RISOLUZIONE DI SISTEMI LINEARI: METODI DIRETTI (Richiami su matrici e vettori)

In questa lezione introduttiva al secondo nucleo, richiameremo alcuni concetti fondamentali su matrici e vettori, riprendendo aspetti essenziali dell'algebra lineare.

Un vettore $\boldsymbol{a}$ è un elemento dello spazio euclideo n dimensionale $\mathbb{R}^n$ (dove $\mathbb{R}$ è lo spazio dei numeri reali). Con questa notazione intendiamo un elemento composto da più componenti

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

Questa notazione, $\boldsymbol{a} \in \mathbb{R}^n$, rappresenta un vettore colonna, ovvero un vettore le cui componenti sono disposte verticalmente. Parleremo anche di vettori trasposti, indicati con $\boldsymbol{a}^T$ (vettore trasposto di $\boldsymbol{a}$), che rappresentano un vettore riga

$$\boldsymbol{a}^T = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix},$$

dove le componenti del vettore sono ordinate in senso orizzontale.

L'oggetto matematico A, denominato matrice ed indentificato con una lettera maiuscola,

$$A \in \mathbb{R}^{n\times n},$$

indicherà una tabella ordinata di n righe ed n colonne del tipo

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

dove le righe vengono identificate con il primo indice (detto indice di riga) e le colonne le indichiamo con il secondo indice (detto indice colonna). Quindi, ad esempio, a_{11} indica l'elemento della prima riga e della prima colonna, a_{21} seconda riga e prima colonna e a_{1n} prima riga ed n-essima colonna, e così via.
In maniera compatta scriveremo

$$A = (a_{ij}),$$

mettendo tra parentesi la componente generica della matrice A, dove i, j indicheranno rispettivamente l'indice riga (i) e l'indice colonna (j).
La trasposta di una matrice si ottiene scambiando le righe con le colonne: le righe diventano colonne e le colonne diventano righe

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ \vdots & & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix},$$

ed in maniera compatta

$$A^T = (a_{ij})^T = (a_{ji})$$

dove abbiamo messo in evidenza lo scambio di indici di riga e di colonna.
Nell'insieme dei numeri reali $\mathbb{R}$, un elemento fondamentale per determinare la lunghezza o la misura di un numero (scalare) è il suo modulo, ovvero il valore assoluto

$$a \in \mathbb{R}, \text{la sua misura è } |a|.$$

Vogliamo introdurre concetti analoghi per misurare vettori e matrici. Il concetto che rappresenta la lunghezza di un vettore è la cosiddetta norma del vettore. Un esempio ben noto di norma, tratto dalla geometria euclidea, è la norma euclidea. Consideriamo, ad esempio, un vettore $\boldsymbol{a} \in \mathbb{R}^3$ (ovvero, $n = 3$), quindi un vettore nello spazio tridimensionale, avente tre componenti

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}.$$

La somma dei quadrati delle componenti sotto radice quadrata definisce la cosiddetta lunghezza (o misura) del vettore

$$\|\boldsymbol{a}\|_2 = (a_1^2 + a_2^2 + a_3^2)^{\frac{1}{2}}.$$

Quest'ultima è anche detta norma due del vettore. Generalizzando per un vettore nello spazio n dimensionale $\boldsymbol{a} \in \mathbb{R}^n$, abbiamo

$$\|\boldsymbol{a}\|_2 = (a_1^2 + a_2^2 + \cdots + a_n^2)^{\frac{1}{2}}.$$

Quindi, possiamo associare alla norma due del vettore il modulo/lunghezza del vettore stesso

$$\|\boldsymbol{a}\|_2 \Longrightarrow |a|.$$

Questo non è l'unico modo per rappresentare la lunghezza di un vettore e quindi per misurarlo. Si può infatti definire la norma uno di un vettore, che corrisponde alla somma dei moduli delle sue componenti

$$\|\boldsymbol{a}\|_1 = |a_1| + |a_2| + \cdots + |a_n|.$$

Possiamo generalizzare i concetti di norma uno e norma due introducendo la norma p, dove p è un qualunque numero reale compreso tra 1 e infinito.

$$1 \le p \le \infty.$$

La norma p di un vettore è data dalla radice p-esima della somma dei moduli delle sue componenti elevati alla potenza p

$$\|\boldsymbol{a}\|_p = \left(\sum_{k=1}^{n} |a_k|^p\right)^{\frac{1}{p}}.$$

Possiamo ritrovare l'espressione della norma uno e della norma due, se utilizziamo in questa definizione $p = 1$ per la norma uno e $p = 2$ per la norma due (o norma euclidea).

