Università LUISS Guido Carli, A.A. 2010-2011 (II semestre)

Metodi Matematici per la Finanza

Dott. Davide Vergni

Introduzione all'algebra lineare

Indice

Capitolo 1

Introduzione all'algebra lineare

1.0.1 Obiettivi principali del capitolo

In questo capitolo verranno introdotti gli elementi principali di algebra lineare, ossia lo spazio dei vettori (generalizzandolo successivamente agli spazi vettoriali) e l'algebra delle matrici. Nel formalismo matematico moderno, la costruzione di una teoria parte dalla definizione degli oggetti e dalle operazioni che possono essere effettuate sugli stessi. La struttura di una teoria matematica non deriva tanto dalla tipologia degli oggetti sui quali si lavora, quanto dalle operazioni che è possibile effettuare sugli stessi.

1.0.2 Notazione

Lettere minuscole in grassetto (i.e., $\mathbf{u}$) rappresentano vettori, lettere maiuscole con cappello (i.e., $\hat{A}$) rappresentano matrici, lettere maiuscole in semigrassetto (i.e., $\mathbb{V}$) rappresentano spazi vettoriali.

1.1 Spazio dei vettori

Il primo oggetto matematico che definiamo sarà il vettore. In quanto estensione di un oggetto fisico ben rappresentato nella realtà del mondo che ci circonda, esistono varie definizioni tutte equivalenti di vettore.

1.1.1 Definizione di vettore

- Partendo dai numeri reali come rappresentazione dei punti della retta reale $\mathbb{R}$, possiamo definire i vettori come rappresentazione dei punti del piano $\mathbb{R}^2$ o dello spazio $\mathbb{R}^3$.

- Abbiamo esperienza dalla vita di tutti i giorni dei concetti di velocità, di gravità, di forza. Ognuna di queste grandezze può essere identificata da una intensità, una direzione e un verso, in altre parole un vettore rappresentato da un segmento orientato la cui lunghezza è la sua intensità.

- Un'applicazione di tipo economico può essere quella che identifica lo stato economico di un paese con un vettore con tante componenti quante sono le variabili di interesse (PIL, inflazione, tasso ufficiale di sconto, ecc. ecc.).

- Una definizione più astratta e analitica è quella che identifica i vettori come elementi dell'insieme $\mathbb{R}^n$ ottenuto come prodotto cartesiano ($\times$) di n insiemi $\mathbb{R}$:

  $$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} \qquad n \text{ volte}$$

  Pertanto, indicheremo con $\mathbb{R}^n$ lo spazio dei vettori con n componenti (o anche *in n dimensioni*) e indicheremo il generico vettore $\mathbf{u} \in \mathbb{R}^n$ con

  $$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \qquad \text{ovvero} \qquad \mathbf{u} = \{u_i\} \qquad \text{con } i = 1 \dots n\,.$$

  Un vettore è quindi identificato dall'insieme (ordinato) dei suoi elementi (o componenti). Per selezionare l'elemento $i-$esimo del vettore, verrà usata la notazione

  $$(\mathbf{u})_i = u_i\,.$$

  Una volta fissati gli elementi, sarà fissato anche il vettore. Nell'effettuare le operazioni con i vettori, spesso verranno date le regole di calcolo per le componenti.

Due vettori sono uguali se hanno lo stesso numero di componenti (cioè la stessa dimensione) e se le componenti di uguale posto coincidono. Dati $\mathbf{u}$ e $\mathbf{v} \in \mathbb{R}^n$ scriviamo quindi $\mathbf{u} = \mathbf{v}$ se $\forall\ i = 1 \dots n\ \ u_i = v_i$. Se, per qualche i, succede che $u_i \neq v_i$ allora i due vettori saranno diversi e scriveremo $\mathbf{u} \neq \mathbf{v}$.

1.1.2 Operazioni con i vettori

Passiamo quindi alla definizione di alcune operazioni elementari che ci permetteranno di lavorare con i vettori.

Somma

La somma tra due vettori è la prima operazione che viene introdotta e può essere vista come un'applicazione che associa a due elementi di $\mathbb{R}^n$ un terzo elemento

$$+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$$

ottenuto sommando tutte le componenti corrispondenti:

Definizione 1.1.1. *Presi due qualsiasi vettori* $\mathbf{u}$ *e* $\mathbf{v} \in \mathbb{R}^n$ *definiamo il vettore* $\mathbf{w} = \mathbf{u} + \mathbf{v} \in \mathbb{R}^n$ *come*

$$(\mathbf{w})_i = (\mathbf{u} + \mathbf{v})_i = u_i + v_i\,. \tag{1.1}$$

O, in notazione estesa

$$\mathbf{w} = \mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1 + u_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix}$$

Questa definizione di somma corrisponde esattamente alla regola di somma geometrica dei vettori nota come regola del parallelogramma.

Le proprietà della somma così definite sono

s1) Associatività: $\forall\, \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$:

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$$

s2) Commutatività: $\forall\, \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$:

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$$

s3) Esistenza dell'elemento neutro: $\exists!\, \mathbf{0} \in \mathbb{R}^n \,|\, \forall\, \mathbf{u} \in \mathbb{R}^n$:

$$\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}\,.$$

Il vettore $\mathbf{0}$ è il vettore che ha tutti gli elementi pari a 0.

s4) Esistenza dell'elemento opposto: $\forall\, \mathbf{u} \in \mathbb{R}^n\ \exists!\, (-\mathbf{u}) \in \mathbb{R}^n$

$$\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$$

Prodotto per scalare

Un'altra operazione fondamentale per lavorare con i vettori è il prodotto per scalare, ossia la moltiplicazione di un numero reale (uno scalare) per un vettore

$$\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n\,.$$

Questa operazione si effettua moltiplicando per lo scalare tutte le componenti del vettore:

Definizione 1.1.2. *Preso un qualsiasi vettore $\mathbf{u} \in \mathbb{R}^n$ e un qualsiasi scalare $\alpha \in \mathbb{R}$ definiamo il vettore $\mathbf{w} = \alpha \cdot \mathbf{u} \in \mathbb{R}^n$ (o più semplicemente $\mathbf{w} = \alpha\mathbf{u}$) come*

$$(\mathbf{w})_i = (\alpha\mathbf{u})_i = \alpha u_i\,. \tag{1.2}$$

O, in notazione estesa:

$$\alpha\mathbf{u} = \begin{pmatrix} \alpha u_1 \\ \alpha u_2 \\ \vdots \\ \alpha u_n \end{pmatrix}$$

Questa definizione di prodotto per scalare corrisponde geometricamente all'espansione (o alla contrazione se $|\alpha| < 1$) dell'ampiezza del vettore $\mathbf{v}$. Se $\alpha < 0$ allora cambierà anche il verso del vettore.

Le proprietà del prodotto per scalare appena definito sono

m1) Associatività: $\forall\, \alpha, \beta \in \mathbb{R}$ e $\forall\, \mathbf{u} \in \mathbb{R}^n$:

$$\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u} = \alpha\beta\mathbf{u}$$

m2) Commutatività: $\forall\, \alpha \in \mathbb{R}$ e $\forall\, \mathbf{u} \in \mathbb{R}^n$:

$$\alpha\mathbf{u} = \mathbf{u}\alpha$$

m3) Esistenza dell'elemento neutro: $\exists!\, 1 \in \mathbb{R} \,|\, \forall\, \mathbf{u} \in \mathbb{R}^n$:

$$1\mathbf{u} = \mathbf{u}$$

m4) Costruzione dell'elemento opposto: $\forall\, \mathbf{u} \in \mathbb{R}^n\ \exists!\, (-1) \in \mathbb{R}$:

$$-1\mathbf{u} + \mathbf{u} = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$$

m5) Distributività rispetto alla somma tra scalari: $\forall\, \alpha, \beta \in \mathbb{R}$ e $\forall\, \mathbf{u} \in \mathbb{R}^n$:

$$(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$$

m6) Distributività rispetto alla somma tra vettori: $\forall\, \alpha \in \mathbb{R}$ e $\forall\, \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$:

$$\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$$

1.1.3 Combinazioni lineari

Con le operazioni appena definite sullo spazio vettoriale $\mathbb{R}^n$, somma tra vettori e prodotto per scalari, possiamo costruire l'operazione fondamentale che si effettua negli spazi vettoriali, la combinazione lineare. Presi due qualsiasi vettori $\mathbf{u}$ e $\mathbf{v} \in \mathbb{R}^n$ e due qualsiasi scalari α e $\beta \in \mathbb{R}$ il vettore

$$\mathbf{w} = \alpha\mathbf{u} + \beta\mathbf{v}$$

si dice combinazione lineare di $\mathbf{u}$ e $\mathbf{v}$ con coefficienti α e β.
Questa definizione può essere facilmente generalizzata a k vettori con k diversi coefficienti:

Definizione 1.1.3. *Presi k vettori qualsiasi $\mathbf{u}_k \in \mathbb{R}^n$ e k numeri qualsiasi $\alpha_i \in \mathbb{R}$, con $i = 1 \ldots k$, definiamo combinazione lineare di $\mathbf{u}_i$ con coefficienti α_i il vettore*

$$\mathbf{v} = \sum_{i=1}^{k} \alpha_i \mathbf{u}_i = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots \alpha_k\mathbf{u}_k \tag{1.3}$$

Lo spazio $\mathbb{R}^n$ è chiuso rispetto alle operazioni di somma e prodotto per scalare, ossia all'operazione di combinazione lineare. Questo significa che presi dei vettori di $\mathbb{R}^n$ ed effettuando con essi delle combinazioni lineari, il vettore risultante appartiene ancora a $\mathbb{R}^n$. Tale proprietà è una diretta conseguenza delle definizioni 1.1.2 e 1.2: infatti, se $\mathbf{u}$ e $\mathbf{v} \in \mathbb{R}^n$ e α e $\beta \in \mathbb{R}$, allora $\alpha\mathbf{u} = \mathbf{w}_1 \in \mathbb{R}^n$ e $\beta\mathbf{v} = \mathbf{w}_2 \in \mathbb{R}^n$, e $\mathbf{w}_1 + \mathbf{w}_2 \in \mathbb{R}^n$, per cui $\alpha\mathbf{u} + \beta\mathbf{v} \in \mathbb{R}^n$.
Attraverso le combinazioni lineari è possibile costruire particolari sottoinsiemi dello spazio vettoriale.

