = \sum_{h=1}^{\ell} a_{ih} b_{hj}. \tag{7}$$

Il prodotto tra matrici è chiamato anche prodotto righe per colonne in quanto l'elemento in posizione (i, j) del prodotto è determinato dagli elementi in riga i del primo fattore e in colonna j del secondo.

Osserviamo che affinché sia possibile calcolare il prodotto tra due matrici, è necessario che il numero di colonne della prima matrice coincida col numero di righe della seconda. Questo equivale a dire che la dimensione del codominio di g deve essere la stessa del dominio di f.

Vediamo nel dettaglio come calcolare questo prodtto. Per farlo, consideriamo due matrici **A** ed **B** di dimensione $m\times\ell$ ed $\ell\times n$ rispettivamente. L'elemento c_{ij} in riga i e colonna j del prodotto tra **A** e **B**, $\mathbf{C} = \mathbf{AB}$ si calcola considerando la i-sima riga di **A** e la j-sima riga di **B** tramite la formula (7). Consideriamo quindi questi due elementi:

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{i\ell} \end{bmatrix}, \qquad \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{\ell j} \end{bmatrix}. \tag{8}$$

Per calcolare l'elemento c_{ij} dobbaimo sommare i prodotti componente-per-componente di questi due vettori:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{i\ell}b_{\ell j}.$$

Questa quantità è detta anche *prodotto scalare* tra i due vettori in (8). Torneremo tra poco su questa nomenclatura.

Osservazione 2.7. Il prodotto tra matrici **non** è, in genere, commutativo.

Esempio 2.10. Consideriamo le due matrici quadrate $\mathbf{A}, \mathbf{B} \in \mathbb{R}^3$ definite come

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Allora, il prodotti **AB** e **BA** risultano essere

$$\mathbf{AB} = \begin{bmatrix} 10 & 9 & 6 \\ 12 & 11 & 6 \\ 8 & 10 & 6 \end{bmatrix}, \qquad \mathbf{BA} = \begin{bmatrix} 8 & 9 & 7 \\ 10 & 13 & 13 \\ 6 & 6 & 6 \end{bmatrix}.$$

Esempio 2.11. Ovviamente, non è impossibile che due matrici commutino. Ad esempio, siano **A**, **B** definite come

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 1 & 3 \\ 1 & 3 & 2 \end{bmatrix},$$

allora i prodotti **AB** e **BA** coincidono

$$\mathbf{AB} = \mathbf{BA} = \begin{bmatrix} 10 & 13 & 13 \\ 13 & 10 & 13 \\ 13 & 13 & 10 \end{bmatrix}.$$

Definiamo ora l'elemento neutro per la moltiplicazione tra matrici.

Definizione 2.10 (Matrice identica / identità)**.** La matrice identica (detta anche matrice identità) è una matrice quadrata $\mathbf{Id}_n = [\delta_{ij}] \in \mathcal{M}_{n,n}(\mathbb{K})$ le cui uniche componenti non nulle si trovano sulla *diagonale principale*, ed hanno tutte valore uno:

$$\delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}, \qquad \mathbf{Id}_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}$$

La matrice identià è chiamata in questo modo perché per ogni matrice $\mathbf{A} \in \mathcal{M}_{m,n}(\mathbb{K})$ si ha

$$\mathbf{Id}_m\mathbf{A} = \mathbf{A}\mathbf{Id}_n = \mathbf{A}.$$

Un altro motivo per cui la matrice identità ha questo nome risiede nel fatto che è la matrice determinata dalla funzione identità tra due spazi spazi vettoriali identici definiti dalla stessa base. Infatti, si consideri $V = W \cong \mathbb{K}^n$ e sia $\mathscr{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base di V e W. Allora, la matrice $\mathbf{A} = [a_{ij}]_{i,j=1,2,\ldots,n}$ associata alla mappa identità $f(\mathbf{v}) = \mathrm{id}(\mathbf{v}) = \mathbf{v}$ deve soddisfare

$$f(\mathbf{v}_j) = \mathbf{v}_j = \sum_{i=1}^{n} a_{ij}\mathbf{v}_j.$$

E, poiché i vettori in $\mathscr{B}$ sono linearmente indipendenti, allora si deve avere

$$a_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}.$$

Una sorta di "generalizzazione" del concetto di matrice identità è quello di matrice diagonale.

Definizione 2.11. Una matrice quadrata $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{K})$ si dice una *matrice diagonale* quando tutte le componenti a_{ij} per cui $i \neq j$ sono nulle.

Esempio 2.12. Le seguenti sono matrici diagonali:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \qquad \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Pur non rispettando la proprietà commutativa, la moltiplicazione tra matrici soddisfa diverse proprietà distributive e associative.

Proposizione 2.8. *Siano* $\mathbf{A}, \mathbf{B}$*, ed* $\mathbf{C}$ *tre matrici, con dimensioni compatibili per le operazioni che menzioneremo, e siano* λ *e* μ *due scalari. Valgono le seguenti proprietà:*

1. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$;

2. $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$;

3. $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$;

4. $\lambda(\mathbf{AB}) = (\lambda\mathbf{A})\mathbf{B} = \mathbf{A}(\lambda\mathbf{B})$;

5. $(\lambda+\mu)\mathbf{A} = \lambda\mathbf{A}+\mu\mathbf{A}$;

6. $(\lambda\mu)\mathbf{A} = \lambda(\mu\mathbf{A})$.

Omettiamo la dimostrazione in quanto queste proprietà discendono dalle euivalenti proprietà della composizione di funzioni lineari. Alternativamente, è sufficiente esprimere la componenti (i, j) delle matrici a destra e sinistra del segno di uguaglianza e verificare che coincidono.

Osservazione 2.8. L'insieme delle matrici quadrate di dimensione n forma un anello non commutativo (si veda la definizione 1.3) rispetto alla somma tra matrici ed al prodotto riga-colonna.

La moltiplicazione tra matrici è una delle operazioni più importanti (se non la più importante) nell'algebra lineare. Prima di approfondire ulteriormente questa operazione però, introduciamo una nuova operazione che ci permetterà, tra le altre cose, di semplificare di molto la notazione che stiamo utilizzando.

Definizione 2.12 (Matrice trasposta)**.** Sia $\mathbf{A} = \left[a_{ij}\right] \in \mathcal{M}_{m,n}(\mathbb{K})$. La *trasposta* di $\mathbf{A}$, denotata con $\mathbf{A}^\intercal$, è la matrice $\mathbf{A}^\intercal = \left[\hat{a}_{ij}\right] \in M_{n,m}(\mathbb{K})$ i cui coefficienti in posizione (i, j) sono i coefficienti in posizione (j, i) di $\mathbf{A}$:

$$\left[\hat{a}_{ij}\right]_{i=1,2,\dots,m}^{j=1,2,\dots,n} = \left[a_{ji}\right]_{j=1,2,\dots,n}^{i=1,2,\dots,m}$$

Esempio 2.13. Sia $\mathbf{A}$ definita come

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix},$$

allora la matrice $\mathbf{A}^\intercal$ è

$$\mathbf{A}^\intercal = \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}.$$

Definizione 2.13 (Matrice simmetrica)**.** Una matrice quadrata $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{K})$ si dice *simmetrica* quando coincide con la sua trasposta: $\mathbf{A} = \mathbf{A}^\intercal$.

Esempio 2.14. La matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 3 & 3 & 0 \\ 1 & 3 & 1 & 4 \\ 2 & 0 & 4 & 0 \end{bmatrix}$$

è simmetrica.

L'operazione di trasposizione soddisfa le seguenti proprietà.

Proposizione 2.9. *Siano* $\mathbf{A}, \mathbf{B} \in M_{m,n}(\mathbb{K})$ *e* $\mathbf{C} \in M_{n,\ell}(\mathbb{K})$ *tre matrici e* $\lambda \in \mathbb{K}$ *uno scalare. Allora, le seguenti identità sono vere:*

1. $(\mathbf{A}^\top)^\top = \mathbf{A}$;
2. $(\mathbf{A}+\mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top$;
3. $(\lambda\mathbf{A})^\top = \lambda(\mathbf{A}^\top)$;
4. $(\mathbf{AC})^\top = \mathbf{C}^\top\mathbf{A}^\top$.

Omettiamo la dimostrazione in quanto è sufficiente esprimere la componenti (i,j) delle matrici a destra e sinistra del segno di uguaglianza e verificare che coincidono.

Un caso particolare di prodotto tra matrici è il prodotto matrice-vettore, ovvero quando la matrice $\mathbf{B}$ è costituita da una singola colonna, $\mathbf{B} \in \mathbb{K}^{n\times 1}$.
Consideriamo quindi un vettore $\mathbf{v} = [v_1, v_2, \dots, v_n]^\top \in \mathbb{K}^n$ ed una matrice $\mathbf{A} = \left[a_{ij}\right] \in \mathbb{K}^{m\times n}$. Il prodotto $\mathbf{Av}$ è un vettore (colonna) $\mathbf{w} \in \mathbb{K}^m$ definito come

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

Abbiamo che se consideriamo come funzione f la moltiplicazione di una matrice per un vettore: $f : V \to W, \mathbf{v} \mapsto \mathbf{Av}$, questa è una funzione lineare. Inoltre, la matrice associata all'omomorfismo f rispetto alle basi canoniche di $\mathbb{K}^n$ e $\mathbb{K}^m$ è esattamente la matrice $\mathbf{A}$.

Un altro caso particolare di moltiplicazione tra matrici è il prodotto di un vettore riga per un vettore colonna, che risulta in uno scalare:

$$\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = v_1 w_1 + v_2 w_2 + \ldots + v_n w_n \in \mathbb{K}.$$

Da questa osservazione, abbiamo che possiamo definire il prodotto scalare tra due vettori.

Definizione 2.14. Siano $\mathbf{v}$ e $\mathbf{w}$ due vettori ad n componenti con valori in un campo $\mathbb{K}$, $\mathbf{v}, \mathbf{w} \in \mathbb{K}^n$. Il *prodotto scalare* di $\mathbf{v}$ con $\mathbf{w}$, denotato come $\mathbf{v}\cdot\mathbf{w}$ o $\langle \mathbf{v}, \mathbf{w}\rangle$, è lo scalare

$$\mathbf{v}\cdot\mathbf{w} = \langle \mathbf{v}, \mathbf{w}\rangle = \sum_{i=1}^{n} v_i w_i.$$

Grazie a questa definizione, abbiamo che in un prodotto matrice-matrice, possiamo interpretare la componente in posizione (i,j) del prodotto come il prodotto scalare tra la riga i della prima matrice e la colonna j della seconda.

Introduciamo ora il concetto di matrice inversa. Per fare ciò, osserviamo prima che l'esistenza di un elemento neutro per l'operazione di moltiplicazione tra matrici rende naturale definire il concetto di inversa.

Definizione 2.15 (Matrice invertibile e matrice inversa)**.** Sia $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{K})$ una matrice quadrata. Questa matrice è detta *invertibile* se esiste una matrice $\mathbf{B} \in \mathcal{M}_{n,n}(\mathbb{K})$ tale per cui $\mathbf{AB} = \mathbf{BA} = \mathbf{Id}_n$. In tal caso, la matrice $\mathbf{B}$ è detta l'*inversa* di $\mathbf{A}$, e viene indicata con $\mathbf{A}^{-1}$.

Si noti che solo le matrici quadrate possono essere invertibili. Infatti, affinché un omomorfismo $f : V \to W$ sia invertibile, questo omomorfismo deve essere biiettivo, e quindi un isomorfismo. Ma l'esistenza di un isomorfismo tra due spazi vettoriali V e W implica che i due spazi vettoriali siano isomorfi e che quindi abbiano la stessa dimensione (si veda il corollario 2.1) e, di conseguenza, la matrice associata all'isomorfismo f deve avere ugual numero di righe e di colonne.

Alcuni testi presentano i concetti di *inversa destra* ed *inversa sinistra* per una generica matrice rettangolare $\mathbf{A} \in \mathcal{M}_{m,n}(\mathbb{K})$, mostrando successivamente che per matrici quadrate queste due matrici coincidono (quando esistono). Noi invece considereremo soltanto il concetto di inversa di una matrice quadrata, senza così bisogno di specificare se sia l'inversa destra o sinistra.

L'operazione di 'inversione' di una matrice soddisfa le seguenti proprietà.

Proposizione 2.10. *Siano* $\mathbf{A}, \mathbf{B} \in \mathcal{M}_{n,n}(\mathbb{R})$ *due matrici invertibili. Allora valgono la seguenti identià*

1. $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$;
2. $(\lambda\mathbf{A})^{-1} = \lambda^{-1}\mathbf{A}^{-1}$;
3. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$;
4. $(\mathbf{A}^{-1})^\intercal = (\mathbf{A}^\intercal)^{-1}$.

Anche in questo caso omettiamo la dimostrazione in quanto questi risultati seguono immediatamente dalle proprietà degli omomorfismi associati ad $\mathbf{A}$ e $\mathbf{B}$.
In particolare, il punto 4. della precedente proposizione ci permette di utilizzare la scrittura $\mathbf{A}^{-\intercal} = (\mathbf{A}^{-1})^\intercal = (\mathbf{A}^\intercal)^{-1}$ in modo non ambiguo, in quanto fare prima la trasposizione o l'inversione di una matrice non cambia il risultato.

Più avanti nel corso, vedremo un algoritmo su come calcolare l'inversa di una matrice quadrata.

Poiché ad una matrice è associata una ed una sola funzione lineare (fissata una base degli spazi di dominio e codominio) abbiamo che possiamo estendere i concetti di *nucleo*, *immagine*, *nullità*, e *rango* anche alle matrici.

Definizione 2.16 (Nucleo e immagine di una matrice)**.** Sia $\mathbf{A} \in \mathcal{M}_{n,m}(\mathbb{K})$ una matrice su un campo $\mathbb{K}$. Allora, il nucleo e l'immagine di $\mathbf{A}$ sono definiti, rispettivamente come

$$\begin{aligned} &\ker\mathbf{A} \doteq \left\{\mathbf{v} \in \mathbb{K}^m : \mathbf{Av} = \mathbf{0}\right\}, \\ &\mathrm{Im}\mathbf{A} \doteq \left\{\mathbf{w} \in \mathbb{K}^n : \exists\mathbf{v} \in \mathbb{K}^m, \mathbf{w} = \mathbf{A}\right\} \end{aligned}$$

Poiché l'applicazione $f : V \to W, \mathbf{v} \mapsto \mathbf{Av}$ è un omomorfismo, abbiamo che il nucleo e l'immagine di una matrice, così come già dimostrato per le funzioni lineari, sono sottospazi vettoriali. Ha quindi senso definire la nullità ed il rango di una matrice come le dimensioni di questi spazi.

Definizione 2.17 (Rango e nullità di una matrice)**.** Sia $\mathbf{A}$ una matrice $m \times n$ a coefficienti in $\mathbb{K}$, e sia $f : \mathbb{K}^n \to \mathbb{K}^m$ la funzione lieneare data dalla moltiplicazione della matrice $\mathbf{A}$ per un vettore colonna:

$f : \mathbf{v} \mapsto \mathbf{Av}$. Definiamo il *rango* e la *nullità* della matrice $\mathbf{A}$ come il rango e la nullità della funzione f, rispettivamente:

$$\mathrm{rk}(\mathbf{A}) \doteq \mathrm{rk}(f) = \dim \mathrm{Im}(f) \tag{9a}$$

$$\mathrm{null}(\mathbf{A}) \doteq \mathrm{null}(f) = \dim \ker(f) \tag{9b}$$

Poiché la nullità ed il rango di una matrice sono definiti come la nullità ed il rango di un particolare omomorfismo, si ha che il teorema di nullità + rango 2.1 si può estendere alle matrici.

Teorema 2.2. *Sia* $\mathbf{A} \in \mathcal{M}_{m,n}(\mathbb{K})$ *una matrice su un campo* $\mathbb{K}$*. Allora vale la seguente identità:*

$$\mathrm{rk}(\mathbf{A}) + \mathrm{null}(\mathbf{A}) = n.$$

Abbiamo che il rango di una matrice (e quindi della funzione lineare associata) può essere calcolato come il numero di colonne linearmente indipendenti della matrice stessa.

Proposizione 2.11. *Il rango di una matrice* $\mathbf{A}$ *è il massimo numero di colonne linearmente indipendentidi* $\mathbf{A}$.

◇ *Dimostrazione.* Per la funzione $f : \mathbf{v} \mapsto \mathbf{Av}$, l'immagine $\mathrm{Im}(f)$ è data dallo spazio vettoriale generato dalle colonne della matrice $\mathbf{A}$. Infatti, un insieme di generatori di $\mathrm{Im}\mathbf{A} = \mathrm{Im} f$ è data dalle immagini di una base di $\mathbb{K}^n$ tramite f. Consideriamo quindi la base canonica $\{\mathbf{e}_j\}_{j=1,2,\ldots,n}$ di $\mathbb{K}^n$. Si ha che $f(\mathbf{e}_j) = \mathbf{Ae}_j$. Il risultato di questo prodotto matrice-per-vettore è la $j-$sima colonna di mA. Quindi, l'insieme delle colonne delle matrici $\mathbf{A}$ genera l'immagine $\mathrm{Im}\mathbf{A}$.
Ricordiamo che il rango di $\mathbf{A}$ coincide, per definizione, con la dimensione dell'immagine $\mathrm{Im}\mathbf{A}$. Questa dimensione è pari al numero di vettori linearmente indipendenti in un suo insieme di generatori. Quindi, la dimensione dell'immagine di $\mathbf{A}$ coincide col numero di colonne linearmente indipendenti di $\mathbf{A}$. □

Grazie alla proposizione 2.11, si ha che alcuni testi definiscono il rango di una matrice come il massimo numero di colonne linearmente indipendenti di $\mathbf{A}$. Per questo motivo, questa quantità viene anche chiamata *rango per colonne*.
In modo simile, si ha che il *rango per righe* di una matrice $\mathbf{A}$ viene definito come il massumo numero di righe linearmente indipendenti di $\mathbf{A}$.

In realtà, abbiamo che il rango per righe ed il rango per colonne di una matrice $\mathbf{A}$ coincidano. Questo segue dal prossimo risultato.

Teorema 2.3. *Il rango (per colonne) di una matrice* $\mathbf{A} \in \mathcal{M}_{m,n}(\mathbb{K})$ *su un campo* $\mathbb{K}$ *coincide col rango della sua trasposta* $\mathbf{A}^\mathsf{T}$*:*

$$\mathrm{rk}(\mathbf{A}) = \mathrm{rk}(\mathbf{A}^\mathsf{T}).$$

Corollario 2.3. *Il rango per colonne di una matrice* $\mathbf{A} \in \mathcal{M}_{m,n}(\mathbb{K})$ *coincide col rango per righe della stessa matrice.*

Dimostrazione. Banalmente, il rango per righe di una matrice $\mathbf{A}$ coincide col rango per colonne della trasposta $\mathbf{A}^\mathsf{T}$. Poiché, grazie al teorema 2.3, i ranghi per colonne di $\mathbf{A}$ ed $\mathbf{A}^\mathsf{T}$ coincidono, si ottiene la tesi. □

Poiché il rango per righe ed il rango per colonne coincidono, da qui in poi utilizzeremo semplicemente il termine *rango*.

Il concetto di rango permette di darci una condizione equivalente all'essere invertibili.

Proposizione 2.12. *Sia* $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$ *una matrice quadrata su un campo* $\mathbb{K}$. *Allora* $\mathbf{A}$ *è invertibile se e solo se la matrice ha rango massimo:* $\mathrm{rk}\mathbf{A} = n$.

La dimostrazione segue dal fatto che un endomorfismo è biiettivo se e solo se il nucleo coincide con lo spazio nullo (o, equivalentemente, se l'immagine è l'intero codominio). Questo coincide col fatto che il rango della matrice sia massimo.

2.2.2 Cambi di Base

Come abbiamo spesso rimarcato dall'inizio della nostra discussione sulle matrici, si ha che, fissata una funzione lineare $f : V \to W$, la matrice associata dipende dalla scelta della base degli spazi V e W. In questa parte, studieremo nel dettaglio cosa accade alla matrice $\mathbf{A}$ quando cambiamo queste basi.

Consideriamo quindi due spazi vettoriali, V e W, di dimensione n ed m rispettivamente. Sia $f : V \to W$ un omomorfismo. Siano inoltre $\mathscr{B}_V = \{\mathbf{v}_i\}_{i=1,2,\dots,n}$ e $\mathscr{B}'_V = \{\mathbf{v}'_i\}_{i=1,2,\dots,n}$ due basi di V, e $\mathscr{B}_W = \{\mathbf{w}_i\}_{i=1,2,\dots,n}$ e $\mathscr{B}'_W = \{\mathbf{w}'_i\}_{i=1,2,\dots,n}$ due basi di W. Infine, indichiamo con $\mathbf{A} = \left[a_{ij}\right]$ la matrice associata af f rispetto alle basi $\mathscr{B}_V$ e $\mathscr{B}_W$, e con $\mathbf{A}' = \left[a'_{ij}\right]$ la matrice associata af f rispetto alle basi $\mathscr{B}'_V$ e $\mathscr{B}'_W$. Ricordiamo che, per definizione, valgono le seguenti identità:

$$f(\mathbf{v}_j) = \sum_{i=1}^{n} a_{ij}\mathbf{w}_i, \qquad f(\mathbf{v}'_j) = \sum_{i=1}^{n} a'_{ij}\mathbf{w}'_i, \qquad \forall j = 1,2,\dots,m.$$

Indichiamo ora con $\alpha : V \to \mathbb{K}^n$ l'isomorfismo che associa ad ogni vettore $\mathbf{v}$ la n−upla $(\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{K}^n$ delle sue coordinate rispetto alla base $\mathscr{B}_V$. In modo simile, indichiamo con $\alpha' : V \to \mathbb{K}^n$ l'isomorfismo che associa a $\mathbf{v}$ la n−upla $(\lambda'_1, \lambda'_2, \dots, \lambda'_n)$ di coordinate rispetto alla base $\mathscr{B}'_V$.
In modo analogo, indichiamo con $\beta, \beta' : W \to \mathbb{K}^m$ gli isomorfismi che associano ad un vettore $\mathbf{w} \in W$ le coordinate $(\mu_1, \mu_2, \dots, \mu_m)$ e $(\mu'_1, \mu'_2, \dots, \mu'_m)$ in $\mathbb{K}^m$ rispetto alle basi $\mathscr{B}_W$ e $\mathscr{B}'_W$.

Consideriamo la composizione di α' con l'inverso di α, denotato con α^{-1}. Notiamo che questo inverso esiste in quanto α è un isomorfismo, ovvero una mappa lineare invertibile. La mappa $F_P = \alpha' \circ \alpha^{-1}$ è un isomorfismo da $\mathbb{K}^n$ in $\mathbb{K}^n$, a cui corrisponde una matrice $\mathbf{P}$.

$$\begin{tikzcd} & V \arrow[dl, "\alpha"'] \arrow[dr, "\alpha'"] & \\ \mathbb{K}^n \arrow[rr, "F_P"'] & & \mathbb{K}^n \end{tikzcd}$$

In modo analogo, consideriamo la composizione di β' con l'inverso di β, denotato con β^{-1}. La mappa $F_Q = \beta' \circ \beta^{-1}$ è un isomorfismo da $\mathbb{K}^m$ in $\mathbb{K}^m$, a cui corrisponde una matrice $\mathbf{Q}$.

$$\begin{tikzcd} & W \arrow[dl, "\beta"'] \arrow[dr, "\beta'"] & \\ \mathbb{K}^m \arrow[rr, "F_Q"'] & & \mathbb{K}^m \end{tikzcd}$$

Le matrici $\mathbf{P}$ ed $\mathbf{Q}$ sono sicuramente invertibili poiché le funzioni corrispondenti F_P ed F_Q sono isomorfismi.

Calcoliamo ora le matrici $\mathbf{P}$ ed $\mathbf{Q}$. Iniziamo con $\mathbf{P}$, la matrice associata all'isomorfismo $F_P : \mathbb{K}^n \to \mathbb{K}^n$. La $j-$sima colonna di $\mathbf{P}$ è l'immagine del $j-$simo vettore della base canonica, $\mathbf{e}_j = [0, \dots, 0, 1, 0, \dots, 0]^\top$. Abbiamo quindi, per definizione di F_P che

$$F_P(\mathbf{e}_j) = \alpha'(\alpha^{-1}(\mathbf{e}_j)) = \alpha'(\mathbf{v}_j),$$

dove, ricordiamo, $\mathbf{v}_j$ è il $j-$simo elemento della base $\mathscr{B}_V$ di V. Il vettore $\alpha'(\mathbf{v}_j)$, che costituisce la $j-$sima colonna di $\mathbf{P}$ consiste nelle coordinate di $\mathbf{v}_j$ rispetto alla base $\mathscr{B}'_V$.
Abbiamo quindi che le colonne della matrice $\mathbf{P}$ sono le coordinate dei vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ della base $\mathscr{B}_V$ calcolate però rispetto alla base $\mathscr{B}'_V$.

Osservazione 2.9. Con un ragionamento analogo, si potrebbero costruire le colonne dell'inversa di $\mathbf{P}$, $\mathbf{P}^{-1}$ scambiando i ruoli di $\mathscr{B}_V$ e $\mathscr{B}'_V$. In altre parole, le colonne di $\mathbf{P}^{-1}$ sono le coordinate dei vettori $\mathbf{v}'_1, \mathbf{v}'_2, \dots, \mathbf{v}'_n$ rispetto alla base $\mathscr{B}_V$.

In modo simile a quanto fatto per $\mathbf{P}$, possiamo costruire la matrice $\mathbf{Q}$ associata all'isomorfismo F_Q come la matrice le cui colonne sono le coordinate dei vettori $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ della base $\mathscr{B}_W$ rispetto alla base $\mathscr{B}'_W$.

Se ora quindi denotiamo con $F_A : \mathbb{K}^n \to \mathbb{K}^m$ ed $F_{A'} : \mathbb{K}^n \to \mathbb{K}^m$ le applicazioni lineari definite come la moltiplicazione per $\mathbf{A}$ ed $\mathbf{A}'$ rispettivamente, abbiamo la seguente identità:

$$\begin{tikzcd}
\mathbb{K}^n \arrow[rrr, "F_A"] \arrow[ddd, "F_P"'] & & & \mathbb{K}^m \arrow[ddd, "F_Q"] \\
 & & & \\
 & V \arrow[uul, "\alpha"'] \arrow[r, "f"] \arrow[dl, "\alpha'"] & W \arrow[uur, "\beta"] \arrow[dr, "\beta'"'] & \\
\mathbb{K}^n \arrow[rrr, "F_{A'}"'] & & & \mathbb{K}^m
\end{tikzcd}$$

$$F_{A'} \circ F_P = F_Q \circ F_A,$$

che equivale alla seguente identità matriciale:

$$\mathbf{A}'\mathbf{P} = \mathbf{Q}\mathbf{A},$$

che può essere riscritta nei seguenti due modi:

$$\mathbf{A}' = \mathbf{Q}\mathbf{A}\mathbf{P}^{-1}, \qquad \mathbf{A} = \mathbf{Q}^{-1}\mathbf{A}'\mathbf{P}.$$

Queste due espressioni permettono di determinare la matrice $\mathbf{A}'$ di un'applicazione lineare $f : V \to W$ rispetto alle basi $\mathscr{B}'_V$ di V e $\mathscr{B}'_W$ di W quando è nota la matrice $\mathbf{A}$ di f rispetto alle basi $\mathscr{B}_V$ e $\mathscr{B}_W$, e le matrici del cambio di base $\mathbf{P}$ ed $\mathbf{Q}$.

Esempio 2.15. Vediamo un esempio su come calcolare la matrice del cambio di base. Sia $V = \mathbb{R}^3$ e consideriamo le due basi

$$\mathscr{B}_V = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}, \qquad \mathscr{B}'_V = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Vogliamo calcolare la matrice $\mathbf{P}$ del cambio di base da $\mathscr{B}_V$ a $\mathscr{B}'_V$.