Infine, possiamo introdurre un altro tipo di norma, chiamata norma infinito, che è definita come il massimo tra i valori assoluti delle componenti del vettore

$$\|\boldsymbol{a}\|_\infty = \max_{1 \le k \le n} |a_k|,$$

ovvero il massimo, preso su tutte le componenti dalla prima alla n-esima.

Abbiamo quindi diversi esempi di norma, ciascuno dei quali definisce in modo distinto la lunghezza del vettore $\boldsymbol{a}$. Verifichiamo con un esempio che calcolando la lunghezza del vettore $\boldsymbol{a}$ di componenti $1, 2, -3$

$$a = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix},$$

quest'ultimo abbia norme differenti. La norma due è la somma delle componenti al quadrato in radice quadrata

$$\|\boldsymbol{a}\|_2 = (1 + 4 + 9)^{\frac{1}{2}} = \sqrt{14} \simeq 3.74,$$

la norma uno è la somma dei valori assoluti delle componenti

$$\|\boldsymbol{a}\|_1 = |1| + |2| + |-3| = 6,$$

e la norma infinito è il massimo dei valori assoluti delle componenti

$$\|\boldsymbol{a}\|_\infty = \max(|1|, |2|, |-3|) = 3.$$

Introduciamo ora un altro concetto importante: il prodotto scalare euclideo. Si tratta di una relazione che mette in rapporto due vettori e produce un numero reale. In particolare, se $\boldsymbol{a}$ e $\boldsymbol{b}$ sono due vettori di $\mathbb{R}^n$, il loro prodotto scalare sarà indicato con $(\boldsymbol{a}, \boldsymbol{b})$

$$\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n \qquad (\boldsymbol{a}, \boldsymbol{b}) = \sum_{k=1}^{n} a_k\, b_k = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \boldsymbol{a}^T \boldsymbol{b},$$

ovvero, il prodotto scalare corrisponde alla somma dei prodotti delle componenti corrispondenti dei due vettori. In termini vettoriali, il prodotto scalare può essere espresso anche come il prodotto tra il vettore trasposto di $\boldsymbol{a}$ e il vettore $\boldsymbol{b}$, cioè come il prodotto di un vettore riga per un vettore colonna (detto anche prodotto riga per colonna). Facciamo un esempio: consideriamo il vettore $\boldsymbol{a}$ (di componenti $1, 3$) e il vettore $\boldsymbol{b}$ (di componenti $2, -2$).

$$\boldsymbol{a} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 2 \\ -2 \end{bmatrix} \qquad (\boldsymbol{a}, \boldsymbol{b}) = 1 \cdot 2 + 3 \cdot -2 = -4.$$

Il prodotto scalare $(\boldsymbol{a}, \boldsymbol{b})$ è quindi la prima componente a_1 per la prima componente b_1, più la seconda componente a_2 per la seconda componente b_2, e quindi risulta -4.

Notiamo che la norma due di un vettore, ovvero la norma euclidea di un vettore, non è altro che il prodotto scalare del vettore $\boldsymbol{a}$ con sé stesso, preso sotto la radice quadrata

$$\|\boldsymbol{a}\|_2 = (\boldsymbol{a}, \boldsymbol{a})^{\frac{1}{2}} = \left(\sum_{k=1}^{n} |a_k|^2 \right)^{\frac{1}{2}}.$$

Pertanto, la norma euclidea si definisce come la norma indotta dal prodotto scalare euclideo, nel senso che il prodotto scalare di un vettore con sé stesso fornisce esattamente il quadrato della norma del vettore.

Abbiamo definito il concetto di norma di vettore, e ora vogliamo definire analogamente il concetto di norma di matrice. In altre parole, desideriamo introdurre uno strumento che ci permetta di misurare le matrici in qualche modo.