Sottoinsieme lineare generato da un vettore

Dato un qualsiasi vettore $\mathbf{u} \in \mathbb{R}^n$ possiamo generare un sottoinsieme proprio di $\mathbb{R}^n$ (vedremo nel seguito che tale sottoinsieme viene anche definito sottospazio vettoriale) considerando tutti i vettori che siano multipli di $\mathbf{u}$:

$$C(\mathbf{u}) = \{\mathbf{w} \in \mathbb{R}^n \,|\, \exists\, \alpha \in \mathbb{R} \text{ con } \mathbf{w} = \alpha\mathbf{u}\},$$

Dal punto di vista geometrico l'insieme $C(\mathbf{u})$ è costituito da tutti i vettori paralleli a $\mathbf{u}$.

Sottoinsieme piano generato da due vettori

Data una coppia di vettori $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ possiamo generare un sottoinsieme proprio di $\mathbb{R}^n$ (anche in questo caso si tratta di un sottospazio vettoriale) sfruttando tutte le possibili combinazioni lineari di questi due vettori:

$$C(\mathbf{u}, \mathbf{v}) = \{\mathbf{w} \in \mathbb{R}^n \,|\, \exists\, \alpha, , \beta \in \mathbb{R} \text{ con } \mathbf{w} = \alpha\mathbf{u} + \beta\mathbf{v}\},$$

Dal punto di vista geometrico l'insieme $C(\mathbf{u}, \mathbf{v})$ è costituito da tutti i vettori complanari alla coppia di vettori $\mathbf{u}$ e $\mathbf{v}$ (due vettori identificano sempre un piano, a meno di non essere paralleli).

Combinazione lineare convessa di più vettori

Dati k vettori, $\mathbf{u}_k \in \mathbb{R}^n$, e k numeri reali, $c_i \in \mathbb{R}$, tali che $c_1 + c_2 + \cdots + c_k = 1$ con $c_i \in [0, 1]\, \forall\, i = 1 \ldots k$ definiamo combinazione lineare convessa degli $\mathbf{u}_i$ l'insieme V

$$V = \left\{ \mathbf{w} \in \mathbb{R}^n \,|\, \mathbf{w} = \sum_{i=1}^{k} c_i \mathbf{u}_i \;\forall c_i \in [0, 1] \text{ con } c_1 + c_2 + \cdots + c_k = 1 \right\}.$$

L'insieme V è costituito dal poligono convesso che ha i vertici negli estremi dei vettori $\mathbf{u}_k$.

1.1.4 Indipendenza lineare

Abbiamo visto come sia possibile costruire degli insiemi semplicemente sommando tra loro dei vettori, o moltiplicandoli per dei numeri reali. A questo punto è abbastanza naturale chiedersi se un vettore può sempre essere ottenuto come combinazione lineare di altri vettori. La risposta è nella definizione di indipendenza lineare.

Definizione 1.1.4. *Dati k vettori $\mathbf{u}_k$ essi si dicono linearmente indipendenti se non esiste nessun insieme di k numeri $\alpha_i \in \mathbb{R}$ tali che*

$$\sum_{i=1}^{k} \alpha_i \mathbf{u}_i = \mathbf{0}$$

a meno che non sia $\alpha_i = 0 \ \forall \ i = 1 \dots k$

In altri termini, k vettori si dicono linearmente indipendenti se non esiste nessuna combinazione lineare di essi (a coefficienti non tutti nulli) che dia come risultato il vettore nullo.
Se k vettori non sono linearmente indipendenti, ossia se è possibile costruirne una combinazione lineare (a coefficienti non tutti nulli) che dia come risultato il vettore nullo, allora si dicono **linearmente dipendenti**. Questo implica che è possibile esprimere un vettore come combinazione lineare degli altri $k-1$. Infatti, considerando per esempio solo tre vettori, $\mathbf{u}, \mathbf{v}$ e $\mathbf{w}$, se esistono tre numeri, α, β e γ, tali che

$$\alpha\mathbf{u} + \beta\mathbf{v} + \gamma\mathbf{w} = \mathbf{0} \ \Rightarrow \ \mathbf{u} = -\frac{\beta}{\alpha}\mathbf{v} - \frac{\gamma}{\alpha}\mathbf{w}\,.$$

Avendo supposto $\alpha \neq 0$ abbiamo potuto esprimere $\mathbf{u}$ in funzione degli altri vettori, ma, ovviamente, avremmo potuto esprimere qualsiasi vettore in funzione degli altri non appena il suo coefficiente nella combinazione lineare fosse stato diverso da zero. Questo semplice risultato implica il seguente risultato: **se k vettori sono linearmente indipendenti, non c'è nessun modo di ottenerne uno qualsiasi come combinazione lineare dei restanti.**

1.1.5 Rango di un insieme di vettori

Una definizione che si basa sul concetto di indipendenza lineare e che ci tornerà utile nel seguito è quella di rango.

Definizione 1.1.5. *Dato un insieme qualsiasi di vettori di $\mathbb{R}^n$ $V = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ definiamo rango di V e scriveremo $Rg(V)$ il numero massimo di vettori linearmente indipendenti appartenenti a V.*

Facciamo qui notare che il numero massimo di vettori linearmente indipendenti che si possono trovare in $\mathbb{R}^n$ è n. Questa affermazione è abbastanza elementare se considerate il modo in cui abbiamo costruito $\mathbb{R}^n$, ossia come prodotto cartesiano di n insiemi $\mathbb{R}$ (ogni insieme $\mathbb{R}$ porta una dimensione aggiuntiva). Però vedremo nella prossima sezione come il numero massimo di vettori linearmente indipendenti stabilisce la dimensione di uno spazio.

1.2 Spazi vettoriali

L'insieme dei vettori di $\mathbb{R}^n$ dotato delle operazioni di somma e prodotto per scalare che abbiamo definito, viene anche chiamato spazio vettoriale. Visto che su molti

insiemi di elementi diversi (non solo vettori di $\mathbb{R}^n$) è possibile introdurre le due operazioni matematiche di somma e prodotto per scalare verificando le relative proprietà s1-s4 e m1-m6, rispettivamente, è possibile estendere il concetto di spazio di vettori anche nel caso di elementi astratti (ossia che non siano vettori "geometrici" di $\mathbb{R}^n$). Tali spazi vengono chiamati **spazi vettoriali**. La struttura matematica di spazio vettoriale deriva dalle operazioni che definiamo sugli elementi piuttosto che dagli elementi in se che formano lo spazio.

In questa sezione, daremo alcune definizioni preliminari con delle semplici applicazioni del concetto di spazio vettoriale principalmente nel caso di vettori di $\mathbb{R}^n$.

1.2.1 Definizione di spazio e sottospazio vettoriale

Definizione 1.2.1. *Dato un generico insieme di elementi $\mathbb{V}$ diciamo che forma uno spazio vettoriale se in esso è possibile definire due operazioni: la somma, $+$,*

$$+ : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$$

che associa a due elementi di $\mathbb{V}$ un altro elemento di $\mathbb{V}$ verificando le proprietà s1-s4, e un prodotto per scalare, $\cdot$,

$$\cdot : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$$

che associa a un elemento di $\mathbb{V}$ e a un elemento di $\mathbb{R}$ un elemento di $\mathbb{V}$ verificando le proprietà m1-m6.

Abbiamo già visto come lo spazio vettoriale $\mathbb{R}^n$ sia chiuso rispetto alle operazioni di somma e prodotto per scalare (in generale è chiuso rispetto alle combinazioni lineari). Questa è la caratteristica fondamentale che deve avere ogni spazio vettoriale: è fondamentale che le combinazioni lineari definite su un certo insieme di elementi $\mathbb{V}$ siano tali da ottenere come risultato nuovi elementi ma sempre appartenenti all'insieme di partenza: $\mathbb{V}$ è uno spazio vettoriale se presi k vettori qualsiasi $\mathbf{u}_i \in \mathbb{V}$ e k numeri qualsiasi $\alpha_i \in \mathbb{R}$, con $i = 1 \dots k$, la combinazione lineare dei k vettori con i k coefficienti continua ad appartenere a $\mathbb{V}$:

$$\sum_{i=1}^{k} \alpha_i \mathbf{u}_i = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots \alpha_k \mathbf{u}_k \in \mathbb{V}$$

1.2.2 Esempi di spazi vettoriali

È abbastanza semplice dare esempi di spazi vettoriali: bisogna cercare degli insiemi in cui esiste già l'operazione di somma interna e prodotto per scalare (e se esiste, verificherà anche le proprietà s1-s4, m1-m6), e che siano chiusi rispetto a combinazioni lineari.

- Definiamo $P_n(x)$ lo spazio dei polinomi di dimensione minore o uguale a n e dimostriamo che esso è uno spazio vettoriale. L'operazione di somma è quella usuale del calcolo algebrico così come il prodotto per scalare. La chiusura

rispetto a queste operazione si dimostra osservando che da qualsiasi combinazione lineare di polinomi di ordine minore o uguale a n si ottiene un polinomio di grado minore o uguale a n. Molto diverse sarebbero state le cose se avessimo scelto come insieme lo spazio $Q_n(x)$ dei polinomi di grado esattamente uguale a n. In questo caso non è difficile costruire una combinazione lineare tale da uscire da $Q_n(x)$: dati $u, v \in Q_2(x)$ tali che $u = x^2 + 1$ e $v = x^2 + x + 1$ si ha che $w = v - u = x \notin Q_2(x)$.

