

UNIVERSITÀ DEGLI STUDI DI FIRENZE  
Dipartimento di Matematica per le Decisioni

# MetM@t

## Metodi e Modelli Lineari



PierLuigi Zezza <sup>1</sup>

Versione 1.0 del 13 Dicembre 2007

---

<sup>1</sup><http://www.dmd.unifi.it/zezza>

PierLuigi Zezza  
Università degli Studi di Firenze  
Dipartimento di Matematica per le Decisioni  
Via Lombroso 6/17 - 50134 FIRENZE, ITALIA  
Tel. +39 055 4796830  
Fax +39 055 4796800  
e-mail: [pierluigi.zezza@unifi.it](mailto:pierluigi.zezza@unifi.it)  
<http://www.dmd.unifi.it/zezza>

# Indice

<b>Prefazione</b>	<b>5</b>
<b>I Metodi lineari</b>	<b>9</b>
<b>1 Vettori e matrici</b>	<b>11</b>
1.1 Vettori . . . . .	12
1.1.1 Operazioni fra vettori . . . . .	12
1.1.2 Ordinamenti . . . . .	14
1.1.3 Il duale . . . . .	15
1.2 Matrici . . . . .	16
1.2.1 Operazioni fra matrici. . . . .	17
1.2.2 Matrici a scala e simmetriche . . . . .	19
1.3 Spazi vettoriali . . . . .	20
1.3.1 Spazio vettoriale . . . . .	21
1.3.2 Sottospazi . . . . .	25
Semispazi . . . . .	28
<b>2 Sistemi lineari</b>	<b>29</b>
2.1 Risolviamo un sistema lineare . . . . .	31
2.1.1 Un Metodo Risolutivo . . . . .	33
2.1.2 L'algoritmo di Gauss . . . . .	38
2.1.3 Possibili risultati finali della procedura di riduzione . . . . .	39
Esempi . . . . .	39
2.2 Ancora sull'algoritmo di Gauss . . . . .	40
2.2.1 L'algoritmo di Gauss-Jordan . . . . .	41
2.2.2 Matrici inverse e soluzioni con diversi termini noti . . . . .	43
2.3 Considerazioni storiche . . . . .	46

<b>3</b>	<b>Complementi sugli spazi vettoriali</b>	<b>49</b>
3.1	Sottospazi e sottospazi affini . . . . .	49
	Iperpiani . . . . .	49
3.2	Relazioni con la geometria Euclidea . . . . .	55
<b>II</b>	<b>Modelli lineari</b>	<b>65</b>
<b>4</b>	<b>Modelli matematici</b>	<b>67</b>
4.1	Vettori e matrici . . . . .	67
4.1.1	Teoria del consumatore . . . . .	67
	Teoria del consumatore . . . . .	69
4.1.2	Ordinamenti parziali* . . . . .	69
4.1.3	Modelli lineari di produzione . . . . .	70
4.2	Sistemi lineari . . . . .	74
4.2.1	Mercati finanziari . . . . .	74
	Alcune considerazioni sul significato del modello . . . . .	77
4.2.2	Un mercato finanziario . . . . .	78
	Verifica della completezza . . . . .	79
	Determinazione del vettore dei prezzi . . . . .	79
	Costruzione dei titoli elementari . . . . .	80
	Riflessione sul significato della non esistenza di un arbitraggio . . . . .	80
	Mercati incompleti . . . . .	82
	Arbitraggi* . . . . .	83
<b>III</b>	<b>Soluzioni</b>	<b>87</b>
	Soluzioni degli Esercizi . . . . .	90

# Prefazione

Obiettivo di questa parte del testo è la presentazione di una classe di modelli che sono di vasta applicazione e la cui analisi può essere effettuata con metodi elementari.

Il compito che abbiamo davanti richiede una descrizione abbastanza dettagliata. Quando vogliamo analizzare un qualche fenomeno naturale o sociale utilizzando la matematica, dobbiamo cercare di individuare quali sono le proprietà che caratterizzano il fenomeno, queste proprietà devono poter essere quantificate e cioè devono poter essere espresse con dei numeri. Non è quindi possibile fare un modello matematico per decidere se un quadro è bello oppure no, poiché il giudizio estetico dipende da molti fattori che non sono quantificabili.

Facciamo un esempio molto semplice legato al vostro quotidiano in cui la modellizzazione è possibile. Vogliamo descrivere la transazione economica fra un cliente che fa la spesa ed il supermercato dove la spesa viene effettuata, in altre parole vogliamo fare il conto della spesa, per fare questo dobbiamo determinare il *valore* del carrello con il quale il cliente si presenta alla cassa. La proprietà essenziale da determinare è il valore della merce; questa viene rappresentata dal numero ottenuto moltiplicando il prezzo unitario della merce per la quantità da acquistare. Dal punto di vista del conto finale quantità eguali di due merci che hanno lo stesso prezzo sono eguali, esse sono quindi rappresentate dallo stesso simbolo (numero). Dimentichiamo per un attimo il nostro problema e consideriamo l'insieme dei numeri, noi sappiamo che con i numeri è possibile effettuare delle operazioni, la somma ed il prodotto, *l'insieme dei numeri reali con queste due operazioni è il nostro modello*. Torniamo al nostro problema cosa accade alla cassa: i prezzi vengono moltiplicati per un numero, la quantità di quella merce che volete acquistare, ed i risultati vengono fra loro sommati, il risultato è il valore della spesa che avete fatto, pagate ed andate via; il modello ha così esaurito il suo compito: ha rappresentato un fenomeno sociale, come avviene la transazione fra cliente e supermercato. Notate che lo scontrino che vi viene dato rappresenta esattamente questo modello.

In questo esempio il fenomeno che ci interessa è rappresentato da un unico nu-

mero che chiamiamo *stato del sistema*, vedremo molto presto che modelli un po' più complicati necessitano di rappresentazione diversa. Nell'insieme dei numeri sono definite delle operazioni che rappresentano delle operazioni che effettivamente possiamo fare sugli stati, l'insieme degli stati con le operazioni che si possono fare si chiama *spazio degli stati* ed infine le semplici operazioni che abbiamo fatto per calcolare il conto della spesa, moltiplicare per un numero e sommare i risultati si chiamano *combinazione lineare*. Dobbiamo fare un ulteriore commento sull'esempio che abbiamo analizzato. Abbiamo considerato come spazio degli stati, per il modello in esame, l'insieme  $\mathbb{R}$  di tutti i numeri reali; questo, in effetti, non è necessario poiché prezzi e quantità sono sempre numeri razionali ed inoltre, sempre positivi. Ma l'aver preso lo spazio degli stati più grande del necessario non comporta nessun inconveniente, vedremo degli esempi in cui l'allargamento può generare delle *soluzioni* che non sono accettabili perché non corrispondono a degli stati realmente possibili.

Rimaniamo nel supermercato e supponiamo di voler studiare un problema del tipo: fra tutti i possibili carrelli che contengono almeno una certa quantità fissata di carne, verdure, ... , qual'è quello che costa meno? In questo caso lo stato del sistema non è più rappresentabile da un solo numero, per rappresentare un carrello qualsiasi dobbiamo fare una lista delle quantità delle diverse merci che contiene. Ogni carrello è identificato da questa lista e quindi lo stato del sistema è una lista. Il resto della modellizzazione richiede opportune notazioni per essere chiara, più avanti nel testo vedrete problemi simili a questo e sarete in grado di risolverli.

Oggetto di questa parte del testo è lo studio di modelli del tipo

$$\begin{array}{ll} \text{Ottimizzare} & f(x) \\ & x \in M \end{array}$$

dove la variabile di stato  $x$  appartiene ad uno *spazio lineare*, la funzione che descrive l'obiettivo è una funzione *lineare* ed il vincolo  $M$  è descritto da un numero finito di equazioni *lineari*. Cosa voglia dire *lineare* in questo contesto sarà spiegato in seguito, ma, essenzialmente, vuol dire, che le uniche operazioni ammesse sono la somma ed il prodotto per un numero.

L'ultima cosa che vogliamo sottolineare è che la variabile di stato  $x$  è, da una parte, un simbolo che maneggiamo secondo regole algebriche e dall'altra rappresenta uno stato particolare del fenomeno che stiamo analizzando; ad esempio, il numero 7 nello scontrino del supermercato viene sommato agli altri secondo le regole dell'algebra, ma corrisponde ad una scatola di biscotti o ad 80 grammi di formaggio. Da questo punto di vista le idee di spazio e di punto dello spazio sono sostanzialmente diverse dalle idee intuitive di spazio e di punto della vostra intuizione geometrica. Quello che vedremo insieme è che invece l'intuizione che avete

dello spazio è rappresentativa di **tutti** questi modelli e che se prescindiamo dal significato dei simboli le operazioni che dobbiamo effettuare sono elementari.

Abbiamo utilizzato ripetutamente la parola simbolo per sottolineare che per fare la somma ed il prodotto di numeri bisogna soltanto conoscere le tabelline, non c'è alcun bisogno di sapere che cosa sono i numeri o cosa rappresentano, una *stupidissima* calcolatrice è perfettamente in grado di farlo.

La presentazione è volutamente concreta e non formale e si ispira a quella che nel mondo anglo-sassone è chiamata *Numerical Linear Algebra*. Il nostro obiettivo è quello di fornire dei metodi matematici che siano utili sia per trovare che per studiare le soluzioni di modelli corrispondenti a problemi concreti. Sono stati quindi privilegiati i metodi ai concetti astratti, senza però rinunciare al rigore. Come ulteriore testo di riferimento si veda ad esempio [?].



**Parte I**

**Metodi lineari**



# Capitolo 1

## Vettori e matrici

*Would you tell me please, which way I ought to go from here? said Alice. That depends a good deal on where you want to get to, said the Cat. I don't much care where, said Alice. Then it doesn't matter which way you go, said the Cat. ... so long as I get somewhere said Alice. Oh, you're sure to do that, if only you walk long enough said the Cat.*

LEWIS CARROL, *Alice in Wonderland*.

Come abbiamo detto nell'introduzione l'insieme dei numeri reali o meglio lo spazio  $\mathbb{R}$  dei numeri reali con le operazioni che vi sono definite, non è sufficiente per descrivere con abbastanza dettagli modelli della realtà economica, sociale e fisica. In questo capitolo introdurremo altri insiemi e descriveremo le operazioni che possiamo effettuare con gli elementi di questi insiemi, operazioni che saranno *simili* a quelle fra numeri reali. Nella seconda parte, per sostanziare questa parte più astratta descriveremo alcuni modelli di tipo economico e sociale che utilizzano questi concetti per la loro formalizzazione matematica.

Iniziamo descrivendo dei concetti elementari. Così come un singolo numero serve a misurare una singola proprietà di un oggetto come la temperatura dell'acqua del motore della vostra auto (novanta [gradi]), allo stesso modo, una lista di numeri serve a descrivere un insieme ordinato di quantità o caratteristiche. In molte applicazioni economiche e non solo economiche, si deve operare con questi insiemi ordinati di numeri che, per il loro significato, devono essere concepiti come un'unica entità.

## 1.1 Vettori

Un **vettore** è una lista e ordinata di numeri reali in cui sono essenziali sia l'ordine che il numero di elementi che la compongono. Indichiamo un vettore con una colonna verticale di numeri, ordinati a partire dall'alto, e chiamiamo ciascuno di questi numeri **componente** del vettore. Il numero delle componenti che formano un vettore si chiama **dimensione del vettore**. Per comodità le componenti vengono numerate da 1 a  $p$ , dove  $p$  è la dimensione del vettore. Se  $x$  è un vettore,  $x_2$  indica quindi la sua seconda componente. Possiamo pensare ad un vettore come ad un elemento dell'insieme

$$\mathbb{R}^p := \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{p \text{ volte}}$$

e quindi  $x \in \mathbb{R}^p$  vuol dire

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}.$$

Si chiama **vettore nullo** quello con tutte le componenti eguali a zero e lo indicheremo con  $0 \in \mathbb{R}^p$ , *un vettore non nullo ha quindi almeno una componente diversa da zero*.

**Nota.** Nel Capitolo 1.3 vedremo che il nome vettore viene utilizzato in generale per gli elementi di un insieme che ha proprietà analoghe a quelle che stiamo descrivendo per  $\mathbb{R}^p$ .

### 1.1.1 Operazioni fra vettori

Nello spazio  $\mathbb{R}^p$  dei vettori  $p$ -dimensionali sono definite due operazioni elementari.

• **L'addizione di due vettori**, che ha per risultato un altro vettore detto **somma** e che viene fatta componente per componente, si sommano cioè le componenti corrispondenti.

$$(x, y) \mapsto x + y := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_p + y_p \end{pmatrix}$$

per esempio se

$$x = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \quad \text{e} \quad y = \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} \quad \text{allora} \quad x + y = \begin{pmatrix} 0 \\ 4 \\ 1 \end{pmatrix}.$$

• La **moltiplicazione di un vettore per un numero**, che ha per risultato un vettore ed è anch'essa definita componente per componente, si effettua cioè moltiplicando per il numero ciascuna componente del vettore.

$$(\alpha, x) \mapsto \alpha x := \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_p \end{pmatrix}$$

per esempio se

$$\alpha = 2 \quad \text{e} \quad x = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} \quad \text{allora} \quad \alpha x = \begin{pmatrix} 2 \\ -4 \\ 6 \end{pmatrix}.$$

Nel seguito chiameremo **combinazione lineare** di due vettori  $x, y \in \mathbb{R}^p$  con **coefficienti**  $\alpha, \beta \in \mathbb{R}$  l'espressione

$$\alpha x + \beta y$$

ottenuta *combinando* le due operazioni precedenti.

Nello spazio  $\mathbb{R}^p$  è possibile definire un'ulteriore operazione.

• Il **prodotto scalare fra due vettori**, indicato con  $\langle \cdot, \cdot \rangle$  che ha per risultato un numero.

$$(x, y) \mapsto \langle x, y \rangle := x_1 y_1 + x_2 y_2 + \cdots + x_p y_p = \sum_{i=1}^p x_i y_i. \quad (1.1)$$

Per esempio se

$$x = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \quad \text{e} \quad y = \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} \quad \text{allora} \quad \langle x, y \rangle = -4 + 3 - 2 = -3.$$

Il prodotto scalare gode della *proprietà commutativa* cioè

$$\langle x, y \rangle = \langle y, x \rangle, \quad \text{per ogni } x, y \in \mathbb{R}^p$$

ed è *lineare rispetto a ciascun argomento*, cioè

$$\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle, \quad \text{per ogni } x, y \in \mathbb{R}^p, \alpha, \beta \in \mathbb{R},$$

come è semplice verificare direttamente dalla definizione.

Se  $n = 1$  il prodotto scalare coincide con l'usuale operazione di prodotto fra numeri reali. Per  $n > 1$  l'analogia con il prodotto fra due numeri reali non vale più,

ad esempio il prodotto scalare fra due vettori può essere zero senza che nessuno dei due sia il vettore nullo, ad esempio se

$$x = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \quad \text{e} \quad y = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

allora

$$\langle x, y \rangle = -1 - 1 + 2 = 0.$$

Se  $\langle x, y \rangle = 0$  i due vettori si dicono **ortogonali**.

**Teorema 1.1.** *Il solo vettore che è ortogonale a tutti gli altri è il vettore nullo.*

*Dimostrazione.* Supponiamo che sia

$$\langle a, x \rangle = 0 \quad \text{per ogni} \quad x \in \mathbb{R}^p,$$

vogliamo dimostrare che  $a = 0$ .

Abbiamo che  $\langle a, x \rangle = 0$  anche per  $x = a$  e quindi

$$0 = \langle a, a \rangle = \sum_{i=1}^p a_i^2,$$

ma la somma di quadrati è zero se e solo se tutti gli addendi sono zero e quindi  $a_i = 0$ ,  $i = 1, 2, \dots, p$ , ed  $a$  è quindi il vettore nullo.  $\square$

### 1.1.2 Ordinamenti

Molti dei problemi che incontrerete nelle scienze economiche e sociali richiedono di scegliere fra molte alternative possibili quella *migliore*. La parola migliore sottintende la possibilità di confrontare fra di loro le diverse situazioni o, in altre parole, di ordinarle secondo un certo criterio. Ad esempio, facendo la spesa potete decidere di acquistare i prodotti che costano di meno o potete decidere di aprire un conto corrente nella banca che vi offre il più alto tasso di interesse. In questi due esempi la scelta è semplice perché dovete scegliere fra numeri reali, nel primo caso il più piccolo e nel secondo il più grande. Se lo stato del sistema è più complicato ed è rappresentato da un vettore questo può non essere possibile in quanto non è possibile dare un **ordinamento totale** in  $\mathbb{R}^p$  compatibile con le operazioni, con totale si intende che dati due vettori è sempre possibile confrontarli fra di loro.

E' però possibile definire **ordinamento parziale**. Per definirlo procediamo nel seguente modo. Sia

$$\mathbb{R}_+^p = \{x \in \mathbb{R}^p : x_i \geq 0, i = 1, 2, \dots, p\}$$

e siano  $x, y \in \mathbb{R}^p$ . Diremo che

$$x \geq y \quad \text{se e solo se} \quad x - y \in \mathbb{R}_+^p$$

ovvero se le componenti di  $x$  sono tutte maggiori od eguali delle corrispondenti componenti di  $y$ , inoltre, diremo che

$$x > y \quad \text{se e solo se} \quad x_i > y_i, \quad i = 1, \dots, p$$

ossia tutte le componenti di  $x$  sono maggiori delle corrispondenti componenti di  $y$ . In  $\mathbb{R}_+^p$  questo **non** è l'unico modo per determinare un ordine parziale (vedi il Paragrafo 4.1.2), nella letteratura economica questi ordinamenti si chiamano **ordinamenti Paretiani** da Vilfredo Pareto<sup>1</sup>. L'ordinamento si chiama parziale perché a differenza dal caso  $p = 1$  esistono vettori che non sono confrontabili fra di loro, ad esempio se

$$x = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \quad \text{e} \quad y = \begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix}$$

allora  $x > y$  mentre se prendiamo

$$y_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

allora  $x$  e  $y_1$  non sono confrontabili.

### 1.1.3 Il duale

Accanto ai vettori colonna che abbiamo descritto consideriamo anche i **vettori riga**

$$q = (q_1 \quad q_2 \quad \dots \quad q_p).$$

Indichiamo l'insieme dei vettori riga con  $\mathbb{R}^{p*}$ , questo spazio si chiama **duale** di  $\mathbb{R}^p$ , e quindi  $q \in \mathbb{R}^{p*}$  vuol dire che  $q$  è un vettore riga con  $p$  componenti. Ovviamente il fatto che siano scritti in modo differente non cambia nulla delle proprietà che abbiamo descritto per i vettori colonna. I motivi di questa doppia notazione nascono dalle diverse interpretazioni che possiamo dare ai due tipi di vettori e che vedremo successivamente.

<sup>1</sup><http://it.wikipedia.org/wiki/Vilfredo.Pareto>

Se consideriamo un vettore riga  $q = (q_1 \ q_2 \ \dots \ q_p)$  ed un vettore colonna  $k$ , entrambi con  $p$  componenti, possiamo definire il **prodotto fra una riga ed una colonna della stessa lunghezza** nel seguente modo

$$qk := q_1 k_1 + q_2 k_2 + \dots + q_p k_p = \sum_{i=1}^p q_i k_i. \quad (1.2)$$

Qualche volta avremo bisogno di scrivere un vettore riga come colonna e viceversa. Chiameremo questa operazione come **trasposizione** di un vettore  $x$  e la indichiamo con  $x^\top$  che chiameremo il vettore **trasposto** di  $x$ , ad esempio se

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \quad \text{allora} \quad x^\top = (x_1 \ x_2 \ \dots \ x_p)$$

e se

$$\pi = (\pi_1 \ \pi_2 \ \dots \ \pi_p) \quad \text{allora} \quad \pi^\top = \begin{pmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_p \end{pmatrix}.$$

Scrivere un vettore  $x \in \mathbb{R}^p$  come una colonna non è conveniente da un punto di vista tipografico, adatteremo quindi una notazione piuttosto diffusa scrivendo un vettore colonna sempre come una riga ma con la seguente convenzione: *in un vettore colonna gli elementi sono separati da virgole mentre in un vettore riga sono separati da spazi* e quindi

$$x = (3, 0, 1)$$

è un vettore colonna, mentre

$$q = (3 \ 0 \ 1)$$

è un vettore riga.

