Lezioni di Calcolo Numerico
Lezione 06: Sistemi lineari

Alberto Tibaldi

5 maggio 2018

Indice

Introduzione

L'argomento *sistemi lineari* è talmente importante da non richiedere tante presentazioni: ogni volta che si devono risolvere equazioni che per qualche motivo sono tra loro *accoppiate*, dobbiamo risolvere un sistema. Per fare un esempio, ogni volta che si cerca di *dare in pasto* un'equazione differenziale a un computer, la si approssima in modo da trasformarla in un sistema di equazioni. Per entrare nel merito,

- un **sistema** è un insieme di equazioni tra loro accoppiate e che devono essere risolte **simultaneamente**;
- un **sistema lineare** è un sistema in cui ciascuna delle equazioni è **lineare**, ovvero non contiene potenze, prodotti tra le variabili o funzioni non lineari di vario tipo (logaritmi, esponenziali...). Per esempio, il sistema

  $$\begin{cases} x + y + 2 = 0 \\ xy + 3 = 0 \end{cases}$$

  è non lineare, dal momento che contiene il prodotto xy. D'altra parte,

  $$\begin{cases} x + y - 5 = 0 \\ 4x + 7y + 9 = 0 \end{cases}$$

  è un tipico esempio di sistema lineare.

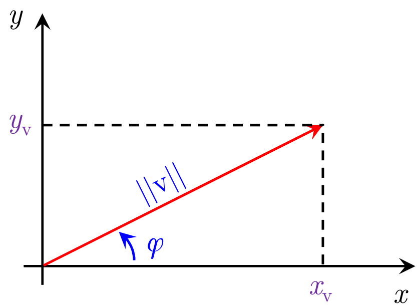

Figura 1: Rappresentazione geometrica di un vettore (freccia rossa) in un sistema di assi cartesiano avente componenti x_v, y_v, o equivalentemente rappresentato in coordinate *polari* (parenti di *modulo*, *direzione*, *verso*).

Il concetto di sistema lineare è indissolubilmente legato a quello di matrice: in effetti, al fine di risolvere i sistemi lineari, la procedura è prima di tutto determinare la *matrice associata* al sistema e, poi, manipolarla. Lo scopo di questa parte del testo sarà presentare dei metodi *furbi* per risolvere sistemi lineari in diverse condizioni. Finché non verrà esplicitamente detto diversamente, ci concentreremo su sistemi a n equazioni e n incognite aventi una e una sola soluzione, dunque associati a matrici quadrate con determinante diverso da zero. Verso la fine dell'argomento studieremo la soluzione di sistemi sovradimensionati, ossia con più equazioni che incognite.

Ci concentreremo sui **sistemi lineari densi**, ossia le cui matrici non contengono molti zeri, e con **metodi diretti**, che cioè risolvono il sistema *in un colpo solo* o comunque *con un numero di passi noto a priori*. Per dare un'idea, risolvere un sistema lineare denso mediante un metodo diretto è assolutamente ragionevole con matrici di sistema 1000×1000, cioè con 1000 righe e 1000 colonne, utilizzando un PC di fascia media. Il mio *record* in questo senso è stato, con MATLAB®e una workstation (da 32 GiB di RAM), un sistema 30000×30000, ma già si era al limite. Quando si raggiungono dimensioni così grandi, è necessario utilizzare tecniche di cui non parleremo, cercando se possibile di sfruttare l'eventuale sparsità[1] della matrice o di ricorrere a metodi *iterativi*.

1 Concetti preliminari sui sistemi lineari

Per via del fortissimo legame presente tra sistemi lineari e vettori/matrici, in questa sezione riprenderemo e/o introdurremo alcuni concetti su di essi.

1.1 Il concetto di norma

La parola *vettore* viene pronunciata molto frequentemente in vari contesti apparentemente disgiunti: algebra lineare, geometria, fisica, informatica. Buone notizie: tutto sommato, in tutti questi contesti si parla sempre della stessa cosa, ma studiata da punti di vista un po' diversi.

- Dal punto di vista dell'algebra lineare e da quello informatico, un vettore viene definito come un insieme di numeri disposti in riga o in colonna.
- Dal punto di vista geometrico e da quello fisico, un vettore viene visto come un segmento al quale si attribuisce un modulo, una direzione e un verso; in particolare, in geometria viene sempre disegnato come una freccia che parte dall'origine degli assi, mentre in fisica si parla di *vettori applicati*[2] con un significato abbastanza simile.

In alcune situazioni, è possibile identificare con grande facilità la connessione tra tutti questi punti di vista. Pensiamo per esempio a $\mathbb{R}^2$, quindi a un vettore $\underline{v}$ con due componenti reali. Ciò che si può fare è attribuire alla prima e alla seconda componente del vettore, x_v e y_v, il significato geometrico di coordinate di un punto sul piano cartesiano. A questo punto, è possibile ottenere la

[1] con un sistema *sparso* (di quelli con tanti zeri che non studieremo), un sistema 150000×150000 si risolve abbastanza tranquillamente con un normale computer in circa un secondo

[2] per esempio, studiando le forze applicate su un punto le frecce partono da esso.

rappresentazione geometrica del vettore congiungendo mediante una freccia l'origine del nostro sistema di riferimento con il punto disegnato sul piano. La rappresentazione per componenti cartesiane $[x_\mathrm{v} \;\; y_\mathrm{v}]$ non è l'unica ammissibile: in Fisica è comune sentir parlare, come anticipato poco fa, di *modulo*, *direzione* e *verso*. Queste tre quantità sono facilmente ottenibili, mediante un po' di trigonometria, a partire dalle componenti:

- si può identificare il *modulo* del vettore mediante la sua *lunghezza*, calcolabile mediante il teorema di Pitagora:

$$L_\mathrm{v} = \sqrt{x_\mathrm{v}^2 + y_\mathrm{v}^2}; \tag{1}$$

- si può identificare la *direzione* per esempio usando l'angolo φ compreso tra il vettore $\underline{v}$ e l'asse x:

$$\varphi = \arctan\left(\frac{y_\mathrm{v}}{x_\mathrm{v}}\right);$$

- il *verso* è semplicemente il *segno* di questo vettore: positivo se punta verso l'alto, negativo se punta verso il basso.

