Corso triennale di Informatica

Sistemi lineari sovra-determinati e sotto-determinati

R. Vermiglio

DIPARTIMENTO DI SCIENZE MATEMATICHE, INFORMATICHE E FISICHE

Università degli Studi di Udine

Outline

- Sistemi sovra-determinati
- Sistemi sotto-derminati

Soluzione di sistemi lineari con matrice rettangolare

Vogliamo trovare un vettore $x \in \mathbb{R}^n$ che è soluzione del sistema lineare

$$Ax = b$$

\# di INCOGNITE

con $A \in \mathbb{R}^{m \times n}$ matrice **rettangolare**, i.e. $\mathbf{m} \neq \mathbf{n}$ e $b \in \mathbb{R}^m$ termine noto.

\# di EQUAZIONI

Il sistema si definisce

- **sotto-determinato** se $m < n$ e la matrice A risulta "grassa"

  $A = $ [$m < n$] m

  → poche equazioni
  tante incognite → sotto determinato

- **sovra-determinato** se $m > n$ e la matrice risulta "magra"

  $A = $ [$m > n$]

  → molte equazioni
  poche incognite → sovradeterminato

Soluzione di sistemi lineari sovra-determinati

Tale sistema è caratterizzato da più equazioni (m) che incognite (n) in generale non ha soluzione. Vale

Theorem

Il sistema ha soluzione se e soltanto se $b \in \overset{\text{SOTTOSPAZIO}}{span}\{a^{(1)}, \ldots, a^{(n)}\}$*, dove* $a^{(j)} \in \mathbb{R}^m$ *è la colonna j-esima di A*, $j = 1, \ldots, n$.

colonna di A

ogmi colonna è um vettore m x 1

righe

chiamiamo SOLUZIONE di un problema SOVRADETERMINATO il vettore $\bar{x}$ t.c. il RESIDUO $b-A\bar{x}$ ha il valore più piccolo possibile

Ci proponiamo di trovare il vettore $\bar{x}$ che minimizza la norma due del residuo (soluzione del sistema nel senso dei **minimi quadrati**):

$$||r(\bar{x})||_2 = ||b - A\bar{x}||_2 \leq ||r(x)||_2 = ||b - Ax||_2, \ \forall x \in \mathbb{R}^n$$

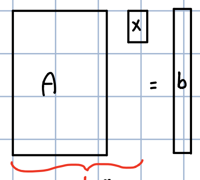
Supponiamo che il $\mathrm{rango}(A) = n$. Vale

Theorem

$\bar{x}$ è soluzione nel senso dei minimi quadrati se e soltanto se

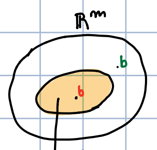
$$r(\bar{x}) = b - A\bar{x} \perp a^{(j)}, j = 1, \ldots, n.$$

$a^{(J)} \in \mathbb{R}^m$ colonna J-esima di A



$$\underbrace{A}_{\sum_{J=1}^{m} a^{(J)} \cdot x_J} x = b$$

Questo sistema ammette SOLUZIONE sse $b \in$ SPAN $\{a^{(0)}, a^{(1)}, \dots, a^{(m)}\}$ con $a^{(i)}$ la i-esima colonna (vettore colonna) di A



$\to \{Ax \mid x \in \mathbb{R}^n\}$

$\to$ ha dimensione m se le colonne sono LIN. INDIP o inferiore se non sono INDIP.

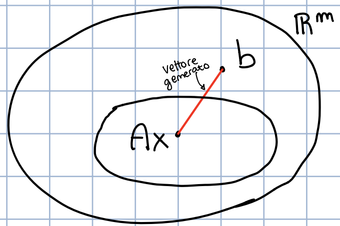
e b può appartenervi come no: - b appartiene $\to \exists$ soluz.

- b NON appartiene $\to \nexists$ soluz. (non riesco a raggiungerlo con tutte le combinaz. possibili)

Si definisce SOLUZIONE di un sistema sovradeterminato un elemento del sottospazio (●) che ha DISTANZA MINIMA da b

si definisce, per la distanza, la norma 2 con generico vettore Ax

$\|\cdot\|_2 \Rightarrow \|b - Ax\|_2$

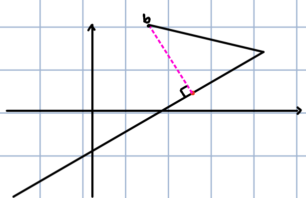


faccio variare x per trovare $\tilde{x}$, se esiste, t.c. $\underbrace{\|b - A\tilde{x}\|_2}_{\text{DISTANZA MINIMA}} = \min_{x \in \mathbb{R}^n} \|b - Ax\|_2$