Il legame fra il prodotto scalare (1.1) ed il prodotto righe per colonne (1.2) è descritto dalla relazione

$$\langle x, y \rangle \equiv x^\top y = y^\top x,$$

che potete verificare facilmente.

## 1.2 Matrici

In alcuni problemi una lista di numeri non è sufficiente a descrivere gli oggetti del nostro studio (stato del sistema) dobbiamo definire quindi qualcosa di più generale.

Una **matrice** è una collezione bidimensionale di numeri. Per quanto riguarda gli ordinamenti si segue la stessa convenzione della lettura delle lingue europee. Iniziamo da *in alto a sinistra* e terminiamo *in basso a destra*. Il primo ordinamento è dall'alto in basso e definisce il numero di **righe**, il secondo da sinistra a destra e definisce il numero di **colonne**, se questi due numeri sono eguali la matrice si chiama **quadrata**.

Una matrice  $A$  con due righe e tre colonne si scrive quindi come

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}.$$

*L'insieme delle matrici a coefficienti reali con  $m$  righe ed  $n$  colonne si indica con  $M_{m \times n}(\mathbb{R})$ .*

### 1.2.1 Operazioni fra matrici.

Anche in questo spazio sono definite delle operazioni elementari analoghe a quelle definite sui vettori di  $\mathbb{R}^p$ .

- La **somma fra matrici** è definita componente per componente per cui si sommano fra di loro le componenti con gli stessi indici.
- Il **prodotto di una matrice per un numero** in cui ciascuna componente della matrice viene moltiplicata per quel numero.

Per definire un prodotto fra due matrici  $A, B$  possiamo vedere la prima come una collezione di righe e la seconda come una collezione di colonne ed eseguire il prodotto righe per colonne, che già conosciamo, in tutti i modi possibili riportando i risultati su una nuova matrice che chiameremo prodotto delle due. Evidentemente ciò è possibile solo se il numero delle colonne di  $A$  è eguale al numero di righe di  $B$ .

- Il **prodotto righe per colonne fra due matrici**  $A \in M_{p \times m}(\mathbb{R})$  e  $B \in M_{m \times n}(\mathbb{R})$  ha per risultato una matrice  $C \in M_{p \times n}(\mathbb{R})$  il cui elemento  $c_{ij}$  si ottiene facendo il prodotto della  $i$ -ma riga della matrice  $A$  con la  $j$ -ma colonna della matrice  $B$  ossia

$$c_{ij} = \sum_{k=1}^m a_{ik} b_{kj}.$$

Se, ad esempio,

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

allora

$$AB = \begin{pmatrix} 5 & 9 & 11 \\ 0 & 7 & 5 \end{pmatrix},$$

in cui ad esempio l'elemento  $AB_{23}$  è ottenuto moltiplicando la seconda riga per la terza colonna. È importante notare come il ruolo di righe e colonne nel prodotto fra matrici sia diverso, così che l'ordine in cui vengono prese le due matrici da moltiplicare è di cruciale importanza: in generale si può fare il prodotto  $AB$  ma non si può fare il prodotto  $BA$  (vedi l'esempio precedente in cui il prodotto  $BA$  non è definito). Inoltre, anche se è possibile effettuare entrambi i prodotti, il prodotto fra matrici non è, in generale, commutativo, ovvero  $AB \neq BA$ .

**Esempio.** Consideriamo le due matrici quadrate

$$A = \begin{pmatrix} 4 & 1 & -5 \\ 4 & -4 & -2 \\ 4 & -4 & 1 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} -2 & 2 & -7 \\ 4 & -7 & 7 \\ -3 & 3 & 5 \end{pmatrix}.$$

Se facciamo i due prodotti  $AB$  e  $BA$ , otteniamo

$$AB = \begin{pmatrix} 11 & -14 & -46 \\ -18 & 30 & -66 \\ -27 & 39 & -51 \end{pmatrix}$$

mentre

$$BA = \begin{pmatrix} -28 & 18 & -1 \\ 16 & 4 & 1 \\ 20 & -35 & 14 \end{pmatrix}.$$

Un vettore (colonna) può essere visto come una matrice con una sola colonna, in questo caso il secondo indice (quello di colonna) non è necessario, analogamente un vettore riga può essere visto come una matrice con una sola riga. Si noti come il prodotto righe per colonne di un vettore riga per un vettore colonna, definito in precedenza, è un caso particolare del prodotto fra una matrice  $1 \times p$  (vettore riga) ed una  $p \times 1$  (vettore colonna).

Analogamente a quanto fatto per i vettori si definisce la **trasposta**  $A^T$  di una matrice  $A$  come la matrice che si ottiene scambiando i ruoli di righe e colonne, ad esempio se

$$A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & -1 \end{pmatrix} \quad \text{allora} \quad A^T = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \end{pmatrix}.$$

Ciò significa che se  $a_{ij}$  è il generico elemento di  $A$ , posto  $B = A^T$ , allora il generico elemento  $b_{ij}$  di  $B$  è dato da  $b_{ij} = a_{ji}$ . Le proprietà elementari dell'operazione di trasposizione sono le seguenti

- $(aA + bB)^T = aA^T + bB^T$   
Questa proprietà è di verifica immediata tramite la definizione
- $(AB)^T = B^T A^T$   
Dalla definizione segue che i due prodotti si scrivono come

$$(AB)^T = \sum a_{jk} b_{ki}$$

e

$$B^T A^T = \sum b_{ki} a_{jk}$$

e sono quindi eguali

- $(A^T)^T = A$   
Questa proprietà è di verifica immediata tramite la definizione

### 1.2.2 Matrici a scala e simmetriche

Nel seguito, saranno di particolare importanza alcune matrici che mostrano una distribuzione particolare di elementi nulli. Premettiamo una definizione, il primo elemento non nullo di ogni riga, se c'è, si chiama **pivot**.

#### Matrici a scala.

Una matrice si dice **a scala** quando

- Se vi è una riga nulla anche tutte le righe sottostanti sono nulle.*
- Il pivot di ogni riga non nulla si trova più a destra (ha un indice di colonna maggiore) del pivot della riga precedente*

Per esempio la matrice

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

è a scala ed ha come pivot  $a_{11} = 1, a_{22} = 1, a_{33} = -2$ . Le matrici a scala non sono necessariamente quadrate. Ad esempio le matrici

$$S = \begin{pmatrix} 2 & -2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad T = \begin{pmatrix} -1 & 1 & 3 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

sono matrici a scala ed hanno come pivot, rispettivamente, 2, 1, -1 e -1, 2, 1, mentre le matrici

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

non sono a scala.

ESERCIZIO 1.1. Spiegare perché le tre matrici  $A, B, C$  non sono a scala.

### Matrici simmetriche e diagonali

Una matrice quadrata  $A \in M_{p \times p}(\mathbb{R})$  si dice **simmetrica** se  $A = A^T$  e questo vuole dire che  $a_{ij} = a_{ji}$  per ogni  $i, j$ . Si noti che non vi sono restrizioni sui valori degli elementi diagonali di una matrice simmetrica.

Una matrice quadrata si dice *diagonale* se tutti gli elementi che non appartengono alla diagonale sono zero, ad esempio la matrice

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

è diagonale. Fra le matrici diagonali ha particolare importanza la matrice **identità**  $I_p$  che ha solo elementi eguali ad uno sulla diagonale.

$$I_p = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

## 1.3 Spazi vettoriali

Anche se il nostro obiettivo finale rimane quello di descrivere un metodo che ci permetta di trovare le soluzioni di sistemi lineari di equazioni, come quelli che abbiamo visto, risulterà particolarmente utile introdurre un ambiente astratto nel quale sarà semplice descrivere i metodi risolutivi di questi problemi, metodi che sarà possibile applicare anche ad altri e nuovi problemi. Molti degli oggetti matematici di cui abbiamo parlato come i vettori, le matrici hanno lo stesso comportamento rispetto ad alcune operazioni elementari (somma, prodotto per un numero). Un comportamento simile hanno anche altre classi di oggetti, come i polinomi, le funzioni continue, quelle derivabili etc..

### 1.3.1 Spazio vettoriale

Un insieme  $V$  si dice **spazio vettoriale** su  $\mathbb{R}$  se in  $V$  sono definite due operazioni: somma e prodotto per un numero reale in modo che i suoi elementi, che chiameremo **vettori** possano essere sommati fra di loro e moltiplicati per un numero reale in modo che valgano le seguenti proprietà:

- **Proprietà associativa e commutativa della somma**

$$(v + w) + z = w + (w + z) \quad \text{per ogni } v, w, z \in V$$

$$v + w = w + v \quad \text{per ogni } v, w \in V.$$

- Esiste in  $V$  l'**elemento neutro** della somma, che indichiamo con  $0$ , tale che

$$v + 0 = v \quad \text{per ogni } v \in V.$$

- Esiste in  $V$  l'**inverso rispetto all'addizione** di ogni elemento  $v \in V$  che indicheremo con  $-v$  e che chiameremo l'opposto di  $v$ . Esso ha la proprietà che

$$v - v = 0.$$

- Per ogni  $v \in V$  valgono le due relazioni

$$1v = v, \quad 0v = 0,$$

dove  $1, 0$  sono l'unità e lo zero di  $\mathbb{R}$ .

- **Proprietà distributiva del prodotto rispetto alla somma**

$$a(v + w) = av + aw \quad \text{per ogni } v, w \in V, a \in \mathbb{R}$$

Siano  $v_1, \dots, v_p$  dei vettori in  $V$  e  $\alpha_1, \dots, \alpha_p$  dei numeri reali, allora il vettore

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_p v_p = \sum_{i=1}^p \alpha_i v_i$$

si dice **combinazione lineare** di  $v_1, \dots, v_p$  con coefficienti  $\alpha_1, \dots, \alpha_p$ .

#### Esempio.

1. L'insieme  $P_k[x]$  dei polinomi nella variabile  $x$  di grado minore o eguale a  $k$  è uno spazio vettoriale rispetto alle usuali operazioni di somma fra polinomi e di prodotto per un numero.

2. L'insieme  $\mathbb{R}^p$  delle  $p$ -ple ordinate di numeri reali è uno spazio vettoriale.
3. L'insieme  $M_{m \times n}(\mathbb{R})$  delle matrici con  $m$  righe ed  $n$  colonne è uno spazio vettoriale.
4. L'insieme  $C^1(a, b)$  delle funzioni  $\phi : (a, b) \rightarrow \mathbb{R}$  derivabili sull'intervallo aperto  $(a, b)$ , o quello delle funzioni continue sullo stesso intervallo, che indicheremo con  $C^0(a, b)$ , è uno spazio vettoriale rispetto alle usuali operazioni di somma di funzioni e di prodotto di una funzione per un numero ed inoltre  $C^1(a, b) \subset C^0(a, b)$ .

Uno dei concetti fondamentali negli spazi vettoriali è quello di indipendenza lineare, diamo allora la seguente definizione, i vettori  $v_1, \dots, v_k \in V$  si dicono **linearmente dipendenti** se esiste una loro combinazione lineare con coefficienti  $\alpha_i$  non tutti nulli tale che

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = 0.$$

Si dicono **linearmente indipendenti** se questo non è possibile.

Questa definizione risponde alla seguente domanda: *dati i vettori  $v_1, \dots, v_k \in V$  è possibile esprimere uno di essi come combinazione lineare degli altri?*

Se i vettori sono linearmente dipendenti allora la risposta è positiva, infatti se

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = 0.$$

e, per esempio,  $\alpha_j \neq 0$  allora

$$v_j = - \sum_{i=1, i \neq j}^k \frac{\alpha_i}{\alpha_j} v_i.$$

### Esempio.

1. Il vettore  $v_1$  è linearmente dipendente se e solo se  $v = 0$ . Lo studente verifichi questa affermazione tramite la definizione.
2. Due vettori sono linearmente dipendenti se e solo se uno è multiplo dell'altro e quindi, in particolare, due vettori di  $\mathbb{R}^3$  sono dipendenti se e solo se stanno sulla stessa retta passante per l'origine.
3. Tre vettori di  $\mathbb{R}^3$  sono dipendenti se e solo se stanno sullo stesso piano passante per l'origine.

4. I vettori in  $\mathbb{R}^p$ 

$$e_1 = (1, 0, \dots, 0) \quad e_2 = (0, 1, \dots, 0) \quad \dots \quad e_p = (0, 0, \dots, 1)$$

sono linearmente indipendenti, come è facile verificare direttamente dalla definizione.

5. I vettori  $p : x \mapsto x_1^2 + x_2, q : x \mapsto x_1 - x_2$  sono indipendenti poiché, per il principio di identità dei polinomi, il polinomio

$$(\alpha p + \beta q)(x) = \alpha x_1^2 + \beta x_1 + (\alpha - \beta)x_2$$

è zero se e solo se sia  $\alpha$  che  $\beta$  sono zero.

Negli spazi vettoriali un vettore può essere individuato dalle sue *coordinate*, ovvero una lista di numeri che esprimano la sua relazione con alcuni vettori di riferimento. Per esempio per individuare una casella degli scacchi data una lettera fra a e h ed un numero fra 1 e 8, oppure potete descrivere i polinomi di secondo grado  $ax^2 + bx + c$  tramite i loro coefficienti  $(a, b, c)$ . Per poter scegliere in modo efficace questi vettori di riferimento dobbiamo essere in grado di escludere che uno dei vettori scelti  $v_1, \dots, v_k \in V$ , si possa scrivere esso stesso come combinazione lineare degli altri. Questo è importante perché siamo interessati ad un insieme  $K$  di vettori con il minor numero di elementi possibile e se, ad esempio, abbiamo che

$$v = 4v_1 - 3v_2 + v_3 \quad \text{e} \quad v_2 = 3v_1 - v_3$$

allora possiamo scrivere

$$v = -5v_1 + 4v_3$$

e il vettore  $v_2$  è superfluo per descrivere il vettore  $v$ . Inoltre vogliamo scegliere  $v_1, \dots, v_k \in V$  in modo che ogni altro vettore si scriva in modo **unico** come loro combinazione lineare mentre nell'esempio precedente il vettore  $v$  poteva essere scritto in due modi diversi.

Consideriamo ora al problema di trovare opportuni vettori di riferimento rispetto ai quali scrivere tutti gli altri.

Diremo che uno spazio vettoriale  $V$  ha **dimensione finita** se esiste un insieme di vettori  $v_1, \dots, v_k \in V$  che generino tutto lo spazio  $V$ , ossia tali che ogni vettore  $v \in V$  possa essere scritto come combinazione lineare dei  $v_i$ ,

$$\text{per ogni } v \in V, \text{ esistono } \alpha_i \in \mathbb{R}, i = 1, \dots, k, : \sum_{i=1}^k \alpha_i v_i = v.$$

Se i vettori  $v_i$  sono linearmente indipendenti allora questo insieme si dice **base** di  $V$ . Notate che le due proprietà dicono, rispettivamente che se

$$B := \{v_1, \dots, v_k\}$$

è una base di  $V$  allora

- I vettori della base  $B$  sono sufficienti per descrivere tutti i vettori dello spazio  $V$  come loro combinazione lineare.
- I vettori della base  $B$  sono indipendenti e sono quindi il minimo numero possibile di vettori con la precedente proprietà.

**Esempio.**

1. I vettori  $e_1, e_2, \dots, e_p \in \mathbb{R}^p$  dell'esempio precedente sono una base, che si chiama base canonica di  $\mathbb{R}^p$ .
2. I vettori  $v_1 = 1, v_2 = x, v_3 = x^2$  sono una base di  $P_2[x]$ .
3. Lo spazio vettoriale  $P[x]$  di tutti i polinomi non ha dimensione finita

Si può dimostrare che il numero di elementi di una base è lo stesso per tutte le basi e questo numero  $k$  si chiama **dimensione** dello spazio vettoriale e scriveremo

$$\dim V = k.$$

**Teorema 1.2.** *Sia  $V$  uno spazio vettoriale a dimensione finita allora tutte le basi hanno lo stesso numero di elementi.*

**Approfondimento.** Dimostrazione del Teorema 1.2

*Dimostrazione.* Siano  $A = \{a_1, \dots, a_p\}$  e  $B = \{b_1, \dots, b_m\}$  due basi e sia  $m \geq p$  facciamo vedere che deve essere necessariamente  $p = m$ .

I vettori dell'insieme  $W_1 = \{a_1, b_1, \dots, b_m\}$  sono linearmente dipendenti poiché  $B$  è una base ed in particolare  $a_1$  può essere scritto come combinazione lineare non nulla degli altri vettori di  $W_1$

$$a_1 = \sum_{i=1}^m \beta_i b_i$$

supponiamo che sia  $\beta_1 \neq 0$  (altrimenti possiamo cambiare l'ordine con cui sono scritti i vettori) e allora anche  $b_1$  può essere scritto come combinazione lineare non nulla degli altri vettori di  $W_1$ , possiamo quindi eliminarlo da  $W_1$  (dipende dagli altri) ed ottenere una base

$$B_1 = \{a_1, b_2, \dots, b_m\}.$$

Questo procedimento può essere continuato fino ad ottenere la base

$$B_p = \{a_1, \dots, a_p, b_{p+1}, \dots, b_m\}.$$

Poiché  $A$  è già una base questo insieme può esserla anche  $B_p$  soltanto se gli elementi  $\{b_{p+1}, \dots, b_m\}$  non ci sono, ovvero, soltanto se  $p = m$ .  $\square$

Avendo a disposizione una base dello spazio vettoriale possiamo quindi identificare un vettore tramite il modo con cui è rappresentato tramite quella base. Precisamente se  $B = \{v_1, \dots, v_p\}$  è una base di  $V$  e  $v \in V$ , allora dalla prima proprietà delle basi segue che esistono  $p$  numeri reali  $\alpha_1, \alpha_2, \dots, \alpha_p$ , tali che

$$v = \sum_{i=1}^p \alpha_i v_i,$$

inoltre dalla seconda proprietà segue che questi coefficienti sono univocamente determinati. In effetti se  $v = \sum_{i=1}^p \beta_i v_i$ , allora per differenza abbiamo

$$(0 = \sum_{i=1}^p (\alpha_i - \beta_i) v_i,$$

e quindi, poiché i  $v_i$  formano una base, deve essere  $\alpha_i = \beta_i$ .

I coefficienti  $\alpha_i$  si chiamano coordinate di  $v$  rispetto alla base  $B$ . Una volta fissata la base  $B$  possiamo identificare  $v$  con il vettore di  $\mathbb{R}^p$

$$[v]_B = (\alpha_1, \alpha_2, \dots, \alpha_p).$$

Notate che da questo punto di vista i vettori di  $\mathbb{R}^3$  e i polinomi di secondo grado sono la stessa cosa. Abbiamo cioè realizzato l'obiettivo di questo paragrafo. Una volta fissata la base contano soltanto le coordinate rispetto a quella base e quindi se due spazi vettoriali hanno la stessa dimensione sono indistinguibili relativamente alle operazioni di somma e di prodotto per un numero. Questo non è vero per le altre operazioni che sono sostanzialmente diverse: per esempio il prodotto fra due polinomi omogenei di grado due (forme quadratiche) non corrisponde al prodotto delle corrispondenti matrici quadrate.

### 1.3.2 Sottospazi

Un insieme  $W \subseteq V$  si dice **sottospazio vettoriale** o semplicemente **sottospazio** di  $V$  se  $W$  non è vuoto ed è **chiuso** rispetto alle combinazioni lineari, cioè per ogni  $w_1, w_2 \in W$  e per ogni  $\alpha_1, \alpha_2 \in \mathbb{R}$  abbiamo che

$$\alpha_1 w_1 + \alpha_2 w_2 \in W.$$

In altre parole chiuso rispetto alle combinazioni lineari vuol dire che effettuando le operazioni definite nello spazio vettoriale rimaniamo nell'insieme  $W$ . A volte chiameremo i sottospazi **varietà lineari**.