È possibile riferirsi alla Fig. 1 al fine di visualizzare meglio i concetti appena introdotti.

Per poter dire di *conoscere completamente* il vettore $\underline{v} \in \mathbb{R}^2$, sono necessari 2 numeri, sia che si parli delle componenti cartesiane, sia che si parli di quelle polari. Tuttavia, può accadere che, per un motivo o per un altro, si decida di sacrificare parte di queste informazioni ed *estrarre*, a partire da $\underline{v}$, un qualche numero che in qualche modo lo rappresenti. Questo potrebbe essere per esempio il caso della *lunghezza*: potremmo dire che del vettore ci interessa solo sapere il modulo, ma non come sia diretto. Al fine di produrre dei numeri che rappresentino le più interessanti proprietà del vettore, è possibile ricorrere a diverse definizioni di **norma**[3]. Si consideri da ora un vettore colonna $\underline{x} \in \mathbb{R}^n$ definito come

$$\underline{x} = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \end{bmatrix}^\mathrm{T}.$$

In questo testo considereremo in particolare tre diverse definizioni di norma per $\underline{x}$.

- La norma 1, definita come la somma dei moduli delle componenti del vettore

$$||\underline{x}||_1 = \sum_{i=1}^{n} |x_i|\,.$$

- La norma 2, definita come la radice della somma dei quadrati delle componenti

$$||\underline{x}||_2 = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

  Questa è una delle norme più significative, dal momento che rappresenta un'estensione al caso n-dimensionale del teorema di Pitagora; infatti, per $n = 2$, questa degenera al classico teorema di Pitagora (1). In altre parole, questa si può vedere come una generalizzazione del concetto di lunghezza del vettore. Esiste un modo diverso per scrivere la norma 2, basato sul prodotto righe per colonne. Infatti, questa si può anche scrivere come

$$||\underline{x}||_2 = \sqrt{\underline{x}^\mathrm{T}\underline{x}}.$$

  Dato per esempio un vettore in $\mathbb{R}^3$, si avrebbe:

[3] questo è già stato fatto in precedenza: quando abbiamo avuto bisogno, per questioni di maneggiabilità dei dati, di ricavare un singolo numero a partire da una funzione o da un vettore, abbiamo dovuto fatto ricorso al concetto di norma infinito

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \underline{x}^{\mathrm{T}} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix},$$

e, quindi,

$$\underline{x}^{\mathrm{T}}\underline{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1^2 + x_2^2 + x_3^2 = \sum_{i=1}^{3} x_i^2.$$

Calcolando quindi la radice di questo, si arriva alla definizione appena scritta.

- La già citata norma infinito, definita come il massimo dei moduli delle componenti del vettore:

$$\|\underline{x}\|_\infty = \max_{1 \le i \le n} |x_i|.$$

Il concetto di norma è anche applicabile a una matrice $\underline{\underline{A}} \in \mathbb{R}^{n,n}$. Sia a_{ij} l'elemento della matrice $\underline{\underline{A}}$ sulla riga i e sulla colonna j. Allora,

- La norma 1 della matrice $\underline{\underline{A}}$ si può scrivere come

$$\left\|\underline{\underline{A}}\right\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}|.$$

  Spiegato a parole, questo significa, per ogni j-esima colonna, calcolare la somma dei suoi moduli, e trovare il massimo tra queste. Questa si può anche spiegare come: considerare la matrice $\underline{\underline{A}}$ come un impilamento di n vettori colonna; calcolare la norma 1 di ciascuno dei vettori colonna, e ottenere quindi un singolo vettore riga contenente tutte le norme 1 delle varie colonne; di queste, trovare la massima.

- La norma infinito della matrice $\underline{\underline{A}}$ si può scrivere come

$$\left\|\underline{\underline{A}}\right\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|.$$

  Si tratta un po' di un'operazione opposta rispetto alla norma 1 di matrice: considero la matrice come l'impilamento di tanti vettori riga, e di ciascuno di questi calcolo la norma 1, ottenendo quindi un vettore colonna che contiene su ciascuna componente la norma 1 della riga di $\underline{\underline{A}}$ corrispondente; di queste, calcolo poi la massima.

- Esiste un'estensione del concetto di norma 2 anche per quanto riguarda le matrici, e questa viene detta **norma spettrale**; sarà tuttavia discussa più avanti nel testo, nel capitolo relativo ad autovalori e autovettori.

Abbiamo introdotto quindi i concetti di norma di vettori e di matrici; tuttavia, in generale, non è detto che queste siano *compatibili*. Si dice che una norma di vettore e una norma di matrici sono compatibili se vale la relazione

$$\left\|\underline{\underline{A}}\,\underline{x}\right\| \le \left\|\underline{\underline{A}}\right\| \|\underline{x}\|. \tag{2}$$

Si noti che in (2) non sono stati specificati i *numeri al pedice*: la relazione deve essere valida per norme di vettori e matrici di tipo corrispondente, per esempio

$$\left\|\underline{\underline{A}}\,\underline{x}\right\|_1 \le \left\|\underline{\underline{A}}\right\|_1 \|\underline{x}\|_1.$$

