Algebra Lineare, a.a. 2018-19

CDL in informatica

Alessandro Berarducci

Dipartimento di Matematica
stanza 216, primo piano

Slides in costruzione e da sistemare.
In continuo aggiornamento

- Home page docente: http://people.dm.unipi.it/berardu/
- Home page del corso: https://elearning.di.unipi.it/
- Selezionare: Corso di Laurea in Informatica L-31, e poi "Matematica Discreta e Algebra Lineare 2018-19" (non 2017-18!).
- password: mdal2019

Matematica Discreta e Algebra Lineare

Ricevimento: ore 18, Stanza 216 del Dipartimento di Matematica, o per appuntamento.

Lunedì	11-13	MD
Martedì	14-16	AL
Mercoledì	9-11	MD
Giovedì	14-16	AL

Esame scritto e orale; il superamento dei compitini esonera dallo scritto.

Sezioni

Equazioni lineari

Definizione

Un'equazione lineare è un'equazione della forma

$$a_1x_1 + \ldots + a_nx_n = b$$

dove $x_1, \ldots, x_n$ sono incognite, e i coefficienti $a_1, \ldots, a_n, b$ sono numeri dati. Se $b = 0$ l'equazione si dice omogenea. In genere b si chiama "il termine noto" (è il termine che non contiene incognite).

- $3x = 0$ (lineare omogenea)
- $3x = 4$ (lineare non omogenea)
- $3x + 4y + 5z = 0$ (lineare omogenea)
- $3x + 4y + 5z = -4$ (lineare non omogenea)
- $3x = 4y$ (lineare omogenea perché equivale a $3x - 4y = 0$)
- $3x^2 + 2x + 4 = 0$ (quadratica, non lineare)
- $xy = 4$ (non lineare)
- $2^x + x = y$ (non lineare)

Zero, una o infinite soluzioni

Un'equazione lineare in una variabile ha zero, una o infinite soluzioni in $\mathbb{R}$.

- $3x = 4$. Unica soluzione $x = 4/3$.
- $0x = 4$. Nessuna soluzione.
- $0x = 0$. Ogni valore di x va bene, infinite soluzioni.
- $x^2 = 9$ ha due soluzioni $x = \pm 3$, ma non è lineare.

Il campo dei coefficienti

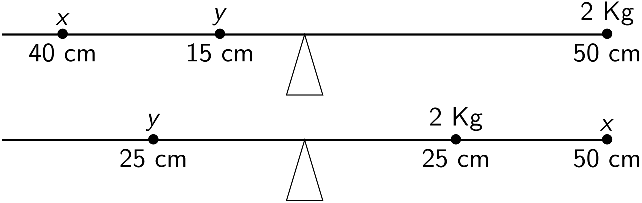
1. Nel modulo di Algebra Lineare considereremo prevalentemente equazioni lineari
$$a_1x_1 + \ldots + a_nx_n = b$$
dove $a_1, \ldots, a_n, b$ sono in $\mathbb{R}$ e le soluzioni x_i vanno cercate in $\mathbb{R}$.
2. Più in generale Invece che su $\mathbb{R}$ considerereremo equazioni su $\mathbb{Q}, \mathbb{C}$ o su un altro "campo" come ad esempio $\mathbb{Z}/(p)$ con p primo.
3. Su $\mathbb{Z}$ bisogna fare più attenzione perché le divisioni non sono sempre possibili.
4. Ad esempio $3x = 2$ ha soluzione in $\mathbb{Q}$ (prendo $x = \frac{2}{3}$) ma non in $\mathbb{Z}$.
5. L'equazione
$$15x + 5y = 4$$
ha infinite soluzioni in $\mathbb{R}$. Infatti posso scegliere y liberamente (ad esempio $y = 43$) e prendere $x = \frac{4-5y}{3}$.
6. L'equazione $15x + 5y = 4$ non ha soluzioni in $\mathbb{Z}$.
7. Lo studio delle equazioni lineari su $\mathbb{Z}$ viene fatto nel modulo di Matematica Discreta.

Sistemi lineari

Un sistema di equazioni lineari è un insieme di equazioni lineari che devono essere risolte simultaneamente per gli stessi valori delle variabili.

La leva di Archimede

Trovare i pesi x e y supponendo che nelle due pesate la bilancia sia in equilibrio (esempio tratto dal libro di Hefferon).



Legge di Archimede: il momento di una forza misura la tendenza di una forza a provocare una rotazione. Il momento generato da un peso p posto a distanza h dal fulcro è pari al prodotto ph. Per l'equilibrio serve che la somma dei momenti a sinistra del fulcro sia uguale a quella a destra.

$$\begin{cases} 40x + 15y = 50 \cdot 2 \\ 25y = 25 \cdot 2 + 50x \end{cases}$$

hi

Soluzione del sistema per sostituzione

- $$\begin{cases} 40x + 15y = 50 \cdot 2 \\ 25y = 25 \cdot 2 + 50x \end{cases}$$
- Dalla seconda ricaviamo $y = \frac{50+50x}{25} = 2 + 2x$
- Sostituendo nella prima $40x + 15(2 + 2x) = 100$.
- Diventa $40x + 30x = 100 - 30$.
- $x = \frac{70}{70} = 1$.
- $y = 2 + 2x = 4$.
- I pesi (in Kg) sono $x = 1$ e $y = 4$.

Zero, una o infinite soluzioni

Un sistema lineare su $\mathbb{R}$ ha zero, una o infinite soluzioni.

- $\begin{cases} 3x + 4y = 0 \\ x + 4y = 2 \end{cases}$ unica soluzione: $x = -1, y = 3/4$.
- $\begin{cases} 3x + 4y = 0 \\ 3x + 4y = 2 \end{cases}$ nessuna soluzione
- $\begin{cases} 3x + 4y = 2 \\ 6x + 8y = 4 \end{cases}$ infinite soluzioni.
  Le due equazioni si equivalgono. Scelgo y liberamente e $x = \frac{2-4y}{3}$.

- Le due equazioni del primo sistema rappresentano due rette che si intersecano nel punto $(-1, 3/4)$.
- Nel secondo sistema abbiamo le equazioni di due rette parallele.
- Nel terzo abbiamo due rette coincidenti.

Sistemi e matrici

Un sistema lineare può essere rappresentato da una matrice contenente i coefficienti delle incognite e i termini noti:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 5 \\ x_1 + x_2 + 2x_3 + 2x_4 = 2 \\ 2x_1 + 2x_2 + 3x_3 + 3x_4 = 3 \end{cases} \qquad \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 5 \\ 1 & 1 & 2 & 2 & 2 \\ 2 & 2 & 3 & 3 & 3 \end{array}\right]$$

A sinistra della barra verticale c'è la matrice dei coefficienti. Se includo i termini noti ottengo la matrice completa.
Se i termini noti sono tutti uguali a zero il sistema si chiama omogeneo. In questo caso per semplificare la notazione ometto i termini noti nella rappresentazione matriciale:

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

Algoritmo di Gauss: mosse di riga

Due sistemi sono equivalenti se hanno le stesse soluzioni. Le seguenti "mosse" transformano la matrice di un sistema nella matrice di un sistema equivalente.

1. Moltiplicare una riga della matrice per un numero diverso da zero (corrisponde a moltiplicare i coefficienti dell'equazione per un numero diverso da zero).
2. Sottrarre da una riga un multiplo di un'altra riga.
3. Scambiare due righe.

L'algorimo di Gauss consiste nel semplificare il sistema con una scelta oculata di queste mosse. Prima di spiegarlo in dettaglio diamo alcuni esempi.

Discutere il sistema omogeneo

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

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 \\ 2 & 2 & 3 & 3 \end{bmatrix} \xrightarrow{\substack{R_2 - R_1 \\ R_3 - 2R_1}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$\xrightarrow{R_3 - R_2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_3 + x_4 = 0 \end{cases}$$

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ \qquad\qquad\; x_3 + x_4 = 0 \end{cases}$$

Scegliamo x_2 e x_4 liberamente, e otteniamo $x_3 = -x_4$ e $x_1 = -x_2$.
Infinite soluzioni:

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

Discutere il sistema non omogeneo

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 5 \\ x_1 + x_2 + 2x_3 + 2x_4 = 2 \\ 2x_1 + 2x_2 + 3x_3 + 3x_4 = 3 \end{cases} \qquad \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 5 \\ 1 & 1 & 2 & 2 & 2 \\ 2 & 2 & 3 & 3 & 3 \end{array}\right]$$

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 5 \\ 1 & 1 & 2 & 2 & 2 \\ 2 & 2 & 3 & 3 & 3 \end{array}\right] \xrightarrow{\begin{array}{c} R_2 - R_1 \\ R_3 - 2R_1 \end{array}} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 5 \\ 0 & 0 & 1 & 1 & -3 \\ 0 & 0 & 1 & 1 & -7 \end{array}\right]$$

$$\xrightarrow{R_3 - R_2} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 5 \\ 0 & 0 & 1 & 1 & -3 \\ 0 & 0 & 0 & 0 & -4 \end{array}\right] \qquad \begin{cases} x_1 + x_2 + x_3 + x_4 = 5 \\ x_3 + x_4 = -3 \\ 0x_4 = -4 \end{cases}$$

Nessuma soluzione.

Per quali b_1, b_2, b_3 il sistema ha soluzione?

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = b_1 \\ x_1 + x_2 + 2x_3 + 2x_4 = b_2 \\ 2x_1 + 2x_2 + 3x_3 + 3x_4 = b_3 \end{cases} \qquad \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & b_1 \\ 1 & 1 & 2 & 2 & b_2 \\ 2 & 2 & 3 & 3 & b_3 \end{array}\right]$$

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & b_1 \\ 1 & 1 & 2 & 2 & b_2 \\ 2 & 2 & 3 & 3 & b_3 \end{array}\right] \xrightarrow{\begin{array}{c} R_2 - R_1 \\ R_3 - 2R_1 \end{array}} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & b_1 \\ 0 & 0 & 1 & 1 & b_2 - b_1 \\ 0 & 0 & 1 & 1 & b_3 - 2b_1 \end{array}\right]$$

$$\xrightarrow{R_3 - R_2} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & b_1 \\ 0 & 0 & 1 & 1 & b_2 - b_1 \\ 0 & 0 & 0 & 0 & b_3 - b_2 - b_1 \end{array}\right] \qquad \begin{cases} x_1 + x_2 + x_3 + x_4 = b_1 \\ x_3 + x_4 = b_2 - b_1 \\ 0 = b_3 - b_2 - b_1 \end{cases}$$

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = b_1 \\ \qquad\qquad x_3 + x_4 = b_2 - b_1 \\ \qquad\qquad\qquad 0 = b_3 - b_2 - b_1 \end{cases}$$

- Ha soluzione $\iff b_3 - b_2 - b_1 = 0$.
- Quando ha soluzione ne ha infinite: x_4 e x_2 sono libere;
- $x_3 = b_2 - b_1 - x_4$;
- $x_1 = b_1 - x_2 - x_3 - x_4 = b_1 - x_2 - (b_2 - b_1 - x_4) - x_4 = 2b_1 - b_2 - x_2$.
- $$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2b_1 - b_2 - x_2 \\ x_2 \\ b_2 - b_1 - x_4 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2b_1 - b_2 \\ 0 \\ b_2 - b_1 \\ 0 \end{bmatrix}$$

Discutere il sistema al variare del parametro a

$$\begin{cases} x + 2y - 3z = 4 \\ 4x + y + 2z = 6 \\ x + 2y + (a^2 - 19)z = a, \end{cases}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 4 & 1 & 2 & 6 \\ 1 & 2 & a^2-19 & a \end{array}\right] \xrightarrow{R_2 - 4R_1} \left[\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & -7 & 14 & -10 \\ 1 & 2 & a^2-19 & a \end{array}\right]$$

$$\xrightarrow{R_3 - R_1} \left[\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & -7 & 14 & -10 \\ 0 & 0 & a^2-16 & a-4 \end{array}\right] \qquad \begin{cases} x + 2y - 3z = 4 \\ \quad - 7y + 14z = -10 \\ \qquad (a^2 - 16)z = a - 4 \end{cases}$$

$$\begin{cases} x + 2y - 3z = 4 \\ \quad - 7y + 14z = -10 \\ \qquad (a^2 - 16)z = a - 4 \end{cases}$$

- Se $a = -4$, la terza equazione diventa $0z = -8$ e non ci sono soluzioni.
- Se $a = 4$, la terza equazione diventa $0z = 0$ e il sistema ha infinite soluzioni, una per ogni scelta di z.
- Se $a \neq \pm 4$, il termine $a^2 - 16$ è diverso da zero e abbiamo un'unica soluzione.

Attenzione a non dividere per zero!

La terza equazione $(a^2 - 16)z = a - 4$ può indurre nel seguente errore.

1. $(a^2 - 16)z = (a - 4)(a + 4)z = a - 4$;
2. Semplifico e ottengo $(a + 4)z = 1$.
3. Quindi per $a = 4$, ho $8z = 1$ e l'unica soluzione è $z = 1/8$.
4. D'altra parte per $a = 4$ l'equazione di partenza diventa $(4^2 - 16)z = 4 - 4$, ovvero $0z = 0$, che ha infinite soluzioni.

L'errore sta nel fatto che nel passaggio di "semplificazione" ho diviso per $a - 4$, ma questo è lecito solo nei casi in cui $a - 4 \neq 0$, ovvero se $a \neq 4$.

Quando ci sono dei parametri si rischia di dividere per zero senza accorgersene.

Forma ridotta

$$\begin{cases} & & & & 9x_5 = 5 \\ & 8x_2 & & +4x_4 & +2x_5 = 6 \\ & & 2x_3 & +9x_4 & +3x_5 = 7 \\ 3x_1 & & +4x_3 & +5x_4 & = 8 \\ & & & & 0x_5 = k \end{cases} \qquad \left[\begin{array}{ccccc|c} 0 & 0 & 0 & 0 & 9 & 5 \\ 0 & 8 & 0 & 4 & 2 & 6 \\ 0 & 0 & 2 & 9 & 3 & 7 \\ 3 & 0 & 4 & 5 & 0 & 8 \\ 0 & 0 & 0 & 0 & 0 & k \end{array}\right].$$

- Un coefficiente di una matrice è un pivot di riga se è il primo coefficiente diverso da zero della sua riga.
- Una matrice si dice in forma ridotta se i pivot di riga sono su colonne diverse.
- Un sistema la cui matrice dei coefficienti è in forma ridotta è risolubile se e solo se non vi sono equazioni impossibili del tipo $0x = k$ con $k \neq 0$. In altre parole, vi sono lo stesso numero di pivot nella matrice dei coefficienti (la parte a sinistra della barra) e nella matrice completa.
- Se il sistema ha soluzione, le variabili nelle colonne senza pivot sono libere e il loro valori possono essere scelti a piacere.
- Nell'esempio il sistema è risolubile se $k = 0$. I valori delle variabili si determinano procedendo da destra a sinistra: $x_5 = 5/9$, x_4 lo scelgo liberamente, poi trovo x_3, ecc.