**Esempio.**

1.  $P_k[x]$  è un sottospazio dello spazio vettoriale  $C^1$  delle funzioni derivabili, in quanto i polinomi sono derivabili.
2.  $P_3[x]$  è un sottospazio di  $P_5[x]$  per la definizione di grado per un polinomio.
3. L'insieme  $P_p^o[x]$  dei polinomi omogenei di grado  $p$  nella variabile  $x$  è un sottospazio di  $P_p[x]$ .
4. Gli esempi più facilmente visualizzabili di sottospazi sono i sottospazi di  $\mathbb{R}^3$ , e cioè:
  1. Le rette che passano dall'origine, fig. 1.1 pag. 26.

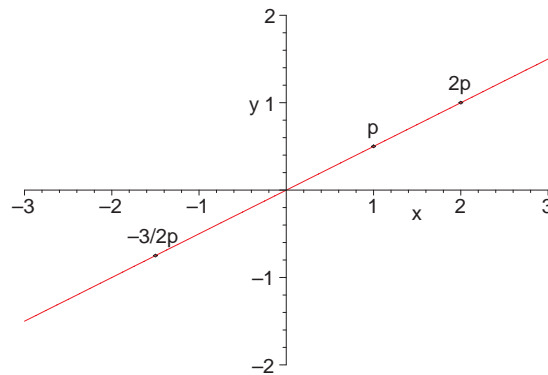


Figura 1.1: Retta per l'origine: multipli di un vettore

2. I piani che passano dall'origine, fig. 1.2 pag. 27.
3.  $\{0\}$  e  $\mathbb{R}^3$  stesso.
5. La semiretta  $A = [1, +\infty)$  è un sottoinsieme di  $\mathbb{R}$  ma non è un sottospazio poiché  $v_1 = 1, v_2 = 2 \in A$  ma la loro combinazione lineare  $v_1 - v_2 = 1 - 2 = -1 \notin A$ .

Dato un certo insieme di vettori  $v_1, \dots, v_k \in V$  è possibile costruire un sottospazio di  $V$  considerando l'insieme di tutte le loro combinazioni lineari. Questo insieme costituisce un sottospazio  $W$  detto **sottospazio generato** da  $v_1, \dots, v_k \in V$  e si scrive

$$W = \left\{ v \in V \mid v = \sum_{i=1}^k \alpha_i v_i, \alpha_i \in \mathbb{R} \right\}$$

ed è il più piccolo sottospazio che contiene  $\{v_1, \dots, v_k\}$ .

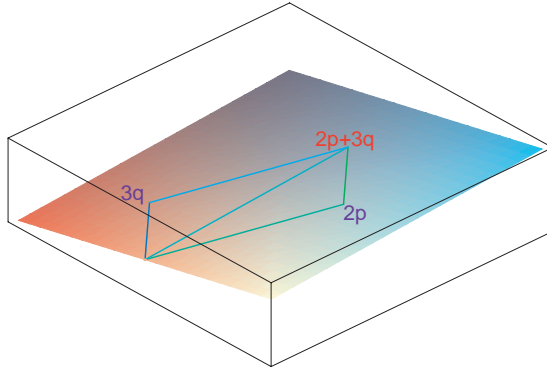


Figura 1.2: Piano per l'origine

**Esempio.**

1. In  $\mathbb{R}^3$  il sottospazio generato da un vettore non nullo è una retta passante per l'origine mentre quello generato da due vettori non nulli che non siano paralleli è un piano per l'origine.
2. Dati i tre monomi  $p_0, p_1, p_2$  con

$$p_0(x) = 1, p_1(x) = x, p_2(x) = x^2,$$

il sottospazio generato da questi tre vettori è lo spazio dei polinomi di grado minore o eguale a due nella variabile  $x \in \mathbb{R}$  ossia  $P_2[x]$ .

Sia  $X$  uno spazio lineare, un sottoinsieme  $V \subseteq X$  si chiama **sottospazio di dimensione  $k$** , se  $V$  è generato da  $k$  vettori  $v_1, v_2, \dots, v_k$  linearmente indipendenti, ovvero

$$V = \text{Span}\{v_1, v_2, \dots, v_k\}.$$

**Definizione 1.3.** Chiameremo **sottospazio affine** o **varietà affine di dimensione  $k$**  di uno spazio vettoriale  $V$  il traslato di un sottospazio di dimensione  $k$ , ovvero il sottoinsieme  $V_0 \subseteq V$  è un sottospazio affine se esistono un sottospazio  $W$  di dimensione  $k$  ed un vettore  $v_0$  tali che

$$V_0 = v_0 + W : \{v \in V \mid v = v_0 + w, w \in W\}.$$

ESERCIZIO 1.2. Dimostrare che se  $V_0$  un sottospazio affine ottenuto come il traslato di un sottospazio  $W$  allora **per ogni**  $\hat{v} \in V_0$  si ha che  $V_0 = \hat{v} + W$ .

**Definizione 1.4.** Se  $\dim X = p$  e  $V$  è una varietà affine di dimensione  $k$ , allora diremo che  $p - k$  è la **codimensione** della varietà  $V$  e scriveremo

$$\text{codim } V = p - k.$$

Le varietà affini di codimensione uno si chiamano **iperpiani**.

**Esempio.** I sottospazi affini di  $\mathbb{R}^3$  sono

1. Le rette che passano per un punto assegnato  $v_0$ .
2. I piani che passano per un punto assegnato  $v_0$ .

### Semispazi

Sia  $a \in \mathbb{R}^{p*}$  un vettore non nullo e  $b \in \mathbb{R}$ , abbiamo detto che l'insieme

$$\{x \in \mathbb{R}^p : ax = b\}$$

è un iperpiano. Tramite questo iperpiano possiamo dividere  $\mathbb{R}^p$  in due sottoinsiemi che diremo **semispazi** chiusi

$$\{x \in \mathbb{R}^p : ax \geq b\}$$

$$\{x \in \mathbb{R}^p : ax \leq b\}.$$

Il vettore  $n := \frac{a^\top}{\|a\|}$  si chiama **normale interna** al primo semispazio e **normale esterna** al secondo. Il motivo di questo nome risiede nel seguente fatto: partendo da un punto  $x_0$  sull'iperpiano muoviamoci nella direzione di  $n$ , consideriamo, cioè, il punto  $x_0 + \varepsilon n$  per  $\varepsilon > 0$ , allora questo punto è interno al primo semispazio, infatti

$$a(x_0 + \varepsilon n) = ax_0 + \varepsilon an = b + \varepsilon > b.$$

Riassumendo:

l'insieme delle soluzioni dell'equazione

$$ax = b, \quad a \in \mathbb{R}^{p*}, a \neq 0,$$

è un iperpiano, mentre l'insieme delle soluzioni della disequazione

$$ax \leq b, \quad a \in \mathbb{R}^{p*}, a \neq 0,$$

è un semispazio.

Se consideriamo più in generale un sistema di disequazioni

$$Ax \leq b$$

l'insieme delle sue soluzioni è l'intersezione di un numero finito di semispazi e si dirà un **poliedro**.

AGGIUNGERE UN ESEMPIO

## Capitolo 2

# Sistemi lineari

*Can you do Division? Divide a loaf by a knife  
... what's the answer to that? said the White  
Queen. I suppose ... Alice was beginning, but  
the Red Queen answered for her. Bread-and-  
butter, of course.*

LEWIS CARROL, *Through the Looking Glass*.

In molti problemi siamo chiamati a scegliere fra molte alternative possibili quella *migliore*, in questi modelli, spesso, non tutte le scelte sono possibili perché dobbiamo rispettare dei *vincoli*. Supponete ad esempio di essere il gestore di un fondo bilanciato di investimento e di dover decidere come investire un certo capitale  $C$ , potete acquistare due diversi tipi di obbligazioni e due diversi tipi di azioni, mantenendo fra i due investimenti un rapporto assegnato di due a uno, in altre parole dovete investire in azioni un terzo del capitale che avete a disposizione. Se  $x_1, x_2$  sono le quantità delle obbligazioni da acquistare al prezzo di  $\pi_1, \pi_2$ , rispettivamente, e  $x_3, x_4$  quelle delle azioni che hanno prezzi  $\pi_3, \pi_4$  allora dovete soddisfare le equazioni

$$\begin{aligned}\pi_1 x_1 + \pi_2 x_2 + \pi_3 x_3 + \pi_4 x_4 &= C \\ 2(x_1 + x_2) &= x_3 + x_4,\end{aligned}$$

la prima delle due equazioni dice che il valore dei titoli acquistati è eguale al capitale che dovete investire mentre la seconda dice che avete rispettato la proporzione fra i due diversi tipi di investimenti. Per sapere quali scelte sono possibili dovete essere in grado di caratterizzare in modo semplice l'insieme delle soluzioni di un sistema di questo genere che si chiama **sistema lineare** in quanto le uniche opera-

zioni che facciamo con le incognite sono quelle di moltiplicarle per un numero e di sommarle.

Obiettivo quindi di questo capitolo è lo studio dei sistemi lineari di equazioni.

Iniziamo con le notazioni, prendiamo, a titolo di esempio, il sistema lineare

$$2x_1 - 3x_2 + 2x_3 = 2 \quad (2.1)$$

$$x_1 - 3x_2 + 5x_3 = 1, \quad (2.2)$$

utilizzando le notazioni del capitolo precedente esso può essere scritto come

$$Ax = b \quad (2.3)$$

dove

$$A = \begin{pmatrix} 2 & -3 & 2 \\ 1 & -3 & 5 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

$A$  è la matrice dei coefficienti,  $x$  il vettore delle incognite e  $b$  il vettore dei termini noti. Notate come il numero delle colonne di  $A$  corrisponda al numero delle incognite mentre il numero delle righe a quello delle equazioni.

È conveniente riferirsi ad  $A$  come alla **matrice incompleta** del sistema e alla matrice  $(A|b)$  ottenuta aggiungendo alle colonne di  $A$  il vettore  $b$ ,

$$(A|b) = \left( \begin{array}{ccc|c} 2 & -3 & 2 & 2 \\ 1 & -3 & 5 & 1 \end{array} \right),$$

come alla **matrice completa** del sistema; la riga separatrice verticale serve a ricordarci il significato diverso dell'ultima colonna, nelle prime colonne ci sono i coefficienti delle variabili mentre nell'ultima i termini noti delle equazioni.

È importante notare che se indichiamo con  $w_1, w_2, w_3$  le tre colonne della matrice  $A$  il sistema lineare può essere scritto come

$$x_1 w_1 + x_2 w_2 + x_3 w_3 = b$$

è quindi trovare una soluzione del sistema lineare corrisponde al problema di trovare i coefficienti della combinazione lineare dei tre vettori  $w_1, w_2, w_3$  che genera il vettore  $b$ .

Per una generica matrice  $A$  ed un generico vettore  $b$  la matrice completa sarà

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

$(A|b) \in M_{m \times (n+1)}(\mathbb{R})$  ed essa rappresenta un sistema lineare di  $m$  equazioni in  $n$  incognite.

Le domande a cui vogliamo dare una risposta sono due

1. Il sistema lineare (2.3) ammette soluzioni ?
2. Se ne ammette, quante sono ?

Procederemo nel modo seguente: tramite degli esempi illustreremo un metodo esplicito per rispondere a queste domande e formalizzeremo dopo le risposte teoriche.

## 2.1 Risolviamo un sistema lineare

Questa sezione è dedicata all'illustrazione di una procedura iterativa (algoritmo) che permette di decidere, in un numero finito di passi, se un sistema lineare ammette soluzioni o no e nel primo caso di calcolarle, questo algoritmo prende il suo nome dal matematico Carl F. Gauss. Iniziamo con alcuni esempi di sistemi lineari nei quali è facile vedere se ci sono soluzioni o no e, nel primo caso, determinare quante sono.

**Esempio.** Sia dato il sistema

$$\begin{aligned} 2x_1 - 3x_2 + x_3 &= 0 \\ x_2 - 2x_3 &= 1 \\ -3x_3 &= -2. \end{aligned}$$

Esso ha una struttura particolarmente semplice che consente di determinare l'unica soluzione per sostituzione a partire dall'ultima equazione

$$x_3 = \frac{2}{3}$$

sostituiamo il valore di  $x_3$  nella seconda equazione

$$x_2 = 1 + 2x_3 = \frac{7}{3}$$

sostituiamo i valori di  $x_2, x_3$  nella prima equazione

$$x_1 = \frac{3x_2 - x_3}{2} = \frac{19}{6}.$$

Per evidenziare la particolare struttura di questo sistema che ci ha permesso di risolverlo facilmente scriviamolo in forma matriciale

$$\begin{pmatrix} 2 & -3 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}.$$

La caratteristica fondamentale di questo sistema è che la matrice incompleta è a scala e questo consente di risolvere il corrispondente sistema lineare per **sostituzione inversa**, si chiama sostituzione inversa perché procediamo dall'ultima equazione alla prima. Inoltre in questo esempio la soluzione che troviamo è unica e questo è *dovuto al fatto che abbiamo un pivot*  $\{a_{11} = 2, a_{22} = 1, a_{33} = -3\}$  *in corrispondenza di ciascuna variabile.*

La variabili nella cui colonna appare un pivot si chiamano **variabili di base**.

Consideriamo un sistema la cui matrice completa è la seguente

$$\left( \begin{array}{cccc|c} 1 & 2 & 3 & 5 & 6 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{array} \right).$$

Il procedimento di sostituzione a partire dall'ultima equazione dà il seguente risultato

$$\begin{aligned} x_4 &= 2 \\ x_3 &= 1 - 2x_4 = -3 \\ x_1 &= -2x_2 - 3x_3 - 5x_4 + 6 = 5 - 2x_2 \end{aligned}$$

abbiamo cioè trovato infinite soluzioni che dipendono dalla scelta di un parametro,  $x_2$ , che può assumere qualsiasi valore reale. Questo è *dovuto al fatto che il numero dei pivot*,  $a_{11} = 1, a_{23} = 1, a_{33} = 1$ , *è inferiore al numero delle incognite e che la seconda colonna, corrispondente alla variabile  $x_2$  non contiene alcun pivot.*

La variabili nella cui colonna non appare un pivot si chiamano **variabili non di base** o **variabili libere**.

Se il sistema lineare è rappresentato dalla matrice completa

$$A = \left( \begin{array}{ccc|c} 1 & -2 & 3 & 1 \\ 0 & 2 & 1 & 3 \\ 0 & 0 & 0 & 2 \end{array} \right)$$

allora il sistema non ammette soluzioni in quanto l'ultima equazione

$$0 = 2$$

è impossibile. Questo è *dovuto alla presenza di un pivot*,  $a_{34} = 2$ , *nella colonna dei termini noti.*

### 2.1.1 Un Metodo Risolutivo

Vediamo ora come possiamo ridurre un generico sistema lineare ad una forma come quella descritta nel paragrafo precedente. Supponiamo di voler risolvere il seguente sistema lineare di tre equazioni in tre incognite la cui matrice completa non è a scala

$$\begin{aligned} 2x_1 - 3x_2 + x_3 &= 0 \\ 3x_1 &\quad - 2x_3 = 1 \\ x_2 - 3x_3 &= -2 \end{aligned}$$

Per risolvere questo sistema applichiamo una procedura ricorsiva detta **riduzione di Gauss** che riduce questo sistema alla forma a scala descritta in precedenza. *La variabile  $x_1$  è presente nella prima equazione, eliminiamo  $x_1$  dalle successive equazioni sommando a queste un opportuno multiplo della prima.* Questa operazione è ammissibile perché sommiamo equazioni fra di loro.

- *Sommiamo alla seconda equazione la prima moltiplicata per  $-3/2$*

$$\begin{aligned} 2x_1 - 3x_2 + x_3 &= 0 \\ \frac{9}{2}x_2 - \frac{7}{2}x_3 &= 1 \\ x_2 - 3x_3 &= -2 \end{aligned}$$

- *La variabile  $x_1$  non compare nella terza equazione e quindi non dobbiamo fare nulla.*

A questo punto la variabile  $x_1$  compare solo nella prima equazione. Possiamo continuare in modo analogo: *La variabile  $x_2$  compare nella seconda equazione, eliminiamo  $x_2$  dalle successive equazioni sommando a queste un opportuno multiplo della seconda.*

- *Sommiamo alla terza equazione la seconda moltiplicata per  $-2/9$*

$$\begin{aligned} 2x_1 - 3x_2 + x_3 &= 0 \\ \frac{9}{2}x_2 - \frac{7}{2}x_3 &= 1 \\ -\frac{20}{9}x_3 &= -\frac{20}{9}. \end{aligned}$$

Il sistema a cui siamo pervenuti è ora in forma a scala ed è evidentemente equivalente al sistema iniziale (ha cioè lo stesso insieme di soluzioni). Possiamo ricavare

la soluzione per sostituzione inversa a partire dall'ultima equazione ottenendo

$$\begin{aligned}x_3 &= 1 \\x_2 &= \frac{2}{9} \left( \frac{7}{2}x_3 + 1 \right) = 1 \\x_1 &= \frac{1}{2}(3x_2 - x_3) = 1.\end{aligned}$$

Per evidenziare il procedimento che abbiamo seguito riscriviamolo in notazione matriciale.

Sia  $A$  la matrice dei coefficienti delle incognite e  $b$  il vettore dei termini noti

$$A = \begin{pmatrix} 2 & -3 & 1 \\ 3 & 0 & -2 \\ 0 & 1 & -3 \end{pmatrix} \quad \text{e} \quad b = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$$

Costruiamo la matrice completa  $(A|b)$

$$(A|b) = \left( \begin{array}{ccc|c} 2 & -3 & 1 & 0 \\ 3 & 0 & -2 & 1 \\ 0 & 1 & -3 & -2 \end{array} \right)$$

Iniziamo con la prima iterazione dell'algoritmo di Gauss.

Partiamo dalla prima riga, l'elemento  $a_{11} = 2$  è il coefficiente non nullo della variabile che vogliamo eliminare dalle righe che stanno sotto quella considerata (la prima) e si chiama **pivot** poiché facciamo perno su di lui per effettuare l'eliminazione

- *Sommiamo la prima riga moltiplicata per  $-3/2$  alla seconda*

$$\left( \begin{array}{ccc|c} 2 & -3 & 1 & 0 \\ 0 & \frac{9}{2} & -\frac{7}{2} & 1 \\ 0 & 1 & -3 & -2 \end{array} \right)$$

Il coefficiente  $-\frac{3}{2}$  si ottiene dividendo l'elemento che si vuole far diventare zero, 3, per il pivot, 2, e cambiando di segno. Andiamo alla riga successiva e cerchiamo il primo elemento non nullo ossia  $a_{22} = \frac{9}{2}$ ; questo è il nuovo pivot che usiamo per cancellare gli elementi della seconda colonna sotto il pivot

- *Sommiamo la seconda riga moltiplicata per  $-2/9$  alla terza*

$$\left( \begin{array}{ccc|c} 2 & -3 & 1 & 0 \\ 0 & \frac{9}{2} & -\frac{7}{2} & 1 \\ 0 & 0 & -\frac{20}{9} & -\frac{20}{9} \end{array} \right)$$

Ricordando che le colonne di  $A$  sono i coefficienti delle incognite e che l'ultima colonna è quella dei termini noti, per sostituzione si ottiene

$$x_1 = 1 \quad x_2 = 1 \quad x_3 = 1.$$

In questo esempio semplice non abbiamo incontrato alcuna difficoltà. Vediamo quali tipi di problemi si possono presentare.

**1.** Per effettuare il passo  $i$ -mo dell'algoritmo è necessario che il coefficiente della variabile  $x_i$  nella riga  $i$ -ma, quello che abbiamo chiamato pivot, sia diverso da zero; se questo non accade possiamo scambiare due righe come nel seguente esempio.