Iniziamo calcolando la mappa lineare $\alpha : V \to \mathbb{R}^3$ che ad un generico vettore $\mathbf{v} = [x, y, z]^\top \in V$ associa le coordinate $[\lambda_1, \lambda_2, \lambda_3]^\top \in \mathbb{R}^3$ di $\mathbf{v}$ rispetto a $\mathscr{B}_V$. Per fare ciò, esprimiamo il vettore $\mathbf{v}$ come combinazione lineare degli elementi della base $\mathscr{B}_V$:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \lambda_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Le coordinate λ_i si calcolano risolvendo il sistema lineare

$$\begin{cases} \lambda_1 + \lambda_2 = x \\ \lambda_1 = y \\ \lambda_2 + \lambda_3 = z \end{cases} \Leftrightarrow \begin{cases} \lambda_1 = y \\ \lambda_2 = x - \lambda_1 \\ \lambda_3 = z - \lambda_2 \end{cases} \Leftrightarrow \begin{cases} \lambda_1 = y \\ \lambda_2 = x - y \\ \lambda_3 = -x + y + z \end{cases},$$

da cui si ottiene l'omomorfismo

$$\alpha : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} y \\ x - y \\ -x + y + z \end{bmatrix}.$$

In modo simile, calcoliamo l'omomorfismo $\alpha' : V \to \mathbb{R}^3$ che ad ogni vettore $\mathbf{v} \in V$ associa le coordinate $[\mu_1, \mu_2, \mu_3]^\top \in \mathbb{R}^3$ di $\mathbf{v}$ rispetto, questa volta, a $\mathscr{B}'_V$. Quindi, scriviamo $\mathbf{v}$ come combinazione lineare degli elementi della base $\mathscr{B}'_V$:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mu_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \mu_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \mu_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},$$

da cui otteniamo il sistema lineare

$$\begin{cases} \mu_1 + \mu_3 = x \\ \mu_2 = y \\ \mu_2 + \mu_3 = z \end{cases} \Leftrightarrow \begin{cases} \mu_1 = x - \mu_3 \\ \mu_2 = y \\ \mu_3 = z - \mu_2 \end{cases} \Leftrightarrow \begin{cases} \mu_1 = x - (z - y) \\ \mu_2 = y \\ \mu_3 = z - y \end{cases} \Leftrightarrow \begin{cases} \mu_1 = x + y - z \\ \mu_2 = y \\ \mu_3 = -y + z \end{cases}.$$

Quindi, l'applicazione $\alpha' : V \to \mathbb{R}^3$ è data da

$$\alpha' : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} x + y - z \\ y \\ -y + z \end{bmatrix}.$$

Il prossimo passo consiste nel calcolare l'omomorfismo $\alpha^{-1} : \mathbb{R}^3 \to V$ inverso di α. Per fare ciò, cerchiamo il vettore $[a, b, c]^\top$ tale per cui

$$\alpha\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Dobbiamo quindi trovare i valori di $a, b,$ e c, scritti rispetto a $x, y,$ e z tali per cui

$$\alpha\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = \begin{bmatrix} b \\ a-b \\ -a+b+c \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Dobbiamo quindi risolvere il seguente sistema lineare nelle incognite $a, b,$ e c:

$$\begin{cases} b = x \\ a-b = y \\ -a+b+c = z \end{cases} \Leftrightarrow \begin{cases} b = x \\ a = y+b \\ c = z+a-b \end{cases} \Leftrightarrow \begin{cases} b = x \\ a = y+x \\ c = z+(y+x)-x \end{cases} \Leftrightarrow \begin{cases} b = x \\ a = x+y \\ c = y+z \end{cases},$$

da cui l'omomorfismo inverso risulta essere

$$\alpha^{-1}: \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} x+y \\ x \\ y+z \end{bmatrix}.$$

Per sicurezza, verifichiamo che la composizione tra α e il suo omomorfismo inverso sia effettivamente la mappa identità. Per fare ciò, valutiamo la composizione $\alpha^{-1} \circ \alpha$ (equivalentemente, si ptrebbe valutare la composizione opposta, $\alpha \circ \alpha^{-1}$):

$$\alpha^{-1}\left(\alpha\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right)\right) = \alpha^{-1}\left(\begin{bmatrix} y \\ x-y \\ -x+y+z \end{bmatrix}\right) = \begin{bmatrix} y+(x-y) \\ y \\ (x-y)+(-x+y+z) \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Ora che abbiamo calcolato gli omomorfismi α^{-1} ed α', possiamo calcolare la funzione lineare $F_P = \alpha' \circ \alpha^{-1}$:

$$\alpha'\left(\alpha^{-1}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right)\right) = \alpha'\left(\begin{bmatrix} x+y \\ x \\ y+z \end{bmatrix}\right) = \begin{bmatrix} (x+y)+x-(y+z) \\ x(y+z)-x \end{bmatrix} = \begin{bmatrix} 2x+z \\ x \\ y+z \end{bmatrix}.$$

Abbiamo quindi che la mappa del cambio di base F_P è definita come segue:

$$F_P: \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} 2x+z \\ x \\ y+z \end{bmatrix}.$$

Ora, calcoliamo la matrice $\mathbf{P}$ del cambio di base. Questa matrice ha come colonne i vettori ottenuti applicando l'omomorfismo F_P alla base canonica $\{\mathbf{e}_j\}_{j=1,2,3}$. Calcoliamo quindi questi vettori:

$$F_P(\mathbf{e}_1) = F_P\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \qquad F_P(\mathbf{e}_2) = F_P\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad F_P(\mathbf{e}_3) = F_P\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

In conclusione, la matrice $\mathbf{P}$ del cambio di base è la matrice

$$\mathbf{P} = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Esempio 2.16. Vediamo un esempio che coinvolga l'intero diagramma. A tal scopo, consideriamo $V = \mathbb{R}^3$, $W = \mathbb{R}^2$, e le basi

$$\mathscr{B}_V = \left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}, \qquad \mathscr{B}'_V = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\},$$

$$\mathscr{B}_W = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}, \qquad \mathscr{B}'_W = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$

Infine, consideriamo l'omomorfismo $f : V \to W$ definito come

$$f : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} y+z \\ x-z \end{bmatrix}.$$

Le matrici $\mathbf{A}, \mathbf{A}' \in \mathscr{M}_{2,3}(\mathbb{K})$ dell'omomorfismo f rispetto alle basi $\mathscr{B}_V, \mathscr{B}_W$ e $\mathscr{B}'_V, \mathscr{B}'_W$ rispettivamente sono

$$\mathbf{A} = \begin{bmatrix} -1 & 0 & 0 \\ 3 & 1 & 1 \end{bmatrix}, \qquad \mathbf{A}' = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}.$$

Calcoliamo ora la matrice $\mathbf{P} \in \mathscr{M}_{3,3}(\mathbb{K})$ del cambio di base tra $\mathscr{B}_V$ e $\mathscr{B}'_V$. Gli omomorfismi α, α', e α^{-1} sono dati da

$$\alpha : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} z-x \\ x \\ x+y-z \end{bmatrix}, \qquad \alpha' : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \qquad \alpha^{-1} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} y \\ x+z \\ x+y \end{bmatrix}.$$

L'omomorfismo $F_P = \alpha' \circ \alpha^{-1}$ è dato da

$$\alpha' \circ \alpha^{-1} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} y \\ x+z \\ x+y \end{bmatrix},$$

da cui si ottiene la matrice $\mathbf{P}$

$$\mathbf{P} = \left[F_P(\mathbf{e}_j) \right]_{j=1,2,3} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Passiamo ora al calcolo della matrice $\mathbf{Q} \in \mathscr{M}_{2,2}(\mathbb{K})$ del cambio di base tra $\mathscr{B}_W$ e $\mathscr{B}'_W$. Gli omomorfismi β, β', e β^{-1} sono dati da

$$\beta : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} y \\ x-y \end{bmatrix}, \qquad \beta' : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x \\ y \end{bmatrix}, \qquad \beta^{-1} : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x+y \\ x \end{bmatrix}.$$

L'omomorfismo $F_Q = \beta' \circ \beta^{-1}$ è dato da

$$\beta' \circ \beta^{-1} : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x+y \\ x \end{bmatrix},$$

da cui si ottiene la matrice **Q**

$$\mathbf{Q} = \left[F_Q(\mathbf{e}_j)\right]_{j=1,2} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Possiamo ora verificare l'identità $\mathbf{A}'\mathbf{P} = \mathbf{QA}$:

$$\mathbf{A}'\mathbf{P} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ -1 & 0 & 0 \end{bmatrix}$$

$$\mathbf{QA} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ -1 & 0 & 0 \end{bmatrix}.$$

3 Sistemi Lineari

In questa parte, affronteremo lo studio dei sistemi lineari, analizzando in particolare come la teoria delle matrici e dei vettori si può utilizzare per risolvere tali sistemi.

Consideriamo quindi un generico sistema di m equazioni lineari con n incognite $x_1, x_2, \ldots, x_n$.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \tag{10}$$

Notiamo che se indichiamo rispettivamente con $\mathbf{A} \in \mathbb{K}^{m\times n}$ la matrice dei coefficienti, $\mathbf{v} \in \mathbb{K}^n$ il vettore delle incognite, e con $\mathbf{b} \in \mathbb{K}^m$ il vettore dei termini noti:

$$\mathbf{A} = \left[a_{ij}\right] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Il sistema lineare in (10) si può dunque riescrive nel seguente modo:

$$\mathbf{Ax} = \mathbf{b}. \tag{11}$$

Ovvero, estendedno la scrittura, come

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

Definizione 3.1 (Sistema lineare omogeneo)**.** Un sistema lineare nella forma (11) si dice *omogeneo* quando il vettore $\mathbf{b}$ coincide col vettore nullo: $\mathbf{b} = \mathbf{0}$. Equivalentemente, il sistema lineare (10) viene detto omogeneo quando i termini noti sono tutti nulli: $b_i = 0$, $\forall i = 1, 2, \ldots, m$.

I sistemi omogenei sono quindi sistemi del tipo

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\ \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0 \end{cases} \qquad \Leftrightarrow \qquad \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

La definizione di sistemi lineari omogenei permette di ottenere diverse proprietà.

Proposizione 3.1. *L'insieme delle soluzioni di un sistema lineare omogeneo di m equazioni in n incognite a coefficienti in* $\mathbb{K}$ *è un sottospazio vettoriale di* $\mathbb{K}^n$.

Dimostrazione. L'insieme delle soluzioni $\mathbf{x}$ al problema omogeneo $\mathbf{Ax} = \mathbf{0}$ coincide col nucleo $\ker \mathbf{A}$ della matrice $\mathbf{A}$ che abbiamo già dimostrato essere uno spazio vettoriale. □

Proposizione 3.2. *Ogni soluzione del sistema lineare non-omogeneo* $\mathbf{Ax} = \mathbf{b}$ *può essere scritta come la somma di una soluzione particolare del sistema con una soluzione del sistema omogeneo associato.*

Dimostrazione. Per dimostrare questo risultato, si consideri $\mathbf{x}$ la soluzione particolare al sistema lineare $\mathbf{Ax} = \mathbf{b}$, e sia $\mathbf{y}$ la soluzione al problema omogeneo associato: $\mathbf{Ay} = \mathbf{0}$. Poiché vale 'identità

$$\mathbf{A}(\mathbf{x}+\mathbf{y}) = \mathbf{Ax}+\mathbf{Ay} = \mathbf{b}+\mathbf{0} = \mathbf{b},$$

si ha che il vettore $\mathbf{x}+\mathbf{y}$ è anch'esso soluzione al sistema lineare. □

Detto in altri termini, questa proposizione ci dice che l'insieme delle soluzioni $S_{\mathbf{b}}$ di $\mathbf{Ax} = \mathbf{b}$ può essere scritto come

$$S_{\mathbf{b}} = S_{\mathbf{0}} + \tilde{\mathbf{x}} = \left\{\mathbf{y}+\tilde{\mathbf{x}} : A\mathbf{y} = \mathbf{0}\right\},$$

dove $S_{\mathbf{0}}$ è l'insieme delle soluzioni al sistema omogeneo $\mathbf{Ax} = \mathbf{0}$ e $\tilde{\mathbf{x}}$ è una soluzione particolare al sistema non omogeneo $\mathbf{Ax} = \mathbf{b}$.

Osservazione 3.1. Questo risultato è analogo a quanto visto nella proposizione 2.6 per la controimmagine di omomorfismi.

Introduciamo ora una nozione che ci permetterà, grazie a quanto dimostreremo tra poco, di poter verificare se un sistema lineare ammette soluzione. La matrice $\mathbf{A}$ in (11) è detta *matrice del sistema* o *matrice dei coefficienti*. Definiamo la *matrice aumentata* del sistema lineare (11) come la matrice $\mathbf{A}$ con l'aggiunta di una colonna contenente i temrini noti di $\mathbf{b}$.

$$[\mathbf{A}|\mathbf{b}] \doteq \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]. \tag{12}$$

Si noti che la linea verticale che separa gli elementi di $\mathbf{A}$ dagli elementi di $\mathbf{b}$ non ha nessun significato matematico: è solo una notazione per aiutarci a distinguere i coefficienti del sistema lineare dai termini noti dello stesso.

Possiamo ora enunciare il seguente risulttato.

Proposizione 3.3. *Sia* $\mathbf{Ax} = \mathbf{b}$ *un sistema lineare ad* m *equazioni ed* n *incocgnite. Sia* $f_{\mathbf{A}} : \mathbb{K}^n \to \mathbb{K}^m$ *l'omomorfismo dato da* $f_{\mathbf{A}}(\mathbf{x}) = \mathbf{Ax}$. *Allora, i seguenti enunciati sono equivalenti:*

1. *il sistema* $\mathbf{Ax} = \mathbf{b}$ *ammette soluzioni;*
2. *il vettore dei termini noti appartiene all'immagine dell'omomorfismo* $f_{\mathbf{A}}$*:* $\mathbf{b} \in \operatorname{Im} f_{\mathbf{A}}$*;*
3. *il vettore* $\mathbf{b}$ *è combinazione lineare delle colonne di* $\mathbf{A}$*;*
4. *il rango della matrice dei coefficienti è pari al rango della matrice aumentata:* $\operatorname{rk}\mathbf{A} = \operatorname{rk}[\mathbf{A}|\mathbf{b}]$.

Dimostrazione. [1. ⇔ 2.] Se il sistema ammette soluzioni, allora esiste un vettore $\mathbf{x}$ tale per cui $\mathbf{Ax} = f_{\mathbf{A}}(\mathbf{x}) = \mathbf{b}$, da cui $\mathbf{b} \in \operatorname{Im} f_{\mathbf{A}}$. Viceversa, se $\mathbf{b} \in \operatorname{Im} f_{\mathbf{A}}$, allora esiste un $\mathbf{x}$ tale per cui $\mathbf{b} = f_{\mathbf{A}}(\mathbf{x}) = \mathbf{Ax}$.
[2. ⇔ 3.] Basti ricordare che il sottospazio vettoriale $\operatorname{Im} f_{\mathbf{A}}$ è generato dalle colonne di $\mathbf{A}$.
[3. ⇔ 4.] Ricordando che il rango di una matrice è pari al numero di colonne linearmente indipendenti della matrice stessa, abbiamo che se $\mathbf{b}$ è combinazione lineare delle colonne di $\mathbf{A}$, aggiungere la colonna $\mathbf{b}$ alla matrice (ottenendo quindi la matrice aumentata $[\mathbf{A}|\mathbf{b}]$) non ne modifica il rango. Viceversa, se il rango è lo stesso, allora l'aggiunta di $\mathbf{b}$ alla matrice $\mathbf{A}$ non ne ha modificato il rango, quindi il numero di colonne indipendenti di $\mathbf{A}$ è lo stesso di $[\mathbf{A}|\mathbf{b}]$ da cui segue che $\mathbf{b}$ è una combinazione lineare di colonne di $\mathbf{A}$. □

Vediamo ora un fondamentale risultato che permette di dire se un sistema lineare ammette soluzioni basandosi esclusivamente sui ranghi della matrice dei coefficienti $\mathbf{A}$ e della matrice aumentata $[\mathbf{A}|\mathbf{b}]$.

Teorema 3.1 (Rouché - Capelli)**.** *Sia* $\mathbf{Ax} = \mathbf{b}$ *un sistema lineare di m equazioni ed n incognite. Allora, il sistema ammette soluzioni se e solo se* $\operatorname{rk}\mathbf{A} = \operatorname{rk}[\mathbf{A}|\mathbf{b}]$. *Inoltre, in tal caso abbiamo che*

1. *se* $\operatorname{rk}\mathbf{A} = \operatorname{rk}[\mathbf{A}|\mathbf{b}] = n$ *allora la soluzione è unica;*

2. *se* $\operatorname{rk}\mathbf{A} = \operatorname{rk}[\mathbf{A}|\mathbf{b}] < n$ *allora il sistema ammette soluzioni, le quali dipendono da* $n - \operatorname{rk}\mathbf{A} = n - \operatorname{rk}[\mathbf{A}|\mathbf{b}]$ *parametri liberi*

Dimostrazione. Il fatto che i ranghi della matrice dei coefficienti e della matrice aumentata debbano coincidere per avere una soluzione al sistema lineare segue dall'equivalenza dei punti 1. e 4. della Proposizione 3.3.
Notiamo ora che l'insieme $S_{\mathbf{0}}$ delle soluzioni al problema omogeneo $\mathbf{Ax} = \mathbf{0}$ coincide col nucleo della mappa $f_{\mathbf{A}} : \mathbf{x} \mapsto \mathbf{Ax}$: $S_{\mathbf{0}} = \ker f_{\mathbf{A}}$. Denotiamo ora $r \doteq \operatorname{rk}(\mathbf{A})$. Dal Teorema di nullità + rango 2.1 segue che $\dim S_{\mathbf{0}} = n - r$. Quindi, se $n = r$ segue che lo spazio nullo cosiste del solo vettore nullo, $S_{\mathbf{0}} = \ker f_{\mathbf{A}} = \{\mathbf{0}\}$, e, dalla proposizione 3.2 segue che il sistema possiede solo la soluzione particolare $\tilde{\mathbf{x}}$.
Se invece $r < n$, il sottospazio $S_{\mathbf{0}}$ ha dimensione positiva $n - r$. Quindi, gli elementi di $S_{\mathbf{0}}$ pososno essere descritti come combinazione lineare di $n - r$ vettori di base. A questi $n - r$ vettori di base possono essere associati altrettanti coefficienti. □

Esempio 3.1. Si consideri il sistema lineare

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

la cui matrice aumentata è

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 2 & 1 & 0 & 1 & 1 \\ 1 & 1 & -1 & 1 & 0 \\ 0 & 1 & -2 & 1 & -1 \end{array}\right].$$

Per questo sistema lineare, si ha che $\operatorname{rk}\mathbf{A} = \operatorname{rk}[\mathbf{A}|\mathbf{b}] = 2$. L'insieme $S_\mathbf{0}$ ha quindi dimensione delle soluzioni al sistema omogeneo ha quindi dimesnione $4-2=2$ ed è lo spazio vettoriale

$$S_\mathbf{0} = \ker\mathbf{A} = \left\{ \begin{bmatrix} \alpha \\ \beta \\ -\alpha \\ -2\alpha-\beta \end{bmatrix} : \alpha, \beta \in \mathbb{R} \right\} = \left\langle \begin{bmatrix} 1 \\ 0 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix} \right\rangle.$$

Una soluzione particolare al sistema è il vettore $\mathbf{x}' = [0,0,1,1]^\top$. Da cui l'insieme delle soluzioni è

$$\mathbf{x}' + S_\mathbf{0} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} \alpha \\ \beta \\ -\alpha \\ -2\alpha-\beta \end{bmatrix} : \alpha, \beta \in \mathbb{R} \right\}.$$

3.1 Risoluzione di un Sistema Lineare mediante Eliminazione di Gauss

In questa parte, descriveremo un procedimento, chiamato *eliminazione di Gauss*, che può essere visto come il metodo di riduzione per sistemi lineari applicato alla forma matriciale del sistema lineare. In particolare, il metodo di eleminazione di Gauss, che da ora in poi abbrevieremo in 'EG', parte dall'osservazione che la soluzione del sistema lineare non cambia quando particolari manipolazioni del sistema stesso vengono svolte. Tra queste manipolazioni vi sono, ad esempio, lo scambio di "posizione" di due equazioni, moltiplicare entrambi i termini di un'equazione per una costante non nulla, e sommare ad un'equazione un multiplo di un'altra.
L'idea dell'EG è quindi quella di usare queste operazioni (chiamate *operazioni elementari*) per trasformare un sistema di equazioni lineari in uno equivalente, ma di cui sia più semplice calcolare la soluzione, applicando queste operazioni alla matrice aumetata $[\mathbf{A}, \mathbf{b}]$ del sistema. Notiamo infatti che le operazioni elementari hanno un'operazione corrispettiva nella matrice aumentata. In particolare, lo scambio di posizione tra equazioni coincide con lo scambiare di posto due righe della matrice $[\mathbf{A}|\mathbf{b}]$; la moltiplicazione di entrambi i membri di un'equazione per una costante non nulla equivale col moltiplicare tutti gli elementi di una riga per quella costante; e la somma tra un'equazione e un multiplo di un'altra coincide col sommare ad una riga di $[\mathbf{A}|\mathbf{b}]$ il multiplo di un'altra riga della matrice aumentata.

Vediamo prima il procedimento generico. Successivamente, lo applicheremo ad un esempio per rendere il tutto più chiaro. Il primo passaggio consiste nello scrivere un sistema lineare nella forma data da (10) come una singola matrice aumentata come in (12):

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

Assumiamo che nessuna colonna della matrice $\mathbf{A}$ sia nulla. Se così non fosse, allora esisterebbe una colonna $\bar{j}$ in cui tutti gli elementi sono nulli. Ma questo implicherebbe che, in ogni equazione del sistema, la variabile $x_{\bar{j}}$ è sempre moltiplicata per zero, e quindi non 'contribuisce' al sistema. Di conseguenza, un sistema lineare di questo tipo si potrebbe riscrivere come un sistema di $n-1$ incognite.

A questo punto, supponiamo che $a_{11} \neq 0$. Se così non fosse, scambiamo la riga $i = 1$ con una riga $\bar{\imath}$ in cui $a_{\bar{\imath}1} \neq 0$. La prima cosa che vogliamo fare è 'trasformare' l'elemento a_{11} di modo che abbia valore unitario. Per fare ciò, applichiamo una delle operazioni elementari: il moltiplicare una riga di $[\mathbf{A}|\mathbf{b}]$ per una costante, in particolare per la costante ${}^{1}/_{a_{11}}$.

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right] \xrightarrow{(EG)} \left[\begin{array}{cccc|c} a_{11}/a_{11} & a_{12}/a_{11} & \cdots & a_{1n}/a_{11} & b_1/a_{11} \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

$$= \left[\begin{array}{cccc|c} 1 & a'_{12} & \cdots & a'_{1n} & b'_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right].$$

Il prossimo passaggio consiste nel rendere nulle, tramite operazioni elementari, tutte le componeti della prima colonna, ad eccezione della prima. Per fare ciò, sottraiamo ad ogni riga i la riga 1 moltiplicata per la costante a_{i1}. Ovviamente, se l'elemento in riga i e colonna 1 è già nullo, si salta questo passaggio per la suddetta riga:

$$\left[\begin{array}{cccc|c} 1 & a'_{12} & \cdots & a'_{1n} & b'_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

$$\xrightarrow{(EG)} \left[\begin{array}{cccc|c} 1 & a'_{12} & \cdots & a'_{1n} & b'_1 \\ a_{21} - a_{21} & a_{22} - a_{21}a'_{12} & \cdots & a_{2n} - a_{21}a'_{1n} & b_2 - a_{21}b'_1 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} - a_{m1} & a_{m2} - a_{m1}a'_{12} & \cdots & a_{mn} - a_{m1}a'_{1n} & b_m - a_{m1}b'_1 \end{array}\right]$$

$$= \left[\begin{array}{cccc|c} 1 & a'_{12} & \cdots & a'_{1n} & b'_1 \\ 0 & a'_{22} & \cdots & a'_{2n} & b'_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a'_{m2} & \cdots & a'_{mn} & b'_m \end{array}\right].$$

Arrivati a questo punto, la prima incognita appare solo nella prima equazione e possiamo considerare il 'sottosistema' la cui matrice aumentata è:

$$[\mathbf{A}'|\mathbf{b}'] = \left[\begin{array}{cccc|c} a'_{22} & a'_{23} & \cdots & a'_{2n} & b'_2 \\ a'_{32} & a'_{33} & \cdots & a'_{3n} & b'_3 \\ \vdots & \vdots & & \vdots & \vdots \\ a'_{m2} & a'_{m3} & \cdots & a'_{mn} & b'_m \end{array}\right].$$

In modo analogo a quanto fatto prima, dividiamo la prima riga per a'_{22}, ottenendo

$$\left[\begin{array}{cccc|c} a'_{22} & a'_{23} & \cdots & a'_{2n} & b'_2 \\ a'_{32} & a'_{33} & \cdots & a'_{3n} & b'_3 \\ \vdots & \vdots & & \vdots & \vdots \\ a'_{m2} & a'_{m3} & \cdots & a'_{mn} & b'_m \end{array}\right] \xrightarrow{(EG)} \left[\begin{array}{cccc|c} 1 & a''_{23} & \cdots & a''_{2n} & b''_2 \\ a'_{32} & a'_{33} & \cdots & a'_{3n} & b'_3 \\ \vdots & \vdots & & \vdots & \vdots \\ a'_{m2} & a'_{m3} & \cdots & a'_{mn} & b'_m \end{array}\right],$$

e ad ogni riga i per cui $a_{i2} \neq 0$ sottraiamo a_{i2} volte la riga $i = 2$ (la prima del sottosistema):

$$\left[\begin{array}{cccc|c} 1 & a''_{23} & \cdots & a''_{2n} & b''_2 \\ a'_{32} & a'_{33} & \cdots & a'_{3n} & b'_3 \\ \vdots & \vdots & & \vdots & \vdots \\ a'_{m2} & a'_{m3} & \cdots & a'_{mn} & b'_m \end{array}\right] \xrightarrow{(EG)} \left[\begin{array}{cccc|c} 1 & a''_{23} & \cdots & a''_{2n} & b''_2 \\ 0 & a''_{33} & \cdots & a''_{3n} & b''_3 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a''_{m3} & \cdots & a''_{mn} & b''_m \end{array}\right].$$

A questo punto, si riapplicano gli stessi passaggii al sottosistema

$$\left[\begin{array}{ccc|c} a''_{33} & \cdots & a''_{3n} & b''_3 \\ \vdots & & \vdots & \vdots \\ a''_{m3} & \cdots & a''_{mn} & b''_m \end{array}\right].$$

Continuando a iterare in questo modo, si ottiene una matrice "a scala", ovvero in cui, in ciascuna riga, il primo coefficiente non nullo si trova a destra del primo coefficiente non nullo della riga precedente. Poiché questo nuovo sistema lineare ha la stessa soluzione di quello originale, possiamo calcolare tale soluzione tramite sostituzione all'indietro, partendo dall'ultima riga e risalendo lungo la matrice.

Quando si svolgono le operazioni della EG, sarebbe buona norma specificare quale operazione si sta svolgendo. Quindi, si può spiegare nel testo tra un passaggio e l'altro quale operazione si sta svolgendo. Altrimenti, si può 'schematizzare' l'operazione direttamente nei passaggi. In questo modo, la moltiplicazione della riga i per una costante $k \neq 0$ può essere scritta come

$$\mathbf{A} \xrightarrow[r_i \to k r_i]{(EG)} \mathbf{A}'.$$

Sommare alla riga i $k-$volte la riga j può essere denotata come

$$\mathbf{A} \xrightarrow[r_i \to r_i + k r_j]{(EG)} \mathbf{A}'.$$

Infine, sostituire di posto due righe i e j può essere scritto come

$$\mathbf{A} \xrightarrow[r_i \leftrightarrow r_j]{(EG)} \mathbf{A}'.$$

Rimarchiamo il fatto che questa è solo una notazione per rendere chiari i passaggi svolti.