Inoltre, si osservi che, al membro sinistro, si ha $\underline{\underline{A}}\,\underline{x}$, che è un vettore: il prodotto di una matrice per un vettore colonna è ancora un vettore colonna! Dunque, l'unica norma di matrice che entra in gioco nell'espressione (2) è quella al membro destro. Inoltre, essendo la norma semplicemente un numero sia per i vettori, sia per le matrici, questa relazione è una semplice diseguaglianza tra scalari. Tutte le norme che considereremo in questa trattazione sono compatibili. Questa è un'ottima notizia, perché ci permette quindi di sfruttare la diseguaglianza (2) per poter effettuare maggiorazioni o minorazioni di alcuni termini al momento di studiare equazioni matriciali. In questo senso, è utile tenere presente anche la diseguaglianza

$$\left\|\underline{\underline{A}}\,\underline{\underline{B}}\right\| \leq \left\|\underline{\underline{A}}\right\| \left\|\underline{\underline{B}}\right\|. \tag{3}$$

Infine, si sappia che per tutte le norme considerate, data $\underline{\underline{I}}$ la matrice identità, $\left\|\underline{\underline{I}}\right\| = 1$.

1.2 Carrellata delle principali categorie di matrici

Per quanto questa trattazione sia focalizzata su matrici dense, è importante sapere che una matrice potrebbe avere particolari configurazioni, per esempio avere un certo numero di elementi a 0; di conseguenza, i metodi che utilizzeremo per risolvere i sistemi lineari associati a esse, saranno diversi. Proponiamo ora una *carrellata* delle più famose matrici; a questo fine, si consideri una matrice $\underline{\underline{A}} \in \mathbb{R}^{5,5}$, che nella più generale forma si può scrivere come

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix}.$$

Da qui si intende che, quando un elemento a_{ij} è scritto esplicitamente (al posto di 0), questo possa essere non nullo.

- Una matrice si dice **diagonale** se solo gli elementi della diagonale principale sono diversi da zero:

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & 0 & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 & 0 \\ 0 & 0 & a_{33} & 0 & 0 \\ 0 & 0 & 0 & a_{44} & 0 \\ 0 & 0 & 0 & 0 & a_{55} \end{bmatrix},$$

  ossia, se $i \neq j$, si ha che $a_{ij} = 0$.

- Una matrice si dice **triangolare superiore** se tutti gli elementi sotto la diagonale principale (esclusa) sono uguali a zero:

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ 0 & a_{22} & a_{23} & a_{24} & a_{25} \\ 0 & 0 & a_{33} & a_{34} & a_{35} \\ 0 & 0 & 0 & a_{44} & a_{45} \\ 0 & 0 & 0 & 0 & a_{55} \end{bmatrix},$$

  ovvero se, per $i > j$, si ha che $a_{ij} = 0$.

- Una matrice si dice **triangolare inferiore** se tutti gli elementi sopra la diagonale principale (esclusa) sono uguali a zero:

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} & 0 \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix},$$

  ovvero se, per $i < j$, si ha che $a_{ij} = 0$.

- Una matrice si dice **tridiagonale** se solo gli elementi della diagonale principale, della prima codiagonale inferiore e della prima codiagonale superiore sono diversi da zero:

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} & 0 \\ 0 & 0 & a_{43} & a_{44} & a_{45} \\ 0 & 0 & 0 & a_{54} & a_{55} \end{bmatrix},$$

  ovvero se, per $|i - j| > 1$, $a_{ij} = 0$ (in questo caso, per esempio, se considero a_{53}, ho che $|5 - 3| = 2$, che è maggiore di 1, e quindi l'elemento va a 0.

- Una matrice si dice **Hessenberg superiore** se, oltre a essere come la triangolare superiore, ha anche gli elementi della prima codiagonale inferiore diversi da zero.

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ 0 & a_{32} & a_{33} & a_{34} & a_{35} \\ 0 & 0 & a_{43} & a_{44} & a_{45} \\ 0 & 0 & 0 & a_{54} & a_{55} \end{bmatrix},$$

  ovvero se, per $i > j + 1$, $a_{ij} = 0$ (per esempio, $i = 5$, $j = 1$, soddisfa $5 > 1 + 1$, e infatti l'elemento va a 0).

1.3 Operazioni e altre definizioni sulle matrici

Terminata la carrellata delle principali forme di matrici che può capitare di trovare, è opportuno introdurre alcune altre matrici che si distinguono più per alcune proprietà particolari che per il numero o le posizioni di elementi nulli. Ora come ora può non essere chiaro a cosa servano tutte queste definizioni ma, man mano che andremo avanti, le ritroveremo e riutilizzeremo!

1.3.1 Trasposta di una matrice e matrici simmetriche

Prima di tutto, a questo fine, è opportuno ricordare la definizione di **trasposta** $\underline{\underline{A}}^T$ di una matrice $\underline{\underline{A}}$, che si ottiene scambiando gli elementi simmetrici rispetto alla diagonale. Per esempio, a partire dalla nostra $\underline{\underline{A}} \in \mathbb{R}^{5,5}$, si ha:

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix} \Longrightarrow \underline{\underline{A}}^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} & a_{41} & a_{51} \\ a_{12} & a_{22} & a_{32} & a_{42} & a_{52} \\ a_{13} & a_{23} & a_{33} & a_{43} & a_{53} \\ a_{14} & a_{24} & a_{34} & a_{44} & a_{54} \\ a_{15} & a_{25} & a_{35} & a_{45} & a_{55} \end{bmatrix}.$$

Detto ancora in un altro modo, bisogna scrivere, al posto di ciascun a_{ij}, il a_{ji}: scambiare gli indici di righe e di colonne.