**Esempio.** Consideriamo il sistema

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 3 \\ x_3 &= 2 \\ 2x_1 + 3x_2 + 3x_3 &= 4 \end{aligned}$$

la cui matrice completa è

$$\left( \begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 2 & 3 & 3 & 4 \end{array} \right)$$

Il primo pivot è  $a_{11} = 1$ . Annulliamo gli elementi sotto il pivot

• *Sommiamo alla terza riga la prima moltiplicata per  $-2$*

$$\left( \begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 1 & -1 & -2 \end{array} \right)$$

Non possiamo prendere come pivot  $a_{22} = 0$  poiché per effettuare un'iterazione dell'algoritmo dobbiamo dividere per il pivot. Si noti che in effetti non è possibile cancellare  $x_2$  dalla terza equazione usando la seconda equazione. Ma, poiché nella terza riga il coefficiente di  $x_2$  è diverso da zero, scambiamo le due righe. Questo è una operazione permessa perché, ovviamente, la risolubilità di un sistema non dipende dall'ordine in cui sono scritte le equazioni.

• *Scambiamo la terza riga con la seconda*

$$\left( \begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 2 \end{array} \right)$$

A questo punto non è necessaria una ulteriore iterazione prendendo  $a_{22} = 1$  come pivot perché il sistema è già in forma a scala e otteniamo come soluzione

$$x_1 = -1 \quad x_2 = 0 \quad x_3 = 2.$$

2. Può avvenire che tutti gli elementi candidati ad essere il pivot ad una certa iterazione siano zero e in questo caso ci spostiamo sulla colonna successiva. Illustriamo questa possibilità con un esempio.

**Esempio.** Supponiamo che il sistema abbia la seguente matrice completa

$$\left( \begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 2 & 2 & 3 & 4 \end{array} \right)$$

Il pivot iniziale è  $a_{11} = 1$  e la prima iterazione dell'algoritmo di Gauss ci da'

$$\left( \begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & -1 & -2 \end{array} \right),$$

come vedete non è possibile trovare un pivot nella seconda colonna. Ci spostiamo sulla terza e prendiamo come pivot  $a_{23} = 1$ . Ora eliminando  $x_3$  dalla terza riga otteniamo

$$\left( \begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right).$$

L'ultima riga essendo costituita da soli zeri è un'identità verificata per ogni valore delle variabili, come tale può essere eliminata dal sistema. Il verificarsi di tale situazione sta a significare che la terza equazione del sistema iniziale poteva essere scritta come una combinazione lineare delle altre due equazioni.

La soluzione di questo esempio richiede un commento. Dalla seconda riga ricaviamo

$$x_3 = 2$$

mentre dalla prima

$$x_1 = -x_2 - 2x_3 + 3 = -x_2 - 1$$

e quindi il valore di  $x_2$  non è determinato, abbiamo cioè infinite soluzioni che dipendono da un parametro. Per sottolineare questo fatto procediamo nel modo seguente, poniamo  $x_2 = \alpha$  e scriviamo

$$\begin{aligned} x_1 &= -1 - \alpha \\ x_2 &= \alpha \\ x_3 &= -2 \end{aligned}$$

che in forma vettoriale diventa

$$x = \begin{pmatrix} -1 \\ 0 \\ -2 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

notate che il primo vettore si ottiene per  $\alpha = 0$  ed il secondo, una volta cancellati i termini costanti appena ottenuti, si ottiene ponendo  $\alpha = 1$ . Da questa scrittura possiamo affermare che l'insieme delle soluzioni è una retta nello spazio  $\mathbb{R}^3$ .

Modifichiamo leggermente quest'ultimo esempio per illustrare un'altra possibile situazione

**Esempio.** Supponiamo che il sistema abbia la seguente matrice completa

$$\left( \begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 2 & 2 & 3 & 2 \end{array} \right)$$

Il pivot iniziale è  $a_{11} = 1$  e la prima iterazione dell'algoritmo di Gauss ci da'

$$\left( \begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & -1 & -4 \end{array} \right),$$

come si vede non è possibile trovare un pivot nella seconda colonna. Ci spostiamo sulla terza e prendiamo come pivot  $a_{23} = 1$ . Ora eliminando  $x_3$  dalla terza riga otteniamo

$$\left( \begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & -2 \end{array} \right).$$

Ricordando la notazione matriciale del sistema l'ultima riga corrisponde all'equazione impossibile

$$0 = 2$$

e quindi il sistema stesso è impossibile, questo corrisponde al fatto che il vettore dei termini noti  $b$  NON può essere scritto come combinazione lineare delle colonne dei coefficienti delle incognite.

Da quanto abbiamo detto possiamo ora dare una risposta alle due domande che ci siamo posti all'inizio *bullet* Il sistema lineare (2.3) ammette soluzioni ?

La prima risposta è contenuta in un teorema dovuto a Eugène Rouché <sup>1</sup> e riformulato successivamente da Alfredo Capelli <sup>2</sup>

**Teorema 2.1** (Rouché-Capelli). *Il sistema lineare (2.3) ammette soluzione se e solo se il numero di pivot di una ridotta a scala della matrice incompleta  $A$  e di quella completa  $(A|b)$  sono uguali.*

<sup>1</sup><http://www-history.mcs.st-and.ac.uk/history/Biographies/Rouche.html>

<sup>2</sup>[http://it.wikipedia.org/wiki/Alfredo\\_Capelli](http://it.wikipedia.org/wiki/Alfredo_Capelli)

*bullet* Se ne ammette, quante sono?

**Teorema 2.2.** *Sia  $A$  una matrice con  $m$  righe ed  $n$  colonne, sia  $b \in \mathbb{R}^m$  e siano  $k_i, k_c$  il numero dei pivot, rispettivamente, di una ridotta a scala delle matrici incompleta e completa. Allora il sistema  $Ax = b$  è risolubile se e solo se  $k_c = k_i$ , e in questo caso le soluzioni dipendono da  $n - k_c$  parametri. In particolare il sistema è risolubile per ogni  $b \in \mathbb{R}^m$  se e solo se  $k_i = m$ , ( $m \leq n$ ) ed ha un'unica soluzione se e solo se  $k_c = k_i = n$ , ( $n \leq m$ ).*

### 2.1.2 L'algoritmo di Gauss

Diamo ora una descrizione formale completa dell'algoritmo di Gauss per ridurre una matrice con  $m$  righe ed  $n$  colonne in forma a scala. Nella descrizione dell'algoritmo useremo tre operazioni elementari

- $C_{pq}$  : Scambiare la colonna  $p$ -ma con la colonna  $q$ -ma.
- $R_{pq}$  : Scambiare la riga  $p$ -ma con la riga  $q$ -ma.
- $cR_p$  : Moltiplicare la riga  $p$ -ma per il numero  $c \neq 0$ .
- $cR_p + R_q$  : Aggiungere  $c$  volte la riga  $p$ -ma alla riga  $q$ -ma.

Consideriamo una matrice con  $m$  righe e  $n$  colonne.

#### Algoritmo per la riduzione a scala di una matrice.

##### Inizializzazione:

1. Porre  $k = 1$ ,  $k$  è il **contatore di riga**.
2. Porre  $h = 1$ ,  $h$  è il **contatore di colonna**.

##### Iterazione:

1. Se  $k = m$  o  $h = n$  allora **STOP**: la matrice è stata ridotta a scala.
2. Determinare la prima colonna  $j_h$  con un elemento  $a_{kj_h} \neq 0$  con  $j_h \geq h$ . Se  $j_h > h$  effettuare  $C_{hj_h}$ , andare al passo successivo. Se tutti gli elementi della riga sono zero, cercare una riga  $j_k > k$  con elementi diversi da zero ed effettuare  $R_{kj_k}$  e ripetere il passo 2. Altrimenti **STOP**: la matrice è stata ridotta a scala.
3. Effettuare  $-\frac{a_{s,k}}{a_{k,h}}R_k + R_s$ , per  $s > k$ .

4. Porre  $k = k + 1$
5. Ripetere l'iterazione

Quando l'algoritmo si arresta, dopo un numero finito di passi, abbiamo ridotto la matrice in forma a scala.

### 2.1.3 Possibili risultati finali della procedura di riduzione

La discussione sulla risolubilità o meno del sistema dipende dal numero di pivot che abbiamo trovato al termine della riduzione e della loro collocazione, in particolare abbiamo che:

1. Il sistema lineare è risolubile se e solo se non ci sono pivot nella colonna dei termini noti.
2. Quando il sistema è risolubile le soluzioni dipendono da un numero di parametri eguale al numero delle incognite meno il numero dei pivot. In particolare la soluzione è unica se e solo se il numero dei pivot è eguale a quello delle incognite

### Esempi

Consideriamo il sistema lineare

$$\begin{aligned}x_1 - 2x_2 + 3x_3 &= -2 \\x_2 - x_3 &= 2 \\x_1 - 6x_2 + 2x_3 &= 0\end{aligned}$$

la cui matrice completa è

$$A = \left( \begin{array}{ccc|c} 1 & -2 & 3 & -2 \\ 0 & 1 & -1 & 2 \\ 1 & -6 & 2 & 0 \end{array} \right)$$

Il primo elemento della prima riga è diverso da zero e quindi possiamo prendere  $a_{1,1} = 1$  come pivot ed effettuare il primo passo della riduzione di Gauss

$$A = \left( \begin{array}{ccc|c} 1 & -2 & 3 & -2 \\ 0 & 1 & -1 & 2 \\ 0 & -4 & -1 & 2 \end{array} \right)$$

Il secondo elemento della seconda riga è diverso da zero e quindi possiamo prendere  $a_{2,2} = 1$  come pivot ed effettuare il secondo passo della riduzione di Gauss

$$A = \left( \begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & -5 & 10 \end{array} \right)$$

la matrice è ora a scala e possiamo trarre le nostre conclusioni

1. Non vi sono pivot nell'ultima colonna e quindi il sistema ammette soluzione.
2. Vi sono 3 pivot nelle prime  $n - 1 = 3$  colonne è quindi il sistema ammette una ( $\infty^{3-3} = \infty^0 = 1$  sola soluzione.

Modifichiamo ora la matrice del sistema lineare prendendo

$$A = \left( \begin{array}{ccc|c} 0 & 1 & -1 & 2 \\ 1 & -2 & 3 & -2 \\ -3 & 8 & -11 & 7 \end{array} \right)$$

e procediamo come prima effettuando la riduzione di Gauss. Questa volta però poiché il primo elemento della prima riga è nullo dobbiamo effettuare uno scambio di righe

$$A = \left( \begin{array}{ccc|c} 1 & -2 & 3 & -2 \\ 0 & 1 & -1 & 2 \\ -3 & 8 & -11 & 7 \end{array} \right)$$

e dopo possiamo procedere

$$A = \left( \begin{array}{ccc|c} 1 & -2 & 3 & -2 \\ 0 & 1 & -1 & 2 \\ 0 & 2 & -2 & 1 \end{array} \right)$$

$$A = \left( \begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & -3 \end{array} \right)$$

Questa volta vediamo che è presente un pivot nell'ultima colonna e il sistema è quindi impossibile, notate che leggendo l'ultima riga come una equazione abbiamo  $0 = -3$  che è impossibile.

## 2.2 Ancora sull'algorithm di Gauss

Come abbiamo visto nei paragrafi precedenti per ottenere la forma esplicita della soluzione di un sistema lineare è necessario procedere alla sostituzione inversa.

### 2.2.1 L'algoritmo di Gauss-Jordan

Possiamo modificare leggermente l'algoritmo di Gauss in modo da ottenere direttamente l'espressione delle soluzioni. Questa versione più raffinata dell'algoritmo di Gauss va sotto il nome di algoritmo di Gauss-Jordan per il contributo dato da Camille Jordan<sup>3</sup>. Le modifiche sono di piccola entità e le illustreremo con un esempio. Consideriamo la matrice

$$A = \left( \begin{array}{ccccc|c} 4 & -4 & 2 & -2 & 0 & 4 \\ 0 & 0 & 2 & 3 & 0 & 2 \\ -3 & 1 & 1 & 2 & -2 & 0 \end{array} \right)$$

Per ottenere una espressione esplicita delle soluzioni faremo due ulteriori operazioni. La prima delle quali è la seguente *Una volta determinato il pivot, dividiamo la riga di pivot per il pivot stesso, in modo che il nuovo pivot sia eguale a uno.* Nel nostro esempio il primo pivot è  $a_{11} = 4$

- Moltiplichiamo la prima riga per  $1/4$

$$\left( \begin{array}{ccccc|c} 1 & -1 & \frac{1}{2} & \frac{-1}{2} & 0 & 1 \\ 0 & 0 & 2 & 3 & 0 & 2 \\ -3 & 1 & 1 & 2 & -2 & 0 \end{array} \right)$$

Il nuovo pivot è  $a_{11} = 1$ .

- Aggiungiamo alla terza riga la prima moltiplicata per 3

$$\left( \begin{array}{ccccc|c} 1 & -1 & \frac{1}{2} & \frac{-1}{2} & 0 & 1 \\ 0 & 0 & 2 & 3 & 0 & 2 \\ 0 & -2 & \frac{5}{2} & \frac{1}{2} & -2 & 3 \end{array} \right)$$

- Scambiamo la seconda e terza riga

$$\left( \begin{array}{ccccc|c} 1 & -1 & \frac{1}{2} & \frac{-1}{2} & 0 & 1 \\ 0 & -2 & \frac{5}{2} & \frac{1}{2} & -2 & 3 \\ 0 & 0 & 2 & 3 & 0 & 2 \end{array} \right)$$

Il nuovo pivot è  $a_{22} = -2$

- Moltiplichiamo la seconda riga per  $-1/2$

$$\left( \begin{array}{ccccc|c} 1 & -1 & \frac{1}{2} & \frac{-1}{2} & 0 & 1 \\ 0 & 1 & \frac{-5}{4} & \frac{-1}{4} & 1 & \frac{-3}{2} \\ 0 & 0 & 2 & 3 & 0 & 2 \end{array} \right)$$

Ora scatta la seconda differenza fra i due algoritmi:

<sup>3</sup><http://www-history.mcs.st-and.ac.uk/history/Biographies/Jordan.html>

utilizziamo il pivot per rendere zero *tutti* gli elementi sulla colonna di pivot, tranne il pivot stesso,

ricordate che con l'algoritmo di Gauss rendevamo zero solo gli elementi sotto il pivot.

- Aggiungiamo alla prima riga la seconda moltiplicata per 1

$$\left( \begin{array}{ccc|ccc} 1 & 0 & \frac{-3}{4} & \frac{-3}{4} & 1 & \frac{-1}{2} \\ 0 & 1 & \frac{-5}{4} & \frac{-1}{4} & 1 & \frac{-3}{2} \\ 0 & 0 & 2 & 3 & 0 & 2 \end{array} \right)$$

Il nuovo pivot è  $a_{33} = 2$

- Moltiplichiamo la terza riga per  $1/2$

$$\left( \begin{array}{ccc|ccc} 1 & 0 & \frac{-3}{4} & \frac{-3}{4} & 1 & \frac{-1}{2} \\ 0 & 1 & \frac{-5}{4} & \frac{-1}{4} & 1 & \frac{-3}{2} \\ 0 & 0 & 1 & \frac{3}{2} & 0 & 1 \end{array} \right)$$

Il pivot sulla terza riga è ora eguale ad uno, procediamo quindi all'eliminazione degli elementi sopra il pivot

- Aggiungiamo alla prima riga la terza moltiplicata per  $3/4$

$$\left( \begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{2} & 1 & \frac{1}{4} \\ 0 & 1 & \frac{-5}{4} & \frac{-1}{4} & 1 & \frac{-3}{2} \\ 0 & 0 & 1 & \frac{3}{2} & 0 & 1 \end{array} \right)$$

- Aggiungiamo alla seconda riga la terza moltiplicata per  $5/4$

$$\left( \begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{2} & 1 & \frac{1}{4} \\ 0 & 1 & 0 & \frac{13}{4} & 1 & \frac{-1}{4} \\ 0 & 0 & 1 & \frac{3}{2} & 0 & 1 \end{array} \right)$$

Consideriamo ora il sistema nelle variabili  $x_1, x_2, \dots, x_5$  di matrice completa  $A$ . Il procedimento effettuato ha evidenziato le tre variabili  $x_1, x_2, x_3$  i cui valori possiamo determinare in funzione delle altre variabili restanti  $x_4, x_5$ . In altre parole da ciascuna equazione ricavo il valore di una delle variabili senza bisogno di operare alcuna sostituzione.

$$x_1 = -\frac{3}{8}x_4 - x_5 + \frac{1}{4}, \quad x_2 = -\frac{13}{8}x_4 - x_5 - \frac{1}{4}, \quad x_3 = -\frac{3}{2}x_4 + 1$$

Le variabili che abbiamo evidenziato corrispondono alle colonne di pivot e quindi ad una base dell'immagine della funzione lineare associata e si chiamano **variabili**

**di base** mentre le altre le chiameremo **variabili non di base**. Dalla espressione della soluzione vediamo inoltre come la soluzione particolare ottenuta ponendo le variabili non di base eguali a zero abbia come valori della variabili di base esattamente i coefficienti dell'ultima colonna della matrice

$$x_1 = \frac{1}{4}, x_2 = \frac{-1}{4}, x_3 = 1, x_4 = 0, x_5 = 0.$$

Utilizzeremo queste proprietà nella soluzione di problemi di programmazione lineare.

Vediamo ora cosa accade se nella matrice iniziale scambiamo la terza colonna con la quinta, abbiamo fatto un'operazione ammessa perché la soluzione di un sistema non dipende dal nome delle variabili. Se calcoliamo la soluzione otteniamo

$$\left( \begin{array}{ccccc|c} 1 & 0 & 1 & 0 & \frac{-1}{4} & 0 \\ 0 & 1 & 1 & 0 & \frac{-13}{12} & \frac{-4}{3} \\ 0 & 0 & 0 & 1 & \frac{2}{3} & \frac{2}{3} \end{array} \right)$$

da cui ricaviamo, ricordando che abbiamo scambiato il posto di  $x_3$  e  $x_5$

$$x_1 = -x_5 + \frac{1}{4}x_3, \quad x_2 = -x_5 + \frac{13}{12}x_3 - \frac{4}{3}, \quad x_4 = -\frac{2}{3}x_3 + \frac{2}{3},$$

Se poniamo uguali a zero le variabili non di base  $x_3, x_5$  otteniamo una diversa soluzione di base

$$x_1 = 0, x_2 = \frac{-4}{3}, x_3 = 0, x_4 = \frac{2}{3}, x_5 = 0.$$

Le due soluzioni di base che abbiamo trovato rappresentano i coefficienti della combinazione lineare delle colonne che contengono il pivot che mi danno il vettore dei termini noti. Esse sono diverse perché diverse sono le colonne individuate nei due casi che rappresentano due diverse basi dell'immagine della trasformazione lineare associata,

### 2.2.2 Matrici inverse e soluzioni con diversi termini noti

Il metodo di Gauss (Gauss-Jordan) può essere utilizzato anche per la determinazione simultanea di soluzioni dello stesso sistema con diversi termini noti.

