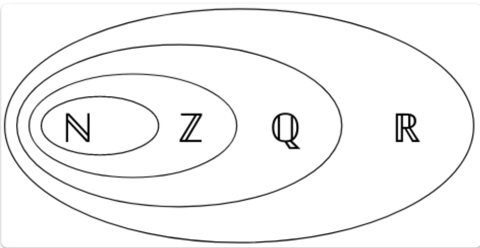
Lezione03

Table of contents

Generalizzazioni algebriche

Vediamo i concetti algebrici fondamentali per apprendere nelle lezioni successive, la matematica dietro i principali algoritmi dell'AI.



L'insieme di numeri verrà rappresentato con la seguente notazione:

- $\mathbb{N}$ l'insieme dei numeri naturali incluso lo 0

$$0, 1, 2, 3, \ldots$$

- $\mathbb{Z}$ l'insieme dei numeri interi, costruiti partendo dai numeri naturali

$$-1, 0, 1, \ldots$$

- $\mathbb{Q}$ l'insieme dei numeri razionali

$$1/2, 5/4, 0.5, \ldots$$

- $\mathbb{R}$ l'insieme dei numeri reali, che comprende tutti gli insiemi visti fino adesso aggiungendo i numeri irrazionali e trascendentali quali

$$\sqrt{2}, \pi, e$$

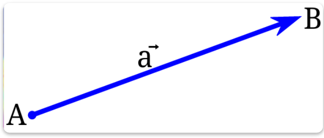
- $\mathbb{C}$ l'insieme dei numeri complessi

$$1 + 3i$$

con $3i$ la parte immaginaria

Con il simbolo $_+$, indicheremo solo i numeri positivi contenuti nell'insieme.

Vettori reali



Fissato $n \in \mathbb{N}_+$, un elemento $\mathbf{x}$ dell'insieme $\mathbb{R}^n$ (spazio normato) viene detto **vettore reale** con la seguente nomenclatura

$$\mathbf{x} \in \mathbb{R}^n \qquad \mathbf{x} = (x_1, x_2, \ldots, x_n) = (x_i)_i^n$$

Il vettore di elementi trattando matrici diventa vettore colonna: il fatto che sia con le virgole in orizzontale è convenzione.

$$(x_1, x_2, \cdots, x_n) := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

I vettori possono essere a loro volta sommati e moltiplicati per vettore o *scalare*:

- elementi moltiplicati $\mathbb{R}^n$, creano elementi contenuti in $\mathbb{R}^n$;
- elementi sommati, anche loro daranno prodotto in $\mathbb{R}^n$

Lunghezza di un vettore: norma euclidea

Compreso il fatto che $\mathbb{R}^n$ sia uno *spazio vettoriale*, per definire la sua lunghezza di un vettore $\mathbf{x}$ usiamo la **norma euclidea**:

$$||\mathbf{x}|| \triangleq \sqrt{\sum_{i=1}^{n} x_i^2}$$

Dimensione vs Lunghezza

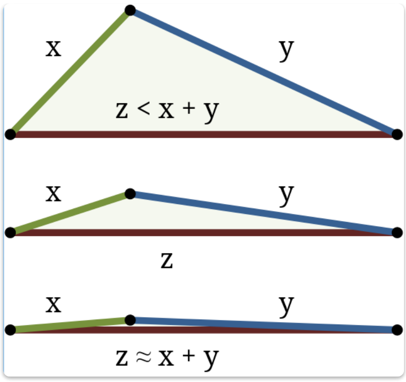
Nella notazione matematica precisa, si dice *dimensione* dello spazio, il numero di componenti del vettore. Da quindi non confondere con la lunghezza, calcolata con il metodo sopra.

Delle radici, usiamo solo la parte positiva, per ogni $\mathbf{x} \in \mathbb{R}^n$ vale quindi:

$$||\mathbf{x}|| \geq 0 \qquad ||\mathbf{x}|| = 0 \iff \mathbf{x} = 0$$

Se fissata la lunghezza a 1, il vettore viene detto *unitario*(o **versore**): ci interessa soltanto la sua direzione in questo caso.

