Calcolo numerico e Matlab

Matrici e sistemi lineari

Claudio Canuto
Dipartimento di Scienze Matematiche - Politecnico di Torino

claudio.canuto@polito.it

Indice

1. Richiami e complementi di algebra lineare

Vettori e di matrici

Un **vettore** $\boldsymbol{x} = (x_i)_{i=1,\ldots,n}$ di ordine n è una tabella di n numeri (reali o complessi) disposti su una riga (*vettore riga*) oppure su una colonna (*vettore colonna*):

$$\boldsymbol{x} = (x_1 \ x_2 \ \cdots \ x_n) \qquad \text{oppure} \qquad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Un **vettore** $\boldsymbol{x} = (x_i)_{i=1,\dots,n}$ di ordine n è una tabella di n numeri (reali o complessi) disposti su una riga (*vettore riga*) oppure su una colonna (*vettore colonna*):

$$\boldsymbol{x} = (x_1 \ x_2 \ \cdots \ x_n) \qquad \text{oppure} \qquad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Se $\boldsymbol{x}$ e $\boldsymbol{y}$ sono vettori entrambi riga o colonna di uguale ordine n, e α, β sono scalari, è definita la combinazione lineare

$$\alpha\boldsymbol{x} + \beta\boldsymbol{y} = (\alpha x_i + \beta y_i)_{i=1,\dots,n}$$

Un **vettore** $\boldsymbol{x} = (x_i)_{i=1,\dots,n}$ di ordine n è una tabella di n numeri (reali o complessi) disposti su una riga (*vettore riga*) oppure su una colonna (*vettore colonna*):

$$\boldsymbol{x} = (x_1 \ x_2 \ \cdots \ x_n) \qquad \text{oppure} \qquad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Se $\boldsymbol{x}$ e $\boldsymbol{y}$ sono vettori entrambi riga o colonna di uguale ordine n, e α, β sono scalari, è definita la combinazione lineare

$$\alpha\boldsymbol{x} + \beta\boldsymbol{y} = (\alpha x_i + \beta y_i)_{i=1,\dots,n}$$

Inoltre se $\boldsymbol{x}$ e $\boldsymbol{y}$ sono vettori reali di uguale ordine n, è definito il *prodotto scalare*

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} x_i y_i \ .$$

Se $\boldsymbol{x}$ è un vettore riga e $\boldsymbol{y}$ è un vettore colonna, il loro prodotto scalare viene indicato semplicemente con $\boldsymbol{x}\boldsymbol{y}$.

Una **matrice** $\boldsymbol{A} = (a_{ij})_{i=1,\dots,m;j=1,\dots n}$ di ordine $m \times n$ è una tabella di $m \times n$ numeri (reali o complessi) disposti su m righe ed n colonne:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Una **matrice** $\boldsymbol{A} = (a_{ij})_{i=1,\ldots,m;j=1,\ldots n}$ di ordine $m \times n$ è una tabella di $m \times n$ numeri (reali o complessi) disposti su m righe ed n colonne:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

- Una matrice dicesi *reale* se tutti i suoi elementi sono reali, e scriviamo $\boldsymbol{A} \in \mathbb{R}^{m \times n}$.

Una **matrice** $\boldsymbol{A} = (a_{ij})_{i=1,\dots,m;j=1,\dots n}$ di ordine $m \times n$ è una tabella di $m \times n$ numeri (reali o complessi) disposti su m righe ed n colonne:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

- Una matrice dicesi *reale* se tutti i suoi elementi sono reali, e scriviamo $\boldsymbol{A} \in \mathbb{R}^{m \times n}$.
- Una matrice dicesi *quadrata* se $m = n$.

Una **matrice** $\boldsymbol{A} = (a_{ij})_{i=1,\dots,m;j=1,\dots n}$ di ordine $m \times n$ è una tabella di $m \times n$ numeri (reali o complessi) disposti su m righe ed n colonne:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

- Una matrice dicesi *reale* se tutti i suoi elementi sono reali, e scriviamo $\boldsymbol{A} \in \mathbb{R}^{m \times n}$.
- Una matrice dicesi *quadrata* se $m = n$.
- La **trasposta** di una matrice reale $\boldsymbol{A}$ di ordine $m \times n$, indicata con $\boldsymbol{A}^T$, è la matrice di ordine $n \times m$ i cui elementi a_{ij}^T soddisfano $a_{ij}^T = a_{ji}$ per ogni i e j.

Una **matrice** $\boldsymbol{A} = (a_{ij})_{i=1,\dots,m;j=1,\dots n}$ di ordine $m \times n$ è una tabella di $m \times n$ numeri (reali o complessi) disposti su m righe ed n colonne:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

- Una matrice dicesi *reale* se tutti i suoi elementi sono reali, e scriviamo $\boldsymbol{A} \in \mathbb{R}^{m \times n}$.
- Una matrice dicesi *quadrata* se $m = n$.
- La **trasposta** di una matrice reale $\boldsymbol{A}$ di ordine $m \times n$, indicata con $\boldsymbol{A}^T$, è la matrice di ordine $n \times m$ i cui elementi a^T_{ij} soddisfano $a^T_{ij} = a_{ji}$ per ogni i e j.
- Una matrice $\boldsymbol{A}$ quadrata reale dicesi *simmetrica* se $\boldsymbol{A}^T = \boldsymbol{A}$, ossia se $a_{ij} = a_{ji}$ per ogni i e j.

Una **matrice** $\boldsymbol{A} = (a_{ij})_{i=1,\ldots,m;j=1,\ldots n}$ di ordine $m \times n$ è una tabella di $m \times n$ numeri (reali o complessi) disposti su m righe ed n colonne:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

- Una matrice dicesi *reale* se tutti i suoi elementi sono reali, e scriviamo $\boldsymbol{A} \in \mathbb{R}^{m\times n}$.
- Una matrice dicesi *quadrata* se $m = n$.
- La **trasposta** di una matrice reale $\boldsymbol{A}$ di ordine $m \times n$, indicata con $\boldsymbol{A}^T$, è la matrice di ordine $n \times m$ i cui elementi a^T_{ij} soddisfano $a^T_{ij} = a_{ji}$ per ogni i e j.
- Una matrice $\boldsymbol{A}$ quadrata reale dicesi *simmetrica* se $\boldsymbol{A}^T = \boldsymbol{A}$, ossia se $a_{ij} = a_{ji}$ per ogni i e j.
- Un vettore riga è una matrice $1 \times n$, mentre un vettore colonna è una matrice $n \times 1$.

Una **matrice** $\boldsymbol{A} = (a_{ij})_{i=1,\dots,m;j=1,\dots n}$ di ordine $m \times n$ è una tabella di $m \times n$ numeri (reali o complessi) disposti su m righe ed n colonne:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

- Una matrice dicesi *reale* se tutti i suoi elementi sono reali, e scriviamo $\boldsymbol{A} \in \mathbb{R}^{m\times n}$.
- Una matrice dicesi *quadrata* se $m = n$.
- La **trasposta** di una matrice reale $\boldsymbol{A}$ di ordine $m \times n$, indicata con $\boldsymbol{A}^T$, è la matrice di ordine $n \times m$ i cui elementi a_{ij}^T soddisfano $a_{ij}^T = a_{ji}$ per ogni i e j.
- Una matrice $\boldsymbol{A}$ quadrata reale dicesi *simmetrica* se $\boldsymbol{A}^T = \boldsymbol{A}$, ossia se $a_{ij} = a_{ji}$ per ogni i e j.

- Un vettore riga è una matrice $1 \times n$, mentre un vettore colonna è una matrice $n \times 1$.
- Se $\boldsymbol{x}$ è un vettore riga, allora $\boldsymbol{x}^T$ è un vettore colonna, e viceversa.

Matrici con strutture particolari

- **Matrice diagonale**: $a_{ij} = 0$ se $i \neq j$

$$\boldsymbol{A} = \begin{pmatrix} * & 0 & 0 & 0 \\ 0 & * & 0 & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \end{pmatrix}$$

Matrici con strutture particolari

- **Matrice diagonale**: $a_{ij} = 0$ se $i \neq j$

$$\boldsymbol{A} = \begin{pmatrix} * & 0 & 0 & 0 \\ 0 & * & 0 & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \end{pmatrix}$$

- **Matrice a banda**: $a_{ij} = 0$ se $|i - j| > m$ (es. **tridiagonale** se $m = 1$)

$$\boldsymbol{A} = \begin{pmatrix} * & * & 0 & 0 \\ * & * & * & 0 \\ 0 & * & * & * \\ 0 & 0 & * & * \end{pmatrix}$$

Matrici con strutture particolari

- **Matrice diagonale**: $a_{ij} = 0$ se $i \neq j$

$$\boldsymbol{A} = \begin{pmatrix} * & 0 & 0 & 0 \\ 0 & * & 0 & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \end{pmatrix}$$

- **Matrice a banda**: $a_{ij} = 0$ se $|i - j| > m$ (es. **tridiagonale** se $m = 1$)

$$\boldsymbol{A} = \begin{pmatrix} * & * & 0 & 0 \\ * & * & * & 0 \\ 0 & * & * & * \\ 0 & 0 & * & * \end{pmatrix}$$

- **Matrice triangolare superiore**: $a_{ij} = 0$ se $i > j$, oppure **inferiore**: $a_{ij} = 0$ se $i < j$

$$\boldsymbol{A} = \begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{pmatrix} \qquad A = \begin{pmatrix} * & 0 & 0 & 0 \\ * & * & 0 & 0 \\ * & * & * & 0 \\ * & * & * & * \end{pmatrix}$$

Indichiamo con $\boldsymbol{a}_{i,:} = (a_{i1} \; a_{i2} \; \cdots \; a_{in})$ la riga i-esima della matrice $\boldsymbol{A}$. Allora possiamo pensare $\boldsymbol{A}$ come un *vettore colonna* di ordine m, i cui elementi sono i vettori riga della matrice:

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_{1,:} \\ \boldsymbol{a}_{2,:} \\ \ldots \\ \boldsymbol{a}_{m,:} \end{pmatrix}.$$

Indichiamo con $\boldsymbol{a}_{i,:} = (a_{i1}\ a_{i2}\ \cdots\ a_{in})$ la riga i-esima della matrice $\boldsymbol{A}$. Allora possiamo pensare $\boldsymbol{A}$ come un *vettore colonna* di ordine m, i cui elementi sono i vettori riga della matrice:

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_{1,:} \\ \boldsymbol{a}_{2,:} \\ \dots \\ \boldsymbol{a}_{m,:} \end{pmatrix}.$$

Simmetricamente, se indichiamo con

$$\boldsymbol{a}_{:,j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{mj} \end{pmatrix}$$

la j-esima colonna della matrice $\boldsymbol{A}$, possiamo pensare $\boldsymbol{A}$ come un *vettore riga* di ordine n, i cui elementi sono i vettori colonna della matrice:

$$\boldsymbol{A} = (\boldsymbol{a}_{:,1}\ \ \boldsymbol{a}_{:,2}\ \ \cdots\ \ \boldsymbol{a}_{:,n}).$$

Se $\boldsymbol{x}$ è un vettore colonna di ordine n, il prodotto $\boldsymbol{Ax}$ è il vettore colonna di ordine m definito come

$$\boldsymbol{Ax} = \begin{pmatrix} \boldsymbol{a}_{1,:}\,\boldsymbol{x} \\ \boldsymbol{a}_{2,:}\,\boldsymbol{x} \\ \dots \\ \boldsymbol{a}_{m,:}\,\boldsymbol{x} \end{pmatrix} = \left(\sum_{j=1}^{n} a_{ij} x_j \right)_{i=1,\dots,m} .$$

Se $\boldsymbol{x}$ è un vettore colonna di ordine n, il prodotto $\boldsymbol{A}\boldsymbol{x}$ è il vettore colonna di ordine m definito come

$$\boldsymbol{A}\boldsymbol{x} = \begin{pmatrix} \boldsymbol{a}_{1,:}\,\boldsymbol{x} \\ \boldsymbol{a}_{2,:}\,\boldsymbol{x} \\ \cdots \\ \boldsymbol{a}_{m,:}\,\boldsymbol{x} \end{pmatrix} = \left(\sum_{j=1}^{n} a_{ij} x_j \right)_{i=1,\dots,m} .$$

Equivalentemente, $\boldsymbol{A}\boldsymbol{x}$ è la combinazione lineare delle colonne di $\boldsymbol{A}$ avente come coefficienti le componenti di $\boldsymbol{x}$, ossia

$$\boldsymbol{A}\boldsymbol{x} = x_1 \boldsymbol{a}_{:,1} + x_2 \boldsymbol{a}_{:,2} + \cdots + x_n \boldsymbol{a}_{:,n} = \sum_{j=1}^{n} x_j \boldsymbol{a}_{:,j} \ .$$

Se $\boldsymbol{x}$ è un vettore colonna di ordine n, il prodotto $\boldsymbol{Ax}$ è il vettore colonna di ordine m definito come

$$\boldsymbol{Ax} = \begin{pmatrix} \boldsymbol{a}_{1,:}\,\boldsymbol{x} \\ \boldsymbol{a}_{2,:}\,\boldsymbol{x} \\ \dots \\ \boldsymbol{a}_{m,:}\,\boldsymbol{x} \end{pmatrix} = \left(\sum_{j=1}^{n} a_{ij} x_j \right)_{i=1,\dots,m} .$$

Equivalentemente, $\boldsymbol{Ax}$ è la combinazione lineare delle colonne di $\boldsymbol{A}$ avente come coefficienti le componenti di $\boldsymbol{x}$, ossia

$$\boldsymbol{Ax} = x_1 \boldsymbol{a}_{:,1} + x_2 \boldsymbol{a}_{:,2} + \cdots + x_n \boldsymbol{a}_{:,n} = \sum_{j=1}^{n} x_j \boldsymbol{a}_{:,j} .$$

Se $\boldsymbol{A}$ è una matrice di ordine $m \times n$ e $\boldsymbol{B}$ è una matrice di ordine $n \times p$, allora è definito il prodotto $\boldsymbol{AB}$ come la matrice di ordine $m \times p$ le cui colonne sono i prodotti $\boldsymbol{Ab}_{:,k}$ per $k = 1, \dots, p$, ossia

$$\boldsymbol{AB} = (\boldsymbol{A}\,\boldsymbol{b}_{:,1} \quad \boldsymbol{A}\,\boldsymbol{b}_{:,2} \quad \cdots \quad \boldsymbol{A}\,\boldsymbol{b}_{:,p}) = \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right)_{i=1,\dots,m;k=1,\dots,p} .$$

Se $\boldsymbol{x}$ è un vettore colonna di ordine n, il prodotto $\boldsymbol{Ax}$ è il vettore colonna di ordine m definito come

$$\boldsymbol{Ax} = \begin{pmatrix} \boldsymbol{a}_{1,:}\,\boldsymbol{x} \\ \boldsymbol{a}_{2,:}\,\boldsymbol{x} \\ \dots \\ \boldsymbol{a}_{m,:}\,\boldsymbol{x} \end{pmatrix} = \left(\sum_{j=1}^{n} a_{ij} x_j \right)_{i=1,\dots,m} .$$

Equivalentemente, $\boldsymbol{Ax}$ è la combinazione lineare delle colonne di $\boldsymbol{A}$ avente come coefficienti le componenti di $\boldsymbol{x}$, ossia

$$\boldsymbol{Ax} = x_1 \boldsymbol{a}_{:,1} + x_2 \boldsymbol{a}_{:,2} + \cdots + x_n \boldsymbol{a}_{:,n} = \sum_{j=1}^{n} x_j \boldsymbol{a}_{:,j} \ .$$

Se $\boldsymbol{A}$ è una matrice di ordine $m \times n$ e $\boldsymbol{B}$ è una matrice di ordine $n \times p$, allora è definito il prodotto $\boldsymbol{AB}$ come la matrice di ordine $m \times p$ le cui colonne sono i prodotti $\boldsymbol{Ab}_{:,k}$ per $k = 1, \dots, p$, ossia

$$\boldsymbol{AB} = (\boldsymbol{A}\,\boldsymbol{b}_{:,1} \quad \boldsymbol{A}\,\boldsymbol{b}_{:,2} \quad \cdots \quad \boldsymbol{A}\,\boldsymbol{b}_{:,p}) = \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right)_{i=1,\dots,m;k=1,\dots,p} .$$

In generale, anche quando $m = n = p$ il prodotto non è commutativo: $\boldsymbol{AB} \neq \boldsymbol{BA}$.

Il determinante $\det \boldsymbol{A}$ di una matrice quadrata $\boldsymbol{A}$ è definito ricorsivamente sull'ordine della matrice.

- Se $\boldsymbol{A} = (a)$ ha ordine 1, si pone $\det \boldsymbol{A} = a$.

Determinante di una matrice quadrata

Il determinante $\det \boldsymbol{A}$ di una matrice quadrata $\boldsymbol{A}$ è definito ricorsivamente sull'ordine della matrice.

- Se $\boldsymbol{A} = (a)$ ha ordine 1, si pone $\det \boldsymbol{A} = a$.
- Sia $\boldsymbol{A}$ di ordine $n > 1$. Indichiamo con $\boldsymbol{A}_{ij}$ la sotto-matrice di $\boldsymbol{A}$ di ordine $n-1$ ottenuta cancellando la riga i-esima e la colonna j-esima di $\boldsymbol{A}$. Si può porre allora

$$\det \boldsymbol{A} = a_{11} \det \boldsymbol{A}_{11} - a_{12} \det \boldsymbol{A}_{12} + \cdots + (-1)^{1+n} a_{1n} \det \boldsymbol{A}_{1n} = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det \boldsymbol{A}_{1j}$$

supponendo di avere gia' calcolato i determinanti delle sotto-matrici.

Determinante di una matrice quadrata

Il determinante $\det \boldsymbol{A}$ di una matrice quadrata $\boldsymbol{A}$ è definito ricorsivamente sull'ordine della matrice.

- Se $\boldsymbol{A} = (a)$ ha ordine 1, si pone $\det \boldsymbol{A} = a$.
- Sia $\boldsymbol{A}$ di ordine $n > 1$. Indichiamo con $\boldsymbol{A}_{ij}$ la sotto-matrice di $\boldsymbol{A}$ di ordine $n-1$ ottenuta cancellando la riga i-esima e la colonna j-esima di $\boldsymbol{A}$. Si può porre allora

$$\det \boldsymbol{A} = a_{11}\det \boldsymbol{A}_{11} - a_{12}\det \boldsymbol{A}_{12} + \cdots + (-1)^{1+n} a_{1n}\det \boldsymbol{A}_{1n} = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det \boldsymbol{A}_{1j}$$

supponendo di avere gia' calcolato i determinanti delle sotto-matrici.
Questo è lo sviluppo del determinante *rispetto alla prima riga*. È possibile svilupparlo rispetto a una qualunque riga o colonna, cioe' si ha, per ogni i oppure per ogni j

$$\det \boldsymbol{A} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det \boldsymbol{A}_{ij} \ , \qquad \text{oppure} \qquad \det \boldsymbol{A} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det \boldsymbol{A}_{ij} \ .$$

Determinante di una matrice quadrata

Il determinante $\det \boldsymbol{A}$ di una matrice quadrata $\boldsymbol{A}$ è definito ricorsivamente sull'ordine della matrice.

- Se $\boldsymbol{A} = (a)$ ha ordine 1, si pone $\det \boldsymbol{A} = a$.
- Sia $\boldsymbol{A}$ di ordine $n > 1$. Indichiamo con $\boldsymbol{A}_{ij}$ la sotto-matrice di $\boldsymbol{A}$ di ordine $n-1$ ottenuta cancellando la riga i-esima e la colonna j-esima di $\boldsymbol{A}$. Si può porre allora

$$\det \boldsymbol{A} = a_{11}\det \boldsymbol{A}_{11} - a_{12}\det \boldsymbol{A}_{12} + \cdots + (-1)^{1+n} a_{1n}\det \boldsymbol{A}_{1n} = \sum_{j=1}^{n} (-1)^{1+j} a_{1j}\det \boldsymbol{A}_{1j}$$

  supponendo di avere gia' calcolato i determinanti delle sotto-matrici.
  Questo è lo sviluppo del determinante *rispetto alla prima riga*. È possibile svilupparlo rispetto a una qualunque riga o colonna, cioe' si ha, per ogni i oppure per ogni j

$$\det \boldsymbol{A} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij}\det \boldsymbol{A}_{ij} \,, \qquad \text{oppure} \qquad \det \boldsymbol{A} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij}\det \boldsymbol{A}_{ij} \,.$$

Il determinante gode delle seguenti proprietà:

- $\det \boldsymbol{A} = \det \boldsymbol{A}^T$
- $\det \boldsymbol{AB} = \det \boldsymbol{A} \det \boldsymbol{B}$
- $\det\left(\mathrm{diag}\,(a_1, a_2, \ldots, a_n)\right) = \prod_{i=1}^{n} a_i$
- $\det \boldsymbol{A}^{-1} = (\det \boldsymbol{A})^{-1}$

Sia $\boldsymbol{A}$ una matrice quadrata di ordine n. Le seguenti definizioni sono equivalenti.

Sia $\boldsymbol{A}$ una matrice quadrata di ordine n. Le seguenti definizioni sono equivalenti.

- **Definizione algebrica di autovalore.** Sia $p_{\boldsymbol{A}}(z) = \det(\boldsymbol{A} - z\boldsymbol{I})$ il *polinomio caratteristico* di $\boldsymbol{A}$, che è un polinomio di grado n nella variabile $z \in \mathbb{C}$. Ogni radice di tale polinomio, ossia ogni numero λ tale che

$$p_{\boldsymbol{A}}(\lambda) = \det(\boldsymbol{A} - \lambda\boldsymbol{I}) = 0,$$

è detto **autovalore** di $\boldsymbol{A}$. Indichiamo con $m_a(\lambda)$ la *molteplicità algebrica* della radice λ.

Sia $\boldsymbol{A}$ una matrice quadrata di ordine n. Le seguenti definizioni sono equivalenti.

- **Definizione algebrica di autovalore.** Sia $p_{\boldsymbol{A}}(z) = \det(\boldsymbol{A} - z\boldsymbol{I})$ il *polinomio caratteristico* di $\boldsymbol{A}$, che è un polinomio di grado n nella variabile $z \in \mathbb{C}$. Ogni radice di tale polinomio, ossia ogni numero λ tale che

$$p_{\boldsymbol{A}}(\lambda) = \det(\boldsymbol{A} - \lambda\boldsymbol{I}) = 0,$$

è detto **autovalore** di $\boldsymbol{A}$. Indichiamo con $m_a(\lambda)$ la *molteplicità algebrica* della radice λ.
Per il Teorema fondamentale dell'Algebra, $\boldsymbol{A}$ possiede n autovalori in campo complesso, contati con le rispettive molteplicità algebriche.

Sia $\boldsymbol{A}$ una matrice quadrata di ordine n. Le seguenti definizioni sono equivalenti.

- **Definizione algebrica di autovalore.** Sia $p_{\boldsymbol{A}}(z) = \det(\boldsymbol{A} - z\boldsymbol{I})$ il *polinomio caratteristico* di $\boldsymbol{A}$, che è un polinomio di grado n nella variabile $z \in \mathbb{C}$. Ogni radice di tale polinomio, ossia ogni numero λ tale che

  $$p_{\boldsymbol{A}}(\lambda) = \det(\boldsymbol{A} - \lambda\boldsymbol{I}) = 0,$$

  è detto **autovalore** di $\boldsymbol{A}$. Indichiamo con $m_a(\lambda)$ la *molteplicità algebrica* della radice λ.
  Per il Teorema fondamentale dell'Algebra, $\boldsymbol{A}$ possiede n autovalori in campo complesso, contati con le rispettive molteplicità algebriche.

- **Definizione geometrica di autovalore.** Un numero $\lambda \in \mathbb{C}$ per il quale esista un vettore $\boldsymbol{w}$ *non nullo* tale che

  $$\boldsymbol{A}\boldsymbol{w} = \lambda\boldsymbol{w}$$

  dicesi **autovalore** di $\boldsymbol{A}$, e $\boldsymbol{w}$ è il corrispondente **autovettore**. Il numero di autovettori linearmente indipendenti di λ è la *molteplicità geometrica* di λ, indicata con $m_g(\lambda)$.

Per ogni autovalore λ di $\boldsymbol{A}$, si ha $m_g(\lambda) \leq m_a(\lambda)$.

Per ogni autovalore λ di $\boldsymbol{A}$, si ha $m_g(\lambda) \leq m_a(\lambda)$.

La condizione che per ogni autovalore si abbia $m_g(\lambda) = m_a(\lambda)$ equivale all'esistenza di n autovettori linearmente indipendenti. In tal caso la matrice $\boldsymbol{A}$ si dice **diagonalizzabile**, per il motivo di seguito illustrato.

Per ogni autovalore λ di $\boldsymbol{A}$, si ha $m_g(\lambda) \leq m_a(\lambda)$.

La condizione che per ogni autovalore si abbia $m_g(\lambda) = m_a(\lambda)$ equivale all'esistenza di n autovettori linearmente indipendenti. In tal caso la matrice $\boldsymbol{A}$ si dice **diagonalizzabile**, per il motivo di seguito illustrato.