Supponiamo di volere risolvere  $k$  sistemi lineari diversi che abbiano la stessa matrice dei coefficienti

$$Ax = b_1, Ax = b_2, \dots, Ax = b_k.$$

Per studiare la risolubilità di questi sistemi procediamo nel seguente modo. Consideriamo la matrice  $(A | b_1 | b_2 | \dots | b_k)$  e utilizziamo l'algoritmo di Gauss per determinarne una ridotta a scala. Da questa possiamo determinare le soluzioni corrispondenti ai diversi termini noti. Vediamo come con un esempio

**Esempio.** Consideriamo la matrice

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

vogliamo risolvere i due sistemi

$$Az = b_i$$

per  $i = 1, 2$  dove

$$b_1 = (1, 2, 0, 1), \quad b_2 = (2, 2, 1, -1)$$

Costruiamo la matrice

$$(A | b_1 | b_2) = \left( \begin{array}{cccc|c|c} 1 & 2 & 0 & -1 & 1 & 2 \\ 2 & -1 & 3 & 0 & 2 & -2 \\ 1 & 2 & -2 & 3 & 0 & 1 \\ -2 & 1 & 2 & 0 & 1 & -1 \end{array} \right)$$

la cui ridotta a scala è

$$\left( \begin{array}{cccc|c|c} 1 & 2 & 0 & -1 & 1 & 2 \\ 0 & -5 & 3 & 2 & 0 & -6 \\ 0 & 0 & -2 & 4 & -1 & -1 \\ 0 & 0 & 0 & 10 & \frac{1}{2} & -\frac{11}{2} \end{array} \right)$$

usando come termini noti prima gli elementi della quinta colonna e poi quelli della sesta, per sostituzione, possiamo trovare le soluzioni dei due sistemi

$$x_1 = \frac{29}{100}, x_2 = \frac{19}{50}, x_3 = \frac{3}{5}, x_4 = \frac{1}{20}$$

$$x_1 = \frac{21}{100}, x_2 = \frac{31}{50}, x_3 = \frac{-3}{5}, x_4 = \frac{-11}{20}$$

Consideriamo ora il problema di trovare l'inversa di una matrice quadrata  $A$ , ovvero il problema di trovare una matrice  $A^{-1}$  tale che

$$A A^{-1} = A^{-1} A = Id.$$

Per capire come fare a costruirla notiamo che l'eguaglianza  $A A^{-1} = Id$  può essere letta nel modo seguente; indichiamo con  $v_i, i = 1, \dots, p$  le colonne di  $A^{-1}$  allora questa eguaglianza è equivalente ai sistemi

$$A v_i = e_i, i = 1, \dots, p$$

e quindi i vettori  $v_i$  sono le soluzioni di questi sistemi e possiamo procedere come abbiamo fatto in precedenza. Sia

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

una matrice quadrata e sia

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

la matrice identità. Costruiamo la matrice completa

$$\left( \begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & -1 & 2 & 0 & 1 & 0 \\ 0 & 0 & -3 & 0 & 0 & 1 \end{array} \right)$$

la cui ridotta a scala è

$$\left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 4/3 \\ 0 & 1 & 0 & 1 & 0 & 2/3 \\ 0 & 0 & 1 & 0 & 0 & -1/3 \end{array} \right)$$

e quindi le ultime tre colonne rappresentano la matrice inversa

$$M^{-1} := \begin{pmatrix} 1 & 1 & 4/3 \\ 1 & 0 & 2/3 \\ 0 & 0 & -1/3 \end{pmatrix}$$

Vediamo ora come dalla ridotta a scala di una matrice sia possibile dedurre una base per l'immagine di  $A$ . Nel illustrare il significato del Teorema ?? abbiamo messo in evidenza come l'algoritmo di Gauss effettui una riduzione per righe, mentre il nostro obiettivo è quello di effettuare una riduzione per colonne. Questo si ottiene facilmente applicando l'algoritmo di Gauss alla trasposta della matrice.

**Esempio.** Consideriamo la matrice  $A$

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

Se applichiamo l'algoritmo di Gauss alla trasposta della matrice  $A$  e prendiamo la trasposta del risultato, otteniamo

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

da cui otteniamo una base ottenuta dalle prime tre colonne.

### 2.3 Considerazioni storiche

In questo paragrafo diamo alcune curiosità a proposito degli argomenti trattati e alcune indicazioni storiche.

La prima apparizione nella storia della matematica dei sistemi di equazioni lineari risale molto indietro nel tempo

*Ci sono due campi, la cui area totale è di 1800 metri quadrati. Uno produce grano in proporzione di  $2/3$  di staio per metro quadrato mentre l'altro produce grano in proporzione di  $1/2$  staio per metro quadrato. Se il raccolto totale è di 1100 stai, quant'è l'area di ciascun campo ?*

Questo problema è rappresentabile con un sistema di due equazioni lineari in due incognite ed è descritto su una tavoletta di creta Babilonese ed è databile intorno al 300 AC. Un problema del tutto analogo è descritto nel libro *Nove Capitoli sull'Arte Matematica* scritto in Cina durante la Dinastia Han (200–100 AC). L'ignoto autore non solo scrive i coefficienti in una tavola (matrice) ma descrive anche un metodo in cui sommando opportunamente le colonne della tavola si giunge alla soluzione (algoritmo di Gauss.)

Lo studio dei sistemi lineari riprende in Europa con Girolamo Cardano<sup>4</sup> che in *Ars Magna* (1545) descrive una regola per risolvere i sistemi due per due. A

<sup>4</sup><http://www-history.mcs.st-and.ac.uk/history/Biographies/Cardan.html>

posteriori possiamo dire che era molto vicino alla definizione di determinante. Bisogna attendere la fine del seicento perché questo concetto venga precisato, esso appare in Giappone con Kowa T. Seki <sup>5</sup> e in Europa con Gottfried Leibnitz <sup>6</sup> contemporaneamente.

L'algorithmo di eliminazione che abbiamo descritto fu usato da Carl F. Gauss (Brunswick 1777, Gottingen 1855) in un suo lavoro in cui studiava l'orbita dell'asteroide Pallade. Usando osservazioni sull'orbita di Pallade fatte fra il 1803 ed il 1809, Gauss ottenne un sistema di sei equazioni in sei incognite e dette un metodo sistematico per la soluzione di questi problemi.

Più tardi Carl G. Jacobi <sup>7</sup> studiò il concetto di determinante dal punto di vista delle funzioni lineari. Il suo punto di vista è quello che noi attualmente usiamo.

---

<sup>5</sup><http://www-history.mcs.st-and.ac.uk/history/Biographies/Seki.html>

<sup>6</sup><http://www-history.mcs.st-and.ac.uk/history/Biographies/Leibniz.html>

<sup>7</sup><http://www-history.mcs.st-and.ac.uk/history/Mathematicians/Jacobi.html>



## Capitolo 3

# Complementi sugli spazi vettoriali

### 3.1 Sottospazi e sottospazi affini

Le varietà affini di codimensione uno si chiamano **iperpiani**.

Le varietà affini possono essere descritte in modi diversi che ne sottolineano caratteristiche diverse; vediamo alcune descrizioni nell'ambiente  $X = \mathbb{R}^2, \mathbb{R}^3$ .

#### Iperpiani

• Un iperpiano in  $\mathbb{R}^2$  è una varietà affine di codimensione uno e quindi, poiché lo spazio ha dimensione due, ha essa stessa dimensione uno ed è quindi interpretabile come una retta nel senso della geometria euclidea. Se vogliamo rappresentarla tramite una equazione lineare scriviamo

$$\alpha_1 x_1 + \alpha_2 x_2 = b$$

o meglio in forma matriciale poniamo  $c := (\alpha_1 \ \alpha_2)$  e scriviamo

$$c x = b$$

se interpretiamo  $c$  come un vettore riga scriveremo

$$\langle c, x \rangle = b.$$

Un primo modo di rappresentarlo è come **grafico di una funzione affine**; ricordiamo che una funzione affine è una funzione la cui espressione algebrica è un

polinomio di grado uno, cioè una funzione che può essere espressa come somma di una funzione costante ed una lineare.

Consideriamo ad esempio la funzione affine da  $\mathbb{R}$  in  $\mathbb{R}$ ,  $x_1 \mapsto 1 - 2x_1$ ; il suo grafico è la retta di equazione

$$x_2 = 1 - 2x_1.$$

Possiamo scrivere questa equazione, in modo equivalente, come

$$2x_1 + x_2 = 1;$$

se poniamo  $A := \begin{pmatrix} 2 & 1 \end{pmatrix} \in M_{1,2}(\mathbb{R})$  e  $b = 1 \in \mathbb{R}$ , allora l'ultima equazione si può scrivere come

$$Ax = b,$$

e l'iperpiano può essere visto come l'**insieme delle soluzioni di un sistema lineare**.

Un altro possibile modo di descrivere la retta è tramite la definizione stessa di varietà affine. I punti del grafico hanno coordinate del tipo

$$(x_1, 1 - 2x_1) = (0, 1) + (1, -2)x_1,$$

con  $x_1 \in \mathbb{R}$ . Se poniamo  $v_0 = (0, 1)$  e  $v_1 = (1, -2)$  allora la retta è la varietà affine

$$v_0 + \text{Span}\{v_1\}.$$

• Sia  $p = 3$ ; abbiamo osservato che un iperpiano in  $\mathbb{R}^3$  è interpretabile geometricamente come un piano nel senso della geometria euclidea.

Anche un piano può essere rappresentato come **grafico di una funzione affine**. Consideriamo ad esempio la funzione affine  $(x_1, x_2) \mapsto 3 - 2x_1 + x_2$  da  $\mathbb{R}^2$  in  $\mathbb{R}$ ; il suo grafico è un piano ed i punti  $x \in \mathbb{R}^3$  del suo grafico soddisfano l'equazione

$$x_3 = 3 - 2x_1 + x_2.$$

Possiamo scrivere questa equazione, in modo equivalente, come

$$2x_1 - x_2 + x_3 = 3.$$

Se poniamo  $A := \begin{pmatrix} 2 & -1 & 1 \end{pmatrix} \in M_{1,3}(\mathbb{R})$  e  $b = 3 \in \mathbb{R}$ , allora l'ultima equazione si può scrivere come

$$Ax = b,$$

e il piano può essere visto come l'**insieme delle soluzioni di un sistema lineare**. Un altro possibile modo di descrivere il piano è tramite la definizione stessa di varietà affine. I punti del grafico hanno coordinate

$$(x_1, x_2, 3 - 2x_1 + x_2) = (0, 0, 3) + \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -2 & 1 \end{pmatrix} (x_1, x_2).$$

Se poniamo  $v_0 = (0, 0, 3)$ ,  $v_1 = (1, 0, -2)$  e  $v_2 = (0, 1, 1)$  allora la retta è la varietà affine

$$v_0 + \text{Span}\{v_1, v_2\}.$$

• In generale un iperpiano in  $\mathbb{R}^{p+1}$  è una varietà affine di dimensione  $p$ . Vediamo le varie rappresentazioni possibili; iniziamo rappresentandola come insieme delle soluzioni di un sistema lineare

$$Ax = b,$$

con  $x \in \mathbb{R}^{p+1}$  ed  $A$  matrice di rango uno e quindi descrivibile nel modo più semplice con  $A$  costituita da una sola riga. Sia  $A := (a_1 \ a_2 \ \dots \ a_p \ a_{p+1}) \in M_{1,p+1}$  con almeno un coefficiente  $a_j$  diverso da zero e  $b = a_0$ . L'iperpiano può essere visto come l'**insieme delle soluzioni di questo sistema lineare** di un'equazione in  $p + 1$  incognite. L'ipotesi che la matrice  $A$  abbia rango 1 garantisce che la dimensione dell'insieme delle soluzioni sia esattamente  $p$ . Questo fatto era garantito negli esempi visti in precedenza dalla presenza di un coefficiente eguale ad uno.

Poiché non tutti i coefficienti di  $A$  sono nulli, prendiamone uno diverso da zero, possiamo sempre supporre che sia  $a_1$ , (se non lo fosse riordiniamo le variabili) e che sia eguale ad 1 (se non lo fosse dividiamo per  $a_1 \neq 0$ ). In questo caso l'equazione  $Ax = b$  si scrive come

$$x_1 + \sum_{j=2}^{p+1} a_j x_j = a_0$$

e possiamo risolverla rispetto a  $x_1$  ottenendo

$$x_1 = a_0 - \sum_{j=2}^{p+1} a_j x_j.$$

Quest'ultima equazione può essere vista come la rappresentazione del grafico della funzione affine  $(x_2, x_3, \dots, x_{p+1}) \mapsto a_0 - \sum_{j=2}^{p+1} a_j x_j$  da  $\mathbb{R}^p$  in  $\mathbb{R}$ .

Se infine vogliamo descrivere l'iperpiano come varietà affine, notiamo che i punti del grafico hanno coordinate

$$(x_1, \dots, x_p, x_{p+1}) = (a_0, 0, \dots, 0) - \begin{pmatrix} a_2 & \dots & a_{p+1} \\ & \mathbf{Id}_p & \end{pmatrix} \begin{pmatrix} x_2 \\ \vdots \\ x_{p+1} \end{pmatrix}.$$

Se poniamo  $v_0 = \begin{pmatrix} a_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$  e  $v_i = (a_{i+1}, e_i) \in \mathbb{R}^{p+1}$  dove  $e_1, e_2, \dots, e_p$  sono

i vettori della base canonica, abbiamo che  $v_1, v_2, \dots, v_p$  sono le colonne della matrice

$$\begin{pmatrix} a_2 & \dots & a_{p+1} \\ & \mathbf{Id}_p & \end{pmatrix}.$$

Con queste notazioni possiamo scrivere i punti del grafico come

$$\begin{pmatrix} x_1 \\ \vdots \\ x_p \\ x_{p+1} \end{pmatrix} = \begin{pmatrix} a_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} - (a_2, e_1)x_2 - \dots - (a_{p+1}, e_p)x_{p+1}$$

e quindi l'iperpiano è la varietà affine di dimensione  $p$

$$v_0 + \text{Span}\{v_1, v_2, \dots, v_p\},$$

si noti che, per costruzione stessa, i vettori  $v_1, v_2, \dots, v_p$  sono linearmente indipendenti.

• Un discorso analogo può essere fatto per qualsiasi varietà affine  $V$  di dimensione  $k$  in  $\mathbb{R}^p$ .

Sia  $A \in M_{(p-k) \times p}$  una matrice di rango massimo  $p-k$ . Allora per ogni  $b \in \mathbb{R}^{p-k}$  l'insieme delle soluzioni del sistema lineare

$$Ax = b \tag{3.1}$$

è una varietà affine  $V$  di dimensione  $k$  e codimensione  $p-k$ . Lo studente noti come ciascuna **equazione di (3.1)** rappresenti un iperpiano e quindi l'insieme delle soluzioni è l'intersezione di  $p-k$  iperpiani. Ci riferiamo alla **equazione (3.1)** come alla **rappresentazione implicita della varietà affine  $V$**

**Lo studente verifichi la correttezza questa descrizione utilizzando i Teoremi che conosce e che caratterizzano le soluzioni dei sistemi lineari.**

Supponiamo che  $A$  abbia le prime  $p - k$  colonne indipendenti (se non fosse vero possiamo riordinare le colonne), possiamo scrivere la matrice  $A$  a blocchi

$$A = [B \ C],$$

dove  $B \in M_{(p-k) \times (p-k)}$  è una matrice invertibile e  $C \in M_{(p-k) \times k}$ . Scomponiamo il vettore  $x \in \mathbb{R}^p$  isolando le sue prime  $p - k$  componenti  $x_B \in \mathbb{R}^{p-k}$  e le sue ultime  $k$  componenti  $x_C \in \mathbb{R}^k$ . Il **sistema lineare (3.1)** può essere scritto come

$$[B \ C](x_B, x_C) = b,$$

ossia come l'equazione in  $\mathbb{R}^{p-k}$

$$B x_b + C x_c = b,$$

se moltiplichiamo a sinistra per  $B^{-1}$  otteniamo

$$[Id \ B^{-1}C](x_B, x_C) = B^{-1}b, \quad (3.2)$$

ovvero, risolvendo rispetto a  $x_B$ , otteniamo

$$x_B = B^{-1}b - B^{-1}C x_C, \quad (3.3)$$

e possiamo rappresentare la varietà affine  $V$  come il grafico della funzione affine da  $\mathbb{R}^k$  in  $\mathbb{R}^{p-k}$

$$x_C \mapsto B^{-1}b - B^{-1}C x_C.$$

Questa si chiama **rappresentazione esplicita della varietà affine  $V$** . Lo studente noti che la **rappresentazione (3.2)** del sistema lineare è quella ottenuta con l'algoritmo di Gauss-Jordan sul sistema  $Ax = b$ , nell'ipotesi che le prime  $p - k$  colonne siano indipendenti. La scelta delle colonne indipendenti non è unica e può essere fatta in molti modi diversi; le  $p - k$  variabili  $x_B$  rispetto alle quali il sistema viene risolto si chiamano **variabili di base** perché corrispondono ad una scelta di una base per  $\text{Im } A$ . Le rimanenti  $k$  variabili  $x_C$  si chiamano **variabili non di base o libere** e giocano il ruolo di parametri. Dalla **espressione (3.3)** possiamo facilmente trovare una rappresentazione del varietà affine  $V$ . Infatti  $x = \begin{pmatrix} x_B \\ x_C \end{pmatrix} \in V$  se e solo se

$$\begin{pmatrix} x_B \\ x_C \end{pmatrix} = \begin{pmatrix} B^{-1}b - B^{-1}C x_C \\ x_C \end{pmatrix},$$

cioè

$$\begin{pmatrix} x_B \\ x_C \end{pmatrix} = \begin{pmatrix} B^{-1}b \\ 0_k \end{pmatrix} - \begin{pmatrix} B^{-1}C \\ Id_k \end{pmatrix} x_C$$

e quindi posto  $v_0 = \begin{pmatrix} B^{-1}b \\ 0_k \end{pmatrix} \in \mathbb{R}^p$  e dette  $v_1, v_2, \dots, v_k$  le colonne della matrice  $\begin{pmatrix} B^{-1}C \\ Id_k \end{pmatrix}$  si ha che

$$V = v_0 + \text{Span}\{v_1, v_2, \dots, v_k\};$$

i  $k$  vettori  $v_i$  sono indipendenti per costruzione. La scrittura

$$\begin{pmatrix} x_B \\ x_C \end{pmatrix} = \begin{pmatrix} B^{-1}b \\ 0_k \end{pmatrix} - v_1 x_{p-k+1} - \dots - v_k x_p,$$

si dice **rappresentazione parametrica** di  $V$ . I parametri sono le  $k$  componenti  $x_{p-k+1}, \dots, x_p$  di  $x_C$ .

**Semispazi.** Sia  $a \in \mathbb{R}^{p*}$  un vettore non nullo e  $b \in \mathbb{R}$ , abbiamo detto che l'insieme

$$\{x \in \mathbb{R}^p : ax = b\}$$

è un iperpiano. Tramite questo iperpiano possiamo dividere  $\mathbb{R}^p$  in due sottoinsiemi che diremo **semispazi** chiusi

$$\{x \in \mathbb{R}^p : ax \geq b\}$$

$$\{x \in \mathbb{R}^p : ax \leq b\}.$$

Il vettore  $n := \frac{a^\top}{\|a\|}$  si chiama **normale interna** al primo semispazio e **normale esterna** al secondo. Il motivo di questo nome risiede nel seguente fatto: partendo da un punto  $x_0$  sull'iperpiano muoviamoci nella direzione di  $n$ , consideriamo, cioè, il punto  $x_0 + \varepsilon n$  per  $\varepsilon > 0$ , allora questo punto è interno al primo semispazio, infatti

$$a(x_0 + \varepsilon n) = ax_0 + \varepsilon an = b + \varepsilon > b.$$

Riassumendo:

l'insieme delle soluzioni dell'equazione

$$ax = b, \quad a \in \mathbb{R}^{p*}, a \neq 0,$$

è un iperpiano, mentre l'insieme delle soluzioni della disequazione

$$ax \leq b, \quad a \in \mathbb{R}^{p*}, a \neq 0,$$

è un semispazio.

Se consideriamo più in generale un sistema di disequazioni

$$Ax \leq b$$

l'insieme delle sue soluzioni è l'intersezione di un numero finito di semispazi e si dirà un **poliedro**.

AGGIUNGERE UN ESEMPIO

## 3.2 Relazioni con la geometria Euclidea

In questo paragrafo interpretiamo alcuni dei concetti definiti in  $\mathbb{R}^p$  in termini della usuale geometria euclidea in modo da permettere allo studente di utilizzare la sua intuizione geometrica del piano e dello spazio per visualizzare gli oggetti di cui stiamo parlando.

Consideriamo lo spazio vettoriale  $\mathbb{R}^p$  dotato di un prodotto scalare rispetto al quale il riferimento  $\Sigma = (0 = (0, \dots, 0), \hat{v})$  sia un riferimento ortogonale.