Esistono matrici che non cambiano una volta che viene applicata l'operazione di trasposizione; queste matrici vengono dette **simmetriche**. Questa cosa è possibile quando gli elementi simmetrici rispetto alla diagonale principale sono uguali tra loro. Questo è il caso della seguente matrice $\underline{\underline{A}}$:

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{12} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{13} & a_{23} & a_{33} & a_{34} & a_{35} \\ a_{14} & a_{24} & a_{34} & a_{44} & a_{45} \\ a_{15} & a_{25} & a_{35} & a_{45} & a_{55} \end{bmatrix},$$

in cui si è cercato di evidenziare con gli stessi colori gli elementi simmetrici[4]. In sintesi, il fatto che applicare la trasposizione non cambia niente, e quindi la **simmetria** della matrice, si può esprimere come

[4]purtroppo si avevano a disposizione pochi colori e alcuni sono stati ripetuti

$$\underline{\underline{A}}^{\mathrm{T}} - \underline{\underline{A}} = 0.$$

1.3.2 Inversa di una matrice, matrici ortogonali e matrici di permutazione

Data una matrice $\underline{\underline{A}}$, a patto che il suo determinante $\det\{\underline{\underline{A}}\}$ sia diverso da zero, è possibile definire la matrice inversa $\underline{\underline{A}}^{-1}$ tale per cui

$$\underline{\underline{A}}^{-1}\underline{\underline{A}} = \underline{\underline{A}}\,\underline{\underline{A}}^{-1} = \underline{\underline{I}}.$$

Un'altra definizione che ci tornerà molto utile più avanti è quella di **matrice ortogonale**. Una matrice $\underline{\underline{A}}$ viene detta **ortogonale** se

$$\underline{\underline{A}}^{\mathrm{T}}\underline{\underline{A}} = \underline{\underline{A}}\,\underline{\underline{A}}^{\mathrm{T}} = \underline{\underline{I}}.$$

Si noti che, nel caso di matrici ortogonali, $\underline{\underline{A}}^{\mathrm{T}} = \underline{\underline{A}}^{-1}$.

Come noto dai rudimenti di Geometria, la parola *ortogonale* è un po' un sinonimo di *perpendicolare*. Tuttavia questo concetto, più che alle matrici, sembra essere applicabile ai vettori: per esempio, si dice che due vettori sono ortogonali se da un punto di vista geometrico formano un angolo retto (pensiamo a $\mathbb{R}^2$...), o, generalizzando, se il loro prodotto scalare è nullo. In questo secondo senso, la matrice è ortogonale se ciascuna delle proprie colonne, trasposta e trasformata dunque in una riga, ha prodotto scalare uguale a zero con tutte le altre, tranne che per sé stessa.

Per chiarire un po' meglio questo concetto, è opportuno focalizzarci su una particolare classe di matrici ortogonali: le **matrici di trasposizione**. Considerando per esempio $\mathbb{R}^{3,3}$, la matrice

$$\underline{\underline{P}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

è una matrice di permutazione. Il motivo per cui matrici di questo tipo vengono chiamate in questo modo è legato a come esse agiscono su un vettore. Immaginiamoci di considerare $\underline{x} = [x_1 \;\; x_2 \;\; x_3]^{\mathrm{T}}$; allora, calcolando il prodotto righe per colonne, otteniamo

$$\underline{\underline{P}}\,\underline{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_3 \\ x_2 \end{bmatrix},$$

ovvero, uno scambio tra la seconda e la terza componente di $\underline{x}$! Le matrici di permutazione sono infatti ottenute a partire dalla matrice identità $\underline{\underline{I}}$, e si scambiano alcune sue righe; in questo modo si ottiene una matrice che, moltiplicata da sinistra per un vettore colonna, produce un vettore con le medesime componenti, ma scambiate di posizione: **permutate**!

Però: cosa c'entrano queste matrici con le matrici ortogonali? Beh, è semplice dimostrare, su questo esempio, che le matrici di permutazione appartengono alle matrici ortogonali! Infatti, calcolando, si dimostra che

$$\underline{\underline{P}}\,\underline{\underline{P}}^{\mathrm{T}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}_{\underline{\underline{P}}}\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}_{\underline{\underline{P}}^{\mathrm{T}}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\underline{\underline{I}}}!$$

Questa proprietà è, ovviamente, generale per matrici di permutazione $\underline{\underline{P}} \in \mathbb{R}^{n,n}$!

1.3.3 Matrici a diagonale dominante (per righe/colonne)

Data una matrice $\underline{\underline{A}} \in \mathbb{R}^{n,n}$, essa viene definita **a diagonale dominante per righe** se

$$|a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} \left|a_{ij}\right|, \quad i = 1, \ldots, n. \tag{4}$$

Cosa significa questa cosa? Per capirla meglio, proviamo a guardare un esempio pratico:

$$\underline{\underline{A}} = \begin{bmatrix} 4 & 1 & -1 \\ 2 & -7 & 1 \\ 3 & -2 & 9 \end{bmatrix}.$$

Prima di tutto, invece di questa matrice, conviene guardare $\left|\underline{\underline{A}}\right|$, ovvero la matrice dei moduli:

$$\left|\underline{\underline{A}}\right| = \begin{bmatrix} 4 & 1 & 1 \\ 2 & 7 & 1 \\ 3 & 2 & 9 \end{bmatrix}.$$

Adesso, seguendo la definizione (4), verifichiamo che ciascun elemento della diagonale sia più grande della somma di tutti gli elementi sulle righe corrispondenti. In particolare:

- per la prima riga,
$$4 > 1 + 1$$
- per la seconda riga,
$$7 > 2 + 1$$
- per la terza riga,
$$9 > 3 + 2.$$

Dal momento che tutte e tre queste condizioni sono soddisfatte, la matrice $\underline{\underline{A}}$ in questione è a diagonale dominante per righe.