Se indichiamo con $\lambda_1, \lambda_2, \ldots, \lambda_n$ gli autovalori di $\boldsymbol{A}$ ripetuti secondo la loro molteplicità, e con $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ i corrispondenti autovettori, si ha

$$\boldsymbol{A}\boldsymbol{w}_k = \lambda_k \boldsymbol{w}_k, \qquad k = 1, 2, \ldots, n.$$

Tale relazione può essere scritta in forma compatta matriciale come

$$\boldsymbol{A}\boldsymbol{W} = \boldsymbol{W}\boldsymbol{\Lambda}\ ,$$

avendo introdotto la matrice $\boldsymbol{W} = (\boldsymbol{w}_1\ \ \boldsymbol{w}_2\ \ \cdots\ \ \boldsymbol{w}_n)$ le cui colonne sono gli autovettori, e la matrice diagonale $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ avente gli autovalori sulla diagonale principale.

Per ogni autovalore λ di $\boldsymbol{A}$, si ha $m_g(\lambda) \leq m_a(\lambda)$.

La condizione che per ogni autovalore si abbia $m_g(\lambda) = m_a(\lambda)$ equivale all'esistenza di n autovettori linearmente indipendenti. In tal caso la matrice $\boldsymbol{A}$ si dice **diagonalizzabile**, per il motivo di seguito illustrato.

Se indichiamo con $\lambda_1, \lambda_2, \ldots, \lambda_n$ gli autovalori di $\boldsymbol{A}$ ripetuti secondo la loro molteplicità, e con $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ i corrispondenti autovettori, si ha

$$\boldsymbol{A}\boldsymbol{w}_k = \lambda_k \boldsymbol{w}_k, \qquad k = 1, 2, \ldots, n.$$

Tale relazione può essere scritta in forma compatta matriciale come

$$\boldsymbol{A}\boldsymbol{W} = \boldsymbol{W}\boldsymbol{\Lambda}\ ,$$

avendo introdotto la matrice $\boldsymbol{W} = (\boldsymbol{w}_1 \ \ \boldsymbol{w}_2 \ \ \cdots \ \ \boldsymbol{w}_n)$ le cui colonne sono gli autovettori, e la matrice diagonale $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ avente gli autovalori sulla diagonale principale.
Se gli autovettori sono linearmente indipendenti, allora la matrice $\boldsymbol{W}$ è non-singolare e dunque invertibile. Possiamo dunque scrivere la relazione precedente come

$$\boldsymbol{A} = \boldsymbol{W}\boldsymbol{\Lambda}\boldsymbol{W}^{-1} \qquad \text{oppure} \qquad \boldsymbol{\Lambda} = \boldsymbol{W}^{-1}\boldsymbol{A}\boldsymbol{W}$$

il che mostra che la matrice $\boldsymbol{A}$ può essere trasformata nella matrice diagonale $\boldsymbol{\Lambda}$ attraverso la trasformazione di similitudine associata alla matrice $\boldsymbol{W}$.

Matrici reali simmetriche

Se $\boldsymbol{A}$ è una matrice reale simmetrica di ordine n, allora

- tutti gli autovalori sono reali
- la matrice è diagonalizzabile, e precisamente ammette n autovettori ortogonali tra loro.

Matrici reali simmetriche

Se $\boldsymbol{A}$ è una matrice reale simmetrica di ordine n, allora

- tutti gli autovalori sono reali
- la matrice è diagonalizzabile, e precisamente ammette n autovettori ortogonali tra loro.

Normalizzando gli autovettori, abbiamo dunque n autovettori $\boldsymbol{w}_k$, $k = 1, \dots, n$, che soddisfano le n^2 relazioni di ortonormalità

$$(\boldsymbol{w}_k, \boldsymbol{w}_\ell) = \boldsymbol{w}_k^T \boldsymbol{w}_\ell = \delta_{k,\ell}, \qquad 1 \le k, \ell \le n,$$

che equivalgono alla relazione matriciale

$$\boldsymbol{W}^T \boldsymbol{W} = \boldsymbol{I}.$$

Matrici reali simmetriche

Se $\boldsymbol{A}$ è una matrice reale simmetrica di ordine n, allora

- tutti gli autovalori sono reali
- la matrice è diagonalizzabile, e precisamente ammette n autovettori ortogonali tra loro.

Normalizzando gli autovettori, abbiamo dunque n autovettori $\boldsymbol{w}_k$, $k = 1, \dots, n$, che soddisfano le n^2 relazioni di ortonormalità

$$(\boldsymbol{w}_k, \boldsymbol{w}_\ell) = \boldsymbol{w}_k^T \boldsymbol{w}_\ell = \delta_{k,\ell}, \qquad 1 \leq k, \ell \leq n,$$

che equivalgono alla relazione matriciale

$$\boldsymbol{W}^T \boldsymbol{W} = \boldsymbol{I}.$$

Tale relazione caratterizza le cosiddette **matrici ortogonali**. Per esse si ha $\boldsymbol{W}^{-1} = \boldsymbol{W}^T$, e dunque la diagonalizzazione di $\boldsymbol{A}$ si esprime come

$$\boldsymbol{A} = \boldsymbol{W} \boldsymbol{\Lambda} \boldsymbol{W}^T .$$

Dunque gli autovettori di $\boldsymbol{A}$ formano una *base ortonormale* di $\mathbb{R}^n$. Ogni vettore $\boldsymbol{x} \in \mathbb{R}^n$ si esprime come

$$\boldsymbol{x} = \sum_{k=1}^{n} \alpha_k \boldsymbol{w}_k \qquad \text{con} \qquad \alpha_k = (\boldsymbol{x}, \boldsymbol{w}_k) = \boldsymbol{x}^T \boldsymbol{w}_k.$$

Dunque gli autovettori di $\boldsymbol{A}$ formano una *base ortonormale* di $\mathbb{R}^n$. Ogni vettore $\boldsymbol{x} \in \mathbb{R}^n$ si esprime come

$$\boldsymbol{x} = \sum_{k=1}^{n} \alpha_k \boldsymbol{w}_k \qquad \text{con} \qquad \alpha_k = (\boldsymbol{x}, \boldsymbol{w}_k) = \boldsymbol{x}^T \boldsymbol{w}_k .$$

Inoltre, si ha immediatamente

$$\boldsymbol{A}\boldsymbol{x} = \sum_{k=1}^{n} \alpha_k \lambda_k \boldsymbol{w}_k \ .$$

Dunque gli autovettori di $\boldsymbol{A}$ formano una *base ortonormale* di $\mathbb{R}^n$. Ogni vettore $\boldsymbol{x} \in \mathbb{R}^n$ si esprime come

$$\boldsymbol{x} = \sum_{k=1}^{n} \alpha_k \boldsymbol{w}_k \qquad \text{con} \qquad \alpha_k = (\boldsymbol{x}, \boldsymbol{w}_k) = \boldsymbol{x}^T \boldsymbol{w}_k .$$

Inoltre, si ha immediatamente

$$\boldsymbol{A}\boldsymbol{x} = \sum_{k=1}^{n} \alpha_k \lambda_k \boldsymbol{w}_k \ .$$

Matrici simmetriche e definite positive. Ecco una classe particolarmente importante di matrici *reali simmetriche*: una tale matrice $\boldsymbol{A}$ si dice **definita positiva** se vale una delle seguenti condizioni, tra loro equivalenti:

Dunque gli autovettori di $\boldsymbol{A}$ formano una *base ortonormale* di $\mathbb{R}^n$. Ogni vettore $\boldsymbol{x} \in \mathbb{R}^n$ si esprime come

$$\boldsymbol{x} = \sum_{k=1}^{n} \alpha_k \boldsymbol{w}_k \qquad \text{con} \qquad \alpha_k = (\boldsymbol{x}, \boldsymbol{w}_k) = \boldsymbol{x}^T \boldsymbol{w}_k .$$

Inoltre, si ha immediatamente

$$\boldsymbol{A}\boldsymbol{x} = \sum_{k=1}^{n} \alpha_k \lambda_k \boldsymbol{w}_k \ .$$

Matrici simmetriche e definite positive. Ecco una classe particolarmente importante di matrici *reali simmetriche*: una tale matrice $\boldsymbol{A}$ si dice **definita positiva** se vale una delle seguenti condizioni, tra loro equivalenti:

- la *forma quadratica* $Q(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$ associata ad $\boldsymbol{A}$ soddisfa $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} > 0$ per ogni vettore $\boldsymbol{x} \neq \boldsymbol{0}$;

Dunque gli autovettori di $\boldsymbol{A}$ formano una *base ortonormale* di $\mathbb{R}^n$. Ogni vettore $\boldsymbol{x} \in \mathbb{R}^n$ si esprime come

$$\boldsymbol{x} = \sum_{k=1}^{n} \alpha_k \boldsymbol{w}_k \qquad \text{con} \qquad \alpha_k = (\boldsymbol{x}, \boldsymbol{w}_k) = \boldsymbol{x}^T \boldsymbol{w}_k .$$

Inoltre, si ha immediatamente

$$\boldsymbol{A}\boldsymbol{x} = \sum_{k=1}^{n} \alpha_k \lambda_k \boldsymbol{w}_k \ .$$

Matrici simmetriche e definite positive. Ecco una classe particolarmente importante di matrici *reali simmetriche*: una tale matrice $\boldsymbol{A}$ si dice **definita positiva** se vale una delle seguenti condizioni, tra loro equivalenti:

- la *forma quadratica* $Q(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$ associata ad $\boldsymbol{A}$ soddisfa $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} > 0$ per ogni vettore $\boldsymbol{x} \neq \boldsymbol{0}$;
- tutti gli autovalori di $\boldsymbol{A}$ sono > 0.

Dunque gli autovettori di $\boldsymbol{A}$ formano una *base ortonormale* di $\mathbb{R}^n$. Ogni vettore $\boldsymbol{x} \in \mathbb{R}^n$ si esprime come

$$\boldsymbol{x} = \sum_{k=1}^{n} \alpha_k \boldsymbol{w}_k \qquad \text{con} \qquad \alpha_k = (\boldsymbol{x}, \boldsymbol{w}_k) = \boldsymbol{x}^T \boldsymbol{w}_k .$$

Inoltre, si ha immediatamente

$$\boldsymbol{A}\boldsymbol{x} = \sum_{k=1}^{n} \alpha_k \lambda_k \boldsymbol{w}_k \ .$$

Matrici simmetriche e definite positive. Ecco una classe particolarmente importante di matrici *reali simmetriche*: una tale matrice $\boldsymbol{A}$ si dice **definita positiva** se vale una delle seguenti condizioni, tra loro equivalenti:

- la *forma quadratica* $Q(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$ associata ad $\boldsymbol{A}$ soddisfa $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} > 0$ per ogni vettore $\boldsymbol{x} \neq \boldsymbol{0}$;
- tutti gli autovalori di $\boldsymbol{A}$ sono > 0.

L'equivalenza è facilmente dimostrabile osservando che dall'espressione di $\boldsymbol{A}\boldsymbol{x}$ si ha immediatamente

$$\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} = \sum_{k=1}^{n} \alpha_k^2 \lambda_k \ .$$

- Per una qualunque matrice quadrata $\boldsymbol{A}$, il determinante è il prodotto degli autovalori:

$$\det \boldsymbol{A} = \prod_{k=1}^{n} \lambda_k .$$

È immediato verificare ciò per una matrice diangonalizzabile, in quanto dalla relazione $\boldsymbol{A} = \boldsymbol{W}\boldsymbol{\Lambda}\boldsymbol{W}^{-1}$ otteniamo

$$\det \boldsymbol{A} = \det \boldsymbol{W} \ \det \boldsymbol{\Lambda} \ \det \boldsymbol{W}^{-1} = \det \boldsymbol{\Lambda} = \prod_{k=1}^{n} \lambda_k .$$

- Per una qualunque matrice quadrata $\boldsymbol{A}$, il determinante è il prodotto degli autovalori:

$$\det \boldsymbol{A} = \prod_{k=1}^{n} \lambda_k .$$

È immediato verificare ciò per una matrice diangonalizzabile, in quanto dalla relazione $\boldsymbol{A} = \boldsymbol{W}\boldsymbol{\Lambda}\boldsymbol{W}^{-1}$ otteniamo

$$\det \boldsymbol{A} = \det \boldsymbol{W} \; \det \boldsymbol{\Lambda} \; \det \boldsymbol{W}^{-1} = \det \boldsymbol{\Lambda} = \prod_{k=1}^{n} \lambda_k .$$

- Per una matrice quadrata *triangolare* (superiore o inferiore) $\boldsymbol{A}$, gli autovalori sono dati dagli elementi diagonali, cioè si ha $\lambda_k = a_{kk}$ per ogni $k = 1, \ldots, n$. Infatti si ha

$$\det (\boldsymbol{A} - \lambda \boldsymbol{I}) = \prod_{k=1}^{n} (a_{kk} - \lambda) .$$

Matrice non-singolare (o invertibile)

Una matrice reale quadrata $\boldsymbol{A}$ di ordine n si dice **non-singolare** oppure **invertibile** se gode di una delle seguenti proprietà, tra di loro equivalenti:

Una matrice reale quadrata $\boldsymbol{A}$ di ordine n si dice **non-singolare** oppure **invertibile** se gode di una delle seguenti proprietà, tra di loro equivalenti:

1. Le colonne di $\boldsymbol{A}$ sono vettori linearmente indipendenti.

Matrice non-singolare (o invertibile)

Una matrice reale quadrata $\boldsymbol{A}$ di ordine n si dice **non-singolare** oppure **invertibile** se gode di una delle seguenti proprietà, tra di loro equivalenti:

1. Le colonne di $\boldsymbol{A}$ sono vettori linearmente indipendenti.
2. Il sistema omogeneo $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ ammette solo la soluzione nulla $\boldsymbol{x} = \boldsymbol{0}$.

Una matrice reale quadrata $\boldsymbol{A}$ di ordine n si dice **non-singolare** oppure **invertibile** se gode di una delle seguenti proprietà, tra di loro equivalenti:

1. Le colonne di $\boldsymbol{A}$ sono vettori linearmente indipendenti.
2. Il sistema omogeneo $\boldsymbol{Ax} = \boldsymbol{0}$ ammette solo la soluzione nulla $\boldsymbol{x} = \boldsymbol{0}$.
3. Il nucleo di $\boldsymbol{A}$, $\ker \boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^n \,:\, \boldsymbol{Ax} = \boldsymbol{0}\}$, contiene solo il vettore nullo.

Matrice non-singolare (o invertibile)

Una matrice reale quadrata $\boldsymbol{A}$ di ordine n si dice **non-singolare** oppure **invertibile** se gode di una delle seguenti proprietà, tra di loro equivalenti:

1. Le colonne di $\boldsymbol{A}$ sono vettori linearmente indipendenti.
2. Il sistema omogeneo $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ ammette solo la soluzione nulla $\boldsymbol{x} = \boldsymbol{0}$.
3. Il nucleo di $\boldsymbol{A}$, $\ker \boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^n \ : \ \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$, contiene solo il vettore nullo.
4. Il sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ ammette una soluzione, per ogni $\boldsymbol{b} \in \mathbb{R}^n$.

Matrice non-singolare (o invertibile)

Una matrice reale quadrata $\boldsymbol{A}$ di ordine n si dice **non-singolare** oppure **invertibile** se gode di una delle seguenti proprietà, tra di loro equivalenti:

1. Le colonne di $\boldsymbol{A}$ sono vettori linearmente indipendenti.
2. Il sistema omogeneo $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ ammette solo la soluzione nulla $\boldsymbol{x} = \boldsymbol{0}$.
3. Il nucleo di $\boldsymbol{A}$, $\ker \boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^n \, : \, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$, contiene solo il vettore nullo.
4. Il sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ ammette una soluzione, per ogni $\boldsymbol{b} \in \mathbb{R}^n$.
5. L'immagine di $\boldsymbol{A}$, $\mathrm{Im}\, \boldsymbol{A} = \{\boldsymbol{A}\boldsymbol{x} \, : \, \boldsymbol{x} \in \mathbb{R}^n\}$, è tutto $\mathbb{R}^n$.

Matrice non-singolare (o invertibile)

Una matrice reale quadrata $\boldsymbol{A}$ di ordine n si dice **non-singolare** oppure **invertibile** se gode di una delle seguenti proprietà, tra di loro equivalenti:

1. Le colonne di $\boldsymbol{A}$ sono vettori linearmente indipendenti.
2. Il sistema omogeneo $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ ammette solo la soluzione nulla $\boldsymbol{x} = \boldsymbol{0}$.
3. Il nucleo di $\boldsymbol{A}$, $\ker \boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^n \,:\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$, contiene solo il vettore nullo.
4. Il sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ ammette una soluzione, per ogni $\boldsymbol{b} \in \mathbb{R}^n$.
5. L'immagine di $\boldsymbol{A}$, $\mathrm{Im}\, \boldsymbol{A} = \{\boldsymbol{A}\boldsymbol{x} \,:\, \boldsymbol{x} \in \mathbb{R}^n\}$, è tutto $\mathbb{R}^n$.
6. Il rango di $\boldsymbol{A}$, $\mathrm{rank}(\boldsymbol{A}) = \dim(\mathrm{Im}\, \boldsymbol{A})$, è uguale a n (cioè $\boldsymbol{A}$ ha rango massimo).

Una matrice reale quadrata $\boldsymbol{A}$ di ordine n si dice **non-singolare** oppure **invertibile** se gode di una delle seguenti proprietà, tra di loro equivalenti:

1. Le colonne di $\boldsymbol{A}$ sono vettori linearmente indipendenti.
2. Il sistema omogeneo $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ ammette solo la soluzione nulla $\boldsymbol{x} = \boldsymbol{0}$.
3. Il nucleo di $\boldsymbol{A}$, $\ker \boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^n \,:\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$, contiene solo il vettore nullo.
4. Il sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ ammette una soluzione, per ogni $\boldsymbol{b} \in \mathbb{R}^n$.
5. L'immagine di $\boldsymbol{A}$, $\mathrm{Im}\, \boldsymbol{A} = \{\boldsymbol{A}\boldsymbol{x} \,:\, \boldsymbol{x} \in \mathbb{R}^n\}$, è tutto $\mathbb{R}^n$.
6. Il rango di $\boldsymbol{A}$, $\mathrm{rank}(\boldsymbol{A}) = \dim(\mathrm{Im}\, \boldsymbol{A})$, è uguale a n (cioè $\boldsymbol{A}$ ha rango massimo).
7. Esiste una matrice quadrata $\boldsymbol{B}$ tale che $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A} = \boldsymbol{I}$. Tale matrice è unica e viene indicata con $\boldsymbol{A}^{-1}$ (e detta la *matrice inversa* di $\boldsymbol{A}$).

Una matrice reale quadrata $\boldsymbol{A}$ di ordine n si dice **non-singolare** oppure **invertibile** se gode di una delle seguenti proprietà, tra di loro equivalenti:

1. Le colonne di $\boldsymbol{A}$ sono vettori linearmente indipendenti.
2. Il sistema omogeneo $\boldsymbol{Ax} = \boldsymbol{0}$ ammette solo la soluzione nulla $\boldsymbol{x} = \boldsymbol{0}$.
3. Il nucleo di $\boldsymbol{A}$, $\ker \boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^n \, : \, \boldsymbol{Ax} = \boldsymbol{0}\}$, contiene solo il vettore nullo.
4. Il sistema $\boldsymbol{Ax} = \boldsymbol{b}$ ammette una soluzione, per ogni $\boldsymbol{b} \in \mathbb{R}^n$.
5. L'immagine di $\boldsymbol{A}$, $\mathrm{Im}\, \boldsymbol{A} = \{\boldsymbol{Ax} \, : \, \boldsymbol{x} \in \mathbb{R}^n\}$, è tutto $\mathbb{R}^n$.
6. Il rango di $\boldsymbol{A}$, $\mathrm{rank}(\boldsymbol{A}) = \dim(\mathrm{Im}\, \boldsymbol{A})$, è uguale a n (cioè $\boldsymbol{A}$ ha rango massimo).
7. Esiste una matrice quadrata $\boldsymbol{B}$ tale che $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$. Tale matrice è unica e viene indicata con $\boldsymbol{A}^{-1}$ (e detta la *matrice inversa* di $\boldsymbol{A}$).
8. $\det \boldsymbol{A} \neq 0$.

Matrice non-singolare (o invertibile)

Una matrice reale quadrata $\boldsymbol{A}$ di ordine n si dice **non-singolare** oppure **invertibile** se gode di una delle seguenti proprietà, tra di loro equivalenti:

1. Le colonne di $\boldsymbol{A}$ sono vettori linearmente indipendenti.
2. Il sistema omogeneo $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ ammette solo la soluzione nulla $\boldsymbol{x} = \boldsymbol{0}$.
3. Il nucleo di $\boldsymbol{A}$, $\ker \boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^n \, : \, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$, contiene solo il vettore nullo.
4. Il sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ ammette una soluzione, per ogni $\boldsymbol{b} \in \mathbb{R}^n$.
5. L'immagine di $\boldsymbol{A}$, $\mathrm{Im}\, \boldsymbol{A} = \{\boldsymbol{A}\boldsymbol{x} \, : \, \boldsymbol{x} \in \mathbb{R}^n\}$, è tutto $\mathbb{R}^n$.
6. Il rango di $\boldsymbol{A}$, $\mathrm{rank}(\boldsymbol{A}) = \dim(\mathrm{Im}\, \boldsymbol{A})$, è uguale a n (cioè $\boldsymbol{A}$ ha rango massimo).
7. Esiste una matrice quadrata $\boldsymbol{B}$ tale che $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A} = \boldsymbol{I}$. Tale matrice è unica e viene indicata con $\boldsymbol{A}^{-1}$ (e detta la *matrice inversa* di $\boldsymbol{A}$).
8. $\det \boldsymbol{A} \neq 0$.
9. Tutti gli autovalori di $\boldsymbol{A}$ sono $\neq 0$.

Matrice non-singolare (o invertibile)

Una matrice reale quadrata $\boldsymbol{A}$ di ordine n si dice **non-singolare** oppure **invertibile** se gode di una delle seguenti proprietà, tra di loro equivalenti:

1. Le colonne di $\boldsymbol{A}$ sono vettori linearmente indipendenti.
2. Il sistema omogeneo $\boldsymbol{Ax} = \boldsymbol{0}$ ammette solo la soluzione nulla $\boldsymbol{x} = \boldsymbol{0}$.
3. Il nucleo di $\boldsymbol{A}$, $\ker \boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^n \, : \, \boldsymbol{Ax} = \boldsymbol{0}\}$, contiene solo il vettore nullo.
4. Il sistema $\boldsymbol{Ax} = \boldsymbol{b}$ ammette una soluzione, per ogni $\boldsymbol{b} \in \mathbb{R}^n$.
5. L'immagine di $\boldsymbol{A}$, $\mathrm{Im}\, \boldsymbol{A} = \{\boldsymbol{Ax} \, : \, \boldsymbol{x} \in \mathbb{R}^n\}$, è tutto $\mathbb{R}^n$.
6. Il rango di $\boldsymbol{A}$, $\mathrm{rank}(\boldsymbol{A}) = \dim(\mathrm{Im}\, \boldsymbol{A})$, è uguale a n (cioè $\boldsymbol{A}$ ha rango massimo).
7. Esiste una matrice quadrata $\boldsymbol{B}$ tale che $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$. Tale matrice è unica e viene indicata con $\boldsymbol{A}^{-1}$ (e detta la *matrice inversa* di $\boldsymbol{A}$).
8. $\det \boldsymbol{A} \neq 0$.
9. Tutti gli autovalori di $\boldsymbol{A}$ sono $\neq 0$.
10. La matrice $\boldsymbol{A}^T$, trasposta di $\boldsymbol{A}$, è non-singolare.

Matrice non-singolare (o invertibile)

Una matrice reale quadrata $\boldsymbol{A}$ di ordine n si dice **non-singolare** oppure **invertibile** se gode di una delle seguenti proprietà, tra di loro equivalenti:

1. Le colonne di $\boldsymbol{A}$ sono vettori linearmente indipendenti.
2. Il sistema omogeneo $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ ammette solo la soluzione nulla $\boldsymbol{x} = \boldsymbol{0}$.
3. Il nucleo di $\boldsymbol{A}$, $\ker \boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^n \,:\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$, contiene solo il vettore nullo.
4. Il sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ ammette una soluzione, per ogni $\boldsymbol{b} \in \mathbb{R}^n$.
5. L'immagine di $\boldsymbol{A}$, $\mathrm{Im}\, \boldsymbol{A} = \{\boldsymbol{A}\boldsymbol{x} \,:\, \boldsymbol{x} \in \mathbb{R}^n\}$, è tutto $\mathbb{R}^n$.
6. Il rango di $\boldsymbol{A}$, $\mathrm{rank}(\boldsymbol{A}) = \dim(\mathrm{Im}\, \boldsymbol{A})$, è uguale a n (cioè $\boldsymbol{A}$ ha rango massimo).
7. Esiste una matrice quadrata $\boldsymbol{B}$ tale che $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A} = \boldsymbol{I}$. Tale matrice è unica e viene indicata con $\boldsymbol{A}^{-1}$ (e detta la *matrice inversa* di $\boldsymbol{A}$).
8. $\det \boldsymbol{A} \neq 0$.
9. Tutti gli autovalori di $\boldsymbol{A}$ sono $\neq 0$.
10. La matrice $\boldsymbol{A}^T$, trasposta di $\boldsymbol{A}$, è non-singolare.
11. Le righe di $\boldsymbol{A}$ sono vettori linearmente indipendenti.