- Sia $T(x) = \{a\sin(x) + b\cos(x) \ \forall \ a, b \in \mathbb{R}\}$ lo spazio delle combinazioni lineari delle funzioni trigonometriche $\sin(x)$ e $\cos(x)$. Anche in questo caso è immediato vedere che $T(x)$ forma uno spazio vettoriale. Le operazioni di somma e di prodotto per scalare sono quelle usuali che utilizziamo quando lavoriamo con delle funzioni, e la chiusura si dimostra osservando che da combinazioni lineari di elementi di $T(x)$ non si può mai ottenere qualcosa che non appartenga a $T(x)$.

1.2.3 Sottospazi vettoriali

A partire da un generico spazio vettoriale $\mathbb{V}$ useremo la nozione di combinazione lineare per definire i sottospazi vettoriali.

Definizione 1.2.2. *L'insieme $\mathbb{U} \subseteq \mathbb{V}$ è detto sottospazio vettoriale (di $\mathbb{V}$) se è chiuso rispetto alle operazioni di somma e prodotto per scalare. In generale, comunque presi k vettori $\mathbf{u}_i \in \mathbb{U}$, ogni loro combinazione lineare dovrà continuare ad appartenere a $\mathbb{U}$.*

In particolare, il vettore nullo $\mathbf{0}$, dovrà per forza appartenere a $\mathbb{U}$: infatti il vettore nullo è sempre ottenibile per mezzo di una combinazione lineare a coefficienti tutti nulli. Gli insiemi riportati in 1.1.3 (tranne la combinazione convessa) sono esempi importanti di sottospazi vettoriali. L'immediata generalizzazione di questi due esempi la troviamo nel concetto di Span:

Definizione 1.2.3. *Dati k vettori $\mathbf{u}_i \in \mathbb{V}$ con $i = 1 \dots k$, il sottospazio ottenuto attraverso tutte le possibili combinazioni lineari dei vettori $\mathbf{u}_i$ viene indicato con*

$$Span(\mathbf{u}_1, \mathbf{u}_2 \dots \mathbf{u}_k) = \left\{ \mathbf{v} = \sum_{i=1}^{k} \alpha_i \mathbf{u}_i \ \ con \ \ \alpha_1, \alpha_2, \dots \alpha_k \in \mathbb{R} \right\} \tag{1.4}$$

e viene detto sottospazio generato dai vettori $\mathbf{u}_i$ ovvero span lineare dei vettori $\mathbf{u}_i$.

Alcune proprietà immediate dello Span sono:

a) $\text{Span}(\mathbf{u}_1, \mathbf{u}_2 \dots \mathbf{u}_k)$ è un sottospazio vettoriale di $\mathbb{V}$.

b) Siano S e T due insiemi di vettori. Se $S \subset T$ allora $\text{Span}(S) \subset \text{Span}(T)$.

c) Se $S = \{\mathbf{u}_1, \mathbf{u}_2 \dots \mathbf{u}_k\}$ e $T = \{\mathbf{u}_1, \mathbf{u}_2 \dots \mathbf{u}_k, \mathbf{u}_{k+1}\}$ allora $\text{Span}(S) = \text{Span}(T)$ se e solo se $\mathbf{u}_{k+1}$ è linearmente dipendente dai vettori $\mathbf{u}_1, \mathbf{u}_2 \dots \mathbf{u}_k$, cioè $\mathbf{u}_{k+1} \in \text{Span}(S)$.

1.2.4 Operazioni insiemistiche su sottospazi vettoriali

Essendo i sottospazi vettoriali degli insiemi, possiamo agire su di essi attraverso gli operatori classici della teoria degli insiemi, ossia l'unione, l'intersezione, ecc.

Intersezione

Dati U e V sottospazi vettoriali, la loro intersezione è ancora un sottospazio vettoriale. Infatti, se $\mathbf{u}, \mathbf{v} \in U \cap V$ allora $\mathbf{u}$ e $\mathbf{v}$ dovranno appartenere sia a U che a V così come ogni loro combinazione lineare. Per cui $\alpha\mathbf{u} + \beta\mathbf{v} \in U \cap V$.

Unione

Dati U e V sottospazi vettoriali, la loro unione non è un sottospazio vettoriale. Se infatti consideriamo $U = \text{Span}(\mathbf{u})$ e $V = \text{Span}(\mathbf{v})$ con $\mathbf{u}$ e $\mathbf{v}$ linearmente indipendenti, allora $U \cup V$ conterrà sicuramente sia $\mathbf{u}$ che $\mathbf{v}$ ma non conterrà $\mathbf{u} + \mathbf{v}$ che è una semplice combinazione lineare di $\mathbf{u}$ e $\mathbf{v}$, e quindi $U \cup V$ non può essere un sottospazio vettoriale.

Somma e somma diretta

Visto che nei sottospazi vettoriali di $\mathbb{V}$ è definita un'operazione di somma tra gli elementi dello spazio, possiamo anche definire un'operazione insiemistica di somma come:

Definizione 1.2.4. *Dati due sottospazi vettoriali $\mathbb{U}_1, \mathbb{U}_2 \in \mathbb{V}$ definiamo l'operazione di somma insiemistica tra $\mathbb{U}_1$ e $\mathbb{U}_2$ come l'insieme*

$$\mathbb{U}_3 = \mathbb{U}_1 + \mathbb{U}_2 = \{\mathbf{u}_3 = \mathbf{u}_1 + \mathbf{u}_2 \ \forall \ \mathbf{u}_1 \in \mathbb{U}_1 \ e \ \mathbf{u}_2 \in \mathbb{U}_2\}\,. \tag{1.5}$$

È semplice mostrare, sfruttando la proprietà di chiusura rispetto alle combinazioni lineari, che la somma tra spazi vettoriali è ancora uno spazio vettoriale. Lo spazio vettoriale somma $\mathbb{U}_3 = \mathbb{U}_1 + \mathbb{U}_2$ sarà sempre più grande sia di $\mathbb{U}_1$ che di $\mathbb{U}_2$ (in notazione insiemistica $\mathbb{U}_1 \subset \mathbb{U}_3$ e $\mathbb{U}_2 \subset \mathbb{U}_3$) a meno che $\mathbb{U}_1$ non sia uguale a $\mathbb{U}_2$, in tal caso $\mathbb{U}_1 = \mathbb{U}_2 = \mathbb{U}_3$.
Un caso particolarmente importante si ha quando $\mathbb{U}_1$ e $\mathbb{U}_2$ non hanno elementi in comune (a meno del vettore nullo, che appartiene a tutti i sottospazi vettoriali).

Definizione 1.2.5. *Dati due sottospazi vettoriali $\mathbb{U}_1$ e $\mathbb{U}_2 \in \mathbb{V}$ tali che $\mathbb{U}_1 \cap \mathbb{U}_2 = \{\mathbf{0}\}$, l'operazione di somma tra i sottospazi viene detta somma diretta e si indica con*

$$\mathbb{U}_3 = \mathbb{U}_1 \oplus \mathbb{U}_2 = \{\mathbf{u}_3 = \mathbf{u}_1 + \mathbf{u}_2 \ \ \forall \ \ \mathbf{u}_1 \in \mathbb{U}_1 \ e \ \mathbf{u}_2 \in \mathbb{U}_2\}\,. \tag{1.6}$$

Formula di Grassmann

Per concludere questa sottosezione sulle operazioni insiemistiche su spazi vettoriali dimostriamo la formula di Grassmann che lega tra loro le varie operazioni introdotte:

Theorem 1.2.6. *Dati due sottospazi vettoriali* $\mathbb{U}_1$ *e* $\mathbb{U}_2 \in \mathbb{V}$ *sussiste la seguente relazione*

$$\dim(\mathbb{U}_1) + \dim(\mathbb{U}_2) = \dim(\mathbb{U}_1 + \mathbb{U}_2) + \dim(\mathbb{U}_1 \cap \mathbb{U}_2) \tag{1.7}$$

Nel caso in cui $\mathbb{U}_1 \cap \mathbb{U}_2 = \{\mathbf{0}\}$ *allora si ha*

$$\dim(\mathbb{U}_1 \oplus \mathbb{U}_2) = \dim(\mathbb{U}_1) + \dim(\mathbb{U}_2) \tag{1.8}$$

1.2.5 Sistemi di generatori e basi dei sottospazi vettoriali

Dalla teoria sviluppata e dagli esempi presentati, dovrebbe risultare chiaro che l'operazione fondamentale negli spazi vettoriali è la combinazione lineare, e che il concetto chiave legato a questa operazione è l'indipendenza lineare. Con le combinazioni lineari si costruiscono tutti i vettori dello spazio, e con l'indipendenza lineare si stabilisce quanti vettori siano effettivamente necessari a questa costruzione. Per chiarire con un ulteriore esempio questa affermazione, dati tre vettori $\mathbf{u}, \mathbf{u}, \mathbf{w} \in \mathbb{V}$ se questi sono linearmente dipendenti, almeno uno dei tre non porta "informazione nuova", ossia posso ottenerlo mediante una combinazione lineare degli altri due. Infatti, supponendo che $\mathbf{u}$ e $\mathbf{v}$ sono linearmente indipendenti si avrà Span$(\mathbf{u}, \mathbf{v})$ =Span$(\mathbf{u}, \mathbf{v}, \mathbf{w})$ e inoltre $\mathbf{w}$ potrà essere ottenuto mediante una combinazione lineare di $\mathbf{u}$ e $\mathbf{v}$.
Queste riflessioni portano a tutta una serie di domande:

- qual'è il numero minimo di vettori necessari per costruire uno spazio (o un sottospazio) vettoriale?
- sottospazi vettoriali generati da gruppi di vettori indipendenti e distinti sono necessariamente diversi?
- la dimensione di un sottospazio di $\mathbb{V}$ è legata all'indipendenza dei vettori?

Queste domande riguardano il cuore della struttura degli spazi vettoriali e per iniziare a dare una risposta è necessario introdurre il concetto di base di uno spazio vettoriale.