Disuguaglianza triangolare



$$||\mathbf{x}+\mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$$

Presi 2 vettori in $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, vale la disuguaglianza sopra.
La norma euclidea fa valere la disuguaglianza triangolare.

Caso $n = 2$ per $\mathbb{R}^n$

$$\mathbf{x} = (x_1, 0) \qquad \mathbf{y} = (0, y_2)$$

allora

$$||\mathbf{x}|| + ||\mathbf{y}|| = ||x_1|| + ||y_2|| \geq \sqrt{x_1^2 + y_2^2} = ||\mathbf{x}+\mathbf{y}||$$

Distanza tra vettori: norma della differenza

La distanza tra due vettori è calcolata con la norma della differenza:

$$d(\mathbf{x}, \mathbf{y}) \triangleq ||\mathbf{x}-\mathbf{y}|| = ||\mathbf{y}-\mathbf{x}||$$

Ragionando sulla norma euclidea, siccome è somma di quadrati, cambiare il segno non porta al cambiamento di segno del risultato, portando alla stessa quantità.

Prodotto scalare

Sempre per due vettori $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, si dice prodotto scalare:

$$\mathbf{x} \cdot \mathbf{y} \triangleq \sum_{i=1}^{n} x_i y_i$$

Il numero reale generato non è quindi per forza positivo.
Per ogni $\mathbf{x} \in \mathbb{R}^n$ (spazio con prodotto interno), vale:

$$||\mathbf{x}||^2 = \mathbf{x} \cdot \mathbf{x} \qquad 0 \cdot \mathbf{x} = \mathbf{x} \cdot 0 = 0$$

Disuguaglianza di Cauchy-Schwartz

$$|\mathbf{x} \cdot \mathbf{y}| \leq ||\mathbf{x}|| \cdot ||\mathbf{y}||$$

Ci servirà per approssimare i punti nel *training set* della nostra rete neurale; presi 2 vettori $\mathbf{x}, \mathbf{y} \in \mathbb{R}_n$, ci accorgiamo che il valore assoluto del prodotto dei 2, è minore/uguale al prodotto delle due lunghezze. L'uguaglianza vale solamente quando esiste coefficiente $\lambda \in \mathbb{R}$ tale che $\mathbf{x} = \lambda \mathbf{y}$.

Per una dimensione di grandezza 2, calcoliamo

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2$$

ruotiamo di un angolo $\beta \in \mathbb{R}$

$$\mathbf{x}' = R_\beta \mathbf{x} \qquad \mathbf{y}' = R_\beta \mathbf{y}$$

dove

$$R_\beta = \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}$$

Ottenendo

$$\mathbf{x}' \cdot \mathbf{y}' = \mathbf{x} \cdot \mathbf{y}$$

Ci permette di muovere uno dei 2 vettori, per esempio, sull'asse delle ascisse e usando l'angolo tra i due vettori $\theta \in \mathbb{R}^n$, calcolare il vettore $\mathbf{y}$ con:

$$|\mathbf{x} \cdot \mathbf{y}| = |||\mathbf{x}|| \cdot ||\mathbf{y}|| \cos\theta| \leq ||\mathbf{x}|| \cdot ||\mathbf{y}||$$

Matrice ortogonale

Si dice ortogonale una matrice di ordine $n \in \mathbb{N}_+$ se e soltanto se è invertibile e vale:

$$R^T = R^{-1}$$

Una matrice ortogonale rappresenta isometria (preservata la lunghezza dei vettori), per ogni $\mathbf{x} \in \mathbb{R}^n$, che vale:

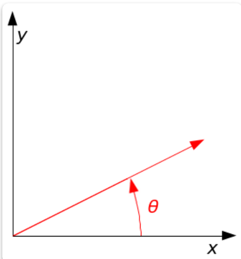
$$||R\mathbf{x}||^2 = ||\mathbf{x}||^2$$

Siccome il determinante può valere solo ± 1 per una matrice ortogonale:

$$1 = (\det R)^2$$

La rotazione non fa mai cambiare il verso dei nostri vettori; prendiamo in esame soltanto il caso in qui il determinante è uguale a 1.