Una matrice reale quadrata $\boldsymbol{A}$ di ordine n si dice **non-singolare** oppure **invertibile** se gode di una delle seguenti proprietà, tra di loro equivalenti:

1. Le colonne di $\boldsymbol{A}$ sono vettori linearmente indipendenti.
2. Il sistema omogeneo $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ ammette solo la soluzione nulla $\boldsymbol{x} = \boldsymbol{0}$.
3. Il nucleo di $\boldsymbol{A}$, $\ker \boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$, contiene solo il vettore nullo.
4. Il sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ ammette una soluzione, per ogni $\boldsymbol{b} \in \mathbb{R}^n$.
5. L'immagine di $\boldsymbol{A}$, $\mathrm{Im}\, \boldsymbol{A} = \{\boldsymbol{A}\boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^n\}$, è tutto $\mathbb{R}^n$.
6. Il rango di $\boldsymbol{A}$, $\mathrm{rank}(\boldsymbol{A}) = \dim(\mathrm{Im}\, \boldsymbol{A})$, è uguale a n (cioè $\boldsymbol{A}$ ha rango massimo).
7. Esiste una matrice quadrata $\boldsymbol{B}$ tale che $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A} = \boldsymbol{I}$. Tale matrice è unica e viene indicata con $\boldsymbol{A}^{-1}$ (e detta la *matrice inversa* di $\boldsymbol{A}$).
8. $\det \boldsymbol{A} \neq 0$.
9. Tutti gli autovalori di $\boldsymbol{A}$ sono $\neq 0$.
10. La matrice $\boldsymbol{A}^T$, trasposta di $\boldsymbol{A}$, è non-singolare.
11. Le righe di $\boldsymbol{A}$ sono vettori linearmente indipendenti.
12. $\det \boldsymbol{A}^T \neq 0$.

Sia $\boldsymbol{x} = (x_i)_{1\le i\le n} \in \mathbb{R}^n$ un vettore colonna avente n componenti reali. Se p è un qualunque numero reale ≥ 1, chiamiamo *norma* p di $\boldsymbol{x}$ l'espressione

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} . \tag{1}$$

Sia $\boldsymbol{x} = (x_i)_{1 \le i \le n} \in \mathbb{R}^n$ un vettore colonna avente n componenti reali. Se p è un qualunque numero reale ≥ 1, chiamiamo *norma* p di $\boldsymbol{x}$ l'espressione

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} . \tag{1}$$

Di particolare importanza sono le norme

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i| \, , \qquad \|\boldsymbol{x}\|_2 = \boldsymbol{x}^T\boldsymbol{x} = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} , \qquad \|\boldsymbol{x}\|_\infty = \max_{1 \le i \le n} |x_i| \, .$$

Sia $\boldsymbol{x} = (x_i)_{1\le i\le n} \in \mathbb{R}^n$ un vettore colonna avente n componenti reali. Se p è un qualunque numero reale ≥ 1, chiamiamo *norma* p di $\boldsymbol{x}$ l'espressione

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} . \tag{1}$$

Di particolare importanza sono le norme

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i| , \qquad \|\boldsymbol{x}\|_2 = \boldsymbol{x}^T \boldsymbol{x} = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} , \qquad \|\boldsymbol{x}\|_\infty = \max_{1\le i\le n} |x_i| .$$

Sia poi $\boldsymbol{A} = (a_{ij})_{1\le i,j\le n} \in \mathbb{R}^{n\times n}$ una matrice quadrata di ordine n. Ad ogni norma di vettore $\|\boldsymbol{x}\|$ è associata una *norma di matrice* $\|\boldsymbol{A}\|$, definita dall'espressione

$$\|\boldsymbol{A}\| = \max_{\boldsymbol{x}\in\mathbb{R}^n \ \boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = \max_{\boldsymbol{x}\in\mathbb{R}^n \ \|\boldsymbol{x}\|=1} \|\boldsymbol{A}\boldsymbol{x}\| . \tag{2}$$

Sia $\boldsymbol{x} = (x_i)_{1 \le i \le n} \in \mathbb{R}^n$ un vettore colonna avente n componenti reali. Se p è un qualunque numero reale ≥ 1, chiamiamo *norma* p di $\boldsymbol{x}$ l'espressione

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} . \tag{1}$$

Di particolare importanza sono le norme

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i| , \qquad \|\boldsymbol{x}\|_2 = \boldsymbol{x}^T \boldsymbol{x} = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} , \qquad \|\boldsymbol{x}\|_\infty = \max_{1 \le i \le n} |x_i| .$$

Sia poi $\boldsymbol{A} = (a_{ij})_{1 \le i,j \le n} \in \mathbb{R}^{n \times n}$ una matrice quadrata di ordine n. Ad ogni norma di vettore $\|\boldsymbol{x}\|$ è associata una *norma di matrice* $\|\boldsymbol{A}\|$, definita dall'espressione

$$\|\boldsymbol{A}\| = \max_{\boldsymbol{x} \in \mathbb{R}^n \ \boldsymbol{x} \neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = \max_{\boldsymbol{x} \in \mathbb{R}^n \ \|\boldsymbol{x}\| = 1} \|\boldsymbol{A}\boldsymbol{x}\| . \tag{2}$$

Dalla definizione, si ha subito che

$$\|\boldsymbol{A}\boldsymbol{x}\| \le \|\boldsymbol{A}\| \, \|\boldsymbol{x}\| \qquad \text{per ogni } \boldsymbol{x} \in \mathbb{R}^n ,$$

e

$$\|\boldsymbol{A}\boldsymbol{B}\| \le \|\boldsymbol{A}\| \, \|\boldsymbol{B}\| , \qquad \|\boldsymbol{I}\| = 1 .$$

In particolare, si ha

$$\|\boldsymbol{A}\|_\infty = \|\boldsymbol{A}^T\|_1 = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$$

In particolare, si ha

$$\|\boldsymbol{A}\|_\infty = \|\boldsymbol{A}^T\|_1 = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$$

e

$$\|\boldsymbol{A}\|_2 = \sqrt{\rho(\boldsymbol{A}^T\boldsymbol{A})}\ ,$$

dove $\rho(\boldsymbol{B})$ indica il *raggio spettrale* di una matrice $\boldsymbol{B}$, ossia

$$\rho(\boldsymbol{B}) = \max\{\,|\lambda|\ :\ \lambda \text{ è autovalore di } \boldsymbol{B}\,\}\ .$$

In particolare, si ha

$$\|\boldsymbol{A}\|_\infty = \|\boldsymbol{A}^T\|_1 = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$$

e

$$\|\boldsymbol{A}\|_2 = \sqrt{\rho(\boldsymbol{A}^T\boldsymbol{A})}\ ,$$

dove $\rho(\boldsymbol{B})$ indica il *raggio spettrale* di una matrice $\boldsymbol{B}$, ossia

$$\rho(\boldsymbol{B}) = \max\{\, |\lambda| \ : \ \lambda \text{ è autovalore di } \boldsymbol{B}\,\}\ .$$

Se $\boldsymbol{A}$ è una matrice simmetrica (e dunque con autovalori tutti reali), si ha

$$\|\boldsymbol{A}\|_2 = \max\{\, |\lambda| \ : \ \lambda \text{ è autovalore di } \boldsymbol{A}\,\}\ .$$

In particolare, si ha

$$\|\boldsymbol{A}\|_\infty = \|\boldsymbol{A}^T\|_1 = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$$

e

$$\|\boldsymbol{A}\|_2 = \sqrt{\rho(\boldsymbol{A}^T\boldsymbol{A})}\ ,$$

dove $\rho(\boldsymbol{B})$ indica il *raggio spettrale* di una matrice $\boldsymbol{B}$, ossia

$$\rho(\boldsymbol{B}) = \max\{\,|\lambda|\ :\ \lambda \text{ è autovalore di } \boldsymbol{B}\,\}\ .$$

Se $\boldsymbol{A}$ è una matrice simmetrica (e dunque con autovalori tutti reali), si ha

$$\|\boldsymbol{A}\|_2 = \max\{\,|\lambda|\ :\ \lambda \text{ è autovalore di } \boldsymbol{A}\,\}\ .$$

Se inoltre $\boldsymbol{A}$ è definita positiva, con autovalori soddisfacenti

$$0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n\ ,$$

posto $\lambda_{\text{min}} = \lambda_1$ e $\lambda_{\text{max}} = \lambda_n$ si ha quindi

$$\|\boldsymbol{A}\|_2 = \lambda_{\text{max}}\ .$$

In particolare, si ha

$$\|\boldsymbol{A}\|_\infty = \|\boldsymbol{A}^T\|_1 = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$$

e

$$\|\boldsymbol{A}\|_2 = \sqrt{\rho(\boldsymbol{A}^T\boldsymbol{A})}\ ,$$

dove $\rho(\boldsymbol{B})$ indica il *raggio spettrale* di una matrice $\boldsymbol{B}$, ossia

$$\rho(\boldsymbol{B}) = \max\{\, |\lambda| \ :\ \lambda \text{ è autovalore di } \boldsymbol{B}\,\}\ .$$

Se $\boldsymbol{A}$ è una matrice simmetrica (e dunque con autovalori tutti reali), si ha

$$\|\boldsymbol{A}\|_2 = \max\{\, |\lambda| \ :\ \lambda \text{ è autovalore di } \boldsymbol{A}\,\}\ .$$

Se inoltre $\boldsymbol{A}$ è definita positiva, con autovalori soddisfacenti

$$0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n\ ,$$

posto $\lambda_{\text{min}} = \lambda_1$ e $\lambda_{\text{max}} = \lambda_n$ si ha quindi

$$\|\boldsymbol{A}\|_2 = \lambda_{\text{max}}\ .$$

Ricordando che

$$\boldsymbol{A}\boldsymbol{w} = \lambda\boldsymbol{w} \quad \Leftrightarrow \quad \boldsymbol{A}^{-1}\boldsymbol{w} = \lambda^{-1}\boldsymbol{w}\ ,$$

si ha quindi

$$\|\boldsymbol{A}^{-1}\|_2 = \frac{1}{\lambda_{\text{min}}}\ .$$

Sia $\boldsymbol{A}$ una matrice quadrata non-singolare. Il numero

$$\text{cond}_p(\boldsymbol{A}) = \|\boldsymbol{A}\|_p \, \|\boldsymbol{A}^{-1}\|_p \tag{3}$$

dicesi *numero di condizionamento* di $\boldsymbol{A}$ (nella norma p).

Il numero di condizionamento di una matrice

Sia $\boldsymbol{A}$ una matrice quadrata non-singolare. Il numero

$$\text{cond}_p(\boldsymbol{A}) = \|\boldsymbol{A}\|_p \, \|\boldsymbol{A}^{-1}\|_p \tag{3}$$

dicesi *numero di condizionamento* di $\boldsymbol{A}$ (nella norma p).
Si ha sempre

$$\text{cond}_p(\boldsymbol{A}) \geq 1 \ .$$

$\boldsymbol{A}$ dicesi *bencondizionata* se $\text{cond}_p(\boldsymbol{A}) \simeq 1$, *malcondizionata* se $\text{cond}_p(\boldsymbol{A}) >> 1$.

Sia $\boldsymbol{A}$ una matrice quadrata non-singolare. Il numero

$$\text{cond}_p(\boldsymbol{A}) = \|\boldsymbol{A}\|_p \, \|\boldsymbol{A}^{-1}\|_p \tag{3}$$

dicesi *numero di condizionamento* di $\boldsymbol{A}$ (nella norma p).
Si ha sempre

$$\text{cond}_p(\boldsymbol{A}) \geq 1 \ .$$

$\boldsymbol{A}$ dicesi *bencondizionata* se $\text{cond}_p(\boldsymbol{A}) \simeq 1$, *malcondizionata* se $\text{cond}_p(\boldsymbol{A}) >> 1$.

Applicazioni. Sia $\boldsymbol{b} \in \mathbb{R}^n$ un vettore non nullo (che rappresenta il "termine noto" di un certo problema, o l'"ingresso" di un sistema fisico), e sia $\boldsymbol{x} \in \mathbb{R}^n$ la soluzione del sistema lineare

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \ ,$$

(che rappresenta la "soluzione" del problema, o l'"uscita" del sistema fisico).

Sia $\boldsymbol{A}$ una matrice quadrata non-singolare. Il numero

$$\text{cond}_p(\boldsymbol{A}) = \|\boldsymbol{A}\|_p \, \|\boldsymbol{A}^{-1}\|_p \tag{3}$$

dicesi *numero di condizionamento* di $\boldsymbol{A}$ (nella norma p).
Si ha sempre

$$\text{cond}_p(\boldsymbol{A}) \geq 1 \; .$$

$\boldsymbol{A}$ dicesi *bencondizionata* se $\text{cond}_p(\boldsymbol{A}) \simeq 1$, *malcondizionata* se $\text{cond}_p(\boldsymbol{A}) >> 1$.

Applicazioni. Sia $\boldsymbol{b} \in \mathbb{R}^n$ un vettore non nullo (che rappresenta il "termine noto" di un certo problema, o l'"ingresso" di un sistema fisico), e sia $\boldsymbol{x} \in \mathbb{R}^n$ la soluzione del sistema lineare

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \; ,$$

(che rappresenta la "soluzione" del problema, o l'"uscita" del sistema fisico).
Supponiamo di conoscere non $\boldsymbol{b}$ ma una sua approssimazione $\tilde{\boldsymbol{b}}$, a causa di vari fattori (errori di misura, errori di rappresentazione numerica, etc.); corrispondentemente, abbiamo una soluzione $\tilde{\boldsymbol{x}}$ definita da

$$\boldsymbol{A}\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{b}} \; ;$$

è ragionevole pensare che $\tilde{\boldsymbol{x}}$ sia una approssimazione di $\boldsymbol{x}$.

Sia $\boldsymbol{A}$ una matrice quadrata non-singolare. Il numero

$$\text{cond}_p(\boldsymbol{A}) = \|\boldsymbol{A}\|_p \, \|\boldsymbol{A}^{-1}\|_p \tag{3}$$

dicesi *numero di condizionamento* di $\boldsymbol{A}$ (nella norma p).
Si ha sempre

$$\text{cond}_p(\boldsymbol{A}) \geq 1 \; .$$

$\boldsymbol{A}$ dicesi *bencondizionata* se $\text{cond}_p(\boldsymbol{A}) \simeq 1$, *malcondizionata* se $\text{cond}_p(\boldsymbol{A}) >> 1$.

Applicazioni. Sia $\boldsymbol{b} \in \mathbb{R}^n$ un vettore non nullo (che rappresenta il "termine noto" di un certo problema, o l'"ingresso" di un sistema fisico), e sia $\boldsymbol{x} \in \mathbb{R}^n$ la soluzione del sistema lineare

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \; ,$$

(che rappresenta la "soluzione" del problema, o l'"uscita" del sistema fisico).
Supponiamo di conoscere non $\boldsymbol{b}$ ma una sua approssimazione $\tilde{\boldsymbol{b}}$, a causa di vari fattori (errori di misura, errori di rappresentazione numerica, etc.); corrispondentemente, abbiamo una soluzione $\tilde{\boldsymbol{x}}$ definita da

$$\boldsymbol{A}\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{b}} \; ;$$

è ragionevole pensare che $\tilde{\boldsymbol{x}}$ sia una approssimazione di $\boldsymbol{x}$.
Sottraendo le due equazioni, abbiamo

$$\boldsymbol{A}\boldsymbol{x} - \boldsymbol{A}\tilde{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x} - \tilde{\boldsymbol{x}}) = \boldsymbol{b} - \tilde{\boldsymbol{b}}$$

da cui si ricava

$$\boldsymbol{x} - \tilde{\boldsymbol{x}} = \boldsymbol{A}^{-1}(\boldsymbol{b} - \tilde{\boldsymbol{b}}).$$

da cui si ricava

$$\boldsymbol{x} - \tilde{\boldsymbol{x}} = \boldsymbol{A}^{-1}(\boldsymbol{b} - \tilde{\boldsymbol{b}}).$$

Prendendo la norma di ambo i membri e maggiorando la norma del secondo membro, otteniamo

$$\|\boldsymbol{x} - \tilde{\boldsymbol{x}}\|_p \leq \|\boldsymbol{A}^{-1}\|_p \|\boldsymbol{b} - \tilde{\boldsymbol{b}}\|_p \ .$$

Questo mostra come l'*errore assoluto* sulla soluzione si controlli attraverso l'*errore assoluto* sui dati.

da cui si ricava

$$\boldsymbol{x} - \tilde{\boldsymbol{x}} = \boldsymbol{A}^{-1}(\boldsymbol{b} - \tilde{\boldsymbol{b}}).$$

Prendendo la norma di ambo i membri e maggiorando la norma del secondo membro, otteniamo

$$\|\boldsymbol{x} - \tilde{\boldsymbol{x}}\|_p \leq \|\boldsymbol{A}^{-1}\|_p \|\boldsymbol{b} - \tilde{\boldsymbol{b}}\|_p \ .$$

Questo mostra come l'*errore assoluto* sulla soluzione si controlli attraverso l'*errore assoluto* sui dati.

Ma è ben noto che l'*errore relativo* è ben più significativo dell'errore assoluto. Per ottenerlo, usiamo la relazione $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$, che ci fornisce

$$\|\boldsymbol{b}\|_p \leq \|\boldsymbol{A}\|_p \|\boldsymbol{x}\|_p \qquad \text{ossia} \qquad \frac{1}{\|\boldsymbol{x}\|_p} \leq \|\boldsymbol{A}\|_p \frac{1}{\|\boldsymbol{b}\|_p}.$$

da cui si ricava

$$\boldsymbol{x} - \tilde{\boldsymbol{x}} = \boldsymbol{A}^{-1}(\boldsymbol{b} - \tilde{\boldsymbol{b}}).$$

Prendendo la norma di ambo i membri e maggiorando la norma del secondo membro, otteniamo

$$\|\boldsymbol{x} - \tilde{\boldsymbol{x}}\|_p \leq \|\boldsymbol{A}^{-1}\|_p \|\boldsymbol{b} - \tilde{\boldsymbol{b}}\|_p \ .$$

Questo mostra come l'*errore assoluto* sulla soluzione si controlli attraverso l'*errore assoluto* sui dati.

Ma è ben noto che l'*errore relativo* è ben più significativo dell'errore assoluto. Per ottenerlo, usiamo la relazione $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$, che ci fornisce

$$\|\boldsymbol{b}\|_p \leq \|\boldsymbol{A}\|_p \|\boldsymbol{x}\|_p \qquad \text{ossia} \qquad \frac{1}{\|\boldsymbol{x}\|_p} \leq \|\boldsymbol{A}\|_p \frac{1}{\|\boldsymbol{b}\|_p}.$$

Combinando le due precedenti disuguaglianze, otteniamo

$$\frac{\|\boldsymbol{x} - \tilde{\boldsymbol{x}}\|_p}{\|\boldsymbol{x}\|_p} \leq \|\boldsymbol{A}\|_p \|\boldsymbol{A}^{-1}\|_p \frac{\|\boldsymbol{b} - \tilde{\boldsymbol{b}}\|_p}{\|\boldsymbol{b}\|_p} = \text{cond}_p(\boldsymbol{A}) \, \frac{\|\boldsymbol{b} - \tilde{\boldsymbol{b}}\|_p}{\|\boldsymbol{b}\|_p} \ .$$

da cui si ricava

$$\boldsymbol{x} - \tilde{\boldsymbol{x}} = \boldsymbol{A}^{-1}(\boldsymbol{b} - \tilde{\boldsymbol{b}}).$$

Prendendo la norma di ambo i membri e maggiorando la norma del secondo membro, otteniamo

$$\|\boldsymbol{x} - \tilde{\boldsymbol{x}}\|_p \leq \|\boldsymbol{A}^{-1}\|_p \|\boldsymbol{b} - \tilde{\boldsymbol{b}}\|_p \ .$$

Questo mostra come l'*errore assoluto* sulla soluzione si controlli attraverso l'*errore assoluto* sui dati.

Ma è ben noto che l'*errore relativo* è ben più significativo dell'errore assoluto. Per ottenerlo, usiamo la relazione $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$, che ci fornisce

$$\|\boldsymbol{b}\|_p \leq \|\boldsymbol{A}\|_p \|\boldsymbol{x}\|_p \qquad \text{ossia} \qquad \frac{1}{\|\boldsymbol{x}\|_p} \leq \|\boldsymbol{A}\|_p \frac{1}{\|\boldsymbol{b}\|_p}.$$

Combinando le due precedenti disuguaglianze, otteniamo

$$\frac{\|\boldsymbol{x} - \tilde{\boldsymbol{x}}\|_p}{\|\boldsymbol{x}\|_p} \leq \|\boldsymbol{A}\|_p \|\boldsymbol{A}^{-1}\|_p \frac{\|\boldsymbol{b} - \tilde{\boldsymbol{b}}\|_p}{\|\boldsymbol{b}\|_p} = \text{cond}_p(\boldsymbol{A}) \, \frac{\|\boldsymbol{b} - \tilde{\boldsymbol{b}}\|_p}{\|\boldsymbol{b}\|_p} \ .$$

Ciò mostra che il **numero di condizionamento della matrice controlla il modo con cui l'errore relativo sui dati si trasforma nell'errore relativo sulla soluzione**.

Se invece supponiamo che il termine noto $\boldsymbol{b}$ sia noto esattamente, mentre la matrice del sistema $\boldsymbol{A}$ sia nota soltanto attraverso una sua approssimazione $\tilde{\boldsymbol{A}}$, con ragionamenti analoghi ai precedenti si perviene al seguente risultato:

Se invece supponiamo che il termine noto $\boldsymbol{b}$ sia noto esattamente, mentre la matrice del sistema $\boldsymbol{A}$ sia nota soltanto attraverso una sua approssimazione $\tilde{\boldsymbol{A}}$, con ragionamenti analoghi ai precedenti si perviene al seguente risultato:

Se $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ e $\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}} = \boldsymbol{b}$ allora vale la maggiorazione

$$\frac{\|\boldsymbol{x} - \tilde{\boldsymbol{x}}\|_p}{\|\tilde{\boldsymbol{x}}\|_p} \leq \mathsf{cond}_p(\boldsymbol{A}) \, \frac{\|\boldsymbol{A} - \tilde{\boldsymbol{A}}\|_p}{\|\boldsymbol{A}\|_p} \ .$$

Anche in questo caso, il numero di condizionamento di $\boldsymbol{A}$ controlla il modo in cui l'errore relativo sulla matrice influenza l'errore relativo sulla soluzione del sistema lineare.

Per una matrice simmetrica e definita positiva, il numero di condizionamento nella norma euclidea si rappresenta come

$$\text{cond}_2(\boldsymbol{A}) = \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} \ . \tag{4}$$

Dunque **una matrice simmetrica e definita positiva è malcondizionata quando i suoi autovalori hanno ordini di grandezza molto diversi tra loro**.

Per una matrice simmetrica e definita positiva, il numero di condizionamento nella norma euclidea si rappresenta come

$$\mathrm{cond}_2(\boldsymbol{A}) = \frac{\lambda_{\mathrm{max}}}{\lambda_{\mathrm{min}}} \, . \tag{4}$$

Dunque **una matrice simmetrica e definita positiva è malcondizionata quando i suoi autovalori hanno ordini di grandezza molto diversi tra loro**.

Esempio

Un classico esempio di matrici (simmetriche e definite positive) molto malcondizionate è costituito dalla famiglia di matrici di Hilbert $\boldsymbol{H}_n$ i cui elementi sono dati da

$$h_{ij} = \frac{1}{i+j-1} \, , \qquad 1 \leq i,j \leq n \, ,$$

(queste matrici sono definite dal comando hilb *di MATLAB). I numeri di condizionamento cond$_2(\boldsymbol{H}_n)$ (stimabili attraverso il comando* cond *di MATLAB) crescono esponenzialmente al crescere di n.*

2. Sistemi lineari e fattorizzazioni LU di matrici

Un esempio: l'opinione di un gruppo di individui

La necessità di risolvere sistemi lineari, spesso di grandi dimensioni, si incontra in svariate applicazioni.
Vediamo qui un semplice esempio, tratto dall'*applicazione della Matematica a problematiche di natura sociale* (un filone della Matematica Applicata che sta avendo crescente interesse).

La necessità di risolvere sistemi lineari, spesso di grandi dimensioni, si incontra in svariate applicazioni.
Vediamo qui un semplice esempio, tratto dall'*applicazione della Matematica a problematiche di natura sociale* (un filone della Matematica Applicata che sta avendo crescente interesse).

Consideriamo un insieme di $n+1$ individui, e supponiamo che ciascuno di loro abbia un'**opinione** su un determinato argomento, rappresentata dal valore di una certa variabile reale x: precisamente, x_i indica l'opinione dell'individuo i-esimo, per $0 \leq i \leq n$.

Valori di x_i negativi indicano che l'individuo i è contrario all'argomento, valori di x_i vicini a 0 indicano che l'individuo i è sostanzialmente indifferente all'argomento, mentre valori di x_i positivi indicano che l'individuo i condivide l'argomento.

Un esempio: l'opinione di un gruppo di individui

La necessità di risolvere sistemi lineari, spesso di grandi dimensioni, si incontra in svariate applicazioni.
Vediamo qui un semplice esempio, tratto dall'*applicazione della Matematica a problematiche di natura sociale* (un filone della Matematica Applicata che sta avendo crescente interesse).