Matrici a scalini per riga

Permutando le righe di una matrice ridotta, posso fare in modo che la matrice sia a scalini, ovvero le eventuali righe con tutti zeri siano in fondo e il pivot di ciascuna riga sia più a sinistra del pivot della riga sottostante (se non nulla).

$$\left[\begin{array}{ccccc|c} 0 & 0 & 0 & 0 & 9 & 5 \\ 0 & 8 & 0 & 4 & 2 & 6 \\ 0 & 0 & 2 & 9 & 3 & 7 \\ 3 & 0 & 4 & 5 & 0 & 8 \\ 0 & 0 & 0 & 0 & 0 & k \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} 3 & 0 & 4 & 5 & 0 & 8 \\ 0 & 8 & 0 & 4 & 2 & 6 \\ 0 & 0 & 2 & 9 & 3 & 7 \\ 0 & 0 & 0 & 0 & 9 & 5 \\ 0 & 0 & 0 & 0 & 0 & k \end{array}\right].$$

Algoritmo di Gauss

Teorema

Con mosse di Gauss ogni matrice si trasforma in una in forma ridotta.

Dimostrazione

- Se vi è una sola riga o se tutti i coefficienti sono zero la matrice è già ridotta.
- Altrimenti considero una riga R che ha il pivot più a sinistra possibile.
- Può essere che vi siano altre righe con il pivot sulla stessa colonna di quello della R, ma sottraendo multipli di R dalle altre righe posso annullarli, e fare quindi in modo che i pivot delle altre righe si spostino più a destra.
- Itero il procedimento per mettere in forma ridotta la matrice privata della riga R.
- Rimetto al suo posto la riga R e ho terminato.

Devo partire dalla colonna più a sinistra affinché nel seguito del procedimento non si ricreino altri pivot su quella colonna.

Discutere il sistema al variare del parametro $k \in \mathbb{R}$

$$\begin{cases} x + ky + (1+4k)z = 1+4k \\ 2x + (k+1)y + (2+7k)z = 1+7k \\ 3x + (k+2)y + (3+9k)z = 1+9k \end{cases}$$

$$\left[\begin{array}{ccc|c} 1 & k & 1+4k & 1+4k \\ 2 & k+1 & 2+7k & 1+7k \\ 3 & k+2 & 3+9k & 1+9k \end{array}\right] \xrightarrow{R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & k & 1+4k & 1+4k \\ 0 & -k+1 & -k & -1-k \\ 3 & k+2 & 3+9k & 1+9k \end{array}\right]$$

$$\xrightarrow{R_3 - 3R_1} \left[\begin{array}{ccc|c} 1 & k & 1+4k & 1+4k \\ 0 & -k+1 & -k & -1-k \\ 0 & -2k+2 & -3k & -2-3k \end{array}\right] \xrightarrow{R_3 - 2R_2} \left[\begin{array}{ccc|c} 1 & k & 4k+1 & 4k+1 \\ 0 & 1-k & -k & -k-1 \\ 0 & 0 & -k & -k \end{array}\right]$$

La matrice è diventata

$$A = \left[\begin{array}{ccc|c} 1 & k & 4k+1 & 4k+1 \\ 0 & 1-k & -k & -k-1 \\ 0 & 0 & -k & -k \end{array}\right].$$

Non conoscendo il valore di k non so se sia ridotta o quanti pivot abbia.

- Se $k \neq 0$ e $k \neq 1$ la matrice ha tre pivot e nessuna variabile libera, quindi c'è un'unica soluzione.
- Se $k = 0$, la matrice A ha due pivot nelle prime due righe e la terza riga ha solo zeri. La terza riga rappresenta l'equazione $0z = 0$ e la z è libera. Il sistema ha infinite soluzioni, una per ogni scelta di z.
- Se $k = 1$, la matrice va ulteriormente ridotta:

$$\text{A} = \left[\begin{array}{ccc|c} 1 & 1 & 5 & 5 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & -1 & -1 \end{array}\right] \xrightarrow{\quad R_3 - R_2 \quad} \left[\begin{array}{ccc|c} 1 & 1 & 5 & 5 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

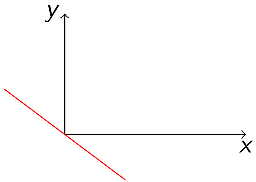
  La terza riga corrisponde all'equazione impossibile $0z = 1$ e non ci sono soluzioni.

Sezioni

Rappresentazione cartesiana e parametrica di una retta passante per l'origine

Retta passante per l'origine:

- Forma cartesiana: $3x + 4y = 0$ (equazione omogenea).
- Forma parametrica: Se pongo $x = 4t$, ho $y = -3t$, quindi $\begin{bmatrix} x \\ y \end{bmatrix} = t \begin{bmatrix} 4 \\ -3 \end{bmatrix}$. Al variare del parametro t si ottengono tutti i punti della retta.



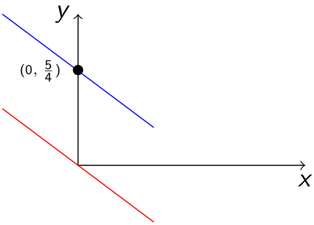
- L'insieme dei punti $(x, y) \in \mathbb{R}^2$ che verificano un'equazione della forma $ax + by = 0$ con almeno uno tra a e b diverso da zero, è una retta passante per l'origine. Se $b \neq 0$ si può scrivere nella forma $y = \frac{a}{b}x$ ed è una retta di pendenza $\frac{a}{b}$.
- Se $b = 0$, ho $ax = 0$ che rappresenta la retta verticale $x = 0$ in $\mathbb{R}^2$.

Rappresentazione cartesiana di una retta

Retta non passante per l'origine:

- Forma cartesiana: $3x + 4y = 5$ (equazione non omogenea).
- Forma parametrica: $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ -7/4 \end{bmatrix} + t \begin{bmatrix} 4 \\ -3 \end{bmatrix}$ (soluzione particolare + soluzioni dell'omogenea associata).

Viene una retta parallela a quella dell'equazione omogenea associata $3x + 4y = 0$.



Rango di una matrice

Definizione

Due matrici si dicono equivalenti per righe se con mosse di Gauss posso trasformare l'una nell'altra. Il rango di una matrice A è il numero di pivot nella matrice ridotta equivalente ad A per righe.

Sezioni

Settimana 25-29 febbraio

Matrici. Applicazione matrice vettore. Somma e prodotto di Matrici. Matrice identità. Matrice inversa.

Struttura delle soluzioni di un sistema lineare: soluzione generale $=$ soluzione particolare $+$ soluzione dell'omogenea. Combinazioni lineari e span di un insieme di vettori in $\mathbb{R}^n$.

Somma di matrici

Una matrice $m \times n$ è una tabella di numeri con m righe e n colonne. Per sommare due matrici della stessa dimensione si sommano i coefficienti nelle posizioni corrispondenti. Ad esempio nel caso di matrici 3×2 abbiamo

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} + \begin{bmatrix} 8 & 1 \\ 1 & 5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 3 \\ 4 & 9 \\ 5 & 7 \end{bmatrix}$$

Scrivo $A = [a_{ij}]$ per dire che a_{ij} è il coefficiente che si trova nella riga i e colonna j della matrice A. In generale $[a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}]$.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{11} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \\ a_{31} + b_{31} & a_{32} + b_{32} \end{bmatrix}$$

Se la matrice avesse più di dieci righe o colonne dovrei mettere una virgola tra gli indici i, j, come ad esempio $a_{12,5}$ (se scrivessi a_{125} non saprei individuare riga e colonna).

Prodotto per uno scalare

Uno scalare è un numero. Posso moltiplicare uno scalare r per una matrice A moltiplicando tutti i coefficienti della matrice per quello scalare.

$$3\begin{bmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{bmatrix} = \begin{bmatrix}3 & 6\\ 9 & 12\\ 15 & 18\end{bmatrix}$$

In generale se $A = [a_{ij}]$ è la matrice,

$$rA = r[a_{ij}] = [ra_{ij}].$$

Matrice trasposta

Data una matrice A di dimensioni $m \times n$, la sua trasposta A^T è la matrice $n \times m$ ottenuta scambiando le righe con le colonne.

$$\text{Se } A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}, \qquad A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^T$$

Il coefficiente in posizione (i, j) di A^T coincide con il coefficiente in posizione (j, i) di A. Ad esempio ,

$$A^T_{31} = A_{13} = 5.$$

Se indico con A^T_{ij} il cofficiente in posizione ij della matrice A, abbiamo

$$A^T_{ij} = A_{ji}$$

Il primo dei due indici indica sempre la riga, non importa se lo scrivo con la lettera i o la j. Nella formula $A^T_{ij} = A_{ji}$, a destra dell'uguaglianza c'è il coefficiente di A nella riga j e colonna i. rr

Vettori riga e vettori colonna

Una matrice di dimensioni $1 \times n$ viene detta vettore riga e una matrice di dimensioni $n \times 1$ viene detta vettore colonna. Nei vettori riga per chiarezza separo gli argomenti con delle virgole.

$$[2, 5] \text{ è un vettore riga } 1 \times 2$$

$$\begin{bmatrix} 2 \\ 5 \end{bmatrix} \text{ è un vettore colonna } 2 \times 1$$

Il trasposto di un vettore riga è un vettore colonna e viceversa:

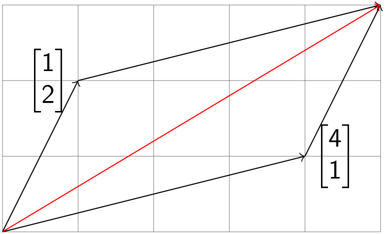
$$[2, 5]^T = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

Trasporre una matrice due volte è come non fare niente:

$$A^{TT} = A.$$

Regola del parallelogrammo

La somma di due vettori colonna in $\mathbb{R}^2$ si può rappresentare disegnando un parallelogrammo (analogamente per i vettori riga).



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

In fisica una forza applicata a $[0,0]$ in direzione $[1,2]$ sommata ad una forza applicata a $[0,0]$ in direzione $[4,1]$ è come una forza applicata a $[0,0]$ in direzione $[5,3]$.

Vettori in $\mathbb{R}^3$

- Un vettore colonna $\begin{bmatrix}1\\2\\3\end{bmatrix}$ (o riga) può rappresentare un punto nello spazio $\mathbb{R}^3$.
- Può anche rappresentare uno spostamento (indicato da una freccia) di 1 passo nella direzione dell'asse x, 2 passi nella direzione y, e 3 passi nella direzione z.
- Se ci troviamo in $\begin{bmatrix}5\\6\\7\end{bmatrix}$ ed effettuiamo lo spostamento $\begin{bmatrix}1\\2\\3\end{bmatrix}$, arriviamo in $\begin{bmatrix}5\\6\\7\end{bmatrix} + \begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix}6\\8\\10\end{bmatrix}$.
- Spostarsi lungo $\begin{bmatrix}1\\2\\3\end{bmatrix}$ e poi lungo $\begin{bmatrix}4\\5\\6\end{bmatrix}$ è come spostarsi lungo $\begin{bmatrix}5\\7\\9\end{bmatrix} = \begin{bmatrix}1\\2\\3\end{bmatrix} + \begin{bmatrix}4\\5\\6\end{bmatrix}$.

Spazi a più dimensioni

Fissate delle unità di misura e degli assi di riferimento,

- Una matrice 2×1 può essere pensata come un punto del piano $\mathbb{R}^2$.
- Una matrice 3×1 come un punto nello spazio $\mathbb{R}^3$.
- Una matrice 4×1 come un punto nello spazio quadridimensionale $\mathbb{R}^4$.
- Una matrice $n \times 1$ come un punto nello spazio ad n dimensioni $\mathbb{R}^n$.
- Ci sono anche le matrici 1×1, ad esempio la matrice $[3]$, che possiamo confondere con il numero 3, e pensare come ad un punto sulla retta numerica $\mathbb{R} = \mathbb{R}^1$.

Invece dei vettori colonna potevamo usare i vettori riga e definire $\mathbb{R}^n$ come l'insieme dei vettori riga $[x_1, \ldots, x_n]$ (successioni di lunghezza n) con $x_1, \ldots, x_n \in \mathbb{R}$. Talvolta conviene usare i vettori riga, talvolta i vettori colonna.

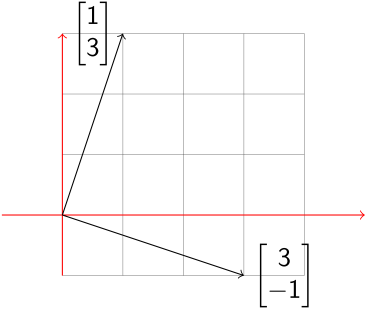
Prodotto di un vettore riga e un vettore colonna

Se $v = [a_1, \ldots, a_n]$ è un vettore riga e $w = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$ è un vettore colonna $n \times 1$, il prodotto $v \cdot w$ è la somma $a_1 b_1 + \ldots + a_n b_n = \sum_{i=1}^{n} a_i b_i$.
Ad esempio

$$[1, 2, 3] \cdot \begin{bmatrix} 0 \\ 5 \\ 6 \end{bmatrix} = 1 \cdot 0 + 2 \cdot 5 + 3 \cdot 6 = 28.$$

Ortogonalità

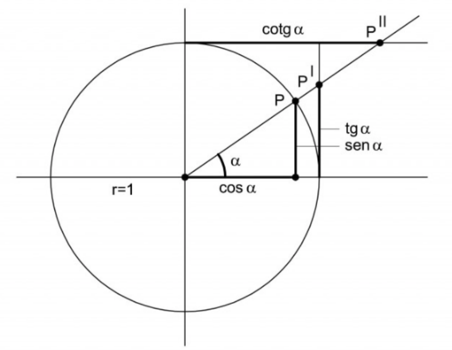
Dati due vettori v_1, v_2 in $\mathbb{R}^2$ (pensati come vettori colonna 1×2) il loro prodotto scalare $\langle v_1, v_2 \rangle \in \mathbb{R}$ è definito come $v_1^T \cdot v_2$. Il prodotto scalare è zero se e solo se i vettori sono ortogonali.



$$\begin{bmatrix} 1 \\ 3 \end{bmatrix}^T \cdot \begin{bmatrix} 3 \\ -1 \end{bmatrix} = [1, 3] \cdot \begin{bmatrix} 3 \\ -1 \end{bmatrix} = 1 \cdot 3 + 3 \cdot (-1) = 0$$

Coseno dell'angolo compreso

1. In generale dati due vettori v_1, v_2 in $\mathbb{R}^2$, il loro prodotto scalare $\langle v_1, v_2 \rangle$ è uguale $|v_1||v_2|\cos(\alpha)$, dove α è l'angolo compreso tra i due vettori.
2. Se i vettori sonon ortogonali, l'angolo è di $\pi/2$ radianti (90 gradi) e il coseno è zero.