Matrice di rotazione



Per le matrici ortogonali, e non solo, si dice matrice di rotazione una matrice R per ogni $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ che soddisfa:

$$R\mathbf{x} \cdot R\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$$

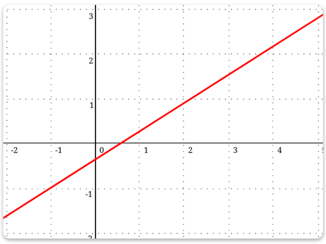
Vettori ortogonali & Dipendenza lineare

Dato $n \in \mathbb{N}_+$, due vettori $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ entrambi diversi da 0, si dicono ortogonali se e soltanto se:

$$\mathbf{x} \cdot \mathbf{y} = 0$$

Chiamiamo la "nuvola" di punti, un cluster di punti che se uniti formano vettori *mutualmente ortogonali*, indicati con la notazione $O \subset \mathbb{R}^n$. Se i punti soddisfano la proprietà allora l'insieme di vettori si dicono linearmente indipendenti.

Rette

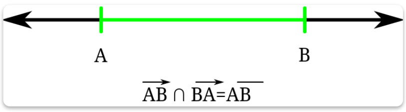


Dati due vettori $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, la retta $\overline{(\mathbf{x}, \mathbf{y})}$ passante per questi:

$$\overline{(\mathbf{x}, \mathbf{y})} = \{\mathbf{v} \in \mathbb{R}^n : \mathbf{v} = \mathbf{x} + \lambda(\mathbf{y} - \mathbf{x}) \wedge \lambda \in \mathbb{R}\}$$

Una retta è un oggetto monodimensionale senza profondità o lunghezza o curvatura, che si estende all'infinito. Si può presentare in dimensioni di grandezza 1, 2 o più.

Segmenti

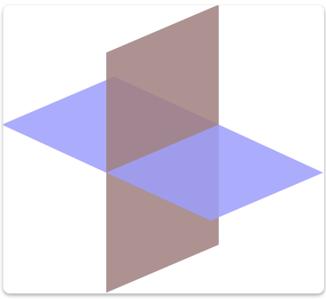


Dati due vettori $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, il segmento $\overline{[\mathbf{x}, \mathbf{y}]}$ che congiunge i due:

$$\overline{[\mathbf{x}, \mathbf{y}]} = \{\mathbf{v} \in \mathbb{R}^n : \mathbf{v} = \mathbf{x} + \lambda(\mathbf{y} - \mathbf{x}) \wedge \lambda \in [0, 1]\}$$

Un segmento viene inteso come l'intersezione di tutti i punti da destra di A in poi, con tutti i punti inclusi da sinistra di B in poi.

Iperpiani

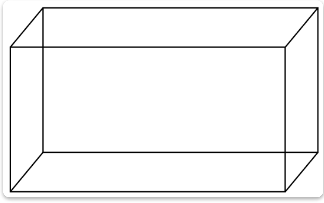


Dato un vettore $\mathbf{x} \in \mathbb{R}^n$ e un versore $n \in \mathbb{R}^n$, l'iperpiano che contiene $\mathbf{x}$ e ha n come versore normale:

$$\mathcal{H}_n(\mathbf{x}) = \{\mathbf{v} \in \mathbb{R}^n : (\mathbf{v} - \mathbf{x}) \cdot n = 0\}$$

L'iperpiano sarebbe il sottospazio con dimensione di grado -1 rispetto all'ambiente preso in considerazione. Per esempio: un piano di dimensione 3 ha come iperpiano uno costruito da due piani, ovvero uno spazio bidimensionale (vedi immagine sopra).