Ciò significa che  $\hat{v} = (v_1, \dots, v_p)$  è una base di  $\mathbb{R}^p$  tale che  $\langle v_i, v_j \rangle = \delta_{i,j}$ , dove  $\delta_{i,j}$  si chiama delta di Kronecker Leopold Kronecker ed è definito come

$$\delta_{i,j} := \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j. \end{cases}$$

In altri termini  $\hat{v}$  è una base ortonormale, cioè i vettori della base hanno norma uno e sono a due a due ortogonali. In questa situazione se

$$\begin{aligned} A &= a_1 v_1 + \dots + a_p v_p \\ B &= b_1 v_1 + \dots + b_p v_p \end{aligned}$$

sono due vettori di  $\mathbb{R}^p$ , allora il prodotto scalare fra  $A$  e  $B$  si esprime come

$$\langle A, B \rangle = \sum_{i=1}^p a_i b_i.$$

In particolare  $\langle A, v_i \rangle = a_i$  cosicché ogni  $A \in \mathbb{R}^p$  può scriversi

$$A = \sum_{i=1}^p \langle A, v_i \rangle v_i. \quad (3.4)$$

Chiamiamo **asse  $i$ -mo di  $\Sigma$**  l'insieme dei punti  $A \in \mathbb{R}^p$  del tipo  $A = \lambda v_i$ ,  $\lambda \in \mathbb{R}$ .

Scopo di questo paragrafo è mostrare come lo spazio  $\mathbb{R}^p$  così equipaggiato, costituisca un modello naturale per lo spazio  $p$ -dimensionale della geometria euclidea (g.e.).

Per comprendere il senso di tale affermazione, rappresentiamo:

- gli assi di  $\Sigma$  come rette due a due ortogonali, nel senso della g.e., che si intersecano in un punto  $O$  e su cui siano fissati i punti (unità)  $V_i$ ;
- il vettore  $A = \sum a_i v_i \in \mathbb{R}^p$ , come il punto (che indicheremo ancora con  $A$ ) di coordinate (in senso g.e.) date da  $(a_1, \dots, a_p)$ ; ciò significa che  $a_i$  è la

misura algebrica, rispetto all'unità  $OV_i$ , della proiezione ortogonale di  $OA$  sull'asse  $i$ -mo, nel senso della geometria elementare. E' conveniente osservare che tale costruzione evidenzia in modo naturale il **segmento orientato**  $OA$ , pensato cioè come percorso da  $O$  ad  $A$  e rappresentato ponendo una freccia su  $A$ .

Vediamo ora come si interpretano geometricamente alcune delle definizioni che sono state date in  $\mathbb{R}^p$  :

### 1. La norma

Essendo  $\hat{v}$  ortonormale, preso  $A = \sum_{i=1}^p a_i v_i$ , per definizione di norma in uno spazio metrico, si ha

$$\|A\| = \sqrt{\langle A, A \rangle} = \sqrt{\sum_{i=1}^p a_i^2}$$

e questa, per il Teorema di Pitagora, è la misura del segmento  $OA$ .

### 2. La somma di vettori

Se  $A = \sum_{i=1}^p a_i v_i$  e  $B = \sum_{i=1}^p b_i v_i$  allora

$$A + B = \sum_{i=1}^p (a_i + b_i) v_i$$

è il quarto vertice del parallelogramma di lati  $OA$ ,  $OB$  come provano semplici considerazioni di geometria elementare. (Lo studente provi a svolgere la dimostrazione.)

### 3. La disuguaglianza triangolare. La disuguaglianza

$$\|A + B\| \leq \|A\| + \|B\|$$

non è altro che una riedizione della nota affermazione della geometria elementare secondo la quale in un triangolo la lunghezza di un lato è minore o uguale della somma delle lunghezze degli altri due: basta infatti considerare il triangolo di vertici  $O$ ,  $A$ ,  $A + B$ .

### 4. Moltiplicazione per uno scalare.

Se  $A = \sum_{i=1}^p a_i v_i$  e  $\lambda \in \mathbb{R}$ , allora

$$\lambda A = \sum_{i=1}^p \lambda a_i v_i$$

è un punto sulla retta per  $O$  e per  $A$ , situato sulla semiretta  $OA$  se e solo se  $\lambda \geq 0$ . In particolare  $-A$  è il simmetrico di  $A$  rispetto ad  $O$ .

### 5. La differenza fra vettori.

L'interpretazione geometrica di tale operazione è immediata considerando che  $A - B = A + (-B)$  e tenendo conto del punto **2** e del punto **4**. Notiamo esplicitamente che i segmenti orientati  $O(A - B)$  e  $BA$  sono paralleli, di uguale lunghezza e identico orientamento.

### 6. La distanza fra vettori.

Essendo

$$d(A, B) = \|A - B\|,$$

la distanza fra due vettori  $A$  e  $B$  corrisponde alla lunghezza, nel senso della g.e., del segmento  $AB$ . Infatti, per il punto **5**, tale lunghezza coincide con quella del segmento  $O(A - B)$  e quest'ultima, per il punto **1**, è proprio  $\|A - B\|$ .

### 7. Il prodotto scalare fra vettori.

Per  $p = 2, 3$

$$\langle A, B \rangle = \|A\| \|B\| \cos \phi \quad (3.5)$$

dove  $\phi$  è l'angolo, nel senso della g.e., fra  $OA$  e  $OB$ .

**Approfondimento.** Dimostrazione della relazione 3.5 Dimostriamo tale affermazione limitandoci al caso  $p = 2$ . Osserviamo che, se  $A = \sum_{j=1}^2 a_j v_j$ ,  $\langle A, v_i \rangle = a_i$  significa che  $\langle A, v_i \rangle$  è la proiezione del segmento  $OA$  sull'asse  $i$ -mo. Pertanto, per la definizione geometrica di coseno, si ha  $\langle A, v_i \rangle = \|A\| \cos \phi_i$  e  $\langle B, v_i \rangle = \|B\| \cos \psi_i$  dove  $\phi_i, \psi_i$  sono rispettivamente gli angoli fra  $OA, OB$  e l'asse  $i$ -mo. Si noti che  $\phi_2 = \pi/2 - \phi_1$  e anche  $\psi_2 = \pi/2 - \psi_1$ .

Se ora  $B = b_1 v_1 + b_2 v_2$ , per la bilinearità del prodotto scalare, si avrà

$$\begin{aligned} \langle A, B \rangle &= \langle A, b_1 v_1 + b_2 v_2 \rangle = b_1 \langle A, v_1 \rangle + b_2 \langle A, v_2 \rangle = \\ &= b_1 \|A\| \cos \phi_1 + b_2 \|A\| \cos(\pi/2 - \phi_1) = \\ &= \|A\| (\langle B, v_1 \rangle \cos \phi_1 + \langle B, v_2 \rangle \sin \phi_1) = \\ &= \|A\| (\|B\| \cos \psi_1 \cos \phi_1 + \|B\| \sin \psi_1 \sin \phi_1) = \\ &= \|A\| \|B\| \cos(\psi_1 - \phi_1) = \\ &= \|A\| \|B\| \cos \phi. \end{aligned}$$

Notiamo esplicitamente che *la proiezione*  $\|A\| \cos \phi$  *di*  $OA$  *su*  $OB$  *si può esprimere tramite*

$$\frac{\langle A, B \rangle}{\|B\|},$$

cioè come prodotto scalare fra  $A$  e il versore  $\frac{B}{\|B\|}$  su  $B$ ; è poi chiaro che *il punto  $H$  proiezione di  $OA$  su  $OB$  è dato da*

$$H = \langle A, \frac{B}{\|B\|} \rangle \frac{B}{\|B\|} = \frac{\langle A, B \rangle}{\|B\|^2} B;$$

pertanto la relazione (3.4) esprime  $A$  come somma delle proiezioni del segmento  $OA$  sugli assi di riferimento. Notiamo anche che il punto  $H$  proiezione di  $PQ$  su  $PS$ , si ottiene facilmente riportando tutto in  $O$ , ossia

$$H - P = \frac{\langle Q - P, S - P \rangle}{\|S - P\|^2} (S - P)$$

e dunque

$$H = P + \frac{\langle Q - P, S - P \rangle}{\|S - P\|^2} (S - P).$$

**Esempio.** Siano  $A = (1, 3)$ ,  $B = (-1, -1)$ ; la proiezione di  $OA$  su  $OB$  è data da  $-2\sqrt{2}$ . Il significato geometrico del valore negativo di tale proiezione è facilmente interpretabile: l'angolo  $A\hat{O}B$  è ottuso. Il punto proiezione  $H$  è individuato da  $H = -2\sqrt{2} \frac{B}{\|B\|} = (2, 2)$ .

### 8. L'angolo fra due vettori.

Per

$$p = 2, 3, \hat{A}B = \arccos \frac{\langle A, B \rangle}{\|A\| \|B\|}$$

corrisponde, a causa del punto 7, all'angolo (g.e.) fra i segmenti  $OA$  ed  $OB$ . In particolare i segmenti  $OA$  ed  $OB$  sono ortogonali (g.e.) se e solo se  $\langle A, B \rangle = 0$  cioè se e solo se sono i vettori  $A$  e  $B$  sono ortogonali. Si noti che

$$|\langle A, B \rangle| = \|A\| \|B\|$$

se e solo se  $\phi = 0, \pi$  cioè se e solo se  $OA$  ed  $OB$  sono paralleli.

### 9. Le rette in $\mathbb{R}^p$ .

Ricordiamo che una retta nello spazio vettoriale  $\mathbb{R}^p$  è stata definita come un insieme del tipo

$$r = \{X \in \mathbb{R}^p : X = X_0 + tL, t \in \mathbb{R}\}$$

dove  $X_0$  è un punto fissato e  $L \in \mathbb{R}^p - \{0\}$  una **direzione**.

Abbiamo già notato, nel punto 4, che  $\{X = tL\}$  descrive una retta per  $O$ ; se sommiamo  $X_0$  ai suoi punti, usando la regola del parallelogramma di cui si è parlato nel punto 2, descriviamo una retta (g.e.) per  $X_0$  parallela alla  $\{X = tL\}$ . Sono ora utili alcune osservazioni:

- a. Per descrivere una stessa retta  $r$  si possono usare più vettori  $L$  perché  $\{X = X_0 + t(\lambda L), t \in \mathbb{R}\} = \{X = X_0 + tL, t \in \mathbb{R}\}$ , qualunque sia  $\lambda \in \mathbb{R} - \{0\}$ .
- b. Se  $r$  è la retta per due punti  $A \neq B$ , una direzione per  $r$  è  $A - B$ . Infatti la retta  $X = B + t(A - B)$  passa per  $B$  ( $t = 0$ ) e per  $A$  ( $t = 1$ ) e da come è scritta segue che ha direzione  $A - B$ . Notiamo che  $r$  si scrive anche  $X = tA + (1-t)B$  e che per  $t \in [0, 1]$  si descrivono i punti del segmento  $AB$ .
- c. In  $\mathbb{R}^2$ , posto  $X = (x, y)$ ,  $X_0 = (x_0, y_0)$ ,  $L = (l_1, l_2)$ , la retta  $r$  è descritta dalle equazioni (parametriche)

$$\begin{cases} x = x_0 + t l_1 \\ y = y_0 + t l_2 \end{cases}$$

Eliminando il parametro  $t$  da tali equazioni, si ottiene una equazione (cartesiana) del tipo

$$ax + by + c = 0.$$

Soffermiamoci sul significato geometrico di tali coefficienti. Poiché la giacitura di  $r$  ha equazione  $ax + by = 0$  e contiene  $L$ , risulta  $al_1 + bl_2 = 0$  ossia, posto  $N = (a, b)$ , si ha  $\langle N, L \rangle = 0$ . Ciò significa che  $(a, b)$  è una direzione ortogonale alla retta  $ax + by + c = 0$ . Il parametro  $c$  è invece legato alla distanza di  $r$  da  $O$  in quanto

$$d(O, r) = \frac{|c|}{\sqrt{a^2 + b^2}} = \frac{|c|}{\|N\|}.$$

- d. La distanza di un punto  $P$  da una retta  $r$  può esprimersi anche fissando un punto  $Q \in r$  qualunque e considerando il valore assoluto della proiezione di  $PQ$  su  $ON$ , ossia

$$d(P, r) = \frac{|\langle Q - P, N \rangle|}{\|N\|}.$$

Il punto  $H$  di minima distanza di  $P$  da  $r$  può essere visto come il punto proiezione di  $PQ$  su  $P(P + N)$  e quindi è dato da

$$H = \frac{\langle Q - P, N \rangle}{\|N\|^2} N + P.$$

### 10. Le direzioni in $\mathbb{R}^p$ .

Dato un vettore non nullo  $v \in \mathbb{R}^p$  allora i vettori  $2v, \pi v$  o, in generale  $\alpha v$  con  $\alpha > 0$  hanno tutti la stessa direzione di  $v$ . A volte vi è la necessità di sceglierne uno come rappresentante, noi sceglieremo quello di lunghezza uno. Chiamiamo **versore** o

**direzione unitaria** in  $\mathbb{R}^p$  un vettore di  $\mathbb{R}^p$  che abbia norma uno. L'insieme di tutti i versori si indica con

$$\mathbf{S}^{p-1} := \{ \mathbf{d} \in \mathbf{R}^p : \|\mathbf{d}\| = 1 \},$$

questo insieme è visualizzabile come una sfera  $(p-1)$ -dimensionale. In particolare

**Esempio.** Abbiamo che:

1.  $S^0 = \{-1, 1\}$ , ossia in  $\mathbb{R}$  ci sono solo due versori, che individuano la direzione positiva e quella negativa, che sono i due vettori  $P = 1$  e  $Q = -1$ , fig. 3.1 pag. 60.



Figura 3.1: Sfera unitaria in dimensione 1

2.  $S^1$  è la circonferenza di centro l'origine e raggio uno, fig. 3.2 pag. 60.

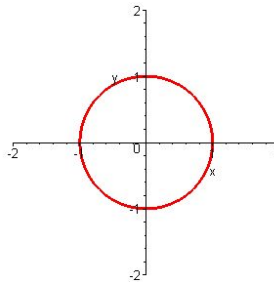


Figura 3.2: Sfera unitaria in dimensione 2

3.  $S^2$  è la superficie della sfera di centro l'origine e raggio uno, fig. 3.3 pag. 61.

In questi ultimi due casi le direzioni sono infinite e ciascuna corrisponde ad un punto della circonferenza (superficie sferica).



Figura 3.3: Sfera unitaria in dimensione 3

Dato un vettore non nullo  $P \in \mathbb{R}^p$  il calcolo del corrispondente versore è molto semplice, lo troviamo come

$$d = \frac{P}{\|P\|}.$$

Ad esempio se  $P = (1, 2)$  abbiamo che  $d = (\sqrt{5}/5, 2\sqrt{5}/5)$ , perché

$$\|P\| = \sqrt{1+4} = \sqrt{5}.$$

### 11. I piani in $\mathbb{R}^p$ .

Ricordiamo che un piano nello spazio vettoriale  $\mathbb{R}^p$  è stato definito come un insieme del tipo

$$\pi = \{X \in \mathbb{R}^p : X = X_0 + t_1 L_1 + t_2 L_2, t_1, t_2 \in \mathbb{R}\}$$

dove  $X_0$  è un punto fissato e  $L_1, L_2 \in \mathbb{R}^p$  sono vettori indipendenti. Poiché  $\{t_1 L_1\}, \{t_2 L_2\}$  sono due distinte rette per  $O$ , la somma di due punti su esse, fatta con la regola del parallelogramma, determina un piano per  $O$  mentre la ulteriore somma con  $X_0$  porta ad un piano per  $X_0$  la cui giacitura è generata da  $L_1, L_2$ .

Poiché  $\text{Rank}\{L_1, L_2\} = 2$ , la condizione di compatibilità per le  $p$  equazioni scalari in cui si decompone l'equazione vettoriale parametrica  $X = X_0 + t_1 L_1 + t_2 L_2$ , conduce a  $p - 2$  equazioni lineari (cartesiane) nelle  $n$  coordinate  $x_1, \dots, x_p$ . Anche in questo contesto può essere utile focalizzare l'attenzione su alcuni fatti:

a. Se  $p = 3$ , un piano  $\pi$  è rappresentato da un'unica equazione cartesiana del tipo

$$ax + by + cz + d = 0.$$

Poiché  $L_1, L_2$  generano la giacitura di  $\pi$ , analogamente a quanto visto per le rette, si ha

$$\langle N, L_1 \rangle = \langle N, L_2 \rangle = 0$$

dove  $N = (a, b, c)$ , ossia  $(a, b, c)$  è una direzione ortogonale al piano  $ax + by + cz + d = 0$ .

b. In  $\mathbb{R}^3$ , le equazioni scalari parametriche di una retta sono del tipo

$$\begin{cases} x = x_0 + t \ell_1 \\ y = y_0 + t \ell_2 \\ z = z_0 + t \ell_3 \end{cases}$$

Eliminando  $t$  si ottengono due equazioni di primo grado in  $x, y, z$  che individuano  $r$  come intersezione di due piani.

c. Si ottiene facilmente un'espressione della distanza di un punto  $P$  da un piano  $\pi$  del tutto analoga a quella ottenuta per la retta. Precisamente risulta

$$d(P, \pi) = \frac{|\langle Q - P, N \rangle|}{\|N\|}$$

dove  $Q$  è un punto di  $\pi$  fissato.

Illustriamo le considerazioni precedenti con due esempi:

**Esempio.**

1. Determinare  $d(P, \pi)$  se  $P = (1, 0, 1)$  e  $\pi : x - 2y + z = 4$ . Si fissa  $Q \in \pi$ , ad esempio  $Q = (0, 0, 4)$  e si calcola

$$d(P, \pi) = \frac{|\langle (1, 0, -3), (1, -2, 1) \rangle|}{\sqrt{6}} = \frac{\sqrt{6}}{3}.$$

2. Trovare l'equazione del piano per  $Q = (0, 1, 1)$  ortogonale a  $N = (1, -1, 1)$ . Tale piano è costituito dagli  $X = (x, y, z) \in \mathbb{R}^3$  tali che

$$\langle (x, y, z) - (0, 1, 1), (1, -1, 1) \rangle = 0,$$

cioè  $x - y + z = 0$ .

**12. L'identità del parallelogramma.**

L'identità

$$\|U + V\|^2 + \|U - V\|^2 = 2 (\|U\|^2 + \|V\|^2)$$

ha una chiara interpretazione geometrica in  $\mathbb{R}^2$  o  $\mathbb{R}^3$ , in quanto  $\|U\|$ ,  $\|V\|$  sono le lunghezze dei lati del parallelogramma di vertici  $U, V, O$  mentre  $\|U+V\|$ ,  $\|U-V\|$  sono le lunghezze delle sue diagonali.

**13. Il determinante di una matrice.**

Proviamo il seguente fatto: *Il determinante di una matrice  $\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$  è uguale, a meno del segno, all'area del parallelogramma di lati  $OA$  e  $OB$  dove  $A = (a_1, a_2)$ ,  $B = (b_1, b_2)$ .* Per ragioni di geometria elementare, tale area è data da

$$\mathcal{A} = \|A\| \|B\| \sin \theta.$$

D'altra parte  $|\sin \theta| = \sqrt{1 - \cos^2 \theta}$  e  $\cos \theta = \frac{\langle A, B \rangle}{\|A\| \|B\|}$  per cui

$$\begin{aligned} \mathcal{A} &= \|A\| \|B\| \sqrt{1 - \frac{\langle A, B \rangle^2}{\|A\|^2 \|B\|^2}} = \\ &= \sqrt{(a_1^2 + a_2^2)(b_1^2 + b_2^2) - (a_1 b_1 + a_2 b_2)^2} = \\ &= \sqrt{(a_1 b_2 - a_2 b_1)^2} = \left| \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \right|. \end{aligned}$$



## **Parte II**

# **Modelli lineari**



## Capitolo 4

# Modelli matematici

*But I don't want to go among mad people, Alice remarked. Oh, you can't help that, said the Cat: we're all mad here. I'm mad. You're mad. How do you know I'm mad? said Alice. You must be, said the Cat, or you wouldn't have come here.*

LEWIS CARROL, *Alice in Wonderland*.

In questo capitolo utilizzeremo la terminologia, i concetti, i teoremi, gli algoritmi che abbiamo imparato nella parte precedente per descrivere e analizzare alcuni modelli matematici della realtà economica.