Ortogonalità

- Dati $v_1, v_2 \in \mathbb{R}^n$ (pensati come vettori colonna $1 \times n$) il loro prodotto scalare $\langle v_1, v_2 \rangle \in \mathbb{R}$ è definito come $v_1^T \cdot v_2$.
- Abbiamo visto che per $n = 2$ i vettori sono ortogonali se e solo se il prodotto scalare è zero. La stessa cosa succede in $\mathbb{R}^3$ (esercizio basato sul teorema di Pitagora).
- Per $n = 2, 3$,
- Se $n > 3$ è impossibile fare le figure, ma per anologia con i casi $n = 2, 3$, diciamo che $v_1, v_2 \in \mathbb{R}^n$ sono ortogonali se e solo se il prodotto scalare è zero.

Prodotto di due matrici conformabili

Due matrici $A = [a_{ij}]$ e $B = [b_{ts}]$ si dicono conformabili (rispetto al prodotto) se il numero delle colonne di A è uguale al numero delle righe di B. In questo caso posso definire il prodotto $C = AB$ come segue. Se

- A è $m \times n$
- B è $n \times k$

allora $C = AB$ è $m \times k$; nella riga i e colonna j di C c'è

$$c_{ij} = [a_{i1}, \dots, a_{in}] \cdot \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} = \sum_{t=1}^{n} a_{it} b_{tj}$$
$$= a_{i1} b_{1j} + a_{i2} b_{2j} + \dots + a_{in} b_{nj}$$

ovvero c_{ij} è il prodotto della la riga i di A e la colonna j di B.

Esempio

Il prodotto di una matrice 2×3 e una 3×4 è una matrice 2×4.

$$\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 5 & 1 & 0 \\ 0 & -2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 4 & 2 \\ 2 & 11 & -2 & -3 \end{bmatrix}$$

Il coefficiente 4 si trova nella riga 1 colonna 3 della matrice prodotto e si ottiene facendo il prodotto

$$4 = [1, 1, 2] \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 1 \cdot 1 + 1 \cdot 1 + 2 \cdot 1$$

della prima riga della prima matrice e la terza colonna della seconda matrice.

Esempio

Il prodotto di una matrice 2×3 e una 3×4 è una matrice 2×4.

$$\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 5 & 1 & 0 \\ 0 & -2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 4 & 2 \\ 2 & 11 & -2 & -3 \end{bmatrix}$$

Il coefficiente 4 si trova nella riga 1 colonna 3 della matrice prodotto e si ottiene facendo il prodotto

$$4 = [1, 1, 2] \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 1 \cdot 1 + 1 \cdot 1 + 2 \cdot 1$$

della prima riga della prima matrice e la terza colonna della seconda matrice.

Esempio

Se moltiplico A per la colonna j di B, ottengo la colonna j di AB:

$$\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 5 & 1 & 0 \\ 0 & -2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 4 & 2 \\ 2 & 11 & -2 & -3 \end{bmatrix}$$

Le colonne di AB si ottengono moltiplicando A per le corrispondenti colonne di B:

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

$$\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 5 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ 11 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \end{bmatrix}$$

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

Base standard di $\mathbb{R}^n$

Considero $\mathbb{R}^n$ come ad un insieme di vettori colonna con n coefficienti reali. La base standard di $\mathbb{R}^n$ è data dalla lista degli n vettori colonna con un coefficiente uguale a uno e tutti gli altri uguali a zero.

- La base standard di $\mathbb{R}^2$ è data dalla lista

$$\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)$$

- La base standard di $\mathbb{R}^3$ è data dalla lista

$$\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right)$$

Nelle liste l'ordine conta: se permuto i vettori della lista ottengo un'altra base, non più quella standard.

Moltiplicando una matrice con n colonne per i vettori della base standard di $\mathbb{R}^n$ ottengo le varie colonne di A.

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

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

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

Il primo elemento della base standard fornisce la prima colonna di A, il secondo la seconda colonna di A, ecc.

Trasposta del prodotto

La trasposta del prodotto è il prodotto delle trasposte in ordine inverso:

$$(AB)^T = B^T A^T$$

Ad esempio se $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -3 \end{bmatrix}$ e $B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, abbiamo

- $$AB = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -3 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$
- $$B^T A^T = [0, 0, 1] \cdot \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 2 & -3 \end{bmatrix} = [2, -3] = (AB)^T$$

Dimostratelo usando la formula che definisce il prodotto di matrici.

Applicazione lineare associata ad una matrice

$$\begin{cases} 2x_1 + 3x_2 + 4x_3 + 5x_4 = b_1 \\ x_1 - x_2 + 2x_3 + 3x_4 = b_2 \\ 6x_1 - 2x_2 - 3x_3 + x_4 = b_3 \end{cases} \qquad \begin{bmatrix} 2 & 3 & 4 & 5 \\ 1 & -1 & 2 & 3 \\ 6 & -2 & -3 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

- Se A è una matrice 3×4 a coefficienti in $\mathbb{R}$, possiamo associare ad A una funzione $L_A : \mathbb{R}^4 \to \mathbb{R}^3$ che prende in input $\mathbf{x} \in \mathbb{R}^4$ e fornisce in output $\mathbf{b} = A \cdot \mathbf{x} \in \mathbb{R}^3$.
- Ad esempio se $A = \begin{bmatrix} 2 & 3 & 4 & 5 \\ 1 & -1 & 2 & 3 \\ 6 & -2 & -3 & 1 \end{bmatrix}$, allora

$$L_A \left(\begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix} \right) = \begin{bmatrix} 8 \\ 5 \\ 11 \end{bmatrix}.$$

- In questo modo possiamo rappresentare i sistemi lineari nella forma compatta $A \cdot \mathbf{x} = \mathbf{b}$.

Sezioni

Esercizio

Risolvere il sistema

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

- Scrivo la matrice del sistema e applico l'algoritmo di Gauss con la seguente miglioria: dopo aver raggiunto la forma ridotta, faccio altre mosse di riga in modo che i pivot siano uguali ad 1 e gli altri coefficienti nelle colonne dei pivot siano zero.
- La matrice è
$$\left[\begin{array}{ccc|c} 2 & -1 & 0 & 1 \\ -1 & -2 & -1 & 0 \\ 0 & -1 & 2 & 0 \end{array}\right]$$
- Nella prossima slide illustro la miglioria. Le mosse aggiuntive sono in rosso. La prima fase (Gauss) parte dalla colonna più a sinistra, la seconda (Jordan) da quella più a destra.

Algoritmo di Gauss - Jordan

$$\left[\begin{array}{ccc|c} 2 & -1 & 0 & 1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 0 \end{array}\right] \xrightarrow{R_2 + \frac{1}{2}R_1} \left[\begin{array}{ccc|c} 2 & -1 & 0 & 1 \\ 0 & \frac{3}{2} & -1 & \frac{1}{2} \\ 0 & -1 & 2 & 0 \end{array}\right]$$

$$\xrightarrow{R_3 + \frac{2}{3}R_2} \left[\begin{array}{ccc|c} 2 & -1 & 0 & 1 \\ 0 & \frac{3}{2} & -1 & \frac{1}{2} \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} \end{array}\right] \xrightarrow{R_2 + \frac{3}{4}R_3} \left[\begin{array}{ccc|c} 2 & -1 & 0 & 1 \\ 0 & \frac{3}{2} & 0 & \frac{3}{4} \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} \end{array}\right]$$

$$\xrightarrow{R_1 + \frac{2}{3}R_2} \left[\begin{array}{ccc|c} 2 & 0 & 0 & \frac{3}{2} \\ 0 & \frac{3}{2} & 0 & \frac{3}{4} \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} \end{array}\right] \xrightarrow{\frac{1}{2}R_1 \text{ ecc.}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{3}{4} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} \end{array}\right]$$

La matrice era 3×3 e ho ottenuto 3 pivot, quindi il sistema ha una ed una sola soluzione, e precisamente $[x_1, x_2, x_3] = [\frac{3}{4}, \frac{1}{2}, \frac{1}{4}]$.

$$\begin{cases} 2x_1 - x_2 = 1 \\ -x_1 - 2x_2 - x_3 = 0 \\ - x_2 + 2x_3 = 0 \end{cases} \qquad \left[\begin{array}{ccc|c} 2 & -1 & 0 & 1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 0 \end{array}\right]$$

$$\Big\downarrow \ \text{Gauss-Jordan}$$

$$\begin{cases} x_1 = \frac{3}{4} \\ x_2 = \frac{1}{2} \\ x_3 = \frac{1}{4} \end{cases} \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{3}{4} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} \end{array}\right]$$

- Se A è una matrice $n \times n$ tale che $A \rightsquigarrow I_n$ con Gauss-Jordan, allora per ogni vettore colonna $c \in \mathbb{R}^n$ di dimensioni $n \times 1$ il sistema $A\vec{x} = c$ ha una unica soluzione $\vec{x} = b \in \mathbb{R}^n$.
- La soluzione si trova facendo la riduzione $[A|c] \rightsquigarrow [I_n|b]$.
- Nel nostro esempio $c = [1, 0, 0]^T$ e $b = [\frac{3}{4}, \frac{1}{2}, \frac{1}{4}]^T$ (la T in alto indica la trasposta, perché lo dovrei mettere in verticale).

Equazioni matriciali

Un'equazione matriciale è una equazione in cui l'incognita da trovare è una matrice, come ad esempio $AX = B$ dove A, B sono matrici date e X è una matrice incognita.

Equazioni matriciali

- Posso risolvere simultaneamente due sistemi $A\vec{x} = c_1$ e $A\vec{x} = c_2$ mettendo entrambi i vettori colonna c_1 e c_2 a destra della barra (funziona anche con più di due).
- Ad esempio sia A una matrice $n \times n$ e consideriamo la matrice $[A|c_1 c_2]$ di dimensioni $n \times (n+2)$ ottenuta aggiungendo le colonne c_1, c_2 a destra della barra.
- Supponiamo che con Gauss - Jordan ottenga $[A|c_1 c_2] \rightsquigarrow [I_n|b_1 b_2]$.
- Allora b_1 è il vettore colonna che risolve il sistema $A\vec{x} = c_1$ e b_2 è il vettore colonna che risolve $A\vec{x} = c_2$.

$$\begin{aligned}
& [A|c_1 c_2] \rightsquigarrow [I_n|b_1 b_2] \quad \text{con mosse di Gauss - Jordan} \\
\iff & [A|c_1] \rightsquigarrow [I_n|b_1] \text{ e } [A|c_2] \rightsquigarrow [I_n|b_2] \quad \text{con le stesse mosse} \\
\iff & Ab_1 = c_1 \wedge Ab_2 = c_2 \\
\iff & A[b_1 b_2] = [c_1 c_2]
\end{aligned}$$

- Quindi $B = [b_1 b_2]$ risolve l'equazione matriciale $AB = C$ dove $C = [c_1, c_2]$
- Se lo applico al caso in cui $C = I_n$, come B trovo l'inversa di A (slide successiva).

Algoritmo per la matrice inversa

Sia A una matrice $n \times n$ e supponiamo $[A|I_n] \rightsquigarrow [I_n|B]$ con Gauss-Jordan, dove I_n la matrice identità $n \times n$ e B è una matrice $n \times n$. Per quanto visto prima B risolve l'equazione matriciale $AB = I_n$.

$$\left[\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 1 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_2 + \frac{1}{2}R_1} \left[\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & \frac{3}{2} & -1 & \frac{1}{2} & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3 + \frac{2}{3}R_2} \left[\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & \frac{3}{2} & -1 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} & \frac{2}{3} & 1 \end{array}\right] \xrightarrow{R_2 + \frac{3}{4}R_3} \left[\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ 0 & \frac{3}{2} & 0 & \frac{3}{4} & \frac{3}{2} & \frac{3}{4} \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} & \frac{2}{3} & 1 \end{array}\right]$$

$$\xrightarrow{R_1 + \frac{2}{3}R_2} \left[\begin{array}{ccc|ccc} 2 & 0 & 0 & \frac{3}{2} & 1 & \frac{1}{2} \\ 0 & \frac{3}{2} & 0 & \frac{3}{4} & \frac{3}{2} & \frac{3}{4} \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} & \frac{2}{3} & 1 \end{array}\right] \xrightarrow{\frac{1}{2}R_1 \text{ ecc.}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{array}\right]$$

- Verificate che $AB = I_n$, ovvero

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Una matrice $n \times n$ si dice invertibile o non singolare se esiste B tale che $AB = I_n$.In tal caso automaticamente $BA = I_n$ (da dimostrare) e B si chiama inversa di A. L'inversa di A se esiste è unica e si indica con la notazione A^{-1}. Quindi

$$AA^{-1} = A^{-1}A = I_n.$$

- Non è detto che con Gauss - Jordan a partire da una matrice A di dimensioni $n \times n$ si arrivi sempre ad I_n. Questo capita solo se con Gauss ottengo un pivot su ogni colonna (e quindi su ogni riga visto che la matrice è quadrata).
- Se ottengo n pivot, allora A è invertibile. L'inversa $B = A^{-1}$ si ottiene facendo $[A|I_n] \rightsquigarrow [I_n|B]$ con Gauss - Jordan.
- Esercizio: se ottengo meno di n pivot, A non è invertibile.