Esempio 3.2. Consideriamo il sistema lineare

$$\begin{cases} 2x_1 - 2x_2 = -2 \\ 2x_1 - 2x_2 + x_3 = 2 \\ x_1 + x_2 - x_3 = 1 \end{cases}$$

La matrice aumentata associata è

$$\left[\begin{array}{ccc|c} 2 & -2 & 0 & -2 \\ 2 & -2 & 1 & 2 \\ 1 & 1 & -1 & 1 \end{array}\right].$$

Il primo passaggio consiste nel dividere la prima riga per due, di modo da ottenere l'elemento in posizione $(1,1)$ pari ad 1. Otteniamo quindi la matrice

$$\left[\begin{array}{ccc|c} 2 & -2 & 0 & -2 \\ 2 & -2 & 1 & 2 \\ 1 & 1 & -1 & 1 \end{array}\right] \xrightarrow[r_1 \to \frac{r_1}{2}]{(EG)} \left[\begin{array}{ccc|c} 1 & -1 & 0 & -1 \\ 2 & -2 & 1 & 2 \\ 1 & 1 & -1 & 1 \end{array}\right].$$

Il prossimo passaggio consiste nel rendere gli elementi in posizione $(2,1)$ ed $(3,1)$ pari a zero. Per fare ciò, sottraiamo alla seconda riga due volte la prima; ed alla terza riga sottraiamo la prima:

$$\left[\begin{array}{ccc|c} 1 & -1 & 0 & -1 \\ 2 & -2 & 1 & 2 \\ 1 & 1 & -1 & 1 \end{array}\right] \xrightarrow[r_2 \to r_2 - 2r_1]{(EG)} \left[\begin{array}{ccc|c} 1 & -1 & 0 & -1 \\ 0 & 0 & 1 & 4 \\ 1 & 1 & -1 & 1 \end{array}\right] \xrightarrow[r_3 \to r_3 - r_1]{(EG)} \left[\begin{array}{ccc|c} 1 & -1 & 0 & -1 \\ 0 & 0 & 1 & 4 \\ 0 & 2 & -1 & 2 \end{array}\right].$$

A questo punto, la componente in $(2,2)$ è nulla. Scambiamo quindi la seconda riga con la terza:

$$\left[\begin{array}{ccc|c} 1 & -1 & 0 & -1 \\ 0 & 0 & 1 & 4 \\ 0 & 2 & -1 & 2 \end{array}\right] \xrightarrow[r_2 \leftrightarrow r_3]{(EG)} \left[\begin{array}{ccc|c} 1 & -1 & 0 & -1 \\ 0 & 2 & -1 & 2 \\ 0 & 0 & 1 & 4 \end{array}\right],$$

e dividiamo la (nuova) seconda riga per due:

$$\left[\begin{array}{ccc|c} 1 & -1 & 0 & -1 \\ 0 & 2 & -1 & 2 \\ 0 & 0 & 1 & 4 \end{array}\right] \xrightarrow[r_2 \to \frac{r_2}{2}]{(EG)} \left[\begin{array}{ccc|c} 1 & -1 & 0 & -1 \\ 0 & 1 & -1/2 & 1 \\ 0 & 0 & 1 & 4 \end{array}\right].$$

Questa matrice aumentata corrisponde al sistema lineare

$$\begin{cases} x_1 - x_2 = -1 \\ x_2 - \frac{1}{2}x_3 = 1 \\ x_3 = 4 \end{cases}$$

Per trovare la soluzione, partiamo dall'ultima equazione che restituisce

$$x_3 = 4.$$

Sostituendo il valore di x_3 nella seconda equazione otteniamo il valore per l'incognita x_2:

$$x_2 - \frac{1}{2}x_3 = 1 \quad \Leftrightarrow \quad x_2 - 2 = 1 \quad \Leftrightarrow \quad x_2 = 3$$

Infine, usiamo i valori trovati per x_2 per calcolare x_1:

$$x_1 - x_2 = -1 \quad \Leftrightarrow \quad x_1 - 3 = -1 \quad \Leftrightarrow \quad x_1 = 2.$$

Esempio 3.3. Si consideri il sistema lineare

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ 2x_1 + x_2 + x_3 = 2 \\ 3x_1 + 3x_2 + 4x_3 = 3 \end{cases}$$

la cui matrice aumentata associata è

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 2 & 1 & 1 & 2 \\ 3 & 3 & 4 & 3 \end{array}\right].$$

La componente in $(1,1)$ è già pari ad uno. Quindi, i primi due passaggi da svolgere consistono nel porre a zero la seconda e terza riga. Per fare ciò, sottraiamo alla seconda riga 2 volte la prima, e sottraiamo alla terza riga 3 volte la prima:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 2 & 1 & 1 & 2 \\ 3 & 3 & 4 & 3 \end{array}\right] \xrightarrow[\substack{r_2 \to r_2 - 2r_1 \\ r_3 \to r_3 - 3r_1}]{(EG)} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -5 & 0 \\ 0 & -3 & -5 & 0 \end{array}\right].$$

A questo punto, la seconda e terza riga coincidono. Quindi, possiamo sottrarre alla terza riga la seconda:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -5 & 0 \\ 0 & -3 & -5 & 0 \end{array}\right] \xrightarrow[r_3 \to r_3 - r_2]{(EG)} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

A questo punto, abbiamo che il valore di x_3 è libero in quanto la terza riga del sistema è nulla, $x_3 = \ell$. Dalla seconda equazione si ottiene invece

$$-3x_2 - 5x_3 = 0 \quad \Leftrightarrow \quad x_2 = -\frac{5}{3}x_3 \quad \Leftrightarrow \quad x_2 = -\frac{5}{3}\ell.$$

Infine, la prima equazione implica

$$x_1 + 2x_2 + 3x_3 = 1 \quad \Leftrightarrow \quad x_1 - 2\frac{5}{3}\ell + 3\ell = 1 \quad \Leftrightarrow \quad x_1 = 1 + \frac{10}{3}\ell - 3\ell \quad \Leftrightarrow \quad x_1 = 1 - \frac{1}{3}\ell.$$

Quindi, abbiamo infinite soluzioni al sistema lineare che dipendono dal parametro $\ell \in \mathbb{R}$:

$$\begin{cases} x_1 = 1 - \frac{\ell}{3} \\ x_2 = -\frac{5\ell}{3} \\ x_3 = \ell \end{cases}$$

Vediamo ora un esempio applicato ad un sistema sottodeterminato

Esempio 3.4. Si consideri il sistema a due equazioni e tre incognite

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 - x_2 - x_3 = 2 \end{cases}$$

la cui matrice aumentata è

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 2 \end{array}\right].$$

Applichiamo EG:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 2 \end{array}\right] \xrightarrow[r_2 \to r_2 - r_1]{(EG)} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & 1 \end{array}\right] \xrightarrow[r_2 \to \frac{r_2}{-2}]{(EG)} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & -\frac{1}{2} \end{array}\right]$$

da cui otteniamo che il valore di x_3 è libero, $x_3 = \lambda$, e la soluzione

$$\begin{cases} x_1 = \frac{3}{2} \\ x_2 = -\frac{1}{2} - \lambda \\ x_3 = \lambda \end{cases}$$

Infine, vediamo cosa accade se applichiamo EG ad un sistema che non ammette soluzioni

Esempio 3.5. Si consideri il sistema

$$\begin{cases} x_1 + x_2 = 1 \\ x_2 - x_3 = 2 \\ x_1 + x_3 = 1 \end{cases}$$

la cui matrice aumentata è

$$[\mathbf{A}|\mathbf{b}] = \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & -1 & 2 \\ 1 & 0 & 1 & 1 \end{array}\right].$$

Applichiamo EG

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & -1 & 2 \\ 1 & 0 & 1 & 1 \end{array}\right] \xrightarrow[r_3 \to r_3 - r_1]{(EG)} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & -1 & 1 & 0 \end{array}\right] \xrightarrow[r_3 \to r_3 + r_2]{(EG)} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 2 \end{array}\right].$$

Se ora riscrivessimo questa matrice aumetata come sistema, otterremo

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

che non ha soluzione.

3.2 Calcolo del Rango di una Matrice Tramite EG

In questa parte, vedremo come usare il processo dell'eliminazione di Gauss per calcolare il rango di una matrice. Definiamo quindi una matrice $\mathbf{A} \in \mathscr{M}_{m,n}(\mathbb{K})$ di dimensione $m \times n$ sul campo $\mathbb{K}$. Ricordiamo

che il rango della matrice $\mathbf{A}$, rk$\mathbf{A}$, è il massimo numero di colonne linearmente indipendenti di $\mathbf{A}$ (proposizion 2.11), e coincide con la dimensione dell'immagine della funzione lineare $f_{\mathbf{A}} : \mathbb{K}^n \to \mathbb{K}^m$ la cui matrice associata rispetto alle basi canoniche di $\mathbb{K}^n$ e $\mathbb{K}^m$ è $\mathbf{A}$.
Ricordiamo inoltre che, in modo simile, si può definire il *rango per righe* di una matrice $\mathbf{A}$ come il massimo numero di righe linearmente indipendenti di $\mathbf{A}$. Inoltre, ricordiamo anche che grazie al corollario 2.3, il rango per righe ed il rango per colonne coincidono.

Vediamo come usare l'EG per caolcolare il rango (per righe) di una matrice $\mathbf{A}$. Iniziamo ricordando che l'EG non altera il numero di righe linearmente indipendenti di $\mathbf{A}$ in quanto tutte le operazioni dell'EG sostituiscono una riga con una combinazione lineare delle stesse righe di $\mathbf{A}$. Di conseguenza, se tramite EG trasformiamo una matrice $\mathbf{A}$ in una matrice con forma a scala $\mathbf{A}'$, abbiamo che il rango di $\mathbf{A}'$ coincide col rango di $\mathbf{A}$.
Questo risultato è particolarmente importante in quanto da una matrice con forma a scala si può facilmente calcolarne il rango usando il seguente risultato.

Proposizione 3.4. *Data una matrice* $\mathbf{A}' \in \mathcal{M}_{m,n}(\mathbb{K})$ *con forma a scala, il suo rango coincide con il numero di righe non nulle.*

◇ *Dimostrazione.* Supponiamo che la matrice $\mathbf{A}$ abbia le prime $r \leq m$ righe non nulle, ed indichiamo con $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ i vettori riga non nulli di $\mathbf{A}$, con $\mathbf{v}_i = [a_{i\,1}, a_{i\,2}, \ldots, a_{i\,n}]$:

$$\mathbf{A} = \begin{bmatrix} a_{1\,1} & a_{1\,2} & \cdots & a_{1\,n} \\ a_{2\,1} & a_{2\,2} & \cdots & a_{2\,n} \\ \vdots & \vdots & & \vdots \\ a_{r\,1} & a_{r\,2} & \cdots & a_{r\,n} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Poiché, per ipotesi, la matrice $\mathbf{A}$ è a scala, il primo coefficiente non nullo di ciascuna riga si trova a destra del primo coefficiente non nullo della riga $i-1$. Quindi, denotiamo con j_1 l'indice di colonna della prima componente non nulla della prima riga di $\mathbf{A}$. Per ipotesi, tutte le altre componenti della colonna j_1 sono nulle. Di conseguenza, non esiste nessuna combinazione lineare delle altre righe che possa risultare nella riga $\mathbf{v}_1$, in quanto nessuna altra riga $\mathbf{v}_2, \mathbf{v}_3 \ldots, \mathbf{v}_r$ può contribuire alla componente in colonna j_1.
In modo analogo, consideriamo una generica riga $\bar{\imath}$, e sia $j_{\bar{\imath}}$ l'indice di colonna della prima componente non nulla della riga $\bar{\imath}$. Questa riga non può essere scritta come combinazione lineare delle altre righe. Infatti, tutte le righe precedenti, se non moltiplicate per zero nella combinazione lineare, restituirebbero un valore non nullo in una posizione $j < j_{\bar{\imath}}$. Ma, poiché l'indice $j_{\bar{\imath}}$ è quello del primo elemento non nullo questo non è possibile. Inoltre, il vettore $\mathbf{v}_{\bar{\imath}}$ non può essere scritto come combinazione lineare delle componenti successive (ovvero con indice di riga superiore a $\bar{\imath}$), con un ragionamento analogo a quanto fatto per la prima riga.

Di conseguenza, l'insieme delle r righe non nulle di $\mathbf{A}$ sono linearmente indipendenti. Quindi, il rango di $\mathbf{A}$ è esattamente r.

□

Riassumendo, data una generica matrice $\mathbf{A}$ possiamo calcolarne il rango come segue. Iniziamo applicando EG trasformando la matrice $\mathbf{A}'$ in una matrice con forma a scala $\mathbf{A}'$. Fatto ciò, è sufficiente contare il numero di righe non nulle di $\mathbf{A}'$ per ottenere il rango.

Esempio 3.6. Si consideri la matrice $\mathbf{A} \in M_{5,6}(\mathbb{K})$ definita come

$$\mathbf{A} = \begin{bmatrix} 2 & -2 & 4 & 6 & 4 & 2 \\ 3 & 2 & 1 & 0 & 1 & 3 \\ 1 & 4 & -3 & -6 & -3 & 1 \\ 0 & 2 & 1 & 2 & -1 & 0 \\ 4 & 4 & -1 & -8 & -1 & 4 \end{bmatrix}.$$

Come primo passaggio, dividiamo la prima riga per 2, di modo da ottenere un 1 in posizione $(1,1)$:

$$\begin{bmatrix} 2 & -2 & 4 & 6 & 4 & 2 \\ 3 & 2 & 1 & 0 & 1 & 3 \\ 1 & 4 & -3 & -6 & -3 & 1 \\ 0 & 2 & 1 & 2 & -1 & 0 \\ 4 & 4 & -1 & -8 & -1 & 4 \end{bmatrix} \xrightarrow[r_1 \to r_1/2]{(EG)} \begin{bmatrix} 1 & -1 & 2 & 3 & 2 & 1 \\ 3 & 2 & 1 & 0 & 1 & 3 \\ 1 & 4 & -3 & -6 & -3 & 1 \\ 0 & 2 & 1 & 2 & -1 & 0 \\ 4 & 4 & -1 & -8 & -1 & 4 \end{bmatrix}.$$

Poi, poniamo le prime componenti delle altre righe i pari a 0 sottrendo ad ognuna a_{i1} volte la prima riga (questo passaggio non è necessario per la quarta riga in quanto la componente in $(4,1)$ è già nulla):

$$\begin{bmatrix} 1 & -1 & 2 & 3 & 2 & 1 \\ 3 & 2 & 1 & 0 & 1 & 3 \\ 1 & 4 & -3 & -6 & -3 & 1 \\ 0 & 2 & 1 & 2 & -1 & 0 \\ 4 & 4 & -1 & -8 & -1 & 4 \end{bmatrix} \xrightarrow[\substack{r_2 \to r_2 - 3r_1 \\ r_3 \to r_3 - r_1 \\ r_5 \to r_5 - 4r_1}]{(EG)} \begin{bmatrix} 1 & -1 & 2 & 3 & 2 & 1 \\ 0 & 5 & -5 & -9 & -5 & 0 \\ 0 & 5 & -5 & -9 & -5 & 0 \\ 0 & 2 & 1 & 2 & -1 & 0 \\ 0 & 8 & -9 & -20 & -9 & 0 \end{bmatrix}.$$

A questo punto, poniamo la componente in $(2,2)$ pari ad 1 dividento per $a_{22} = 5$

$$\begin{bmatrix} 1 & -1 & 2 & 3 & 2 & 1 \\ 0 & 5 & -5 & -9 & -5 & 0 \\ 0 & 5 & -5 & -9 & -5 & 0 \\ 0 & 2 & 1 & 2 & -1 & 0 \\ 0 & 8 & -9 & -20 & -9 & 0 \end{bmatrix} \xrightarrow[r_2 \to r_2/5]{(EG)} \begin{bmatrix} 1 & -1 & 2 & 3 & 2 & 1 \\ 0 & 1 & -1 & -9/5 & -1 & 0 \\ 0 & 5 & -5 & -9 & -5 & 0 \\ 0 & 2 & 1 & 2 & -1 & 0 \\ 0 & 8 & -9 & -20 & -9 & 0 \end{bmatrix}.$$

Ora, poniamo a zero le componenti successive della seconda colonna:

$$\begin{bmatrix} 1 & -1 & 2 & 3 & 2 & 1 \\ 0 & 1 & -1 & -9/5 & -1 & 0 \\ 0 & 5 & -5 & -9 & -5 & 0 \\ 0 & 2 & 1 & 2 & -1 & 0 \\ 0 & 8 & -9 & -20 & -9 & 0 \end{bmatrix} \xrightarrow[\substack{r_3 \to r_3 - 5r_2 \\ r_4 \to r_4 - 2r_2 \\ r_5 \to r_5 - 8r_2}]{(EG)} \begin{bmatrix} 1 & -1 & 2 & 3 & 2 & 1 \\ 0 & 1 & -1 & -9/5 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 28/5 & 1 & 0 \\ 0 & 0 & 1 & -28/5 & 17 & 0 \end{bmatrix}.$$

La terza riga è nulla, quindi la sostituiamo con l'ultima:

$$\begin{bmatrix} 1 & -1 & 2 & 3 & 2 & 1 \\ 0 & 1 & -1 & -9/5 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 28/5 & 1 & 0 \\ 0 & 0 & 1 & -28/5 & 17 & 0 \end{bmatrix} \xrightarrow[r_3 \leftrightarrow r_5]{(EG)} \begin{bmatrix} 1 & -1 & 2 & 3 & 2 & 1 \\ 0 & 1 & -1 & -9/5 & -1 & 0 \\ 0 & 0 & 1 & -28/5 & 17 & 0 \\ 0 & 0 & 3 & 28/5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Ora, azzeriamo la componente in $(4,3)$:

$$\begin{bmatrix} 1 & -1 & 2 & 3 & 2 & 1 \\ 0 & 1 & -1 & -9/5 & -1 & 0 \\ 0 & 0 & 1 & -28/5 & 17 & 0 \\ 0 & 0 & 3 & 28/5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow[r_4 \to r_4 - 3r_3]{(EG)} \begin{bmatrix} 1 & -1 & 2 & 3 & 2 & 1 \\ 0 & 1 & -1 & -9/5 & -1 & 0 \\ 0 & 0 & 1 & -28/5 & 17 & 0 \\ 0 & 0 & 0 & 112/5 & -50 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Infine, concludiamo ponendo ad 1 la componente in posizione $(4,4)$:

$$\begin{bmatrix} 1 & -1 & 2 & 3 & 2 & 1 \\ 0 & 1 & -1 & -9/5 & -1 & 0 \\ 0 & 0 & 1 & -28/5 & 17 & 0 \\ 0 & 0 & 0 & 112/5 & -50 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow[r_4 \to r_4\, 5/112]{(EG)} \begin{bmatrix} 1 & -1 & 2 & 3 & 2 & 1 \\ 0 & 1 & -1 & -9/5 & -1 & 0 \\ 0 & 0 & 1 & -28/5 & 17 & 0 \\ 0 & 0 & 0 & 1 & -250/112 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

La matrice ottenuta ha forma a scala, ed ha quattro righe non nulle. Quindi, il rango di $\mathbf{A}$ è pari a 4: $\text{rk}(\mathbf{A}) = 4$.

3.3 Calcolo della Matrice Inversa tramite EG

In questa parte vedremo un procedimento che utilizza l'Eliminazione di Gauss per calcolare l'inversa $\mathbf{A}^{-1}$ di una matrice quadrata $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$.
Tale procedimento si basa sul fatto che effettuare un'operazione lineare su una matrice equivale a moltiplicare quella stessa matrice, a sinistra, per un'opportuna matrice invertibile.

3.3.1 Matrici Associate alle Operazioni Elementari

Come prima cosa, quindi, occorre chiarire chi siano queste meatrici invertibili che descrivono le operazioni elementari. Iniziamo con lo scambio di due righe i e j della matrice $\mathbf{A}$. La maatrice che descrive questa operazione elementare è la matrice $\mathbf{P}(i,j) = [\rho_{hk}]_{h,k=1,2,\dots,n}$ le cui componenti sono

$$\rho_{hk} = \begin{cases} 1 & \text{se } h = k, h \neq i, h \neq j \\ 1 & \text{se } h = i, k = j \text{ o } h = j, k = i \\ 0 & \text{altrimenti} \end{cases} \tag{13}$$

In pratica, la matrice $\mathbf{P}(i,j)$ è la matrice identità $\mathbf{Id}_n$ in cui la riga i e la riga j sono scate scambiate.

Esempio 3.7. Ad esempio, sia $\mathbf{A}$ una generica matrice 4×4, e consideriamo il prodotto, a sinistra, con la matrice $\mathbf{P}(1,3)$:

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} a_{31} & a_{32} & a_{33} & a_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{11} & a_{12} & a_{13} & a_{14} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

Osservazione 3.2. Poiché la matrice $\mathbf{P}(i,j)$ è la matrice identità con le righe i e j scambiate, abbiamo che il quadrato della matrice coincide con l'identità (in quanto le righe i e j vengono scambiate due volte): $\mathbf{P}^2(i,j) = \mathbf{P}(i,j)\mathbf{P}(i,j) = \mathbf{Id}$. Quindi, la matrice $\mathbf{P}(i,j)$ è invertibile e l'inversa coincide con la matrice stessa:

$$\mathbf{P}^{-1}(i,j) = \mathbf{P}(i,j).$$

La seconda operazione elementare che andiamo a considerare è la moltiplicazione di una riga i di $\mathbf{A}$ per una costante λ. La matrice che descrive questa operazione è la matrice $\mathbf{M}(i,\lambda) = [\mu_{hk}]_{h,k=1,2,\dots,n}$ le cui componenti sono:

$$\mu_{hk} = \begin{cases} 1 & \text{se } h = k, h \neq i \\ \lambda & \text{se } h = k = i \\ 0 & \text{altrimenti} \end{cases} \tag{14}$$

In pratica, la matrice $\mathbf{M}(i,\lambda)$ è la matrice identità in cui la riga i è moltiplicata per λ.

Esempio 3.8. Ad esempio, sia $\mathbf{A}$ una generica matrice 4×4, e consideriamo il prodotto, a sinistra, con la matrice $\mathbf{M}(2,\lambda)$:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ \lambda a_{21} & \lambda a_{22} & \lambda a_{23} & \lambda a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

Osservazione 3.3. L'inverso di una matrice $\mathbf{M}(i,\lambda)$ è la matrice $\mathbf{M}(i,\lambda^{-1})$:

$$\mathbf{M}^{-1}(i,\lambda) = \mathbf{M}(i,\lambda^{-1}).$$

Infatti, la prima matrice moltiplica la $i-$sima riga di $\mathbf{Id}$ per una costante, e la seconda moltiplica la stessa riga per l'inverso di quella stessa costante.

L'ultima operazione elementare consiste nel sommare ad una riga i λ volte la riga j. La matrice associata a questa operazione è la matrice $\mathbf{S}(i,j;\lambda) = [\sigma_{hk}]_{h,k=1,2,\dots,n}$ le cui componenti sono

$$\sigma_{hk} = \begin{cases} 1 & \text{se } h = k \\ \lambda & \text{se } h = i, k = j \\ 0 & \text{altrimenti} \end{cases} \tag{15}$$

Il altre parole, la matrice $\mathbf{S}(i,j;\lambda)$ è la matrice identità in cui lo zero in posizione (i,j) è sostituito dalla costante λ

Esempio 3.9. Ad esempio, sia $\mathbf{A}$ una generica matrice 4×4, e consideriamo il prodotto, a sinistra, con la matrice $\mathbf{S}(2,3;\lambda)$:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \lambda & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} + \lambda a_{31} & a_{22} + \lambda a_{32} & a_{23} + \lambda a_{33} & a_{24} + \lambda a_{34} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

Osservazione 3.4. La matrice inversa a $\mathbf{S}(i,j;\lambda)$ è $\mathbf{S}(i,j;-\lambda)$. Infatti se alla riga i prima sommiamo λ volte la riga j e poi sottraiamo la stessa quantità, otteniamo la matrice di partenza.

$$\mathbf{S}^{-1}(i,j,\lambda) = \mathbf{S}(i,j-\lambda).$$

Abbiamo quindi che una successione di operazioni elementari su una matrice **A** può essere scritta in modo equivalente come una successione di moltiplicazioni a sinistra per delle matrici invertibili come quelle che abbiamo appena definito. Poiché il prodotto di matrici invertibili è ancora invertibile, abbiamo che l'effetto di una successione di operazioni elementari su **A** può essere scritto come il prodotto, a sinistra, di **A** per un'apposita matrice invertibile.

Osservazione 3.5. Se, anzichè moltiplicare a sinistra, moltiplicassimo a destra, otterremmo le stesse operazioni ma applicate alle colonne di **A** anzichè alle righe. Quindi, la matrice $\mathbf{AP}(i,j)$ coincide con la matrice **A** in cui la colonna i e j sono scambiate. Similmente, La matrice $\mathbf{AM}(i,\lambda)$ coincide con la matrice **A** in cui la $i-$sima colonna è moltiplicata per λ. Infine, la matrice $\mathbf{AS}(i,j;\lambda)$ consiste nella matrice **A** in cui alla colonna i è stata sommata λ volte la colonna j.

Vediamo ora come usare quese operazioni, e le relativi matrici, per calcolare l'inversa di una matrice quadrata **A**.

3.3.2 L'Algoritmo

Consideriamo ora una matrice quadrata $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$ e supponiamo sia invertibile. Tramite operazioni elementari, si può trasformare questa matrice in una matrice con forma a scala $\mathbf{A}'$ in cui il primo coefficiente non nullo di ogni riga è pari ad 1. Poiché la matrice **A** è invertibile, il suo rango è pari ad n. Di conseguenza, nella forma a scala non ci sono righe nulle. La matrice $\mathbf{A}'$ ottenuta ha una forma chiamata "triangolare superiore" in cui le componenti sulla diagonale sono tutte pari ad 1:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} \xrightarrow{(EG)} \begin{bmatrix} 1 & a'_{12} & a'_{13} & \cdots & a'_{1n} \\ 0 & 1 & a'_{23} & \cdots & a'_{2n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & a'_{n-1\,n} \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix}.$$

A questo punto è possibile, tramite opportune operazioni elementari, trasformare la matrice $\mathbf{A}'$ nella matrice identità:

$$\begin{bmatrix} 1 & a'_{12} & a'_{13} & \cdots & a'_{1n} \\ 0 & 1 & a'_{23} & \cdots & a'_{2n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & a'_{n-1\,n} \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \xrightarrow{(EG)} \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix}.$$

Abbiamo quindi che tramite operazioni elementari, possiamo trasformare una matrice quadrata invertibile **A** nella matrice identità. Inoltre, poiché la sequenza di operazioni elementari può essere scritta come una moltiplicazione a sinistra per una matrice **B**, abbiamo che questa matrice coincide con l'inversa di **A**: $\mathbf{BA} = \mathbf{Id} \Leftrightarrow \mathbf{B} = \mathbf{A}^{-1}$. Di conseguenza, possiamo calcolare l'inversa di una matrice quadrata invertibile **A** tenendo traccia (vedremo a breve come) delle operazioni elementari svolte per trasformare **A** nella matrice identità.

Un metodo efficace per tenere traccia di queste operazioni è il seguente. Consideriamo una matrice "a blocchi" $[\mathbf{A}|\mathbf{Id}]$ in cui si affiancano la matrice **A** e la matrice identità. Poi, si osservi che il prodotto a sinistra della matrice a blocchi per una matrice **B** è la matrice i cui blocchi sono le singole moltiplicazioni:

$$\mathbf{B}\,[\mathbf{A}|\mathbf{Id}] = [\mathbf{B}\mathbf{A}|\mathbf{B}\,\mathbf{Id}] = [\mathbf{Id}|\mathbf{B}]\,.$$

Questo significa che, una volta trasformata la matrice **A** (nel blocco sinistro) nella matrice identità tramite operazioni elementari, la matrice inversa sarà nel blocco destro.

In conclusione, possiamo quindi calcolare la matrice inversa in questo modo. Iniziamo scrivendo la matrice a blocchi $[\mathbf{A}|\mathbf{Id}]$. Poi, utilizziamo EG per trasformare il blocco a sinistra nella matrice identità. In questo modo, stiamo applicando le stesse operazioni elementari a entrambi i blocchi. Quando il blocco di sinistra coincide con la matrice identità, il blocco a destra sarà la matrice inversa:

$$[\mathbf{A}|\mathbf{Id}] \xrightarrow{(EG)} [\mathbf{Id}|\mathbf{A}^{-1}].$$

Osserviamo che, se **A** non fosse invertibile, allora nessuna sequenza di operazioni elementari può trasformare la matrice **A** nella matrice identica. Quindi, questo algoritmo porta ad avere la matrice identità nel blocco sinistro se e solo se la matrice è invertibile.

Esempio 3.10. Calcoliamo l'inversa della matrice

$$\mathbf{A} = \begin{bmatrix} 2 & 4 & 4 \\ 3 & 4 & 3 \\ 4 & 3 & 1 \end{bmatrix}.$$

Iniziamo scrivendo la matrice a blocchi $[\mathbf{A}|\mathbf{Id}]$:

$$[\mathbf{A}|\mathbf{Id}] = \left[\begin{array}{ccc|ccc} 2 & 4 & 4 & 1 & 0 & 0 \\ 3 & 4 & 3 & 0 & 1 & 0 \\ 4 & 3 & 1 & 0 & 0 & 1 \end{array}\right].$$

Ora, iniziamo ad applicare EG. Il primo obiettivo è quello di porre la matrice **A** in forma a scala. Come primo passaggio, dividiamo per 2 la prima riga:

$$\left[\begin{array}{ccc|ccc} 2 & 4 & 4 & 1 & 0 & 0 \\ 3 & 4 & 3 & 0 & 1 & 0 \\ 4 & 3 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[r_1\to r_1/2]{(EG)} \left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1/2 & 0 & 0 \\ 3 & 4 & 3 & 0 & 1 & 0 \\ 4 & 3 & 1 & 0 & 0 & 1 \end{array}\right].$$

Ora, annulliamo le componenti $(2,1)$ e $(3,1)$ del blocco sinistro:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1/2 & 0 & 0 \\ 3 & 4 & 3 & 0 & 1 & 0 \\ 4 & 3 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[r_2\to r_2-3r_1]{(EG)} \left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1/2 & 0 & 0 \\ 0 & -2 & -3 & -3/2 & 1 & 0 \\ 4 & 3 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[r_3\to r_3-4r_1]{(EG)} \left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1/2 & 0 & 0 \\ 0 & -2 & -3 & -3/2 & 1 & 0 \\ 0 & -5 & -7 & -2 & 0 & 1 \end{array}\right]$$

Ora, poniamo la componente in $(2,2)$ pari ad 1, e successivamente poniamo a 0 la componente in $(3,2)$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1/2 & 0 & 0 \\ 0 & -2 & -3 & -3/2 & 1 & 0 \\ 0 & -5 & -7 & -2 & 0 & 1 \end{array}\right] \xrightarrow[r_2\to -r_2/2]{(EG)} \left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1/2 & 0 & 0 \\ 0 & 1 & 3/2 & 3/4 & -1/2 & 0 \\ 0 & -5 & -7 & -2 & 0 & 1 \end{array}\right] \xrightarrow[r_3\to r_3+5r_2]{(EG)} \left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1/2 & 0 & 0 \\ 0 & 1 & 3/2 & 3/4 & -1/2 & 0 \\ 0 & 0 & 1/2 & 7/4 & -5/2 & 1 \end{array}\right].$$

L'ultimo passaggio per porre il blocco a sinistra in una triangolare superiore con diagonale unitaria consiste nel moltiplicare l'ultima riga per 2:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1/2 & 0 & 0 \\ 0 & 1 & 3/2 & 3/4 & -1/2 & 0 \\ 0 & 0 & 1/2 & 7/4 & -5/2 & 1 \end{array}\right] \xrightarrow[r_3 \to 2r_3]{(EG)} \left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1/2 & 0 & 0 \\ 0 & 1 & 3/2 & 3/4 & -1/2 & 0 \\ 0 & 0 & 1 & 7/2 & -5 & 2 \end{array}\right].$$

A questo punto dobbiamo, tramite EG, annullare le componenti sopra la diagonale di **A**. Iniziamo annullando l'elemento in $(2,3)$:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1/2 & 0 & 0 \\ 0 & 1 & 3/2 & 3/4 & -1/2 & 0 \\ 0 & 0 & 1 & 7/2 & -5 & 2 \end{array}\right] \xrightarrow[r_2 \to r_2 - 3/2 r_3]{(EG)} \left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & -9/2 & 7 & -3 \\ 0 & 0 & 1 & 7/2 & -5 & 2 \end{array}\right]$$

Infine, le componenti $(1,2)$ ed $(1,3)$ vengono annullate sottrraendo alla prima riga 2 volte la terza (per annullare l'elemento in $(1,3)$) e due volte la seconda (per annullare l'elemento in $(2,3)$):

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & -9/2 & 7 & -3 \\ 0 & 0 & 1 & 7/2 & -5 & 2 \end{array}\right] \xrightarrow[r_1 \to r_1 - 2r_3]{(EG)} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -13/2 & 10 & -4 \\ 0 & 1 & 0 & -9/2 & 7 & -3 \\ 0 & 0 & 1 & 7/2 & -5 & 2 \end{array}\right] \xrightarrow[r_1 \to r_1 - 2r_2]{(EG)} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 5/2 & -4 & 2 \\ 0 & 1 & 0 & -9/2 & 7 & -3 \\ 0 & 0 & 1 & 7/2 & -5 & 2 \end{array}\right].$$

Abbiamo quindi ottenuto l'inversa di **A**:

$$\begin{bmatrix} 2 & 4 & 4 \\ 3 & 4 & 3 \\ 4 & 3 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 5/2 & -4 & 2 \\ -9/2 & 7 & -3 \\ 7/2 & -5 & 2 \end{bmatrix}.$$

Esempio 3.11. Vediamo ora cosa succederebbe se la matrice **A non** fosse invertibile. Consideriamo la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \\ 0 & -1 & 1 \end{bmatrix},$$

che risulta nella matrice a blocchi

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right].$$

Iniziamo ad applicare EG per trasformare il blocco di sinistra nella matrice identità

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[r_2 \to r_2 - r_1]{(EG)} \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[r_2 \to -r_2]{(EG)} \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[r_3 \to r_3 + r_2]{(EG)} \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \end{array}\right].$$

Come si può notare, non è possibile utilizzare EG per trasformare il blocco sinistro nella matrice identità.