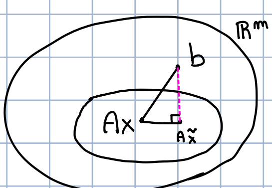
Si chiama soluz. dei minimi quadrati perché usa la NORMA 2, che in generale è $\|x\|_2^2 = \sum_{i=1}^{m} x_i^2$ è la norma euclidea

Associato a questa norma c'è il PRODOTTO SCALARE $\langle x, y \rangle = \sum_{i=1}^{m} x_i y_i \Rightarrow \|x\|_2^2 = \langle x, x \rangle$ N.B. $\langle x, y \rangle = x^T y = y^T x$

e si definisce il concetto di ORTOGONALITA' $x \perp y$ sse $\langle x, y \rangle = 0$



$\to$ Se ho un punto b e una retta e voglio trovare la DISTANZA MINIMA (dalla retta) faccio la PROIEZ. ORTOGONALE



$\to$ deve essere quindi che il vettore $b - Ax$ sia ORTOGONALE a tutti i vettori dello spazio (generato dalle colonne)

Supponendo le colonne lin. INDIPENDENTI così il sottospazio $\{Ax : x \in \mathbb{R}^m\}$ sia $= m$

$\Rightarrow$ $\tilde{x}$ è soluz. di $Ax = b$ (nei minimi quadrati) sse $b - A\tilde{x} \perp a^{(J)}, J = 1, \dots, m$

che vuol dire $\langle a^{(J)}, b - A\tilde{x} \rangle = 0 \Rightarrow a^{(J)T}(b - A\tilde{x}) = 0 \Rightarrow \underbrace{a^{(J)T}}_{A^T} b = \underbrace{a^{(J)T}}_{A^T} A\tilde{x} \Rightarrow \boxed{A^T A x = A^T b}$

DIM. del teo di PITAGORA:

$$\underbrace{b - Ax}_{r(x)} = b - A\tilde{x} + A\tilde{x} - Ax$$

$$= \underbrace{b - A\tilde{x}}_{r(\tilde{x})} + \underbrace{A(\tilde{x} - x)}_{y \in \{Ax \mid x \in \mathbb{R}^n\}}$$ ma dato che sono ORTOGONALI

DISTANZA

$$\|r(x)\|_2^2 = \langle r(x), r(x) \rangle$$ → Teorema di Pitagora

$$= \langle r(\tilde{x}) + y, r(\tilde{x}) + y \rangle$$

$$\leq \|r(\tilde{x})\|_2^2 + \underbrace{2\langle r(\tilde{x}), y \rangle}_{\substack{=0 \\ \text{perchè} \\ \text{ORTOGONALI}}} + \underbrace{\|y\|_2^2}_{\geq 0}$$

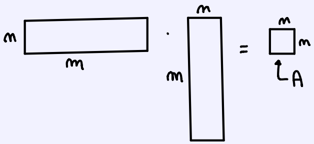
↳ un qualsiasi vettore, perciò la sua norma è ≥ 0

Soluzione di sistemi lineari sovra-determinati, continua

Equazioni normali

Il calcolo di $\bar{x}$ si riconduce quindi alla risoluzione del **sistema di equazioni normali**

$$A^T A x = A^T b$$



che è un sistema quadrato

che richiede i seguenti passi

- calcolo di $M = A^T A$
- calcolo di $d = A^T b$
- risoluzione di $Mx = d$

ma la matrice di questo nuovo sistema è NON SINGOLARE? sí se le colonne sono LIN. INDIP. (cosa che abbiamo assunto in precedenza)

M in realtà è anche SIMMETRICA perché $M^T = (A^T \cdot A)^T = A^T \cdot \underbrace{(A^T)^T}_{A} = A^T \cdot A = M$

M è definita POSITIVA $\rightarrow x^T M x = x^T A^T A x = (Ax)^T A x = \|Ax\|_2^2 > 0$

calcolo solo la METÀ della matrice perché è SIMMETRICA

M è **simmetrica e definita positiva** quindi devo calcolare solo $\frac{n(n+1)}{2}$ elementi e il sistema di equazioni normali di dimensione n può essere risolto con l'algoritmo di Cholesky.