Consideriamo un insieme di $n+1$ individui, e supponiamo che ciascuno di loro abbia un'**opinione** su un determinato argomento, rappresentata dal valore di una certa variabile reale x: precisamente, x_i indica l'opinione dell'individuo i-esimo, per $0 \leq i \leq n$.

Valori di x_i negativi indicano che l'individuo i è contrario all'argomento, valori di x_i vicini a 0 indicano che l'individuo i è sostanzialmente indifferente all'argomento, mentre valori di x_i positivi indicano che l'individuo i condivide l'argomento.

Le opinioni variano con il tempo: io parlo con un mio amico che ha un'opinione molto più positiva della mia sull'argomento, e tendo a migliorare la mia opinione; oppure leggo sul web una pagina che parla in modo critico dell'argomento, e tendo a peggiorare la mia opinione.

Supponiamo che il tempo vari in modo discreto con passo $\Delta t > 0$ attraverso gli istanti $t_k = k\Delta t$, $k \geq 0$, e sia x_i^k l'opinione dell'individuo i al tempo t_k.

Supponiamo che il tempo vari in modo discreto con passo $\Delta t > 0$ attraverso gli istanti $t_k = k\Delta t$, $k \geq 0$, e sia x_i^k l'opinione dell'individuo i al tempo t_k.

Supponendo di conoscere tutte le opinioni x_i^0 al tempo iniziale $t = 0$, un modello che descrive l'evoluzione delle opinioni (**modello di consenso**) è il seguente:

$$x_i^{k+1} = x_i^k + \sum_{\substack{j=0 \\ j \neq i}}^{n} \alpha_{ij}(x_j^k - x_i^k) + b_i^k \ , \qquad 0 \leq i \leq n, \quad k \geq 0, \tag{5}$$

dove

Supponiamo che il tempo vari in modo discreto con passo $\Delta t > 0$ attraverso gli istanti $t_k = k\Delta t$, $k \geq 0$, e sia x_i^k l'opinione dell'individuo i al tempo t_k.

Supponendo di conoscere tutte le opinioni x_i^0 al tempo iniziale $t = 0$, un modello che descrive l'evoluzione delle opinioni (**modello di consenso**) è il seguente:

$$x_i^{k+1} = x_i^k + \sum_{\substack{j=0 \\ j \neq i}}^{n} \alpha_{ij}(x_j^k - x_i^k) + b_i^k \,, \qquad 0 \leq i \leq n, \quad k \geq 0, \tag{5}$$

dove

- il coefficiente $\alpha_{ij} \geq 0$ misura quanto la differenza di opinioni tra l'individuo j e l'individuo i influenza l'opinione dell'individuo i. Se $\alpha_{ij} = 0$, l'individuo i non è influenzato dall'individuo j.

Supponiamo che il tempo vari in modo discreto con passo $\Delta t > 0$ attraverso gli istanti $t_k = k\Delta t$, $k \geq 0$, e sia x_i^k l'opinione dell'individuo i al tempo t_k.

Supponendo di conoscere tutte le opinioni x_i^0 al tempo iniziale $t = 0$, un modello che descrive l'evoluzione delle opinioni (**modello di consenso**) è il seguente:

$$x_i^{k+1} = x_i^k + \sum_{\substack{j=0 \\ j \neq i}}^{n} \alpha_{ij}(x_j^k - x_i^k) + b_i^k \,, \qquad 0 \leq i \leq n, \quad k \geq 0, \tag{5}$$

dove

- il coefficiente $\alpha_{ij} \geq 0$ misura quanto la differenza di opinioni tra l'individuo j e l'individuo i influenza l'opinione dell'individuo i. Se $\alpha_{ij} = 0$, l'individuo i non è influenzato dall'individuo j.
- il termine b_i^k quantifica l'influsso del mondo esterno sull'opinione dell'individuo i al tempo t_k.

Supponiamo che il tempo vari in modo discreto con passo $\Delta t > 0$ attraverso gli istanti $t_k = k\Delta t$, $k \geq 0$, e sia x_i^k l'opinione dell'individuo i al tempo t_k.

Supponendo di conoscere tutte le opinioni x_i^0 al tempo iniziale $t = 0$, un modello che descrive l'evoluzione delle opinioni (**modello di consenso**) è il seguente:

$$x_i^{k+1} = x_i^k + \sum_{\substack{j=0 \\ j\neq i}}^{n} \alpha_{ij}(x_j^k - x_i^k) + b_i^k \,, \qquad 0 \leq i \leq n, \quad k \geq 0, \tag{5}$$

dove

- il coefficiente $\alpha_{ij} \geq 0$ misura quanto la differenza di opinioni tra l'individuo j e l'individuo i influenza l'opinione dell'individuo i. Se $\alpha_{ij} = 0$, l'individuo i non è influenzato dall'individuo j.
- il termine b_i^k quantifica l'influsso del mondo esterno sull'opinione dell'individuo i al tempo t_k.

La (5) è un *sistema dinamico discreto*, di cui è interessante conoscere il comportamento per tempi lunghi, ossia per $k \to \infty$.

Una situazione notevole è quella in cui il sistema converge verso una *configurazione di equilibrio*, o *stato stazionario*, supponendo che l'influsso del mondo esterno sia invariante nel tempo. Indicando con x_i l'opinione dell'individuo i all'equilibrio, si ottiene quindi il sistema algebrico

$$\sum_{\substack{j=0 \\ j \neq i}}^{n} \alpha_{ij}(x_j - x_i) + b_i = 0 \,, \qquad 0 \leq i \leq n. \tag{6}$$

Una situazione notevole è quella in cui il sistema converge verso una *configurazione di equilibrio*, o *stato stazionario*, supponendo che l'influsso del mondo esterno sia invariante nel tempo. Indicando con x_i l'opinione dell'individuo i all'equilibrio, si ottiene quindi il sistema algebrico

$$\sum_{\substack{j=0 \\ j \neq i}}^{n} \alpha_{ij}(x_j - x_i) + b_i = 0 \;, \qquad 0 \leq i \leq n. \tag{6}$$

Tale sistema, però, non è ben posto: aggiungendo una costante a tutte le opinioni, $x_i \to x_i + c$, il sistema non cambia.

Una situazione notevole è quella in cui il sistema converge verso una *configurazione di equilibrio*, o *stato stazionario*, supponendo che l'influsso del mondo esterno sia invariante nel tempo. Indicando con x_i l'opinione dell'individuo i all'equilibrio, si ottiene quindi il sistema algebrico

$$\sum_{\substack{j=0 \\ j\neq i}}^{n} \alpha_{ij}(x_j - x_i) + b_i = 0 \ , \qquad 0 \leq i \leq n. \tag{6}$$

Tale sistema, però, non è ben posto: aggiungendo una costante a tutte le opinioni, $x_i \to x_i + c$, il sistema non cambia.

Per rendere ben posto il sistema, possiamo

Una situazione notevole è quella in cui il sistema converge verso una *configurazione di equilibrio*, o *stato stazionario*, supponendo che l'influsso del mondo esterno sia invariante nel tempo. Indicando con x_i l'opinione dell'individuo i all'equilibrio, si ottiene quindi il sistema algebrico

$$\sum_{\substack{j=0 \\ j \neq i}}^{n} \alpha_{ij}(x_j - x_i) + b_i = 0 \,, \qquad 0 \leq i \leq n. \tag{6}$$

Tale sistema, però, non è ben posto: aggiungendo una costante a tutte le opinioni, $x_i \to x_i + c$, il sistema non cambia.

Per rendere ben posto il sistema, possiamo

- fissare l'opinione di un individuo, ad esempio porre $x_0 = 0$

Una situazione notevole è quella in cui il sistema converge verso una *configurazione di equilibrio*, o *stato stazionario*, supponendo che l'influsso del mondo esterno sia invariante nel tempo. Indicando con x_i l'opinione dell'individuo i all'equilibrio, si ottiene quindi il sistema algebrico

$$\sum_{\substack{j=0 \\ j \neq i}}^{n} \alpha_{ij}(x_j - x_i) + b_i = 0 \; , \qquad 0 \leq i \leq n. \tag{6}$$

Tale sistema, però, non è ben posto: aggiungendo una costante a tutte le opinioni, $x_i \to x_i + c$, il sistema non cambia.

Per rendere ben posto il sistema, possiamo

- fissare l'opinione di un individuo, ad esempio porre $x_0 = 0$
- supporre che il sistema sia *irriducibile*, ossia che ogni individuo sia influenzato da ogni altro, o direttamente o per il tramite di altri individui.

Una situazione notevole è quella in cui il sistema converge verso una *configurazione di equilibrio*, o *stato stazionario*, supponendo che l'influsso del mondo esterno sia invariante nel tempo. Indicando con x_i l'opinione dell'individuo i all'equilibrio, si ottiene quindi il sistema algebrico

$$\sum_{\substack{j=0 \\ j\neq i}}^{n} \alpha_{ij}(x_j - x_i) + b_i = 0\,, \qquad 0 \leq i \leq n. \tag{6}$$

Tale sistema, però, non è ben posto: aggiungendo una costante a tutte le opinioni, $x_i \to x_i + c$, il sistema non cambia.

Per rendere ben posto il sistema, possiamo

- fissare l'opinione di un individuo, ad esempio porre $x_0 = 0$
- supporre che il sistema sia *irriducibile*, ossia che ogni individuo sia influenzato da ogni altro, o direttamente o per il tramite di altri individui.

Sotto queste ipotesi, il sistema diventa:

$$\begin{cases} \displaystyle\sum_{\substack{j=0 \\ j\neq i}}^{n} \alpha_{ij}(x_j - x_i) + b_i = 0\,, \quad 1 \leq i \leq n, \\ x_0 = 0 \end{cases} \tag{7}$$

Se introduciamo i vettori colonna $\boldsymbol{x} = (x_i)_{1 \le i \le n}$, $\boldsymbol{b} = (b_i)_{1 \le i \le n}$ e la matrice quadrata $\boldsymbol{A}$ di ordine n i cui elementi a_{ij} sono dati da

$$a_{ij} = \begin{cases} \displaystyle\sum_{\substack{j=0 \\ j \neq i}}^{n} \alpha_{ij} & \text{se } i = j, \\ -\alpha_{ij} & \text{se } i \neq j, \end{cases}$$

il sistema precedente diventa

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

e la matrice $\boldsymbol{A}$ è non-singolare.

Se introduciamo i vettori colonna $\boldsymbol{x} = (x_i)_{1 \le i \le n}$, $\boldsymbol{b} = (b_i)_{1 \le i \le n}$ e la matrice quadrata $\boldsymbol{A}$ di ordine n i cui elementi a_{ij} sono dati da

$$a_{ij} = \begin{cases} \displaystyle\sum_{\substack{j=0 \\ j \neq i}}^{n} \alpha_{ij} & \text{se } i = j, \\ -\alpha_{ij} & \text{se } i \neq j, \end{cases}$$

il sistema precedente diventa

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

e la matrice $\boldsymbol{A}$ è non-singolare.

La ricerca della configurazione di equilibrio del nostro modello di opinioni di una popolazione di individui è dunque ridotto alla *soluzione di tale sistema algebrico*.

Alcuni casi notevoli di sistemi algebrici

Il nostro obiettivo è trovare metodi efficienti per la risoluzione di un sistema algebrico di n equazioni lineari in n incognite

$$\left\{ \begin{array}{ccccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \dots & + & a_{2n}x_n & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & + & \dots & + & a_{3n}x_n & = & b_3 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & a_{n3}x_3 & + & \dots & + & a_{nn}x_n & = & b_n \end{array} \right.$$

che scriviamo in forma compatta come

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \; .$$

Supponiamo che la matrice $\boldsymbol{A}$ sia non-singolare, e dunque il sistema ammette una e una soluzione per ogni scelta del termine noto $\boldsymbol{b} \in \mathbb{R}^n$.

Alcuni casi notevoli di sistemi algebrici

Il nostro obiettivo è trovare metodi efficienti per la risoluzione di un sistema algebrico di n equazioni lineari in n incognite

$$\left\{\begin{array}{ccccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \dots & + & a_{2n}x_n & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & + & \dots & + & a_{3n}x_n & = & b_3 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & a_{n3}x_3 & + & \dots & + & a_{nn}x_n & = & b_n \end{array}\right.$$

che scriviamo in forma compatta come

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\ .$$

Supponiamo che la matrice $\boldsymbol{A}$ sia non-singolare, e dunque il sistema ammette una e una soluzione per ogni scelta del termine noto $\boldsymbol{b} \in \mathbb{R}^n$.

Osservazione. La *regola di Cramer* **NON deve mai essere usata!** (tranne forse per fare i calcoli a mano con un sistema 3×3...)
Infatti è *mostruosamente inefficiente*: richiede ben $(n+1)!$ operazioni per risolvere il sistema lineare.

Consideriamo nel seguito alcuni casi notevoli di sistemi algebrici.

1. Sistemi diagonali. Sono della forma

$$\begin{cases} a_{11}x_1 & & & & & = & b_1 \\ & a_{22}x_2 & & & & = & b_2 \\ & & a_{33}x_3 & & & = & b_3 \\ & & & \ddots & & & \vdots \\ & & & & a_{nn}x_n & = & b_n \end{cases}$$

ossia la matrice del sistema è diagonale: $\boldsymbol{A} = \boldsymbol{D} = \text{diag}\,(\boldsymbol{a})$, con $\boldsymbol{a} = (a_{11}, a_{22}, \ldots, a_{nn})$ vettore di elementi tutti diversi da 0.

Consideriamo nel seguito alcuni casi notevoli di sistemi algebrici.

1. Sistemi diagonali. Sono della forma

$$\left\{\begin{array}{ccccccc} a_{11}x_1 & & & & & = & b_1 \\ & a_{22}x_2 & & & & = & b_2 \\ & & a_{33}x_3 & & & = & b_3 \\ & & & \ddots & & & \vdots \\ & & & & a_{nn}x_n & = & b_n \end{array}\right.$$

ossia la matrice del sistema è diagonale: $\boldsymbol{A} = \boldsymbol{D} = \text{diag}\,(\boldsymbol{a})$, con $\boldsymbol{a} = (a_{11}, a_{22}, \ldots, a_{nn})$ vettore di elementi tutti diversi da 0.

In tal caso, si ha

$$x_i = b_i/a_{ii}, \qquad 1 \leq i \leq n,$$

ossia in MATLAB

```
x = b./a
```

La risoluzione richiede n operazioni.

2. Sistemi triangolari inferiori. Sono della forma

$$\begin{cases} a_{11}x_1 & & & & & & & & & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & & & & & & & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & & & & & = & b_3 \\ \vdots & & \vdots & & \vdots & & \vdots & & & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & a_{n3}x_3 & + & \dots & + & a_{nn}x_n & = & b_n \end{cases}$$

ossia la matrice $\boldsymbol{A} = \boldsymbol{L}$ è triangolare inferiore (cioè soddisfa $a_{ij} = 0$ se $j > i$) con elementi diagonali tutti diversi da 0.

2. Sistemi triangolari inferiori. Sono della forma

$$\begin{cases} a_{11}x_1 & & & & & & & & & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & & & & & & & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & & & & & = & b_3 \\ \vdots & & \vdots & & \vdots & & \vdots & & & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & a_{n3}x_3 & + & \dots & + & a_{nn}x_n & = & b_n \end{cases}$$

ossia la matrice $\boldsymbol{A} = \boldsymbol{L}$ è triangolare inferiore (cioè soddisfa $a_{ij} = 0$ se $j > i$) con elementi diagonali tutti diversi da 0.

A partire da x_1, si ricava la i-esima incognita x_i dalla i-esima equazione, e si sostituisce il suo valore in tutte le equazioni *successive* alla i-esima (**metodo di sostituzione in avanti**):

$$\begin{aligned} x_1 &= b_1/a_{11} \\ x_i &= \Big(b_i - \sum_{j=1}^{i-1} a_{ij}x_j\Big)/a_{ii} \qquad i = 2, 3, \dots, n. \end{aligned}$$

Si tratta dunque di un procedimento *ricorsivo*.

2. Sistemi triangolari inferiori. Sono della forma

$$\begin{cases} a_{11}x_1 & & & & & & & & & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & & & & & & & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & & & & & = & b_3 \\ \vdots & & \vdots & & \vdots & & \vdots & & & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & a_{n3}x_3 & + & \ldots & + & a_{nn}x_n & = & b_n \end{cases}$$

ossia la matrice $\boldsymbol{A} = \boldsymbol{L}$ è triangolare inferiore (cioè soddisfa $a_{ij} = 0$ se $j > i$) con elementi diagonali tutti diversi da 0.

A partire da x_1, si ricava la i-esima incognita x_i dalla i-esima equazione, e si sostituisce il suo valore in tutte le equazioni *successive* alla i-esima (**metodo di sostituzione in avanti**):

$$\begin{aligned} x_1 &= b_1/a_{11} \\ x_i &= \Big(b_i - \sum_{j=1}^{i-1} a_{ij}x_j\Big)/a_{ii} \qquad i = 2, 3, \ldots, n. \end{aligned}$$

Si tratta dunque di un procedimento *ricorsivo*.

L'algoritmo richiede i operazioni per il calcolo dell'incognita i-esima, e quindi il costo totale è di

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} \sim \frac{n^2}{2} \quad \text{operazioni.}$$

Metodo di sostituzione in avanti (o forward substitution) in MATLAB:

```
for i=1:n
    y(i)=b(i);
    for j=1:i-1
        y(i)=y(i)-A(i,j)*y(j);
    end
    y(i)=y(i)/A(i,i);
end
```

Metodo di sostituzione in avanti (o forward substitution) in MATLAB:

```
for i=1:n
    y(i)=b(i);
    for j=1:i-1
        y(i)=y(i)-A(i,j)*y(j);
    end
    y(i)=y(i)/A(i,i);
end
```

Un'alternativa che sfrutta la capacità di MATLAB di operare in modo ottimizzato direttamente sui vettori (attraverso la libreria BLAS) è la seguente:

```
y(1)=b(1)/A(1,1);
for i=2:n
    y(i)=(b(i)-A(i,1:i-1)*y(1:i-1))/A(i,i);
end
```

3. Sistemi triangolari superiori. Sono della forma

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \dots & + & a_{1n}x_n & = & b_1 \\ & & a_{22}x_2 & + & a_{23}x_3 & + & \dots & + & a_{2n}x_n & = & b_2 \\ & & & & a_{33}x_3 & + & \dots & + & a_{3n}x_n & = & b_3 \\ & & & & & & \vdots & & \vdots & & \vdots \\ & & & & & & & & a_{nn}x_n & = & b_n \end{cases}$$

ossia la matrice $\boldsymbol{A} = \boldsymbol{U}$ è triangolare superiore (cioè soddisfa $a_{ij} = 0$ se $j < i$) con elementi diagonali tutti diversi da 0.

3. Sistemi triangolari superiori. Sono della forma

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \dots & + & a_{1n}x_n & = & b_1 \\ & & a_{22}x_2 & + & a_{23}x_3 & + & \dots & + & a_{2n}x_n & = & b_2 \\ & & & & a_{33}x_3 & + & \dots & + & a_{3n}x_n & = & b_3 \\ & & & & & & \vdots & & \vdots & & \vdots \\ & & & & & & & & a_{nn}x_n & = & b_n \end{cases}$$

ossia la matrice $\boldsymbol{A} = \boldsymbol{U}$ è triangolare superiore (cioè soddisfa $a_{ij} = 0$ se $j < i$) con elementi diagonali tutti diversi da 0.

A partire da x_n, si ricava la i-esima incognita x_i dalla i-esima equazione, e si sostituisce il suo valore in tutte le equazioni *precedenti* alla i-esima (**metodo di sostituzione all'indietro**):

$$\begin{aligned} x_n &= b_n/a_{nn} \\ x_i &= \Big(b_i - \sum_{j=i+1}^{n} a_{ij}x_j\Big)/a_{ii} \qquad i = n-1, n-2, \dots, 2, 1. \end{aligned}$$

3. Sistemi triangolari superiori. Sono della forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1 \\ \phantom{a_{11}x_1 +} a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = b_2 \\ \phantom{a_{11}x_1 + a_{22}x_2 +} a_{33}x_3 + \dots + a_{3n}x_n = b_3 \\ \vdots \\ a_{nn}x_n = b_n \end{cases}$$

ossia la matrice $\boldsymbol{A} = \boldsymbol{U}$ è triangolare superiore (cioè soddisfa $a_{ij} = 0$ se $j < i$) con elementi diagonali tutti diversi da 0.

A partire da x_n, si ricava la i-esima incognita x_i dalla i-esima equazione, e si sostituisce il suo valore in tutte le equazioni *precedenti* alla i-esima (**metodo di sostituzione all'indietro**):

$$\begin{aligned} x_n &= b_n / a_{nn} \\ x_i &= \big(b_i - \sum_{j=i+1}^{n} a_{ij}x_j\big)/a_{ii} \qquad i = n-1, n-2, \dots, 2, 1. \end{aligned}$$

Anche in questo caso abbiamo un procedimento *ricorsivo*, che richiede

$$\frac{n(n+1)}{2} \sim \frac{n^2}{2} \quad \text{operazioni.}$$

Metodo di sostituzione all'indietro (o backward substitution) in MATLAB:

```
for i=n:-1:1
    x(i)=b(i);
    for j=i+1:n
        x(i)=x(i)-A(i,j)*x(j);
    end
    x(i)=x(i)/A(i,i);
end
```

Sfruttando le funzioni ottimizzate di MATLAB:

```
x(n)=b(n)/A(n,n);
for i=n-1:-1:1
    x(i)=(b(i)-A(i,i+1:n)*x(i+1:n))/A(i,i);
end
```

4. Sistemi la cui matrice A è fattorizzata in forma LU.

Supponiamo di conoscere una matrice triangolare inferiore $\boldsymbol{L}$ e una matrice triangolare superiore $\boldsymbol{U}$ tali che la matrice $\boldsymbol{A}$ del sistema si scriva come

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$$

(questo è un esempio di *fattorizzazione* della matrice $\boldsymbol{A}$).

4. Sistemi la cui matrice A è fattorizzata in forma LU.

Supponiamo di conoscere una matrice triangolare inferiore $\boldsymbol{L}$ e una matrice triangolare superiore $\boldsymbol{U}$ tali che la matrice $\boldsymbol{A}$ del sistema si scriva come

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$$

(questo è un esempio di *fattorizzazione* della matrice $\boldsymbol{A}$).

Dalla relazione

$$0 \neq \det \boldsymbol{A} = \det \boldsymbol{L} \det \boldsymbol{U}$$

segue che $\det \boldsymbol{L} \neq 0$, $\det \boldsymbol{U} \neq 0$, cioè $\boldsymbol{L}$ e $\boldsymbol{U}$ sono *non-singolari*.

4. Sistemi la cui matrice A è fattorizzata in forma LU.

Supponiamo di conoscere una matrice triangolare inferiore $\boldsymbol{L}$ e una matrice triangolare superiore $\boldsymbol{U}$ tali che la matrice $\boldsymbol{A}$ del sistema si scriva come

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$$

(questo è un esempio di *fattorizzazione* della matrice $\boldsymbol{A}$).

Dalla relazione

$$0 \neq \det \boldsymbol{A} = \det \boldsymbol{L} \det \boldsymbol{U}$$

segue che $\det \boldsymbol{L} \neq 0$, $\det \boldsymbol{U} \neq 0$, cioè $\boldsymbol{L}$ e $\boldsymbol{U}$ sono *non-singolari*.

Il sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, cioè $\boldsymbol{L}\boldsymbol{U}\boldsymbol{x} = \boldsymbol{b}$, equivale ai due sistemi

$$\boldsymbol{L}\boldsymbol{y} = \boldsymbol{b} \qquad \text{e} \qquad \boldsymbol{U}\boldsymbol{x} = \boldsymbol{y}$$

che possono essere risolti in cascata mediante una *sostituzione in avanti seguita da una sostituzione all'indietro*.

4. Sistemi la cui matrice A è fattorizzata in forma LU.

Supponiamo di conoscere una matrice triangolare inferiore $\boldsymbol{L}$ e una matrice triangolare superiore $\boldsymbol{U}$ tali che la matrice $\boldsymbol{A}$ del sistema si scriva come

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$$

(questo è un esempio di *fattorizzazione* della matrice $\boldsymbol{A}$).

Dalla relazione

$$0 \neq \det \boldsymbol{A} = \det \boldsymbol{L} \det \boldsymbol{U}$$

segue che $\det \boldsymbol{L} \neq 0$, $\det \boldsymbol{U} \neq 0$, cioè $\boldsymbol{L}$ e $\boldsymbol{U}$ sono *non-singolari*.

Il sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, cioè $\boldsymbol{L}\boldsymbol{U}\boldsymbol{x} = \boldsymbol{b}$, equivale ai due sistemi

$$\boldsymbol{L}\boldsymbol{y} = \boldsymbol{b} \qquad \text{e} \qquad \boldsymbol{U}\boldsymbol{x} = \boldsymbol{y}$$

che possono essere risolti in cascata mediante una *sostituzione in avanti seguita da una sostituzione all'indietro*.

Il numero totale di operazioni è

$$\frac{n(n+1)}{2} + \frac{n(n+1)}{2} = n(n+1) \sim n^2 \; .$$

5. Sistemi per i quali si conosce una fattorizzazione LU di una permutazione della matrice associata.

Non tutte le matrici non-singolari $\boldsymbol{A}$ ammettono una fattorizzazione $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$ con $\boldsymbol{L}$ e $\boldsymbol{U}$ non-singolari.