Osservazione 3.6. Poiché l'inversa di una matrice può essere **sempre** calcolata utilizzando EG, abbiamo che ogni matrice invertibile (e la sua inversa) può essere scritta come prodotto di matrici elemetari $\mathbf{P}(i,j), \mathbf{M}(i,\lambda)$, e $\mathbf{S}(i,j;\lambda)$.

3.4 Interpretazione Geometrica di un Sistema Lineare

◊ In questa parte, daremo un'interpretazione geometrica del concetto di sistema lineare. Per fare ciò, odbbiamo prima introdurre alcune definizioni e risultati. Il primo concetto che introduciamo è quello di *ortogonalità* tra vettori.

Definizione 3.2. Due vettori in uno stesso spazio vettoriale, $\mathbf{v}, \mathbf{w} \in V$, si dicono *ortogonali* quando il loro prodotto scalare è nullo: $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. Viene solitamente denotato con la scrittura $\mathbf{v} \perp \mathbf{w}$.

Esempio 3.12. I vettori

$$\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

sono ortogonali, infatti

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^{3} v_i\, w_i = v_1 w_1 + v_2 w_2 + v_3 w_3 = 1 - 1 + 0 = 0.$$

Proposizione 3.5. *Sia V uno spazio vettoriale finitamente generato, con dimensione pari ad n, su un campo $\mathbb{K}$, $\dim V = n$, e sia $\mathbf{v}$ un suo elemento non nullo, $\mathbf{v} \in V \setminus \{\mathbf{0}\}$. Allora, l'insieme W dei vettori ortogonali a $\mathbf{v}$,*

$$W = \mathbf{v}^{\perp} = \{\mathbf{w} \in V : \langle \mathbf{v}, \mathbf{w} \rangle = 0\}$$

è uno spazio vettoriale di dimensione $n-1$: $\dim W = \dim V - 1$.

Prima di dimostrare questo risultato, osserviamo che, se avessimo posto $\mathbf{v} = \mathbf{0}$, allora lo spazio W coinciderebbe con lo spazio V: $W = V$.
Vediamo adesso la dimostrazione del risultato.

Dimostrazione. Il fatto che W sia uno spazio vettoriale segue dalle identità

$$\langle \mathbf{v}, \alpha \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle, \qquad \langle \mathbf{v}, \mathbf{w} + \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle \langle \mathbf{v}, \mathbf{u} \rangle,$$

per ogni $\alpha \in \mathbb{K}$, $\mathbf{v}, \mathbf{w}, \mathbf{u} \in V$.

La dimensione di W si calcola notando che il prodotto scalere tra due vettori può essere visto come il prodotto matrice-vettore in cui la matrice è una sola riga. Si ha quindi che l'immagine di questa matrice ha dimensione 1, se non tutti i coefficienti sono nulli. Ma poiché l'insieme dei vettori ortogonali coincide col nucleo di $\mathbf{A}$, si ha che, per il teorema di nullità + rango 2.1, la dimensione del nucleo è: $\mathrm{null}\mathbf{A} = \dim V - \mathrm{rk}\mathbf{A}$ da cui

$$\dim \mathbf{v}^{\perp} = \dim V - 1.$$

□

Possiamo quindi osservare subito che se l'insieme dei vettori che soddisfa l'$i-$simo vincolo dato di un sistema lineare omogeneo (ovvero la sua $i-$sima equazione) è lo spazio dei vettori ortogonali alla $i-$sima riga della matrice dei coefficienti $\mathbf{A} \in \mathcal{M}_{m,n}(\mathbb{K})$, che da ora in poi denoteremo con $\mathbf{r}_i$.

Ricordiamo ora che l'insieme delle soluzioni ad un sistema lineare è costituito da tutti i vettori $\mathbf{x} \in \mathbb{K}^n$ che soddisfano tutte le equazioni del sistema. Detto in un altro modo, la soluzione $\mathbf{x}$ al sistema omogeneo $\mathbf{Ax} = \mathbf{0}$ è un vettore ortogonale a tutte le righe $\mathbf{r}_i$ di $\mathbf{A}$.
Detto in un altro modo, lo spazio delle soluzioni al problema lineare omogeneo è l'intersezione degli spazi dei vettori ortogonali alle righe della matrice dei coefficienti:

$$\left\{\mathbf{x} \in \mathbb{K}^n : \mathbf{Ax} = \mathbf{0}\right\} = \mathbf{r}_1{}^\perp \cap \mathbf{r}_2{}^\perp \cap \ldots \cap \mathbf{r}_m{}^\perp = \bigcap_{i=1}^m \mathbf{r}_i{}^\perp.$$

Di conseguenza, la soluzione al sistema è unica (ed è il vettore nullo) quando l'intersezione degli spazi ortogonali è lo spazio nullo:

$$\bigcap_{i=1}^m \mathbf{r}_i{}^\perp = \{\mathbf{0}\} \qquad \Leftrightarrow \qquad \dim\left(\bigcap_{i=1}^m \mathbf{r}_i{}^\perp\right) = 0.$$

Al contrario, la soluzione non è unica quando questa intersezione non è lo spazio nullo. Ed in questo caso, l'insieme delle soluzioni è lo spazio dell'intersezione di tutti gli spazi ortogonali. Inoltre, la dimensione di questo spazio coincide col numero di parametri liberi della soluzione al sistema.

Esempio 3.13. Consideriamo il sistema lineare omogeneo a tre equazioni e tre incognite dove la matrice dei coefficienti è

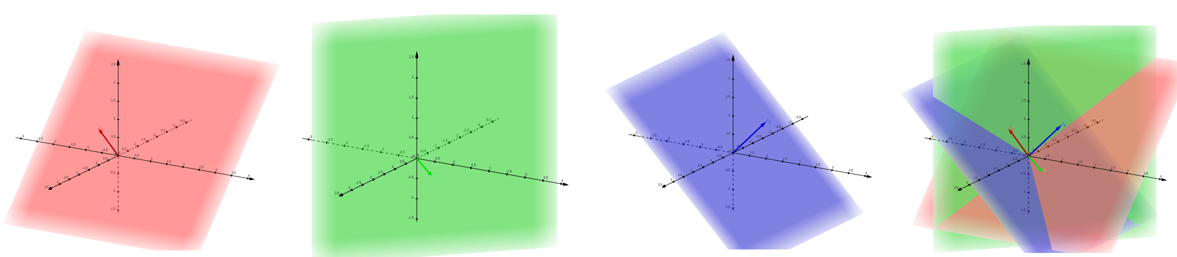
$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix},$$

e l'unica soluzione è il vettore nullo. Abbiamo, con la notazione introdotta prima,

$$\mathbf{r}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{r}_1{}^\perp = \left\langle \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\rangle,$$

$$\mathbf{r}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \qquad \mathbf{r}_2{}^\perp = \left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\rangle,$$

$$\mathbf{r}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \qquad \mathbf{r}_3{}^\perp = \left\langle \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\rangle.$$



Esempio 3.14. Consideriamo il sistema lineare omogeneo a tre equazioni e tre incognite dove la matrice dei coefficienti è

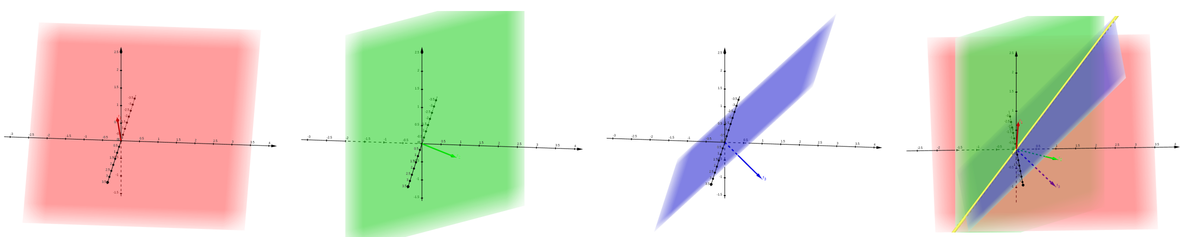
$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix},$$

e lo spazio delle soluzioni è generato dal vettore $[1,-1,-1]^\intercal$. Abbiamo, con la notazione introdotta prima,

$$\mathbf{r}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{r}_1{}^{\perp} = \left\langle \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\rangle,$$

$$\mathbf{r}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \qquad \mathbf{r}_2{}^{\perp} = \left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\rangle,$$

$$\mathbf{r}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \qquad \mathbf{r}_3{}^{\perp} = \left\langle \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\rangle.$$



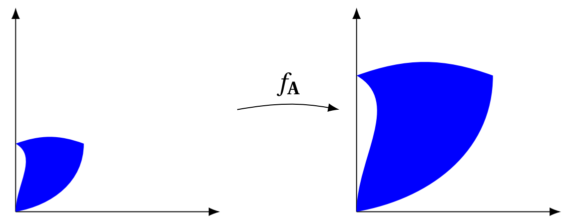
4 Determinante

Prima di iniziare a parlare di 'determinante', conviene dare un'interpretazione geometrica di endomorfismo. Per fare ciò, dobbiamo immaginare di applicare un endomorfismo ad un insieme di vettori. In particolare, consideriamo l'endomorfismo $f_{\mathbf{A}} : \mathbf{x} \mapsto \mathbf{Ax}$ e vediamo come diverse matrici $\mathbf{A}$ risultano in diverse 'trasformazioni' di parti di piano (ci concentriamo quindi, per questo ragionamento, al caso $n = 2$, ma si estende a qualsiasi $n \in \mathbb{N}$).

Una matrice del tipo

$$\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

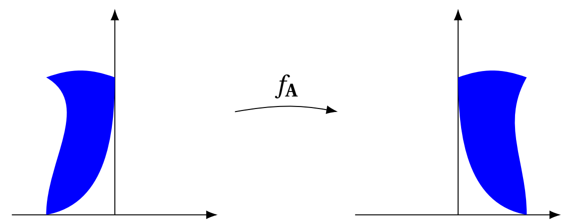
risulta in un "raddoppio" di ogni vettore dentro all'insieme, scalando l'insieme per 2.



Una matrice del tipo

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

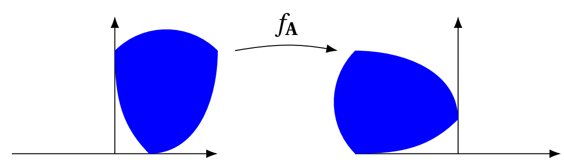
risulta in una "riflessione" di ogni vettore dentro all'insieme rispetto all'asse y.



Una matrice del tipo

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

risulta in una "rotazione" di 90°, rispetto all'origine, di ogni vettore dentro all'insieme.



Desideriamo quindi un operatore che descriva, in un qualche senso, come la matrice $\mathbf{A}$ trasformi l'area (con $n = 2$, il volume per $n = 3$) di un insieme. Ad esempio, per la prima matrice, ci aspettiamo un valore di 4 (l'area viene quadruplicata quando le dimensioni raddoppiano); per la seconda ci aspettiamo un valore di -1 (l'area non cambia ma, in qualche modo, cambia segno); ed un valore di 1 per il terzo caso (l'area non cambia in nessun modo a seguito di una rotazione).

In questa parte daremo la definizione di *determinante* di una matrice quadrata, e studieremo le proprietà di questo operatore.

Prima di fare ciò, notiamo il fatto che le matrici associate alle operazioni elementari che abbiamo definito in Sezione 3.3.1 hanno particolari effetti quando utilizzate per trasformare un vettore, e questi effetti si notano in particolare quando si trasformano porzioni intere del dominio (come un triangolo nel piano, o un tetraedro nello spazio)

Ad esempio, la marice $\mathbf{P}(i, j)$ è una riflessione. Quindi, se applichiamo questa matrice ad ogni punto interno ad un triangolo, otterremo un altro triangolo della stessa area. In modo simile, la matrice $\mathbf{M}(i, \lambda)$ moltiplica l'area (o il volume) dell'insieme per λ; mentre la matrice $\mathbf{S}(i, j; \lambda)$ non cambia l'area dell'insieme.

Queste osservazioni motiveranno le proprietà che utilizzeremo per definire l'operatore determinate nella definizione 4.1.

4.1 La Funzione Determinante

Prima di mostrare come si calcola il determinante di una matrice quadrata, daremo una definizione più generica di determinante, e mostreremo dopo che esiste essenzialmente una sola funzione che soddisfa tale operazione. Per fare ciò ci rifaremo alle matrici correlate alle operazioni elementari $\mathbf{P}(i,j)$ in (13), $\mathbf{M}(i,\lambda)$ in (14), ed $\mathbf{S}(i,j;\lambda)$ in (15).

Definizione 4.1 (Determinante)**.** Un'applicazione $\varphi^{(n)} : \mathcal{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ si dice una *funzione determinante* se, per ogni matrice quadrata $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{K})$ ed ogni costante $\lambda \in \mathbb{K}$, soddisfa le seguenti identità:

D1. $\varphi^{(n)}(\mathbf{P}(i,j)\mathbf{A}) = -\varphi^{(n)}(\mathbf{A})$;

D2. $\varphi^{(n)}(\mathbf{M}(i,\lambda)\mathbf{A}) = \lambda\varphi^{(n)}(\mathbf{A})$;

D3. $\varphi^{(n)}(\mathbf{S}(i,j;\lambda)\mathbf{A}) = \varphi^{(n)}(\mathbf{A})$.

In altre parole, la proprietà D1. ci dice che se scambiamo due righe di una matrice, allora il determinante cambia di segno. La proprietà D2. implica che se moltiplichiamo una riga per una costante, il determinante viene moltiplicato per quella stessa costante. Infine, la proprietà D3. implica che sommare ad una riga un multiplo di un'altra non cambia il valore del determinante.

La funzione determinante soddisfa le seguenti proprietà.

Proposizione 4.1. *Sia $\varphi^{(n)} : \mathcal{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ una funzione determinante e sia $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{K})$ una matrice quadrata. Allora:*

1. *se una riga di $\mathbf{A}$ è nulla, allora il determinante è nullo;*

2. *se una matrice $\mathbf{A}$ non ha rango massimo, allora il determinante è nullo.*

◇ *Dimostrazione.* Per dimostrare il punto 1. assumiamo che la matrice $\mathbf{A}$ abbia la $i-$sima riga nulla. Allora abbiamo che, per ogni $\lambda \in \mathbb{K}$, $\mathbf{M}(i,\lambda)\mathbf{A} = \mathbf{A}$ in quanto stiamo moltiplicando una riga nulla per una costante. Ma, poiché dalla proprietà D2., $\varphi^{(n)}(\mathbf{M}(i,\lambda)\mathbf{A}) = \lambda\varphi^{(n)}(\mathbf{A})$. Quindi vale l'identità

$$\lambda\varphi^{(n)}(\mathbf{A}) = \varphi^{(n)}(\mathbf{A}), \quad \forall \lambda \in \mathbb{K},$$

che implica $\varphi^{(n)}(\mathbf{A}) = 0$, dimostrando il punto 1.

Per dimostrare il punto 2. si osservi che, applicando EG alla matrice $\mathbf{A}$ per ottenere una matrice con forma a scala $\mathbf{A}'$, si ottiene che esiste uno scalare $\tilde{\lambda} \neq 0$ tale per cui

$$\varphi^{(n)}(\mathbf{A}') = \tilde{\lambda}\varphi^{(n)}(\mathbf{A}).$$

Se la matrice $\mathbf{A}$ non ha rango massimo, allora la matrice con forma a scala $\mathbf{A}'$ ha una riga nulla, il cui determinante è nullo (dal punto 1.). Da questo segue che anche il determinante di $\mathbf{A}$ è nullo. □

Mostriamo ora che esiste essenzialmente un'unica funzione determinante, a meno di costanti.

Teorema 4.1. *Sia $\alpha \in \mathbb{K}$ uno scalare. Allora esiste al più una funzione determinante $\varphi^{(n)} : \mathcal{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ tale per cui $\varphi^{(n)}(\mathbf{Id}_n) = \alpha$.*

◇ *Dimostrazione.* Dalla proprietà 4.1 l'applicazione $\varphi^{(n)}$ ha valore non nullo solo per matrici invertibili. Dall'osservazione 3.6, abbiamo che ogni matrice invertibile $\mathbf{A}$ si può scrivere come prodotto di matrici elementari. Da questo segue, grazie alla proprietà del determinante nella definizione 4.1 che

$$\begin{aligned}
&\varphi^{(n)}(\mathbf{P}(i,j)\,\mathbf{Id}_n) = -\varphi^{(n)}(\mathbf{Id}_n) = -\alpha;\\
&\varphi^{(n)}(\mathbf{M}(i,\lambda)\,\mathbf{Id}_n) = \lambda\varphi^{(n)}(\mathbf{Id}_n) = \lambda\,\alpha;\\
&\varphi^{(n)}(\mathbf{S}(i,j;\lambda)\mathbf{Id}_n) = \varphi^{(n)}(\mathbf{Id}_n) = \alpha.
\end{aligned}$$

Di conseguenza, il determinante può essere univocamente determinanto conoscendo le operazioni elementari che definiscono $\mathbf{A}$, ed il valore del determinante per la matrice identità. □

Questo teorema ci permette di definire in modo univoco il determinante $\det : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ come la funzione determinante $\varphi^{(n)}$ per cui la matrice identià abbia determinante unitario:

$$\det : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}, \quad \det : \mathbf{A} \mapsto \frac{\varphi^{(n)}(\mathbf{A})}{\varphi^{(n)}(\mathbf{Id}_n)}.$$

Vediamo ora un importante risultato che mette in relazione il determinante di un prodotto di matrici col prodotto dei determinanti.

Teorema 4.2 (Binet)**.** *Sia $\varphi : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ è una funzione determinante tale per cui $\varphi(\mathbf{Id}_n) = 1$, e siano $\mathbf{A},\mathbf{B} \in \mathscr{M}_{n,n}(\mathbb{K})$. Allora*

$$\varphi(\mathbf{A}\mathbf{B}) = \varphi(\mathbf{A})\,\varphi(\mathbf{B}).$$

Dimostrazione. Consideriamo una matrice $\mathbf{C} \in \mathscr{M}_{n,n}(\mathbb{K})$ e definiamo una mappa $\psi : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ come segue: $\psi(\mathbf{X}) = \varphi(\mathbf{C}\mathbf{X})$. Allora valgono le seguenti identità:

$$\begin{aligned}
&\psi(\mathbf{P}(i,j)\mathbf{X}) = \varphi(\mathbf{P}(i,j)\mathbf{C}\mathbf{X}) = -\varphi(\mathbf{C}\mathbf{X}) = -\psi(\mathbf{X}),\\
&\psi(\mathbf{M}(i,\lambda)\mathbf{X}) = \varphi(\mathbf{M}(i,\lambda)\mathbf{C}\mathbf{X}) = \lambda\varphi(\mathbf{C}\mathbf{X}) = \lambda\psi(\mathbf{X}),\\
&\psi(\mathbf{S}(i,j;\lambda)\mathbf{X}) = \varphi(\mathbf{S}(i,j;\lambda)\mathbf{C}\mathbf{X}) = \varphi(\mathbf{C}\mathbf{X}) = \psi(\mathbf{X}).
\end{aligned}$$

Quidni, anche la mappa ψ è una funzione determinante per cui $\psi(\mathbf{Id}) = \varphi(\mathbf{C})$.

Se ora consideriamo una funzione $\psi' : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ definita come $\psi'(\mathbf{X}) = \varphi(\mathbf{C})\,\varphi(\mathbf{X})$. In modo analogo a quanto fatto prima, si ha che anche ψ' è una funzione determinante per cui $\psi'(\mathbf{X}) = \varphi(\mathbf{C})\,\varphi(\mathbf{X})$.

Poiché dal teorema 4.1 segue che, fissato il valore assegnato alla matrice identità, il determinante è unico, si ha che le mappe ψ e ψ' coincidono: $\psi \equiv \psi'$.

Calcolando per $\mathbf{C} = \mathbf{A}$ ed $\mathbf{X} = \mathbf{B}$ abbiamo la tesi. □

Si ha inoltre il fatto che le funzione determinante è invariante per traslazioni dell'argomento.

Proposizione 4.2. *Sia $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$ e sia $\varphi : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ una funzione determinante. Allora*

$$\varphi(\mathbf{A}^\top) = \varphi(\mathbf{A}).$$

◇ *Dimostrazione.* Se $\mathbf{A}$ non è invertibile, allora nemmeno la sua trasposta lo è, e il determinante di entrambe è pari a zero.

Altrimenti, la matrice $\mathbf{A}$ si può scrivere come prodotto di matrici elementari $\mathbf{E}_i$: $\mathbf{A} = \mathbf{E}_1 \mathbf{E}_2 \cdots \mathbf{E}_n$ da cui, utilizzando il Teorema di Binet 4.2

$$\begin{aligned}
\varphi(\mathbf{A}^\top) &= \varphi\left((\mathbf{E}_1 \mathbf{E}_2 \cdots \mathbf{E}_n)^\top\right) \\
&= \varphi\left(\mathbf{E}_n^\top \mathbf{E}_{n-1}^\top \cdots \mathbf{E}_1^\top\right) \\
&= \varphi(\mathbf{E}_n^\top)\, \varphi(\mathbf{E}_{n-1}^\top) \cdots \varphi(\mathbf{E}_1^\top)
\end{aligned}$$

Ora, poiché per le matrici elementari valgono le seguenti identità:

$$\mathbf{P}^\top(i,j) = \mathbf{P}(i,j), \qquad \mathbf{M}^\top(i,\lambda) = \mathbf{M}(i,\lambda), \qquad \mathbf{S}^\top(i,j;\lambda) = \mathbf{S}(j,i;\lambda),$$

abbiamo la tesi:

$$\varphi(\mathbf{A}^\top) = \varphi(\mathbf{E}_n^\top)\, \varphi(\mathbf{E}_{n-1}^\top) \cdots \varphi(\mathbf{E}_1^\top) = \varphi(\mathbf{E}_n)\, \varphi(\mathbf{E}_{n-1}) \cdots \varphi(\mathbf{E}_1) = \varphi(\mathbf{A}).$$

□

Osservazione 4.1. Poiché il determinante è invariante per trasposizioni, abbiamo che il determinante può essere definito in modo equivalente a quanto fatto nella definizione 4.1 come un'applicazione $\varphi^{(n)} : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ che, per ogni matrice quadrata $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$ ed ogni costante $\lambda \in \mathbb{K}$, soddisfa le seguenti identità:

d1. $\varphi^{(n)}(\mathbf{A}\,\mathbf{P}(i,j)) = -\varphi^{(n)}(\mathbf{A})$;

d2. $\varphi^{(n)}(\mathbf{A}\mathbf{M}(i,\lambda)) = \lambda\, \varphi^{(n)}(\mathbf{A})$;

d3. $\varphi^{(n)}(\mathbf{A}\,\mathbf{S}(i,j;\lambda)) = \varphi^{(n)}(\mathbf{A})$.

Ricordiamo (osservazione 3.5) che una moltiplicazione a sinistra per una matrice elementare opera sulle righe, mentre una moltiplicazione a destra opera sulle colonne.

Vediamo ora un risultato fondamentale riguardante il determinante di una trasposta.

Proposizione 4.3. *Sia* $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$ *una matrice invertibile e sia* $\varphi : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ *una funzione determinante tale per cui* $\varphi(\mathbf{Id}) = 1$. *Allora*

$$\varphi(\mathbf{A}^{-1}) = \varphi(\mathbf{A})^{-1}.$$

Dimostrazione. Il risultato segue immediatamente dal teorema di Binet 4.2:

$$1 = \varphi(\mathbf{Id}) = \varphi(\mathbf{A}\mathbf{A}^{-1}) = \varphi(\mathbf{A})\, \varphi(\mathbf{A}^{-1}).$$

□

4.2 Una Formula per il Determinante

Fin'ora, abbiamo dimostrato che, se esistesse una funzione determinante per ongi n, allora questa funzione è unica (a meno di costanti). Abbiamo inoltre dimostrato diversi risultati riguardanti la funzione determinante.

Il seguente risultato mostra che questa funzione esiste per ogni valore $n \in \mathbb{N}_+$.

Teorema 4.3 (Laplace)**.** *Per ogni* $n = 1,2,3,\dots$ *esiste una funzione determinante* $\varphi^{(n)} : \mathscr{M}_{n,n}(\mathbb{K}) \to \mathbb{K}$ *tale per cui* $\varphi^{(n)}(\mathbf{Id}) = 1$.

◇ *Dimostrazione.* Dimostriamo il risultato per induzione.

La tesi è ovvia per $n = 1$ definendo $\varphi^{(1)}$ come la mappa identità: $\varphi^{(1)} : [a_{11}] \mapsto a_{11}$.

Supponiamo ora $n > 1$ e di avere a disposizione una funzione determinante $\varphi^{(n-1)}$. Questo significa che data una matrice $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$, possiamo calcolare il determinante per ogni sottomatrice $\mathbf{A}_{ij}$ ottenuta eliminando una riga i ed una colonna j. Definiamo ora

$$\varphi(\mathbf{A}) = \varphi^{(n)}(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \varphi^{(n-1)}(\mathbf{A}_{ij}), \tag{16}$$

e verifichiamo che sia una funzione determinante.