Scatole



Dati due vettori $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, si dice **scatola** (*aperta*):

$$(\mathbf{x}, \mathbf{y}) = \{\mathbf{v} \in \mathbb{R}^n : \underline{v}_1 < v_1 < \overline{v}_2 \wedge \underline{v}_2 < v_2 < \overline{v}_2 \wedge \cdots \wedge \underline{v}_n < v_n < \overline{v}_n\}$$

Dove per ogni $1 \leq i \leq n$,

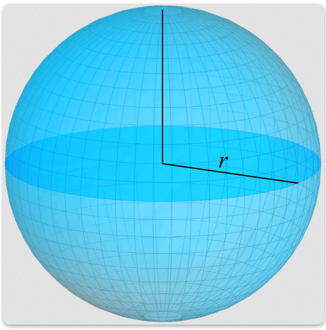
$$\underline{v}_i = \mathtt{min}\{x_i, y_i\} \qquad \text{e} \qquad \overline{v}_i = \mathtt{max}\{x_i, y_i\}$$

Mentre una **scatola** *chiusa*:

$$(\mathbf{x}, \mathbf{y}) = \{\mathbf{v} \in \mathbb{R}^n : \underline{v}_1 \leq v_1 \leq \overline{v}_2 \wedge \underline{v}_2 \leq v_2 \leq \overline{v}_2 \wedge \cdots \wedge \underline{v}_n \leq v_n \leq \overline{v}_n\}$$

La differenza sta nel $<$ o $\leq$: una scatola i cui vettori limite vengono esclusi ($<$), si dice scatola aperta, una i cui vettori limite inclusi ($\leq$) si dice chiusa. Tutti i punti inclusi all'interno della scatola, rappresentano la scatola stessa.

Palle



Dato vettore $\mathbf{x} \in \mathbb{R}^n$ e uno scalare $r \in \mathbb{R}_+$, si dice **palla** *aperta* centrata in $\mathbf{x}$ e di raggio r:

$$\mathcal{B}_r(\mathbf{x}) = \{\mathbf{v} \in \mathbb{R}^n : ||\mathbf{v} - \mathbf{x}|| < r\}$$

Mentre si dice **palla** *chiusa*:

$$\mathcal{B}_r(\mathbf{x}) = \{\mathbf{v} \in \mathbb{R}^n : ||\mathbf{v} - \mathbf{x}|| \leq r\}$$

Anche quì vale la stessa logica della scatola: tutti i punti inclusi dalla palla, sono componenti della palla stessa, che la descrivono.

Insiemi aperti, chiusi, convessi, limitati e compatti

Viste le nozioni imparate in cima, riguardo gli insiemi, ne introduciamo altre per gli insiemi di numeri in $\mathbb{R}, \mathbb{R}^2, \mathbb{R}^3$:

- $L \subseteq R^n$ si dice *limitato* se esiste $r \in \mathbb{R}_+$ e un vettore $\mathbf{x} \in \mathbb{R}^n$ t.c. $L \subseteq \mathcal{B}_r(\mathbf{x})$
  Un insieme è limitato se riesco a costruire una palla che include tutto. Un esempio in cui non succede è quando dobbiamo generare palle sempre più grandi per contenere i vettori che sforano l'insieme.
- $A \subseteq R^n$ si dice *aperto* se per ogni $\mathbf{x} \in A$ esiste $r \in \mathbb{R}_+$ t.c. $\mathcal{B}_r(\mathbf{x}) \subseteq A$
  Pensiamo agli intervalli di $\mathbb{R}$: quando diciamo che un intervallo è aperto? Quando non sono inclusi i bordi, i punti estremi, dicendo che qualsiasi punto interno lo possiamo usare come centro di un'altra palla di raggio r.
- $C \subseteq R^n$ si dice *chiuso* se esiste $A \subseteq \mathbb{R}^n$ aperto t.c. $A \cup C = \mathbb{R}^n$
  Come faccio a dire che un intervallo è chiuso? Perché ne esistono due aperti che uniti formano quello

chiuso (vedere esempio del segmento).

- $K \subseteq \mathbb{R}^n$ di dice *compatto* se è chiuso e limitato
- $V \subseteq \mathbb{R}^n$ si dice *convesso* se per ogni coppia di vettori $\mathbf{x}, \mathbf{y} \in V$ vale che $\overline{[\mathbf{x}, \mathbf{y}]} \subseteq V$
  Immaginiamo un insieme convesso (disegnandolo) e prendiamo 2 punti qualsiasi creando un segmento che li unisce: se il segmento è dentro l'insieme allora l'insieme è convesso, se non lo è allora altrimenti.

02/03/2023