Allo stesso modo, è possibile proporre una definizione di matrici a diagonali dominanti per colonne. Come si può immaginare, seguendo l'esempio appena svolto, invece di fare la somma dei moduli degli elementi fuori-diagonale delle righe, lo faremo con quelli delle colonne. Tuttavia, non è garantito che una matrice a diagonale dominante per righe lo sia anche per colonne; ad esempio, nella matrice $\underline{\underline{A}}$ appena considerata, si ha che, per la prima colonna,

$$4 \not> 2 + 3,$$

quindi $\underline{\underline{A}}$ è a diagonale dominante per righe, ma non per colonne.

1.3.4 Matrici simmetriche definite positive

Consideriamo, almeno preliminarmente (per poi riprenderla in seguito), un'ultima definizione: quella di **matrice simmetrica definita positiva**. Data una matrice $\underline{\underline{A}}$ simmetrica, essa si dice **definita positiva** se, **per ogni** $\underline{x}$ vettore colonna non identicamente nullo,

$$\underline{x}^{\mathrm{T}} \underline{\underline{A}}\, \underline{x} > 0.$$

Discutiamo un momento questa definizione. Prima di tutto, $\underline{\underline{A}}\, \underline{x}$ è un vettore colonna, dal momento che $\underline{x}$ è un vettore colonna e gli si sta applicando da sinistra una matrice. Quindi, moltiplicando il vettore colonna $(\underline{\underline{A}}\, \underline{x})$ al vettore riga $\underline{x}^{\mathrm{T}}$ da destra, il risultato sarà un singolo numero: uno scalare. Per questo motivo, ha senso chiedersi se questo sia maggiore o minore di zero.

Detto questo, questa definizione è abbastanza inutile, scritta in questo modo: è **impossibile da verificare**! Infatti, visto che questa deve essere verificata **per ogni vettore non nullo**, noi dovremmo, per quanto ne sappiamo ora, provare qualsiasi vettore esistente, e vedere se il numero che esce è sempre maggiore di 0. Più avanti studieremo di nuovo questa condizione, cosa implica, e un modo furbo per verificarla.

Per ora, chiudo con un piccolo *spoiler*: dimostreremo che, data una matrice $\underline{\underline{B}} \in \mathbb{R}^{n,n}$, la matrice $\underline{\underline{A}}$ ottenuta come

$$\underline{\underline{A}} = \underline{\underline{B}}^{\mathrm{T}} \underline{\underline{B}},$$

con $\det\{\underline{\underline{B}}\} \neq 0$, è simmetrica definita positiva.

1.4 Matrice associata a un sistema lineare

È ora giunto il momento più opportuno per visualizzare lo stretto legame esistente tra sistemi lineari e matrici. In buona sostanza, questo legame è costituito dal concetto di prodotto righe per colonne. Si consideri per esempio un sistema lineare 4×4:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 = b_3 \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4 = b_4. \end{cases}$$

Di questo, ci si concentri sulla prima equazione:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 = b_1.$$

Le incognite sono gli x_i, il termine noto è il b_1, e gli altri termini noti sono i coefficienti che moltiplicano le x_i. Sfruttando la definizione di prodotto righe per colonne, è possibile riscrivere questa equazione come

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = b_1.$$

La stessa operazione si può ripetere per le altre righe, e, una volta impilate, ci permette di riscrivere il sistema lineare in forma matriciale come

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}}_{\underline{\underline{A}}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\underline{x}} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}}_{\underline{b}},$$

o, in forma compatta,

$$\underline{\underline{A}}\,\underline{x} = \underline{b}.$$

1.4.1 Esempio con una matrice 2×2

Al fine di completare la comprensione di questo argomento, si provi a scrivere in forma matriciale il sistema

$$\begin{cases} 2x + y = 4 \\ 3y + 4x = 5. \end{cases}$$

La prima cosa da fare, è decidere in che ordine mettere le incognite; per esempio, possiamo definire il vettore $\underline{x}$ nella forma

$$\underline{x} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

In questo modo, ha quindi senso, per chiarire, riscrivere il sistema mettendo in ordine le incognite:

$$\begin{cases} 2x + y = 4 \\ 4x + 3y = 5. \end{cases}$$

A questo punto è immediato scrivere il sistema in forma matriciale:

$$\begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}.$$

1.5 Condizionamento di un sistema lineare

Precedentemente, era stato introdotto il problema del condizionamento. Questo genere di analisi è applicabile anche ai sistemi lineari nella forma

$$\underline{\underline{A}}\,\underline{x} = \underline{b}.$$

La soluzione del sistema lineare è data dal vettore $\underline{x}$, ottenuto a partire dai dati $\underline{\underline{A}}$ e $\underline{b}$. A questo punto, per studiare il problema del condizionamento, ricordiamo che è necessario perturbare i dati, ottenendo $\overline{\underline{\underline{A}}}$, $\overline{\underline{b}}$, e, facendo tutti i conti in aritmetica esatta su questi dati perturbati, si ottiene una soluzione $\overline{\underline{x}}$. Quindi, ci si pone domande sull'errore relativo ε_x, cercando di stimarlo a partire dagli errori introdotti sui dati in seguito alla loro perturbazione. Si noti che tutti questi errori sono calcolati a partire da vettori e/o matrici; di conseguenza, utilizzare il modulo sarebbe improprio: finiremmo per avere un sacco di numeri, e non capire esattamente a cosa servano. Per questo motivo, la regola generale è che, **quando si ha a che fare con vettori o matrici, il segno di modulo va sostituito a quello di norma**. I conti che stiamo per fare funzionano sia con la norma 1, sia con la norma 2, sia con la norma infinito, quindi non lo specificheremo ulteriormente ai pedici dei segni di norma. Per esempio, per l'errore sulla soluzione, abbiamo

$$\varepsilon_x = \frac{\|\underline{x} - \overline{\underline{x}}\|}{\|\underline{x}\|}.$$