Notazioni per le matrici

- Indico con 0 una matrice composta interamente di zeri.
- Indico con I una matrice quadrata con 1 sulla diagonale e zero altrove.
- Ricordiamo che la somma $X + Y$ di due matrici è definita solo quando X, Y hanno le stesse dimensioni.
- Il prodotto XY è definito solo quando il numero di colonne di X è uguale al numero delle righe di Y.
- Se r è uno scalare (un numero), rX si ottiene moltiplicando tutti gli elementi della matrice X per r.
- X^T è la trasposta di X, ottenuta scambiando righe e colonne.

Nella slide seguente riassumo alcune proprietà.

Proprietà delle matrici

- $X + Y = Y + X$ (la somma è commutativa)
- $X + (Y + Z) = (X + Y) + Z$ (la somma è associativa)
- $0 + X = X + 0 = X$ (elemento neutro della somma)
- $X \cdot (Y \cdot Z) = (X \cdot Y) \cdot Z$ (il prodotto è associativo)
- $I \cdot X = X \cdot I = X$ se I, X sono quadrate della stessa dimensione.
- $XX^{-1} = X^{-1}X = I$ se X è quadrata ed invertibile.
- Se X, Y sono $n \times n$ invertibili, $(XY)^{-1} = Y^{-1}X^{-1}$.
- $I^{-1} = I$.
- $(X + Y)Z \quad = Z(X + Y) \quad = XZ + YZ$ (leggi distributive)
- $(rX)Y = X(rY) = rXY$ se r è uno scalare.
- Se X è invertibile e $r \neq 0$ è uno scalare, $(rX)^{-1} = \frac{1}{r}X^{-1}$.
- rI è una matrice quadrata con tutti r sulla diagonale e zero altrove. Essa commuta con qualsiasi altra matrice delle stesse dimensioni, ovvero $(rI)X = X(rI) = rX$.
- $(X + Y)^T = X^T + Y^T$.
- $(XY)^T = Y^TX^T$.

Proprietà dell'inversa

- Se A è invertibile, posso risolvere il sistema
$$A\vec{x} = b,$$
facendo
$$\vec{x} = A^{-1}b,$$
come fosse un'equazione numerica.
- Infatti da $A\vec{x} = b$ ottengo $A^{-1}A\vec{x} = A^{-1}b$ e $A^{-1}A\vec{x} = I_n\vec{x} = \vec{x}$.
- Se A è invertibile la soluzione del sistema è unica: $A\vec{x} = b \iff \vec{x} = A^{-1}b$.
- Una matrice di dimensione 1×1 si comporta come un numero: $[3]^{-1} = [\frac{1}{3}]$.
- Se A è una matrice diagonale (tutti elementi uguali a zero fuori della diagonale), l'inversa esiste se gli elementi sulla diagonale sono tutti diversi da zero e si calcola facendo gli inversi di quegli elementi:
$$\begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{4} \end{bmatrix}$$
- La matrice $\begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}$ è diagonale ma non è invertibile.

Moltiplicazione a blocchi

La moltiplicazione AB si può fare "a blocchi", ovvero si spezzano A e B in blocchi fatti di altre matrici di dimensioni compatibili e si moltiplicano come se al posto dei blocchi ci fossero dei numeri.

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$$

I blocchi A_{ij} potrebbero ad esempio essere matrici 2×2 che messe insieme formano una matrice 4×4 e analogamente per i B_{ij}. Conviene fare la divisione a blocchi se alcuni blocchi sono tutti zero.

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} & 0 \\ 0 & A_{22}B_{22} \end{bmatrix}$$

Se le matrici sono quadrate vale anche la formula

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22}^{-1} \end{bmatrix}$$

come per le matrici diagonali, salvo che sulla diagonale invece che degli scalari vi sono della matrici invertibili.

Inversa di una matrice 2×2

- Se $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ allora $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.
- Il numero $ad - bc$ si chiama determinante di A, scritto $\det(A)$.
- L'inversa di A esiste se e solo se $\det(A) \neq 0$.

Sezioni

Anelli

$\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}/(n)$ e le matrici $n \times n$ sono esempi di "anelli".

Definizione

Un anello è un insieme su cui sono definite delle operazioni $+, \cdot, -$ e due costanti $0, 1$ con le seguenti proprietà:

- $x + y = y + x$ *(la somma è commutativa)*
- $x + (y + z) = (x + y) + z$ *(la somma è associativa)*
- $0 + x = x + 0 = x$ *(elemento neutro della somma)*
- $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ *(il prodotto è associativo)*
- $1 \cdot x = x \cdot 1 = x$ *(elemento neutro del prodotto)*
- $(x + y)z = xz + yz \quad$ *e* $\quad z(x + y) = zx + zy$ *(leggi distributive)*
- $x + (-x) = 0$ *(opposto)*

Un'anello è commutativo se inoltre

- $x \cdot y = y \cdot x$ *(commutatività del prodotto)*

La sottrazione $x - y$ si definisce come $x + (-y)$.

Si noti che $\mathbb{N}$ non è un anello perché non ha l'opposto. Le matrici $n \times n$ sono un anello non commutativo (dove 1 è la matrice identità).

Negli anelli commutativi valgono tutti le proprietà della somma e del prodotto a cui siamo abituati, come ad esempio $0 \cdot x = 0$. Non le ho elencate tutte perché si deducono da quelle date.

Campi

Definizione

Un campo *è un anello commutativo che ha anche un'operazione $x \mapsto x^{-1}$ chiamata "reciproco" o "inverso moltiplicativo" tale che, per ogni $x \neq 0$, si ha*

$$x^{-1} \cdot x = x \cdot x^{-1} = 1.$$

0^{-1} *non può essere definito perché in un anello non c'è nessun numero che moltiplicato per* 0 *dia* 1.

La divisione è definita come $\frac{x}{y} = xy^{-1}$ ma si preferisce usare la notazione xy^{-1} se non sono nei campi usuali $\mathbb{R}$ o $\mathbb{C}$.

Esempi

- $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ sono campi perché ogni elemento diverso da zero ha un reciproco. In altre parole nei campi si può fare la divisione.
- $\mathbb{Z}$ è un anello ma non è un campo (si può fare la divisione "con resto" ma non la divisione esatta).
- $\mathbb{N}$ non è nemmeno un anello perché gli manca la sottrazione (è un semianello).
- $\mathbb{Z}/(n)$ è un anello; se n è primo è anche un campo (lo vedremo).
- Le matrici $n \times n$ (quadrate) sono un anello (non commutativo), ma non sono un campo.

Sezioni

Spazi vettoriali

Le matrici $m \times n$ su $\mathbb{R}$ sono un esempio di "spazio vettoriale" su $\mathbb{R}$.

Definizione

Sia $\mathbb{K}$ un campo. Gli elementi di $\mathbb{K}$ verranno chiamati scalari. Uno spazio vettoriale su $\mathbb{K}$ è un insieme V di oggetti, che chiameremo vettori sui quali siano definite le seguenti operazioni:

- *Se $v, w \in V$, si ha $v + w \in V$;*
- *$v + w = w + v$;*
- *$(u + v) + w = u + (v + w)$. Lo scrivo $u + v + w$.*
- *Esiste un vettore $\mathbf{0}$ tale che $\mathbf{0} + v = v + \mathbf{0} = v$ per qualsiasi $v \in V$.*
- *Dato un vettore v esiste un vettore $-v$ tale che $v + (-v) = \mathbf{0}$.*
- *Se $r \in \mathbb{K}$ e $v \in V$, allora $rv \in V$.*
- *Dati $r_1, r_2 \in \mathbb{K}$ e $v \in V$, $(r_1 r_2)v = r_1(r_2 v)$. Lo scrivo $r_1 r_2 v$.*
- *Se r è lo scalare -1, rv è la stessa cosa di $-v$.*
- *Se r è lo scalare 0, $0v$ è il vettore $\mathbf{0}$.*
- *Se $r_1, r_2 \in \mathbb{K}$ e $v \in V$, $(r_1 + r_2)v = r_1 v + r_2 v$.*
- *Se $r \in \mathbb{K}$ e $v, w \in V$, ho $r(v + w) = rv + rw$.*

Esempi

- Per ogni $n \in \mathbb{N}$, l'insieme $\mathbb{R}^n$ (pensato come vettori riga $1 \times n$ o vettori colonna $n \times 1$), è uno spazio vettoriale su $\mathbb{R}$.
- Per ogni $m, n \in \mathbb{N}$, l'insieme $M_{m,n}(\mathbb{R})$ delle matrici $m \times n$ è uno spazio vettoriale su $\mathbb{R}$.
- $\mathbb{R}$ stesso è uno spazio vettoriale su $\mathbb{R}$, dove gli elementi di $\mathbb{R}$ giocano sia il ruolo di vettori che di scalari.
- L'insieme $\mathbb{R}[x]$ dei polinomi a coefficienti in $\mathbb{R}$, i cui elementi sono espressioni della forma $a_0 + a_1x + \ldots + a_nx^n$ con $a_0, \ldots, a_n \in \mathbb{R}$, è uno spazio vettoriale su $\mathbb{R}$.
- Fissato $d \in \mathbb{N}$, l'insieme $\mathbb{R}[x]^{\leq d}$ dei polinomi di grado $\leq d$, sono uno spazio vettoriale su $\mathbb{R}$.

Precisazione

Quando diciamo che qualcosa è uno spazio vettoriale dovremmo sempre specificare come sono definite le operazioni. Ad esempio i polinomi $\mathbb{R}[x]$ sono uno spazio vettoriale rispetto alla usuale somma di polinomi fatta sommando i coefficienti delle potenze corrispondenti di x.

Sezioni

Combinazioni lineari e span

Definizione

- *Sia V uno spazio vettoriale su $\mathbb{K}$ e siano $v_1, \ldots, v_n$ elementi di V. Una combinazione lineare dei vettori $v_1, \ldots, v_n$ è un vettore $v \in V$ che si ottiene con una scrittura del tipo*

$$v = r_1 v_1 + \ldots + r_n v_n$$

  dove $r_1, \ldots, r_n \in \mathbb{K}$.
- *Scriviamo $v \in \mathrm{span}(A)$ se v si ottiene come combinazione lineare di vettori appartenenti all'insieme A.*

Proposizione

- *Se* $v \in \operatorname{span}(\{v1, \ldots, v_n\})$, *allora i vettori* $v, v_1, \ldots, v_n$ *sono linearmente dipendenti.*
- *Se* $v_0, v_1, \ldots, v_n$ *sono linearmente dipendenti, allora uno dei* v_i *è nello span degli altri.*

Dimostrazione

- Infatti se $v = r_1 v_1 + \ldots + r_n v_n$, la combinazione lineare $1v + (-r_1)v_1 + \ldots + (-r_n)v_n$ è uguale a zero (e non è la combinazione con tutti zeri perché il coefficiente di v è 1).
- Per il secondo punto, se $r_0 v_0 + \ldots + r_n v_n = 0$ e almeno uno dei r_i è $\neq 0$; dividendo per r_i e portando v_i dall'altro lato dell'uguaglianza vedo che v_i è nello span degli altri.

Sottospazi

Dato un sottoinsieme W di uno spazio vettoriale V, diciamo che W è un sottospazio di V, se ogni combinazione lineare di elementi di W è ancora in W. Questo equivale a dire che:

1. la somma di due elementi di W sia ancora in W;
2. Se $r \in \mathbb{K}$ e $v \in W$, allora $rv \in W$;
3. Il vettore $\mathbf{0}$ di V appartiene al sottoinsieme W.

Esempi

- L'insieme dei polinomi di grado esattamente d non è un sottospazio di $\mathbb{R}[x]$ perché la somma di due polinomi di grado d può non avere grado d (ad es. $(2x^3 + 5x + 1) + (-2x^3 + 4x^2) = 4x^2 + 5x + 1$).
- Dato $d \in \mathbb{N}$, l'insieme $\mathbb{R}[x]^{\leq d}$ dei polinomi di grado $\leq d$ è un sottospazio dello spazio vettoriale $\mathbb{R}[x]$.

Esercizi

Fare gli esercizi sugli spazi vettoriali nella raccolta di esercizi Berarducci-Papini.

Sezioni

Dimensione

Teorema

Sia A è una matrice quadrata $n \times n$ di rango n a coefficienti nel campo $\mathbb{K}$ allora il sistema $A\mathbf{x} = b$ ha una e una sola soluzione $\mathbf{x} \in \mathbb{K}^n$.

- $\mathbf{x}$ è un vettore colonna di variabili $x_1, \ldots, x_n$ (in colonna), mentre b è un vettore colonna di termini noti in $b_1, \ldots, b_n$.
- Se la matrice A ha rango n, una volta ridotta a scalini ha n pivot, ed essendo quadrata ha un pivot su ogni riga e ogni colonna.
- Per risolvere il sistema dobbiamo ridurre a scalini la matrice completa $[A|b]$.
- Siccome la forma ridotta di A ha un pivot su ogni colonna, non vi sono variabili libere e vi è una ed una sola soluzione.

Teorema

Un sistema di equazioni lineari con più variabili che equazioni o non ha soluzione o ne ha infinite (se è omogeneo ne ha infinite).

- Sia m il numero di equazioni ed $n > m$ il numero delle variabili.
- Rappresento il sistema nella forma $Ax = b$ dove A è una matrice $m \times n$, x è il vettore $n \times 1$ delle variabili, e b è il vettore $n \times 1$ dei termini noti.
- Riduco $[A|b]$ a scalini e controllo se il rango della matrice completa è uguale al rango della matrice dei coefficienti.
- Se i due ranghi sono diversi, vuol dire che nella forma ridotta $[A'|b']$ vi è un'equazione impossibile (tipo $0x = 1$) e il sistema non ha soluzione.
- Supponendo che i ranghi sono uguali vi è almeno una soluzione. Visto vi sono più variabili che equazioni, almeno una delle variabili è libera (ovvero nella sua colonna non vi sono pivot).
- Siccome posso fissare arbitrariamente il valore delle variabili libere vi sono infinite soluzioni.