Definizione 1.2.7. *Un insieme ordinato di vettori* $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ *è una base per lo spazio vettoriale* $\mathbb{V}$ *se*

- *sono linearmente indipendenti:* $\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = 0 \iff \alpha_i = 0 \ \forall \ i$;
- *generano tutto lo spazio: Span*$(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \mathbb{V}$.

Nella definizione non si fà alcun riferimento a come devono essere questi vettori (ovviamente a parte la loro indipendenza lineare ed il fatto che generano tutto lo spazio), per cui insiemi di vettori diversi possono essere una base per lo stesso spazio vettoriale.

1.2.6 Base canonica di $\mathbb{R}^n$

Tuttavia esistono delle basi privilegiate attraverso le quali è più "semplice" o "immediato" generare i vettori dello spazio. In $\mathbb{R}^n$ questa base è quella che si definisce **canonica** ed è formata dai seguenti vettori:

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots, \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Utilizzando i vettori di questa base la generazione dei vettori di $\mathbb{R}^n$ è banale, in quanto, dato un generico vettore $\mathbf{u} \in \mathbb{R}^n$, ogni coefficiente reale che va moltiplicato per i vettori della base nella combinazione lineare che genera $\mathbf{u}$, è pari alla componente di $\mathbf{u}$ presente nella coordinata corrispondente. In formule,

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \Rightarrow \mathbf{u} = u_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + u_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + u_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \Rightarrow$$

$$\mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + \cdots + u_n\mathbf{e}_n = \sum_{i=1}^{n} u_i\mathbf{e}_i \,.$$

1.2.7 Basi e dimensione di spazi vettoriali

Il concetto di base è strettamente legato al concetto di dimensione di uno spazio vettoriale. Si ha infatti il seguente fondamentale teorema sulla dimensione degli spazi vettoriali

Theorem 1.2.8. *Due basi diverse per uno stesso spazio vettoriale devono contenere lo stesso numero di vettori.*

Utilizzando questo teorema (di cui ometteremo la dimostrazione), e visto che la base canonica di $\mathbb{R}^n$ è composta da n elementi possiamo affermare che ogni base di $\mathbb{R}^n$ è composta da n vettori, e quindi che la dimensione dello spazio vettoriale è pari al numero di vettori di base. Questo risultato è valido anche nel caso di spazi vettoriali astratti. Ossia possiamo introdurre la seguente:

Definizione 1.2.9. *Dato lo spazio vettoriale $\mathbb{V}$, definiamo dimensione di $\mathbb{V}$ come il numero massimo di vettori linearmente indipendenti che è possibile scegliere in $\mathbb{V}$, e scriveremo*

$$dim(\mathbb{V}) = n \,.$$

1.2.8 Unicità della decomposizione

Data una qualsiasi base $\mathbb{E} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ per uno spazio vettoriale $\mathbb{V}$ e dato un qualsiasi vettore $\mathbf{v} \in \mathbb{V}$ la decomposizione di $\mathbf{v}$ rispetto alla base $\mathbb{E}$ è unica. Siano infatti α_i i coefficienti della decomposizione di $\mathbf{v}$:

$$\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n = \sum_{i=1}^{n} \alpha_i\mathbf{u}_i \,.$$

Se questa decomposizione non fosse unica, allora esisterebbero altri coefficienti β_i tali che

$$\mathbf{v} = \beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \cdots + \beta_n\mathbf{u}_n = \sum_{i=1}^{n} \beta_i\mathbf{u}_i \,.$$

Allora, sottraendo queste due uguaglianze otterremo

$$\sum_{i=1}^{n} (\alpha_i - \beta_i)\mathbf{u}_i = 0 \,,$$

e visto che i vettori $\mathbf{u}_i$ sono linearmente indipendenti l'unica possibilità perché questa uguaglianza sia verificata è che $\alpha_i = \beta_i \;\forall\; i = 1 \ldots n$. I valori α_i vengono detti **coordinate** di $\mathbf{v}$ rispetto alla base $\mathbb{E}$. Solo nel caso in cui $\mathbb{V} = \mathbb{R}^n$ ed $\mathbb{E}$ sia la base canonica allora c'è una corrispondenza tra le coordinate spaziali di un vettore di $\mathbb{R}^n$ e le coordinate rispetto alla base. In generale le coordinate geometriche di un vettore di $\mathbb{R}^n$ possono differire dalle coordinate vettoriali. Nel caso di spazi vettoriali astratti non esiste una corrispondenza tra coordinate e geometria. Infatti, nel caso degli spazi vettoriali astratti non sarà possibile associare in generale ai vettori un'interpretazione geometrica e si potranno definire solo le coordinate dei vettori rispetto a una determinata base.

1.2.9 Basi e rappresentazione dei vettori

Fissare una base in uno spazio vettoriale equivale a fissare un sistema di riferimento rispetto al quale *guardare* i vettori dello spazio. Mentre possiamo considerare i vettori come entità astratte, le coordinate dei vettori rispetto alla base sono una **rappresentazione** di questi vettori che dipende dalla base scelta. Anche se lo spazio vettoriale è composto da elementi di un insieme astratto, una volta stabilita una base, allora è possibile identificare univocamente i vettori con le coordinate che questi hanno rispetto alla base.
Dato $\mathbb{V}$ uno spazio vettoriale di dimensione n, identifichiamo con $\mathbb{E}$ una possibile base:

$$\mathbb{E} = (\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n)$$

dove i vari $\mathbf{u}_i$ sono i vettori di base. Quindi, ogni vettore $\mathbf{x} \in \mathbb{V}$ può essere rappresentato in modo univoco a partire dai vettori di base:

$$\mathbf{x} = x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_n\mathbf{u}_n = \sum_{i=1}^{n} x_i\mathbf{u}_i \,. \tag{1.9}$$

Se consideriamo $\mathbb{E}$ come una matrice con una riga ed n colonne (un vettore riga), dove ogni singolo elemento di colonna è un vettore di base (possibilmente anche un elemento astratto, tipo $\sin x$, x^2, o un semplice vettore della base canonica di $\mathbb{R}^n$) allora possiamo scrivere, utilizzando il prodotto righe per colonne:

$$\mathbf{x} = x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_n\mathbf{u}_n = \mathbb{E}\mathbf{x}_{\mathbb{E}} \text{ con } \mathbf{x}_E = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{1.10}$$

dove $\mathbf{x}_{\mathbb{E}}$ è il vettore con le componenti di $\mathbf{x}$ rispetto alla base $\mathbb{E}$, ossia **la rappresentazione vettoriale** di $\mathbf{x}$ rispetto alla base $\mathbb{E}$. Cambiando base, cambiano le coordinate di $\mathbf{x}$ rispetto alla base, e quindi cambia la sua rappresentazione vettoriale, anche se il vettore astratto resta sempre lo stesso.
Quindi, una volta fissata una base in uno spazio vettoriale (anche astratto) è fissata una corrispondenza biunivoca tra i vettori dello spazio vettoriale e i vettori di $\mathbb{R}^n$:

$$\mathbf{x} = \mathbb{E}\mathbf{x}_{\mathbb{E}} \;\rightarrow\; \mathbf{x} \iff \mathbf{x}_{\mathbb{E}}$$

dove $\mathbf{x} \in \mathbb{V}$ e $\mathbf{x}_{\mathbb{E}} \in \mathbb{R}^n$.

1.3 Calcolo matriciale

Interrompiamo momentaneamente l'esposizione della teoria sugli spazi vettoriali per introdurre il calcolo matriciale. Se l'importanza dell'introduzione dei vettori è abbastanza palese, se non altro per la struttura tridimensionale dello spazio in cui viviamo (per identificare qualsiasi punto nel nostro spazio fisico abbiamo bisogno di tre coordinate, ossia di un vettore di $\mathbb{R}^3$) l'importanza dell'introduzione delle matrici sarà chiarito solo nel seguito, motivato dal loro utilizzo. Inizialmente ci accontenteremo di pensare alle matrici come a tabelle di numeri.

1.3.1 Definizione di matrice e primi esempi

Definizione 1.3.1. *Una matrice $n \times m$, che indicheremo con $\hat{A}$ è una tabella di numeri composta da n righe $\times$ m colonne.*

In alcuni contesti sarà utile pensare a una matrice $n \times m$ come a un insieme di m vettori colonna appartenenti a $\mathbb{R}^n$ (ovvero a un insieme di n vettori riga appartenenti a $\mathbb{R}^m$).

Una rappresentazione tabellare di una matrice che evidenzia gli elementi è la seguente:

$$\hat{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}.$$

Per identificare l'elemento alla riga i e alla colonna j di una matrice useremo la notazione

$$(\hat{A})_{ij} = a_{ij}\,.$$

Tipi di matrici

Una matrice $\hat{A}$ si dice

- quadrata se $n = m$,
- rettangolare se $n \neq m$,
- simmetrica se quadrata e se $(\hat{A})_{ij} = (\hat{A})_{ji}$,
- vettore (o vettore colonna) se $n = 1$,
- vettore riga se $m = 1$,
- scalare se $n = m = 1$.

In generale definiamo $\mathcal{M}(n, m)$ come l'insieme di tutte le matrici $n \times m$. Vedremo poi nel seguito che $\mathcal{M}(n, m)$ è anche uno spazio vettoriale.

1.3.2 Operazioni con le matrici

Procedendo come nel caso dei vettori, introdurremo prima la somma tra matrici e dopo il prodotto tra numeri e matrici. Successivamente vedremo come sarà possibile, in alcuni casi, definire un prodotto tra matrici.