Allo stesso modo, per $\underline{\underline{A}}$ e $\underline{b}$, avremo:

$$\varepsilon_A = \frac{\left\|\underline{\underline{A}} - \overline{\underline{\underline{A}}}\right\|}{\left\|\underline{\underline{A}}\right\|}, \qquad \varepsilon_b = \frac{\|\underline{b} - \overline{\underline{b}}\|}{\|\underline{b}\|}.$$

È possibile dimostrare il seguente teorema: dato per ipotesi

$$\left\|\underline{\underline{A}} - \overline{\underline{\underline{A}}}\right\| < \frac{1}{2\left\|\underline{\underline{A}}^{-1}\right\|},$$

il sistema perturbato

$$\overline{\underline{\underline{A}}}\,\overline{\underline{x}} = \overline{\underline{b}}$$

ammette una e una sola soluzione, e

$$\varepsilon_x \le 2\kappa(\underline{\underline{A}})(\varepsilon_A + \varepsilon_b),$$

dove

$$\kappa(\underline{\underline{A}}) = \left\|\underline{\underline{A}}\right\| \left\|\underline{\underline{A}}^{-1}\right\|$$

è detto **numero di condizionamento del sistema lineare**. Si ribadisce che questi risultati e i prossimi valgono per le norme 1, 2, ∞.

È possibile spendere qualche parola in più, su questo numero di condizionamento. Infatti, se ricordiamo la proprietà (3), che ci permette di scrivere:

$$\left\|\underline{\underline{A}}\right\| \left\|\underline{\underline{A}}^{-1}\right\| \ge \left\|\underline{\underline{A}}\,\underline{\underline{A}}^{-1}\right\| = \left\|\underline{\underline{I}}\right\| = 1.$$

Questo dimostra che il numero di condizionamento di un sistema lineare non può essere più piccolo di 1. In particolare, se esso è prossimo a 1 (il caso migliore), il sistema si dice **ben condizionato**, ovvero, a piccole perturbazioni dei dati corrispondono piccole perturbazioni della soluzione $\overline{\underline{x}}$. Al contrario, se $\kappa(\underline{\underline{A}}) \gg 1$, piccole perturbazioni dei dati **potrebbero** portare enormi perturbazioni sulla soluzione.

1.5.1 Interpolazione polinomiale: reprise!

È ora giunto il momento di chiudere definitivamente l'argomento *interpolazione polinomiale*, rispondendo a una delle grandi domande ancora aperte dalle lezioni precedenti: come funziona il comando `polyfit` di MATLAB®?

Ciò che sappiamo è che MATLAB®lavora sulla forma monomiale, dal momento che `polyfit` permette di ottenere i corrispettivi coefficienti. Ripassiamo dunque questa forma monomiale:

$$p_n(x) = c_1 x^n + x_2 x^{n-1} + ... + c_n x + c_{n+1}.$$

Perché il polinomio sia interpolante, esso deve soddisfare le condizioni di interpolazione, che sono

$$p_n(x_i) = y_i, \qquad i = 1, ..., n+1.$$

Questo permette di scrivere il seguente sistema lineare:

$$\begin{cases} c_1 x_1^n + c_2 x_1^{n-1} + ... + c_n x_1 + c_{n+1} = y_1 \\ c_1 x_2^n + c_2 x_2^{n-1} + ... + c_n x_2 + c_{n+1} = y_2 \\ \vdots \\ c_1 x_n^n + c_2 x_n^{n-1} + ... + c_n x_n + c_{n+1} = y_n \\ c_1 x_{n+1}^n + c_2 x_{n+1}^{n-1} + ... + c_n x_{n+1} + c_{n+1} = y_{n+1}, \end{cases}$$

che, in forma matriciale, si può scrivere come

$$\underbrace{\begin{bmatrix} x_1^n & x_1^{n-1} & ... & x_1 & 1 \\ x_2^n & x_2^{n-1} & ... & x_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^n & x_n^{n-1} & ... & x_n & 1 \\ x_{n+1}^n & x_{n+1}^{n-1} & ... & x_{n+1} & 1 \end{bmatrix}}_{\underline{\underline{V}}} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \\ c_{n+1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \\ y_{n+1} \end{bmatrix}.$$

Si noti che il sistema è lineare anche se contiene questi x_i^n, perché le incognite del sistema sono i $\{c_i\}$. In ogni caso, se il sistema non fosse stato lineare, non saremmo mai stati in grado di scriverlo in forma matriciale! Detto questo, a partire da un vettore di punti $\underline{x}$ avente $n+1$ componenti, è possibile definire la cosiddetta matrice di Vandermonde nella forma di $\underline{\underline{V}}$. MATLAB®permette di calcolare a partire da un vettore $\underline{x}$ la matrice di Vandermonde mediante il comando `vander`. Sciaguratamente, è risaputo che i sistemi lineari associati alla matrice di Vandermonde siano mal condizionati, e, per questo, talvolta MATLAB®si lamenta quando si cerca di usare il comando `polyfit`. MATLAB®permette di calcolare, nelle varie norme di nostro interesse (1, 2, infinito) il numero di condizionamento di un sistema mediante il comando `cond`. In particolare, data `A` la variabile che contiene la matrice del sistema lineare, `cond(A)` restituisce il numero di condizionamento in norma 2, mentre `cond(A,1)` e `cond(A,inf)` in norma 1 e infinito.

1.5.2 Un altro esempio di sistema mal condizionato

L'*altro* esempio *famoso* di sistema mal condizionato è quello associato alla matrice di Hilbert. La matrice di Hilbert appare al momento di utilizzare il metodo dei minimi quadrati, che però in questo testo verrà descritto in maniera leggermente diversa. Esiste un comando MATLAB®che permette di generare una matrice di Hilbert con grande facilità: data `n` la dimensione della matrice (numero di righe e numero di colonne: la matrice di Hilbert è quadrata!), `H = hilb(n)` permette di generare la suddetta matrice e memorizzarla nella variabile `H`.