Teorema

back *Nello $\mathbb{R}$-spazio vettoriale $\mathbb{R}^n$, non posso avere più di n vettori linearmente indipendenti. (La stessa cosa vale sostituendo $\mathbb{R}$ con un campo qualsiasi $\mathbb{K}$.)*

- Penso ad $\mathbb{R}^n$ come all'insieme dei vettori colonna $n \times 1$.
- Dati k vettori colonna $v_1, \ldots, v_k \in \mathbb{R}^n$ con $k > n$ considero la matrice $A = [v_1|\ldots|v_k]$ (di dimensioni $n \times k$) che ha i v_i come sue colonne.
- Il sistema omogeneo $A\mathbf{x} = \mathbf{0}$ ha n equazioni e k variabili ($\mathbf{x}$ è il vettore colonna delle variabili $x_1, \ldots, x_k$ e $\mathbf{0}$ è un vettore colonna di n zeri).
- Siccome il sistema è omogeneo e ha più variabili che equazioni esso ha infinite soluzioni. In particolare posso trovare una soluzione con almeno un x_i diverso da zero.
- Scrivere $A\mathbf{x} = \mathbf{0}$ è la stessa cosa che scrivere $x_1 v_1 + \ldots + x_k v_k = \mathbf{0}$; poichè esiste una soluzione con almeno un $x_i \neq 0$, i vettori $v_1, \ldots, v_k$ sono linearmente dipendenti.

Corollario

n vettori indipendenti di $\mathbb{R}^n$ generano $\mathbb{R}^n$ (e quindi formano una base).

- Sia $v_1, \dots, v_n$ una lista di n vettori indipendenti in $\mathbb{R}^n$.
- Se per assurdo esistesse un $v \in \mathbb{R}^n$ che non appartiene a $\text{span}(\{v_1, \dots, v_n\})$, gli $n+1$ vettori $v, v_1, \dots, v_n$ sarebbero indipendenti.
- Questo è assurdo perché abbiamo visto che in $\mathbb{R}^n$ non possono esserci $n+1$ vettori indipendenti vedi.

Sezioni

Equivalenza per righe

Definizione

Due matrici $m \times n$ si dicono equivalenti per righe se posso trasformare l'una nell'altra con mosse di Gauss di riga.

Osservazione

- *Se A e B sono equivalenti per riga, i sistemi $A\mathbf{x} = \mathbf{0}$ e $B\mathbf{x} = \mathbf{0}$ sono equivalenti, ovvero hanno le stesse soluzioni $\mathbf{x}$.*
- *La stessa cosa vale per i sistemi non omogenei $A\mathbf{x} = \mathbf{c}$, salvo che dobbiamo considerare le matrici complete $[A|\mathbf{c}]$.*
- *Ogni matrice si può ridurre a scalini per riga con l'algoritmo di Gauss.*

Matrici ridotte per colonne

Una matrice si dice ridotta per colonne se la sua trasposta è ridotta per righe.

Scalini per righe

$$\begin{bmatrix} 1 & * & * & * & * \\ 0 & 0 & 6 & * & * \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Scalini per colonne

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 0 & 0 & 0 \\ * & 6 & 0 & 0 \\ * & * & 0 & 0 \\ * & * & 3 & 0 \end{bmatrix}$$

Data una colonna non nulla, il suo coefficiente non nullo più in alto è il pivot di colonna. Una matrice è ridotta per colonne se non vi sono due pivot-colonna sulla stessa riga.

Equivalenza per colonne

Definizione

Due matrici A e B si dicono equivalenti per colonna, se posso trasfomare A in B con mosse di colonna (moltiplicare una colonna per uno scalare, scambiare due colonne, aggiungere ad una colonna un'altra moltiplicata per uno scalare).

Proposizione

Con mosse di colonna ogni matrice si può trasformare in una matrice ridotta per colonna (permutando le colonna posso anche renderla a scalini per colonna).

- Se siete abituati a lavorare per riga, potete trasporre la matrice, ridurre per riga, e trasporre il risultato.
- Altrimenti agite direttamente sulle colonne.

Teorema

In una matrice a scalini per righe, le righe non nulle sono indipendenti. In una matrice a scalini per colonna le colonne non nulle sono indipendenti.

- Faccio il caso colonna, l'altro è analogo.
- Consideriamo una matrice a scalini per colonna A e siano $C_1, \dots, C_n$ le sue colonne non nulle, con C_1 quella con il pivot più in alto e C_m quella con il pivot più in basso.
- Consideriamo una combinazione lineare di $C_1, \dots, C_n$ che dia come risultato una colonna nulla $x_1 C_1 + \dots + x_n C_n = \mathbf{0}$.
- Questo può capitare solo se $x_1 = 0$, perché in corrispondenza del pivot di C_1 le altre colonne hanno degli zeri.
- Quindi mi ritrovo con $x_2 C_2 + \dots + x_n C_n = \mathbf{0}$, e ragionando per induzione mostro che tutti gli x_i sono zero, dunque le colonne sono indipendenti.

Esempio

Le colonne di

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 7 & 0 \\ 3 & 8 & 0 \\ 4 & 9 & 2 \end{bmatrix}$$

sono indipendenti. Infatti supponiamo che

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 7 \\ 8 \\ 9 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Prima mostro che $x_1 = 0$, poi che $x_2 = 0$, infine che $x_3 = 0$.

Proposizione

Un insieme con meno di n vettori non può generare $\mathbb{R}^n$.

- Supponiamo che $k < n$ e consideriamo lo span di k vettori in $\mathbb{R}^n$.
- Sia $A = [v_1|\dots|v_k]$ la matrice $n \times k$ che ha come colonne le coordinate di questi vettori.
- Con mosse di colonna riduco A ad una matrice A' a scalini per colonna. Lo span delle colonne non è cambiato.
- Siccome $n > k$, il numero di righe è maggiore del numero delle colonne; è allora possibile formare una nuova matrice $[A'|\mathbf{c}]$ aggiungendo ad A' una nuova colonna $\mathbf{c}$ che ha il suo pivot ad altezza diversa, cosicché $[A'|\mathbf{c}]$ continua ad essere ridotta per colonne.
- Le colonne non nulle di una matrice ridotta per colonna sono indipendenti, quindi il vettore colonna $\mathbf{c} \in \mathbb{R}^n$ non è nello span delle colonne di A', e pertanto nemmeno di quelle di A.

Teorema

Ciascuna base di $\mathbb{R}^n$ contiene esattamente n vettori.

- Non può contenerne di meno altrimenti non genererebbero tutto $\mathbb{R}^n$.
- Non può contenerne di più altrimenti non sarebbero indipendenti.

Sezioni

Coordinate

- Sia V uno spazio vettoriale di dimensione n su un campo $\mathbb{K}$ e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V
  (una base è un insieme ordinato: scambiando di posto v_1 e v_2 otteniamo un'altra base).
- Dato $v \in V$ visto che i vettori della base generano V possiamo scrivere v come combinazione lineare

$$v = r_1 v_1 + \dots + r_n v_n.$$

  Poi vedremo che visto che i vettori della base sono indipendenti c'è un solo modo di scegliere gli r_i.
- I coefficienti della combinazione lineare sono le coordinate di v rispetto alla base $\mathcal{B}$.
- Associando a $v \in V$ il vettore colonna $v_{\mathcal{B}} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$ delle sue coordinate rispetto alla base $\mathcal{B}$ possiamo trattare V come fosse $\mathbb{K}^n$.
- più precisamente, la bigezione $v \in V \mapsto v_{\mathcal{B}} \in \mathbb{K}^n$ è un isomorfismo da V a $\mathbb{K}^n$. Le operazioni in V corrispondono alle omologhe operazioni in $\mathbb{K}^n$: $(v + w)_{\mathcal{B}} = v_{\mathcal{B}} + w_{\mathcal{B}}$ e

Esempio: matrici simmetriche 2×2

1. Sia V è lo spazio delle matrici simmetriche 2×2.
2. Una matrice è simmetrica se è uguale alla sua trasposta, quindi deve avere la forma $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$.
3. Possiamo scrivere
$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$
4. Quindi una base di V è data dall'insieme ordinato
$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$
5. Rispetto a questa base $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ ha coordinate $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$
6. Fissando una base possiamo trattare lo spazio delle matrici 2×2 come fosse $\mathbb{R}^3$, o detto più precisamente lo spazio delle matrici simmetriche 2×2 è *isomorfo* ad $\mathbb{R}^3$.

- Se come V prendiamo $\mathbb{K}^n$ e come base la base standard, la distinzione tra vettori e coordinate diventa quasi invisibile.
- Ad esempio il vettore $v = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \in \mathbb{R}^2$ ha come coordinate rispetto alla base standard $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ di $\mathbb{R}^2$ il vettore colonna $v_{\mathcal{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, perché posso scrivere

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

- Invece rispetto alla base $\mathcal{C} = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$, lo stesso vettore $v = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \in \mathbb{R}^2$ ha coordinate $v_{\mathcal{C}} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ perché posso scrivere

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Unicità delle coordinate

Sia V uno spazio vettoriale e $v_1, \ldots, v_n$ una sua base.

- Dato $v \in V$, voglio dimostare che le sue coordinate sono uniche, ovvero posso scrivere v come combinazione dei v_i in un solo modo.
- Più precisamente se

$$v = r_1 v_1 + \ldots + r_n v_n = b_1 v_1 + \ldots + b_n v_n,$$

  allora $r_i = b_i$ per ogni i.
- Per dimostrarlo considero la differenza $\mathbf{0} = (r_1 - b_1)v_1 + \ldots + (r_n - b_n)v_n$,
- se un r_i fosse diverso dal corrispondente b_i, avrei una combinazione dei vettori $v_1, \ldots, v_n$ uguale a zero con il coefficiente $r_i - b_i$ di v_i diverso da zero.
- Questo contraddice il fatto che i v_i erano indipendenti.

Invarianza della dimensione

- In modo analogo a quanto visto per $\mathbb{R}^n$ (o $\mathbb{K}^n$ con $\mathbb{K}$ campo), si dimostra che se V è uno spazio con una base di n vettori $\mathcal{B} = \{v_1, \ldots, v_n\}$, ogni altra base di V contiene esattamente n vettori.
- Per dimostrarlo ci si riconduce al caso di $\mathbb{K}^n$ associando a ciascun $v \in V$ il vettore colonna $v_{\mathcal{B}} \in \mathbb{K}^n$ delle sue coordinate.
- Allo stesso modo di dimostra che in uno spazio V di dimensione n, meno di n vettori non generano V e più di n non sono indipendenti.
- Si dimostra inoltre che n vettori indipendenti di V generano V, e che n vettori che generano V sono anche indipendenti.
- Da tutto ciò si deduce che se V è uno spazio di dimensione finita n e W è un sottospazio di V della stessa dimensione di V, allora $W = V$.

Esercizio

Siano $v_1, \dots, v_k$ vettori indipendenti di uno spazio V e sia $v \in V$. Le seguenti affermazioni sono equivalenti:

1. $v \notin \text{span}\{v_1, \dots, v_k\}$.
2. $v_1, \dots, v_k, v$ *sono indipendenti.*

- Assumo (1) e mostro (2). Se $v_1, \dots, v_k, v$ fossero dipendenti, ci sarebbe una combinazione lineare $r_1 v_1 + \dots + r_k v_k + r_{k+1} v = \mathbf{0}$ con almeno un r_i diverso da zero. Se r_{k+1} fosse diverso da zero, portando $r_{k+1} v$ dall'altro lato dell'uguaglianza con segno cambiato e dividendo per $-r_{k+1}$ ottengo $v \in \text{span}\{v_1, \dots, v_k\}$ contro le ipotesi.
- Assumo (2) e mostro (1). Se per assurdo $v \in \text{span}\{v_1, \dots, v_k\}$ esiste una combinazione lineare $v = a_1 v_1 + \dots + a_k v_k$.
  Portando v dall'altro lato ottengo
  $\mathbf{0} = -1v + a_1 v_1 + \dots + a_k v_k$, contraddicendo il fatto che $v, v_1, \dots, v_k$ sono indipendenti.

torna a completamento

Teorema del completamento ad una base

Teorema

Sia V uno spazio vettoriale di dimensione finita n e sia $\mathcal{B} = \{v_1, \dots, v_k\}$ un insieme ordinato di vettori linearmente indipendenti di V. Allora $k \leq n$ ed esistono $n - k$ vettori di V che aggiunti a quelli di $\mathcal{B}$ formano una base di V (se $k = n$ i vettori di $\mathcal{B}$ già formano una base).

- Se $k = n$ allora $v_1, \dots, v_k$ generano V e formano una base.
- Se $k < n$ sappiamo che $v_1, \dots, v_k$ non possono generare V e e quindi esiste almeno un vettore w_1 di V con $w_1 \notin \operatorname{span}\{v_1, \dots, v_k\}$. Esercizio
- $\implies v_1, \dots, v_k, w_1$ continuano ad essere indipendenti
- posso continuare nello stesso modo ottenendo una lista di vettori indipendenti $v_1, \dots, v_k, w_1, \dots, , w_i$ più lunga possibile.
- Se la lista non è ulteriormente prolungabile vuol dire che i vettori formano una base ed $i = n - k$.

◂ ritorno a nucleo-immagine. ◂ ritorno a Grassmann.

Algoritmo per completare ad una base di $\mathbb{K}^n$

- Sia $V = \mathbb{K}^n$ e siano $v_1, \ldots, v_k \in \mathbb{K}^n$ indipendenti, con $k < n$.
- Formo la matrice $M = [v_1, \ldots, v_k]$ (di dimensioni $n \times k$) che ha come colonne $v_1, \ldots, v_k$.
- Sia $W = \operatorname{span}\{v_1, \ldots, v_k\} \subseteq \mathbb{K}^n$. Siccome i v_i sono indipendenti, $\dim(W) = k$.
- Con mosse di colonna riduco M ad una matrice a scalini per colonna $M' = [v'_1, \ldots, v'_k]$.
- Le mosse di colonna non cambiano lo span delle colonne: $\implies \operatorname{span}\{v_1, \ldots, v_k\} = \operatorname{span}\{v'_1, \ldots, v'_k\} = W$.
- Nessun v'_i può essere zero altrimenti W avrebbe dimensione $< k$.
- Siccome M' ha più righe che colonne, posso aggiungere altri $n - k$ vettori in modo che la matrice quadrata $M'' = [v'_1, \ldots, v'_k, w_1, \ldots, w_{n-k}]$ (di dimensioni $n \times n$) sia ancora ridotta per colonne. (Come w_i posso scegliere opportuni vettori della base standard.)
- $w_1, \ldots, w_{n-k}$ sono indipendenti e non sono nello span di W, quindi aggiunti a $v_1, \ldots, v_k$ formano una base di V.