Iniziamo con la proprietà d1. L'idea sarebbe quella di definire una matrice $\mathbf{B} = [b_{ij}] = \mathbf{AP}(i,j)$ per una coppia di indici k,ℓ. Tuttavia, questo caso generale richiede una quantità enorme di indici differenti. Invece, dimostriamo la proprietà d1. per il caso $\mathbf{B} = [b_{ij}] = \mathbf{AP}(1,2)$. Il caso generico si dimostra in modo identico. Abbiamo

$$b_{1j} = \begin{cases} a_{12} & \text{se } j = 1 \\ a_{11} & \text{se } j = 2 \\ a_{1j} & \text{altrimenti} \end{cases} \qquad \mathbf{B}_{1j} = \begin{cases} \mathbf{A}_{12} & \text{se } j = 1 \\ \mathbf{A}_{11} & \text{se } j = 2 \\ \mathbf{A}_{1j}\mathbf{P}(1,2) & \text{altrimenti} \end{cases}$$

da cui

$$\begin{aligned}
\varphi(\mathbf{B}) &= \sum_{j=1}^{n} (-1)^{1+j} b_{1j} \varphi^{(n-1)}(\mathbf{B}_{ij}) \\
&= (-1)^{1+i} b_{11} \varphi^{(n-1)}(\mathbf{B}_{11}) + (-1)^{1+2} b_{12} \varphi^{(n-1)}(\mathbf{B}_{12}) + \sum_{j=3}^{n} (-1)^{1+j} b_{1j} \varphi^{(n-1)}(\mathbf{B}_{1j}) \\
&= a_{12} \varphi^{(n-1)}(\mathbf{A}_{12}) - a_{11} \varphi^{(n-1)}(\mathbf{A}_{11}) + \sum_{j=3}^{n} (-1)^{1+j} a_{1j} \varphi^{(n-1)}(\mathbf{A}_{1j}\mathbf{P}(1,2)) \\
&= -\left(a_{11} \varphi^{(n-1)}(\mathbf{A}_{11}) - a_{12} \varphi^{(n-1)}(\mathbf{A}_{12})\right) + (-1) \sum_{j=3}^{n} (-1)^{1+j} a_{1j} \varphi^{(n-1)}(\mathbf{A}_{1j}) \\
&= -(-1) \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \varphi^{(n-1)}(\mathbf{A}_{1j}) \\
&= -\varphi(\mathbf{A}).
\end{aligned}$$

Dimostriamo ora che vale la proprietà d2. Consideriamo, a tale scopo, la matrice $\mathbf{B} = \mathbf{AM}(k,\lambda)$ da cui

$$b_{1j} = \begin{cases} a_{1j} & \text{se } j \neq k \\ \lambda a_{1k} & \text{se } j = k \end{cases} \qquad \mathbf{B}_{1j} = \begin{cases} \mathbf{A}_{1j}\mathbf{M}(k,\lambda) & \text{se } j \neq k \\ \mathbf{A}_{1k} & \text{se } j = k \end{cases}$$

Abbiamo quindi

$$
\begin{aligned}
\varphi(\mathbf{B}) &= \sum_{j=1}^{n} (-1)^{1+j} b_{1\,j} \varphi^{(n-1)}(\mathbf{B}_{1\,j}) \\
&= (-1)^{1+k} \lambda a_{1\,k} \varphi^{(n-1)}(\mathbf{A}_{1\,k}) + \sum_{\substack{j=1 \\ j\neq k}}^{n} a_{1\,j} \varphi^{(n-1)}(\mathbf{A}_{1\,j}\mathbf{M}(k,\lambda)) \\
&= \lambda \sum_{j=1}^{n} a_{1\,j} \varphi^{(n-1)} \mathbf{A}_{1\,j} \\
&= \lambda \varphi(\mathbf{A}).
\end{aligned}
$$

Dimostriamo ora la proprietà d3. Come per la proprietà d1. dimostriamo solo il caso particolare $\mathbf{B} = \mathbf{AS}(2,1;\lambda)$ per evitare di introdurre un numero eccessivo di indici. Abbiamo

$$
b_{1\,j} = \begin{cases} a_{1\,1} + \lambda a_{1\,2} & \text{se } j = 1 \\ a_{1\,j} & \text{se } j > 1 \end{cases}
\qquad\qquad
\mathbf{B}_{1\,j} = \begin{cases} \mathbf{A}_{1\,1} & \text{se } j = 1 \\ \mathbf{X} & \text{se } j = 2 \\ \mathbf{A}_{[1\,j]} S(2,1;\lambda) & \text{se } j > 2 \end{cases}
$$

per una opportuna matrice $\mathbf{X}$ che studieremo tra poco. Calcoliamo ora il determinante di $\mathbf{B}$.

$$
\begin{aligned}
\varphi(\mathbf{B}) &= \sum_{j=1}^{n} (-1)^{1+j} b_{1\,j} \varphi^{(n-1)}(\mathbf{B}_{i\,j}) \\
&= (-1)^{1+1}(a_{1\,1} + \lambda a_{1\,2}) \varphi^{(n-1)}(\mathbf{A}_{1\,1}) + (-1)^{1+2} \varphi^{(n-1)}(\mathbf{X}) + \sum_{j=3}^{n} (-1)^{1+j} a_{1\,j} \varphi^{(n-1)}(\mathbf{A}_{1\,j}\mathbf{S}(2,1;\lambda)) \\
&= \varphi(\mathbf{A}) + \lambda a_{1\,2} \varphi^{(n-1)}(\mathbf{A}_{1\,1}) + (-1)^{1+2} a_{1\,2} \varphi^{(n-1)}(\mathbf{X}) - (-1)^{1+2} a_{1\,2} \varphi^{(n-1)}(\mathbf{A}_{1\,2}).
\end{aligned}
$$

Quindi, per dimostrare la tesi, dobbiamo dimostrare che

$$
\varphi^{(n-1)}(\mathbf{X}) = \varphi^{(n-1)}(\mathbf{A}_{1\,2}) + \lambda \varphi^{(n-1)}(\mathbf{A}_{1\,1}).
$$

Più in generale, dimostreremo che, per $m = n-1$, il determinante $\varphi^{(m)}$ è lineare nella prima colonna, ovvero che, date tre matrici

$$
\mathbf{X}' = \begin{bmatrix} \mathbf{x}' & \mathbf{x}_2 & \cdots & \mathbf{x}_m \end{bmatrix}, \qquad \mathbf{X}'' = \begin{bmatrix} \mathbf{x}'' & \mathbf{x}_2 & \cdots & \mathbf{x}_m \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} \mathbf{x}' + \mathbf{x}'' & \mathbf{x}_2 & \cdots & \mathbf{x}_m \end{bmatrix}
$$

si ha

$$
\varphi^{(m)}(\mathbf{X}) = \varphi^{(m)}(\mathbf{X}') + \varphi^{(m)}(\mathbf{X}''). \tag{17}
$$

Ricordiamo che è sufficiente dimostrare questa proprietà per dire che il determinante è lineare nella prima colonna in quanto già sappiamo che

$$
\varphi^{(m)}\left(\begin{bmatrix} \lambda\mathbf{x} & \mathbf{x}_2 & \cdots & \mathbf{x}_m \end{bmatrix}\right) = \lambda \varphi^{(m)}\left(\begin{bmatrix} \mathbf{x} & \mathbf{x}_2 & \cdots & \mathbf{x}_m \end{bmatrix}\right).
$$

Sicuramente (17) è vera per $m = 1$: la funzione $\varphi^{(1)}$ è l'identità. Operiamo quindi per induzione su m, assumento che l'assunzione sia vera per $m-1$. Denotiamo con x_1, x'_1, ed x''_1 i coefficienti in prima riga di

$\mathbf{x}$, $\mathbf{x}'$, e $\mathbf{x}''$ rispettivamente. Inoltre, chiamiamo $\mathbf{y}$, $\mathbf{y}'$, $\mathbf{y}''$, e $\mathbf{y}_j$ i vettori ottenuti rimuovendo il coefficiente della prima riga di $\mathbf{x}$, $\mathbf{x}'$, $\mathbf{x}''$, e $\mathbf{x}_j$ rispettivamente.
Quando applichiamo $\varphi^{(m)}$ ad $\mathbf{X}$, dobbiamo cancellare via via una colonna, oltre alla rima riga. Se cancelliamo la prima colonna, rimane la matrice

$$\begin{bmatrix} \mathbf{y}_2 & \mathbf{y}_3 & \cdots & \mathbf{y}_m \end{bmatrix},$$

mentre se cancelliamo la seconda colonna si ottiene la matrice

$$\begin{bmatrix} \mathbf{y}' + \mathbf{y}'' & \mathbf{y}_3 & \cdots & \mathbf{y}_m \end{bmatrix},$$

e così anche per le successive, alle quali si applica l'ipotesi induttiva. Svolgendo i calcoli si ottiene la tesi. □

Osservazione 4.2. In questa dimostrazione, abbiamo definito ricorsivamente il determinante utilizzando l'equazione (16). In realtà, si può sviluppare il determinante su una qualsiasi riga di $\mathbf{A}$ tramite la *formula di Laplace*.

$$\varphi^{(n)}(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \varphi^{(n-1)}(\mathbf{A}_{ij}).$$

È facile vedere che si può passare da una formulazione all'altra considerando una sequenza di permutazioni di righe tramite matrici $\mathbf{P}(i,j)$.

Facciamo notare che la dimostrazione di questo risultato si ottiene per induzione **definendo** una funzione determinante. Di conseguenza, oltre a dimostrare l'esistenza (unica) di una funzione determinante, abbiamo anche la formulazione per poterla calcolare. In particolare, si definisce il determinante di una matrice 1×1 come il valore dell'unica componente della matrice stessa:

$$\varphi^{(1)}\left(\begin{bmatrix} a_{11} \end{bmatrix}\right) = a_{11}.$$

Poi, si definisce il determinante di una generica matrice $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$ come segue:

$$\varphi^{(n)}(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \varphi^{(n-1)}(\mathbf{A}_{ij}),$$

dove la matrice $\mathbf{A}_{ij}$ è la *sottomatrice* ottenuta togliendo da $\mathbf{A}$ la $i-$sima riga e $j-$sima colonna. Notiamo che il determinante si può calcolare considerando qualsiasi riga i. Semplificando la notazione, abbiamo che il determinante è definito dalle seguenti relazioni:

$$\det\left(\begin{bmatrix} a_{11} \end{bmatrix}\right) = a_{11}, \tag{18a}$$

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det \mathbf{A}_{ij}. \tag{18b}$$

Osservazione 4.3. Poiché il determinante non cambia per trasposizione (proposizione 4.2), abbiamo che la sommatoria si può svolgere lungo l'indice di riga i per ogni colonna j. In questo modo, anizchè usare (18b), possiamo calcolare il determinante di una matrice in modo equivalente come

$$\det(\mathbf{A}) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det \mathbf{A}_{ij}. \tag{19}$$

Esempio 4.1. Calcoliamo il determinante della matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 1 & 2 & 0 \end{bmatrix}.$$

Iniziamo col calcolo (in cui introduciamo una nuova notazione: la matrice tra due linee verticali denota il determinante di quella matrice):

$$\begin{aligned}
\det \mathbf{A} &= \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 1 & 2 & 0 \end{vmatrix} \\
&= (1)^{1+1} \begin{vmatrix} 1 & 3 \\ 2 & 0 \end{vmatrix} + (1)^{1+2} 2 \begin{vmatrix} 0 & 3 \\ 1 & 0 \end{vmatrix} + (1)^{1+3} 1 \begin{vmatrix} 0 & 1 \\ 1 & 2 \end{vmatrix} \\
&= \begin{vmatrix} 1 & 3 \\ 2 & 0 \end{vmatrix} - 2 \begin{vmatrix} 0 & 3 \\ 1 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 1 \\ 1 & 2 \end{vmatrix}.
\end{aligned}$$

Ora, dobbiamo calcolare i determinanti delle sottomatrici

$$\begin{vmatrix} 1 & 3 \\ 2 & 0 \end{vmatrix} = 1 \cdot 0 - 3 \cdot 2 = -6 \qquad \begin{vmatrix} 0 & 3 \\ 1 & 0 \end{vmatrix} = 0 \cdot 0 - 3 \cdot 1 = -3 \qquad \begin{vmatrix} 0 & 1 \\ 1 & 2 \end{vmatrix} = 0 \cdot 2 - 1 \cdot 1 = -1;$$

da cui otteniamo

$$\det \mathbf{A} = -6 - 2(-3) + (-1) = -1.$$

4.3 Calcolo del Determinante

Le equazioni (18) ci permettono di calcolare il determinante. Tuttavia, trattandosi di un metodo ricorsivo, per matrici relativamente grandi richiederebbe numerosi passaggi. In questa parte, vedremo alcuni metodi per calcolare il determinante in modo veloce.

Il primo metodo è notare che per una matrice 2×2 è molto facile calcolare il determinante. Abbiamo infatti

$$\det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = (1) a_{11} \begin{vmatrix} a_{22} \end{vmatrix} + (-1) a_{12} \begin{vmatrix} a_{21} \end{vmatrix} = a_{11} a_{22} - a_{12} a_{21}.$$

Quindi, in una matrice 2×2, il determinante è pari al prodotto degli elementi della diagonale sottratto del prodotto degli elementi della anti-diagonale.

In modo simile, si può trovare una semplice formula, detta *formula di Sarrus*, per calcolare il determinante di una matrice 3×3:

$$\begin{aligned}
\det\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\right) &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
&= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\
&= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\
&= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - \\
&\qquad (a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33} + a_{13}a_{22}a_{31})
\end{aligned}$$

Nonostante sembri una formula complicata da ricordare, c'è un metodo 'grafico' per ricordarla. Per ottenere questo metodo, è necessario definire la seguente matrice

$$\tilde{\mathbf{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{bmatrix}.$$

Allora, il determinante di **A** si ottiene sommando i prodotti sulle tre diagonali e sottraendo i prodotti sulle tre anti-diagonali. La seguente colorazione dovrebbe aiutare a capire come si faccia.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{bmatrix} \qquad \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{bmatrix}$$

$$\det \mathbf{A} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - (a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33} + a_{13}a_{22}a_{31}).$$

4.3.1 Calcolo del Determinante Tramite EG

L'eliminazione di Gauss può essere usata per calcolare il determinante anche di grandi matrici senza bisogno di svolgere troppi calcoli. Il modo per fare ciò segue dall'osservazione che per una matrice trangolare (superiore), il determinate è pari al prodotto degli elementi sulla diagonale. Infatti, si consideri una generica matrice triangolare superiore $\mathbf{U} \in \mathscr{M}_{n,n}(\mathbb{K})$

$$\mathbf{U} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{1n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix}.$$

Possiamo calcolare il determinante usando la formula (19) applicata alla prima colonna ottenendo, poiché tutti gli elementi della colonna tranne il primo sono nulli,

$$\det(\mathbf{U}) = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{vmatrix}.$$

Il determinante a destra si può scrivere in modo simile, ottenendo

$$\det(\mathbf{U}) = a_{11}a_{22} \begin{vmatrix} a_{33} & a_{34} & \cdots & a_{3n} \\ 0 & a_{44} & \cdots & a_{4n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{vmatrix}.$$

Proseguendo in questo modo arriviamo alla conclusione

$$\det(\mathbf{U}) = a_{11}a_{22}\cdots a_{nn} = \prod_{i=1}^{n} a_{ii}.$$

Quindi, per calcolare il determinante di una matrice, si può applicare EG tenendo traccia di tutte le volte che un'operazione elementare risulta in una modifica del determinante stesso. Arrivati ad una matrice triangolare, si moltipicano i valori sulla diagonale e si moltiplicano per i fattori introdotti durante EG.

Vediamo un esempio su come usare EG per calcolare il determinante di una matrice.

Esempio 4.2. Calcoliamo il determinante di

$$\mathbf{A} = \begin{bmatrix} -2 & 0 & 2 & 2 & -3 \\ -1 & -1 & 1 & 1 & -2 \\ 0 & -1 & 1 & 2 & 1 \\ 1 & 1 & 0 & -2 & -1 \\ 2 & 1 & -1 & 2 & 0 \end{bmatrix}$$

usando EG.

Iniziamo annullando gli elementi della prima colonna, escluso quello in prima riga:

$$\begin{vmatrix} -2 & 0 & 2 & 2 & -3 \\ -1 & -1 & 1 & 1 & -2 \\ 0 & -1 & 1 & 2 & 1 \\ 1 & 1 & 0 & -2 & -1 \\ 2 & 1 & -1 & 2 & 0 \end{vmatrix} \xrightarrow[\substack{r_2 \to 2r_2 \\ r_4 \to 2r_4}]{(EG)} \frac{1}{2} \cdot \frac{1}{2} \cdot \begin{vmatrix} -2 & 0 & 2 & 2 & -3 \\ -2 & -2 & 2 & 2 & -4 \\ 0 & -1 & 1 & 2 & 1 \\ 2 & 2 & 0 & -4 & -2 \\ 2 & 1 & -1 & 2 & 0 \end{vmatrix} \xrightarrow[\substack{r_2 \to r_2 - r_1 \\ r_4 \to r_4 + r_1 \\ r_5 \to r_5 + r_1}]{(EG)} \frac{1}{2} \cdot \frac{1}{2} \cdot \begin{vmatrix} -2 & 0 & 2 & 2 & -3 \\ 0 & -2 & 0 & 0 & -1 \\ 0 & -1 & 1 & 2 & 1 \\ 0 & 2 & 2 & -2 & -5 \\ 0 & 1 & 1 & 4 & -3 \end{vmatrix}$$

Ora, annulliamo le ultime tre componenti della seconda colonna

$$\frac{1}{4} \cdot \begin{vmatrix} -2 & 0 & 2 & 2 & -3 \\ 0 & -2 & 0 & 0 & -1 \\ 0 & -1 & 1 & 2 & 1 \\ 0 & 2 & 2 & -2 & -5 \\ 0 & 1 & 1 & 4 & -3 \end{vmatrix} \xrightarrow[\substack{r_3 \to 2r_3 \\ r_5 \to 2r_5}]{(EG)} \frac{1}{4} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \begin{vmatrix} -2 & 0 & 2 & 2 & -3 \\ 0 & -2 & 0 & 0 & -1 \\ 0 & -2 & 2 & 4 & 2 \\ 0 & 2 & 2 & -2 & -5 \\ 0 & 2 & 2 & 8 & -6 \end{vmatrix} \xrightarrow[\substack{r_3 \to r_3 - r_2 \\ r_4 \to r_4 + r_2 \\ r_5 \to r_5 + r_2}]{(EG)} \frac{1}{4} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \begin{vmatrix} -2 & 0 & 2 & 2 & -3 \\ 0 & -2 & 0 & 0 & -1 \\ 0 & 0 & 2 & 4 & 3 \\ 0 & 0 & 2 & -2 & -6 \\ 0 & 0 & 2 & 8 & -7 \end{vmatrix}.$$

Andiamo ora ad annullare le ultime due componenti della terza colonna

$$\frac{1}{16} \cdot \begin{vmatrix} -2 & 0 & 2 & 2 & -3 \\ 0 & -2 & 0 & 0 & -1 \\ 0 & 0 & 2 & 4 & 3 \\ 0 & 0 & 2 & -2 & -6 \\ 0 & 0 & 2 & 8 & -7 \end{vmatrix} \xrightarrow[\substack{r_4 \to r_4 - r_3 \\ r_5 \to r_5 - r_3}]{(EG)} \frac{1}{16} \cdot \begin{vmatrix} -2 & 0 & 2 & 2 & -3 \\ 0 & -2 & 0 & 0 & -1 \\ 0 & 0 & 2 & 4 & 3 \\ 0 & 0 & 0 & -6 & -9 \\ 0 & 0 & 0 & 4 & -10 \end{vmatrix}.$$

Infine, annulliamo la componente in $(5,4)$:

$$\frac{1}{16} \cdot \begin{vmatrix} -2 & 0 & 2 & 2 & -3 \\ 0 & -2 & 0 & 0 & -1 \\ 0 & 0 & 2 & 4 & 3 \\ 0 & 0 & 0 & -6 & -9 \\ 0 & 0 & 0 & 4 & -10 \end{vmatrix} \xrightarrow[r_4 \to \frac{r_4}{3}]{(EG)} \frac{1}{16} \cdot 3 \cdot \begin{vmatrix} -2 & 0 & 2 & 2 & -3 \\ 0 & -2 & 0 & 0 & -1 \\ 0 & 0 & 2 & 4 & 3 \\ 0 & 0 & 0 & -2 & -3 \\ 0 & 0 & 0 & 4 & -10 \end{vmatrix} \xrightarrow[r_5 \to r_5 + 2r_4]{(EG)} \frac{1}{16} \cdot 3 \cdot \begin{vmatrix} -2 & 0 & 2 & 2 & -3 \\ 0 & -2 & 0 & 0 & -1 \\ 0 & 0 & 2 & 4 & 3 \\ 0 & 0 & 0 & -2 & -3 \\ 0 & 0 & 0 & 0 & -16 \end{vmatrix}.$$

A questo punto, il determinante di $\mathbf{A}$ si può calcolare in questo modo:

$$\det(\mathbf{A}) = \frac{3}{16} \prod_{i=1}^{5} a_{ii} = \frac{3}{16} \cdot (-2) \cdot (-2) \cdot (2) \cdot (-2) \cdot (-16) = 48.$$

4.4 Uso del Determinante per Risolvere i Sistemi Lineari

In questa parte vedremo come il determinante può essere usato per risolvere un sistema lineare ad n equazioni ed n incognite nella forma $\mathbf{Ax} = \mathbf{b}$, dove $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{K})$ e $\mathbf{b}, \mathbf{x} \in \mathbb{K}^n$.

Prima di arrivare a questo risultato (conosciuto come *teorema di Cramer*) abbiamo bisogno di definire il concetto di complemento algebrico e di matrice aggiunta.

Definizione 4.2 (Complemento algebrico)**.** Data una matrice quadrata $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{K})$, la quantità

$$a^\star_{ij} \doteq (-1)^{i+j} \det(\mathbf{A}_{ij})$$

è detta il *complemento algebrico* (o *cofattore*) dell'elemento a_{ij}, dove, ricordiamo $\mathbf{A}_{ij}$ denota la matrice $\mathbf{A}$ privata della $i-$sima riga e $j-$sima colonna.

Osservazione 4.4. Il determinante di una matrice può essere scritto come

$$\det \mathbf{A} = \sum_{j=1}^{n} a_{ij} a^\star_{ij}.$$

Definizione 4.3 (Matrice aggiunta)**.** Data una matrice quadrata $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{K})$, chiamiamo *matrice aggiunta* di $\mathbf{A}$ la matrice $\operatorname{adj}(\mathbf{A}) = \mathbf{A}^\star$ come la trasposta della matrice contenente i complementi algebrici di $\mathbf{A}$:

$$\operatorname{adj}(\mathbf{A}) = \mathbf{A}^\star = \left[a_{ij}\right]^\star_{i,j=1,2,\ldots,n} \doteq \left(\left[a^\star_{ij}\right]_{i,j=1,2,\ldots,n}\right)^\intercal = \left[a^\star_{ji}\right]_{i,j=1,2,\ldots,n}.$$

Vediamo un esempio di matrice aggiunta.

Esempio 4.3. Calcoliamo la matrice aggiunta di

$$\mathbf{A} = \begin{bmatrix} -2 & -1 & -2 \\ 1 & -2 & -1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Iniziamo calcolando i complementi algebrici:

$$a^\star_{11} = (-1)^{1+1}\begin{vmatrix} -2 & -1 \\ 0 & 0 \end{vmatrix} = 0, \qquad a^\star_{12} = (-1)^{1+2}\begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} = -1, \qquad a^\star_{13} = (-1)^{1+3}\begin{vmatrix} 1 & -2 \\ 1 & 0 \end{vmatrix} = 2,$$
$$a^\star_{21} = (-1)^{2+1}\begin{vmatrix} -1 & -2 \\ 0 & 0 \end{vmatrix} = 0, \qquad a^\star_{22} = (-1)^{2+2}\begin{vmatrix} -2 & -2 \\ 1 & 0 \end{vmatrix} = 2, \qquad a^\star_{23} = (-1)^{2+3}\begin{vmatrix} -2 & -1 \\ 1 & 0 \end{vmatrix} = -1,$$
$$a^\star_{31} = (-1)^{3+1}\begin{vmatrix} -1 & -2 \\ -2 & -1 \end{vmatrix} = -3, \qquad a^\star_{32} = (-1)^{3+2}\begin{vmatrix} -2 & -2 \\ 1 & -1 \end{vmatrix} = -4, \qquad a^\star_{33} = (-1)^{3+3}\begin{vmatrix} -2 & -1 \\ 1 & -2 \end{vmatrix} = 5.$$

Abbiamo quindi che la matrice aggiunta $\mathbf{A}^\star$ è

$$\mathbf{A}^\star = \begin{bmatrix} 0 & 0 & -3 \\ -1 & 2 & -4 \\ 2 & -1 & 5 \end{bmatrix}.$$

Dimostriamo il seguente risultato che permette di scrivere l'inversa di una matrice utilizzando la matrice aggiunta.

Proposizione 4.4. *Sia* $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{K})$ *una matrice invertibile, e sia* $\mathbf{A}^\star$ *la sua matrice aggiunta:* $\mathbf{A}^\star = \mathrm{adj}(\mathbf{A})$. *Allora vale la seguente identità:*

$$\mathbf{A}\mathbf{A}^\star = \det(\mathbf{A})\mathbf{Id}_n. \tag{20}$$

Dimostrazione. Iniziamo osservando che l'identità (20) equivale a

$$\sum_{h=1}^{n} a_{ih} a^\star_{jh} = \begin{cases} \det(\mathbf{A}) & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}.$$

Usando la definizione di complemento algebrico, questa equazione diventa

$$\sum_{h=1}^{n} a_{ih} \left((-1)^{h+j} \left|\mathbf{A}_{jh}\right|\right) = \begin{cases} \det(\mathbf{A}) & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}. \tag{21}$$

Ora, dimostriamo che questa equazione è vera.

Per $i = j$, il lato sinistro dell'equazione diventa

$$\sum_{h=1}^{n} \left((-1)^{h+i} a_{ih} |A_{ih}|\right),$$

che è esattamente il determinato di $\mathbf{A}$ (calcolato lungo la i−sima riga). Quindi, l'identità (21) è stata verificata per il caso $i = j$.

Per $i \neq j$ invece, l'equazione

$$\sum_{h=1}^{n} (-1)^{h+j} a_{ih} \left|\mathbf{A}_{jh}\right|$$

può essere interpretata come il calcolo, secondo la j−sima riga, del determinante della matrice ottenuta partendo da $\mathbf{A}$ e sostituendo la j−sima riga con una copia della i−sima:

$$\sum_{h=1}^{n} (-1)^{h+j} a_{ih} \left|\mathbf{A}_{jh}\right| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j-11} & a_{j-12} & \cdots & a_{j-1n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{j+11} & a_{j+12} & \cdots & a_{j+1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Poiché la riga i e j coincidono, esse sono linearmente dipendenti ed il determinante è nullo (proposizione 4.1). Questo dimostra che l'identità (20) è vera anche per $i \neq j$, concludendo la dimostrazione. □

Notiamo che l'identità (20) implica che l'inversa di una matrice si può calcolare come

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\mathbf{A}^{\star}.$$

Questi nuovi concetti e risultati ci permettono di enunciare e dimostrare il teorema di Cramer. Denotiamo con $\mathbf{A}(\mathbf{b}, j)$ la matrice $\mathbf{A}$ in cui la $j-$sima colonna di $\mathbf{A}$ è sostituita dal vettore dei termini noti $\mathbf{b}$.

Teorema 4.4 (Cramér)**.** *Sia* $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$ *una matrice quadrata invertibile e sia* $\mathbf{b} \in \mathbb{K}^n$ *un vettore con n componenti. Sia inoltre* $\mathbf{x} = [x_1, x_2, \ldots, x_n]^\top$ *il vettore delle soluzioni del sistema lineare* $\mathbf{Ax} = \mathbf{b}$*. Allora*

$$x_k = \frac{\det\big(\mathbf{A}(\mathbf{b}, k)\big)}{\det(\mathbf{A})}, \qquad k = 1, 2, \ldots, n. \tag{22}$$

Dimostrazione. Poiché, per ipotesi, la matrice $\mathbf{A}$ è invertibile, allora il sistema lineare $\mathbf{Ax} = \mathbf{b}$ ammette un'unica soluzione data da $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Dalla proposizione 4.4, abbiamo che $\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\mathbf{A}^{\star}$. Quindi, sviluppando il prodotto righe per colonne di

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \frac{1}{\det(\mathbf{A})}\mathbf{A}^{\star}\mathbf{b}$$

lungo la riga k abbiamo

$$\begin{aligned} x_k &= \frac{1}{\det(\mathbf{A})}\sum_{h=1}^{n} a^{\star}_{hk} b_h \\ &= \frac{1}{\det(\mathbf{A})}\sum_{h=1}^{n} (-1)^{h+k} |\mathbf{A}_{hk}|\, b_h \end{aligned}$$

da cui la tesi. □

Esempio 4.4. Consideriamo il sistema a due equazioni e due incognite

$$\begin{cases} -x_1 + x_2 = 1 \\ -2x_1 + x_2 = 1 \end{cases}.$$

Il sistema lineare si riscive come $\mathbf{Ax} = \mathbf{b}$ con

$$\mathbf{A} = \begin{bmatrix} -1 & 1 \\ -2 & 1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Allora, le matrici di interesse sono

$$\mathbf{A} = \begin{bmatrix} -1 & 1 \\ -2 & 1 \end{bmatrix}, \qquad \mathbf{A}(\mathbf{b}, 1) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad \mathbf{A}(\mathbf{b}, 2) = \begin{bmatrix} -1 & 1 \\ -2 & 1 \end{bmatrix};$$

da cui

$$\det(\mathbf{A}) = 1, \qquad \det(\mathbf{A}(\mathbf{b}, 1)) = 0, \qquad \det(\mathbf{A}(\mathbf{b}, 2)) = 1.$$

La soluzione è dunque:

$$x_1 = \frac{\det(\mathbf{A}(\mathbf{b}, 1))}{\det(\mathbf{A})} = 0, \qquad x_2 = \frac{\det(\mathbf{A}(\mathbf{b}, 2))}{\det(\mathbf{A})} = 1.$$

Esempio 4.5. Consideriamo il sistema lineare

$$\begin{cases} 2x_1 - x_2 - 2x_3 = 1 \\ -x_1 - x_2 + x_3 = 1 \\ -x_1 + 2x_2 - x_3 = 2 \end{cases},$$

in cui la matrice dei coefficienti **A** ed il vettore dei termini noti sono

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & -2 \\ -1 & -1 & 1 \\ -1 & 2 & -1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}.$$

Calcoliamo le matrici $\mathbf{A}(\mathbf{b}, i)$ per $i = 1, 2, 3$:

$$\mathbf{A}(\mathbf{b},1) = \begin{bmatrix} 1 & -1 & -2 \\ 1 & -1 & 1 \\ 2 & 2 & -1 \end{bmatrix}, \qquad \mathbf{A}(\mathbf{b},2) = \begin{bmatrix} ,2 & 1 & -2 \\ -1 & 1 & 1 \\ -1 & 2 & -1 \end{bmatrix}, \qquad \mathbf{A}(\mathbf{b},3) = \begin{bmatrix} 2 & -1 & 1 \\ -1 & -1 & 1 \\ -1 & 2 & 2 \end{bmatrix}.$$

Abbiamo quindi che la soluzione $\mathbf{x} = [x_1, x_2, x_3]^\top$ ha componenti

$$x_1 = \frac{\begin{vmatrix} 1 & -1 & -2 \\ 1 & -1 & 1 \\ 2 & 2 & -1 \end{vmatrix}}{\begin{vmatrix} 2 & -1 & -2 \\ -1 & -1 & 1 \\ -1 & 2 & -1 \end{vmatrix}} = \frac{-12}{6} = -2,$$

$$x_2 = \frac{\begin{vmatrix} 2 & 1 & -2 \\ -1 & 1 & 1 \\ -1 & 2 & -1 \end{vmatrix}}{\begin{vmatrix} 2 & -1 & -2 \\ -1 & -1 & 1 \\ -1 & 2 & -1 \end{vmatrix}} = \frac{-6}{6} - 1,$$

$$x_3 = \frac{\begin{vmatrix} 2 & -1 & 1 \\ -1 & -1 & 1 \\ -1 & 2 & 2 \end{vmatrix}}{\begin{vmatrix} 2 & -1 & -2 \\ -1 & -1 & 1 \\ -1 & 2 & -1 \end{vmatrix}} = \frac{-12}{6} = -2.$$

5 Autovettori, Autovalori, e Diagonalizzazione

In questa parte di corso, risponderemo alla seguente domanda:

> Dato un endomorfismo (ovvero un omomorfismo da uno spazio vettoriale in se stesso), è possibile trovare una base dello spazio rispetto alla quale la matrice associata all'endomorfismo sia una matrice diagonale?