Somma

La somma di due o più matrici $n \times m$ è un'applicazione che associa a due elementi di $\mathcal{M}(n, m)$ un terzo elemento sempre in $\mathcal{M}(n, m)$

$$+ : \mathcal{M}(n, m) \times \mathcal{M}(n, m) \to \mathcal{M}(n, m)$$

ottenuto sommando tutte le componenti omogenee delle due matrici iniziali:

Definizione 1.3.2. *Prese due qualsiasi matrici $n \times m$, $\hat{A}, \hat{B} \in \mathcal{M}(n, m)$, definiamo $\hat{C} \in \mathcal{M}(n, m)$*

$$(\hat{C})_{ij} = (\hat{A} + \hat{B})_{ij} = c_{ij} = a_{ij} + b_{ij}. \tag{1.11}$$

Nella rappresentazione tabellare estesa questo equivale a scrivere

$$\hat{C} = \hat{A} + \hat{B} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1m} + b_{1m} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2m} + b_{2m} \\ \cdots & \cdots & \ddots & \cdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \cdots & a_{nm} + b_{nm} \end{pmatrix}$$

Le proprietà della somma così definite sono

s1) Associatività: $\forall\, \hat{A}, \hat{B}, \hat{C} \in \mathcal{M}(n,m)$:

$$(\hat{A} + \hat{B}) + \hat{C} = \hat{A} + (\hat{B} + \hat{C})$$

s2) Commutatività: $\forall\, \hat{A}, \hat{B} \in \mathcal{M}(n,m)$:

$$\hat{A} + \hat{B} = \hat{B} + \hat{A}$$

s3) Esistenza dell'elemento neutro: $\exists!\, \hat{0} \in \mathcal{M}(n,m) \,|\, \forall\, \hat{A} \in \mathbb{R}^n$:

$$\hat{A} + \hat{0} = \hat{0} + \hat{A} = \hat{A}$$

s4) Esistenza dell'elemento opposto: $\forall\, \hat{A} \in \mathcal{M}(n,m)\ \exists!\, (-\hat{A}) \in \mathcal{M}(n,m)$

$$\hat{A} + (-\hat{A}) = (-\hat{A}) + \hat{A} = \hat{0}$$

La matrice $\hat{0}$ è la matrice con tutti gli elementi pari a 0.

Prodotto per scalare

Un'altra operazione fondamentale per lavorare con le matrici è il prodotto per scalare, ossia la moltiplicazione di un numero semplice (uno scalare) per una matrice:

$$\cdot : \mathbb{R} \times \mathcal{M}(n,m) \to \mathcal{M}(n,m)\,.$$

Questa operazione si effettua moltiplicando per lo scalare tutte le componenti della matrice.

Definizione 1.3.3. *Presa una qualsiasi matrice $\hat{A} \in \mathcal{M}(n,m)$ e un qualsiasi scalare $\alpha \in \mathbb{R}$ definamo $\hat{B} \in \mathcal{M}(n,m)$*

$$(\hat{B})_{ij} = (\alpha \hat{A})_{ij} = b_{ij} = \alpha a_{ij}. \tag{1.12}$$

Nella rappresentazione tabellare questo equivale a scrivere

$$\alpha\hat{A} = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1m} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2m} \\ \cdots & \cdots & \ddots & \cdots \\ \alpha a_{n1} & \alpha a_{n2} & \cdots & \alpha a_{nm} \end{pmatrix}$$

Le proprietà del prodotto per scalare appena definito sono

m1) Associatività: $\forall\, \alpha, \beta \in \mathbb{R}$ e $\forall\, \hat{A} \in \mathcal{M}(n,m)$:

$$\alpha(\beta\hat{A}) = (\alpha\beta)\hat{A} = \alpha\beta\hat{A}$$

m2) Commutatività: $\forall\, \alpha \in \mathbb{R}$ e $\forall\, \hat{A} \in \mathcal{M}(n,m)$:

$$\alpha\hat{A} = \hat{A}\alpha$$

m3) Esistenza dell'elemento neutro: $\exists!\, 1 \in \mathbb{R} \,|\, \forall\, \hat{A} \in \mathcal{M}(n,m)$:

$$1\hat{A} = \hat{A}$$

m4) Costruzione dell'elemento opposto: $\forall\, \hat{A} \in \mathcal{M}(n,m)\ \exists!\, (-1) \in \mathbb{R}$:

$$-1\hat{A} + \hat{A} = (-\hat{A}) + \hat{A} = \mathbf{0}$$

m5) Distributività rispetto alla somma tra scalari: $\forall\, \alpha, \beta \in \mathbb{R}$ e $\forall\, \hat{A} \in \mathcal{M}(n,m)$:

$$(\alpha + \beta)\hat{A} = \alpha\hat{A} + \beta\hat{A}$$

m6) Distributività rispetto alla somma tra matrici: $\forall\, \alpha \in \mathbb{R}$ e $\forall\, \hat{A}, \hat{B} \in \mathcal{M}(n,m)$:

$$\alpha(\hat{A} + \hat{B}) = \alpha\hat{A} + \alpha\hat{B}$$

Mentre è stato semplice nel caso dei vettori dare un'interpretazione geometrica alle operazioni di somma e di prodotto per scalare, nel caso delle matrici quest'interpretazione viene a mancare e le operazioni sono definite in modo puramente analitico.

Spazio delle matrici come spazio vettoriale

Visto che le proprietà di somma tra matrici e di prodotto per scalare verificano esattamente le stesse proprietà viste per i vettori, possiamo affermare che lo spazio delle matrici $\mathcal{M}(n,m)$ è uno spazio vettoriale.
Inoltre con una semplice riorganizzazione degli indici possiamo mostrare come $\mathcal{M}(n,m)$ è equivalente a $\mathbb{R}^{nm}$ ossia allo spazio dei vettori di dimensione n per m:

$$\mathbf{v} \in \mathbb{R}^{nm} \Rightarrow \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{nm} \end{pmatrix} \Longrightarrow \hat{A}(\mathbf{v}) = \begin{pmatrix} v_1 & v_{n+1} & \cdots & v_{(m-1)n+1} \\ v_2 & v_{n+2} & \cdots & v_{(m-1)n+2} \\ \cdots & \cdots & \ddots & \cdots \\ v_n & v_{2n} & \cdots & v_{mn} \end{pmatrix}$$

Questa equivalenza di tipo *qualitativo* non implica necessariamente un'equivalenza di tipo *strutturale* tra i due spazi. O, per meglio dire, per quanto riguarda la struttura di spazio vettoriale, $\mathcal{M}(n,m)$ è assolutamente equivalente a $\mathbb{R}^{nm}$. Ma, visto che è possibile definire negli spazi di matrici un'operazione di prodotto che non ha equivalenti nel caso dei vettori, questa darà allo spazio delle matrici delle sua caratteristiche peculiari.

Prodotto tra matrici

Negli spazi delle matrici, in alcuni casi, è possibile definire anche un'operazione interna simile all'operazione di prodotto tra scalari. Per introdurre il prodotto tra matrici dobbiamo prima definire il concetto di matrici conformabili. Diremo che due matrici, $\hat{A}$ e $\hat{B}$, sono **conformabili** se il numero di colonne di $\hat{A}$ è uguale al numero

di righe di $\hat{B}$. In questo caso è possibile definire un'operazione di prodotto tra matrici come un'applicazione che associa a due matrici conformabili una terza matrice:

$$\cdot : \mathcal{M}(n,l) \times \mathcal{M}(l,m) \to \mathcal{M}(n,m)\,.$$

Il simbolo utilizzato per il prodotto tra matrici è lo stesso utilizzato per il prodotto scalare-matrice, ma sarà sempre chiaro dal contesto a quale prodotto ci si riferisce. Anzi, molto spesso, il segno di prodotto (quale esso sia) viene eliminato del tutto, essendo sottointeso che quando due oggetti matematici (scalare, vettore, matrice) sono scritti uno accanto all'altro, c'è sempre un'operazione di prodotto in mezzo.

Definizione 1.3.4. *Siano $\hat{A} \in \mathcal{M}(n,l)$ e $\hat{B} \in \mathcal{M}(l,m)$ due matrici conformabili. Il prodotto matriciale $\hat{A}\hat{B}$ è una matrice $\hat{C} \in \mathcal{M}(n,m)$ definita come*

$$(\hat{C})_{ij} = (\hat{A}\cdot\hat{B})_{ij} = (\hat{A}\hat{B})_{ij} = c_{ij} = \sum_{k=1}^{l} a_{ik}b_{kj}. \tag{1.13}$$

Risulta evidente da questa definizione che, anche se è possibile calcolare il prodotto $\hat{A}\hat{B}$, non è detto che sia possibile calcolare il prodotto $\hat{B}\hat{A}$, e, nel caso in cui si possa fare (per esempio se $\hat{A} \in \mathcal{M}(n,m)$ e $\hat{B} \in \mathcal{M}(m,n)$ oppure se $\hat{A},\hat{B} \in \mathcal{M}(n,n)$) non è detto che sia $\hat{A}\hat{B} = \hat{B}\hat{A}$, anzi, in generale, sarà $\hat{A}\hat{B} \neq \hat{B}\hat{A}$ sia per quanto riguarda le dimensioni delle matrici prodotto, sia, nel caso in cui $\hat{A},\hat{B} \in \mathcal{M}(n,n)$, per quanto riguarda le componenti delle matrici. Quindi nel prodotto tra matrici è fondamentale l'ordine con il quale lo si effettua, ed è per questo che si dice che il prodotto tra matrici non è commutativo.