Data una matrice $A \in \mathbb{R}^{n \times n}$, ovvero una matrice di n righe e n colonne (una matrice quadrata), indichiamo con $\|\boldsymbol{a}\|$ la norma di un vettore (quindi una norma che si può definire per ogni vettore $\boldsymbol{a} \in \mathbb{R}^n$). In corrispondenza della norma di un vettore introduciamo la norma della matrice. Più precisamente, definiamo norma della matrice A il massimo valore assunto dalla norma di $A\boldsymbol{a}$, dove $\boldsymbol{a}$ è un vettore di $\mathbb{R}^n$, la cui norma si uguale a uno

$$\|A\| = \max_{\|\boldsymbol{a}\|=1} ||A\boldsymbol{a}||.$$

È importante comprendere che stiamo introducendo due oggetti diversi: $A\boldsymbol{a}$ è un vettore, per cui ha senso definire la sua norma $||A\boldsymbol{a}||$. Al variare di $\boldsymbol{a}$, otteniamo a priori infiniti valori della norma $||A\boldsymbol{a}||$, quindi infiniti numeri reali. Tra questi numeri reali, prendiamo il massimo. Questo valore massimo è un numero reale che chiameremo norma di A.

In questo modo, abbiamo trovato un modo per dedurre una norma di matrice a partire da una norma di vettore. Questa norma sarà chiamata norma indotta dalla norma vettoriale o norma naturale. Le norme naturali di matrici non sono tutte le possibili norme, ma giocheranno un ruolo molto importante nelle nostre analisi. In particolare, per le norme di vettore che abbiamo introdotto, come ad esempio la norma

p-esima, possiamo applicare questa definizione e ottenere una norma di matrice che chiameremo norma p-esima

$$\|A\|_p = \max_{\|\boldsymbol{a}\|_p=1} \|A\boldsymbol{a}\|_p\,, \qquad 1 \le p \le \infty.$$

Abbiamo quindi trovato un modo per generare una famiglia di norme di matrici: le norme p della matrice A che abbiamo definito norme indotte o norme naturali.
Si può verificare che la norma 1 della matrice A, ovvero la norma associata alla norma 1 vettoriale, non è altro che il massimo, preso sulle colonne, delle somme dei valori assoluti degli elementi della matrice. In altre parole, per ogni colonna di A, calcoliamo la somma dei valori assoluti degli elementi della colonna, e poi prendiamo il massimo di queste somme.

$$\|A\|_1 = \max_k \sum_{j=1}^{n} |a_{jk}|.$$

Quindi fissiamo la colonna k-esima della matrice e andiamo a fare la somma di tutti gli elementi, presi in valore assoluto, presenti su questa colonna. Troviamo tanti numeri reali quante sono le colonne. Prendiamo il massimo su questi numeri reali. Questo massimo è esattamente la norma 1 dalla matrice A.
Ripetendo lo stesso ragionamento, ma scambiando fra loro il ruolo delle righe e delle colonne, troveremo la norma infinito

$$\|A\|_\infty = \max_j |\sum_{k=1}^{n} |a_{kj}|.$$

Dunque, la norma infinito è calcolata prima facendo le somme sulle colonne e poi prendendo il massimo sulle righe.
Vediamo un esempio su una matrice molto semplice ($A \in \mathbb{R}^{2\times 2}$):

$$A = \begin{bmatrix} 1 & -3 \\ 5 & 4 \end{bmatrix},$$
$$\|A\|_1 = \max\{|1| + |5|, |-3| + |4|\} = \max\{6{,}7\} = 7,$$
$$\left\||A|\right\|_\infty = \max\{|1| + |-3|, |5| + |4|\} \max\{4{,}9\} = 9.$$

La norma uno ci porta, prima, a fare la somma degli elementi sulla prima colonna $1 + 5 = 6$ e sulla seconda colonna $3 + 4 = 7$ (dove stiamo prendendo il valore assoluto delle componenti) e poi troviamo il massimo $\max(6, 7) = 7$, ricavando la norma uno di A. Per la norma infinito facciamo il processo inverso, ovvero prima la somma sulle righe $1 + 3 = 4$ e $5 + 4 = 9$ poi il massimo $\max(4, 9) = 9$.

Introdurremo adesso il concetto di sfera unitaria e richiameremo alcune proprietà delle norme vettoriali, del prodotto scalare e delle norme matriciali.