5.1 Autovalori ed autovettori

Come prima cosa, introdurremo, in questa parte, i concetti di autovalore ed autovettore di un endomorfismo.

Consideriamo uno spazio vettoriale V di dimensione n, $\dim V = n$, su un campo $\mathbb{K}$, e sia $f : V \to V$ una funzione lineare. Sappiamo che, fissata una base di V, è possibile associare ad f una matrice quadrata di dimensione $n \times n$. Ovviamente, a basi diverse corrispondono matrici diverse. Ci chiediamo dunque se sia possibile trovare una base di V in modo tale che la matrice associata ad f sia una matrice diagonale. Affinchè ciò sia possibile, abbiamo bisogno di definire una base $\mathscr{B}_V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ di V tale per cui $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ per ogni $i = 1, 2, \dots, n$ e per opportuni scalari $\lambda_i \in \mathbb{K}$. La matrice associata ad una tale funzione lineare assumerebbe la seguente forma diagonale:

$$\mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}.$$

Possiamo quindi naturalmente dare la seguente definizione:

Definizione 5.1 (Endomorfismo diagonalizzabile)**.** Un endomorfismo $f : V \to V$ è detto *diagonalizzabile* se esiste una base di V tale che la matrice associata a f rispetto a tale base è una matrice diagonale.

Ricordando la definizione 2.6 di matrici simili, possiamo dare anche la seguente definizione.

Definizione 5.2 (Matrice diagonalizzabile)**.** Sia $\mathbf{A}$ una matrice quadrata, $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$. Questa matrice è detta *diagonalizzabile* se è simile ad una matrice diagonale, ovvero se esiste una matrice invertibile $\mathbf{S} \in \mathscr{M}_{n,n}(\mathbb{K})$ ed una matrice diagonale $\mathbf{D} \in \mathscr{M}_{n,n}(\mathbb{K})$ tale che $\mathbf{A} = \mathbf{SDS}^{-1}$.

Il prossimo passo consiste quindi nel cercare di determinare quali sono le condizioni per cui una matrice (o, equivalentemente, un endomorfismo) sia diagonalizzabile. Iniziamo dando la seguente definizione.

Definizione 5.3 (Autovalore, autovettore, e spettro)**.** Uno scalare $\lambda \in \mathbb{K}$ ed un vettore non nullo $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ si dicono, rispettivamente, *autovalore* ed *autovettore* per un endomorfismo $f : V \to V$ se $f(\mathbf{v}) = \lambda \mathbf{v}$.

Analogamente, uno scalare $\lambda \in \mathbb{K}$ ed un vettore non nullo $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ si dicono, rispettivamente, *autovalore* ed *autovettore* per una matrice quadrata $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$ se

$$\mathbf{Av} = \lambda \mathbf{v}. \tag{23}$$

Per entrambi i casi, lo *spettro* (di un endomorfismo o di una matrice) è l'insieme di tutti gli autovalori (dell'endomorfismo o della matrice).

Da ora in avanti, dato uno scalare $\lambda \in \mathbb{K}$ indicheremo con $\lambda : V \to V$ l'endomorfismo

$$\lambda \mathrm{id} : \mathbf{v} \mapsto \lambda \mathbf{v}.$$

Se consideriamo come base di V la base canonica, abbiamo che la matrice associata all'endomorfismo λ è

$$\lambda \mathbf{Id} = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda \end{bmatrix}.$$

Da ora in poi, parleremo solo di autovalori ed autovettori per una matrice, ma tutti i risultati che presentiamo sono validi anche per gli autovalori ed autovettori di una funzione lineare.

Notiamo ora che l'equazione (23) può essere riscritta come

$$(\mathbf{A} - \lambda \mathbf{Id})\mathbf{v} = \mathbf{0}.$$

Quindi, si ha che l'insieme degli autovettori relativi all'autovalore λ (insieme al vettore nullo) è il nucleo della matrice $\mathbf{A} - \lambda \mathbf{Id}$:

$$V_\lambda = \ker(\mathbf{A} - \lambda \mathbf{Id}) = \{\mathbf{v} \in V : \mathbf{A}\mathbf{v} = \lambda \mathbf{v}\}.$$

Da questa osservazione segue immediatamente il seguente risultato.

Proposizione 5.1. *Per ogni autovalore λ di una matrice $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{K})$ l'insieme V_λ degli autovettori relativi a λ,*

$$\{\mathbf{v} \in V : \mathbf{A}\mathbf{v} = \lambda \mathbf{v}\},$$

è un sottospazio vettoriale di V ed è detto autospazio *di $\mathbf{A}$ relativo a λ.*

Osservazione 5.1. Se una matrice $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{K})$ soddisfa il fatto che l'autospazio di tutti gli autovettori di $\mathbf{A}$ ha la stessa dimensione di $\mathbb{K}^n$,

$$\dim \left(\sum_{\lambda_i \text{ autovalori di } \mathbf{A}} V_{\lambda_i} \right) = \dim \left(\{\mathbf{v} \in \mathbb{K}^n : \exists \lambda, \mathbf{A}\mathbf{v} = \lambda \mathbf{v}\} \right) = n,$$

allora $\mathbf{A}$ è diagonalizzabile.

Equivalentemente, abbiamo che una matrice $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{K})$ è diagonalizzabile se e solo se esiste una base di $\mathbb{K}^n$ formata da autovettori di $\mathbf{A}$.

Prima di iniziare lo studio delle proprietà e dei risultati relativi agli autovalori ed autovettori, vediamo come determinarli. Iniziamo subito a notare che, affinché la matrice $\mathbf{A}$ ammetta degli autovalori, abbiamo bisogno che il nucleo di $\mathbf{A} - \lambda \mathbf{Id}$ non sia lo spazio nullo: $\ker(\mathbf{A} - \lambda \mathbf{Id}) \neq \{\mathbf{0}\}$. Questo equivale a dire che il determinante di $\mathbf{A} - \lambda \mathbf{Id}$ deve essere nullo.
Di conseguenza abbiamo che λ è un autovalore di $\mathbf{A}$ se e solo se

$$\det(\mathbf{A} - \lambda \mathbf{Id}) = 0.$$

Diamo quindo ora la seguente definizione, che ci permetterà di proseguire il ragionamento.

Definizione 5.4. Sia $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$ una matrice quadrata su un campo $\mathbb{K}$, e sia x un'incognita. Il *polinomio caratteristico* di $\mathbf{A}$ è

$$\mathscr{P}_{\mathbf{A}}(x) \doteq \det(\mathbf{A} - x\,\mathbf{Id}).$$

Il polinomio caratteristico è di fondamentale importanza in quanto non cambia tra matrici simili.

Proposizione 5.2. *Sia $f : V \to V$ un endomorfismo di V e siano $\mathbf{A}$ e $\mathbf{A}'$ due matrici associate ad f rispetto a due basi diverse (equivalentemente, siano $\mathbf{A}$ e $\mathbf{A}'$ due matrici simili). Allora $\mathscr{P}_{\mathbf{A}}(x) = \mathscr{P}_{\mathbf{A}'}(\lambda)$.*

Dimostrazione. Ricordiamo che due matrici $\mathbf{A}$ e $\mathbf{A}'$ sono simili se esiste una matrice invertibile $\mathbf{S}$ tale che $\mathbf{A}' = \mathbf{SAS}^{-1}$. In tal caso, si ha

$$\begin{aligned}
\det(\mathbf{A}' - x\,\mathbf{Id}) &= \det(\mathbf{SAS}^{-1} - x\,\mathbf{Id}) \\
&= \det(\mathbf{SAS}^{-1} - x\mathbf{SIdS}^{-1}) \\
&= \det(\mathbf{S}(\mathbf{A} - x\,\mathbf{Id})\mathbf{S}^{-1}) \\
&= \det\mathbf{S}\,\det(\mathbf{A} - x\,\mathbf{Id})\,\det(\mathbf{S}^{-1})
\end{aligned}$$

Poiché per una matrice invertibile $\mathbf{S}$ si ha $\det(\mathbf{S}^{-1}) = (\det\mathbf{S})^{-1}$ segue che

$$\det(\mathbf{A}' - x\,\mathbf{Id}) = \det(\mathbf{A} - \lambda\,\mathbf{Id}).$$

□

Grazie a questo risultato possiamo definire il polinomio caratteristico di un endomorfismo f come il polinomio caratteristico di una qualsiasi delle matrici associate all'endomorfismo stesso.

Quindi, abbiamo che λ è un autovalore di una matrice $\mathbf{A}$ se e solo se è una radice del polinomio caratteristico.

Esempio 5.1. Calcoliamo il polinomio caratteristico della matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & -2 \end{bmatrix}.$$

Calcoliamo dunque il polinomio caratteristico di $\mathbf{A}$:

$$\mathscr{P}_{\mathbf{A}}(x) = \begin{vmatrix} 1-\lambda & 0 \\ 1 & -2-\lambda \end{vmatrix} = (1-\lambda)(-2-\lambda) - 0\cdot 1 = -(1-\lambda)(2+\lambda)$$

Quindi, gli autovalori di $\mathbf{A}$ sono $\lambda_1 = +1$ e $\lambda_2 = -2$.

Osservazione 5.2. Si noti che, in genere, potrebbe succedere che gli autovalori di una matrice su un campo $\mathbb{K}$ non siano elementi dello stesso campo. Ad esempio, si consideri la matrice a coefficienti in $\mathbb{R}$

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Gli autovalori x di questa matrice devono soddisfare

$$\det(\mathbf{A} - x\,\mathbf{Id}) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1,$$

da cui si ricava che gli autovalori sono puramente immaginari

$$\lambda \in \{\pm\mathbf{i}\} \in \mathbb{C} \setminus \mathbb{R}.$$

Osservazione 5.3. La proposizoone 5.2 ci dice che due matrici simili hanno lo stesso polinomio caratteristico. L'implicazione inversa non è sempre vera. Infatti, le matrici $\mathbf{A}, \mathbf{B} \in \mathcal{M}_{n,n}(\mathbb{K})$ definite come

$$\mathbf{A} = \alpha \mathbf{Id} = \begin{bmatrix} \alpha & 0 & \cdots & 0 \\ 0 & \alpha & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \alpha \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} \alpha & 1 & 0 & \cdots & 0 \\ 0 & \alpha & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 & 0 \\ \vdots & & \ddots & \alpha & 1 \\ 0 & \cdots & \cdots & 0 & \alpha \end{bmatrix},$$

hanno lo stesso polinomio caratteristico,

$$\mathscr{P}_{\mathbf{A}}(\lambda) = \mathscr{P}_{\mathbf{B}}(\lambda) = (\alpha - \lambda)^n,$$

ma non sono simili in quando un multiplo della matrice identità è simile solo ad un multiplo della matrice identità.

Una volta che sono noti gli autovalori λ di una matrice $\mathbf{A}$, il calcolo degli autovettori $\mathbf{v}$ si riduce alla soluzione del sistema lineare

$$(\mathbf{A} - \lambda\, \mathbf{Id})\mathbf{v} = \mathbf{0}.$$

Esempio 5.2. Calcoliamo gli autovalori ed autovettori della matrice $\mathbf{A} \in \mathcal{M}_{3,3}(\mathbb{K})$ definita come

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & -1 & 3 \end{bmatrix}.$$

Iniziamo calcolando il polinomio caratteristico

$$\begin{aligned} \mathscr{P}_{\mathbf{A}}(\lambda) &= \begin{vmatrix} 3-\lambda & -1 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & -1 & 3-\lambda \end{vmatrix} \\ &= (3-\lambda)(2-\lambda)(3-\lambda) - (2-\lambda) \\ &= -(\lambda^3 - 8\lambda^2 + 20\lambda - 16) \\ &= -(\lambda - 2)^2(\lambda - 4) \end{aligned}$$

Quindi, gli autovalori sono $\lambda_1 = 2$ e $\lambda_2 = 4$.

Calcoliamo ora gli autovettori, partendo dall'autovalore λ_1:

$$\left[\mathbf{A} - \lambda_1 \mathbf{Id} \;\middle|\; \mathbf{0}\right] = \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 \end{array}\right] \xrightarrow[r_3 \to r_3 - r_1]{(EG)} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Gli autovettori relativi a λ_1 quindi hanno forma

$$\begin{bmatrix} \beta - \alpha \\ \beta \\ \alpha \end{bmatrix}, \qquad \alpha, \beta \in \mathbb{K}.$$

Di conseguenza, l'autospazio relativo a λ_1 è lo spazio

$$V_{\lambda_1} = \left\langle \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\rangle.$$

Passiamo ora al calcolo dell'autospazio relativo a λ_2:

$$\left[\mathbf{A} - \lambda_2 \mathbf{Id} \;\middle|\; \mathbf{0}\right] = \left[\begin{array}{ccc|c} -1 & -1 & 1 & 0 \\ 0 & -2 & 0 & 0 \\ 1 & -1 & -1 & 0 \end{array}\right] \xrightarrow[r_3 \to r_3 + r_1]{(EG)} \left[\begin{array}{ccc|c} -1 & -1 & 1 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & -2 & 0 & 0 \end{array}\right] \xrightarrow[r_3 \to r_3 - r_2]{(EG)} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Gli autovettori relativi a λ_2 hanno forma

$$\begin{bmatrix} \alpha \\ 0 \\ \alpha \end{bmatrix}, \qquad \alpha \in \mathbb{K};$$

da cui l'autospazio relativo a λ_2 è

$$V_{\lambda_2} = \left\langle \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\rangle.$$

Poiché $V_{\lambda_1} + V_{\lambda_2} = \mathbb{K}^n$, si ha che la matrice $\mathbf{A}$ è diagonalizzabile.

Nell'esempio precedente, lo zero $\lambda_1 = 2$ è doppio. È dunque naturale dare la prossima definizione.

Definizione 5.5 (Molteplicità algebrica e geometrica di un autovalore)**.** Sia $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$ una matrice quadrata su un campo $\mathbb{K}$ e sia $\mathscr{P}_{\mathbf{A}}(x)$ il suo polinomio caratteristico. Sia inoltre λ un autovalore di $\mathbf{A}$.
Allora, la *molteplicità algebrica* di λ è la molteplicità della radice λ per il polinomio $\mathscr{P}_{\mathbf{A}}(x)$, ovvero il più grande intero $m \in \mathbb{N}$ per cui il polinomio caratteristico $\mathscr{P}_{\mathbf{A}}(x)$ divide $(x - \lambda)^m$.
Inoltre, la *molteplicità geometrica* dell'autovalore λ è la dimensione dell'autospazio relativo a λ $V_\lambda \doteq \ker(\mathbf{A} - \lambda\, \mathbf{Id}_n)$.

Proposizione 5.3. *Sia* $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$ *una matrice quadrata su un campo* $\mathbb{K}$*. Allora autovettori relativi ad autovalori distinti sono linearmente indipendenti.*
In altre parole, siano $\lambda_1, \lambda_2, \ldots, \lambda_r$ *autovalori a due a due distinti di* $\mathbf{A}$*, e siano* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ *i relativi autovettori (ovvero il vettore* $\mathbf{v}_i$ *è l'autovettore relativo all'autovalore* λ_i*). Allora l'insieme* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ *sono linearmente indipendenti.*

◊ *Dimostrazione.* Dimostriamo il risultato per induzione sul valore r.

Il caso $n = 1$ è banale in quanto il vettore $\mathbf{v}_1$ è, per definizione, non nullo, ed è quindi linearmente indipendente.

Consideriamo adesso un generico r ed assumiamo che la tesi sia vera per $r - 1$ vettori. Consideriamo ora la combinazione lineare

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_{r-1} \mathbf{v}_{r-1} + \alpha_r \mathbf{v}_r = \mathbf{0}. \tag{24}$$

Poiché λ_i è l'autovalore relativo all'autovettore $\mathbf{v}_i$, abbiamo che la precedente equazione si può riscrivere come

$$\alpha_1\lambda_1\mathbf{v}_1 + \alpha_2\lambda_2\mathbf{v}_2 + \ldots + \alpha_{r-1}\lambda_{r-1}\mathbf{v}_{r-1} + \alpha_r\lambda_r\mathbf{v}_r = \mathbf{0}. \tag{25}$$

A questo punto, possiamo moltiplicare l'equazione (24) per λ_r e sottrarre (25), ottenendo

$$\alpha_1(\lambda_r - \lambda_1)\mathbf{v}_1 + \alpha_2\lambda_2\mathbf{v}_2 + \ldots + \alpha_{r-1}(\lambda_r - \lambda_{r-1})\mathbf{v}_{r-1} = \mathbf{0}.$$

Per l'ipotesi induttiva, quest'ultima equazione è soddisfatta solo nel caso i coefficienti α_i siano tutti nulli (in quanto le differenze $\lambda_r - \lambda_i$ sono sempre non nulle in quanto gli autovalori sono distinti). Di conseguenza, l'equazione (24) si può riscrivere come

$$\alpha_r\mathbf{v}_r = \mathbf{0}.$$

da cui segue $\alpha_r = 0$. Quindi, l'equazione (24) è soffisfatta solo quando tutti i coefficienti α_i sono nulli. Di conseguenza, i vettori $\mathbf{v}_i$, $i = 1, 2, \ldots, r$ sono linearmente indipendenti. □

La molteplicità algebrica e geometrica di un autovalore soddisfano la seguente proprietà.

Proposizione 5.4. *Sia $f : V \to V$ un endomorfismo su uno spazio vettoriale V di dimesnione finita su un campo $\mathbb{K}$, e sia $\lambda \in \mathbb{K}^n$ un autovalore di f con molteplicità algebrica pari ad m. Allora la molteplicità geometrica dell'autovalore è minore o uguale alla molteplicità algebrica:*

$$\dim V_\lambda \leq m.$$

Dimostrazione. Denotiamo con r la dimensione dell'autospazio relativo ad un autovalore λ di f, $r = \dim V_\lambda$, e consideriamo una base $\mathscr{B}_\lambda = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ di V_λ, e completiamola ad una base di V: $\mathscr{B}_V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ Allora, la matrice $\mathbf{A}$ associata ad f rispetto alla base $\mathscr{B}_V$ assume la seguente forma a blocchi

$$\mathbf{A} = \begin{bmatrix} \lambda\mathbf{Id}_r & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{bmatrix},$$

dove $\mathbf{B} \in \mathscr{M}_{r,n-r}(\mathbb{K})$, $\mathbf{C} \in \mathscr{M}_{n-r,n-r}(\mathbb{K})$ sono due generiche matrici, e $\mathbf{0} \in \mathscr{M}_{n-r,r}(\mathbb{K})$ è una matrice nulla.

Il polinomio caratteristico di $\mathbf{A}$ è dunque

$$\mathscr{P}_f(x) = (\lambda - x)^r \det(\mathbf{C} - x\,\mathbf{Id}).$$

Quindi, λ è una radice di molteplicità maggiore o uguale ad r del polinomio $\mathscr{P}_f(x)$. Da questo segue che $m \geq r = \dim V_\lambda$. □

Siamo ora in grado di dimostrare un risultato che da una condizione necessaria e sufficiente al fatto che un endomorfismo sia diagonalizzabile.

Teorema 5.1 (della diagonalizzazione)**.** *Sia $f : V \to V$ un endomorfismo di uno spazio vettoriale V di dimensione n su un campo $\mathbb{K}$. Siano $\lambda_1, \lambda_2, \ldots, \lambda_r$ gli autovalori dell'endomorfismo f e denotiamo con $m_1, m_2, \ldots, m_r$ le rispettive moltelicità algebriche. Allora, f è diagonalizzabile se e solo se*

1. *la somma delle molteplicità algebriche è pari alla dimensione di V: $m_1 + m_2 + \ldots + m_r = n = \dim V$;*

2. *per ogni autovalore λ_i, la sua molteplicità geometrica coincide con la molteplicità algebrica.*

Dimostrazione. Dimostriamo prima che un endomorfismo diagonalizzabile implichi i punti 1. e 2. Se f è diagonalizzabile, allora esiste una base di V rispetto alla quale la matrice associata ad f assume una forma a blocchi:

$$\begin{bmatrix} \lambda_1 \mathbf{Id}_{m_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \lambda_2 \mathbf{Id}_{m_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \lambda_r \mathbf{Id}_{m_r} \end{bmatrix}.$$

Utilizziamo questa matrice per calcolare il polinomio caratteristico di f, che è dunque pari a

$$\mathscr{P}_f(x) = (\lambda_1 - x)^{m_1}(\lambda_2 - x)^{m_2} \cdots (\lambda_r - x)^{m_r}.$$

Segue quindi che il grado di $\mathscr{P}_f$, che è pari ad n, è pari anche alla somma delle moteplicità algebriche m_i:

$$n = \deg \mathscr{P}_f = m_1 + m_2 + \ldots + m_r.$$

Inoltre, poiché f è diagonalizzabile, esiste una base di V costituita da autovettori di f. Quindi

$$\dim V_{\lambda_1} + \dim V_{\lambda_2} + \ldots \dim V_{\lambda_r} = n = \dim V.$$

Viceversa, supponiamo che esisteano r autovalori $\lambda_1, \lambda_2, \ldots, \lambda_r$ di f con molteplicità algebriche, rispettivamente, $m_1, m_2, \ldots, m_r$ tali che $m_1 + m_2 + \ldots + m_r = n$; e, che per ogni autovalore λ_i di f, la sua moltelicità algebrica e geometrica coincidano. Poiché, per ipotesi, $\dim V_{\lambda_i} = m_i$, allora esiste una base $\mathscr{B}_{\lambda_i} = \{\mathbf{v}_1^{(i)}, \mathbf{v}_2^{(i)}, \ldots, \mathbf{v}_{m_i}^{(i)}\}$ per ogni $i = 1, 2, \ldots, r$. Inoltre, poiché, per ipotesi, $m_1 + m_2 + \ldots m_r = n$, allora l'insieme

$$\left\{\mathbf{v}_1^{(1)}, \ldots, \mathbf{v}_{m_1}^{(1)}, \mathbf{v}_1^{(2)}, \ldots, \mathbf{v}_{m_2}^{(2)}, \ldots, \mathbf{v}_1^{(r)}, \ldots, \mathbf{v}_{m_r}^{(r)}\right\}$$

contiene esattamente n vettori. Inoltre, dalla proposizione 5.3, abbiamo che tutti questi vettori sono linearmente indipendenti. Di conseguenza, questo insieme costituisce una base di V. Esiste dunque una base di V formata da autovettori di f. Quindi, f è diagonalizzabile. □

Osservazione 5.4. Un endomorfismo $f : V \to V$ di uno spazio vettoriale di dimensione finita su un campo $\mathbb{K}$ è diagonalizzabile se e solo se f possiede autovalori $\lambda_1, \lambda_2, \ldots, \lambda_r$ tali per cui

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \ldots \oplus V_{\lambda_r}.$$

Se una matrice $\mathbf{A}$ è diagonalizzabile, allora la matrice invertibile $\mathbf{S} \in \mathscr{M}_{n,n}(\mathbb{K})$ ed $\mathbf{D} \in \mathscr{M}_{n,n}(\mathbb{K})$ tali per cui abbiamo (si veda la definizione 5.2) $\mathbf{A} = \mathbf{SDS}^{-1}$ sono, rispettivamente, la matrice le cui colonne sono gli autovalori di $\mathbf{A}$ e la matrice diagonale la cui diagonale è formata dagli autovalori di $\mathbf{A}$. Questo segue dal fatto che l'identità $\mathbf{Av} = \lambda\mathbf{v}$ equivale alla scrittura $\mathbf{AS} = \mathbf{DS}$ 'valutata' in una particolare colonna.

Ovviamente, l'ordine deve essere lo stesso per le due matrici (ovvero, la j−sima colonna di $\mathbf{S}$ deve essere l'autovettore relativo all'autovalore in posizione (j, j) di $\mathbf{D}$):

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}, \qquad \mathbf{S} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ \vdots & \vdots & & \vdots \end{bmatrix}.$$

Si noti che, nella scrittura sopra, $\mathbf{v}_i \neq \mathbf{v}_j$ anche nel caso in cui $\lambda_i = \lambda_j$. Più in generale, se $\mathbf{A}$ è diagonalizzabile, $\mathbf{S}$ deve essere invertibile. Da questo segue che nessuna delle colonne $\mathbf{v}_i$ è combinazione lineare delle altre.

Vediamo un esempio di diagonalizzazione di una matrice.

Esempio 5.3. Si consideri la matrice $\mathbf{A}$ definita come

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}.$$

Il polinomio caratteristico è $\mathscr{P}_{\mathbf{A}}(\lambda) = (2-\lambda)^2(1-\lambda)$. Di conseguenza, gli autovalori sono $\lambda_1 = 2$ con molteplicità algebrica pari a 2, e $\lambda_2 = 1$ con molteplicità algebrica pari ad 1.
Cercando gli autovettori, si ottengono i seguenti autospazi:

$$V_{\lambda_1} = \left\langle \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\rangle, \qquad V_{\lambda_2} = \left\langle \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\rangle.$$

Abbiamo quindi che entrambi gli autovalori hanno molteplicità algebrica e geometrica equivalenti. Quindi $\mathbf{A}$ è diagonalizzabile.
Per verificarlo definiamo le matrici $\mathbf{D}$ ed $\mathbf{S}$ come segue

$$\mathbf{D} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{S} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix},$$

ed osserviamo che

$$\mathbf{SDS}^{-1} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix} = \mathbf{A}.$$

Vediamo ora un esempio di una matrice non diagonalizzabile.

Esempio 5.4. Consideriamo la matrice

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & 0 \\ 0 & 2 & 0 \\ -1 & 0 & 3 \end{bmatrix},$$

i cui autovalori sono $\lambda_1 = 2$ con molteplicità algebrica pari a 2, e $\lambda_2 = 3$ con molteplicità algebrica pari a 1. L'autospazio V_{λ_1} è

$$V_{\lambda_1} = \left\langle \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\rangle.$$

Si ha quindi che la molteplicità algebrica di λ_1 è *strettamente* minore della sua molteplicità algebrica. Quindi $\mathbf{A}$ non è diagonalizzabile.

5.2 La Forma Canonica di Jordan

◇ Sia V uno spazio vettoriale di dimensione n sul campo dei complessi $\mathbb{C}$, e sia $f : V \to V$ una funzione lineare. Per ogni $\lambda \in \mathbb{C}$ indicheremo con $\lambda : V \to V$ l'applicazione lineare $\mathbf{v} \mapsto \lambda\mathbf{v}$ (ovvero $\lambda \equiv \lambda \mathrm{id}_V$). Rimarchiamo il fatto che questo è un abuso di notazione in quanto indicheremo con lo stesso simbolo uno scalare ed un'applicazione lineare. Tuttavia, il contesto permetterà di aver chiaro cosa intenderemo di volta in volta.

Ricordiamo che $\lambda \in \mathbb{C}$ è un autovalore di f se $\ker(f - \lambda) \neq \{\mathbf{0}\}$, e che un vettore non nullo $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ è un autovettore relativo all'autovalore λ se $\mathbf{v} \in \ker(f - \lambda)$, ovver, equivalentemente, se $f(\mathbf{v}) = \lambda\mathbf{v}$. Inoltre, ricordiamo che il sottospazio $\ker(f - \lambda)$ di V è detto l'*autospazio* di f relativo a λ. Infine, ricordiamo che tutti i risultati che presentermo relativamente agli omomorfismi si possono scrivere in modo identico per le matrici.

Generalizziamo ora la nozione di autovettore. Per ogni interro $m \in \mathbb{N}_{>0}$ indichiamo con $(f - \lambda)^m$ l'omomorfismo $f - \lambda$ composto con se stesso m volte:

$$(f - \lambda)^m \doteq \underbrace{(f - \lambda) \circ (f - \lambda) \circ \cdots \circ (f - \lambda)}_{m \text{ volte}}.$$

Definizione 5.6 (Autovettore generalizzato)**.** Sia $\lambda \in \mathbb{C}$ un autovalore di $f : V \to V$. Un vettore non nullo $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ è detto *autovettore generalizzato* di f, relativo all'autovalore λ se esiste un intero $m > 0$ tale che $\mathbf{v} \in \ker(f - \lambda)^m$.
Il minimo valore di m per cui $\mathbf{v} \in \ker(f - \lambda)^m$ è detto il *periodo* di $\mathbf{v}$ (nel caso in cui il periodo sia unitario, $m = 1$, $\mathbf{v}$ è un autovettore ordinario).