Esempio

Vedere dispense Gaiffi-Di Martino pag. 59, esempio 2.20.

Estrazione di una base

Teorema

Dato uno spazio vettoriale W, e sia $V = \operatorname{span}\{v_1, \ldots, v_m\} \subseteq W$. Esiste allora un sottoinsieme di $v_1, \ldots, v_m$ che forma una base di V. (In particolare $\dim(V) \leq m$.)

- Basta togliere da $v_1, \ldots, v_m$ quei vettori che sono nello span dei precedenti.
- Perché funziona? Lo span non cambia, quindi i vettori che rimangono generano ancora V.
- Per costruzione nessuno di quelli che rimane è nello span dei precedenti, quindi sono indipendenti (esercizio) e formano una base di V.

Stesso spazio nullo $\implies$ stesso rango

Teorema

Se due matrici A, B della stessa dimensione $m \times n$ hanno lo stesso spazio nullo, allora hanno anche lo stesso rango.

- Supponiamo per fissare le idee che le matrici siano 6×5 e scriviamo $A = [v_1, v_2, v_3, v_4, v_5]$ e $B = [w_1, w_2, w_3, w_4, w_5]$ per indicarne le colonne.
- Se applico A al vettore colonna $[x_1, x_2, x_3, x_4, x_5]^T$ ottengo $x_1 v_1 + x_2 v_2 + x_3 v_3 + x_4 v_4 + x_5 v_5$.
- Scelgo alcune colonne, ad esempio la prima, la terza e la quinta, e osservo che $A[x_1, 0, x_3, 0, x_5]^T = x_1 v_1 + x_3 v_3 + x_5 v_5$. Analogamente $B[x_1, 0, x_3, 0, x_5]^T = x_1 w_1 + x_3 w_3 + x_5 w_5$.
- Se le due matrici hanno lo stesso spazio nullo, $[x_1, 0, x_3, 0, x_5]^T \in \ker A \iff [x_1, 0, x_3, 0, x_5]^T \in \ker B$.
- Quindi $x_1 v_1 + x_3 v_3 + x_5 v_5 = \mathbf{0} \iff x_1 w_1 + x_3 w_3 + x_5 w_5 = \mathbf{0}$.
- La stessa cosa succede se si scelgono altre colonne. Quindi se certe colonne di A sono dipendenti/indipendenti lo sono anche le corrispondenti colonne di B e viceversa.
- Visto che il rango è dato dal massimo numero di colonne indipendenti, ne segue che A e B hanno lo stesso rango.

Il rango non cambia con mosse di riga o colonna

Corollario

Se transformo A in un'altra matrice B facendo mosse sia di riga sia di colonna, il rango di A è uguale al rango di B.

- Le mosse di colonna non cambiano lo span delle colonne, e quindi nemmeno la dimensione dello span delle colonne, che è proprio il rango.
- Le mosse di riga non cambiano lo spazio nullo, ma per il teorema precedente "stesso spazio nullo $\implies$ stesso rango", quindi non cambia nemmeno il rango.

◂ ritorna a estrazione base, ◂ ritorna a rango righe = rango colonne

Algoritmo per estrarre una base

Il seguente algoritmo funziona per spazi di vettori colonna. Per altri spazi (spazi di matrici, spazi di polinomi, ecc.) possiamo fissare una base e prendere i vettori colonna delle coordinate.

- Sia $V = \text{span}\{v_1, \ldots, v_m\} \subseteq \mathbb{K}^n$, dove $\mathbb{K}^n$ = vettori colonna $n \times 1$ con coefficienti nel campo $\mathbb{K}$.
- Per estrarre da $v_1, \ldots, v_m$ una base di V scrivo la matrice $M = [v_1, \ldots, v_m]$ che ha come colonne le coordinate dei v_i.
- Riduco M ad una matrice a scalini $M' = [v'_1, \ldots, v'_m]$ con mosse di riga.
- Se i pivot di M' sono nelle colonne $v'_{i_1}, v'_{i_2}, \ldots, v'_{i_k}$, allora M' ha rango k e $v'_{i_1}, \ldots, v'_{i_k}$ sono una base del suo spazio delle colonne.
- Siccome le mosse di riga (o colonna) non cambiano il rango e le dipendenze lineari tra le colonne rango non cambia, anche M ha rango k e $v_{i_1}, v_{i_2}, \ldots, i_k$ sono una base del suo spazio delle colonne.
- Queste colonne formano una base di V.

Per una dimostrazione leggermente diversa vedi Gaiffi-Di Martino, Oss. 2.51 p. 74.

Rango delle righe = rango delle colonne

Ricordiamo che data una matrice $m \times n$, il suo spazio delle righe è lo span dei vettori riga della matrice (un sottospazio di $\mathbb{K}^m$), e il suo spazio delle colonne è lo span dei suoi vettori colonna (un sottospazio di $\mathbb{K}^n$).

- Se una matrice A è a scalini per riga, le righe dove si trovano i pivot sono una base dello spazio delle righe e le colonne dove si trovano i pivot sono una base dello spazio delle colonne.
- Quindi almeno nel caso delle matrici a scalini, il rango per riga (dimensione dello span delle righe) è uguale al rango per colonna (dimensione dello span delle colonne).
- Detto in altri termini, il massimo numero di righe indipendenti, è uguale al massimo numero di colonne indipendenti.
- Funziona anche se la matrice non è a scalini perché con mosse di riga la posso ridurre a scalini senza alterare né lo spazio nullo (e quindi il rango delle righe) né il rango (che per definizione è la dimensione dello span delle colonne).

rango non cambia

Intersezione di sottospazi

Definizione

Dati due sottospazi A, B di uno spazio vettoriale V, la loro intersezione $A \cap B$ è data dai vettori di V che appartengono sia ad A sia a B, ovvero $A \cap B = \{v \in V \mid v \in A \wedge v \in B\}$.

Sezioni

Somma di sottospazi

Definizione

Dati due sottospazi A, B di uno spazio vettoriale V, definiamo la loro somma $A + B$ come l'insieme dei vettori di V che si possono scrivere come somma di un vettore di A ed uno di B:

$$\begin{aligned} A + B &= \{v \in V \mid \exists a \in A, b \in B : v = a + b\} \\ &= \{a + b \mid a \in A, b \in B\} \end{aligned}$$

Esercizio

- *Se* $A = \mathrm{span}\{v_1, \dots, v_m\}$ *e* $B = \mathrm{span}\{w_1, \dots, w_k\}$, *allora*

$$A + B = \mathrm{span}\{v_1, \dots, v_m, w_1, \dots, w_k\}.$$

- $A + B$ *è un sottospazio di* V.
- $A + B$ *include sia* A *sia* B.
- *se invece della somma prendo l'unione* $A \cup B = \{v \in V \mid v \in A \vee v \in B\}$ *non ottengo in generale un sottospazio. Ad esempio l'unione di due rette distinte in* $\mathbb{R}^2$ *passanti per l'orgine non è un sottospazio di* $\mathbb{R}^2$.

Somma di tre o più sottospazi

Definizione

La somma di tre sottospazi A, B, C di V è definita come l'insieme dei vettori di V che si possono scrivere come somma di un vettore di A uno di B e uno di C.
Analogamente definiamo la somma di n sottospazi

$$A_1 + \ldots + A_n = \{v_1 + \ldots + v_n \mid v_1 \in A_1, \ldots, v_n \in A_n\}$$

Osservazione

Se
$A = \text{span}\{v_1, \ldots, v_m\}$,
$B = \text{span}\{w_1, \ldots, w_k\}$,
$C = \text{span}\{u_1, \ldots, u_n\}$,
allora $A + B + C = \text{span}\{v_1, \ldots, v_m, w_1, \ldots, w_k, u_1, \ldots, u_n\}$.

Esempio

- $\mathbb{R}^3$ può essere visto come $A_1 + A_2 + A_3$ dove
$A_1 = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$, $A_2 = \text{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$, $A_3 = \text{span}\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.
- Se penso ad A_1, A_2, A_3 come agli assi x, y e z rispettivamente, allora $A_1 + A_2$ è il piano xy, $A_1 + A_3$ è il piano xz e $A_2 + A_3$ è il piano yz.
- L'intersezione tra il piano xy e il piano yz è dato dall'asse x:

$$(A_1 + A_2) \cap (A_2 + A_3) = A_1.$$

Somme dirette

Definizione

La somma $A + B$ di due sottospazi di uno spazio V è diretta se comunque prendo dei vettori non nulli $v \in A$ e $w \in B$, si ha che $v + w$ sono linearmente indipendenti. Se la somma $A + B$ è diretta, scrivo anche $A \oplus B$ invece di $A + B$.

Proposizione

La somma $A + B$ è diretta se e solo se $A \cap B = \{\mathbf{0}\}$.

- Se $\mathbf{0} \neq u \in A \cap B$, esistono $\mathbf{0} \neq v \in A$ e $\mathbf{0} \neq w \in B$ linearmente dipendenti: basta prendere $v = u, w = u$.
- Viceversa se $A \cap B = \mathbf{0}$, dati $\mathbf{0} \neq v \in A$, $\mathbf{0} \neq w \in B$ e una combinazione lineare $av + bw = \mathbf{0}$, devo mostrare che $a = b = 0$. A tal fine osservo che $av = -bw \in A \cap B$, e quindi $av = bw = \mathbf{0}$, da cui deduco $a = b = 0$.

Somma diretta di tre o più sottospazi

- Una somma $A + B + C$ di tre sottospazi di V è diretta se dati tre vettori diversi da zero $u \in A, v \in B, w \in C$, si ha sempre che u, v, w sono linearmente indipendenti.
- Definizione analoga vale per quattro o più sottospazi: la somma $A_1 + \ldots + A_n$ è diretta se prendendo da ciascun A_i un vettore v_i diverso da zero, si ha che $v_1, \ldots, v_n$ sono lineamente indipendenti.
- Se la somma è diretta, scriviamo anche $A_1 \oplus \ldots \oplus A_n$ invece di $A_1 + \ldots + A_n$.

Esempio

- I sottospazi di $\mathbb{R}^3$ dati da $A = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix} \right\}$ e $B = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 6 \end{bmatrix} \right\}$ sono in somma diretta.
- $A + B = A \oplus B = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 6 \end{bmatrix} \right\}$.
- Quando la somma è diretta, mettendo insieme le basi dei singoli spazi si ottiene una base della somma (se la somma non è diretta si ottiene solamente un'insieme di generatori). Ne segue che la dimensione della somma è la somma delle dimensioni
- Nel nostro esempio $\dim(A + B) = \dim(A) + \dim(B) = 2 + 1 = 3$.
- Siccome $A + B$ è incluso in $\mathbb{R}^3$ ed ha la stessa dimensione, deve coincidere con $\mathbb{R}^3$.

Esempio

- Due piani in $\mathbb{R}^3$ non sono mai in somma diretta, perché altrimenti la loro somma dovrebbe avere dimensione 4, il che non è possibile dovendo essere inclusa in $\mathbb{R}^3$.
- In $\mathbb{R}^4$ due spazi di dimensione 2 possono essere in somma diretta: ad esempio $A = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$ e $B = \text{span}\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ sono in somma diretta.

Dimensione della somma diretta di sottospazi

Teorema

Dati dei sottospazi $A_1, \dots, A_n$ *di uno spazio vettoriale* V*, la somma* $A_1 + \dots + A_n$ *è diretta se e solo se* $\dim(A_1 + \dots + A_n) = \dim(A_1) + \dots + \dim(A_n)$.

Dimostrazione

Dire che $A_1, \dots, A_n$ sono in somma diretta equivale a dire che posso scegliere delle basi di $A_1, \dots, A_n$ in modo tale che, mettendo insieme i vettori delle varie basi, ottengo un insieme di vettori linearmente indipendente (che quindi forma una base di $A_1 + \dots + A_n$).

Formula di Grassmann

Se abbiamo due sottospazi A, B di uno spazio V, abbiamo la seguente formula, che vale sia quando la somma è diretta sia quando non lo è.

Teorema

$\dim(A + B) = \dim(A) + \dim(B) - \dim(A \cap B)$.

Dimostrazione

1. Partiamo da una base $u_1, \ldots, u_k$ di $A \cap B$ ed estendiamola ad una $u_1, \ldots, u_k, v_1 \ldots, v_n$ di A e ad una base $u_1, \ldots, u_k, w_1, \ldots w_m$ di B. completamento ad una base
2. Basta dimostrare che una base di $A + B$ è data da $u_1, \ldots, u_k, v_1 \ldots, v_n, w_1, \ldots w_m$.
3. Chiaramente questi vettori generano $A + B$.
4. Per mostrarne l'indipendenza considero una combinazione $x_1 v_1 + \ldots + x_n v_n + y_1 u_1 + \ldots + y_k u_k + z_1 w_1 + \ldots + z_m w_m = \mathbf{0}$ e mostro che tutti i coefficienti devono essere zero.
5. $x_1 v_1 + \ldots + x_n v_n$ è in A, ma è anche in B visto che è uguale a $-(y_1 u_1 + \ldots + y_k u_k + z_1 w_1 + \ldots + z_m w_m)$, quindi è in $A \cap B$.
6. Posso allora scrivere $x_1 v_1 + \ldots + x_n v_n$ come combinazione lineare $a_1 u_1 + \ldots + a_k u_k$.
   $\implies x_1 v_1 + \ldots + x_n v_n - a_1 u_1 - \ldots - a_k u_k = \mathbf{0}$.
7. Tutti i coefficienti devono essere zero perché $v_1, \ldots, v_n, u_1, \ldots, u_k$ sono indipendenti.
8. In particolare gli x_i sono zero; ragionando simmetricamente, anche gli z_i sono zero.
9. La (4) si riduce a $y_1 u_1 + \ldots + y_k u_k = \mathbf{0}$, quindi anche gli y_i sono zero visto che $u_1, \ldots, u_k$ sono indipendenti.