Soluzione di sistemi lineari sovra-determinati, continua

Difficoltà numeriche per le equazioni normali

Un aspetto numericamente delicato riguarda il calcolo di M : gli errori di arrotondamento possono far perdere alla matrice la proprietà di essere definita positiva. Esempio:

↳ ovvero invertibile

ϵ positivo → quindi le colonne sono lin. INDIP.

$$A = \begin{pmatrix} 1 & 1 \\ \epsilon & 0 \\ 0 & \epsilon \end{pmatrix} \rightsquigarrow A^T A = \begin{pmatrix} 1+\epsilon^2 & 1 \\ 1 & 1+\epsilon^2 \end{pmatrix}$$

che risulta singolare se $\epsilon^2 < u$.

→ PRECISIONE di MACCHINA

e quando $|\epsilon^2| < u$ allora $\epsilon^2 \to 0$ e quindi $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ SINGOLARE

Anche il condizionamento può peggiorare in quanto vale

$$(cond_2(A) = ||A^\dagger||||A||)^2 \approx cond_2(M)$$

dove $A^\dagger = (A^T A)^{-1} A^T$ è la **pseudoinversa** di A

Soluzione di sistemi lineari sovra-determinati, continua

Vogliamo costruire un approccio alternativo trasformando il sistema lineare in un altro che abbia la **stessa soluzione** ma che sia **facilmente** risolvibile.

Quali sono le **trasformazioni** che non modificano la soluzione di un sistema lineare sovra-determinato nel senso dei minimi quadrati?

Quali sono i casi **facilmente** risolvibili?

Soluzione di sistemi lineari sovra-determinati

Casi semplici

CASI SEMPLICE 1:

$A = R$ con R **triangolare superiore**, i.e.

$$R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix} \begin{array}{l} \leftarrow \mathbb{R}^{n \times n} \\ \leftarrow \mathbb{R}^{(m-n) \times n} \end{array}$$

con R_1 triangolare superiore. Sia

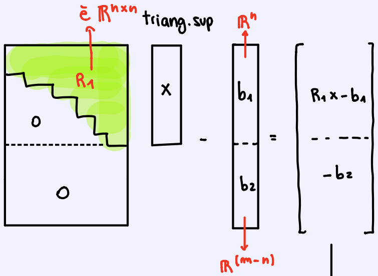
$$b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \begin{array}{l} \leftarrow \mathbb{R}^{n} \\ \leftarrow \mathbb{R}^{m-n} \end{array},$$

allora $\bar{x}$ è soluzione nel senso dei minimi quadrati se

$$R_1 \bar{x} = b_1$$

ed inoltre

$$||r(\bar{x})||_2 = ||b - A\bar{x}||_2 = ||b_2||_2$$



per trovare x che minimizza questo vettore (basta solo guardare in alto)

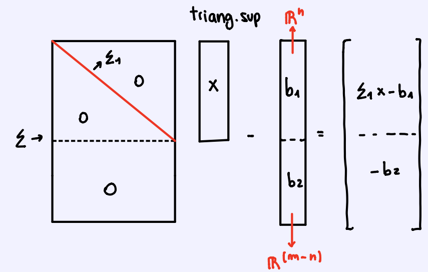
$R_1 x = b_1 \Rightarrow$ e trovo $\tilde{x}$, il residuo

$\| A\tilde{x} - b\| = \left\| \begin{bmatrix} 0 \\ -b_2 \end{bmatrix} \right\| = \|b_2\|_2$

Soluzione di sistemi lineari sovra-determinati

Casi semplici

CASO SEMPLICE 2:



$A = \Sigma$ con Σ **diagonale**, i.e.

$$\Sigma = \begin{pmatrix} \Sigma_1 \\ 0 \end{pmatrix} \begin{array}{l} \leftarrow \mathbb{R}^{n \times n} \\ \leftarrow \mathbb{R}^{(m-n) \times n} \end{array}$$

con $\Sigma_1 \in \mathbb{R}^{n \times n}$ ~~triangolare~~ DIAGONALE superiore. Sia

$$b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \begin{array}{l} \leftarrow \mathbb{R}^{n} \\ \leftarrow \mathbb{R}^{m-n} \end{array},$$

allora $\bar{x}$ è soluzione nel senso dei minimi quadrati se

$$\Sigma_1 \bar{x} = b_1$$

ed inoltre

$$||r(\bar{x})||_2 = ||b - A\bar{x}||_2 = ||b_2||_2$$

Soluzione di sistemi lineari sovra-determinati
Trasformazioni possibili

matrici la cui INVERSA è data dalla trasposta
↳ facilmente invertibili

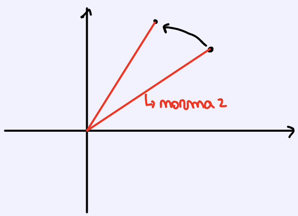
Qualsiasi trasformazione descritta da una matrice **ortogonale** $Q \in \mathbb{R}^{m \times m}$, i.e. $QQ^T = Q^T Q = I$.