Si noti che valgono le seguenti inclusioni

$$\ker(f - \lambda) \subset \ker(f - \lambda)^2 \subset \ldots \subset \ker(f - \lambda)^m \subset \ldots \subset V.$$

Da ora in avanti, indicheremo con

$$V_\lambda = \bigcup_{m>0} \ker(f - \lambda)^m$$

il sottospazio di V costituito dagli autovettori generalizzati relativi all'autovalore λ.

Osservazione 5.5. Si noti che esiste un $r \in \mathbb{N}$ sufficientemente grande per cui $V_\lambda = \ker(f - \lambda)^r$.

Lemma 5.1. *Sia* $\mathbf{v}$ *un autovettore generalizzato per un omomorfismo f relativo all'autovalore λ, e sia m il periodo di* $\mathbf{v}$. *Allora gli m vettori*

$$\mathbf{v}, (f - \lambda)(\mathbf{v}), (f - \lambda)^2(\mathbf{v}), \ldots, (f - \lambda)^{m-1}(\mathbf{v})$$

sono linearmente indipendenti.

Dimostrazione. Iniziamo notando che l'ipotesi implica il fatto che $\mathbf{v} \in \ker(f-\lambda)^m$ ma $\mathbf{v} \notin \ker(f-\lambda)^{m-1}$. Consideriamo ora una combinazione lineare

$$\alpha_0\mathbf{v} + \alpha_1(f-\lambda)(\mathbf{v}) + \alpha_2(f-\lambda)^2(\mathbf{v}) + \ldots + (f-\lambda)^{m-1}(\mathbf{v}) = \mathbf{0} \tag{26}$$

ed applichiamo la funzione lineare $(f-\lambda)^{m-1}$ ad ambo i lati dell'equazione (ricordando che $(f-\lambda)^m(\mathbf{v}) = \mathbf{0}$), ottenendo

$$\alpha_0(f-\lambda)^{m-1}(\mathbf{v}) + \mathbf{0} = \mathbf{0}$$

da cui si ottiene $\alpha = 0$, ottenendo che l'equazione (26) si riduce a

$$\alpha_1(f-\lambda)(\mathbf{v}) + \alpha_2(f-\lambda)^2(\mathbf{v}) + \ldots + (f-\lambda)^{m-1}(\mathbf{v}) = \mathbf{0}.$$

Applichiamo ora a questa nuova identità la mappa lineare $(f-\lambda)^{m-2}$, ottenendo

$$\alpha_1(f-\lambda)^{m-1}(\mathbf{v}) + \mathbf{0} = \mathbf{0},$$

da cui segue $\alpha_1 = 0$.
Continuando in questo modo, si dimostra che tutti i coefficienti α_i sono nulli. □

Da questo risultato si ottiene il fatto che ogni autovettore generalizzato di $f : V \to V$ ha periodo minore od uguale alla dimensione dello spazio V. Infatti, se un autovettore generalizzato $\mathbf{v}$ avesse periodo $m > n = \dim V$, gli m vettori $\mathbf{v}, (f-\lambda)(\mathbf{v}), (f-\lambda)^2(\mathbf{v}), \ldots, (f-\lambda)^{m-1}(\mathbf{v})$ sarebbero linearmente indipendenti. Ma ciò è assurdo poiché $m > n = \dim V$. Quindi si ha

$$V_\lambda = \ker(f-\lambda)^n.$$

Il prossimo risultato è una generalizzazione della proposizione **??** per cui ad autovalori distinti corrispondono autovettori linearmente indipendenti.

Proposizione 5.5. *Siano V uno spazio vettoriale di dimensione n, ed $f : V \to V$ un endomorfismo. Siano inoltre $\lambda_1, \lambda_2, \ldots, \lambda_r$ autovalori di f a due a due distinti. Per ogni $i = 1, 2, \ldots, r$, sia $\mathbf{v}_i \in V$ un autovettore generalizzato relativo all'autovalore λ_i. Allora i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ sono linearmente indipendenti.*

Dimostrazione. Per ogni $i = 1, 2, \ldots, r$, indichiamo con m_i il periodo di $\mathbf{v}_i$ e consideriamo l'endomorfismo $(f-\lambda_i)^{m_i-1}$.
Osserviamo che, per ogni i e j, gli endomorfismi $(f-\lambda_i)^s$ e $(f-\lambda_j)^t$ commutano tra loro per ogni $s, t \geq 1$. Infatti, le potenze di f commutano tra loro ed f commuta con la moltiplicazione per ogni scalare λ (poiché f è lineare). Inoltre, per ogni $i = 1, 2, \ldots, r$, il vettore $\mathbf{w}_i = (f-\lambda_i)^{m_i-1}(\mathbf{v}_i)$ è un autovettore di f relativo all'autovalore λ_i. Si ha infatti

$$(f-\lambda_i)(\mathbf{w}_i) = (f-\lambda_i) \circ (f-\lambda_i)^{m_i-1}(\mathbf{v}_i) = (f-\lambda_i)^{m_i}(\mathbf{v}_i) = \mathbf{0}.$$

Da ciò discende che

$$(f-\lambda_j)^m(\mathbf{w}_i) = (\lambda_i - \lambda_j)^m \mathbf{w}_i,$$

per ogni $i, j = 1, 2, \ldots, r$ ed ogni $m \geq 1$.

Consideriamo ora una combinazione lineare

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r = \mathbf{0}, \tag{27}$$

ed applichiamo ad entrambi i lati dell'equazione l'endomorfismo

$$(f-\lambda_1)^{m_1-1} \circ (f-\lambda_2)^{m_2} \circ \ldots \circ (f-\lambda_r)^{m_r},$$

ottenendo

$$\alpha_1(\lambda_1-\lambda_2)^{m_2}(\lambda_1-\lambda_3)^{m_3}\cdots(\lambda_1-\lambda_r)^{m_r}(f-\lambda_1)^{m_1-1}(\mathbf{v}_1) = \mathbf{0}.$$

Poiché gli autovalori $\lambda_1, \lambda_2, \ldots, \lambda_r$ sono a due a due distinti, si ha quindi che $\alpha_1 = 0$.

La combinazione lineare (27) si riduce quindi a

$$\alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \ldots + \alpha_r\mathbf{v}_r = \mathbf{0}.$$

Applicando ora l'endomorfismo

$$(f-\lambda_2)^{m_2-1} \circ (f-\lambda_3)^{m_3} \circ \ldots \circ (f-\lambda_r)^{m_r},$$

si ottiene

$$\alpha_2(\lambda_2-\lambda_3)^{m_3}(\lambda_2-\lambda_4)^{m_4}\cdots(\lambda_2-\lambda_r)^{m_r}(f-\lambda_1)^{m_2-1}(\mathbf{v}_2) = \mathbf{0},$$

da cui anche il coefficiente α_2 è nullo: $\alpha_2 = 0$. Iterando in questo modo, si dimostra che tutti i coefficienti della combinazione lineare sono nulli: $\alpha_i = 0, i = 1, 2, \ldots, r$. □

Ricordiamo che un endomorfismo $f : V \to V$ potrebbe non essere diagonalizzabile, il che è equivalente al fatto che gli autovettori di f non generano lo spazio V. Per gli autovettori generalizzati, questo problema non si presenta.

Proposizione 5.6. *Sia V uno spazio vettoriale di dimensione finita sul campo dei compressi $\mathbb{C}$. Per ogni endomorfismo $f : V \to V$, esiste una base di V costituita da autovettori generalizzati di f.*

Dimostrazione. Dimostriamo per induzione sulla dimensione di V.

Se $\dim V = 1$, una funzione lineare $f : V \to V$ è la funzione di moltiplicazione per uno scalare. Quindi, ogni vettore $\mathbf{v} \in V$ è autovettore di f.

Supponiamo ora che il risultato sia vero per tutti gli spazi vettoriali di dimensione strettamente minore di $n = \dim V$. Sia $\lambda \in \mathbb{C}$ un autovalore di f e sia $V_\lambda = \ker(f-\lambda)^n$ il sottospazio di V costituito dagli autovettori generalizzati di f relativi all'autovalore λ. Se $V = V_\lambda$, la dimostrazione è conclusa. Altrimenti, chiamiamo $W = \operatorname{Im}(f-\lambda)^n$. Vogliamo dimostrare che $V = V_\lambda \oplus W_\lambda$.

Poiché $\dim V_\lambda + \dim W_\lambda = \dim V$, è sufficiente dimostrare, grazie alla formula di Grassmann 1.1, che l'intersezione è lo spazio nullo: $V_\lambda \cap W_\lambda = \{\mathbf{0}\}$. Consideriamo quindi un elemento $\mathbf{v}$ dell'intersezione, $\mathbf{v} \in V_\lambda \cap W_\lambda$. Poiché $\mathbf{v} \in W_\lambda = \operatorname{Im}(f-\lambda)^n$, si ha che esiste un vettore $\mathbf{u} \in V$ tale che $\mathbf{v} = (f-\lambda)^n(\mathbf{u})$. Ma, poiché $\mathbf{v} \in V_\lambda = \ker(f-\lambda)^n$, abbiamo che $(f-\lambda)^n(\mathbf{v}) = (f-\lambda)^{2n}(\mathbf{u}) = \mathbf{0}$. Questo significa che anche $\mathbf{u}$ è un autovettore generalizzato di f relativo all'autovalore λ. Inoltre, poiché il periodo di ogni autovettore generalizzato è minore o uguale alla dimensione n di V, si deve avere $(f-\lambda)^n(\mathbf{u})$, ovvero $\mathbf{v} = \mathbf{0}$.

Dimostriamo ora che $F(W_\lambda) \subset W_\lambda$. Sia $\mathbf{v} \in W_\lambda$, allora esiste un $\mathbf{u} \in V$ tale per cui $\mathbf{v} = (f-\lambda)^n(\mathbf{u})$. Si ha, poiché f commuta con $(f-\lambda)^n$, che vale la seguente identità:

$$f(\mathbf{v}) = f\big((f-\lambda)^n(\mathbf{u})\big) = (f-\lambda)^n\big(f(\mathbf{u})\big).$$

Ciò significa che $f(\mathbf{v}) \in W_\lambda$.

Consideriamo ora lo spazio vettoriale W_λ dotato della restrizione della funzione lineare f. Poiché $\dim W_\lambda < \dim V$ per l'ipotesi induttiva si ha che W_λ è generato da autovettori generalizzati di f. □

Quanto visto fin'ora si può riassumere nel seguente risultato.

Teorema 5.2. *Siano V uno spazio vettoriale ed $f : V \to V$ un endomorfismo. Siano $\lambda_1, \lambda_2, \ldots, \lambda_r$ gli autovalori di f, a due a due distinti. Sia inoltre*

$$\mathscr{P}(x) = (x-\lambda_1)^{\ell_1}(x-\lambda_2)^{\ell_2}\cdots(x-\lambda_r)^{\ell_r}$$

il polinomio caratteristico di f. Allora si ha:

1. *$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \ldots \oplus V_{\lambda_r}$, ove V_{λ_i} è il sottospazio vettoriale costituito dagli autovettori generalizzati di f relativi all'autovalore λ_i;*

2. *ogni V_{λ_i}* è stabile *per f, ovvero $f(V_{\lambda_i}) \subset V_{\lambda_i}$ per ogni $i = 1, 2, \ldots, r$;*

3. *$\dim V_{\lambda_i} = \ell_i$, da cui $V_{\lambda_i} = \ker(f-\lambda_i)^n = \ker(f-\lambda_i)^{\ell_i}$ per ogni $i = 1, 2, \ldots, r$.*

In particolare, il punto 3. afferma che il periodo di un autovettore generalizzato di f relativo all'autovalore λ_i è minore o uguale all'esponente ℓ_i con cui il fattore $(x-\lambda_i)$ compare nel polinomio caratteristico di f. Ovvero, il periodo di ogni autovettore generalizzato è minore o uguale alla molteplicità algebrica dell'autovalore corrispondente.

Dai punti 1. e 2. segue il fatto che, se scegliamo una base $\{\mathbf{w}_1^{(i)}, \mathbf{w}_2^{(i)}, \ldots, \mathbf{w}_{d_i}^{(i)}\}$ di V_{λ_i} per $i = 1, 2, \ldots, r$, allora l'insieme

$$\left\{\mathbf{w}_1^{(1)}, \mathbf{w}_2^{(1)}, \ldots, \mathbf{w}_{d_1}^{(1)}, \mathbf{w}_1^{(2)}, \mathbf{w}_2^{(2)}, \ldots, \mathbf{w}_{d_2}^{(2)}, \ldots, \mathbf{w}_1^{(r)}, \mathbf{w}_2^{(r)}, \ldots, \mathbf{w}_{d_r}^{(r)}\right\}$$

è una base di V rispetto alla quale la matrice di f è una matrice a blocchi della forma

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{A}_r \end{bmatrix},$$

ove ciascuna matrice $\mathbf{A}_i$ è la matrice della restrizione di f al sottospazio V_{λ_i} rispetto alla base $\{\mathbf{w}_1^{(i)}, \mathbf{w}_2^{(i)}, \ldots, \mathbf{w}_{d_i}^{(i)}\}$.

Concentriamoci ora sulla restrizione di f ad un singolo sottospazio V_{λ_i}. Sia $m_i \leq d_i = \dim V_{\lambda_i}$ il massimo dei periodi degli elementi di V_{λ_i}. Valgono le seguenti inclusioni (proprie)

$$\ker(f-\lambda_i) \subsetneq \ker(f-\lambda_i)^2 \subsetneq \ldots \subsetneq \ker(f-\lambda_i)^{m_i} = V_{\lambda_i}.$$

Possiamo quindi scegliere una base di V_{λ_i} come segue. Iniziamo scegliendo una base di $\ker(f-\lambda_i)$. Completiamola poi ad una base di $\ker(f-\lambda_i)^2$. Completiamo questa ad una base di $\ker(f-\lambda_i)^3$. Iteriamo così fino ad ottenere una base $\{\mathbf{w}_1^{(i)}, \mathbf{w}_2^{(i)}, \dots, \mathbf{w}_{d_i}^{(i)}\}$ di V_{λ_i}.

Osserviamo ora che se $\mathbf{v} \in \ker(f-\lambda_i)^k$ si ha $f(\mathbf{v}) = \lambda_i \mathbf{v} + (f-\lambda_i)(\mathbf{v})$, e $(f-\lambda_i)(\mathbf{v}) \in \ker(f-\lambda_i)^{k-1}$. Da ciò segue che, per ogni $j = 1, 2, \dots, d$,

$$f\left(\mathbf{w}_j^{(i)}\right) = \lambda_i \mathbf{w}_j^{(i)} + \sum_{h=1}^{j-1} \alpha_h \mathbf{w}_h^{(i)}.$$

Di conseguenza, la matrice $\mathbf{A}_i$ di f rispetto alla base $\{\mathbf{w}_1^{(i)}, \mathbf{w}_2^{(i)}, \dots, \mathbf{w}_{d_i}^{(i)}\}$ è una matrice triangolare superiore in cui tutti gli elementi sulla diagonale sono pari a λ_i:

$$\mathbf{A}_i = \begin{bmatrix} \lambda_i & \star & \cdots & \star \\ 0 & \lambda_i & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \cdots & 0 & \lambda_i \end{bmatrix}.$$

Osservazione 5.6. Quanto visto fino ad ora implica che ogni matrice quadrata sul campo dei complessi $\mathbb{C}$ è simile ad una matrice diagonale a blocchi

$$\begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{A}_r \end{bmatrix},$$

in cui i blocchi $\mathbf{A}_i$ sono le matrici traignolari superiori in cui gli elementi sulla diagonale principale sono gli autovalori λ_i di $\mathbf{A}$.

Possiamo tuttavia ottenere un risultato migliore, ma dobbiamo prima introdurre alcuni nuovi concetti.

Definizione 5.7 (Polinomio minimo)**.** Sia V uno spazio vettoriale ef $f : V \to V$ un endomorfismo. Sia inoltre $\mathscr{P}(x) = (x-\lambda_1)^{\ell_1}(x-\lambda_2)^{\ell_2} \cdots (x-\lambda_r)^{\ell_r}$ il polinomio caratteristico di f, ove gli autovalori λ_i sono a due a due distinti. Sia inoltre, per goni $i = 1, 2, \dots, r$, $m_i \leq \ell_i$ il massimo dei periodi degli elementi di V_{λ_i}. Allora il polinomio

$$\mathscr{Q}(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2} \cdots (x-\lambda_r)^{m_r}$$

è detto il *polinomio minimo* dell'endomorfismo f di V.

Si noti che, banalmente, le radici del polinomio minimo coincidono con le radici del polinomio caratteristico. Inoltre, il polinomio minimo divide il polinomio caratteristico.

Possiamo ora fornire un criterio di diagonalizzabilità di un endomorfismo in termini del suo polinomio minimo.

Proposizione 5.7. *Sia V uno spazio vettoriale e $f : V \to V$ un endomorfismo. Allora, f è diagonalizzabile se e solo se il suo polinomio minimo è prodotto di fattori lineari distinti.*

Dimostrazione. Dire che il polinomio minimo di f è prodotto di fattori lineari distinti equivale a dire che i periodi $m_1, m_2, \dots, m_r$ degli autovettori generalizzati sono tutti pari ad 1. Questo equivale a dire che gli autovettori generalizzati sono autovettori 'in senso classico'. Dal puto 1. del teorema 5.2 si deduce l'esistenza di una base di V costituita da autovettori di f, il che equivale a dire che f è diagonalizzabile. □

Possiamo ora dimostrare il seguente risultato.

Teorema 5.3 (Hamilton-Cayley)**.** *Siano V uno spazio vettoriale ed $f : V \to V$ un endomorfismo. Sia inoltre $\mathscr{Q}(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_r)^{m_r}$ il polinomio minimo di f. Allora l'endomorfismo $\mathscr{Q}(f)$ è l'endomorfismo nullo:*

$$\mathscr{Q}(f) = \mathbf{0}.$$

Dimostrazione. $\mathscr{Q}(f)$ è l'endomorfismo di V dato da

$$\mathscr{Q}(f) = (f-\lambda_1)^{m_1} \circ (f-\lambda_2)^{m_2} \circ \cdots \circ (f-\lambda_r)^{m_r}.$$

Se $\mathbf{v} \in V_{\lambda_i}$ per un qualche $i = 1, 2, \dots, r$, si ha $(f-\lambda_i)^{m_i}(\mathbf{v}) = \mathbf{0}$, da cui $\mathscr{Q}(f)(\mathbf{v}) = \mathbf{0}$. Poiché ogni vettore di V si può scrivere come combinazione lineare di vettori appartenenti ai vari sottospazi V_{λ_i}, si ha $Q(f)(\mathbf{v}) = \mathbf{0}$. □

Da questo teorema segue immediatamente il seguente corollario.

Corollario 5.1. *Sia $\mathbf{A}$ una matrice quadrata di dimensione n, e sia $\mathscr{Q}(x)$ il suo polinomio minimo. Allora $\mathscr{Q}(\mathbf{A}) = 0$. Inoltre, anche il polinomio caratteristico di $\mathbf{A}$, $\mathscr{P}(x)$, applicato ad $\mathbf{A}$ è la matrice nulla: $\mathscr{P}(\mathbf{A}) = \mathbf{0}$.*

Dimostrazione. La prima affermazione è ovvia poiché ad ogni endomorfismo si può associare una matrice quadrata.

La seconda affermazione segue dal fatto che il polinomio minimo $\mathscr{Q}(x)$ divide il polinomio caratteristico $\mathscr{P}(x)$. □

Vediamo ora come possiamo migliorare la scelta della base di ciascun sottospazio V_{λ_i} al fine di semplificare ulteriormente la forma delle matrici $\mathbf{A}_i$ che costituiscono i blocchi diagonali della matrice $\mathbf{A}$ di f.

Focalizziamoci su un singolo autovalore che indichiamo semplicemente come λ, e sul relativo sottospazio di autovettori generalizzati V_λ. Denotiamo con m il massimo dei periodi degli elementi di V_λ e consideriamo la catena di inclusioni (proprie)

$$\ker(f-\lambda) \subsetneq \ker(f-\lambda)^2 \subsetneq \dots \subsetneq \ker(f-\lambda)^m = V_\lambda.$$

Per ogni $j = 1, 2, \dots, m$, denotiamo con d_j la dimensione del nucleo dell'omomorfismo $(f-\lambda)$ composto j volte con se stesso: $d_j = \dim\ker(f-\lambda)^j$.

Consideriamo ora un vettore non nullo

$$\mathbf{v} \in \ker(f-\lambda)^m \setminus \ker(f-\lambda)^{m-1}.$$

Allora, gli m vettori

$$\mathbf{w}_m = \mathbf{v}, \quad \mathbf{w}_{m-1} = (f-\lambda)(\mathbf{v}), \quad \mathbf{w}_{m-2} = (f-\lambda)^2(\mathbf{v}), \quad \dots, \quad \mathbf{w}_1 = (f-\lambda)^{m-1}(\mathbf{v}),$$

sono linearmente indipendenti (dal lemma 5.1). Notiamo inoltre che

$$(f-\lambda)(\mathbf{w}_1) = (f-\lambda)^m(\mathbf{v}) = \mathbf{0},$$

ovvero

$$f(\mathbf{w}_1) = \lambda \mathbf{w}_1,$$

mentre

$$(f-\lambda)(\mathbf{w}_j) = (f-\lambda)(f-\lambda)^{m-j}(\mathbf{v}) = (f-\lambda)^{m-j+1}(\mathbf{v}) = \mathbf{w}_{j-1},$$

cioè, per $j = 2, 3, \ldots, m$

$$f(\mathbf{w}_j) = \lambda \mathbf{w}_j + \mathbf{w}_{j-1}.$$

Ciò significa che f induce un endomorfismo del sottospazio di V_λ, generato dai vettori $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m$, la cui matrice rispetto alla base $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ di tale sottospazio è la seguente matrice quadrata di dimensione m:

$$\mathbf{J}_\lambda = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix}.$$

Una matrice di questa forma è detta *blocco di Jordan* di ordine m relativo all'autovalore λ.

Sia ora $s = d_m - d_{m-1}$. Ricordiamo che $s \geq 1$. Vogliamo dimostrare che la matrice della restrizione di f al sottospazio V_λ contiene, lungo la diagonale principale, s blocchi di Jordan J_λ di ordine m. A tal fine consideriamo s vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s \in \ker(f-\lambda)^m \setminus \ker(f-\lambda)^{m-1}$ tali per cui

$$\ker(f-\lambda)^m = \ker(f-\lambda)^{m-1} \oplus \langle \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s \rangle .$$

Applicando a ciasuno dei vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ il procedimento descritto in precedenza, otteniamo il seguente insieme di vettori

$$\begin{array}{lllll} \mathbf{v}_1, & (f-\lambda)(\mathbf{v}_1), & (f-\lambda)^2(\mathbf{v}_1), & \ldots, & (f-\lambda)^{m-1}(\mathbf{v}_1), \\ \mathbf{v}_2, & (f-\lambda)(\mathbf{v}_2), & (f-\lambda)^2(\mathbf{v}_2), & \ldots, & (f-\lambda)^{m-1}(\mathbf{v}_2), \\ & & \vdots & & \\ \mathbf{v}_s, & (f-\lambda)(\mathbf{v}_s), & (f-\lambda)^2(\mathbf{v}_s), & \ldots, & (f-\lambda)^{m-1}(\mathbf{v}_s), \end{array} \tag{28}$$

dove ciascuna delle righe corrisponde ad un blocco di Jordan di dimensione m.

Dobbiamo ora dimostrare che tutti questi vettori sono linearmente indipendenti. Consideriamo a tal fine una combinazione lineare

$$\begin{array}{l} \alpha_1^{(0)}\mathbf{v}_1 + \alpha_2^{(0)}\mathbf{v}_2 + \ldots + \alpha_s^{(0)}\mathbf{v}_s + \\ \alpha_1^{(1)}(f-\lambda)(\mathbf{v}_1) + \alpha_2^{(1)}(f-\lambda)(\mathbf{v}_2) + \ldots + \alpha_s^{(1)}(f-\lambda)(\mathbf{v}_s) + \\ \alpha_1^{(2)}(f-\lambda)^2(\mathbf{v}_1) + \alpha_2^{(2)}(f-\lambda)^2(\mathbf{v}_2) + \ldots + \alpha_s^{(2)}(f-\lambda)^2(\mathbf{v}_s) + \\ \qquad \ldots \\ \alpha_1^{(m-1)}(f-\lambda)^{m-1}(\mathbf{v}_1) + \alpha_2^{(m-1)}(f-\lambda)^{m-1}(\mathbf{v}_2) + \ldots + \alpha_s^{(m-1)}(f-\lambda)^{m-1}(\mathbf{v}_s) = \mathbf{0}. \end{array}$$

Se ora teniamo il termine $\alpha_1^{(0)}\mathbf{v}_1 + \alpha_2^{(0)}\mathbf{v}_2 + \ldots + \alpha_s^{(0)}\mathbf{v}_s$ a sinistra e portiamo il resto alla destra dell'uguale, si deduce che

$$\alpha_1^{(0)}\mathbf{v}_1 + \alpha_2^{(0)}\mathbf{v}_2 + \ldots + \alpha_s^{(0)}\mathbf{v}_s \in \ker(f-\lambda)^{m-1},$$

da cui segue

$$\alpha_1^{(0)} = \alpha_2^{(0)} = \ldots = \alpha_s^{(0)} = 0.$$

La combinazione lineare si riduce quindi a

$$\begin{array}{c} \alpha_1^{(1)}(f-\lambda)(\mathbf{v}_1) + \alpha_2^{(1)}(f-\lambda)(\mathbf{v}_2) + \ldots + \alpha_s^{(1)}(f-\lambda)(\mathbf{v}_s) + \\ \alpha_1^{(2)}(f-\lambda)^2(\mathbf{v}_1) + \alpha_2^{(2)}(f-\lambda)^2(\mathbf{v}_2) + \ldots + \alpha_s^{(2)}(f-\lambda)^2(\mathbf{v}_s) + \\ \ldots \\ \alpha_1^{(m-1)}(f-\lambda)^{m-1}(\mathbf{v}_1) + \alpha_2^{(m-1)}(f-\lambda)^{m-1}(\mathbf{v}_2) + \ldots + \alpha_s^{(m-1)}(f-\lambda)^{m-1}(\mathbf{v}_s) = \mathbf{0}. \end{array} .$$

Procedendo in modo identico a prima, si ha

$$\alpha_1^{(1)}(f-\lambda)(\mathbf{v}_1) + \alpha_2^{(1)}(f-\lambda)(\mathbf{v}_2) + \ldots + \alpha_s^{(1)}(f-\lambda)(\mathbf{v}_s) \in \ker(f-\lambda)^{m-2},$$

da cui

$$\alpha_1^{(1)}(\mathbf{v}_1) + \alpha_2^{(1)}(\mathbf{v}_2) + \ldots + \alpha_s^{(1)}(\mathbf{v}_s) \in \ker(f-\lambda)^{m-1}.$$

Di conseguenza, anche i coefficienti $\alpha_j^{(1)}$, $j = 1, 2, \ldots, s$ sono nulli:

$$\alpha_1^{(1)} = \alpha_2^{(1)} = \ldots = \alpha_s^{(1)} = 0.$$

Proseguendo in questo modo, si ottiene che tutti i coefficienti $\alpha_j^{(i)}$ sono nulli:

$$\alpha_j^{(i)} = 0, \qquad \forall\, i = 1, 2, \ldots, m-1, \quad \forall\, j = 1, 2, \ldots, s,$$

dimostrando l'indipendenza lineare.

Osserviamo ora che se $d_m - d_{m-1} = s$ allora $d_i - d_{i-1} \geq s$ per ogni $i = 1, 2, \ldots, m$.[3] Infatti, dato che i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ hano periodo m, i vettori $(f-\lambda)^{m-1}(\mathbf{v}_1), (f-\lambda)^{m-1}(\mathbf{v}_2), \ldots, (f-\lambda)^{m-1}(\mathbf{v}_s)$ hanno periodi i e, come abbiamo appena dimostrato, sono linearmente indipendenti.