Esempio

- I sottospazi di $\mathbb{R}^3$ dati da $A = \text{span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\2\\3\end{bmatrix}\right\}$ e

  $B = \text{span}\left\{\begin{bmatrix}1\\1\\6\end{bmatrix}, \begin{bmatrix}1\\3\\4\end{bmatrix}\right\}$ **non** sono in somma diretta.
- $A + B = \text{span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\2\\3\end{bmatrix}, \begin{bmatrix}1\\1\\6\end{bmatrix}, \begin{bmatrix}1\\3\\4\end{bmatrix}\right\}$, ma l'ultimo vettore è nello span dei precedenti, quindi possiamo anche scrivere $A + B = \text{span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\2\\3\end{bmatrix}, \begin{bmatrix}1\\1\\6\end{bmatrix}\right\}$. I tre vettori tra le graffe formano una base di $A + B$, quindi $\dim(A + B) = 3$.
- Per la formula di Grassmann:
  $\dim(A + B) = \dim(A) + \dim(B) - \dim(A \cap B)$.
- Nel nostro caso $\dim(A + B) = 2 + 2 - 1 = 3$.

Sezioni

Definizione

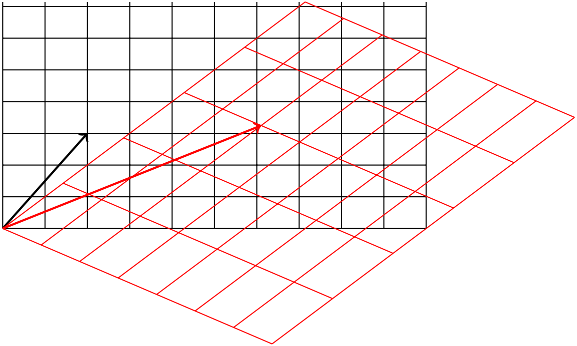
Dati due spazi vettoriali V, W su un campo $\mathbb{K}$ e una funzione $f : V \to W$, diciamo che f è lineare se

- $f(\mathbf{0}_V) = \mathbf{0}_W$ *(dove $\mathbf{0}_V$ è il vettore zero dello spazio V ecc.)*
- $f(r_1 v_1 + \ldots + r_n v_n) = r_1 f(v_1) + \ldots + r_n f(v_n)$.

I seguenti termini sono sinomimi: funzione lineare, applicazione lineare, trasformazione lineare.

Applicazioni lineari: interpretazione geometrica

Data un'applicazione lineare iniettiva $f : \mathbb{R}^2 \to \mathbb{R}^2$, se gli input si spostano lungo una retta (nera nel disegno), gli output corrispondenti si spostano lungo una retta (rossa nel disegno). Rette parallele vanno a finire in rette parallele. Stessa cosa avviene per funzioni lineari da $\mathbb{R}^m$ ad $\mathbb{R}^n$ con $m, n \leq 3$.



Se la f non è iniettiva, una retta potrebbe andare a finire in un singolo punto, come in una proiezione di $\mathbb{R}^2$ sull'asse x dove le rette verticali si schiacciano in un singolo punto dell'asse x.

Matrice associata ad un'applicazione lineare

Siano

- $V =$ spazio vettoriale di dimensione m con base fissata $\mathcal{B} = \{v_1, \ldots, v_m\}$
- $W =$ spazio vettoriale di dimensione n con base $\mathcal{C} = \{w_1, \ldots, w_n\}$.

Teorema

Data un'applicazione lineare $f : V \to W$ posso associare ad f una matrice $A = f_{\mathcal{C}}^{\mathcal{B}}$ di dimensioni $n \times m$ in modo tale che

$$f(v) = w \iff A \cdot b = c,$$

dove $b = v_{\mathcal{B}} \in \mathbb{K}^m$ è il vettore colonna delle coordinate di v e $c = w_{\mathcal{C}} \in \mathbb{K}^n$ è il vettore colonna delle coordinate di w rispetto alle basi scelte.

Dimostrazione

Basta prendere la matrice A che ha come i-esima colonna le coordinate di v_i rispetto alla base $\mathcal{C}$.

Sezioni

Nucleo e Immagine

Definizione

Sia $f : V \to W$ un'applicazione lineare con dominio V e codominio W.

1. *Il nucleo* $\ker(f)$ *di f è i sottoinsieme di V dato dai vettori $v \in V$ tali che $f(v) = \mathbf{0}$:*

$$\{v \in V \mid f(v) = \mathbf{0}\}$$

2. *L'immagine* $\mathrm{Im}(f)$ *è il sottoinsieme di W dato da tutte le possibili immagini $f(v) \in W$ dei vettori v di V:*

$$\mathrm{Im}\, f = \{f(v) \mid v \in V\}$$

Scriveremo anche $f(V)$ invece di $\mathrm{Im}\, f$ *per indicare l'immagine.*

Nucleo e Immagine come sottospazi

Proposizione

Sia $f : V \to W$ un'applicazione lineare.

1. *Il nucleo* $\ker(f) = \{v \in V \mid f(v) = \mathbf{0}\}$ *di f è un sottospazio vettoriale del dominio V.*
2. *L'immagine* $\operatorname{Im}(f) = \{f(v) \mid v \in V\}$ *di f è un sottospazio vettoriale del codominio W .*

Dimostrazione

1. Dobbiamo verificare che $\ker(f)$ sia chiuso rispetto alla somma di vettori e alla moltiplicazione di un vettore per uno scalare $a \in \mathbb{K}$.
   Somma: siano u e v in $\ker(f)$. Allora $f(u) = f(v) = \mathbf{0}$ e quindi $f(u + v) = f(u) + f(v) = \mathbf{0} + \mathbf{0} = \mathbf{0}$.
   Prodotto per uno scalare: se $a \in \mathbb{K}$, $f(au) = af(u) = a\mathbf{0} = \mathbf{0}$.
2. Stessa cosa per $\operatorname{Im}(f)$.
   Somma: siano $w_1, w_2 \in \operatorname{Im}(f)$. Allora esistono $v_1, v_2 \in V$ tali che $f(v_1) = w_1, f(v_2) = w_2$. Ne segue che $w_1 + w_2 = f(v_1) + f(v_2) = f(v_1 + v_2) \in \operatorname{Im}(f)$.
   Prodotto per uno scalare: se $a \in \mathbb{K}$,

Un'applicazione lineare è determinata dalla sua restrizione ad una base

Dimostrazione

Sia $f : V \to W$ un'applicazione lineare. Se $v_1, \ldots, v_n$ è una base di V e conosco $f(v_1), \ldots, f(v_n)$, posso determinare $f(v)$ per qualsiasi $v \in V$.

- Scrivo $v = a_1 v_1 + \ldots + a_n v_n$
- Osservo che

$$\begin{aligned} f(v) &= f(a_1 v_1 + \ldots + a_n v_n) \\ &= a_1 f(v_1) + \ldots + a_n f(v_n) \end{aligned}$$

Example

Sia $f : \mathbb{R}^2 \to \mathbb{R}^3$ un'applicazione lineare tale che

$$f\left(\begin{bmatrix}1\\2\end{bmatrix}\right) = \begin{bmatrix}3\\4\\5\end{bmatrix} \quad \text{e} \quad f\left(\begin{bmatrix}2\\3\end{bmatrix}\right) = \begin{bmatrix}6\\0\\2\end{bmatrix}$$

Calcolate $f\left(\begin{bmatrix}8\\13\end{bmatrix}\right)$.

- $\begin{bmatrix}8\\13\end{bmatrix} = 2\begin{bmatrix}1\\2\end{bmatrix} + 3\begin{bmatrix}2\\3\end{bmatrix}$.
- $\implies f\left(\begin{bmatrix}8\\13\end{bmatrix}\right) = 2\begin{bmatrix}3\\4\\5\end{bmatrix} + 3\begin{bmatrix}6\\0\\2\end{bmatrix} = \begin{bmatrix}24\\8\\16\end{bmatrix}$.
-

f è iniettiva se e solo se $\ker(f) = \{\mathbf{0}\}$

Ricordiamo che $\ker(f) = \{v \in V \mid f(v) = \mathbf{0}\}$.

Teorema

Sia $f : V \to W$ un'applicazione lineare. Allora $\ker(f) = \{\mathbf{0}\} \iff f$ *è iniettiva.*

- Assumo $\ker(f) = \mathbf{0}$. Da $f(u) = f(v)$ segue $f(u - v) = \mathbf{0}$, quindi $u - v \in \ker(f)$. Siccome $\ker(f) = \mathbf{0}$ ottengo $u - v = \mathbf{0}$, quindi $u = v$. Ho mostrato che f è iniettiva.
- Viceversa se $\ker(f) \neq \mathbf{0}$, sia $\mathbf{0} \neq v \in \ker(f)$. Sia v sia $\mathbf{0}$ vengono mappati in $\mathbf{0}$ da f, quindi f non è iniettiva.

$\dim(f(V)) \leq \dim(V)$

Proposizione

Sia $f : V \to W$ un'applicazione lineare, e sia $v_1, \ldots, v_n$ una base di V. Allora $f(v_1), \ldots, f(v_n)$ generano $f(V) = \operatorname{Im} f$. In particolare $\dim f(V) \leq \dim(V)$.

- Sia $w \in \operatorname{Im} f$.
- Per definizione di immagine, esiste $v \in V$ tale che $f(v) = w$.
- Posso scrivere $v = a_1 v_1 + \ldots + a_n v_n$.
- Quindi

$$\begin{aligned} w &= f(v) \\ &= f(a_1 v_1 + \ldots a_n v_n) \\ &= a_1 f(v_1) + \ldots + a_n f(v_n) \end{aligned}$$

f iniettiva $\implies \dim f(V) = \dim(V)$

Proposizione

Sia $f : V \to W$ un'applicazione lineare, e siano $v_1, \ldots, v_n$ vettori indipendenti di V. Se f è iniettiva, $f(v_1), \ldots, f(v_n)$ sono vettori indipendenti di W, quindi $\dim(\operatorname{Im} f) \geq \dim V$.

- Supponiamo che $a_1 f(v_1) + \ldots + a_n f(v_n) = \mathbf{0}$, dove gli a_i sono scalari.
- Basta dimostrare che $a_1 = \ldots = a_n = 0$.
- Per linearità, $f(a_1 v_1 + \ldots + a_n v_n) = \mathbf{0}$.
- Quindi $a_1 v_1 + \ldots + a_n v_n \in \ker(f) = \{\mathbf{0}\}$, ovvero $a_1 v_1 + \ldots + a_n v_n = \mathbf{0}$.
- Siccome i v_i sono indipendenti, $a_1 = \ldots = a_n = 0$.

Dimensione del nucleo e dell'immagine

Teorema

Sia $f : V \to W$ un'applicazione lineare. Allora
$\dim(\ker f) + \dim(\operatorname{Im} f) = \dim(V)$.

- Sia $n = \dim(V)$ e sia $k \leq n$ la dimensione di $\ker(f) \subseteq V$.
- Fisso una base $\{v_1, \dots, v_k\}$ di $\ker f$ e la completo ad una base $\{v_1, \dots, v_k, u_1, \dots, u_{n-k}\}$ di V (completamento).
- Gli $f(v_i)$ sono zero quindi non contribuiscono all'immagine:

$$\begin{aligned} \operatorname{Im} f &= \operatorname{span}\{f(v_1), \dots f(v_k), f(u_1), \dots, f(u_{n-k})\} \\ &= \operatorname{span}\{f(u_1), \dots, f(u_{n-k})\} \end{aligned}$$

- Devo mostrare $\dim(\operatorname{Im} f) = n - k$; basta far vedere che gli $f(u_i)$ sono indipendenti.
- Sia $a_1 f(u_1) + \ldots + a_{n-k} f(u_{n-k}) = \mathbf{0}$; devo mostrare che $a_1 = \ldots = a_{n-k} = 0$.
- Per linearità $f(a_1 u_1 + \ldots + a_{n-k} u_{n-k}) = \mathbf{0}$; quindi $a_1 u_1 + \ldots + a_{n-k} u_{n-k} \in \ker f = \operatorname{span}(v_1, \dots, v_k)$.
- Quindi esistono degli scalari $b_1, \dots, b_k$ con $a_1 u_1 + \ldots + a_{n-k} u_{n-k} + b_1 v_1 + \ldots + b_k v_k = \mathbf{0}$.
- Ma gli u_i, v_j sono indipendenti, quindi gli a_i, b_j sono zero.

Sezioni

Autovalori e autovettori di un'endomorfismo lineare

Definizione

Un'endomorfismo è un'applicazione lineare $f : V \to V$ da uno spazio vettoriale V a se stesso.

- *Diciamo che $v \in V$ è un autovettore dell'endomorfismo $f : V \to V$ se $v \neq \mathbf{0}$ ed esiste uno scalare $\lambda \in \mathbb{K}$ tale che $f(v) = \lambda v$.*
- *Lo scalare λ viene chiamato autovalore di f relativo all'autovettore v.*
- *In generale se λ è uno scalare, l'insieme $V_\lambda = \{v \in V \mid f(v) = \lambda v\}$ è un sottospazio di V (esercizio) chiamato autospazio relativo a λ.*
- *Quindi λ è un autovalore di f (relativo a qualche autovettore) se e solo se $V_\lambda \neq \{\mathbf{0}\}$.*

Autovalori e autovettori di una matrice

- Gli autovalori e autovettori di una matrice quadrata $n \times n$ A sono quelli relativi all'applicazione lineare $L_A : \mathbb{K}^n \to \mathbb{K}^n$ associata ad A.
- Quindi $v \in \mathbb{K}^n$ è un autovettore di A se $v \neq \mathbf{0}$ ed esiste un $\lambda \in \mathbb{K}$ (detto autovalore di A relativo a v) tale che $Av = \lambda v$.
- L'autospazio V_λ è l'insieme dei vettori $v \in \mathbb{K}^n$ tali che $Av = \lambda v$.

Autovalori di una matrice diagonale

Osservazione

Una matrice diagonale $A \in M_{n\times n}(\mathbb{K})$ ha come autovalori gli elementi della base standard di $\mathbb{K}^n$, e come autovalori gli elementi sulla diagonale.