Prima di mostrare un segmento di codice contenente un esempio pratico, è opportuno ricordare qual è il nostro punto di arrivo: questa sezione è finalizzata a introdurre metodi numerici atti a risolvere sistemi lineari. Per questo motivo, un po' come per ogni tipo di problema che si vuole affrontare, è opportuno disporre di sistemi **dei quali conosciamo la soluzione**. In questo modo, potremmo calcolare l'errore relativo rispetto a un riferimento che sappiamo essere esatto, e quindi per esempio quantificare il legame tra condizionamento ed errore sulla soluzione.

Allora, un tipo di esercizio che affronteremo spesso, è: **applicare un metodo per la soluzione di un sistema lineare, del quale sappiamo già che la soluzione è il vettore unitario**. L'idea è: considerando una matrice fornitaci dal testo del problema, ricavare quel vettore di termini noti **tale per cui la soluzione sia uguale al vettore con tutti uni**: $\underline{x} = [1 \;\; 1 \;\; \ldots \;\; 1]^{\mathrm{T}}$. Considerando per esempio una matrice $\underline{\underline{A}} \in \mathbb{R}^{4,4}$

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix},$$

qual è il vettore $\underline{b}$ tale per cui il sistema lineare $\underline{\underline{A}}\,\underline{x} = \underline{b}$ abbia soluzione pari al vettore contenente tutti uni? Beh, nulla di più semplice: abbiamo appena scritto che $\underline{\underline{A}}\,\underline{x} = \underline{b}$, quindi, per calcolare questo vettore, dobbiamo semplicemente scrivere

$$\underline{b} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Ma quindi, la prima componente del vettore sarà il prodotto riga per colonna della prima riga di $\underline{\underline{A}}$ per il vettore $[1 \;\; 1 \;\; 1 \;\; 1]^{\mathrm{T}}$:

$$b_1 = a_{11} \times 1 + a_{12} \times 1 + a_{13} \times 1 + a_{14} \times 1 = \sum_{j=1}^{4} a_{1j}.$$

Come si può intuire, in generale, si avrà, per una matrice $\mathbb{R}^{n,n}$,

$$b_i = \sum_{j=1}^{n} a_{ij}. \tag{5}$$

A questo punto, possiamo proporre un segmento di codice in cui si cerca di mettere in correlazione l'errore relativo sulla soluzione del sistema causato dal condizionamento della matrice. In particolare, questo è il primo segmento di codice che mostrerà come risolvere un sistema lineare mediante MATLAB®. Data la matrice di sistema memorizzata nella variabile `A`, e il termine noto memorizzato nella variabile `b`, come nella notazione $\underline{\underline{A}}\,\underline{x} = \underline{b}$, la soluzione del sistema matriciale si può ricavare scrivendo `x = A\b`.

Viene ora riportato lo script che risolve il problema appena descritto.

```
clear
close all
clc

ord = [2 3 4 5 6 7 8 9 10 11 12 13 14]; % ordini della matrice da considerare

for n = 1:length(ord) % ciclo da 1 al numero di elementi di ord
    A = hilb(ord(n)); % in questo modo A viene definita come la matrice di
                      % Hilbert n x n
    b = sum(A,2); % il comando sum usato in questo modo produce, a partire
                  % dalla matrice A, un vettore colonna ottenuto sommando
                  % gli elementi di ciascuna riga; questo corrisponde con
                  % la formula (5) del testo.
    x = A\b; % trovo la soluzione x del sistema; in virtu` di quanto
             % abbiamo fatto nella riga precedente, questa dovra` essere un
             % vettore che in qualche modo "somigli" al vettore unitario:
             % se il numero di condizionamento e` basso (e dunque il
             % sistema ben condizionato), questo sara` molto simile al
             % vettore con tutti 1
    u = ones(ord(n),1); % vettore con tutti 1 di dimensione n x 1.
    errrel(n) = norm(u-x,inf)/norm(u,inf); % errore relativo in norma infinito
    Kinf(n) = cond(A,inf);
end

format short e % metto un format che permetta di apprezzare l'errore

[errrel.' Kinf.']

% guardando i risultati, risulta evidente come al crescere del numero di
% condizionamento, l'errore relativo cresca inesorabilmente
```

2 Soluzione di sistemi lineari di forma particolare

Incominceremo da adesso a studiare i metodi numerici che permettono di risolvere sistemi lineari. Il nostro scopo sarà studiare diversi metodi, cercando di suggerire quale possa essere il più adatto nelle varie situazioni che studieremo, in particolare a seconda della forma del sistema e quindi della matrice ad esso associata. Per *adatto* intendiamo quale sia quello che ci permetta di ottenere la soluzione minimizzando il numero di operazioni elementari: ottimizzando il costo computazionale. Stimeremo quindi per ciascuna di queste tecniche il numero di operazioni necessarie per portarle a termine, come funzione della dimensione n della matrice (che, a meno di specificare diversamente, sarà sempre quadrata, e quindi $n \times n$).

2.1 Soluzione di sistemi diagonali

La prima classe di *sistemi* lineari che studieremo è quella dei sistemi diagonali. Si tratta quindi di sistemi associati a matrici diagonali, per esempio quindi nella forma

$$\begin{cases} a_{11}x_1 = b_1 \\ a_{22}x_2 = b_2 \\ a_{33}x_3 = b_3 \\ \qquad \vdots \end{cases} \tag{6}$$

Francamente, io non riesco nemmeno a chiamarli *sistemi* con tanta convinzione, non perché io desideri in alcun modo sminuirli, quanto perché l'idea di sistema che ho in mente contiene implicitamente un accoppiamento tra le varie equazioni. Qui, semplicemente, abbiamo a che fare con n equazioni, tra loro indipendenti, e che quindi possono essere risolte ciascuna separatamente dalle altre. In particolare, la componente i-esima del vettore soluzione $\underline{x}$ si può scrivere come

$$x_i = \frac{b_i}{a_{ii}}. \tag{7}$$

Questo esempio è molto utile perché, nella sua semplicità, permette di introdurre il concetto di *costo computazionale*, ovvero il numero di operazioni elementari (somme, sottrazioni, prodotti, divisioni) necessarie durante l'esecuzione di un algoritmo.