Ad esempio, se $a_{11} = 0$ tale fattorizzazione non può esistere, perchè se esistesse si avrebbe

$$a_{11} = \sum_{j=1}^{n} \ell_{1j} u_{j1} = \ell_{11} u_{11}$$

e dunque necessariamente $\ell_{11} = 0$ oppure $u_{11} = 0$.

5. Sistemi per i quali si conosce una fattorizzazione LU di una permutazione della matrice associata.

Non tutte le matrici non-singolari $\boldsymbol{A}$ ammettono una fattorizzazione $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$ con $\boldsymbol{L}$ e $\boldsymbol{U}$ non-singolari.

Ad esempio, se $a_{11} = 0$ tale fattorizzazione non può esistere, perchè se esistesse si avrebbe

$$a_{11} = \sum_{j=1}^{n} \ell_{1j} u_{j1} = \ell_{11} u_{11}$$

e dunque necessariamente $\ell_{11} = 0$ oppure $u_{11} = 0$.

In tal caso, però, possiamo scambiare la prima equazione del sistema algebrico con un'altra equazione, diciamo la r-esima, tale che $a_{r1} \neq 0$; una tale equazione esiste certamente, perchè se fosse $a_{i1} = 0$ per ogni i, la matrice $\boldsymbol{A}$ sarebbe singolare.

5. Sistemi per i quali si conosce una fattorizzazione LU di una permutazione della matrice associata.

Non tutte le matrici non-singolari $\boldsymbol{A}$ ammettono una fattorizzazione $\boldsymbol{A} = \boldsymbol{LU}$ con $\boldsymbol{L}$ e $\boldsymbol{U}$ non-singolari.

Ad esempio, se $a_{11} = 0$ tale fattorizzazione non può esistere, perchè se esistesse si avrebbe

$$a_{11} = \sum_{j=1}^{n} \ell_{1j} u_{j1} = \ell_{11} u_{11}$$

e dunque necessariamente $\ell_{11} = 0$ oppure $u_{11} = 0$.

In tal caso, però, possiamo scambiare la prima equazione del sistema algebrico con un'altra equazione, diciamo la r-esima, tale che $a_{r1} \neq 0$; una tale equazione esiste certamente, perchè se fosse $a_{i1} = 0$ per ogni i, la matrice $\boldsymbol{A}$ sarebbe singolare.

Questo scambio di equazioni implica uno scambio tra la prima e la r-esima riga della matrice $\boldsymbol{A}$, in modo che l'elemento nella prima riga e prima colonna della nuova matrice sia $\neq 0$, e dunque la fattorizzazione $\boldsymbol{LU}$ di tale matrice non sia impedita.

5. Sistemi per i quali si conosce una fattorizzazione LU di una permutazione della matrice associata.

Non tutte le matrici non-singolari $\boldsymbol{A}$ ammettono una fattorizzazione $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$ con $\boldsymbol{L}$ e $\boldsymbol{U}$ non-singolari.

Ad esempio, se $a_{11} = 0$ tale fattorizzazione non può esistere, perchè se esistesse si avrebbe

$$a_{11} = \sum_{j=1}^{n} \ell_{1j} u_{j1} = \ell_{11} u_{11}$$

e dunque necessariamente $\ell_{11} = 0$ oppure $u_{11} = 0$.

In tal caso, però, possiamo scambiare la prima equazione del sistema algebrico con un'altra equazione, diciamo la r-esima, tale che $a_{r1} \neq 0$; una tale equazione esiste certamente, perchè se fosse $a_{i1} = 0$ per ogni i, la matrice $\boldsymbol{A}$ sarebbe singolare.

Questo scambio di equazioni implica uno scambio tra la prima e la r-esima riga della matrice $\boldsymbol{A}$, in modo che l'elemento nella prima riga e prima colonna della nuova matrice sia $\neq 0$, e dunque la fattorizzazione $\boldsymbol{L}\boldsymbol{U}$ di tale matrice non sia impedita.

Faremo vedere tra poco che *per ogni matrice non-singolare $\boldsymbol{A}$ si può trovare una permutazione delle sue righe* (che corrisponde a scrivere le equazioni del sistema in un ordine diverso da quello iniziale) *tale che la nuova matrice ammetta una fattorizzazione $\boldsymbol{L}\boldsymbol{U}$*.

In base a questo risultato, esiste una matrice non singolare $\boldsymbol{P}$, detta *matrice di permutazione* (vedi le prossime slides), una matrice triangolare inferiore $\boldsymbol{L}$ e una matrice triangolare superiore $\boldsymbol{U}$ tali che

$$\boldsymbol{P}\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U} \ .$$

In base a questo risultato, esiste una matrice non singolare $\boldsymbol{P}$, detta *matrice di permutazione* (vedi le prossime slides), una matrice triangolare inferiore $\boldsymbol{L}$ e una matrice triangolare superiore $\boldsymbol{U}$ tali che

$$\boldsymbol{P}\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U} \ .$$

Dalla relazione

$$0 \neq \det \boldsymbol{P} \det \boldsymbol{A} = \det \boldsymbol{L} \det \boldsymbol{U}$$

segue ancora che $\det \boldsymbol{L} \neq 0$, $\det \boldsymbol{U} \neq 0$, cioè $\boldsymbol{L}$ e $\boldsymbol{U}$ sono non-singolari.

In base a questo risultato, esiste una matrice non singolare $\boldsymbol{P}$, detta *matrice di permutazione* (vedi le prossime slides), una matrice triangolare inferiore $\boldsymbol{L}$ e una matrice triangolare superiore $\boldsymbol{U}$ tali che

$$\boldsymbol{P}\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U} \ .$$

Dalla relazione

$$0 \neq \det \boldsymbol{P} \det \boldsymbol{A} = \det \boldsymbol{L} \det \boldsymbol{U}$$

segue ancora che $\det \boldsymbol{L} \neq 0$, $\det \boldsymbol{U} \neq 0$, cioè $\boldsymbol{L}$ e $\boldsymbol{U}$ sono non-singolari.

Il sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ equivale al sistema

$$\boldsymbol{P}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{P}\boldsymbol{b} \qquad \text{ossia} \qquad \boldsymbol{L}\boldsymbol{U}\boldsymbol{x} = \boldsymbol{P}\boldsymbol{b}.$$

Tale sistema può essere risolto, come precedentemente, attraverso una sostituzione in avanti seguita da una sostituzione all'indietro, cioè risolvendo in cascata i due sistemi

$$\boldsymbol{L}\boldsymbol{y} = \boldsymbol{P}\boldsymbol{b} \qquad \text{e} \qquad \boldsymbol{U}\boldsymbol{x} = \boldsymbol{y}.$$

In base a questo risultato, esiste una matrice non singolare $\boldsymbol{P}$, detta *matrice di permutazione* (vedi le prossime slides), una matrice triangolare inferiore $\boldsymbol{L}$ e una matrice triangolare superiore $\boldsymbol{U}$ tali che

$$\boldsymbol{PA} = \boldsymbol{LU} \ .$$

Dalla relazione

$$0 \neq \det \boldsymbol{P} \det \boldsymbol{A} = \det \boldsymbol{L} \det \boldsymbol{U}$$

segue ancora che $\det \boldsymbol{L} \neq 0$, $\det \boldsymbol{U} \neq 0$, cioè $\boldsymbol{L}$ e $\boldsymbol{U}$ sono non-singolari.

Il sistema $\boldsymbol{Ax} = \boldsymbol{b}$ equivale al sistema

$$\boldsymbol{PAx} = \boldsymbol{Pb} \qquad \text{ossia} \qquad \boldsymbol{LUx} = \boldsymbol{Pb}.$$

Tale sistema può essere risolto, come precedentemente, attraverso una sostituzione in avanti seguita da una sostituzione all'indietro, cioè risolvendo in cascata i due sistemi

$$\boldsymbol{Ly} = \boldsymbol{Pb} \qquad \text{e} \qquad \boldsymbol{Ux} = \boldsymbol{y}.$$

Il numero totale di operazioni è ancora

$$n(n+1) \sim n^2$$

in quanto il prodotto $\boldsymbol{Pb}$ rappresenta solo una permutazione delle componenti di $\boldsymbol{b}$, e quindi non comporta operazioni algebriche.

Matrici di permutazione. Siano $k \neq r$ interi compresi tra 1 e n. Una *matrice di permutazione semplice* $\boldsymbol{P}_{[kr]}$ è ottenuta dalla matrice identità scambiando tra loro le righe k ed r:

$$\boldsymbol{P}_{[kr]} = \begin{array}{cc} & \begin{array}{ccccccccc} & & & & k & & r & & \end{array} \\ \begin{array}{c} \\ \\ \\ \\ k \\ \\ r \\ \\ \\ \end{array} & \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ & & & & \vdots & & \vdots & & \\ 0 & 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ & & & & \vdots & & \vdots & & \\ 0 & 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ & & & & \vdots & & \vdots & & \\ 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \end{array}.$$

Matrici di permutazione. Siano $k \neq r$ interi compresi tra 1 e n. Una *matrice di permutazione semplice* $\boldsymbol{P}_{[kr]}$ è ottenuta dalla matrice identità scambiando tra loro le righe k ed r:

$$\boldsymbol{P}_{[kr]} = \begin{array}{c} \\ \\ \\ \\ k \\ \\ r \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccccccc} & & & & k & & r & & \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ & & & & \vdots & & \vdots & & \\ 0 & 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ & & & & \vdots & & \vdots & & \\ 0 & 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ & & & & \vdots & & \vdots & & \\ 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \end{array}.$$

- L'applicazione di $\boldsymbol{P}_{[kr]}$ a un vettore colonna $\boldsymbol{b}$, formando il vettore $\boldsymbol{P}_{[kr]}\boldsymbol{b}$, scambia tra loro le componenti di indice k ed r di $\boldsymbol{b}$.

Matrici di permutazione. Siano $k \neq r$ interi compresi tra 1 e n. Una *matrice di permutazione semplice* $\boldsymbol{P}_{[kr]}$ è ottenuta dalla matrice identità scambiando tra loro le righe k ed r:

$$\boldsymbol{P}_{[kr]} = \begin{array}{c} \\ \\ \\ \\ \\ k \\ \\ r \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccccccc} & & & & k & & r & & \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ & & & & \vdots & & \vdots & & \\ 0 & 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ & & & & \vdots & & \vdots & & \\ 0 & 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ & & & & \vdots & & \vdots & & \\ 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \end{array} .$$

- L'applicazione di $\boldsymbol{P}_{[kr]}$ a un vettore colonna $\boldsymbol{b}$, formando il vettore $\boldsymbol{P}_{[kr]}\boldsymbol{b}$, scambia tra loro le componenti di indice k ed r di $\boldsymbol{b}$.
- L'applicazione di $\boldsymbol{P}_{[kr]}$ a una matrice $\boldsymbol{A}$, formando la matrice $\boldsymbol{P}_{[\boldsymbol{kr}]}\boldsymbol{A}$, scambia tra loro le *righe* di indice k ed r di $\boldsymbol{A}$.
  Invece, formando la matrice $\boldsymbol{A}\boldsymbol{P}_{[kr]}$ si scambiano tra loro le *colonne* di indice k ed r di $\boldsymbol{A}$.

Matrici di permutazione. Siano $k \neq r$ interi compresi tra 1 e n. Una *matrice di permutazione semplice* $\boldsymbol{P}_{[kr]}$ è ottenuta dalla matrice identità scambiando tra loro le righe k ed r:

$$\boldsymbol{P}_{[kr]} = \begin{array}{c c} & \begin{array}{ccccccccc} & & & & k & & r & & \end{array} \\ \begin{array}{c} \\ \\ \\ \\ k \\ \\ r \\ \\ \\ \end{array} & \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ & & & & \vdots & & \vdots & & \\ 0 & 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ & & & & \vdots & & \vdots & & \\ 0 & 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ & & & & \vdots & & \vdots & & \\ 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \end{array} .$$

- L'applicazione di $\boldsymbol{P}_{[kr]}$ a un vettore colonna $\boldsymbol{b}$, formando il vettore $\boldsymbol{P}_{[kr]}\boldsymbol{b}$, scambia tra loro le componenti di indice k ed r di $\boldsymbol{b}$.
- L'applicazione di $\boldsymbol{P}_{[kr]}$ a una matrice $\boldsymbol{A}$, formando la matrice $\boldsymbol{P}_{[\boldsymbol{kr}]}\boldsymbol{A}$, scambia tra loro le *righe* di indice k ed r di $\boldsymbol{A}$.
  Invece, formando la matrice $\boldsymbol{A}\boldsymbol{P}_{[kr]}$ si scambiano tra loro le *colonne* di indice k ed r di $\boldsymbol{A}$.
- Si ha $\boldsymbol{P}_{[kr]}^T = \boldsymbol{P}_{[kr]}$, ed inoltre $\boldsymbol{P}_{[kr]}\boldsymbol{P}_{[kr]} = \boldsymbol{I}$, cioè $\boldsymbol{P}_{[kr]}^{-1} = \boldsymbol{P}_{[kr]}$.

Una *generica permutazione delle righe* della matrice $\boldsymbol{A}$ può essere espressa nella forma

$$\boldsymbol{P}\boldsymbol{A} = \underbrace{\boldsymbol{P}_{[n-1,r_{n-1}]} \cdots \boldsymbol{P}_{[3,r_3]}\boldsymbol{P}_{[2,r_2]}\boldsymbol{P}_{[1,r_1]}}_{\boldsymbol{P}} \boldsymbol{A}$$

dove $r_k \geq k$ e

- se $r_k = k$, $\boldsymbol{P}_{[k,k]}$ è per definizione la matrice identità,
- se $r_k > k$, $\boldsymbol{P}_{[k,r_k]}$ è la matrice di permutazione semplice che scambia la riga corrente di posto k con quella di posto r_k.

Una *generica permutazione delle righe* della matrice $\boldsymbol{A}$ può essere espressa nella forma

$$\boldsymbol{P}\boldsymbol{A} = \underbrace{\boldsymbol{P}_{[n-1,r_{n-1}]} \cdots \boldsymbol{P}_{[3,r_3]}\boldsymbol{P}_{[2,r_2]}\boldsymbol{P}_{[1,r_1]}}_{\boldsymbol{P}}\boldsymbol{A}$$

dove $r_k \geq k$ e

- se $r_k = k$, $\boldsymbol{P}_{[k,k]}$ è per definizione la matrice identità,
- se $r_k > k$, $\boldsymbol{P}_{[k,r_k]}$ è la matrice di permutazione semplice che scambia la riga corrente di posto k con quella di posto r_k.

Ricordando che se $\boldsymbol{B}$, $\boldsymbol{C}$ sono matrici quadrate invertibili vale la formula

$$(\boldsymbol{B}\boldsymbol{C})^{-1} = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1},$$

si ha

$$\boldsymbol{P}^{-1} = \boldsymbol{P}^{-1}_{[1,r_1]}\boldsymbol{P}^{-1}_{[2,r_2]}\boldsymbol{P}^{-1}_{[3,r_3]} \cdots \boldsymbol{P}^{-1}_{[n-1,r_{n-1}]}.$$

Una *generica permutazione delle righe* della matrice $\boldsymbol{A}$ può essere espressa nella forma

$$\boldsymbol{PA} = \underbrace{\boldsymbol{P}_{[n-1,r_{n-1}]} \cdots \boldsymbol{P}_{[3,r_3]}\boldsymbol{P}_{[2,r_2]}\boldsymbol{P}_{[1,r_1]}}_{\boldsymbol{P}} \boldsymbol{A}$$

dove $r_k \geq k$ e

- se $r_k = k$, $\boldsymbol{P}_{[k,k]}$ è per definizione la matrice identità,
- se $r_k > k$, $\boldsymbol{P}_{[k,r_k]}$ è la matrice di permutazione semplice che scambia la riga corrente di posto k con quella di posto r_k.

Ricordando che se $\boldsymbol{B}$, $\boldsymbol{C}$ sono matrici quadrate invertibili vale la formula

$$(\boldsymbol{BC})^{-1} = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1},$$

si ha

$$\boldsymbol{P}^{-1} = \boldsymbol{P}^{-1}_{[1,r_1]}\boldsymbol{P}^{-1}_{[2,r_2]}\boldsymbol{P}^{-1}_{[3,r_3]} \cdots \boldsymbol{P}^{-1}_{[n-1,r_{n-1}]}.$$

Nota. A livello implementativo, sovente una matrice di permutazione non viene realmente costruita, ma l'informazione in essa contenuta viene codificata da un vettore di n interi (*puntatore*) che indica quale sia la nuova posizione di ciascuna riga della matrice originaria.

Il **metodo di eliminazione di Gauss con pivoting parziale** può essere visto in due modi diversi (ma tra loro collegati):

Il **metodo di eliminazione di Gauss con pivoting parziale** può essere visto in due modi diversi (ma tra loro collegati):

- trasforma un sistema algebrico $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ con matrice $\boldsymbol{A}$ quadrata non-singolare in un sistema equivalente (cioè con la stessa soluzione) $\boldsymbol{U}\boldsymbol{x} = \boldsymbol{c}$ con matrice $\boldsymbol{U}$ triangolare superiore e non-singolare.

Il **metodo di eliminazione di Gauss con pivoting parziale** può essere visto in due modi diversi (ma tra loro collegati):

- trasforma un sistema algebrico $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ con matrice $\boldsymbol{A}$ quadrata non-singolare in un sistema equivalente (cioè con la stessa soluzione) $\boldsymbol{U}\boldsymbol{x} = \boldsymbol{c}$ con matrice $\boldsymbol{U}$ triangolare superiore e non-singolare.

- produce i fattori $\boldsymbol{L}$, $\boldsymbol{U}$, $\boldsymbol{P}$ della fattorizzazione $\boldsymbol{P}\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$ relativa a una matrice quadrata non-singolare $\boldsymbol{A}$.

Il **metodo di eliminazione di Gauss con pivoting parziale** può essere visto in due modi diversi (ma tra loro collegati):

- trasforma un sistema algebrico $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ con matrice $\boldsymbol{A}$ quadrata non-singolare in un sistema equivalente (cioè con la stessa soluzione) $\boldsymbol{U}\boldsymbol{x} = \boldsymbol{c}$ con matrice $\boldsymbol{U}$ triangolare superiore e non-singolare.

- produce i fattori $\boldsymbol{L}$, $\boldsymbol{U}$, $\boldsymbol{P}$ della fattorizzazione $\boldsymbol{P}\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$ relativa a una matrice quadrata non-singolare $\boldsymbol{A}$.

Nel seguito, descriviamo come il metodo opera su un sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, ma parallelamente indichiamo come vengono costruite le matrici $\boldsymbol{P}$, $\boldsymbol{L}$ e $\boldsymbol{U}$.

Sia n il numero di equazioni e di incognite del sistema algebrico, cioè l'ordine della matrice $\boldsymbol{A}$.

- Il metodo di Gauss si compone di $n-1$ passi: ad ogni passo si riduce di 1 il numero di equazioni e di incognite su cui il metodo opera.

Sia n il numero di equazioni e di incognite del sistema algebrico, cioè l'ordine della matrice $\boldsymbol{A}$.

- Il metodo di Gauss si compone di $n - 1$ passi: ad ogni passo si riduce di 1 il numero di equazioni e di incognite su cui il metodo opera.
- A sua volta, ogni passo si suddivide in due fasi:

Sia n il numero di equazioni e di incognite del sistema algebrico, cioè l'ordine della matrice $\boldsymbol{A}$.

- Il metodo di Gauss si compone di $n-1$ passi: ad ogni passo si riduce di 1 il numero di equazioni e di incognite su cui il metodo opera.
- A sua volta, ogni passo si suddivide in due fasi:
    1. la ricerca del cosiddetto *elemento pivot* sulla prima colonna della matrice dei coefficienti del sistema algebrico corrente, e l'eventuale conseguente *scambio di due equazioni*;

Sia n il numero di equazioni e di incognite del sistema algebrico, cioè l'ordine della matrice $\boldsymbol{A}$.

- Il metodo di Gauss si compone di $n-1$ passi: ad ogni passo si riduce di 1 il numero di equazioni e di incognite su cui il metodo opera.
- A sua volta, ogni passo si suddivide in due fasi:
  1. la ricerca del cosiddetto *elemento pivot* sulla prima colonna della matrice dei coefficienti del sistema algebrico corrente, e l'eventuale conseguente *scambio di due equazioni*;
  2. l'*eliminazione* della prima incognita del sistema algebrico corrente da tutte le equazioni successive alla prima.

Il fondamento teorico del metodo di eliminazione di Gauss è dato dalle seguenti due *proprietà dei sistemi lineari*:

Il fondamento teorico del metodo di eliminazione di Gauss è dato dalle seguenti due *proprietà dei sistemi lineari*:

- la soluzione rimane invariata se si scambiano tra loro due equazioni del sistema;

Il fondamento teorico del metodo di eliminazione di Gauss è dato dalle seguenti due *proprietà dei sistemi lineari*:

- la soluzione rimane invariata se si scambiano tra loro due equazioni del sistema;
- la soluzione rimane invariata se si sostituisce ad un'equazione del sistema una combinazione lineare dell'equazione stessa con un'altra equazione.

Il fondamento teorico del metodo di eliminazione di Gauss è dato dalle seguenti due *proprietà dei sistemi lineari*:

- la soluzione rimane invariata se si scambiano tra loro due equazioni del sistema;
- la soluzione rimane invariata se si sostituisce ad un'equazione del sistema una combinazione lineare dell'equazione stessa con un'altra equazione.

Nel seguito, descriviamo il primo passo del metodo, che ci porta dal sistema originale di ordine n a un sistema algebrico ridotto di $n-1$ equazioni e incognite. Reiterando tale procedura, si completa la realizzazione del metodo.

Successivamente, esemplifichiamo l'esecuzione del metodo per un sistema algebrico di ordine 4.

Consideriamo il sistema

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \dots & + & a_{2n}x_n & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & + & \dots & + & a_{3n}x_n & = & b_3 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & a_{n3}x_3 & + & \dots & + & a_{nn}x_n & = & b_n \end{cases}$$

Uno almeno dei coefficienti a_{i1} della prima colonna è diverso da 0, altrimenti il sistema sarebbe singolare.

Consideriamo il sistema

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \dots & + & a_{2n}x_n & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & + & \dots & + & a_{3n}x_n & = & b_3 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & a_{n3}x_3 & + & \dots & + & a_{nn}x_n & = & b_n \end{cases}$$

Uno almeno dei coefficienti a_{i1} della prima colonna è diverso da 0, altrimenti il sistema sarebbe singolare.

Cerchiamo un elemento di modulo massimo:

$$|a_{r,1}| = \max_{1 \le i \le n} |a_{i1}|$$

che chiamiamo *elemento pivot*.

Tale scelta è motivata da considerazioni di *stabilità numerica*, ossia dalla volontà di limitare la propagazione degli errori di arrotondamento dovuti all'aritmetica finita della macchina.

Evidenziamo l'elemento pivot:

$$\left\{\begin{array}{ccccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \dots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \dots & + & a_{2n}x_n & = & b_2 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\color{red}{a_{r1}}x_1 & + & a_{r2}x_2 & + & a_{r3}x_3 & + & \dots & + & a_{rn}x_n & = & b_r \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
a_{n1}x_1 & + & a_{n2}x_2 & + & a_{n3}x_3 & + & \dots & + & a_{nn}x_n & = & b_n
\end{array}\right.$$

Evidenziamo l'elemento pivot:

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \color{red}{a_{r1}}x_1 & + & a_{r2}x_2 & + & a_{r3}x_3 & + & \dots & + & a_{rn}x_n & = & b_r \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & a_{n3}x_3 & + & \dots & + & a_{nn}x_n & = & b_n \end{cases}$$

Se $r \neq 1$, scambiamo la prima equazione con la r-esima:

$$\begin{cases} \color{red}{a_{r1}x_1} & \color{red}{+} & \color{red}{a_{r2}x_2} & \color{red}{+} & \color{red}{a_{r3}x_3} & \color{red}{+} & \color{red}{\dots} & \color{red}{+} & \color{red}{a_{rn}x_n} & \color{red}{=} & \color{red}{b_r} \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \color{blue}{a_{11}x_1} & \color{blue}{+} & \color{blue}{a_{12}x_2} & \color{blue}{+} & \color{blue}{a_{13}x_3} & \color{blue}{+} & \color{blue}{\dots} & \color{blue}{+} & \color{blue}{a_{1n}x_n} & \color{blue}{=} & \color{blue}{b_1} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & a_{n3}x_3 & + & \dots & + & a_{nn}x_n & = & b_n \end{cases}$$

Evidenziamo l'elemento pivot:

$$\left\{\begin{array}{ccccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \dots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \dots & + & a_{2n}x_n & = & b_2 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\color{red}{a_{r1}}x_1 & + & a_{r2}x_2 & + & a_{r3}x_3 & + & \dots & + & a_{rn}x_n & = & b_r \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
a_{n1}x_1 & + & a_{n2}x_2 & + & a_{n3}x_3 & + & \dots & + & a_{nn}x_n & = & b_n
\end{array}\right.$$

Se $r \neq 1$, scambiamo la prima equazione con la r-esima:

$$\left\{\begin{array}{ccccccccccc}
a_{r1}x_1 & + & a_{r2}x_2 & + & a_{r3}x_3 & + & \dots & + & a_{rn}x_n & = & b_r \\
a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \dots & + & a_{2n}x_n & = & b_2 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \dots & + & a_{1n}x_n & = & b_1 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
a_{n1}x_1 & + & a_{n2}x_2 & + & a_{n3}x_3 & + & \dots & + & a_{nn}x_n & = & b_n
\end{array}\right.$$

In questo modo, resta definita la *matrice di permutazione semplice* $\boldsymbol{P}_{[1r]}$.