L'operazione di prodotto tra matrici verifica le seguenti proprietà:

p1) Associatività: date tre matrici conformabili $\hat{A} \in \mathcal{M}(n,k)$, $\hat{B} \in \mathcal{M}(k,l)$, $\hat{C} \in \mathcal{M}(l,m)$ si ha
$$(\hat{A}\hat{B})\hat{C} = \hat{A}(\hat{B}\hat{C})\,.$$

p2) Distributività rispetto alla somma: date tre matrici $\hat{A} \in \mathcal{M}(n,k)$, $\hat{B},\hat{C} \in \mathcal{M}(k,m)$ si ha
$$\hat{A}(\hat{B}+\hat{C}) = \hat{A}\hat{B} + \hat{A}\hat{C}\,.$$
Ovvero, analogamente, date tre matrici $\hat{A},\hat{B} \in \mathcal{M}(n,k)$, $\hat{C} \in \mathcal{M}(k,m)$ si ha
$$(\hat{A}+\hat{B})\hat{C} = \hat{A}\hat{C} + \hat{B}\hat{C}\,.$$

p3) Esistenza dell'elemento neutro: data una qualsiasi matrice $\hat{A} \in \mathcal{M}(n,m)$ possiamo definire un elemento neutro destro $\hat{\mathfrak{I}}_d \in \mathcal{M}(m,m)$ e un elemento neutro sinistro $\hat{\mathfrak{I}}_s \in \mathcal{M}(n,n)$ tali che
$$\hat{A}\hat{\mathfrak{I}}_d = \hat{\mathfrak{I}}_s\hat{A} = \hat{A}\,.$$
Definiamo
$$(\hat{\mathfrak{I}})_{ij} = \delta_{ij}$$
dove δ_{ij} è la funzione delta di Kronecker che vale 1 se $i = j$ o 0 se $i \neq j$. Ossia la matrice identità $\hat{\mathfrak{I}}$ è una matrice quadrata con tutti gli elementi nulli, tranne quelli sulla diagonale che valgono 1. In generale se $n \neq m$ le matrici $\hat{\mathfrak{I}}_d$ e $\hat{\mathfrak{I}}_s$ sono diverse.

Queste sono le uniche proprietà strutturali del prodotto tra matrici. Rispetto alle proprietà del prodotto tra numeri reali salta subito agli occhi, come abbiamo già anticipato, l'assenza della proprietà commutativa, $\hat{A}\hat{B} = \hat{B}\hat{A}$ e l'assenza dell'elemento inverso. Vedremo che in alcuni casi è possibile recuperare queste proprietà, che, in generale, non sono verificate dall'operazione di prodotto tra matrici.

Trasposizione

Essendo le matrici delle tabelle di numeri, le possibilità di combinarne gli elementi per ottenere delle "nuove operazioni" (oltre quelle particolarmente semplici appena introdotte) sono praticamente illimitate. Ovviamente le operazioni tra matrici che prenderemo in considerazione dovranno anche essere utili per lo sviluppo della teoria o nelle applicazioni.
Un'operazione che si utilizza spesso nel calcolo matriciale è l'operazione di trasposizione, che consiste nel semplice scambio delle righe con le colonne di una matrice fissata.

Definizione 1.3.5. *Sia $\hat{A} \in \mathcal{M}(n,m)$. Definiamo $\hat{A}^t \in \mathcal{M}(m,n)$ la matrice ottenuta dalla matrice $\hat{A}$ scambiandone le righe con le colonne:*

$$\{\hat{A}^t\}_{ij} = a^t_{ij} = a_{ji} \tag{1.14}$$

Nella rappresentazione tabellare questo equivale a scrivere

$$\hat{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \rightarrow \hat{A}^t = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \ddots & \cdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix}$$

L'operazione di trasposizione può essere composta con le operazione di somma e prodotto tra matrici, verificando le seguenti identità facilmente dimostrabili:

$$(\alpha\hat{A} + \beta\hat{B})^t = \alpha\hat{A}^t + \beta\hat{B}^t \qquad (\alpha\hat{A}\hat{B})t = \alpha\hat{B}^t\hat{A}^t \,.$$

1.3.3 Le matrici quadrate

Un sottocaso molto importante dell'algebra matriciale riguarda le matrici quadrate. Per queste matrici è sempre possibile effettuare l'operazione di prodotto, infatti le matrici quadrate sono sempre conformabili con loro stesse. Quindi lo spazio delle matrici $n \times n$ oltre a essere uno spazio vettoriale (come nel caso dello spazio $\mathcal{M}(n,m)$), dotato quindi delle operazioni di somma e prodotto per scalare, ammette sempre anche l'operazione di prodotto tra matrici.

Definizione 1.3.6. *Lo spazio delle matrici quadrate $\mathcal{M}(n,n)$ con le operazioni di somma, prodotto per scalare e prodotto tra matrici, verificanti le proprietà m1-m6 e p1-p3, rispettivamente, si dice* **algebra delle matrici** $n \times n$.

L'importanza delle matrici quadrate diverrà evidente nello sviluppo della teoria degli operatori lineari su spazi vettoriali. Qui accenneremo solo al fatto che operazioni fondamentali come il calcolo del determinante, della traccia o degli autovalori sarà possibile solo per matrici quadrate. Inoltre, le matrici quadrate forniscono una rappresentazione di applicazioni lineari da spazi vettoriali in se stessi (endomorfismi), oppure una rappresentazione di prodotti scalari su spazi vettoriali.

1.3.4 Il determinante

Questa fondamentale operazione associa a ogni matrice $\hat{A} \in \mathcal{M}(n,n)$ un numero reale, che indicheremo con $\det(\hat{A}) \in \mathbb{R}$:

$$\det : \mathcal{M}(n,n) \to \mathbb{R}\,.$$

L'operazione di determinante trova applicazioni in moltissimi campi della matematica, dall'algebra lineare al calcolo differenziale a più dimensioni, dalla geometria differenziale alla teoria combinatoria. Molteplici sono anche le possibili interpretazioni di questa operazione. Per utilizzare un concetto che abbiamo già introdotto, definiremo e interpreteremo il determinante a partire dall'idea di dipendenza lineare tra vettori.
Dati $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ di coordinate

$$\mathbf{u} = \begin{pmatrix} a \\ c \end{pmatrix} \qquad \mathbf{v} = \begin{pmatrix} b \\ d \end{pmatrix}$$

determiniamo quale condizione devono verificare le componenti a, b, c, d perché $\mathbf{u}$ e $\mathbf{v}$ siano linearmente dipendenti. Se $\mathbf{u}$ e $\mathbf{v}$ devono essere linearmente dipendenti allora esisterà un $\alpha \in \mathbb{R}$ tale che $\mathbf{u} = \alpha\mathbf{v}$. Per cui

$$\begin{pmatrix} a \\ c \end{pmatrix} = \alpha \begin{pmatrix} b \\ d \end{pmatrix} \;\Rightarrow\; \begin{cases} a = \alpha b \\ c = \alpha d \end{cases} \;\Rightarrow\; \begin{cases} \alpha = \dfrac{a}{b} \\ \alpha = \dfrac{c}{d} \end{cases} \;\Rightarrow\; \frac{a}{b} = \frac{c}{d} \;\Rightarrow\; ad - cb = 0\,.$$

Ora, accoppiando i due vettori colonna in una matrice $\hat{A}$ 2×2, definiamo determinante di $\hat{A}$ la quantità $ad - cb$:

Definizione 1.3.7. *Sia $\hat{A} \in \mathcal{M}(2,2)$. Definiamo determinante di $\hat{A}$ la quantità*

$$\det(\hat{A}) = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - cb\,. \tag{1.15}$$

Quindi, due vettori colonna sono linearmente indipendenti se la matrice costruita a partire dalle loro componenti ha determinante diverso da zero. Ovvero, le colonne (o le righe) di una matrice sono linearmente indipendenti se il determinante della matrice è diverso da zero.
Per dare un'altra possibile interpretazione al determinante delle matrici 2×2 è semplice mostrare che il determinante di $\hat{A}$ è pari all'area del parallelogramma aventi come lati i vettori $\mathbf{u}$ e $\mathbf{v}$. Questa interpretazione è ovviamente compatibile con quella

di indipendenza lineare. Infatti due vettori sono linearmente dipendenti possono formare solo un parallelogramma singolare, la cui area è zero.

Partendo dalla definizione di determinante per le matrici 2×2, cercheremo una generalizzazione alle matrici quadrate $n\times n$. Esistono varie definizioni di determinante, e qui citeremo solo la più importante, la definizione assiomatica.

Definizione 1.3.8. *Il determinante è l'unica funzione*

$$\det : \mathcal{M}(n,n) \to \mathbb{R}$$

che verifica le seguenti proprietà:

a) $\det(\hat{\mathbb{I}}) = 1$*;*

b) se modifichiamo una matrice $\hat{A}$ *in modo da ottenere una nuova matrice* $\hat{B}$

1) scambiando due righe di $\hat{A}$*, allora* $\det(\hat{B}) = -\det(\hat{A})$*,*

2) moltiplicando per α *una riga di* $\hat{A}$*, allora* $\det(\hat{B}) = \alpha \det(\hat{A})$*,*

3) sommando una riga (o una colonna) a un'altra di $\hat{A}$*, allora* $\det(\hat{B}) = \det(\hat{A})$.

Le tre modifiche alla struttura di una matrice che abbiamo introdotto nella definizione di determinante (le modifiche, o mosse, $b1$), $b2$), e $b3$)) vengono chiamate mosse di Gauss. Vedremo nel paragrafo dedicato ai sistemi lineari il loro preciso significato. In questa sede, ci limitiamo a osservare come la struttura della funzione determinante è definita proprio dalle sue proprietà di trasformazione rispetto alle mosse di Gauss.

Theorem 1.3.9. *Sia* $\hat{A} \in \mathcal{M}(n,n)$*, allora* $\det(\hat{A}) = 0$ *se*

1) $\hat{A}$ *ha due righe (o due colonne) uguali;*

2) $\hat{A}$ *ha una riga (o colonna) nulla (composta di soli zero);*

3) $\hat{A}$ *ha due righe (o due colonne) linearmente dipendenti;*

4) le n *righe (o colonne) di* $\hat{A}$ *sono linearmente dipendenti.*

Dimostrazione. Prima di tutto consideriamo una matrice, $\hat{A}$, come un insieme di n vettori riga (o analogamente n vettori colonna).

1) Se costruiamo la matrice $\hat{B}$ con la mossa di Gauss b1 che scambia le due righe uguali, si avrà $\hat{A} = \hat{B}$, ma anche $\det(\hat{A}) = -\det(\hat{B})$, per cui $\det(\hat{A}) = 0$.

2) Se $\hat{A}$ ha una riga di tutti zero, usando la mossa di Gauss b3 è facile ottenere una matrice $\hat{B}$ con due colonne uguali, con $\det(\hat{B}) = 0$. Ma visto che $\det(\hat{A}) = \det(\hat{B})$ allora $\det(\hat{A}) = 0$.