Iniziamo con l'introdurre il concetto di sfera unitaria, definito come l'insieme dei vettori di $\mathbb{R}^n$, la cui norma p-esima è uguale ad uno

$$S_p = \left\{\boldsymbol{a} \in \mathbb{R}^n : \|\boldsymbol{a}\|_p = 1\right\}.$$

In altre parole, stiamo cercando di caratterizzare tutti i vettori di $\mathbb{R}^n$ che hanno la stessa norma, e in particolare quelli con norma/misura uguale a uno. Ad esempio, nel caso di $\mathbb{R}^2$, consideriamo vettori in due dimensioni con due componenti. I vertici dei vettori che hanno norma uno (dove la norma uno è la somma dei moduli delle componenti) formano un rombo. Questo significa che ogni vettore che ha origine nel punto di coordinate (0,0) e vertice su uno dei lati del rombo ha norma uguale a 1. In modo simile, possiamo considerare i vettori che hanno norma due uguale a 1, ovvero norma euclidea uguale a 1. Come ci si aspetta, questi vettori formano un cerchio unitario centrato nell'origine. Analogamente, possiamo caratterizzare l'insieme dei punti nel piano $\mathbb{R}^2$ che sono vertici di vettori con norma infinito uguale a 1, ovvero quelli per cui il massimo tra i valori assoluti delle componenti è uguale a uno. In altri termini, si tratta dell'insieme di vettori che hanno una delle componenti (o entrambe) uguale a 1. In questo caso, la sfera unitaria diventa un quadrato.

Abbiamo definito la norma infinito nel modo visto in precedenza perché essa può essere caratterizzata come il limite delle norme p quando p tende all'infinito.

$$\forall \boldsymbol{a} \in \mathbb{R}^n, \lim_{p\to\infty} \|\boldsymbol{a}\|_p = \|\boldsymbol{a}\|_\infty.$$

Qualunque sia il vettore $\boldsymbol{a}$ appartenente a $\mathbb{R}^n$, la sua norma infinito può essere vista come il limite delle norme p quando p tende all'infinito. Questo risultato trova un riscontro visivo nel semplice esempio che abbiamo introdotto sulle sfere unitarie. Infatti, osserviamo che la sfera unitaria con norma infinito è il limite delle altre sfere unitarie di indice p quando p tende all'infinito: S_1 rappresenta la sfera più interna, S_2 quella intermedia, e S_∞ è l'ultima, che corrisponde alla sfera unitaria con norma infinito.

Nella lezione abbiamo visto alcuni esempi particolari di norme di vettore. In maniera rigorosa, la norma di un vettore è una applicazione, ovvero una funzione che mappa l'insieme $\mathbb{R}^n$ (l'insieme dei vettori di dimensione n) nei numeri reali $\mathbb{R}$. In particolare, a ogni vettore $\boldsymbol{a}$ associamo un numero reale, che rappresenta la sua norma, denotata da $\|\boldsymbol{a}\|$

$$\|\cdot\|: \mathbb{R}^n \to \mathbb{R}, \quad \boldsymbol{a} \to \|\boldsymbol{a}\|.$$

Le proprietà di questa applicazione norma sono le seguenti:

1. Positività: $\|\boldsymbol{a}\| \geq 0$; $\|\boldsymbol{a}\| = \boldsymbol{0} \iff \boldsymbol{a} = \boldsymbol{0}$.
   La norma di $\boldsymbol{a}$ è maggiore o uguale a zero per ogni vettore $\boldsymbol{a}$, quindi non è mai negativa. In particolare, è sempre positiva, a meno che il vettore $\boldsymbol{a}$ non sia il vettore nullo (cioè, $\boldsymbol{a} = \boldsymbol{0}$). Questa proprietà è chiamata positività.

2. Omogeneità: $\|c\boldsymbol{a}\| = |c|\|\boldsymbol{a}\|\ \forall c \in \mathbb{R}$.
   La seconda proprietà, chiamata omogeneità, afferma che, se prendiamo un vettore $\boldsymbol{a}$ e lo moltiplichiamo per uno scalare c (un coefficiente reale), otteniamo un nuovo vettore $c\boldsymbol{a}$ la cui norma è data dal valore assoluto di c moltiplicato per la norma di $\boldsymbol{a}$. Questo significa che la norma è omogenea rispetto alla moltiplicazione per scalari.