Indichiamo i coefficienti del nuovo sistema con l'apice $^{(1)}$:

$$\left\{\begin{array}{ccccccccccc}
a_{11}^{(1)}x_1 & + & a_{12}^{(1)}x_2 & + & a_{13}^{(1)}x_3 & + & \dots & + & a_{1n}^{(1)}x_n & = & b_1^{(1)} \\
a_{21}^{(1)}x_1 & + & a_{22}^{(1)}x_2 & + & a_{23}^{(1)}x_3 & + & \dots & + & a_{2n}^{(1)}x_n & = & b_2^{(1)} \\
a_{31}^{(1)}x_1 & + & a_{32}^{(1)}x_2 & + & a_{33}^{(1)}x_3 & + & \dots & + & a_{3n}^{(1)}x_n & = & b_3^{(1)} \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
a_{n1}^{(1)}x_1 & + & a_{n2}^{(1)}x_2 & + & a_{n3}^{(1)}x_3 & + & \dots & + & a_{nn}^{(1)}x_n & = & b_n^{(1)}
\end{array}\right.$$

Fase 2: eliminazione di un'incognita e riduzione dell'ordine del sistema

Indichiamo i coefficienti del nuovo sistema con l'apice $^{(1)}$:

$$\left\{ \begin{array}{ccccccccccc} a_{11}^{(1)}x_1 & + & a_{12}^{(1)}x_2 & + & a_{13}^{(1)}x_3 & + & \dots & + & a_{1n}^{(1)}x_n & = & b_1^{(1)} \\ a_{21}^{(1)}x_1 & + & a_{22}^{(1)}x_2 & + & a_{23}^{(1)}x_3 & + & \dots & + & a_{2n}^{(1)}x_n & = & b_2^{(1)} \\ a_{31}^{(1)}x_1 & + & a_{32}^{(1)}x_2 & + & a_{33}^{(1)}x_3 & + & \dots & + & a_{3n}^{(1)}x_n & = & b_3^{(1)} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}^{(1)}x_1 & + & a_{n2}^{(1)}x_2 & + & a_{n3}^{(1)}x_3 & + & \dots & + & a_{nn}^{(1)}x_n & = & b_n^{(1)} \end{array} \right.$$

- I coefficienti della prima equazione definiscono gli elementi della *prima riga* della matrice triangolare superiore $\boldsymbol{U}$, mentre il termine noto della prima equazione definisce la *prima componente* del vettore $\boldsymbol{c}$ nel sistema equivalente $\boldsymbol{Ux} = \boldsymbol{c}$.
  In altri termini, poniamo:

$$u_{1j} = a_{1j}^{(1)}, \quad 1 \le j \le n, \qquad \text{e} \qquad c_1 = b_1^{(1)}.$$

Indichiamo i coefficienti del nuovo sistema con l'apice $^{(1)}$:

$$\begin{cases} a_{11}^{(1)}x_1 & + & a_{12}^{(1)}x_2 & + & a_{13}^{(1)}x_3 & + & \dots & + & a_{1n}^{(1)}x_n & = & b_1^{(1)} \\ a_{21}^{(1)}x_1 & + & a_{22}^{(1)}x_2 & + & a_{23}^{(1)}x_3 & + & \dots & + & a_{2n}^{(1)}x_n & = & b_2^{(1)} \\ a_{31}^{(1)}x_1 & + & a_{32}^{(1)}x_2 & + & a_{33}^{(1)}x_3 & + & \dots & + & a_{3n}^{(1)}x_n & = & b_3^{(1)} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}^{(1)}x_1 & + & a_{n2}^{(1)}x_2 & + & a_{n3}^{(1)}x_3 & + & \dots & + & a_{nn}^{(1)}x_n & = & b_n^{(1)} \end{cases}$$

- I coefficienti della prima equazione definiscono gli elementi della *prima riga* della matrice triangolare superiore $\boldsymbol{U}$, mentre il termine noto della prima equazione definisce la *prima componente* del vettore $\boldsymbol{c}$ nel sistema equivalente $\boldsymbol{Ux} = \boldsymbol{c}$.
  In altri termini, poniamo:

$$u_{1j} = a_{1j}^{(1)}, \quad 1 \le j \le n, \qquad \text{e} \qquad c_1 = b_1^{(1)}.$$

Ricordando che $a_{11}^{(1)} \neq 0$ per la Fase 1, ora l'idea è quella di usare la prima equazione per esprimere l'incognita x_1 in funzione delle altre incognite $x_2, x_2, \dots, x_n$, e di sostituire tale espressione nelle equazioni successive alla prima.

Ciò può essere realizzato nel modo seguente:

Ciò può essere realizzato nel modo seguente:

- definiamo i *moltiplicatori*

$$m_{i1} = -\frac{a_{i1}^{(1)}}{a_{11}^{(1)}}, \qquad 2 \leq i \leq n,$$

 che per la Fase 1 soddisfano $|m_{i1}| \leq 1$, una proprietà che garantisce la *stabilità numerica* dell'algoritmo.

Ciò può essere realizzato nel modo seguente:

- definiamo i *moltiplicatori*

$$m_{i1} = -\frac{a_{i1}^{(1)}}{a_{11}^{(1)}}, \qquad 2 \le i \le n,$$

  che per la Fase 1 soddisfano $|m_{i1}| \le 1$, una proprietà che garantisce la *stabilità numerica* dell'algoritmo.

- per $2 \le i \le n$, sostituiamo alla i-esima equazione, la combinazione lineare

$$\{\ i\text{-esima equazione}\ \} + m_{i1} * \{\ \text{prima equazione}\ \}$$

  ottenendo un sistema equivalente.

Ciò può essere realizzato nel modo seguente:

- definiamo i *moltiplicatori*

$$m_{i1} = -\frac{a_{i1}^{(1)}}{a_{11}^{(1)}}, \qquad 2 \leq i \leq n,$$

  che per la Fase 1 soddisfano $|m_{i1}| \leq 1$, una proprietà che garantisce la *stabilità numerica* dell'algoritmo.

- per $2 \leq i \leq n$, sostituiamo alla i-esima equazione, la combinazione lineare

$$\{\ i\text{-esima equazione}\ \} + m_{i1} * \{\ \text{prima equazione}\ \}$$

  ottenendo un sistema equivalente.

  In tale combinazione, il coefficiente che moltiplica x_1 è $a_{i1}^{(1)} + m_{i1}a_{11}^{(1)} = 0$, dunque *non compare l'incognita* x_1, mentre gli altri coefficienti sono dati da

$$a_{ij}^{(2)} = a_{ij}^{(1)} + m_{i1}a_{1j}^{(1)}, \qquad 2 \leq i, j \leq n,$$

  e i termini noti sono dati da

$$b_i^{(2)} = b_i^{(1)} + m_{i1}b_1^{(1)}, \qquad 2 \leq i \leq n.$$

In tal modo, si giunge al sistema equivalente

$$\left\{\begin{array}{cccccccccccc}
a_{11}^{(1)}x_1 & + & a_{12}^{(1)}x_2 & + & a_{13}^{(1)}x_3 & + & \dots & + & a_{1n}^{(1)}x_n & = & b_1^{(1)} \\
 & & a_{22}^{(2)}x_2 & + & a_{23}^{(2)}x_3 & + & \dots & + & a_{2n}^{(2)}x_n & = & b_2^{(2)} \\
 & & a_{32}^{(2)}x_2 & + & a_{33}^{(2)}x_3 & + & \dots & + & a_{3n}^{(2)}x_n & = & b_3^{(2)} \\
 & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
 & & a_{n2}^{(2)}x_2 & + & a_{n3}^{(2)}x_3 & + & \dots & + & a_{nn}^{(2)}x_n & = & b_n^{(2)}
\end{array}\right.$$

In tal modo, si giunge al sistema equivalente

$$\begin{cases} a_{11}^{(1)}x_1 & + & a_{12}^{(1)}x_2 & + & a_{13}^{(1)}x_3 & + & \dots & + & a_{1n}^{(1)}x_n & = & b_1^{(1)} \\ & & a_{22}^{(2)}x_2 & + & a_{23}^{(2)}x_3 & + & \dots & + & a_{2n}^{(2)}x_n & = & b_2^{(2)} \\ & & a_{32}^{(2)}x_2 & + & a_{33}^{(2)}x_3 & + & \dots & + & a_{3n}^{(2)}x_n & = & b_3^{(2)} \\ & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ & & a_{n2}^{(2)}x_2 & + & a_{n3}^{(2)}x_3 & + & \dots & + & a_{nn}^{(2)}x_n & = & b_n^{(2)} \end{cases}$$

- I moltiplicatori cambiati di segno definiscono gli elementi della *prima colonna* della matrice triangolare inferiore $\boldsymbol{L}$, che ha gli elementi sulla diagonale tutti uguali a 1. In altri termini, poniamo:

$$\ell_{i1} = \begin{cases} 1 & \text{se } i = 1, \\ -m_{i1} & \text{se } 2 \le i \le n. \end{cases}$$

In tal modo, si giunge al sistema equivalente

$$\left\{\begin{array}{ccccccccccc}
a_{11}^{(1)}x_1 & + & a_{12}^{(1)}x_2 & + & a_{13}^{(1)}x_3 & + & \dots & + & a_{1n}^{(1)}x_n & = & b_1^{(1)} \\
 & & a_{22}^{(2)}x_2 & + & a_{23}^{(2)}x_3 & + & \dots & + & a_{2n}^{(2)}x_n & = & b_2^{(2)} \\
 & & a_{32}^{(2)}x_2 & + & a_{33}^{(2)}x_3 & + & \dots & + & a_{3n}^{(2)}x_n & = & b_3^{(2)} \\
 & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
 & & a_{n2}^{(2)}x_2 & + & a_{n3}^{(2)}x_3 & + & \dots & + & a_{nn}^{(2)}x_n & = & b_n^{(2)}
\end{array}\right.$$

- I moltiplicatori cambiati di segno definiscono gli elementi della *prima colonna* della matrice triangolare inferiore $\boldsymbol{L}$, che ha gli elementi sulla diagonale tutti uguali a 1. In altri termini, poniamo:

$$\ell_{i1} = \begin{cases} 1 & \text{se } i = 1, \\ -m_{i1} & \text{se } 2 \leq i \leq n. \end{cases}$$

A questo punto, siamo pronti ad effettuare un nuovo passo del metodo di Gauss applicato al sistema ridotto costituito dalle ultime $n-1$ equazioni nelle ultime $n-1$ incognite.
E così via...

Al termine degli $n-1$ passi, si giunge al sistema triangolare superiore $\boldsymbol{U}\boldsymbol{x} = \boldsymbol{c}$ dato da

$$\left\{\begin{array}{ccccccccccc}
a_{11}^{(1)}x_1 & + & a_{12}^{(1)}x_2 & + & a_{13}^{(1)}x_3 & + & \dots & + & a_{1n}^{(1)}x_n & = & b_1^{(1)} \\
 & & a_{22}^{(2)}x_2 & + & a_{23}^{(2)}x_3 & + & \dots & + & a_{2n}^{(2)}x_n & = & b_2^{(2)} \\
 & & & & a_{33}^{(3)}x_3 & + & \dots & + & a_{3n}^{(3)}x_n & = & b_3^{(3)} \\
 & & & & & & \ddots & & \vdots & & \vdots \\
 & & & & & & & & a_{nn}^{(n)}x_n & = & b_n^{(n)}
\end{array}\right.$$

Esso viene risolto per *sostituzione all'indietro* (si noti che tutti gli elementi diagonali sono $\neq 0$ per costruzione).

Al termine degli $n-1$ passi, si giunge al sistema triangolare superiore $\boldsymbol{U}\boldsymbol{x} = \boldsymbol{c}$ dato da

$$\left\{ \begin{array}{ccccccccccc} a_{11}^{(1)}x_1 & + & a_{12}^{(1)}x_2 & + & a_{13}^{(1)}x_3 & + & \dots & + & a_{1n}^{(1)}x_n & = & b_1^{(1)} \\ & & a_{22}^{(2)}x_2 & + & a_{23}^{(2)}x_3 & + & \dots & + & a_{2n}^{(2)}x_n & = & b_2^{(2)} \\ & & & & a_{33}^{(3)}x_3 & + & \dots & + & a_{3n}^{(3)}x_n & = & b_3^{(3)} \\ & & & & & & \ddots & & \vdots & & \vdots \\ & & & & & & & & a_{nn}^{(n)}x_n & = & b_n^{(n)} \end{array} \right.$$

Esso viene risolto per *sostituzione all'indietro* (si noti che tutti gli elementi diagonali sono $\neq 0$ per costruzione).

Nel contempo, si è costruita la matrice di permutazione

$$\boldsymbol{P} = \boldsymbol{P}_{[n-1,r_{n-1}]} \cdots \boldsymbol{P}_{[3,r_3]}\boldsymbol{P}_{[2,r_2]}\boldsymbol{P}_{[1,r_1]}$$

e la matrice triangolare inferiore $\boldsymbol{L}$ i cui elementi ℓ_{ij} con $j \leq i$ sono dati da

$$\ell_{ij} = \begin{cases} 1 & \text{se } i = j, \\ -m_{ij} & \text{se } j \leq i \leq n. \end{cases}$$

Si verifica che le matrici $\boldsymbol{P}$, $\boldsymbol{L}$, $\boldsymbol{U}$ costruite dall'algoritmo di Gauss soddisfano la relazione

$$\boldsymbol{P}\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}.$$

- L'algoritmo di eliminazione di Gauss richiede, nel caso di una matrice $\boldsymbol{A}$ piena,

$$\sim \frac{n^3}{3} \quad \text{operazioni.}$$

Strutture particolari della matrice (ad esempio matrici a banda) possono in taluni casi portare a un costo computazionale minore.

- L'algoritmo di eliminazione di Gauss richiede, nel caso di una matrice $\boldsymbol{A}$ piena,

$$\sim \frac{n^3}{3} \text{ operazioni.}$$

  Strutture particolari della matrice (ad esempio matrici a banda) possono in taluni casi portare a un costo computazionale minore.

- Per alcuni tipi di matrici, *è possibile eseguire l'algoritmo senza la Fase 1*, cioè senza effettuare la ricerca dell'elemento di modulo massimo sulla prima colonna della sottomatrice quadrata corrente. Questo perchè l'elemento che viene a trovarsi nella prima riga e prima colonna risulta sempre $\neq 0$, e la sua scelta come elemento pivot non pregiudica la stabilità dell'algoritmo.

- L'algoritmo di eliminazione di Gauss richiede, nel caso di una matrice $\boldsymbol{A}$ piena,

$$\sim \frac{n^3}{3} \ \text{operazioni.}$$

  Strutture particolari della matrice (ad esempio matrici a banda) possono in taluni casi portare a un costo computazionale minore.

- Per alcuni tipi di matrici, *è possibile eseguire l'algoritmo senza la Fase 1*, cioè senza effettuare la ricerca dell'elemento di modulo massimo sulla prima colonna della sottomatrice quadrata corrente. Questo perchè l'elemento che viene a trovarsi nella prima riga e prima colonna risulta sempre $\neq 0$, e la sua scelta come elemento pivot non pregiudica la stabilità dell'algoritmo.

  Esempi di matrici per cui ciò è possibile sono
  - le matrici *a dominanza diagonale per righe* (cioè tali che $|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}|$ per ogni i)
  - le matrici a dominanza diagonale per colonne (cioè tali che $|a_{jj}| > \sum_{i=1, i\neq j}^{n} |a_{ij}|$ per ogni j)
  - le matrici *simmetriche e definite positive* (per le quali però esiste uno specifico algoritmo di fattorizzazione, il **metodo di Choleski** - vedi dopo).

- L'algoritmo di eliminazione di Gauss richiede, nel caso di una matrice $\boldsymbol{A}$ piena,

$$\sim \frac{n^3}{3} \quad \text{operazioni.}$$

Strutture particolari della matrice (ad esempio matrici a banda) possono in taluni casi portare a un costo computazionale minore.

- Per alcuni tipi di matrici, *è possibile eseguire l'algoritmo senza la Fase 1*, cioè senza effettuare la ricerca dell'elemento di modulo massimo sulla prima colonna della sottomatrice quadrata corrente. Questo perchè l'elemento che viene a trovarsi nella prima riga e prima colonna risulta sempre $\neq 0$, e la sua scelta come elemento pivot non pregiudica la stabilità dell'algoritmo.

  Esempi di matrici per cui ciò è possibile sono
  - le matrici *a dominanza diagonale per righe* (cioè tali che $|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}|$ per ogni i)
  - le matrici a dominanza diagonale per colonne (cioè tali che $|a_{jj}| > \sum_{i=1, i\neq j}^{n} |a_{ij}|$ per ogni j)
  - le matrici *simmetriche e definite positive* (per le quali però esiste uno specifico algoritmo di fattorizzazione, il **metodo di Choleski** - vedi dopo).

  Se non si effettua il pivoting, la proprietà di $\boldsymbol{A}$ di essere simmetrica viene preservata in tutte le sottomatrici quadrate generate dall'algoritmo di Gauss, e il costo computazionale si dimezza

$$\sim \frac{n^3}{6} \quad \text{operazioni.}$$

In ogni caso, anche per l'algoritmo con pivoting, la propagazione degli errori di arrotondamento generati nel corso delle operazioni di macchina è legata al buono o cattivo **condizionamento** della matrice $\boldsymbol{A}$.

[*Esempi da sviluppare in aula*]

Un esempio

Applichiamo il metodo di eliminazioni di Gauss al sistema:

$$\begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 2 & 0 & -1 \\ 2 & -1 & 1 & -2 \\ 0 & 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 4 \end{pmatrix}$$

la cui soluzione esatta è $\boldsymbol{x} = (1, 1, 1, 1)^T$.

Applichiamo il metodo di eliminazioni di Gauss al sistema:

$$\begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 2 & 0 & -1 \\ 2 & -1 & 1 & -2 \\ 0 & 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 4 \end{pmatrix}$$

la cui soluzione esatta è $\boldsymbol{x} = (1, 1, 1, 1)^T$.

Consideriamo la *matrice estesa* ottenuta da $\boldsymbol{A}$ aggiungendo come ulteriore colonna il termine noto. Questo perchè le operazioni sul termine noto sono del tutto simili a quelle sugli elementi della matrice.

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

Passo 1

1. L'elemento pivot è $a_{31} = 2$ $\Rightarrow$ scambio delle righe 1 e 3:

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

Passo 1

1. L'elemento pivot è $a_{31} = 2 \;\Rightarrow\;$ scambio delle righe 1 e 3:

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

La corrispondente matrice di permutazione è

$$\boldsymbol{P}_{[13]} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Passo 1

1. L'elemento pivot è $a_{31} = 2 \;\Rightarrow\;$ scambio delle righe 1 e 3:

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

La corrispondente matrice di permutazione è

$$\boldsymbol{P}_{[13]} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

2. I moltiplicatori sono $m_{21} = 0,\; m_{31} = -1/2,\; m_{41} = 0$. I loro opposti vengono memorizzati nella prima colonna della matrice trasformata. Eliminando la prima incognita dalla seconda, terza e quarta equazione, otteniamo

$$\begin{pmatrix} 2 & -1 & 1 & -2 & 0 \\ 0 & 2 & 0 & -1 & 1 \\ 1/2 & 1/2 & -5/2 & 2 & 0 \\ 0 & 2 & 1 & 1 & 4 \end{pmatrix}.$$

Passo 2

1. L'elemento pivot vale 2. Non si effettuano scambi di righe, $\boldsymbol{P}_{[22]} = \boldsymbol{I}$.

Passo 2

1. L'elemento pivot vale 2. Non si effettuano scambi di righe, $\boldsymbol{P}_{[22]} = \boldsymbol{I}$.
2. I moltiplicatori sono $m_{32} = -1/4,\ m_{42} = -1$. Eliminando la seconda incognita dalla terza e quarta equazione, otteniamo

$$\left(\begin{array}{rrrrr} 2 & -1 & 1 & -2 & 0 \\ 0 & 2 & 0 & -1 & 1 \\ 1/2 & 1/4 & -5/2 & 9/4 & -1/4 \\ 0 & 1 & 1 & 2 & 3 \end{array}\right).$$

Passo 2

1. L'elemento pivot vale 2. Non si effettuano scambi di righe, $\boldsymbol{P}_{[22]} = \boldsymbol{I}$.
2. I moltiplicatori sono $m_{32} = -1/4,\ m_{42} = -1$. Eliminando la seconda incognita dalla terza e quarta equazione, otteniamo

$$\left(\begin{array}{rrrrr} 2 & -1 & 1 & -2 & 0 \\ 0 & 2 & 0 & -1 & 1 \\ 1/2 & 1/4 & -5/2 & 9/4 & -1/4 \\ 0 & 1 & 1 & 2 & 3 \end{array}\right).$$

Passo 3

1. L'elemento pivot vale -5/2. Non si effettuano scambi di righe, $\boldsymbol{P}_{[33]} = \boldsymbol{I}$.

Passo 2

1. L'elemento pivot vale 2. Non si effettuano scambi di righe, $\boldsymbol{P}_{[22]} = \boldsymbol{I}$.
2. I moltiplicatori sono $m_{32} = -1/4,\ m_{42} = -1$. Eliminando la seconda incognita dalla terza e quarta equazione, otteniamo

$$\left(\begin{array}{rrrrr} 2 & -1 & 1 & -2 & 0 \\ 0 & 2 & 0 & -1 & 1 \\ 1/2 & 1/4 & -5/2 & 9/4 & -1/4 \\ 0 & 1 & 1 & 2 & 3 \end{array}\right).$$

Passo 3

1. L'elemento pivot vale -5/2. Non si effettuano scambi di righe, $\boldsymbol{P}_{[33]} = \boldsymbol{I}$.
2. L'unico moltiplicatore è $m_{43} = 2/5$. Eliminando la terza incognita dalla quarta equazione, otteniamo

$$\left(\begin{array}{rrrrr} 2 & -1 & 1 & -2 & 0 \\ 0 & 2 & 0 & -1 & 1 \\ 1/2 & 1/4 & -5/2 & 9/4 & -1/4 \\ 0 & 1 & -2/5 & 29/10 & 29/10 \end{array}\right).$$

Infine, risolvendo per sostituzione all'indietro il sistema triangolare superiore $\boldsymbol{U}\boldsymbol{x} = \boldsymbol{c}$ dato da

$$\begin{pmatrix} 2 & -1 & 1 & -2 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & -5/2 & 9/4 \\ 0 & 0 & 0 & 29/10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1/4 \\ 29/10 \end{pmatrix}$$

Infine, risolvendo per sostituzione all'indietro il sistema triangolare superiore $\boldsymbol{U}\boldsymbol{x} = \boldsymbol{c}$ dato da

$$\begin{pmatrix} 2 & -1 & 1 & -2 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & -5/2 & 9/4 \\ 0 & 0 & 0 & 29/10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1/4 \\ 29/10 \end{pmatrix}$$

otteniamo

$$\begin{aligned} x_4 &= (29/10)/(29/10) = 1 \\ x_3 &= (-1/4 - 9/4)/(-5/2) = 1 \\ x_2 &= (1 - (-1))/2 = 1 \\ x_1 &= -(-2 + 1 - 1)/2 = 1. \end{aligned}$$

Infine, risolvendo per sostituzione all'indietro il sistema triangolare superiore $\boldsymbol{Ux} = \boldsymbol{c}$ dato da

$$\begin{pmatrix} 2 & -1 & 1 & -2 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & -5/2 & 9/4 \\ 0 & 0 & 0 & 29/10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1/4 \\ 29/10 \end{pmatrix}$$

otteniamo

$$\begin{aligned} x_4 &= (29/10)/(29/10) = 1 \\ x_3 &= (-1/4 - 9/4)/(-5/2) = 1 \\ x_2 &= (1 - (-1))/2 = 1 \\ x_1 &= -(-2 + 1 - 1)/2 = 1. \end{aligned}$$

Inoltre, abbiamo

$$\boldsymbol{L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1/2 & 1/4 & 1 & 0 \\ 0 & 1 & -2/5 & 1 \end{pmatrix} \quad \text{e} \quad \boldsymbol{P} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Comandi MATLAB

- `x=A\b` calcola $\boldsymbol{x}$ soluzione di $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ applicando il metodo di eliminazione di Gauss con pivoting parziale seguito da una sostituzione all'indietro.

  Se $\boldsymbol{b}$ è una matrice $n \times m$, cioè un insieme di m vettori colonna, allora $\boldsymbol{x}$ è la matrice che raccoglie le m soluzioni corrispondenti. Questo permette di risolvere diversi sistemi lineari aventi la stessa matrice dei coefficienti, eseguendo una sola eliminazione di Gauss.

Comandi MATLAB

- `x=A\b` calcola $\boldsymbol{x}$ soluzione di $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ applicando il metodo di eliminazione di Gauss con pivoting parziale seguito da una sostituzione all'indietro.