3) Chiamando $\mathbf{u}$ e $\mathbf{v}$ le due righe linearmente dipendenti della matrice $\hat{A}$ si dovrà avere $\mathbf{u} = \alpha\mathbf{v}$. Se ora costruiamo una nuova matrice $\hat{B}$ moltiplicando per α la riga corrispondente al vettore $\mathbf{v}$, allora tale matrice avrà due colonne uguali, per cui $\det(\hat{B}) = 0$. Ma $det(\hat{B}) = \alpha \det(\hat{A})$ e quindi (visto che $\alpha \neq 0$, altrimenti la matrice $\hat{A}$ avrebbe una riga composta di soli zero, ricadendo nel caso 2)) $\det(\hat{A}) = 0$.

4) La dimostrazione di questo caso è una ripetizione del caso 3) iterando più mosse di Gauss.

□

1.3.5 Minori e complementi algebrici

La definizione assiomatica di determinante non ci offre un metodo di calcolo efficace che possa essere utilizzato per matrici quadrate di dimensione superiore a 2. Per introdurre una procedure generale di calcolo del determinante è necessario prima introdurre i concetti di minore e di complemento algebrico.

Definizione 1.3.10. *Data una matrice $\hat{A} \in \mathcal{M}(n,n)$ definiamo:*

- *$\hat{A}_{ij}$ la* **sottomatrice** *ottenuta eliminando da $\hat{A}$ la riga i e la colonna j. Pertanto se $\hat{A} \in \mathcal{M}(n,n)$ sarà $\hat{A}_{ij} \in \mathcal{M}(n-1, n-1)$;*
- *$A_{ij} = \det \hat{A}_{ij}$ il* **minore** *(a volte chiamato anche* **minore complementare***) ossia il determinante della sottomatrice $\hat{A}_{ij}$;*
- *$(-1)^{i+j} A_{ij}$ il* **complemento algebrico** *ottenuto moltiplicando il minore A_{ij} per il segno $(-1)^{i+j}$.*

1.3.6 Calcolo del determinante con il metodo di Laplace

Il teorema di Laplace per il calcolo del determinante afferma:

Theorem 1.3.11. *Il determinante di una matrice quadrata $\hat{A} \in \mathcal{M}(n,n)$ è pari alla somma dei prodotti di una qualunque riga (o colonna) per i rispettivi complementi algebrici:*

$$\det \hat{A} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} A_{ij} \qquad \forall j \tag{1.16}$$

$$\det \hat{A} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} A_{ij} \qquad \forall i \tag{1.17}$$

Quindi, per calcolare il determinante di una matrice occorre scegliere una riga (o una colonna) qualsiasi della matrice, ed effettuare la somma dei prodotti degli elementi di riga (o di colonna) per i rispettivi complementi algebrici. Ora, nei complementi algebrici, compare un'altra volta il calcolo del determinante, ma di una

sottomatrice con dimensione inferiore rispetto alla matrice di partenza. Per il calcolo di questo determinante si applica ancora il metodo di Laplace. Per cui tale metodo è iterativo, e si arresta solo una volta raggiunto il calcolo del determinante di una matrice 2x2, per il quale si applica la formula (1.15).

Esempio 1.3.12. *Calcolare il determinante della seguente matrice:*

$$\hat{A} = \begin{pmatrix} 1 & 1 & 2 & 5 \\ 1 & 0 & 2 & 0 \\ 3 & 2 & 6 & 5 \\ 1 & 6 & 2 & 0 \end{pmatrix}$$

Applicando il metodo di Laplace potremmo usare una riga o una colonna qualsiasi, ma sarà conveniente scegliere una riga o una colonna in cui compaiono molti zeri, in modo da non dover calcolare il complemento algebrico relativo a quegli elementi. Sceglieremo pertanto la seconda riga:

$$\det \hat{A} = \sum_{j=1}^{4} (-1)^{2+j} a_{2j} A_{2j} =$$

$$(-1)^3(1) \det \begin{pmatrix} 1 & 2 & 5 \\ 2 & 6 & 5 \\ 6 & 2 & 0 \end{pmatrix} + (-1)^4(0) \det \begin{pmatrix} 1 & 2 & 5 \\ 3 & 6 & 5 \\ 1 & 2 & 0 \end{pmatrix} +$$

$$+(-1)^5(2) \det \begin{pmatrix} 1 & 1 & 5 \\ 3 & 2 & 5 \\ 1 & 6 & 0 \end{pmatrix} + (-1)^6(0) \det \begin{pmatrix} 1 & 1 & 2 \\ 3 & 2 & 6 \\ 1 & 6 & 2 \end{pmatrix}$$

A questo punto dobbiamo calcolare solo il primo e il terzo determinante 3x3. Applicheremo ancora una volta il metodo di Laplace sviluppando per il primo determinante rispetto alla terza riga mentre per il terzo determinante rispetto alla prima riga:

$$\det \hat{A} = -1 \left((-1)^4(6) \det \begin{pmatrix} 2 & 5 \\ 6 & 5 \end{pmatrix} + (-1)^5(2) \det \begin{pmatrix} 1 & 5 \\ 2 & 5 \end{pmatrix} + (-1)^6(0) \det \begin{pmatrix} 1 & 2 \\ 2 & 6 \end{pmatrix} \right) +$$

$$-2 \left((-1)^2(1) \det \begin{pmatrix} 2 & 5 \\ 6 & 0 \end{pmatrix} + (-1)^3(1) \det \begin{pmatrix} 3 & 5 \\ 1 & 0 \end{pmatrix} + (-1)^4(5) \det \begin{pmatrix} 3 & 2 \\ 1 & 6 \end{pmatrix} \right) =$$

$$-1(6(-20) - 2(-5)) - 2(1(-30) - 1(-5) + 5(16)) = 110 - 110 = 0$$

Il risultato è corretto perché la prima e la terza colonna della matrice sono proporzionali, e quindi, per il teorema (1.3.9) il determinante deve essere pari a zero.

Theorem 1.3.13. *Data la matrice* $\hat{A} \in \mathcal{M}(n, n)$*, per l'operazione* $\det(\hat{A})$ *valgono le seguenti proprietà elementari:*

a) $\det(\alpha \hat{A}) = \alpha^n \det(\hat{A})$

b) Il determinante è una funzione lineare rispetto alle combinazioni lineari di ciascuna colonna (o riga). Se rappresentiamo una matrice come insieme di vettori colonna (o riga) si ha

$$\hat{A} = \begin{pmatrix} \begin{pmatrix} \\ \mathbf{u}_1 \\ \\ \end{pmatrix} & \begin{pmatrix} \\ \mathbf{u}_2 \\ \\ \end{pmatrix} & \cdots & \begin{pmatrix} \\ \mathbf{u}_n \\ \\ \end{pmatrix} \end{pmatrix},$$

per cui:

$$\det \begin{pmatrix} \begin{pmatrix} \\ \alpha\mathbf{v} + \beta\mathbf{w} \\ \\ \end{pmatrix} & \begin{pmatrix} \\ \mathbf{u}_2 \\ \\ \end{pmatrix} & \cdots & \begin{pmatrix} \\ \mathbf{u}_n \\ \\ \end{pmatrix} \end{pmatrix} =$$

$$= \alpha \det \begin{pmatrix} \begin{pmatrix} \\ \mathbf{v} \\ \\ \end{pmatrix} & \begin{pmatrix} \\ \mathbf{u}_2 \\ \\ \end{pmatrix} & \cdots & \begin{pmatrix} \\ \mathbf{u}_n \\ \\ \end{pmatrix} \end{pmatrix} + \beta \det \begin{pmatrix} \begin{pmatrix} \\ \mathbf{w} \\ \\ \end{pmatrix} & \begin{pmatrix} \\ \mathbf{u}_2 \\ \\ \end{pmatrix} & \cdots & \begin{pmatrix} \\ \mathbf{u}_n \\ \\ \end{pmatrix} \end{pmatrix}.$$

c) Se $\hat{A}$ è una matrice triangolare si ha

$$\det(\hat{A}) = \prod_{i=1}^{n} a_{ii}$$

d) $\det(\hat{A}^t) = \det(\hat{A})$

e) Teorema di Binet: $\det(\hat{A}\hat{B}) = \det(\hat{A})\det(\hat{B})$

1.3.7 Il rango delle matrici

Abbiamo visto come attraverso il calcolo del determinante è possibile sapere se le righe e le colonne di una matrice sono linearmente indipendenti. Il calcolo del determinante è però possibile solo per matrici quadrate.

Vediamo ora come definire un'operazione che permetta di identificare il numero massimo di righe o di colonne linearmente indipendenti per una generica matrice $n \times m$.

Data una matrice $\hat{A} \in \mathcal{M}(n, m)$ da questa è possibile estrarre molte sottomatrici quadrate di dimensione massima pari a $l = \min(n, m)$, selezionando il medesimo numero di righe e di colonne ed estraendo gli elementi che si trovano agli incroci.

Definizione 1.3.14. *Definiamo* **rango** *(o caratteristica) di una matrice $\hat{A} \in \mathcal{M}(n, m)$ la dimensione massima l per cui esiste una sottomatrice $l \times l$ di $\hat{A}$ con determinante non nullo. Indicheremo con*

$$Rg(\hat{A}) = l$$

il rango della matrice.

A differenza del determinante, da cui possiamo ottenere vari tipi di informazioni, il rango ci fornisce solo un'indicazione di tipo dimensionale, ossia ci fornisce il numero massimo di vettori linearmente indipendenti che possiamo estrarre dalla matrice $\hat{A}$.