Se accade che $d_i - d_{i-1} = s$ per ogni $i = 1, 2, \ldots, m$, si ha

$$\begin{aligned} \dim\ker(f-\lambda) &= s, \\ \dim\ker(f-\lambda)^2 &= 2s, \\ &\vdots \\ \dim\ker(f-\lambda)^m &= m\,s, \end{aligned}$$

[3] indichiamo con $d_0 = \dim\ker(f-\lambda) = 0$.

e, dato che $V_\lambda = \ker(f-\lambda)^m$, gli $m\,s$ vettori in (28) sono una base di V_λ. Abbiamo quinci che la matrice della restrizione di f a V_λ consiste di s blocchi di Jordan di ordine m:

$$\begin{bmatrix} \mathbf{J}_\lambda & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_\lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{J}_\lambda \end{bmatrix}.$$

Se invece di ha $d_i - d_{i-1} > s$ per qualche i, indichiamo con j il massimo indice minore di m tale che $d_j - d_{j-1} = t > s$ e prendiamo dei vettori $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{t-s}$ tali che si abbia

$$\ker(f-\lambda)^j = \ker(f-\lambda)^{j-1} \oplus \left\langle (f-\lambda)^{m-1}(\mathbf{v}_1), (f-\lambda)^{m-1}(\mathbf{v}_2), \dots, (f-\lambda)^{m-1}(\mathbf{v}_s) \right\rangle \oplus \langle \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{t-s} \rangle .$$

Ragionando in mdoo simile a prima, si dimostra che i vettori

$$\begin{array}{lllll} \mathbf{w}_1, & (f-\lambda)(\mathbf{w}_1), & (f-\lambda)^2(\mathbf{w}_1), & \dots, & (f-\lambda)^{j-1}(\mathbf{w}_1), \\ \mathbf{w}_2, & (f-\lambda)(\mathbf{w}_2), & (f-\lambda)^2(\mathbf{w}_2), & \dots, & (f-\lambda)^{j-1}(\mathbf{w}_2), \\ & & \vdots & & \\ \mathbf{w}_{t-s}, & (f-\lambda)(\mathbf{w}_{t-s}), & (f-\lambda)^2(\mathbf{w}_{t-s}), & \dots, & (f-\lambda)^{j-1}(\mathbf{w}_{t-s}) \end{array}$$

sono linearmente indipendenti. Questo implica che la matrice della restrizione di f a V_λ contiene anche, lungo la diagonale principale, $t-s$ blocchi di Jordan relativi all'autovalore λ, di dimensione j (oltre agli s blocchi $\mathbf{J}_\lambda$ di ordine m già menzionati).

Procedendo in modo analogo con gli autovettori generalizzati di periodo via via minore, si arriva a concludere che, in generale, la matrice della restrizione di f a V_λ è una matrice diagonale a blocchi in cui i blocchi diagonali sono dei blocchi di Jordan relativi all'autovalore λ di dimensione minore o uguale ad m, ove m è il massimo dei periodi degli elementi di V_λ. Tale intero m è l'esponente con cui il fattore $(x-\lambda)$ compare nel polinomio minimo di f.

Abbiamo quindi dimostrato il seguente risultato.

Teorema 5.4 (Jordan)**.** *Sia V uno spazio vettoriale di dimensione n sul campo $\mathbb{C}$ e sia $f : V \to V$ un endomorfismo. Sia*

$$\mathscr{P}(x) = (x-\lambda_1)^{\ell_1}(x-\lambda_2)^{\ell_2}\cdots(x-\lambda_r)^{\ell_r}$$

il polinomio caratteristico di f, in cui gli autovalori $\lambda_1, \lambda_2, \dots, \lambda_r$ sono a due a due distinti. Allora, esiste una base di V rispetto alla queale la matrice di f è una matrice a blocchi

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{J}_r \end{bmatrix}$$

ove ogni blocco diagonale $\mathbf{J}_i$ è, a sua volta, una matrice a blocchi del tipo

$$\mathbf{J}_i = \begin{bmatrix} \mathbf{J}^1_{\lambda_i} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}^2_{\lambda_i} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{J}^{s_i}_{\lambda_i} \end{bmatrix}$$

in cui $\mathbf{J}^1_{\lambda_i}, \mathbf{J}^2_{\lambda_i}, \dots, \mathbf{J}^{s_i}_{\lambda_i}$ *sono opportuni blocchi di Jordan relativi all'autovalore* λ_i.

Una matrice $\mathbf{J}$ di questo tipo è detta *forma canonica di Jordan* e una base di V rispetto a cui f ha questa matrice è detta *base di Jordan*.

Esempio 5.5. Sia V uno spazio vettoriale di dimensione 5 su $\mathbb{C}$, e sia $\mathscr{B}_V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_5\}$ una base di V. Sia $f : V \to V$ l'applicazione lineare di matrice

$$\begin{bmatrix} 1 & 0 & -2 & 0 & 0 \\ -2 & 3 & -2 & 0 & 0 \\ 2 & 0 & 5 & 0 & 0 \\ 0 & 2 & 2 & 2 & -1 \\ 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

rispetto alla base $\mathscr{B}_V$. Il polinomio caratteristico di f è

$$\mathscr{P}(x) = \det(\mathbf{A} - x\,\mathbf{Id}_5) = -(x-3)^5.$$

Di conseguenza, f ha un unico autovalore $\lambda = 3$ con molteplicità pari a 5. Il polinomio minimo $\mathscr{Q}(x)$ deve dividere $\mathscr{P}(x)$ ed ha quindi forma

$$\mathscr{Q}(x) = (x-3)^m, \quad 1 \le m \le 5.$$

Osserviamo ora che

$$\mathbf{A} - 3\mathbf{Id} = \begin{bmatrix} -2 & 0 & -2 & 0 & 0 \\ -2 & 0 & -2 & 0 & 0 \\ 2 & 0 & 2 & 0 & 0 \\ 0 & 2 & 2 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \quad (\mathbf{A} - 3\mathbf{Id})^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & -2 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 \end{bmatrix}, \quad (\mathbf{A} - 3\mathbf{Id})^3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Quindi, il polinomio minimo è

$$\mathscr{Q}(x) = (x-3)^3.$$

Questo significa che il massimo periodo degli autovettori generalizzati (relativi all'unico autovalore) di f è $m = 3$. Quindi, abbiamo la catena di inclusioni proprie

$$\ker(f-3) \subsetneq \ker(f-3)^2 \subsetneq \ker(f-3)^3 = V.$$

Si ha inoltre

$$\begin{aligned} &\dim\ker(f-3) = 5 - \operatorname{rk}(\mathbf{A} - 3\mathbf{Id}) = 5 - 3 = 2, \\ &\dim\ker(f-3)^2 = 5 - \operatorname{rk}(\mathbf{A} - 3\mathbf{Id})^2 = 5 - 1 = 4, \\ &\dim\ker(f-3)^3 = \dim V = 5. \end{aligned}$$

Ponendo $d_i = \dim]ker(f-3)^i$, l'intero $s = d_m - d_{m-1} = d_3 - d_2$ è pari a $s = 5 - 4 = 1$. Di conseguenza, la matrice di Jordan $\mathbf{J}$ contiene un blocco di Jordan di dimensione 3, relativo all'autovalore $\lambda = 3$,

$$\begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}.$$

Poiché abbiamo che $\dim\ker(f-3)^2 - \dim\ker(f-3) = 4-2 = 2$, abbiamo $j = 2$ e $d_2 - d_1 = t = 2 > s = 1$, da cui $t - s = 1$ e la matrice di Jordan contiene anche un blocco di Jordan di ordine 2

$$\begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}.$$

In conclusione, la matrice di Jordan di f è

$$\mathbf{J} = \left[\begin{array}{ccc|cc} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{array} \right].$$

Determiniamo ora la base $\mathscr{B}_W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_5\}$ di V rispetto alla quale la matrice di f sia $\mathbf{J}$. Scegliamo il vettore

$$\mathbf{v}_2 \in \ker(f-3)^3 \setminus \ker(f-3)^2$$

e poniamo

$$\mathbf{w}_3 = \mathbf{v}_2, \quad \mathbf{w}_2 = (f-3)(\mathbf{v}_2) = 2\mathbf{v}_4, \quad \mathbf{w}_1 = (f-3)^2(\mathbf{v}_2) = -2\mathbf{v}_4 + 2\mathbf{v}_5.$$

L'insieme $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ formano la parte di base di V responsabile della presenza del blocco di Jordan di dimensione 3.

Per continuare scegliamo un vettore $\mathbf{u}$ tale che

$$\ker(f-3)^2 = \ker(f-3) \oplus \left\langle (f-3)^{3-2}(\mathbf{v}_2) \right\rangle \oplus \langle \mathbf{u} \rangle.$$

Ad esempio, $\mathbf{u} = \mathbf{v}_1$. Quindi, i due vettori mancanti per completare la base $\mathscr{B}_W$ sono

$$\mathbf{w}_5 = \mathbf{v}_1, \quad \mathbf{w}_4 = (f-3)(\mathbf{v}_1) = -2\mathbf{v}_1 - 2\mathbf{v}_2 + 2\mathbf{v}_3.$$

A Esercizi

A.1 Spazi Vettoriali

Esercizio A.1. Si dimostri che l'insieme $2\mathbb{Z} = \{2k : k \in \mathbb{Z}\} = \{0, 2, -2, 4, -4\}$ è un gruppo rispetto alla somma.

Esercizio A.2. Si dimostri che i seguenti insiemi sono spazi vettoriali rispetto alla classica operazione di somma e di moltiplicazione per scalare:

1. L'insieme dei vettori $\mathbf{v} = [0, a, b]$ con $a, b \in \mathbb{R}$;
2. L'insieme dei vettori $\mathbf{v} = [a, b, a+b]$ con $a, b \in \mathbb{R}$;
3. L'insieme dei vettori $\mathbf{v} = [x, y, z, w]$ le cui componenti soddisfano il sistema

$$\begin{cases} x + z = y + w \\ z + w = 0 \end{cases}.$$

Si dimostri invece che linsieme dei vettori $\mathbf{v} = [x, y, z, w]$ le cui componenti soddisfano il sistema

$$\begin{cases} x + z = y + w \\ z + w = 1 \end{cases}$$

non formano uno spazio vettoriale.

Esercizio A.3. Si dica, motivando, se i seguenti insiemi sono linearmente dipendenti o indipendenti. Se sono linearmente dipendenti, si estragga un sottoinsieme di vettori linearmente indipendenti.

$$V_1 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\};$$

$$V_2 = \left\{ \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \right\};$$

$$V_3 = \left\{ \mathscr{P}_1(x) = 1 + x, \mathscr{P}_2(x) = x + x^2, \mathscr{P}_3(x) = 1 + x^2 \right\}.$$

Esercizio A.4. Si considerino i seguenti sottospazi vettoriali di $V = \mathbb{R}^3$:

$$U_1 = \left\{ \begin{bmatrix} a \\ 2a \\ b \end{bmatrix} : a, b \in \mathbb{R} \right\}, \quad U_2 = \left\{ \begin{bmatrix} c \\ c \\ 0 \end{bmatrix} : c \in \mathbb{R} \right\}, \quad U_3 = \left\{ \begin{bmatrix} d \\ e \\ d \end{bmatrix} : d, e \in \mathbb{R} \right\}, \quad U_4 = \left\{ \begin{bmatrix} f \\ 0 \\ f \end{bmatrix} : f \in \mathbb{R} \right\}.$$

Allora

1. Per ogni coppia di questi sottospazi si trovi l'intersezione, la somma, e si dica se la somma è diretta;
2. Per ogni sottospazio, si definisca una base.

Esercizio A.5. Si considerino i seguenti sottospazi vettoriali di $V = \mathbb{R}^3$:

$$U_1 = \left\langle \begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\2\\0\end{bmatrix} \right\rangle, \qquad U_2 = \left\langle \begin{bmatrix}2\\3\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix} \right\rangle, \qquad U_3 = \left\langle \begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix} \right\rangle, \qquad U_4 = \left\langle \begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}2\\2\\1\end{bmatrix} \right\rangle.$$

Allora

1. Per ogni coppia di questi sottospazi si trovi l'intersezione, la somma, e si dica se la somma è diretta;
2. Per ogni sottospazio, si definisca una base.

Esercizio A.6. Per ognuno dei seguenti sistemi lineari S_i, si consideri lo spazio dei vettori delle soluzioni $W_i = \{\mathbf{v}_i = [x, y, z, w] : x, y, z, w \text{ sono soluzioni di } S_i\}$. Per ognuno di questi spazi, si definisca una base $\mathscr{B}_i$ di W_i:

$$S_1 : \begin{cases} x + y + z = 0 \\ x - y + 2z + w = 0 \end{cases}, \qquad S_2 : \begin{cases} x + y + 2z + w = 0 \\ x - y - z + w = 0 \end{cases}, \qquad S_3 : \begin{cases} x + y + 2z - w = 0 \\ x - y - z + 2w = 0 \end{cases},$$

$$S_4 : \begin{cases} x - y + z - w = 0 \\ x + y - z + w = 0 \\ x + 2y - 2z + 2w = 0 \end{cases}, \qquad S_5 : \begin{cases} x - y - z + w = 0 \\ x + 2y - 2z + w = 0 \\ x + y - z + 2w = 0 \end{cases}, \qquad S_6 : \begin{cases} x + 2y + 3z - w = 0 \\ 2x + 2y - z + w = 0 \\ 3x + y - 2z + 2w = 0 \end{cases}.$$

A.2 Applicazioni Lineari e Matrici

Esercizio A.7. Date le seguenti funzioni lineari $f_i : \mathbb{R}^4 \to \mathbb{R}^4$, si verifichi se la funzione è lineare. Se lo è, se ne trovi il nucleo e l'immagine, e si verifichi il teorema di nullità più rango.

$$f_1 : \begin{bmatrix}x\\y\\z\\w\end{bmatrix} \mapsto \begin{bmatrix}x+y\\2y+w\\x\\x+z+w\end{bmatrix}, \quad f_2 : \begin{bmatrix}x\\y\\z\\w\end{bmatrix} \mapsto \begin{bmatrix}x+2z\\y+2z\\x\\x+z+w\end{bmatrix}, \quad f_3 : \begin{bmatrix}x\\y\\z\\w\end{bmatrix} \mapsto \begin{bmatrix}x+1\\x+y\\z\\w\end{bmatrix}, \quad f_4 : \begin{bmatrix}x\\y\\z\\w\end{bmatrix} \mapsto \begin{bmatrix}x-y\\z\\w\\z-w\end{bmatrix}.$$

Esercizio A.8. Date le seguenti funzioni lineari $f_i : \mathbb{R}^3 \to \mathbb{R}^2$, si verifichi se la funzione è lineare. Se lo è, se ne trovi il nucleo e l'immagine, e si verifichi il teorema di nullità più rango.

$$f_1 : \begin{bmatrix}x\\y\\z\end{bmatrix} \mapsto \begin{bmatrix}x+y\\y+z\end{bmatrix}, \qquad f_2 : \begin{bmatrix}x\\y\\z\end{bmatrix} \mapsto \begin{bmatrix}x\\x\end{bmatrix}, \qquad f_3 : \begin{bmatrix}x\\y\\z\end{bmatrix} \mapsto \begin{bmatrix}z\\x+y+z\end{bmatrix}, \qquad f_4 : \begin{bmatrix}x\\y\\z\end{bmatrix} \mapsto \begin{bmatrix}z+1\\x+y+z\end{bmatrix}.$$

Esercizio A.9. Per ognuna dei seguenti omomorfismi, si determini se esso è iniettivo, suriettivo, o biiettivo:

$$f_1 : \mathbb{R}^2 \to \mathbb{R}^3, \qquad f_1 : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x \\ y \\ x+y \end{bmatrix},$$

$$f_2 : \mathbb{R}^3 \to \mathbb{R}^2, \qquad f_2 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} x+y+z \\ y+z \end{bmatrix},$$

$$f_3 : \mathbb{R}^3 \to \mathbb{R}^3, \qquad f_3 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} y+z \\ x+z \\ x+y \end{bmatrix},$$

$$f_4 : \mathbb{R}^2 \to \mathbb{R}^3, \qquad f_4 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} x+y+z \\ x+y \\ z \end{bmatrix}.$$

Esercizio A.10. Si determini per quali valori di $t \in \mathbb{R}$ esiste una funzione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ tale che

$$f\left(\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \qquad f\left(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} t \\ 2t \\ 1 \end{bmatrix}, \qquad \ker f = \mathscr{L}\left(\left\{\begin{bmatrix} 1 \\ t-2 \\ 2 \end{bmatrix} : t \in \mathbb{R}\right\}\right).$$

Esercizio A.11. Si determini per quali valori di $t \in \mathbb{R}$ esiste una funzione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ tale che

$$f\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} t \\ 2t \\ 1 \end{bmatrix}, \qquad f\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \qquad \ker f = \mathscr{L}\left(\left\{\begin{bmatrix} 0 \\ -t^2 \\ 1 \end{bmatrix} : t \in \mathbb{R}\right\}\right).$$

Esercizio A.12. Si determini per quali valori di $t \in \mathbb{R}$ esiste una funzione lineare $f : \mathbb{R}^3 \to \mathbb{R}^3$ tale che

$$f\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \qquad f\left(\begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}, \qquad \ker f = \mathscr{L}\left(\left\{\begin{bmatrix} 1 \\ t^2 \\ 1 \end{bmatrix} : t \in \mathbb{R}\right\}\right).$$

Esercizio A.13. Si calcolino le matrici associate alle funzioni lineari $f_i : V \to W$, $V = W = \mathbb{R}^3$

$$f_1 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} x+y+z \\ x-y-z \\ 2x-y \end{bmatrix}, \qquad f_2 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} x+2y \\ -y-z \\ x-3y \end{bmatrix}, \qquad f_3 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} 2y+z \\ 2x-2y-z \\ -y-2z \end{bmatrix}.$$

rispetto alle basi

$$\mathscr{B}_V = \left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right\}, \qquad \mathscr{B}_W = \left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right\}.$$

Esercizio A.14. Si calcolino le matrici associate alle funzioni lineari $f_i : V \to W$, $V = \mathbb{R}^3, W = \mathbb{R}^2$

$$f_1 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} x+y+z \\ x-y-z \end{bmatrix}, \qquad f_2 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} x+y \\ x-z \end{bmatrix}, \qquad f_3 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} y+z \\ x-y+z \end{bmatrix}.$$

rispetto alle basi

$$\mathscr{B}_V = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}, \qquad \mathscr{B}_W = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}.$$

Esercizio A.15. Si calcolino le matrici associate alle funzioni lineari $f_i : V \to W$, $V = \mathbb{R}^2, W = \mathbb{R}^3$

$$f_1 : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x \\ x+y \\ x+2y \end{bmatrix}, \qquad f_2 : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} y \\ x \\ 2x+y \end{bmatrix}, \qquad f_3 : \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x+y \\ x+2y \\ 3x+2y \end{bmatrix}.$$

rispetto alle basi

$$\mathscr{B}_V = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}, \qquad \mathscr{B}_W = \left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \right\}.$$

Esercizio A.16. Si calcolino i seguenti prodotti tra matrici:

$$\begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 3 & 4 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 5 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 1 & 0 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 & 2 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 \\ 0 & 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 & 1 \\ 2 & 1 & 4 & 1 \\ 2 & 1 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 3 & 1 & 1 \\ 2 & 1 & 4 & 1 \\ 2 & 1 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 \\ 0 & 1 & 2 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 1 & 3 & 2 \\ 2 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 0 & 1 & 3 \\ 4 & 2 & 2 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 6 \\ 3 & 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 4 & 2 & 4 \\ 2 & 3 & 3 & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

Esercizio A.17. Si calcolino il nucleo e l'immagine delle seguenti matrici:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 2 & 2 \\ 1 & 1 & 3 \end{bmatrix},$$

Esercizio A.18. Si calcolino le matrici del cambio di base tra le seguenti coppie di basi

1. $\mathscr{B}_V = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}, \qquad \mathscr{B}'_V = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\};$

2. $\mathscr{B}_V = \left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}, \qquad \mathscr{B}'_V = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\};$

3. $\mathscr{B}_V = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}, \qquad \mathscr{B}'_V = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$

Esercizio A.19. Si consideri il diagramma a destra. Siano $V = \mathbb{R}^3$ e $W = \mathbb{R}^2$. Sia $f : V \to W$ l'omomorfismo

$$f : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} x+y \\ y-z \end{bmatrix}.$$

$$\begin{tikzcd}
\mathbb{K}^n \arrow[rrr, "F_A"] \arrow[dd, "F_P"'] & & & \mathbb{K}^m \arrow[dd, "F_Q"] \\
& V \arrow[ul, "\alpha"'] \arrow[dl, "\alpha'"] \arrow[r, "f"] & W \arrow[ur, "\beta"] \arrow[dr, "\beta'"'] & \\
\mathbb{K}^n \arrow[rrr, "F_{A'}"'] & & & \mathbb{K}^m
\end{tikzcd}$$

Siano $\alpha, \alpha' : V \to \mathbb{R}^3$ gli ismomorfismi che ad ogni vettore $\mathbf{v} \in V$ associa le coordinate rispetto alle due basi

$$\mathscr{B}_V = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}, \qquad \mathscr{B}'_V = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

Siano infine $\beta, \beta' : W \to \mathbb{R}^2$ gli ismomorfismi che ad ogni vettore $\mathbf{w} \in W$ associa le coordinate rispetto alle due basi

$$\mathscr{B}_W = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}, \qquad \mathscr{B}'_W = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}.$$

Si calcolino le matrici $\mathbf{A}, \mathbf{A}'$ dell'omomorfismo f rispetto alle basi $\mathscr{B}_V, \mathscr{B}_W$ e $\mathscr{B}'_V, \mathscr{B}'_W$ rispettivamente; le matrici $\mathbf{P}$ e $\mathbf{Q}$ del cambio di base tra $\mathscr{B}_v, \mathscr{B}'_V$ e $\mathscr{B}_W, \mathscr{B}'_W$ rispettivamente; e si verifichi l'identità $\mathbf{A}'\mathbf{P} = \mathbf{Q}\mathbf{A}$.

A.3 Sistemi Lineari

Esercizio A.20. Si risolvano i seguenti sistemi lineari utilizzando EG

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

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

Esercizio A.21. Si calcoli, tramite EG, il rango delle seguenti matrici:

$$\mathbf{A} = \begin{bmatrix} -1 & 2 & 2 & 1 & -1 \\ -1 & -2 & 0 & -2 & -1 \\ -2 & 0 & 2 & -1 & -2 \\ -2 & 2 & -1 & -1 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & -1 & -1 & -2 & 0 \\ 0 & 2 & -2 & -1 & 0 \\ 0 & 1 & -1 & -1 & -2 \\ 2 & 0 & 1 & -2 & -1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 2 & -1 & 0 & -1 & 1 & 2 \\ 1 & -2 & 0 & -1 & 1 & 1 \\ -1 & 1 & 1 & -2 & 0 & 2 \\ -2 & 2 & 1 & 1 & -1 & 1 \end{bmatrix},$$

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & 1 & 2 & -1 \\ -2 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ -1 & -1 & 0 & 2 & 2 \\ 1 & 1 & 2 & 0 & 1 \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} -1 & -2 & 2 & 2 & -1 \\ -1 & 0 & -2 & -1 & 1 \\ -2 & 2 & -2 & 2 & -1 \\ 0 & -2 & 2 & 2 & -1 \\ -1 & 0 & 0 & 2 & -1 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} -2 & 1 & 2 & 0 & -2 \\ -2 & 0 & -2 & -1 & 0 \\ 2 & -1 & 2 & 2 & -1 \\ -1 & -2 & -1 & 1 & 0 \\ 1 & 1 & -2 & 0 & -1 \end{bmatrix}$$

Esercizio A.22. Calcolare l'inversa delle seguenti matrici:

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 2 & -1 \\ 0 & -1 & 2 & 0 \\ -2 & -1 & 0 & 0 \\ 2 & -1 & 1 & -2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 2 & 0 & 0 \\ 2 & -2 & -1 & 2 \\ 1 & 1 & 0 & -1 \\ -1 & 1 & 0 & -2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 2 & 2 & 0 & 2 \\ 0 & -1 & -1 & 0 \\ 2 & -2 & -2 & -2 \\ 1 & -2 & -2 & -2 \end{bmatrix},$$

$$\mathbf{D} = \begin{bmatrix} 1 & -2 & 0 & 2 \\ 1 & 2 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 2 & 0 & -2 & 0 \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} -1 & 1 & -2 & -2 \\ -1 & -2 & 1 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & -2 & 0 & 2 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & -2 & 2 & 1 \\ -2 & 2 & -1 & -1 \\ 2 & 2 & 1 & 1 \end{bmatrix}$$

A.4 Determinante

Esercizio A.23. Calcolare il determinante det delle seguenti matrici 2×2:

$$\begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}, \quad \begin{bmatrix} 2 & 2 \\ 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 4 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix}, \quad \begin{bmatrix} \lambda & -\lambda \\ \mu & \mu \end{bmatrix}.$$

Esercizio A.24. Calcolare il determinante det delle seguenti matrici 3×3 utilizzando la formula di Sarrus:

$$\begin{bmatrix} 2 & 1 & 2 \\ 1 & 3 & 3 \\ 1 & 3 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 2 & 2 \\ -1 & -1 & -3 \\ -1 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & 2 \\ 2 & 0 & -2 \\ -1 & -1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & -2 & -3 \\ -3 & -3 & 2 \\ 2 & 3 & 2 \end{bmatrix},$$

$$\begin{bmatrix} 1 & -3 & 0 \\ 3 & -2 & 1 \\ -3 & 1 & -2 \end{bmatrix}, \quad \begin{bmatrix} -1 & 3 & -2 \\ -2 & 3 & 2 \\ 3 & 2 & -2 \end{bmatrix}, \quad \begin{bmatrix} 0 & -3 & 0 \\ 3 & 2 & -3 \\ -1 & 2 & 1 \end{bmatrix}, \quad \begin{bmatrix} -2 & 1 & -1 \\ -1 & -2 & 3 \\ -3 & 3 & -2 \end{bmatrix}.$$

Esercizio A.25. Si calcolino i seguenti determinanti

$$
\begin{vmatrix} -2 & -1 & 0 & -1 & 1 \\ -2 & 1 & -1 & 1 & 1 \\ 1 & 2 & 2 & -2 & 2 \\ -3 & 2 & 3 & -2 & -1 \\ 2 & -1 & -3 & 2 & 0 \end{vmatrix}, \qquad
\begin{vmatrix} -2 & -2 & -2 & 0 & 0 \\ 1 & -2 & 0 & 2 & -2 \\ 1 & -1 & -2 & 0 & -2 \\ 0 & 0 & 1 & 1 & -2 \\ 3 & -3 & -1 & -2 & 3 \end{vmatrix}, \qquad
\begin{vmatrix} -2 & 0 & 1 & -1 & 1 \\ -1 & 0 & 3 & 1 & -3 \\ 0 & 2 & 1 & 1 & -1 \\ 3 & -3 & 2 & 2 & -2 \\ 1 & 2 & -2 & 3 & -2 \end{vmatrix},
$$

$$
\begin{vmatrix} -2 & -3 & -1 & 0 & -1 \\ 0 & -3 & 3 & 3 & -3 \\ -3 & 1 & -3 & -3 & 0 \\ 3 & -3 & 3 & -2 & 2 \\ 3 & 0 & -2 & -2 & 0 \end{vmatrix}, \qquad
\begin{vmatrix} -2 & -1 & 0 & -1 & 1 \\ -2 & 1 & -1 & 1 & 1 \\ 1 & 2 & 2 & -2 & 2 \\ -3 & 2 & 3 & -2 & -1 \\ 2 & -1 & -3 & 2 & 0 \end{vmatrix}, \qquad
\begin{vmatrix} 0 & -2 & 2 & 3 & 3 \\ -2 & 0 & 0 & -3 & -1 \\ 2 & 1 & -2 & 0 & -3 \\ 3 & -3 & 0 & 3 & -3 \\ 3 & -2 & -2 & -3 & 1 \end{vmatrix}.
$$

A.5 Autovettori, Autovalori, e Diagonalizzazione

Esercizio A.26. Per le seguenti matrici, si calcolino gli autovalori e gli autovettori. Inoltre, si determino le molteplicità algebriche e geometriche degli autovalori, e si dica se la matrice è diagonalizzabile. Se lo è, si usi il fatto che $\left(\mathbf{PDP}^{-1}\right)^n = \mathbf{PD}^n\mathbf{P}^{-1}$, per calcolare la potenza quinta della matrice.

$$
\mathbf{A} = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix}, \qquad
\mathbf{B} = \begin{bmatrix} 0 & 2 & -1 \\ 1 & 1 & -1 \\ -1 & -2 & 0 \end{bmatrix}, \qquad
\mathbf{C} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 3 \end{bmatrix},
$$

$$
\mathbf{D} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 3 & -2 & 1 & 1 \\ -3 & -1 & -4 & -1 \\ -6 & -2 & -2 & -5 \end{bmatrix}, \qquad
\mathbf{E} = \begin{bmatrix} 2 & 0 & 0 & 4 \\ 0 & 3 & 0 & 5 \\ 0 & 5 & -2 & 5 \\ 4 & 0 & 0 & 2 \end{bmatrix}, \qquad
\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}
$$

$$
\mathbf{G} = \begin{bmatrix} 3 & 12 & -21 \\ -1 & -6 & 13 \\ 0 & -2 & 6 \end{bmatrix}, \qquad
\mathbf{H} = \begin{bmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, \qquad
\mathbf{I} = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}.
$$

B Notazione Matematica

$x \in A$	:	l'elemento x appartiene all'insieme A
$B \subset A$	:	l'insieme B è un sottoinsieme dell'insieme A
$A \times B$	:	prodotto cartesiano tra l'insieme A e l'insieme B
$A \cap B$	:	intersezione tra l'insieme A e l'insieme B
$A \cup B$	:	unione tra l'insieme A e l'insieme B
$A \setminus B$	:	differenza insiemistica tra A e B
$f : A \to B$	:	la funzione f va dall'insieme A all'insieme B
$f : x \mapsto f(x)$	:	la funzione f mappa l'elemento x nell'elemento $f(x)$
$\sum_{i=1}^{n} x_i$	:	sommatoria di tutti gli elementi x_i $(x_1 + x_2 + \ldots + x_n)$
$f \circ g$	:	composizione della funzione f con la funzione g

Indice