  Se $\boldsymbol{b}$ è una matrice $n \times m$, cioè un insieme di m vettori colonna, allora $\boldsymbol{x}$ è la matrice che raccoglie le m soluzioni corrispondenti. Questo permette di risolvere diversi sistemi lineari aventi la stessa matrice dei coefficienti, eseguendo una sola eliminazione di Gauss.

- `[L U P]=lu(A)` calcola i fattori $\boldsymbol{L}$, $\boldsymbol{U}$, $\boldsymbol{P}$ della fattorizzazione $\boldsymbol{P}\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$ relativa alla matrice $\boldsymbol{A}$

  Esempi di uso di tali fattori vengono dati nel seguito.

Comandi MATLAB

- `x=A\b` calcola $\boldsymbol{x}$ soluzione di $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ applicando il metodo di eliminazione di Gauss con pivoting parziale seguito da una sostituzione all'indietro.

  Se $\boldsymbol{b}$ è una matrice $n \times m$, cioè un insieme di m vettori colonna, allora $\boldsymbol{x}$ è la matrice che raccoglie le m soluzioni corrispondenti. Questo permette di risolvere diversi sistemi lineari aventi la stessa matrice dei coefficienti, eseguendo una sola eliminazione di Gauss.

- `[L U P]=lu(A)` calcola i fattori $\boldsymbol{L}$, $\boldsymbol{U}$, $\boldsymbol{P}$ della fattorizzazione $\boldsymbol{P}\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$ relativa alla matrice $\boldsymbol{A}$

  Esempi di uso di tali fattori vengono dati nel seguito.

- `R=chol(A)` calcola il fattore $\boldsymbol{R}$ triangolare superiore della fattorizzazione di Choleski $\boldsymbol{A} = \boldsymbol{R}^T\boldsymbol{R}$ della matrice simmetrica definita positiva $\boldsymbol{A}$.

Il metodo di Choleski

Sia $\boldsymbol{A}$ una matrice quadrata simmetrica e definita positiva.

Il metodo di eliminazione di Gauss non necessita di *pivoting*, quindi fornisce una fattorizzazione

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$$

in cui gli elementi diagonali di $\boldsymbol{L}$ sono tutti uguali a 1, mentre gli elementi diagonali di $\boldsymbol{U}$ sono tutti > 0.

Il metodo di Choleski

Sia $\boldsymbol{A}$ una matrice quadrata simmetrica e definita positiva.

Il metodo di eliminazione di Gauss non necessita di *pivoting*, quindi fornisce una fattorizzazione

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$$

in cui gli elementi diagonali di $\boldsymbol{L}$ sono tutti uguali a 1, mentre gli elementi diagonali di $\boldsymbol{U}$ sono tutti > 0.

Questo, la simmetria di $\boldsymbol{A}$ e alcune (semplici) trasformazioni matriciali dimostrano che $\boldsymbol{A}$ può fattorizzarsi nella forma

$$\boldsymbol{A} = \boldsymbol{R}^T\boldsymbol{R},$$

dove $\boldsymbol{R}$ è una matrice triangolare superiore con elementi diagonali tutti > 0.

Sia $\boldsymbol{A}$ una matrice quadrata simmetrica e definita positiva.

Il metodo di eliminazione di Gauss non necessita di *pivoting*, quindi fornisce una fattorizzazione

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$$

in cui gli elementi diagonali di $\boldsymbol{L}$ sono tutti uguali a 1, mentre gli elementi diagonali di $\boldsymbol{U}$ sono tutti > 0.

Questo, la simmetria di $\boldsymbol{A}$ e alcune (semplici) trasformazioni matriciali dimostrano che $\boldsymbol{A}$ può fattorizzarsi nella forma

$$\boldsymbol{A} = \boldsymbol{R}^T\boldsymbol{R},$$

dove $\boldsymbol{R}$ è una matrice triangolare superiore con elementi diagonali tutti > 0.

Il **metodo di Choleski** calcola gli elementi di $\boldsymbol{R}$ usando ricorsivamente le relazioni

$$a_{ij} = \sum_{k=1}^{n} r_{ik}^T r_{kj} = \sum_{k=1}^{n} r_{ki} r_{kj} = \sum_{k=1}^{i} r_{ki} r_{kj}, \qquad i \leq j,$$

nell'ordine di indici $ij = 11, 12, 22, 31, 32, 33, \ldots, n(n-1), nn$. Il numero di operazioni è

$$\sim \frac{n^3}{6}.$$

- Se dobbiamo risolvere *diversi* sistemi lineari $\boldsymbol{Ax} = \boldsymbol{b}$ con la *stessa* matrice $\boldsymbol{A}$ e con termini noti $\boldsymbol{b}$ che sono disponibili in momenti diversi, conviene fattorizzare la matrice una sola volta (con costo computazionale $\sim n^3/3$) e poi applicare più volte la sostituzione all'indietro (con costo computazionale $\sim n^2/2$).

- Se dobbiamo risolvere *diversi* sistemi lineari $\boldsymbol{Ax} = \boldsymbol{b}$ con la *stessa* matrice $\boldsymbol{A}$ e con termini noti $\boldsymbol{b}$ che sono disponibili in momenti diversi, conviene fattorizzare la matrice una sola volta (con costo computazionale $\sim n^3/3$) e poi applicare più volte la sostituzione all'indietro (con costo computazionale $\sim n^2/2$).

  Nota: se invece si usa il comando `x=A\b` per ogni termine noto, si esegue ogni volta la eliminazione di Gauss, il che non è efficiente!

- Se dobbiamo risolvere *diversi* sistemi lineari $\boldsymbol{Ax} = \boldsymbol{b}$ con la *stessa* matrice $\boldsymbol{A}$ e con termini noti $\boldsymbol{b}$ che sono disponibili in momenti diversi, conviene fattorizzare la matrice una sola volta (con costo computazionale $\sim n^3/3$) e poi applicare più volte la sostituzione all'indietro (con costo computazionale $\sim n^2/2$).

  Nota: se invece si usa il comando `x=A\b` per ogni termine noto, si esegue ogni volta la eliminazione di Gauss, il che non è efficiente!

- Se si vuole calcolare det $\boldsymbol{A}$, possiamo usare la relazione det $\boldsymbol{P}$ det $\boldsymbol{A}$ = det $\boldsymbol{L}$ det $\boldsymbol{U}$ e osservare che
  - det $\boldsymbol{P} = (-1)^s$, dove s è il numero di scambi di righe effettuati nella permutazione;
  - det $\boldsymbol{L} = \prod_{i=1}^n \ell_{ii} = 1$, perchè ogni elemento $\ell_{ii} = 1$;
  - det $\boldsymbol{U} = \prod_{i=1}^n u_{ii}$.

  Dunque

$$\det \boldsymbol{A} = (-1)^s \prod_{i=1}^{n} u_{ii}.$$

- Se si deve calcolare $\boldsymbol{A}^{-1}$ (quando proprio è indispensabile farlo...), si usa la relazione

$$\boldsymbol{A}^{-1}\boldsymbol{P}^{-1} = \boldsymbol{U}^{-1}\boldsymbol{L}^{-1}$$

cioè

$$\boldsymbol{A}^{-1} = \boldsymbol{U}^{-1}\boldsymbol{L}^{-1}\boldsymbol{P}.$$

L'inversa di una matrice triangolare inferiore o superiore può essere calcolata in modo efficiente risolvendo n sistemi lineari (con il metodo di sostituzione) i cui termini noti sono le colonne della matrice identità $\boldsymbol{I}$.

Complessivamente, il numero di operazioni per calcolare $\boldsymbol{A}^{-1}$ è

$$\sim n^3.$$

Alcune applicazioni della fattorizzazione $PA = LU$

- Se si deve calcolare $\boldsymbol{A}^{-1}$ (quando proprio è indispensabile farlo...), si usa la relazione

$$\boldsymbol{A}^{-1}\boldsymbol{P}^{-1} = \boldsymbol{U}^{-1}\boldsymbol{L}^{-1}$$

cioè

$$\boldsymbol{A}^{-1} = \boldsymbol{U}^{-1}\boldsymbol{L}^{-1}\boldsymbol{P}.$$

L'inversa di una matrice triangolare inferiore o superiore può essere calcolata in modo efficiente risolvendo n sistemi lineari (con il metodo di sostituzione) i cui termini noti sono le colonne della matrice identità $\boldsymbol{I}$.

Complessivamente, il numero di operazioni per calcolare $\boldsymbol{A}^{-1}$ è

$$\sim n^3.$$

Nota: Calcolare $\boldsymbol{A}^{-1}$ e poi $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$ per risolvere un sistema lineare $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ è una stupidaggine da non fare! (tranne per sistemi molto piccoli).
Infatti, eseguire il comando `x=inv(A)*b` costa $\sim n^3 + n^2 \sim n^3$ operazioni, mentre eseguire il comando `x=A\b` costa $\sim n^3/3 + n^2/2 \sim n^3/3$ operazioni.

3. Trasformazioni di Householder e fattorizzazioni QR di matrici

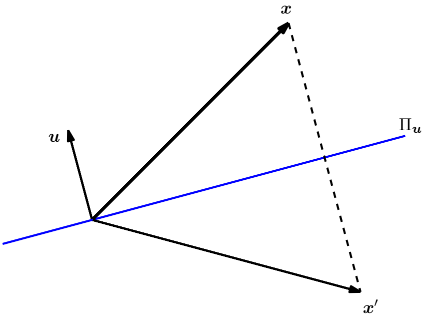
Sia dato un versore $\boldsymbol{u}$ (cioè un vettore di norma euclidea unitaria) e sia $\Pi_{\boldsymbol{u}}$ l'(iper)piano perpendicolare a $\boldsymbol{u}$, ossia il sottospazio di dimensione $n-1$

$$\Pi_{\boldsymbol{u}} = \{\boldsymbol{z} \in \mathbb{R}^n \,:\, \boldsymbol{u}^T \boldsymbol{z} = 0\}.$$

Sia dato un versore $\boldsymbol{u}$ (cioè un vettore di norma euclidea unitaria) e sia $\Pi_{\boldsymbol{u}}$ l'(iper)piano perpendicolare a $\boldsymbol{u}$, ossia il sottospazio di dimensione $n-1$

$$\Pi_{\boldsymbol{u}} = \{\boldsymbol{z} \in \mathbb{R}^n \,:\, \boldsymbol{u}^T \boldsymbol{z} = 0\}.$$

Se interpretiamo $\Pi_{\boldsymbol{u}}$ come uno "specchio", vogliamo costruire la trasformazione che "riflette" un qualunque vettore rispetto a tale specchio, ossia che associa a ogni vettore $\boldsymbol{x}$ la sua immagine speculare $\boldsymbol{x}'$ rispetto a $\Pi_{\boldsymbol{u}}$.

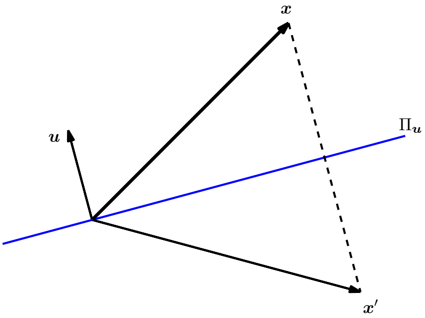


Riflettori elementari

Sia dato un versore $\boldsymbol{u}$ (cioè un vettore di norma euclidea unitaria) e sia $\Pi_{\boldsymbol{u}}$ l'(iper)piano perpendicolare a $\boldsymbol{u}$, ossia il sottospazio di dimensione $n-1$

$$\Pi_{\boldsymbol{u}} = \{\boldsymbol{z} \in \mathbb{R}^n \, : \, \boldsymbol{u}^T \boldsymbol{z} = 0\}.$$

Se interpretiamo $\Pi_{\boldsymbol{u}}$ come uno "specchio", vogliamo costruire la trasformazione che "riflette" un qualunque vettore rispetto a tale specchio, ossia che associa a ogni vettore $\boldsymbol{x}$ la sua immagine speculare $\boldsymbol{x}'$ rispetto a $\Pi_{\boldsymbol{u}}$.



Vedremo presto che questa semplice trasformazione, che risulta sempre numericamente stabile, interviene nella definizione di importanti algoritmi numerici.

Per raggiungere il nostro obiettivo, osserviamo che dato un qualunque vettore $\boldsymbol{x} \in \mathbb{R}^n$, il vettore $(\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u}$ rappresenta la *proiezione ortogonale* di $\boldsymbol{x}$ sulla retta generata da $\boldsymbol{u}$, infatti

$$\boldsymbol{u}^T\left(\boldsymbol{x} - (\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u}\right) = (\boldsymbol{u}^T\boldsymbol{x}) - (\boldsymbol{u}^T\boldsymbol{x})(\boldsymbol{u}^T\boldsymbol{u}) = (\boldsymbol{u}^T\boldsymbol{x}) - (\boldsymbol{u}^T\boldsymbol{x}) = 0.$$

Per raggiungere il nostro obiettivo, osserviamo che dato un qualunque vettore $\boldsymbol{x} \in \mathbb{R}^n$, il vettore $(\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u}$ rappresenta la *proiezione ortogonale* di $\boldsymbol{x}$ sulla retta generata da $\boldsymbol{u}$, infatti

$$\boldsymbol{u}^T\left(\boldsymbol{x} - (\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u}\right) = (\boldsymbol{u}^T\boldsymbol{x}) - (\boldsymbol{u}^T\boldsymbol{x})(\boldsymbol{u}^T\boldsymbol{u}) = (\boldsymbol{u}^T\boldsymbol{x}) - (\boldsymbol{u}^T\boldsymbol{x}) = 0.$$

Notiamo che possiamo scrivere

$$(\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{u}^T\boldsymbol{x}) = (\boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{x} = \boldsymbol{P}_{\boldsymbol{u}}\boldsymbol{x},$$

avendo introdotto la matrice quadrata di ordine n

$$\boldsymbol{P}_{\boldsymbol{u}} = \boldsymbol{u}\boldsymbol{u}^T.$$

Per raggiungere il nostro obiettivo, osserviamo che dato un qualunque vettore $\boldsymbol{x} \in \mathbb{R}^n$, il vettore $(\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u}$ rappresenta la *proiezione ortogonale* di $\boldsymbol{x}$ sulla retta generata da $\boldsymbol{u}$, infatti

$$\boldsymbol{u}^T\Big(\boldsymbol{x} - (\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u}\Big) = (\boldsymbol{u}^T\boldsymbol{x}) - (\boldsymbol{u}^T\boldsymbol{x})(\boldsymbol{u}^T\boldsymbol{u}) = (\boldsymbol{u}^T\boldsymbol{x}) - (\boldsymbol{u}^T\boldsymbol{x}) = 0.$$

Notiamo che possiamo scrivere

$$(\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{u}^T\boldsymbol{x}) = (\boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{x} = \boldsymbol{P}_{\boldsymbol{u}}\boldsymbol{x},$$

avendo introdotto la matrice quadrata di ordine n

$$\boldsymbol{P}_{\boldsymbol{u}} = \boldsymbol{u}\boldsymbol{u}^T.$$

Essa gode delle seguenti proprietà:

Per raggiungere il nostro obiettivo, osserviamo che dato un qualunque vettore $\boldsymbol{x} \in \mathbb{R}^n$, il vettore $(\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u}$ rappresenta la *proiezione ortogonale* di $\boldsymbol{x}$ sulla retta generata da $\boldsymbol{u}$, infatti

$$\boldsymbol{u}^T\Big(\boldsymbol{x}-(\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u}\Big) = (\boldsymbol{u}^T\boldsymbol{x}) - (\boldsymbol{u}^T\boldsymbol{x})(\boldsymbol{u}^T\boldsymbol{u}) = (\boldsymbol{u}^T\boldsymbol{x}) - (\boldsymbol{u}^T\boldsymbol{x}) = 0.$$

Notiamo che possiamo scrivere

$$(\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{u}^T\boldsymbol{x}) = (\boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{x} = \boldsymbol{P_u}\boldsymbol{x},$$

avendo introdotto la matrice quadrata di ordine n

$$\boldsymbol{P_u} = \boldsymbol{u}\boldsymbol{u}^T.$$

Essa gode delle seguenti proprietà:

- è simmetrica, $\boldsymbol{P}_{\boldsymbol{u}}^T = \boldsymbol{P_u}$;

Per raggiungere il nostro obiettivo, osserviamo che dato un qualunque vettore $\boldsymbol{x} \in \mathbb{R}^n$, il vettore $(\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u}$ rappresenta la *proiezione ortogonale* di $\boldsymbol{x}$ sulla retta generata da $\boldsymbol{u}$, infatti

$$\boldsymbol{u}^T\left(\boldsymbol{x} - (\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u}\right) = (\boldsymbol{u}^T\boldsymbol{x}) - (\boldsymbol{u}^T\boldsymbol{x})(\boldsymbol{u}^T\boldsymbol{u}) = (\boldsymbol{u}^T\boldsymbol{x}) - (\boldsymbol{u}^T\boldsymbol{x}) = 0.$$

Notiamo che possiamo scrivere

$$(\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{u}^T\boldsymbol{x}) = (\boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{x} = \boldsymbol{P_u}\boldsymbol{x},$$

avendo introdotto la matrice quadrata di ordine n

$$\boldsymbol{P_u} = \boldsymbol{u}\boldsymbol{u}^T.$$

Essa gode delle seguenti proprietà:

- è simmetrica, $\boldsymbol{P_u}^T = \boldsymbol{P_u}$;
- è una *matrice di proiezione*, cioè soddisfa $\boldsymbol{P_u}\boldsymbol{P_u} = \boldsymbol{P_u}$;

Per raggiungere il nostro obiettivo, osserviamo che dato un qualunque vettore $\boldsymbol{x} \in \mathbb{R}^n$, il vettore $(\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u}$ rappresenta la *proiezione ortogonale* di $\boldsymbol{x}$ sulla retta generata da $\boldsymbol{u}$, infatti

$$\boldsymbol{u}^T\left(\boldsymbol{x} - (\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u}\right) = (\boldsymbol{u}^T\boldsymbol{x}) - (\boldsymbol{u}^T\boldsymbol{x})(\boldsymbol{u}^T\boldsymbol{u}) = (\boldsymbol{u}^T\boldsymbol{x}) - (\boldsymbol{u}^T\boldsymbol{x}) = 0.$$

Notiamo che possiamo scrivere

$$(\boldsymbol{u}^T\boldsymbol{x})\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{u}^T\boldsymbol{x}) = (\boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{x} = \boldsymbol{P_u}\boldsymbol{x},$$

avendo introdotto la matrice quadrata di ordine n

$$\boldsymbol{P_u} = \boldsymbol{u}\boldsymbol{u}^T.$$

Essa gode delle seguenti proprietà:

- è simmetrica, $\boldsymbol{P_u}^T = \boldsymbol{P_u}$;
- è una *matrice di proiezione*, cioè soddisfa $\boldsymbol{P_u}\boldsymbol{P_u} = \boldsymbol{P_u}$;
- ha rango 1, perchè l'immagine è la retta generata da $\boldsymbol{u}$.

Abbiamo dunque la seguente *decomposizione ortogonale* di $\boldsymbol{x}$:

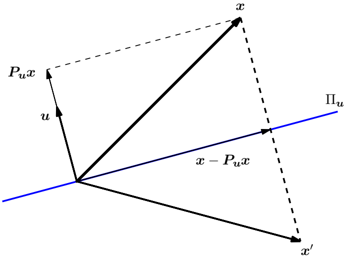
$$\boldsymbol{x} = \boldsymbol{P_u x} + (\boldsymbol{x} - \boldsymbol{P_u x})$$

con $\quad \boldsymbol{P_u x} \perp \Pi_{\boldsymbol{u}} \quad$ e $\quad (\boldsymbol{x} - \boldsymbol{P_u x}) \in \Pi_{\boldsymbol{u}}.$

Abbiamo dunque la seguente *decomposizione ortogonale* di $\boldsymbol{x}$:

$$\boldsymbol{x} = \boldsymbol{P_u x} + (\boldsymbol{x} - \boldsymbol{P_u x})$$

con $\quad \boldsymbol{P_u x} \perp \Pi_{\boldsymbol{u}} \quad$ e $\quad (\boldsymbol{x} - \boldsymbol{P_u x}) \in \Pi_{\boldsymbol{u}}$.



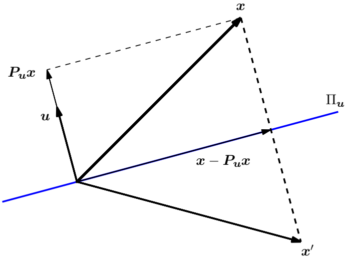
Ora, detta $\boldsymbol{x}'$ l'immagine speculare di $\boldsymbol{x}$ rispetto a $\Pi_{\boldsymbol{u}}$, è immediato convincersi che il vettore differenza $\boldsymbol{x}' - \boldsymbol{x}$ è parallelo a $\boldsymbol{u}$, e precisamente vale

$$\boldsymbol{x}' - \boldsymbol{x} = -2\boldsymbol{P_u x}$$

Abbiamo dunque la seguente *decomposizione ortogonale* di $\boldsymbol{x}$:

$$\boldsymbol{x} = \boldsymbol{P_u x} + (\boldsymbol{x} - \boldsymbol{P_u x})$$

con $\quad \boldsymbol{P_u x} \perp \Pi_{\boldsymbol{u}}$ e $\quad (\boldsymbol{x} - \boldsymbol{P_u x}) \in \Pi_{\boldsymbol{u}}$.



Ora, detta $\boldsymbol{x}'$ l'immagine speculare di $\boldsymbol{x}$ rispetto a $\Pi_{\boldsymbol{u}}$, è immediato convincersi che il vettore differenza $\boldsymbol{x}' - \boldsymbol{x}$ è parallelo a $\boldsymbol{u}$, e precisamente vale

$$\boldsymbol{x}' - \boldsymbol{x} = -2\boldsymbol{P_u x}$$

e dunque

$$\boldsymbol{x}' = \boldsymbol{x} - 2\boldsymbol{P_u x} = (\boldsymbol{I} - 2\boldsymbol{P_u})\boldsymbol{x} = \boldsymbol{H_u x}$$

La matrice

$$\boldsymbol{H_u} = \boldsymbol{I} - 2\boldsymbol{P_u}$$

dicesi **riflettore elementare** o **trasformazione di Householder** associata a $\boldsymbol{u}$.

Valgono per $\boldsymbol{H_u}$ le seguenti proprietà, conseguenza di quelle per $\boldsymbol{P_u}$ e di facile verifica:

Valgono per $\boldsymbol{H_u}$ le seguenti proprietà, conseguenza di quelle per $\boldsymbol{P_u}$ e di facile verifica:

- è simmetrica, $\boldsymbol{H_u}^T = \boldsymbol{H_u}$;

Valgono per $\boldsymbol{H_u}$ le seguenti proprietà, conseguenza di quelle per $\boldsymbol{P_u}$ e di facile verifica:

- è simmetrica, $\boldsymbol{H_u^T} = \boldsymbol{H_u}$;
- è una matrice *involutoria*, cioè soddisfa $\boldsymbol{H_u H_u} = \boldsymbol{I}$;

Valgono per $\boldsymbol{H_u}$ le seguenti proprietà, conseguenza di quelle per $\boldsymbol{P_u}$ e di facile verifica:

- è simmetrica, $\boldsymbol{H}_{\boldsymbol{u}}^T = \boldsymbol{H_u}$;
- è una matrice *involutoria*, cioè soddisfa $\boldsymbol{H_u}\boldsymbol{H_u} = \boldsymbol{I}$;
- dunque è una matrice *ortogonale*, $\boldsymbol{H}_{\boldsymbol{u}}^T\boldsymbol{H_u} = \boldsymbol{I}$;

Valgono per $\boldsymbol{H}_{\boldsymbol{u}}$ le seguenti proprietà, conseguenza di quelle per $\boldsymbol{P}_{\boldsymbol{u}}$ e di facile verifica:

- è simmetrica, $\boldsymbol{H}_{\boldsymbol{u}}^T = \boldsymbol{H}_{\boldsymbol{u}}$;
- è una matrice *involutoria*, cioè soddisfa $\boldsymbol{H}_{\boldsymbol{u}}\boldsymbol{H}_{\boldsymbol{u}} = \boldsymbol{I}$;
- dunque è una matrice *ortogonale*, $\boldsymbol{H}_{\boldsymbol{u}}^T\boldsymbol{H}_{\boldsymbol{u}} = \boldsymbol{I}$;
- dunque ha *numero di condizionamento ottimale* in norma euclidea, $\mathrm{cond}_2(\boldsymbol{H}_{\boldsymbol{u}}) = 1$.

Valgono per $\boldsymbol{H_u}$ le seguenti proprietà, conseguenza di quelle per $\boldsymbol{P_u}$ e di facile verifica:

- è simmetrica, $\boldsymbol{H_u}^T = \boldsymbol{H_u}$;
- è una matrice *involutoria*, cioè soddisfa $\boldsymbol{H_u}\boldsymbol{H_u} = \boldsymbol{I}$;
- dunque è una matrice *ortogonale*, $\boldsymbol{H_u}^T\boldsymbol{H_u} = \boldsymbol{I}$;
- dunque ha *numero di condizionamento ottimale* in norma euclidea, $\text{cond}_2(\boldsymbol{H_u}) = 1$.