### 4.1 Vettori e matrici

Le notazioni introdotte ci permettono di descrivere in modo astratto e significativo alcuni modelli di realtà economiche

#### 4.1.1 Teoria del consumatore

Consideriamo un'economia in cui siano presenti  $n$  beni distinti, l'insieme dei beni detenuti da un consumatore viene, di solito, chiamato il **paniere di beni** posseduto da quel consumatore. Una volta stabilito l'ordine con cui i singoli beni sono considerati e l'unità di misura con cui misurarli, il paniere è rappresentabile da una lista ordinata di  $n$  numeri, ( $n$ -pla), ciascuno dei quali rappresenta la quantità presente

nel paniere del bene che occupa la posizione corrispondente. Il paniere  $q$  potrà quindi essere indicato con la notazione vettoriale

$$q = (q_1, q_2, \dots, q_p),$$

dove  $q_i$  è la quantità del bene  $i$ -mo presente nel paniere, lo *stato* del sistema è il vettore  $q$  e lo *spazio degli stati* è lo spazio  $\mathbb{R}^p$ . Notiamo che non tutti i vettori di  $\mathbb{R}^p$  rappresentano degli stati possibili, questo è vero, in questo modello, solo per i vettori con tutte le componenti non negative. A partire da due panieri, contenenti gli stessi tipi di beni detenuti da due soggetti distinti, (il paniere  $q$  ed il paniere  $s$ ) si può costruire il paniere complessivo ottenuto *unendo le forze* dei due soggetti. Questo è l'operazione che facciamo quando consideriamo il consumo globale di due coniugi sommando i consumi individuali degli stessi beni. Ciò a cui si perviene è un nuovo vettore  $z$ , somma dei due vettori, le cui componenti sono la somma delle componenti corrispondenti di  $q$  e di  $s$ . Ed ancora, a partire da un assegnato paniere  $q$ , si può considerare il paniere ottenibile unendo  $m$  panieri tutti uguali a  $q$ . Questa volta il nuovo superpaniere  $w$  è quello nel quale le quantità dei singoli beni, rappresentate dalle componenti di  $q$ , sono state tutte moltiplicate per  $m$ . Il numero  $m$  per cui moltiplichiamo non è necessariamente un numero intero; se pensate, ad esempio, di dovere versare allo stato una percentuale  $\alpha$  della disponibilità di ciascun bene, il nuovo paniere in vostro possesso diviene  $(1 - \alpha)q$ , le operazioni fra vettori corrispondono quindi a delle operazioni reali fra panieri detenuti da consumatori.

Supponiamo di voler valutare il valore di mercato di un certo paniere di beni  $q$ . Per fare questo dobbiamo conoscere il prezzo unitario  $\pi_i$  a cui è commerciabile sul mercato il bene  $i$ -mo del paniere. Scriviamo il vettore  $\pi$  dei **prezzi** come vettore riga

$$\pi = (\pi_1 \quad \pi_2 \quad \dots \quad \pi_p),$$

ed otteniamo che il **valore del paniere** è

$$\text{valore}(q) = \pi_1 q_1 + \pi_2 q_2 + \dots + \pi_p q_p = \pi q,$$

cioè il valore di un paniere è il prodotto righe per colonne del vettore riga (prezzi) per il vettore colonna (paniere).

I altre parole possiamo dire che, fissato il vettore dei prezzi  $\pi = (\pi_1, \pi_2, \dots, \pi_p)$  degli  $p$  beni, possiamo definire una funzione *valore* :  $\mathbb{R}^p \rightarrow \mathbb{R}$  che associa ad ogni paniere  $q \in \mathbb{R}^p$  il suo valore come

$$\text{valore}(q) = \pi q.$$

*Il vettore dei prezzi rappresenta quindi una funzione che associa ad ogni paniere il suo valore.*

### Teoria del consumatore

Riprendiamo ora il modello di teoria del consumatore visto nel Paragrafo 4.1.1; se indichiamo con  $e_i$  il paniere elementare che contiene solo una quantità unitaria del bene  $i$ -mo, allora ogni paniere  $q^T = (q_1, q_2, \dots, q_p)$  può essere espresso in modo unico come combinazione lineare di panieri elementari che quindi formano una base,

$$q = \sum_{i=1}^p q_i e_i.$$

In pratica questo vuol dire che dobbiamo fissare l'unità di misura delle quantità di ciascun bene. Anche per i prezzi dobbiamo fissare la valuta e l'unità di misura in cui sono espressi.

Abbiamo parlato di ordinamento parziale in  $\mathbb{R}^p$ ,  $p > 1$  nel Paragrafo 1.1.2) usando gli strumenti dell'algebra lineare è possibile definirne altri

#### 4.1.2 Ordinamenti parziali\*

In uno spazio  $\mathbb{R}^p$  è possibile introdurre degli **ordinamenti parziali** basati sul concetto di cono,

**Definizione 4.1.** *Un sottoinsieme  $C$  non vuoto di uno spazio vettoriale  $V$  si chiama cono se*

$$x \in C, \quad \lambda \geq 0 \quad \implies \quad \lambda x \in C$$

*ed il cono  $C$  si chiama acuto se*

$$x \in C, \quad -x \in C \quad \implies \quad x = 0.$$

Possiamo ora dire quando un vettore è più piccolo di un altro questo non è il modo più generale di definire un ordinamento parziale, ma in questo contesto siamo interessati ad ordinamenti compatibili con la struttura di spazio vettoriale nel senso illustrato dalla seguente.

Sia  $C$  un cono convesso in  $V$ , dati due elementi  $x, y \in V$  diremo che

$$x \leq y, \quad \text{se e solo se} \quad y - x \in C,$$

un ordinamento parziale come quello definito soddisfa le seguenti proprietà per ogni  $x, y, z, w \in V$  abbiamo

a.  $x \leq x$  riflessività

b.  $x \leq y, \quad y \leq z \implies x \leq z$  transitività

- c.  $x \leq y, z \leq w \implies x + z \leq y + w$  compatibilità con l'addizione
- d.  $x \leq y, \alpha \in \mathbb{R}, \alpha \geq 0 \implies \alpha x \leq \alpha y$  compatibilità con la moltiplicazione per un numero
- e. se il cono è acuto allora l'ordinamento parziale è **antisimmetrico** e cioè per ogni  $x, y \in V$  abbiamo

$$x \leq y, y \leq x \implies x = y$$

Le prime due proprietà e l'ultima sono le proprietà dell'ordinamento parziale mentre la seconda e la terza ci dicono che questo è compatibile con la struttura di spazio vettoriale, nel senso che è possibile risolvere le disuguaglianze nello stesso modo e con le stesse regole che avete imparato per i numeri reali. Abbiamo già utilizzato questa definizione.

**Esempio.** Nello spazio  $\mathbb{R}^p$  abbiamo definito un ordinamento parziale tramite, il cono  $C$  dei vettori con componenti non negative,

$$C := \mathbb{R}_+^p = \{x \in \mathbb{R}^p : x_i \geq 0, i = 1, 2, \dots, p\}$$

allora  $x \leq y$  vuol dire che  $x_i \leq y_i, i = 1, 2, \dots, p$

### 4.1.3 Modelli lineari di produzione

L'esempio che descriviamo mostra come sia utile la notazione matriciale nel modellizzare alcune situazioni concrete.

Una compagnia mineraria produce sia lignite che antracite e si trova nella favorevole situazione di poter vendere tutto il carbone che riesce a produrre. Attualmente il profitto per la vendita di ogni tonnellata di lignite e di antracite è, rispettivamente, di euro 600 e 500. A causa di diversi tipi di restrizioni, l'impianto di estrazione, quello di cernita e quello di lavaggio possono funzionare per non più di 12, 10 e 8 ore giornaliere. Per la lavorazione di una tonnellata di lignite occorre utilizzare per tre ore l'impianto di estrazione, per tre ore quello di cernita e per quattro ore quello di lavaggio. Per una tonnellata di antracite i corrispondenti tempi sono di quattro, tre e due ore. Il problema è quello di decidere quante tonnellate di ciascun tipo di carbone devono essere prodotte giornalmente per massimizzare il profitto. Per poter analizzare questo problema dobbiamo prima rappresentarlo tramite un modello matematico. Definiamo le seguenti variabili

$x_1$  = produzione giornaliera di lignite in tonnellate

$x_2$  = produzione giornaliera di antracite in tonnellate.

Quindi, per un piano di produzione che preveda  $x_1$  tonnellate di lignite e  $x_2$  di antracite, il profitto, in euro, sarà

$$600x_1 + 500x_2.$$

Possiamo formulare le limitazioni sul tempo di uso dei vari macchinari tramite disequazioni. Consideriamo ad esempio l'utilizzazione dell'impianto di estrazione, se produciamo  $x_1$  tonnellate di lignite al giorno allora il tempo totale di utilizzo della macchina estrattrice per la lignite sarà

$$3x_1 \text{ ore,}$$

per l'antracite otteniamo un tempo totale di utilizzo dell'impianto di estrazione di

$$4x_2 \text{ ore,}$$

ma l'uso complessivo giornaliero di questo impianto non può superare le dodici ore, abbiamo quindi il vincolo

$$3x_1 + 4x_2 \leq 12.$$

Formulazioni analoghe sono ottenibili per il limite di utilizzo degli altri impianti. Otteniamo quindi il seguente modello

$$\text{Massimizzare } 600x_1 + 500x_2$$

sotto i vincoli di disponibilità degli impianti

$$3x_1 + 4x_2 \leq 12$$

$$3x_1 + 3x_2 \leq 10$$

$$4x_1 + 2x_2 \leq 8$$

vogliamo inoltre che

$$x_1 \geq 0, \quad x_2 \geq 0$$

per il loro significato.

Vediamo ora come questo modello si possa descrivere utilizzando le notazioni introdotte in precedenza. Definiamo il vettore (riga) dei **profitti**

$$p = (600 \quad 500)$$

il vettore (colonna) delle **risorse**, nel nostro caso i tempi massimi di uso delle macchine,

$$b = (12, 10, 8)$$

il vettore (colonna) delle **variabili decisionali** che rappresentano la scelta da fare

$$x = (x_1, x_2)$$

ed infine quella che viene chiamata **matrice della tecnologia** i cui coefficienti descrivono le risorse necessarie per una produzione unitaria dei beni allo stato attuale della tecnologia.

$$A = \begin{pmatrix} 3 & 4 \\ 3 & 3 \\ 4 & 2 \end{pmatrix}.$$

Il problema si può scrivere come

$$\text{Massimizzare } px$$

sotto i vincoli

$$\begin{aligned} Ax &\leq b \\ x &\geq 0. \end{aligned}$$

Come avete visto, per descrivere questo semplice modello abbiamo avuto bisogno di vettori riga, colonna e di matrici. Notate come la descrizione matematica di questo problema rientra nel modello astratto che abbiamo descritto nell'introduzione: abbiamo prima definito lo spazi degli stati del sistema, lo spazio lineare  $\mathbb{R}^n$ , abbiamo poi descritto tramite un sistema di disuguaglianze l'insiemi degli stati ammissibili per il problema,  $Ax \leq b$ ,  $x \geq 0$ , ed infine abbiamo dato un criterio di scelta, il valore della produzione giornaliera, che ci permette di confrontare i diversi stati ammissibili, il nostro obiettivo è determinare quello migliore (ottimo). Un modello di questo genere rientra nella categoria dei cosiddetti modelli di programmazione lineare che impareremo a risolvere nella parte dedicata alla **Programmazione Lineare**.

In questo esempio la soluzione ottima è la seguente

$$x_1^* = \frac{4}{5}, \quad x_2^* = \frac{12}{5},$$

cioè la politica ottima consiste nell'estrarre 0.8 tonnellate di lignite e 2.4 tonnellate di antracite per un profitto complessivo di euro 1,680. Si noti che ciò si realizza sfruttando per 12 ore l'impianto di estrazione per 9 ore e 36 minuti quello di cernita e per 8 ore quello di lavaggio.

Consideriamo ora un problema differente ma che risulterà avere uno stretto legame con il precedente. Supponiamo che una holding abbia interesse a noleggiare gli impianti della compagnia per produrre lignite e antracite.

La holding è incerta su quali prezzi orari di noleggio debba proporre alla compagnia per i tre tipi di impianti e propone un prezzo sapendo che la compagnia non accetterà mai di noleggiare i suoi impianti se non ottenendo un profitto almeno pari a quello che avrebbe utilizzandoli in proprio.

Se la holding volesse risolvere il problema di determinare l'offerta minima da fare, dovrebbe sapere qual'è il profitto per tonnellata che la compagnia ha per la lignite e per l'antracite. Supponiamo che la holding conosca questi dati.

Per determinare i prezzi orari minimi di noleggio, definiamo le seguenti variabili

$y_1$  = prezzo orario di noleggio dell'impianto di estrazione

$y_2$  = prezzo orario di noleggio dell'impianto di cernita

$y_3$  = prezzo orario di noleggio dell'impianto di lavaggio.

Ricordando i vari tempi necessari per la lavorazione della lignite, il profitto per la compagnia, se noleggia alla holding le macchine per estrarre una tonnellata di lignite, è di

$$3y_1 + 3y_2 + 4y_3$$

è poiché la compagnia vuole un guadagno almeno pari a quello che avrebbe lavorando in proprio dobbiamo imporre che

$$3y_1 + 3y_2 + 4y_3 \geq 600,$$

in modo analogo il vincolo per l'antracite che è

$$4y_1 + 3y_2 + 2y_3 \geq 500.$$

Vogliamo determinare quale è il prezzo minimo di noleggio giornaliero per i tre impianti. Ricordando il limite massimo di utilizzazione degli impianti, dobbiamo minimizzare

$$12y_1 + 10y_2 + 8y_3,$$

inoltre, per ovvi motivi, i prezzi di noleggio dovranno essere non negativi

$$y_1 \geq 0, \quad y_2 \geq 0, \quad y_3 \geq 0.$$

Se consideriamo il vettore (riga) dei prezzi

$$y = (y_1 \quad y_2 \quad y_3)$$

Il problema si può scrivere come

$$\text{Minimizzare } yb$$

sotto i vincoli

$$\begin{aligned} yA &\geq p \\ y &\geq 0 \end{aligned}$$

La soluzione ottima è

$$y_1 = 80, \quad y_2 = 0, \quad y_3 = 90$$

che corrisponde ad un prezzo giornaliero di noleggio di euro 1.680 eguale al profitto ottenibile utilizzando in proprio gli impianti in modo ottimale. Il motivo per il quale il prezzo di noleggio dell'impianto di cernita è zero è che nella soluzione ottima di uso degli impianti esso non viene usato per tutto il tempo per cui è disponibile e quindi il suo **valore marginale** è zero.

Per di poter risolvere problemi di questo tipo dobbiamo prima essere in grado di descrivere l'insieme degli stati ammissibili ovvero l'insieme delle soluzioni di un sistema di disequazioni.

## 4.2 Sistemi lineari

In questo paragrafo vogliamo far vedere come sia possibile utilizzare strumenti matematici teorici per ottenere informazioni su un modello economico semplice che descrive un mercato finanziario, si tratta di una prima descrizione molto elementare e semplificata che però ci permette già di fare alcune considerazioni.

### 4.2.1 Mercati finanziari

Iniziamo con il descrivere gli oggetti che compongono il mercato. Supponiamo che siano fissate in ordine crescente  $n$  date (future) che indicheremo con  $t_1, \dots, t_p$  e che  $t_0$  rappresenti la data odierna, chiameremo **scadenario** il vettore

$$T = (t_0, t_1, \dots, t_p),$$

e chiameremo **titolo (portafoglio)** un vettore

$$X = (x_0, x_1, \dots, x_p)$$

di  $\mathbb{R}^{p+1}$ , intendendo con questo che il pagamento  $x_0$  effettuato oggi garantisce un flusso di pagamenti futuri  $x_1, x_2, \dots, x_p$  alle scadenze  $t_1, t_2, \dots, t_p$ , poiché non vi sono restrizioni sul segno delle componenti,  $x_j$  può rappresentare un importo effettivamente pagato se  $x_j < 0$  o incassato se  $x_j > 0$ .

Notate come il vettore che descrive il portafoglio ha componenti fra di loro omogenee (importi) che si differenziano per il tempo in cui saranno disponibili, mentre, nel caso di un paniere di beni, le componenti sono non omogenee fra di loro ma disponibili tutte allo stesso momento.

Per costruire un modello astratto di un mercato finanziario faremo le seguenti ipotesi :

- Non ci sono costi di transazione o gravami fiscali
- Ogni titolo può essere trattato in qualsiasi quantità.
- Sono consentite vendite allo scoperto, si possono cioè vendere titoli che non si possiedono.

Dal punto di vista matematico queste ipotesi significano semplicemente che se ci sono due titoli sul mercato allora posso costruire un portafoglio che sia ottenuto come combinazione lineare di questi due. È chiaro che in questo modo portafogli (insiemi di titoli in diverse quantità) o titoli diventano indistinguibili. Questo tipo di ipotesi ci permette di introdurre la seguente definizione.

Dati  $k$  titoli  $X_1, X_2, \dots, X_k$  con  $k \leq p$  chiameremo **mercato finanziario** il sottospazio vettoriale di  $\mathbb{R}^{p+1}$

$$M = \text{Span}\{X_1, X_2, \dots, X_k\}.$$

Per rendere il modello corrispondente alla realtà sarà necessario fare la seguente ipotesi :

- Nel mercato non ci sono **arbitraggi**, ovvero non esistono titoli non nulli rappresentati da vettori con tutte le componenti non negative, in altri termini supporremo che l'unico vettore (portafoglio)

$$m \in M \cap \mathbb{R}_+^{p+1},$$

sia il vettore (portafoglio) nullo

Il significato economico di questa ipotesi è palese. Se non fosse soddisfatta potremmo costruire un portafoglio con flussi di pagamento tutti non negativi (ed almeno uno positivo) e quindi ad un esborso non negativo alla data odierna corrisponderebbe un flusso di pagamenti non negativi futuri, ipotesi auspicabile ma in genere non realizzabile.

Nel mercato possono essere presenti dei titoli dal significato particolare sono i titoli chiamati ZCB (zero coupon bond) ovvero TCN (titoli a cedola nulla) essi sono rappresentati da un vettore con *due sole componenti non nulle e con la seconda delle quali eguale ad uno* e quindi, in ipotesi di non arbitraggio, l'altra componente deve essere negativa.

Questi titoli corrispondono a titoli realmente presenti sul mercato (BOT, buoni ordinari del Tesoro) che hanno una caratteristica ben precisa che viene descritta dalla seguente ipotesi. Sia  $X \in M$  uno ZCB e sia  $p$  la sua prima componente non nulla allora

$$0 < -p < 1.$$

Questa ipotesi è a volte chiamata ipotesi di impazienza e può essere descritta come: *il prezzo al tempo  $t$  di una lira esigibile al tempo  $T > t$  è minore di uno ovvero una lira disponibile oggi è preferita ad una lira disponibile domani.*

Di una certa rilevanza è la dimensione di  $M$ : se è eguale ad  $p$ , ovvero se  $M$  è un iperpiano, diremo che il mercato è **completo** altrimenti lo chiameremo **incompleto**.

Il teorema principale che caratterizza i mercati finanziari è il seguente:

**Teorema 4.2.** *Sia  $M$  un mercato completo senza arbitraggi allora esiste un sistema di prezzi di equilibrio, ovvero esiste un vettore  $\pi = (\pi_0, \pi_1, \dots, \pi_p) \in \mathbb{R}^{p+1*}$ ,  $\pi > 0$  tali che il valore di ogni portafoglio sia zero, cioè per ogni  $X \in M$  si abbia*

$$\pi X = 0.$$

*Il vettore dei prezzi è unico a meno di costanti moltiplicative. Se vale l'ipotesi di impazienza allora i prezzi sono decrescenti ovvero*

$$\pi_i > \pi_j \quad \text{se } i < j.$$

Per identificare in modo unico il vettore dei prezzi di equilibrio supporremo sempre che  $\pi_0 = 1$  e questo vuol semplicemente dire che il valore del titolo è espresso nella stessa unità di misura del suo prezzo.

Spesso l'ipotesi che  $\pi X = 0$  per ogni  $X \in M$  viene descritta dicendo che il prezzo del portafoglio  $X$  è **equo**.