Ad esempio sia $A = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$. I vettori della base standard

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

sono autovettori di A con autovalori 3 e 4 rispettivamente:

$$\begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} = -1 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Una rotazione di 90 gradi non ha autovettori reali

- Sia $f : \mathbb{R}^2 \to \mathbb{R}^2$ l'applicazione lineare data da una rotazione di 90 gradi ($= \pi/2$ radianti) in senso antiorario intorno all'origine.
- Abbiamo $f\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}0\\1\end{bmatrix}$ e $f\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}-1\\0\end{bmatrix}$
- Rispetto alla base standard f è rappresentata dalla matrice $\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}$ (le colonne sono le immagini dei vettori della base).
- Gli autovettori di f sono quegli $\begin{bmatrix}x\\y\end{bmatrix}$ non nulli tali che per qualche $\lambda \in \mathbb{R}$ si ha $\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \lambda\begin{bmatrix}x\\y\end{bmatrix}$,
- L'equazione equivale a $-y = \lambda x$ e $x = \lambda y$
- $\implies x = -\lambda^2 x$ e $y = -\lambda^2 y$.
- Uno tra x ed y deve essere diverso da zero. Se $x \neq 0$ ottengo $1 = -\lambda^2$ ovvero $\lambda = \pm\sqrt{-1}$, un numero complesso!
- Non vi sono soluzioni reali, quindi $f : \mathbb{R}^2 \to \mathbb{R}^2$ non ha autovettori. Si vedeva anche geometricamente: ogni vettore viene ruotato, mentre gli autovettori non vengono ruotati.

Rotazione di θ radianti

- Sia $f : \mathbb{R}^2 \to \mathbb{R}^2$ l'applicazione lineare definita da una rotazione in senso antiorario del piano $\mathbb{R}^2$ di un angolo θ.
- Rispetto alla base standard la matrice di f è $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.
- Per quali valori di θ la f ha degli autovettori?
- Verificate che gli unici angoli possibili sono $\theta = 0$ o $\theta = \pi$ (possiamo aggiungere a θ multipli di 2π).

Rotazione intorno all'asse z

- Sia $f : \mathbb{R}^3 \to \mathbb{R}^3$ una rotazione di θ radianti in senso antiorario intorno all'asse verticale z (l'asse verticale è lo span di $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$).
- Rispetto alla base standard di $\mathbb{R}^3$, la matrice di f è
$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
- Ad esempio se $\theta = \pi/2$, la matrice è
$$\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
- Per qualsiasi θ, i vettori della forma $v = \begin{bmatrix} 0 \\ 0 \\ a \end{bmatrix}$ con $a \neq 0$ in $\mathbb{R}$ sono autovettori con autovalore 1, ovvero $f(v) = v$.
- Per $\theta = \pi/2$ non vi sono altri autovettori.
- In particolare non è possibile trovare una base di $\mathbb{R}^3$ fatta di autovettori di f.

Matrici diagonalizzabili

Definizione

Una matrice quadrata $A = [a_{ij}]$ si dice diagonale se gli elementi fuori dalla diagonale sono tutti zero, ovvero $a_{ij} = 0$ per $i \neq j$.
Una matrice quadrata si dice diagonalizzabile se è coniugata ad una matrice diagonale, ovvero se esiste una matrice invertibile P, detta matrice diagonalizzante di A, tale che $P^{-1}AP$ è diagonale.

Esempio

La matrice $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ è diagonalizzabile. Una matrice diagonalizzante è $P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Infatti $P^{-1}AP = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}$

Chiaramente se A è già diagonale, è diagonalizzabile con $P = \mathrm{I}$.

Esiste base di autovettori $\implies$ diagonalizzabile

Teorema

Supponiamo che una matrice $A \in M_{n\times n}(\mathbb{K})$ abbia degli autovettori $v_1, \dots, v_n \in \mathbb{K}^n$ che formano una base di $\mathbb{K}^n$ e sia $P = [v_1, \dots, v_n]$ la matrice che ha come i-esima colonna le coordinate di v_i. Allora P diagonalizza M.

- Se λ_i l'autovalore di v_i, allora $Av_i = \lambda_i v$ e abbiamo

$$\begin{aligned} AP &= A[v_1, \dots, v_n] \\ &= [Av_1, \dots, Av_n] \\ &= [\lambda_1 v_1, \dots, \lambda_n v_n] \\ &= [v_1, \dots, v_n]\operatorname{diag}(\lambda_1, \dots, \lambda_n) \\ &= PD \end{aligned}$$

  dove $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ è la matrice diagonale i cui elementi sulla diagonale sono $\lambda_1, \dots, \lambda_n$. esempio

- Siccome $v_1, \dots, v_n$ sono indipendenti, $P = [v_1, \dots, v_n]$ è invertibile e possiamo riscrivere $AP = PD$ come $P^{-1}AP = D$.

Il seguente esempio illustra la formula

$$[v_1, \dots, v_n] \operatorname{diag}(\lambda_1, \dots, \lambda_n) = [\lambda_1 v_1, \dots, \lambda_n v_n]$$

Esempio

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 a & \lambda_2 b \\ \lambda_1 c & \lambda_2 d \end{bmatrix}$$

torna a diagonalizzabile

Diagonalizzabile $\Longrightarrow$ esiste base di autovettori

Proposizione

Se $A \in M_{n\times n}(\mathbb{K})$ è diagonalizzabile, allora esiste una base di $\mathbb{K}^n$ composta da autovettori di A

- Se $P^{-1}AP = D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, i vettori $e_1, \dots, e_n$ della base standard sono autovettori di D.
- Dico che i vettori $v_1 = Pe_1, \dots, v_n = Pe_n$ sono autovettori di A e formano una base di $\mathbb{K}^n$.
- Verifico che sono autovettori:

$$\begin{aligned} Av_i &= PDP^{-1}v_i \\ &= PDP^{-1}Pe_i \\ &= PDe_i \\ &= P\lambda_i e_i \\ &= \lambda_i Pe_i \\ &= \lambda_i v_i \end{aligned}$$

- Verifico che sono una base: questo segue dal fatto che un'applicazione invertibile P manda una base in una base.

Sezioni

Somma di applicazioni lineari

Definizione

Date due applicazioni lineari $f : V \to W$ e $g : V \to W$, possiamo definire la loro somma $f + g : V \to W$ come quell'applicazione che dato un input v fornisce l'output $f(v) + g(v)$

$$(f + g)(v) = f(v) + g(v)$$

Fissate delle basi di V e W, se f è rappresentata dalla matrice A e g è rappresentata dalla matrice B, allora $f + g$ è rappresentata dalla somma $A + B$ delle due matrici.

Definizione

Dato uno spazio vettoriale V, sia $I : V \to I$ l'applicazione lineare definita da $I(v) = v$ per ogni $v \in V$.

Se fissiamo una base $\mathcal{B}$ di V, la matrice di I rispetto alla base $\mathcal{B}$ in partenza e in arrivo è la matrice identità.

Autospazio come kernel

Proposizione

Dato un endomorfismo $f : V \to V$, *un vettore* $v \neq 0$ *è un autovettore di* f *associato all'autovalore* $\lambda \in \mathbb{K}$ *se e solo se* $v \in \ker(f - \lambda I)$.

Dimostrazione

$$
\begin{aligned}
& & f(v) &= \lambda v \\
\iff & & f(v) - \lambda v &= \mathbf{0} \\
\iff & & (f - \lambda \mathrm{I})(v) &= \mathbf{0} \\
\iff & & v \in \ker(f - \lambda \mathrm{I}) &
\end{aligned}
$$

dove I è l'applicazione lineare definita da $\mathrm{I}(v) = v$ ed $f - \lambda \mathrm{I}$ è l'applicazione definita da $(f - \lambda \mathrm{I})(v) = f(v) - \lambda \mathrm{I}(v) = f(v) - \lambda v$.

Polinomio caratteristico di una matrice

Data una matrice $n \times n$ A, il suo polinomio caratteristico $p_A(x)$ è un polinomio di grado n definito come

$$p_A(x) = \det(A - x\mathrm{I})$$

dove I è la matrice identità $n \times n$. Ad esempio se $A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$, allora

$$\begin{aligned} p_A(x) &= \det\left(\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} - x \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} 2-x & 3 \\ 4 & 5-x \end{bmatrix}\right) \\ &= (2-x)(5-5) - 12 \\ &= x^2 - 7x - 2 \end{aligned}$$

Esercizio: perché viene sempre un polinomio di grado n? rcjh

Commento

Il polinomio caratteristico $p_A(x)$ è talvolta definito come $\det(x\mathrm{I} - A)$ invece che come $\det(A - x\mathrm{I})$.
Le due definizioni differiscono solo per il segno:

$$\det(x\mathrm{I} - A) = \pm \det(A - x\mathrm{I}).$$

Più precisamente, data una matrice $n \times n$ A,

$$\begin{array}{rcl} n \text{ pari} & \Longrightarrow & \det(x\mathrm{I} - A) = \det(A - x\mathrm{I}) \\ n \text{ dispari} & \Longrightarrow & \det(x\mathrm{I} - A) = -\det(A - x\mathrm{I}) \end{array}$$

Visto che a noi interesseranno solo le radici del polinomio caratteristico, possiamo usare indifferentemente l'una o l'altra definizione perché le radici sono le stesse.
Con la definizione $\det(x\mathrm{I} - A)$ viene un polinomio monico (il coefficiente del termine di grado più alto è 1), mentre con l'altra definizione il coefficiente direttivo è ± 1.

Traccia di una matrice

Definizione

La traccia $\operatorname{tr}(A)$ *di una matrice quadrata è la somma degli elementi sulla sua diagonale.*

Nonostante il prodotto di matrici non commuti, si ha:

Teorema

Se A, B *sono matrici* $n \times n$,

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

Polinomio caratteristico, traccia e determinante

Proposizione

Se A è una matrice 2×2, allora

$$p_A(x) = x^2 - \mathrm{tr}(A)x + \det(A).$$

Dimostrazione

Se $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, abbiamo

$$\begin{aligned} p_A(x) &= \det \begin{pmatrix} a - x & b \\ c & d - x \end{pmatrix} \\ &= (a - x)(d - x) - bc \\ &= x^2 - (a + d)x + ad - bc \\ &= x^2 - \mathrm{tr}(A)x + \det(A) \end{aligned}$$

Polinomio caratteristico, traccia e determinante

Proposizione

Se A è una matrice 3×3, allora

$$p_A(x) = -x^3 + \mathrm{tr}(A)x^2 + cx + \det(A)$$

per qualche scalare c.

Caso $n \times n$

Proposizione

Con una dimostrazione simile si dimostra che il il polinomio caratteristico $p_A(x)$ di una matrice $n \times n$ ha la forma

$$p_A(x) = (-1)^n x^n + (-1)^{n-1}\operatorname{tr}(A)x^{n-1} + \ldots + \det(A)$$

ovvero il termine noto è $\det(A)$ *(che è anche uguale a $p_A(0)$) e il coefficiene di x^{n-1} è* $\pm\operatorname{tr}(A)$.
Nel caso di una matrice 3×3 otteniamo

$$p_A(x) = \det(A - xI) = -x^3 + \operatorname{tr}(A)x^2 + cx + \det(A)$$

per qualche scalare c.

Matrici simili o coniugate

Definizione

Due matrici quadrate $n \times n$ A e B si dicono coniugate o simili se esiste una matrice invertibile M tale che $M^{-1}AM = B$.

Il fatto che l'inversa di M sia a sinistra invece che a destra è irrilevante, perché possiamo riscrivere $M^{-1}AM = B$ come $NAN^{-1} = B$ con $N = M^{-1}$.

Proposizione

La similitudine tra matrici è una relazione di equivalenza.

- Se A è simile a B, B è simile ad A. Infatti da $M^{-1}AM = B$, moltiplicando a sinistra per M e a destra per M^{-1}, otteniamo $A = MBM^{-1} = N^{-1}BN$ con $N = M^{-1}$.
- Una matrice è simile a se stessa: $\mathrm{I}^{-1}A\mathrm{I} = \mathrm{I}A\mathrm{I} = A$.
- Se A è simile a B e B è simile a C, allora A è simile a C. Infatti se $M^{-1}AM = B$ e $N^{-1}BN = C$, allora sostituendo B com $M^{-1}AM$ nell'equazione per C, ottengo $C = N^{-1}M^{-1}AMN = (MN)^{-1}A(MN)$.

Vogliamo dimostrare che matrici simili hanno lo stesso polinomio caratteristico. Occorre richiamare alcuni fatti:

- La matrice $x\mathrm{I}$ commuta con tutte le matrici: $(x\mathrm{I})M = M(x\mathrm{I}) = xM$.
- Ne segue che $M^{-1}(x\mathrm{I})M = (x\mathrm{I})M^{-1}M = x\mathrm{I}$.
- Per la legge distributiva

$$\begin{aligned} M^{-1}(x\mathrm{I} - A)M &= (M^{-1}x\mathrm{I} - M^{-1}A)M \\ &= M^{-1}(x\mathrm{I})M - M^{-1}AM \\ &= x\mathrm{I} - M^{-1}AM \end{aligned}$$

Matrici simili hanno lo stesso polinomio caratteristico

Proposizione

Se $M^{-1}AM = B$, allora $p_A(x) = p_B(x)$.

$$\begin{aligned}
p_B(x) &= \det(xI - B) \\
&= \det\left(x\mathrm{I} - M^{-1}AM\right) \\
&= \det(M^{-1}(x\mathrm{I} - A)M) \\
&= \det(M^{-1})\det(x\mathrm{I} - A)\det(M) \\
&= \frac{1}{\det M}\det(x\mathrm{I} - A)\det(M) \\
&= \det(x\mathrm{I} - A) \\
&= p_A(x)
\end{aligned}$$

Polinomio caratteristico di un endomorfismo

Definizione

Dato uno spazio vettoriale V su $\mathbb{K}$ di dimensione n ed un endomorfismo $f : V \to V$, fissiamo una base $\mathcal{B} = \{v_1, \dots, v_n\}$ e scriviamo la matrice $A = [f]_{\mathcal{B}}^{\mathcal{B}}$ di f rispetto alla base $\mathcal{B}$. Il polinomio caratteristico $p_f(x)$ di f è definito come

$$p_f(x) = \det(A - xI).$$

Se scegliamo un'altra base $\mathcal{C}$, la matrice $B = [f]_{\mathcal{C}}^{\mathcal{C}}$ viene coniugata alla matrice $A = [f]_{\mathcal{B}}^{\mathcal{B}}$, e quindi il polinomio caratteristico non cambia:

$$p_f(x) = p_A(x) = p_B(x)$$

Base di autovettori $\implies$ diagonalizzabile

Corollario

Supponiamo che un'endomorfismo $f : V \to V$ *abbia degli autovettori* $v_1, \dots, v_n$ *che formano una base* $\mathcal{B}$ *di* V *e sia*