Osservazione importante. Tutte le matrici $\boldsymbol{Q}$ *ortogonali* soddisfano

$$\text{cond}_2(\boldsymbol{Q}) = 1.$$

Valgono per $\boldsymbol{H_u}$ le seguenti proprietà, conseguenza di quelle per $\boldsymbol{P_u}$ e di facile verifica:

- è simmetrica, $\boldsymbol{H_u}^T = \boldsymbol{H_u}$;
- è una matrice *involutoria*, cioè soddisfa $\boldsymbol{H_u}\boldsymbol{H_u} = \boldsymbol{I}$;
- dunque è una matrice *ortogonale*, $\boldsymbol{H_u}^T\boldsymbol{H_u} = \boldsymbol{I}$;
- dunque ha *numero di condizionamento ottimale* in norma euclidea, $\text{cond}_2(\boldsymbol{H_u}) = 1$.

Osservazione importante. Tutte le matrici $\boldsymbol{Q}$ *ortogonali* soddisfano

$$\text{cond}_2(\boldsymbol{Q}) = 1.$$

Infatti:

- Per ogni $\boldsymbol{x} \in \mathbb{R}^n$ si ha $\|\boldsymbol{Qx}\|_2 = \|\boldsymbol{x}\|_2$, in quanto

$$\|\boldsymbol{Qx}\|_2^2 = (\boldsymbol{Qx})^T\boldsymbol{Qx} = \boldsymbol{x}^T(\boldsymbol{Q}^T\boldsymbol{Q})\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{x} = \|\boldsymbol{x}\|_2^2;$$

Valgono per $\boldsymbol{H_u}$ le seguenti proprietà, conseguenza di quelle per $\boldsymbol{P_u}$ e di facile verifica:

- è simmetrica, $\boldsymbol{H_u^T} = \boldsymbol{H_u}$;
- è una matrice *involutoria*, cioè soddisfa $\boldsymbol{H_u H_u} = \boldsymbol{I}$;
- dunque è una matrice *ortogonale*, $\boldsymbol{H_u^T H_u} = \boldsymbol{I}$;
- dunque ha *numero di condizionamento ottimale* in norma euclidea, $\text{cond}_2(\boldsymbol{H_u}) = 1$.

Osservazione importante. Tutte le matrici $\boldsymbol{Q}$ *ortogonali* soddisfano

$$\text{cond}_2(\boldsymbol{Q}) = 1.$$

Infatti:

- Per ogni $\boldsymbol{x} \in \mathbb{R}^n$ si ha $\|\boldsymbol{Qx}\|_2 = \|\boldsymbol{x}\|_2$, in quanto

$$\|\boldsymbol{Qx}\|_2^2 = (\boldsymbol{Qx})^T\boldsymbol{Qx} = \boldsymbol{x}^T(\boldsymbol{Q}^T\boldsymbol{Q})\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{x} = \|\boldsymbol{x}\|_2^2;$$

- dunque si ha $\|\boldsymbol{Q}\|_2 = 1$, in quanto

$$\|\boldsymbol{Q}\|_2 = \max_{\boldsymbol{x}\in\mathbb{R}^n \;\; \boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{Qx}\|_2}{\|\boldsymbol{x}\|_2} = 1;$$

Valgono per $\boldsymbol{H}_{\boldsymbol{u}}$ le seguenti proprietà, conseguenza di quelle per $\boldsymbol{P}_{\boldsymbol{u}}$ e di facile verifica:

- è simmetrica, $\boldsymbol{H}_{\boldsymbol{u}}^T = \boldsymbol{H}_{\boldsymbol{u}}$;
- è una matrice *involutoria*, cioè soddisfa $\boldsymbol{H}_{\boldsymbol{u}}\boldsymbol{H}_{\boldsymbol{u}} = \boldsymbol{I}$;
- dunque è una matrice *ortogonale*, $\boldsymbol{H}_{\boldsymbol{u}}^T\boldsymbol{H}_{\boldsymbol{u}} = \boldsymbol{I}$;
- dunque ha *numero di condizionamento ottimale* in norma euclidea, $\text{cond}_2(\boldsymbol{H}_{\boldsymbol{u}}) = 1$.

Osservazione importante. Tutte le matrici $\boldsymbol{Q}$ *ortogonali* soddisfano

$$\text{cond}_2(\boldsymbol{Q}) = 1.$$

Infatti:

- Per ogni $\boldsymbol{x} \in \mathbb{R}^n$ si ha $\|\boldsymbol{Q}\boldsymbol{x}\|_2 = \|\boldsymbol{x}\|_2$, in quanto

$$\|\boldsymbol{Q}\boldsymbol{x}\|_2^2 = (\boldsymbol{Q}\boldsymbol{x})^T\boldsymbol{Q}\boldsymbol{x} = \boldsymbol{x}^T(\boldsymbol{Q}^T\boldsymbol{Q})\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{x} = \|\boldsymbol{x}\|_2^2;$$

- dunque si ha $\|\boldsymbol{Q}\|_2 = 1$, in quanto

$$\|\boldsymbol{Q}\|_2 = \max_{\boldsymbol{x}\in\mathbb{R}^n \ \boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{Q}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} = 1;$$

- analogamente, si ha $\|\boldsymbol{Q}^T\|_2 = 1$;

Valgono per $\boldsymbol{H_u}$ le seguenti proprietà, conseguenza di quelle per $\boldsymbol{P_u}$ e di facile verifica:

- è simmetrica, $\boldsymbol{H_u}^T = \boldsymbol{H_u}$;
- è una matrice *involutoria*, cioè soddisfa $\boldsymbol{H_u}\boldsymbol{H_u} = \boldsymbol{I}$;
- dunque è una matrice *ortogonale*, $\boldsymbol{H_u}^T\boldsymbol{H_u} = \boldsymbol{I}$;
- dunque ha *numero di condizionamento ottimale* in norma euclidea, $\mathrm{cond}_2(\boldsymbol{H_u}) = 1$.

Osservazione importante. Tutte le matrici $\boldsymbol{Q}$ *ortogonali* soddisfano

$$\mathrm{cond}_2(\boldsymbol{Q}) = 1.$$

Infatti:

- Per ogni $\boldsymbol{x} \in \mathbb{R}^n$ si ha $\|\boldsymbol{Qx}\|_2 = \|\boldsymbol{x}\|_2$, in quanto

$$\|\boldsymbol{Qx}\|_2^2 = (\boldsymbol{Qx})^T\boldsymbol{Qx} = \boldsymbol{x}^T(\boldsymbol{Q}^T\boldsymbol{Q})\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{x} = \|\boldsymbol{x}\|_2^2;$$

- dunque si ha $\|\boldsymbol{Q}\|_2 = 1$, in quanto

$$\|\boldsymbol{Q}\|_2 = \max_{\boldsymbol{x}\in\mathbb{R}^n \ \boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{Qx}\|_2}{\|\boldsymbol{x}\|_2} = 1;$$

- analogamente, si ha $\|\boldsymbol{Q}^T\|_2 = 1$;
- dunque concludiamo che $\mathrm{cond}_2(\boldsymbol{Q}) = \|\boldsymbol{Q}\|_2\|\boldsymbol{Q}^T\|_2 = 1$.

Un esempio (notevole) di scelta della direzione di riflessione

Proprietà. Dati due vettori $\boldsymbol{y}$ e $\boldsymbol{z}$ con $\boldsymbol{y} \neq \boldsymbol{z}$ ma $\|\boldsymbol{y}\|_2 = \|\boldsymbol{z}\|_2$, esiste una trasformazione di Householder $\boldsymbol{H_u}$ che riflette $\boldsymbol{y}$ in $\boldsymbol{z}$.

Proprietà. Dati due vettori $\boldsymbol{y}$ e $\boldsymbol{z}$ con $\boldsymbol{y} \neq \boldsymbol{z}$ ma $\|\boldsymbol{y}\|_2 = \|\boldsymbol{z}\|_2$, esiste una trasformazione di Householder $\boldsymbol{H_u}$ che riflette $\boldsymbol{y}$ in $\boldsymbol{z}$.

Il versore $\boldsymbol{u}$ è definito come

$$\boldsymbol{u} = \frac{\boldsymbol{y} - \boldsymbol{z}}{\|\boldsymbol{y} - \boldsymbol{z}\|_2}.$$

Proprietà. Dati due vettori $\boldsymbol{y}$ e $\boldsymbol{z}$ con $\boldsymbol{y} \neq \boldsymbol{z}$ ma $\|\boldsymbol{y}\|_2 = \|\boldsymbol{z}\|_2$, esiste una trasformazione di Householder $\boldsymbol{H_u}$ che riflette $\boldsymbol{y}$ in $\boldsymbol{z}$.

Il versore $\boldsymbol{u}$ è definito come

$$\boldsymbol{u} = \frac{\boldsymbol{y} - \boldsymbol{z}}{\|\boldsymbol{y} - \boldsymbol{z}\|_2}.$$

Infatti

$$\boldsymbol{H_u y} \;=\; \boldsymbol{y} - 2(\boldsymbol{u}^T \boldsymbol{y})\boldsymbol{u} \;=\; \boldsymbol{y} - \underbrace{\frac{2(\boldsymbol{y} - \boldsymbol{z})^T \boldsymbol{y}}{\|\boldsymbol{y} - \boldsymbol{z}\|_2^2}}_{=1}(\boldsymbol{y} - \boldsymbol{z}) \;=\; \boldsymbol{y} - (\boldsymbol{y} - \boldsymbol{z}) \;=\; \boldsymbol{z}$$

Proprietà. Dati due vettori $\boldsymbol{y}$ e $\boldsymbol{z}$ con $\boldsymbol{y} \neq \boldsymbol{z}$ ma $\|\boldsymbol{y}\|_2 = \|\boldsymbol{z}\|_2$, esiste una trasformazione di Householder $\boldsymbol{H}_{\boldsymbol{u}}$ che riflette $\boldsymbol{y}$ in $\boldsymbol{z}$.

Il versore $\boldsymbol{u}$ è definito come

$$\boldsymbol{u} = \frac{\boldsymbol{y} - \boldsymbol{z}}{\|\boldsymbol{y} - \boldsymbol{z}\|_2}.$$

Infatti

$$\boldsymbol{H}_{\boldsymbol{u}}\boldsymbol{y} \;=\; \boldsymbol{y} - 2(\boldsymbol{u}^T\boldsymbol{y})\boldsymbol{u} \;=\; \boldsymbol{y} - \underbrace{\frac{2(\boldsymbol{y} - \boldsymbol{z})^T\boldsymbol{y}}{\|\boldsymbol{y} - \boldsymbol{z}\|_2^2}}_{=1}(\boldsymbol{y} - \boldsymbol{z}) \;=\; \boldsymbol{y} - (\boldsymbol{y} - \boldsymbol{z}) \;=\; \boldsymbol{z}$$

in quanto, ricordando che $\|\boldsymbol{y}\|_2 = \|\boldsymbol{z}\|_2$, si ha

$$2(\boldsymbol{y} - \boldsymbol{z})^T\boldsymbol{y} = 2\|\boldsymbol{y}\|_2^2 - 2\boldsymbol{z}^T\boldsymbol{y} = \|\boldsymbol{y}\|_2^2 - 2\boldsymbol{z}^T\boldsymbol{y} + \|\boldsymbol{z}\|_2^2 = \|\boldsymbol{y} - \boldsymbol{z}\|_2^2.$$

Applicazione:
come annullare tutte le componenti di un vettore, tranne una

Mediante una trasformazione di Householder, possiamo riflettere un vettore dato in un vettore parallelo al primo vettore della base canonica di $\mathbb{R}^n$.

Applicazione: come annullare tutte le componenti di un vettore, tranne una

Mediante una trasformazione di Householder, possiamo riflettere un vettore dato in un vettore parallelo al primo vettore della base canonica di $\mathbb{R}^n$.

Precisamente, sia $\boldsymbol{y} \neq \boldsymbol{0}$ il vettore dato. Poniamo $\sigma = +\|\boldsymbol{y}\|_2$ oppure $\sigma = -\|\boldsymbol{y}\|_2$ (discutiamo tra un momento quale segno scegliere). Definiamo

$$\boldsymbol{z} = \sigma \boldsymbol{e}_1 = \sigma \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \sigma \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Applicazione:
come annullare tutte le componenti di un vettore, tranne una

Mediante una trasformazione di Householder, possiamo riflettere un vettore dato in un vettore parallelo al primo vettore della base canonica di $\mathbb{R}^n$.

Precisamente, sia $\boldsymbol{y} \neq \boldsymbol{0}$ il vettore dato. Poniamo $\sigma = +\|\boldsymbol{y}\|_2$ oppure $\sigma = -\|\boldsymbol{y}\|_2$ (discutiamo tra un momento quale segno scegliere). Definiamo

$$\boldsymbol{z} = \sigma \boldsymbol{e}_1 = \sigma \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \sigma \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Allora, la matrice ortogonale $\boldsymbol{Q} = \boldsymbol{H}_{\boldsymbol{u}}$ introdotta nella slide precedente relativamente ai vettori $\boldsymbol{y}$ e $\boldsymbol{z}$, soddisfa

$$\boldsymbol{Q}\boldsymbol{y} = \boldsymbol{z}.$$

Applicazione:
come annullare tutte le componenti di un vettore, tranne una

Mediante una trasformazione di Householder, possiamo riflettere un vettore dato in un vettore parallelo al primo vettore della base canonica di $\mathbb{R}^n$.

Precisamente, sia $\boldsymbol{y} \neq \mathbf{0}$ il vettore dato. Poniamo $\sigma = +\|\boldsymbol{y}\|_2$ oppure $\sigma = -\|\boldsymbol{y}\|_2$ (discutiamo tra un momento quale segno scegliere). Definiamo

$$\boldsymbol{z} = \sigma \boldsymbol{e}_1 = \sigma \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \sigma \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Allora, la matrice ortogonale $\boldsymbol{Q} = \boldsymbol{H}_{\boldsymbol{u}}$ introdotta nella slide precedente relativamente ai vettori $\boldsymbol{y}$ e $\boldsymbol{z}$, soddisfa

$$\boldsymbol{Q}\boldsymbol{y} = \boldsymbol{z}.$$

Il vettore $\boldsymbol{u}$ è ottenuto normalizzando il vettore $\boldsymbol{y} - \sigma \boldsymbol{e}_1 = (y_1 - \sigma, y_2, y_3, \ldots, y_n)^T$.

Per evitare gli effetti negativi della cancellazione numerica, il segno di σ è scelto in modo discorde rispetto a quello di y_1.

La fattorizzazione QR di una matrice

Mediante successive trasformazioni di Householder, possiamo fattorizzare una matrice quadrata $\boldsymbol{A}$ nella forma

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$$

dove $\boldsymbol{Q}$ è una matrice ortogonale e $\boldsymbol{R}$ è una matrice triangolare superiore.

Mediante successive trasformazioni di Householder, possiamo fattorizzare una matrice quadrata $\boldsymbol{A}$ nella forma

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$$

dove $\boldsymbol{Q}$ è una matrice ortogonale e $\boldsymbol{R}$ è una matrice triangolare superiore.

A tale scopo, scriviamo la matrice come $\boldsymbol{A} = (\boldsymbol{a}_{:,1} \quad \boldsymbol{a}_{:,2} \quad \cdots \quad \boldsymbol{a}_{:,n})$, dove $\boldsymbol{a}_{:,j}$ indica il j-esimo vettore colonna di $\boldsymbol{A}$.

Sia $\boldsymbol{Q}_1$ la trasformazione di Householder, definita nella slide precedente, che manda il primo vettore colonna di $\boldsymbol{A}$ (certamente non nullo, essendo la matrice non-singolare) in un multiplo del primo vettore della base canonica, precisamente sia

$$\boldsymbol{Q}_1 \boldsymbol{a}_{:,1} = r_{11} \boldsymbol{e}_1 = \begin{pmatrix} r_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \text{(con } \; r_{11} \neq 0).$$

La fattorizzazione QR di una matrice

Mediante successive trasformazioni di Householder, possiamo fattorizzare una matrice quadrata $\boldsymbol{A}$ nella forma

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$$

dove $\boldsymbol{Q}$ è una matrice ortogonale e $\boldsymbol{R}$ è una matrice triangolare superiore.

A tale scopo, scriviamo la matrice come $\boldsymbol{A} = (\boldsymbol{a}_{:,1} \quad \boldsymbol{a}_{:,2} \quad \cdots \quad \boldsymbol{a}_{:,n})$, dove $\boldsymbol{a}_{:,j}$ indica il j-esimo vettore colonna di $\boldsymbol{A}$.

Sia $\boldsymbol{Q}_1$ la trasformazione di Householder, definita nella slide precedente, che manda il primo vettore colonna di $\boldsymbol{A}$ (certamente non nullo, essendo la matrice non-singolare) in un multiplo del primo vettore della base canonica, precisamente sia

$$\boldsymbol{Q}_1 \boldsymbol{a}_{:,1} = r_{11} \boldsymbol{e}_1 = \begin{pmatrix} r_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad (\text{con} \quad r_{11} \neq 0).$$

Indichiamo le immagini degli altri vettori colonna come

$$\boldsymbol{Q}_1 \boldsymbol{a}_{:,j} = \begin{pmatrix} r_{1j} \\ a_{2j}^{(1)} \\ \vdots \\ a_{nj}^{(1)} \end{pmatrix}.$$

Dunque dopo questo primo passo otteniamo

$$\boldsymbol{Q}_1\boldsymbol{A} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix},$$

Dunque dopo questo primo passo otteniamo

$$Q_1 A = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix},$$

che possiamo scrivere in forma compatta a blocchi come

$$Q_1 A = \begin{pmatrix} r_{11} & \boldsymbol{r}_1 \\ \mathbf{0} & \widetilde{A}^{(1)} \end{pmatrix}.$$

Dunque dopo questo primo passo otteniamo

$$\boldsymbol{Q}_1\boldsymbol{A} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix},$$

che possiamo scrivere in forma compatta a blocchi come

$$\boldsymbol{Q}_1\boldsymbol{A} = \begin{pmatrix} r_{11} & \boldsymbol{r}_1 \\ \boldsymbol{0} & \widetilde{\boldsymbol{A}}^{(1)} \end{pmatrix}.$$

Nel secondo passo, applichiamo la procedura precedente alla matrice $\widetilde{\boldsymbol{A}}^{(1)}$ di ordine $n-1$, che è ancora non-singolare. Precisamente, sia $\widetilde{\boldsymbol{Q}}_2$ la matrice di Householder tale che trasforma la prima colonna di $\widetilde{\boldsymbol{A}}^{(1)}$ in un vettore di tipo

$$\begin{pmatrix} r_{22} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{n-1}$$

con $r_{22} \neq 0$.

Se definiamo

$$\boldsymbol{Q}_2 = \begin{pmatrix} 1 & \boldsymbol{0}^T \\ \boldsymbol{0} & \widetilde{\boldsymbol{Q}}_2 \end{pmatrix},$$

otteniamo

$$\boldsymbol{Q}_2\boldsymbol{Q}_1\boldsymbol{A} = \begin{pmatrix} r_{11} & \boldsymbol{r}_1 \\ \boldsymbol{0} & \widetilde{\boldsymbol{Q}}_2\widetilde{\boldsymbol{A}}^{(1)} \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix},$$

Se definiamo

$$\boldsymbol{Q}_2 = \begin{pmatrix} 1 & \boldsymbol{0}^T \\ \boldsymbol{0} & \widetilde{\boldsymbol{Q}}_2 \end{pmatrix},$$

otteniamo

$$\boldsymbol{Q}_2\boldsymbol{Q}_1\boldsymbol{A} = \begin{pmatrix} r_{11} & \boldsymbol{r}_1 \\ \boldsymbol{0} & \widetilde{\boldsymbol{Q}}_2\widetilde{\boldsymbol{A}}^{(1)} \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix},$$

che possiamo scrivere in forma compatta a blocchi come

$$\boldsymbol{Q}_2\boldsymbol{Q}_1\boldsymbol{A} = \begin{pmatrix} \boldsymbol{R}_{11} & \boldsymbol{R}_{12} \\ \boldsymbol{0} & \widetilde{\boldsymbol{A}}^{(2)} \end{pmatrix}.$$

con $\boldsymbol{R}_{11} \in \mathbb{R}^{2\times 2}$, $\boldsymbol{R}_{12} \in \mathbb{R}^{2\times(n-2)}$ e $\widetilde{\boldsymbol{A}}^{(2)} \in \mathbb{R}^{(n-2)\times(n-2)}$.

Dopo $n-1$ passi di questo tipo, otteniamo

$$\boldsymbol{Q}_{n-1}\cdots\boldsymbol{Q}_2\boldsymbol{Q}_1\boldsymbol{A} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & r_{33} & \cdots & r_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r_{nn} \end{pmatrix} = \boldsymbol{R},$$

dove, per $k \geq 2$,

$$\boldsymbol{Q}_k = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0}^T \\ \boldsymbol{0} & \widetilde{\boldsymbol{Q}}_k \end{pmatrix} \qquad \text{con} \quad \boldsymbol{I} \in \mathbb{R}^{(k-1)\times(k-1)} \quad \text{e} \quad \widetilde{\boldsymbol{Q}}_k \in \mathbb{R}^{(n-k+1)\times(n-k+1)}.$$

Gli elementi diagonali della matrice $\boldsymbol{R}$ sono tutti $\neq 0$.

Dopo $n-1$ passi di questo tipo, otteniamo

$$\boldsymbol{Q}_{n-1}\cdots\boldsymbol{Q}_2\boldsymbol{Q}_1\boldsymbol{A} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & r_{33} & \cdots & r_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r_{nn} \end{pmatrix} = \boldsymbol{R},$$

dove, per $k \geq 2$,

$$\boldsymbol{Q}_k = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0}^T \\ \boldsymbol{0} & \widetilde{\boldsymbol{Q}}_k \end{pmatrix} \qquad \text{con} \;\; \boldsymbol{I} \in \mathbb{R}^{(k-1)\times(k-1)} \;\; \text{e} \;\; \widetilde{\boldsymbol{Q}}_k \in \mathbb{R}^{(n-k+1)\times(n-k+1)}.$$

Gli elementi diagonali della matrice $\boldsymbol{R}$ sono tutti $\neq 0$.

Ricordando le proprietà delle matrici di Householder, è facile vedere che $\boldsymbol{Q}_k^{-1} = \boldsymbol{Q}_k^T = \boldsymbol{Q}_k$ e dunque

$$\boldsymbol{A} = \underbrace{\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_{n-1}}_{\boldsymbol{Q}}\boldsymbol{R} = \boldsymbol{Q}\boldsymbol{R}$$

con $\boldsymbol{Q}^{-1} = \boldsymbol{Q}^T$.

- Il numero di operazioni necessarie per calcolare la fattorizzazione $\boldsymbol{QR}$ di una matrice quadrata di ordine n è

$$\sim \frac{2n^3}{3}.$$

- Il numero di operazioni necessarie per calcolare la fattorizzazione $\boldsymbol{QR}$ di una matrice quadrata di ordine n è
$$\sim \frac{2n^3}{3}.$$
- Pertanto, se si vuole risolvere un sistema lineare $\boldsymbol{Ax} = \boldsymbol{b}$, *non è conveniente calcolare la fattorizzazione* $\boldsymbol{QR}$ *di* $\boldsymbol{A}$, in quanto il metodo di eliminazione di Gauss (o fattorizzazione $\boldsymbol{LU}$) risolve il problema con $\sim n^3/3$ operazioni.

- Il numero di operazioni necessarie per calcolare la fattorizzazione $\boldsymbol{QR}$ di una matrice quadrata di ordine n è
$$\sim \frac{2n^3}{3}.$$
- Pertanto, se si vuole risolvere un sistema lineare $\boldsymbol{Ax} = \boldsymbol{b}$, *non è conveniente calcolare la fattorizzazione* $\boldsymbol{QR}$ *di* $\boldsymbol{A}$, in quanto il metodo di eliminazione di Gauss (o fattorizzazione $\boldsymbol{LU}$) risolve il problema con $\sim n^3/3$ operazioni.
- Come vedremo, la fattorizzazione $\boldsymbol{QR}$ risulta invece di fondamentale importanza per risolvere altri tipi di problemi algebrici, a causa della sua stabilità numerica.

"Householder reflections have impeccable numerical credentials"

(*Cleve Moler, padre di Matlab*)

- Il numero di operazioni necessarie per calcolare la fattorizzazione $\boldsymbol{QR}$ di una matrice quadrata di ordine n è

$$\sim \frac{2n^3}{3}.$$

- Pertanto, se si vuole risolvere un sistema lineare $\boldsymbol{Ax} = \boldsymbol{b}$, *non è conveniente calcolare la fattorizzazione $\boldsymbol{QR}$ di $\boldsymbol{A}$*, in quanto il metodo di eliminazione di Gauss (o fattorizzazione $\boldsymbol{LU}$) risolve il problema con $\sim n^3/3$ operazioni.

- Come vedremo, la fattorizzazione $\boldsymbol{QR}$ risulta invece di fondamentale importanza per risolvere altri tipi di problemi algebrici, a causa della sua stabilità numerica.

"Householder reflections have impeccable numerical credentials"

(*Cleve Moler, padre di Matlab*)

Comando MATLAB

- `[Q R]=qr(A)` calcola i fattori $\boldsymbol{Q}$ e $\boldsymbol{R}$ della fattorizzazione $\boldsymbol{A} = \boldsymbol{QR}$ relativa alla matrice $\boldsymbol{A}$