#### **Approfondimento.**

*Dimostrazione.* La prima parte del teorema afferma che esiste una funzione lineare  $\pi : \mathbb{R}^{p+1} \rightarrow \mathbb{R}$  tale che  $\pi X = 0$ , per ogni  $x \in M$ . Per trovare la funzione  $\pi$  procediamo nel seguente modo, sia  $\{v_1, v_2, \dots, v_p\}$  una base di  $M$ , e sia  $A$  la matrice che ha per colonne i vettori  $v_i$ ,  $i = 1, 2, \dots, p$  per costruzione abbiamo che

$$\text{Rank } A = p.$$

Il vettore  $\pi \in \mathbb{R}^{p*}$  ha la seguente proprietà

$$\pi A = 0 \quad \text{ovvero} \quad A^T \pi^T = 0 \quad (4.1)$$

Le seconda delle (4.1) rappresenta un sistema lineare omogeneo in  $p + 1$  incognite e poiché  $\text{Rank } A = \text{Rank } A^T$  il sistema ha infinite soluzioni che dipendono da un parametro e quindi  $\pi$  è determinato a meno di una costante moltiplicativa.

Facciamo vedere che dall'ipotesi di non arbitraggio segue che tutte le componenti di  $\pi$  sono tutte dello stesso segno.

- Se  $\pi_i = 0$  allora  $\pi_i e_i = 0$  e  $e_i$  è un arbitraggio.
- Se  $\pi_i \pi_j < 0$  allora  $-\pi_j e_i + \pi_i e_j \in M$  e poiché  $\pi(-\pi_j e_i + \pi_i e_j) = -\pi_j \pi_i + \pi_i \pi_j = 0$ , allora è un arbitraggio.

Dobbiamo infine verificare che, grazie all'ipotesi di impazienza, i prezzi sono decrescenti. Per  $k = 1, \dots, p$ , consideriamo i seguenti titoli

$$Y_k = (0, \dots, 0, -\frac{\pi_k}{\pi_{k-1}}, 1, 0, \dots, 0)$$

↑  
k-mo

essi appartengono al mercato e sono titoli a cedola nulla e quindi per l'ipotesi di impazienza abbiamo

$$\frac{\pi_k}{\pi_{k-1}} < 1, \quad k = 1, \dots, p$$

ovvero

$$\pi_k < \pi_{k-1}, \quad k = 1, \dots, p$$

e quindi i prezzi sono decrescenti □

### Alcune considerazioni sul significato del modello

Nelle ipotesi del teorema è possibile scegliere una base di  $M$  in modo che sia composta solo da titoli a cedola nulla con scadenze crescenti questo si può realizzare considerando i seguenti  $n$  vettori (titoli)

$$b_i = (-\pi_i, 0, \dots, 0, 1, 0, \dots, 0), i = 1, \dots, p$$

↑  
i-mo

essi sono  $n$  vettori linearmente indipendenti, come è immediato verificare, e che appartengono ad  $M$  poiché soddisfano  $\pi b_i = 0$ , essi sono quindi una base per  $M$  e rappresentano  $n$  titoli a cedola nulla con scadenza  $t_i$ ,  $i = 1, 2, \dots, p$  che generano tutti i titoli del mercato, li chiameremo **titoli elementari**.

Per evidenziare il loro significato poniamo

$$\pi_i = v(t_0, t_i), i = 1, 2, \dots, p$$

e cioè  $\pi_i$  è il prezzo del titolo  $b_i$ , e rappresenta il valore oggi di una lira esigibile al tempo  $t_i$ , per questo motivo questi  $n$  numeri si chiamano **prezzi a pronti** e complessivamente il vettore  $\Pi$  si chiama **struttura per scadenza dei prezzi a pronti**.

Consideriamo ora un titolo generico  $X \in M$ , abbiamo che

$$0 = \pi X = x_0 + \sum_{k=1}^p x_k v(t_0, t_k)$$

e quindi il valore attualizzato di ogni titolo è nullo, ovvero il suo prezzo  $x_0$  è eguale al valore attualizzato del flusso futuro di pagamenti. In maniera analoga possiamo trovare analoghi prezzi su qualsiasi intervallo. Consideriamo i titoli

$$Y_k = (0, \dots, 0, -\frac{\pi_k}{\pi_{k-1}}, 1, 0, \dots, 0)$$

$\uparrow$   
 k-mo

essi rappresentano delle operazione a termine e sono anch'essi dei Titoli a Cedola Nulla che a fronte di un pagamento al tempo  $t_{k-1}$  garantiscono l'incasso di una lira la tempo  $t_k$ . Possiamo porre

$$\frac{\pi_k}{\pi_{k-1}} = v(t_0, t_{k-1}, t_k)$$

ovvero

$$v(t_0, t_{k-1}, t_k) = \frac{v(t_0, t_k)}{v(t_0, t_{k-1})}$$

interpretandolo come il valore al tempo  $t_{k-1}$  di una lira esigibile al tempo  $t_k$ , valore fissato al tempo  $t_0$ , i prezzi di questi titoli si chiamano **prezzi a termine**.

#### 4.2.2 Un mercato finanziario

In questo paragrafo, descriviamo tramite un esempio .....

**Verifica della completezza**

Consideriamo un mercato finanziario generato dai seguenti titoli

$$T_1 := (-2, -2, 3, 0, 10), T_2 := (-1, 0, 0, 4, 0) T_3 := (-2, 2, 0, -3, 1.25)$$

$$T_4 := (-1, -2, 3, 4, 0) T_5 := (-3, -2, 6, -15, 28.75)$$

per verificare se si tratta di un mercato completo dobbiamo vedere se questi cinque titoli descrivono un iperpiano in  $R^5$ . Essenzialmente dobbiamo vedere se lo spazio generato da questi vettori ha dimensione quattro. Verifichiamo quanti di questi vettori sono indipendenti, per fare questo costruiamo la matrice

$$(T_1, T_2, T_3, T_4, T_5)$$

se effettuiamo la riduzione di Gauss otteniamo la seguente matrice, senza effettuare nessuno cambio di colonne

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

e quindi solo i primi quattro vettori sono indipendenti mentre il quinto è una loro combinazione lineare. Essi generano un iperpiano in uno spazio di dimensione cinque ed il mercato è completo.

**Determinazione del vettore dei prezzi**

Poiché il mercato è completo possiamo determinare il vettore dei prezzi, ovvero un vettore riga  $\Pi$ , tale che

$$\Pi T_i = 0, \text{ per } i = 1, 2, 3, 4$$

Se costruiamo la matrice  $A$  che ha per colonne i quattro vettori allora la condizione appena descritta può essere scritta come

$$\Pi T_i = A^T \Pi^T = 0$$

e quest'ultimo è un sistema lineare omogeneo di matrice incompleta

$$A^T := \begin{pmatrix} -2 & -2 & 3 & 0 & 10 \\ -1 & 0 & 0 & 4 & 0 \\ -2 & 2 & 0 & -3 & 8.75 \\ -1 & -2 & 3 & 4 & 0 \end{pmatrix}$$

che può essere ridotta a scala tramite l'algoritmo di Gauss ottenendo

$$\begin{pmatrix} -2 & -2 & 3 & 0 & 10 \\ 0 & 1 & -1.5 & 4 & -5 \\ 0 & 0 & 3 & -19 & 18.75 \\ 0 & 0 & 0 & 8 & -10 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

poiché abbiamo una variabile libera il sistema ha infinite soluzioni dipendenti da un parametro (la variabile libera). Se assegnamo a quest'ultima il valore  $\Pi_5 = 1$  otteniamo la soluzione

$$(2.5, 1.667, 1.25, 1)$$

Le componenti del vettore dei prezzi sono positive e quindi il mercato è senza arbitraggi. Il vettore dei prezzi  $\Pi$  è univocamente determinato dalla condizione di normalizzazione per la quale la sua prima componente è uno

$$\Pi := (1, 0.5, 0.333, 0.25, 0.2)$$

I prezzi sono decrescenti e quindi vale l'ipotesi di impazienza.

### Costruzione dei titoli elementari

Possiamo prendere come titoli rappresentativi del Mercato dei Titoli a Cedola Nulla elementari che generino il mercato stesso, essi sono ZCB con prezzo  $-\Pi_i$  e scadenza  $t_i$ . Essi sono

$$b_1 = (0.5, 1, 0, 0, 0), \quad b_2 = (0.33, 0, 1, 0, 0)$$

$$b_3 = (0.25, 0, 0, 1, 0), \quad b_4 = (0.2, 0, 0, 0, 1)$$

Essi sono chiaramente indipendenti.

### Riflessione sul significato della non esistenza di un arbitraggio

Poniamoci ora il seguente problema, supponiamo di costruirci un portafoglio  $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$  a partire dai titoli  $T_i$ ,  $i = 1, 2, \dots, T_5$  che ci garantisce quindi il flusso di pagamenti

$$\sum_{i=1}^5 \alpha_i T_i$$

quali tipi di flussi di pagamenti possiamo ottenere?

In termini matematici, sia  $B$  la matrice che descrive i titoli che sono nel portafoglio

$$B := \begin{bmatrix} -2 & -1 & -2 & -1 & -3 \\ -2 & 0 & 2 & -2 & -2 \\ 3 & 0 & 0 & 3 & 6 \\ 0 & 4 & -3 & 4 & -15 \\ 10 & 0 & 8.75 & 0 & 28.75 \end{bmatrix}$$

e sia  $\alpha \in \mathbb{R}^{p+1}$  il vettore che descrive le quantità dei titoli presenti nel portafoglio e che individua il portafoglio stesso il portafoglio  $\alpha$  genera il flusso di pagamenti

$$\begin{pmatrix} -2\alpha_1 - \alpha_2 - 2\alpha_3 - \alpha_4 - 3\alpha_5 \\ -2\alpha_1 + 2\alpha_3 - 2\alpha_4 - 2\alpha_5 \\ 3\alpha_1 + 3\alpha_4 + 6\alpha_5 \\ 4\alpha_2 - 3\alpha_3 + 4\alpha_4 - 15\alpha_5 \\ 10\alpha_1 + 8.75\alpha_3 + 28.75\alpha_5 \end{pmatrix}$$

Il problema che ci siamo posti può essere formulato come: per quali vettori  $Y$  il sistema  $B\alpha = Y$  ammette soluzioni. La risposta dovrebbe essere chiara da quanto detto sino ad ora. Il sistema è risolubile se e solo se  $\Pi Y = 0$  poiché dire che  $B\alpha = Y$  corrisponde a dire che  $Y$  è combinazione lineare delle colonne di  $B$  ed è quindi un elemento del mercato e quindi appartiene al  $\ker \Pi$ . Inoltre poiché abbiamo visto che il quinto titolo dipende linearmente dagli altri, possiamo sempre prendere  $\alpha_5 = 0$ .

Sia no

$$Y_1 := [-4, 2, 3, 4, 5], Y_2 := [1, 2, 3, 4, 5]$$

abbiamo

$$\langle \Pi, Y_1 \rangle = 0, \langle \Pi, Y_2 \rangle = 5$$

e quindi per il primo vettore il problema è risolubile per il secondo no. Infatti

$$[\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5] = [-1.25 - 2u, 0.25 + 3u, 2 - t_1, 2.25, t_1]$$

la soluzione dipende da un parametro che possiamo scegliere in modo che  $\Pi_5 = 0$ , basta sceglierlo in modo che

$$u = 0$$

sostituiamo nella espressione trovata per la soluzione

$$P = [1.25, 0.25, 2, 2.25, 0]$$

Passiamo al secondo vettore

Non ci sono soluzioni e non potrebbero essercene perché questo vettore rappresenta un arbitraggio avendo tutte le componenti positive.

### Mercati incompleti

**I prezzi in un mercato incompleto.** Consideriamo ora il caso in cui i titoli che generano il mercato non siano in numero sufficiente a garantire la completezza del mercato stesso. Scegliamo tre dei cinque vettori dell'esempio precedente e procediamo in modo del tutto analogo.

Costruiamo la matrice che ha per righe i tre vettori

$$A := \begin{pmatrix} -2 & -1 & -2 \\ -2 & 0 & 2 \\ 3 & 0 & 0 \\ 0 & 4 & -3 \\ 10 & 0 & 8.75 \end{pmatrix}$$

In questo caso le soluzioni del sistema omogeneo dipendono da due parametri e possono essere scritte come combinazione lineare dei due vettori

$$c := (4, 5.5, 6.333, 1, 0)$$

$$d := (0, 4.375, 6.5, 0, 1)$$

I vettori dei prezzi si ottengono dalle soluzioni che abbiamo trovato come combinazione lineare

$$\text{prezzi} := [4b, -4.375a + 5.5b, -6.5a + 6.333b, b, a]$$

E' facile determinare  $b$  poiché il primo coefficiente deve essere eguale ad uno e quindi  $b = \frac{1}{4}$

$$\Pi_0 := (1, -4.375a + 1.375, -6.5a + 1.583, 0.25, a)$$

vedete quindi che il vettore dei prezzi non é univocamente determinato. Per ottenere una vettore dei prezzi che garantisca la condizione di non arbitraggio e l'ipotesi di impazienza la condizione che dobbiamo imporre è quella di decrescenza e nonnegatività dei prezzi, cioè

$$\begin{aligned} -4.375a + 1.375 &\leq 1 \\ -6.5a + 6.333b &\leq -4.375a + 1.375 \\ 0.25 &\leq -6.5a + 6.333b \\ a &\leq 0.25 \\ 0 &\leq a \end{aligned}$$

questo insieme di disequazioni è risolto per

$$0.111 \leq a \leq 0.213$$

come vedete al variare di  $a$  abbiamo infiniti prezzi che soddisfano tutte le nostre ipotesi.

### Arbitraggi\*

**Determinazione di un portafoglio di arbitraggio.** In questo paragrafo consideremo un mercato in cui esistono arbitraggi

Consideriamo un mercato finanziario generato dai seguenti titoli

$$T_1 := [-1, 2, 0, 0, 0], T_2 := [-2, 2, -2, 0, 4]$$

$$T_3 := [-1, -1, 0, 2, 0], T_4 := [2, -1, 0, 0, 1]$$

si tratta di un mercato completo perché i quattro titoli sono indipendenti in uno spazio di dimensione cinque e generano quindi un iperpiano. Per determinare il vettore dei prezzi dobbiamo determinare il nucleo dell'operatore che ha i vettori che descrivono i titoli come righe (vedi esempio precedente)

$$A := \begin{bmatrix} -1 & -2 & -1 & 2 \\ 2 & 2 & -1 & -1 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 4 & 0 & 1 \end{bmatrix}$$

$$\Pi := [2, 1, -7, 1.5, -3]$$

come vedete il vettore dei prezzi non ha tutte le componenti positive e quindi siamo in presenza di arbitraggi.

Vogliamo ora determinare tutti i possibili arbitraggi, ovvero le soluzioni (non nulle) non negative del seguente problema

$$2x_1 + x_2 - 7x_3 + 1.5x_4 - 3x_5 = 0$$

è facile vedere come trovare una soluzione. *Il procedimento di costruzione degli arbitraggi è lo stesso che abbiamo seguito nel fare la dimostrazione del Teorema che ci garantisce che in ipotesi di non esistenza di arbitraggi il vettore dei prezzi aveva tutte le componenti positive.*

Ad esempio poiché  $x_1$  ed  $x_3$  hanno coefficienti di segno opposto che sono rispettivamente 2,  $-7$  allora il vettore che ha tutte le componenti zero tranne la prima e la terza i cui valori sono 7, 2 è un arbitraggio. Infatti sia

$$a_1 := [7, 0, 2, 0, 0]$$

allora

$$0$$

Per chiarire il significato di questa costruzione. Consideriamo una faccia del cono delle direzioni positive, essa è individuata dalle combinazioni lineari positive di due elementi  $e_i, e_j$  della base canonica. Se cerchiamo un vettore dell'iperpiano che appartiene alla faccia individuata da  $e_1, e_3$  dobbiamo risolvere l'equazione

$$2x_1 - 7x_3 = 0$$

che ha come soluzione quella che abbiamo descritto. Ovviamente l'intersezione non è vuota se e solo se i due coefficienti  $\Pi_i, \Pi_j$  sono uno positivo e l'altro non negativo, in effetti, l'equazione non ha soluzioni positive. In questo modo possiamo costruire molti arbitraggi il loro numero è esattamente eguale al prodotto del numero dei coefficienti positivi per il numero dei coefficienti non negativi. Sui tratta di ripetere quindi il precedente ragionamento molte volte questo è fatto dalla procedura

$$\begin{aligned} & \left[ \frac{7}{9}, 0, \frac{2}{9}, 0, 0 \right], \left[ \frac{3}{5}, 0, 0, 0, \frac{2}{5} \right] \\ & \left[ 0, \frac{7}{8}, \frac{1}{8}, 0, 0 \right], \left[ 0, \frac{3}{4}, 0, 0, \frac{1}{4} \right] \\ & \left[ 0, 0, \frac{3}{17}, \frac{14}{17}, 0 \right], \left[ 0, 0, 0, \frac{2}{3}, \frac{1}{3} \right] \end{aligned}$$

Abbiamo trovato gli arbitraggi *elementari*, tutti gli altri si possono determinare da questi per combinazione lineare convessa cioè ogni altro arbitraggio è esprimibile come  $\sum_{i=1}^6 \lambda_i ARB_i$  con  $0 \leq \lambda_i, i = 1, \dots, 6$

Determiniamo ora i portafogli che li generano. Sia

$$A := \begin{bmatrix} -1 & -2 & -1 & 2 \\ 2 & 2 & -1 & -1 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 4 & 0 & 1 \end{bmatrix}$$

i portafogli di arbitraggio elementari sono

$$P_1 := \left[ \frac{1}{3}, \frac{-1}{9}, 0, \frac{4}{9} \right]$$

$$P_2 := \left[ \frac{1}{5}, 0, 0, \frac{2}{5} \right]$$

$$P_3 := \left[ \frac{5}{8}, \frac{-1}{16}, 0, \frac{1}{4} \right]$$

$$P_4 := \left[ \frac{1}{2}, 0, 0, \frac{1}{4} \right]$$

$$P_5 := \left[ \frac{8}{17}, \frac{-3}{34}, \frac{7}{17}, \frac{6}{17} \right]$$

$$P_6 := \left[ \frac{1}{3}, 0, \frac{1}{3}, \frac{1}{3} \right]$$

questi sono portafogli di arbitraggio nel senso che le loro componenti indicano le quantità dei diversi titoli che dobbiamo acquistare per generare il flusso di pagamenti

$ARB_i$



**Parte III**  
**Soluzioni**



# Esercizi e Quiz

In questo capitolo sono contenute le soluzioni .....

## Soluzioni degli Esercizi

**Esercizio 1.1.** La matrice  $A$  non è scala poiché sotto il primo elemento non nullo della seconda riga  $a_{22} = 3$  vi è un elemento diverso da zero,  $a_{32} = 1$ .

La matrice  $B$  non è scala poiché sotto il primo elemento non nullo della seconda riga  $a_{23} = 2$ , in una colonna alla sua sinistra, vi è un elemento diverso da zero,  $a_{32} = 1$ .

La matrice  $C$  non è scala poiché la prima riga è nulla mentre quelle sottostanti non lo sono.

Esercizio 1.1

**Esercizio 1.2.** Poiché  $\hat{v} \in V_0$ , dalla definizione di sottospazio segue che esiste un vettore  $v_0$  tale che possiamo scrivere

$$\hat{v} = v_0 + w_0$$

per qualche  $w_0 \in W$ . Consideriamo ora un vettore  $x$  della forma  $x = \hat{v} + w$  allora possiamo scrivere

$$x = \hat{v} + w = v_0 + w_0 + w = v_0 + (w_0 + w)$$

poiché la somma di due vettori di  $W$ ,  $(w_0 + w)$  appartiene ancora a  $W$  abbiamo che  $x \in V_0$ .

Esercizio 1.2

# Elenco delle figure

1.1	Retta per l'origine: multipli di un vettore . . . . .	26
1.2	Piano per l'origine . . . . .	27
3.1	Sfera unitaria in dimensione 1 . . . . .	60
3.2	Sfera unitaria in dimensione 2 . . . . .	60
3.3	Sfera unitaria in dimensione 3 . . . . .	61