In questo senso, la definizione che abbiamo dato di rango di matrice è perfettamente analoga a quella data di rango di vettori nel paragrafo 1.1.5. Infatti se prendiamo m vettori di $\mathbb{R}^n$, se chiamiamo $\hat{A} \in \mathcal{M}(n,m)$ la matrice che ha per righe questi vettori, e se chiamiamo l il rango di $\hat{A}$, ci dovrà essere una sottomatrice $l \times l$ di $\hat{A}$ a determinante non nullo. Questa sottomatrice sarà composta da l elementi di l vettori che, visto che il determinante è non nullo, saranno necessariamente linearmente indipendenti. Dal punto di vista geometrico, il rango può anche essere visto come la dimensione del sottospazio vettoriale che possiamo generare con le righe o le colonne di una data matrice $\hat{A}$. Infatti, se abbiamo $l = \mathrm{Rg}(\hat{A})$ vettori indipendenti nella matrice $\hat{A}$ allora anche la dimensione del sottospazio generato da questi vettori sarà pari a l.

L'operazione di calcolo del rango di una matrice è sicuramente più elaborata del calcolo del determinante. Se abbiamo una matrice $\hat{A}$ 4×3 da questa possiamo estrarre 4 matrici 3×3 e 24 matrici 2×2. Se riusciamo a trovare subito una matrice 3×3 a determinante non nullo, bene, allora il $\mathrm{Rg}(\hat{A}) = 3$ e abbiamo concluso. Ma se tutte le matrici 3×3 che possiamo estrarre da $\hat{A}$ sono a determinante nullo, allora dobbiamo iniziare a provare con le matrici 2×2. Se la matrice di partenza è ancora più grande è evidente come i calcoli possano farsi lunghi. Esiste però un teorema (che non dimostreremo) che ci può essere di aiuto:

Theorem 1.3.15. *Condizione necessaria e sufficiente affinché una matrice abbia rango l è che esista una sottomatrice $\hat{B}(l) \in \mathcal{M}(l,l)$ a determinante non nullo, e che tutte le sottomatrici $l+1 \times l+1$ che contengono $\hat{B}(l)$ (o che, equivalentemente, la orlano) abbiano tutte determinante nullo.*

1.3.8 La matrice inversa

Abbiamo visto come lo spazio delle matrici quadrate, a differenza degli spazi vettoriali semplici, ammette sempre un'operazione di prodotto tra matrici. Per analogia al caso dei numeri reali, dove per ogni $x \neq 0$ esiste sempre un numero y inverso di x pari a $y = 1/x$ tale che $xy = 1$ (l'elemento neutro dell'operazione prodotto), vediamo se nel caso delle matrici quadrate è possibile introdurre una matrice inversa e quindi, in qualche senso, definire l'operazione di divisione tra matrici.

Definizione 1.3.16. *Data una matrice $\hat{A} \in \mathcal{M}(n,n)$ diciamo che la matrice $\hat{A}^{-1}$ è la sua inversa se $\hat{A}^{-1}\hat{A} = \hat{A}\hat{A}^{-1} = \hat{\mathbb{1}}$*

Questa è solo una definizione che non permette di sapere né se una matrice possa ammettere un'inversa, né come calcolarla.

Theorem 1.3.17. *Data una qualsiasi matrice $\hat{A} \in \mathcal{M}(n,n)$ essa ammette inversa se e solo se* $\det(\hat{A}) \neq 0$. *Gli elementi di tale matrice si calcolano come*

$$\left(\hat{A}^{-1}\right)_{ij} = a^{-1}_{ij} = \frac{(-1)^{i+j} A_{ji}}{\det(\hat{A})}\,.$$

Dimostrazione. È importante notare subito come nella definizione di matrice inversa compare il minore che appartiene alla posizione ji e non ij (è come se per i

complementi algebrici prendessimo quelli relativi alla matrice trasposta $\hat{A}^t$). Ora, se $\det(\hat{A}) = 0$, non è possibile applicare la formula e quindi definire una matrice inversa. Nel caso in cui $\det(\hat{A}) \neq 0$ verifichiamo che $\hat{A}\hat{A}^{-1} = \hat{\mathbb{I}}$:

$$(\hat{A}\hat{A}^{-1})_{ij} = \sum_{k=1}^{n} a_{ik} \frac{(-1)^{j+k} A_{jk}}{\det(\hat{A})} = \frac{1}{\det(\hat{A})} \sum_{k=1}^{n} (-1)^{j+k} a_{ik} A_{jk} \,.$$

Ora, considerando l'espressione a destra, se $i = j$ abbiamo nell'operazione di somma la definizione di determinante, e quindi $(\hat{A}\hat{A}^{-1})_{ii} = 1$, mentre se $i \neq j$ è come se calcolassimo il determinante di una matrice con due righe uguali (se $i \neq j$ la riga a_{ik} è contenuta anche nei minori A_{jk}) e quindi $(\hat{A}\hat{A}^{-1})_{ij \neq i} = 0$. □

1.3.9 Applicazione: sistemi lineari $n \times n$

Come prima applicazione del calcolo matriciale vediamo come sia possibile calcolare la soluzione dei sistemi lineari quadrati non singolari (ossia con matrice associata a determinante non nullo).

Dato il sistema di n equazioni in n incognite:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots & \vdots \quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{cases}$$

una volta definite

$$\hat{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

dove $\hat{A}$ è la matrice dei coefficienti, $\mathbf{x}$ il vettore delle incognite e $\mathbf{b}$ il vettore dei termini noti, può essere scritto come

$$\hat{A}\mathbf{x} = \mathbf{b} \,.$$

Se il sistema è non singolare, ossia se $\det(\hat{A}) \neq 0$, allora, moltiplicando a sinistra entrambi i membri per l'inversa di $\hat{A}$, si ottiene $\hat{A}^{-1}(\hat{A}\mathbf{x} = \mathbf{b}) \rightarrow \hat{A}^{-1}\hat{A}\mathbf{x} = \hat{A}^{-1}\mathbf{b} \rightarrow$

$$\mathbf{x} = \hat{A}^{-1}\mathbf{b} \,, \tag{1.18}$$

che è l'unica soluzione del sistema. In termini di componenti si ha:

$$x_i = \sum_{k=1}^{n} (\hat{A}^{-1})_{ik} b_k = \frac{1}{\det(\hat{A})} \sum_{k=1}^{n} (-1)^{i+k} b_k A_{ki} \,.$$

Ora, il termine di somma può essere interpretato come il determinante della matrice $\hat{A}$ in cui è stata sostituita la colonna $i-$esima con il vettore $\mathbf{b}$, ossia la formula di

Cramer: chiamando $\hat{B}_i$ la matrice ottenuta sostituendo alla colonna $i-$esima di $\hat{A}$ il vettore dei termini noti:

$$\hat{B}_i = \begin{pmatrix} a_{11} & a_{12} & \cdots & b_1 & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & b_2 & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & b_n & \cdots & a_{nm} \end{pmatrix}.$$

si ha

$$x_i = \frac{\det(\hat{B}_i)}{\det(\hat{A})} \tag{1.19}$$

È importante osservare che la teoria dei sistemi lineari quadrati possiede una peculiare interpretazione nell'ambito della teoria degli spazi vettoriali. Se rappresentiamo la matrice dei coefficienti con insieme di vettori (colonna) si ha:

$$\hat{A}\mathbf{x} = \mathbf{b} \quad \rightarrow \quad \left(\begin{pmatrix} \\ \mathbf{u}_1 \\ \\ \end{pmatrix} \quad \begin{pmatrix} \\ \mathbf{u}_2 \\ \\ \end{pmatrix} \quad \cdots \quad \begin{pmatrix} \\ \mathbf{u}_n \\ \\ \end{pmatrix} \right) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Un momento di riflessione sull'operazione di prodotto righe per colonne, porta a semplificare l'espressione appena scritta in:

$$\hat{A}\mathbf{x} = x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_n\mathbf{u}_n = \mathbf{b}$$

che nel contesto degli spazi vettoriali ci dice che il sistema $\hat{A}\mathbf{x} = \mathbf{b}$ ha soluzione se e solo se $\mathbf{b}$ può essere ottenuto come combinazione lineare a coefficienti x_i dei vettori $\mathbf{u}_i$ che compongono le colonne della matrice $\hat{A}$. Se il determinante di $\hat{A}$ è diverso da zero, allora i vettori $\mathbf{u}_i$ sono linearmente indipendenti ed esisterà sicuramente una ed una sola combinazione lineare (unicità della decomposizione, vedi il paragrafo 1.2.8) da cui ottenere qualsiasi vettore $\mathbf{b}$. Nel caso in cui il determinante di $\hat{A}$ è pari a zero, allora ci sono soluzioni al sistema solo se il vettore $\mathbf{b}$ è linearmente dipendente dalle colonne della matrice $\hat{A}$. Per stabilire questa eventualità non si dovrà far altro che calcolare il rango della matrice $\hat{A}$ (che determina il numero di vettori colonna linearmente indipendenti presenti in $\hat{A}$) ed il rango della matrice $\hat{B} = (\hat{A}|\mathbf{b})$, ossia la matrice $\hat{A}$ a cui è stata aggiunta la colonna dei termini noti. Se il rando di $\hat{B}$ è pari al rango di $\hat{A}$ allora il vettore dei termini noti è linearmente dipendente dai vettori colonna che compongono la matrice $\hat{A}$. Ma di questo parleremo più approfonditamente alla fine del prossimo paragrafo.

Come ultima considerazione, le mosse di Gauss presentate nella definizione 1.3.8 acquistano molta chiarezza riflettendo sul loro significato in termini di sistemi lineari. Senza entrare troppo nel dettaglio è chiaro che la soluzione di un sistema non può dipendere dal fatto che io cambi di posto due equazioni (mossa 1) oppure che io sommi tra loro due equazioni (mossa 3), oppure moltiplichi una equazione per α. Tutte queste mosse devono riflettersi in maniera semplice sul determinante dei coefficienti del sistema, che, in ultima analisi, ci fornisce le soluzioni del sistema.

Per la teoria generale delle soluzioni di sistemi lineari (anche a determinante nullo o nel caso di sistemi non quadrati) rimandiamo alla fine del prossimo capitolo dopo aver introdotto ed analizzato le funzioni lineari su spazi vettoriali.