3. Disuguaglianza triangolare: $\|\boldsymbol{a} + \boldsymbol{b}\| \leq \|\boldsymbol{a}\| + \|\boldsymbol{b}\|$.
   La proprietà, detta disuguaglianza triangolare, afferma che, se prendiamo un vettore $\boldsymbol{a}$ e gli sommiamo un altro vettore $\boldsymbol{b}$, otteniamo un nuovo vettore $\boldsymbol{a} + \boldsymbol{b}$ la cui norma è minore o uguale alla somma delle norme di $\boldsymbol{a}$ e $\boldsymbol{b}$. Questo concetto ha una chiara interpretazione nella geometria euclidea, in particolare nel contesto della norma due. Infatti, quando sommiamo due vettori, secondo la nota regola del parallelogramma, otteniamo un vettore che congiunge i vertici dei due vettori originali, e il modulo di questo vettore è minore o uguale alla somma dei moduli degli altri due vettori. Questa proprietà si generalizza a qualsiasi tipo di norma e viene chiamata disuguaglianza triangolare o disuguaglianza del triangolo.

Ricordiamo ora una proprietà rilevante del prodotto scalare, la disuguaglianza di Cauchy-Schwarz. La disuguaglianza di Cauchy-Schwarz afferma che il valore assoluto del prodotto scalare tra due vettori $\boldsymbol{a}$ e $\boldsymbol{b}$ è minore o uguale al prodotto delle loro norme euclidee

$$|(\boldsymbol{a},\boldsymbol{b})| \leq \|\boldsymbol{a}\|_2\|\boldsymbol{b}\|_2 \qquad \forall \boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^n.$$

Facciamo un esempio che verifica la disuguaglianza di Schwarz su un caso molto semplice:

$$\boldsymbol{a} = \begin{bmatrix}1\\3\end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix}2\\-2\end{bmatrix}, \qquad (\boldsymbol{a},\boldsymbol{b}) = 1\cdot 2 - 3\cdot 2 = -4,$$
$$\|\boldsymbol{a}\|_2 = \sqrt{10}, \qquad \|\boldsymbol{b}\|_2 = \sqrt{8},$$
$$4 = |(\boldsymbol{a},\boldsymbol{b})| \leq \|\boldsymbol{a}\|_2\|\boldsymbol{b}\|_2 = \sqrt{80} \simeq 8.94.$$

Il prodotto scalare di $\boldsymbol{a}$ con $\boldsymbol{b}$ è -4. La norma euclidea dei vettori $\boldsymbol{a}$ e $\boldsymbol{b}$ è rispettivamente $\sqrt{10}$ e $\sqrt{8}$. Dunque, abbiamo che 4, che è il valore assoluto del prodotto scalare, è minore uguale del prodotto delle norme $\sqrt{80}$, circa 9.

Mostriamo, infine, le proprietà generali che permettono di definire una norma di matrice. Ritroveremo proprietà analoghe a quelle viste nel caso dei vettori. La norma è un'applicazione dallo spazio $\mathbb{R}^{n\times n}$ allo spazio $\mathbb{R}$, che associa a ogni matrice un numero reale, che chiamiamo norma (o norma matriciale) di A

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

La norma matriciale soddisfa quattro proprietà:

4. Positività: $\|A\| \geq 0$; $\|A\| = 0 \iff A = 0$.
   La norma di A è maggiore o uguale a zero ed è uguale a zero solo se la matrice è identicamente nulla, ovvero se tutti gli elementi della matrice sono nulli.
1. Omogeneità: $\|cA\| = |c|\|A\| \; \forall c \in \mathbb{R}$.
   La norma di $\|cA\|$, con c numero reale, è il modulo di c per la norma di A.
2. Disuguaglianza triangolare: $\|A+B\| \leq \|A\| + \|B\|$.
   La norma di $A+B$, con A e B matrici, è minore o uguale alla somma delle norme di A e B, ovvero soddisfa la disuguaglianza triangolare.
3. Cauchy-Schwarz: $\|AB\| \leq \|A\|\|B\|$.
   Questa proprietà, che vale per le norme indotte o naturali (ovvero per tutte quelle norme che derivano spontaneamente da una norma vettoriale), afferma che, se abbiamo due matrici, A e B, e ne facciamo il prodotto riga per colonna (e ricordiamo che, in particolare, se le due matrici sono quadrate, il prodotto è una matrice quadrata della stessa dimensione), la norma di AB è minore o uguale alla norma di A per la norma di B. Questa disuguaglianza ricorda molto da vicino la disuguaglianza di Cauchy-Schwarz che abbiamo visto per il prodotto scalare dei vettori.

Queste proprietà si scrivono esattamente come nel caso dei vettori, con la differenza che dobbiamo attribuire significati diversi alle notazioni.