Nel caso dei sistemi lineari diagonali, dal momento che ogni equazione può essere risolta semplicemente con la divisione (7), sono sufficienti n operazioni aritmetiche e quindi il costo computazionale è lineare: questo è il **meglio assoluto** che si possa fare con un *sistema lineare*.

2.2 Soluzione numerica di sistemi lineari triangolari (superiori)

La seconda categoria di sistemi lineari che impareremo a risolvere è quella dei **sistemi lineari triangolari**, ossia associati a matrici triangolari. Concentriamoci per esempio sui sistemi triangolari superiori, e, al fine di chiarire, su un esempio di sistema a 3 equazioni e 3 incognite:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ \qquad\quad\;\; a_{22}x_2 + a_{23}x_3 = b_2 \\ \qquad\qquad\qquad\quad a_{33}x_3 = b_3. \end{cases}$$

Come si può intuire, la matrice associata al sistema avrà una forma triangolare superiore. Aldilà di questo, si osservi prima di tutto che l'ultima equazione è disaccoppiata da tutte le altre, ovvero si può risolvere utilizzando l'espressione (7):

$$x_3 = \frac{b_3}{a_{33}}.$$

A questo punto, si consideri la penultima equazione: essa, è accoppiata solamente all'ultima, che abbiamo appena risolto, ma non alle precedenti! Dunque, a partire da

$$a_{22}x_2 + a_{23}x_3 = b_2,$$

possiamo risolvere rispetto a x_2, dal momento che **tutte le altre quantità sono note**; si avrà, quindi,

$$x_2 = \frac{b_2 - a_{23}x_3}{a_{22}}. \tag{8}$$

L'unica equazione che rimane è la prima, la quale è accoppiata con tutte le successive (ma non ne ha prima); tuttavia, abbiamo già risolto tutte le successive equazioni, e quindi, a partire da

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1,$$

è lecito risolvere rispetto a x_1, poiché tutte le altre quantità sono note; quindi, avremo:

$$x_1 = \frac{b_1 - a_{12}x_2 - a_{13}x_3}{a_{11}}. \tag{9}$$

Una volta chiarito questo esempio, è possibile proporre una ricetta generale per la soluzione di sistemi triangolari superiori.

1. Il primo passo, che risolvo solo una volta, consiste nel trovare la soluzione dell'ultima equazione, quella disaccoppiata da tutte le altre:

$$x_n = \frac{b_n}{a_{nn}}$$

2. A questo punto, devo procedere a ritroso; dato un indice i che varia da $n-1$ a 1, devo eseguire i seguenti passi:

   (a) calcolo un termine s_i definito come

$$s_i = \sum_{j=i+1}^{n} a_{ij}x_j$$

   questo rappresenta la generalizzazione dei termini evidenziati in colore rosso nelle equazioni (8) e (9) per matrici $n \times n$.

   (b) calcolo la componente i-esima del vettore soluzione come:

$$x_i = \frac{b_i - s_i}{a_{ii}}.$$

Se provassimo a contare il numero di operazioni necessarie per risolvere uno di questi sistemi, scopriremmo che il costo computazionale sarebbe circa pari a $n^2/2$.

Viene ora riportato uno script che implementa l'algoritmo appena descritto.

```
function x = fLez06_SistemaTriSup(A,b)

n = length(b); % dimensione della matrice di sistema
x = zeros(n,1); % inizializzo il vettore soluzione con tutti zeri

x(n) = b(n)/A(n,n); % primo passo: risolvo l'equazione indipendente dalle altre

for i = n-1:-1:1 % ciclo esterno: i va da n-1 a 1, a passi di -1

    % calcolo del s_i : valore di s alla i-esima iterazione
    s = 0;
    for j = i+1:n
        s = s+A(i,j)*x(j);
    end

    % calcolo della componente i-esima del vettore delle soluzioni
    x(i) = (b(i)-s)/A(i,i);

end
```

Si tenga presente che MATLAB®permette di effettuare in maniera più astuta (ottimizzata) il calcolo del termine s_i. In effetti, è possibile rimpiazzare il segmento di codice

```
s = 0;
for j = i+1:n
    s = s+A(i,j)*x(j);
end
```

semplicemente con

```
s = A(i,i+1:n)*x(i+1:n)
```

che effettua le operazioni di somma e moltiplicazione utilizzando il prodotto riga per colonna `*` di MATLAB®. Si sappia tuttavia che questo è un dettaglio implementativo, che sfrutta il fatto che i prodotti riga per colonna sono implementati *built-in* in MATLAB®, e quindi più veloci per il semplice fatto che sono stati ottimizzati a livello di programmazione; dal punto di vista del costo computazionale, il numero di operazioni effettuato è sempre lo stesso!

Un metodo del tutto analogo può essere applicato anche alle matrici triangolari inferiori, tenendo però conto delle differenze con i sistemi triangolari superiori. Nei sistemi triangolari inferiori, la prima equazione è quella disaccoppiata da tutte le altre, e quindi si dovrà procedere *in avanti*, anziché *all'indietro*, sfruttando il fatto che la i-esima equazione è accoppiata solo alle precedenti, e non alle successive, al contrario di come capita nei sistemi triangolari superiori.