la norma 2 non cambia se ruoto il vettore o se lo proietto

Vale infatti

norma che non viene modificato

- $||Qx||_2 = ||x||_2, \quad \forall x \in \mathbb{R}^m$
- $||QB||_2 = ||B||_2 \quad \forall B \in \mathbb{R}^{m \times m}$

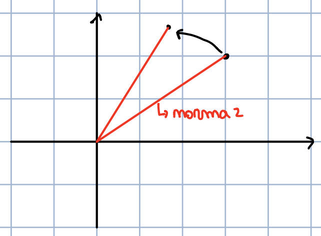


Esempio

La matrice quadrata di rotazione di un angolo α di dimensione $n = 2$

$$Q = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

↳ è una MATRICE ORTOGONALE



la norma 2 non cambia se ruoto il vettore o se lo proietto

$$\|Qx\|_2^2 = \langle Qx, Qx \rangle = (Qx)^T Qx = x^T \underbrace{Q^T Q}_{I} x = x^T x = \|x\|_2^2$$

La fattorizzazione QR

Riusciamo a ricondurci dalla matrice A ad una matrice TRIANGOLARE attraverso delle TRASFORMAZIONI ORTOGONALI

Vale

Theorem

Ogni matrice $A \in \mathbb{R}^{m\times n}$ ammette una fattorizzazione QR i.e.

$$A = QR$$

con $Q \in \mathbb{R}^{m\times m}$ **ortogonale** *e $R \in \mathbb{R}^{m\times n}$* **triangolare superiore**.

Data $A = QR$ osservando che

perchè Q è ortogonale e il residuo non cambia

$$||b - Ax||_2 = ||Q^T(b - Ax)||_2 = ||\overset{Q^Tb}{d} - \overset{Q^TAx}{Rx}||_2$$

con $d = Q^T b$, ci si riconduce al caso di una matrice triangolare superiore.

La fattorizzazione QR, continua

I passi computazionali da compiere sono

- calcolo $A = QR$
- calcolo di $d = Q^T b = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$
- risoluzione di $R_1 x = d_1$
- $||r(\bar{x})||_2 = ||d_2||_2$ → calcolo del RESIDUO

Il costo è dominato dalla fattorizzazione QR ed è stimato in $n^2(m + n/3)$ flop

Decomposizione a valori singolari *SVD*

Un'altra TRASFORMAZIONE, che mi porta alla forma DIAGONALE

↳ come con Gauss Jordan devo eliminare gli elementi che stanno sopra e sotto la diagonale

Vale

Theorem

Ogni matrice $A \in \mathbb{R}^{m\times n}$ ammette una decomposizione a valori singolari SVD i.e.

$$A = U\Sigma V^T$$

*con $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ **ortogonali** e $\Sigma \in \mathbb{R}^{m\times n}$ **diagonale**.*

Gli elementi di Σ si chiamano **valori singolari** e sono non negativi. Data $A = U\Sigma V^T$ osservando che

moltiplico il RESIDUO per la trasposta di U perché so che se è ORTOGONALE il RESIDUO NON CAMBIA

$$||b - Ax||_2 = ||U^T(b - Ax)||_2 = ||d - \underbrace{\Sigma y}_{U^TAx = V^Tx = y}||_2$$

con $d = U^T b$ e $y = V^T x$, ci si riconduce al caso di una matrice diagonale.

$$\underbrace{U^T b}_{d} - U^T A x = d - \Sigma \underbrace{V^T x}_{y} = d - \Sigma y$$

SVD, continua

I passi computazionali da compiere sono

- calcolo $A = U\Sigma V^T$ → parte più onerosa
- calcolo di $d = U^T b = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$
- risoluzione di $\Sigma_1 y = d_1$
- calcolo di $\bar{x} = Vy$
- $||r(\bar{x})||_2 = ||d_2||_2$

Remark

Il calcolo di $\bar{x}$ richiede solo il vettore d_1, che può essere ottenuto con la matrice $U_1 \in \mathbb{R}^{m\times n}$ (economy size) formata dalle prime n colonne di U. Le matrici U_1, Σ_1 e V forniscono la decomposizione SVD ridotta (economy size) e la soluzione del problema è data da

$$\bar{x} = V\Sigma_1^{-1}U_1^T b.$$

Altri utilizzi della *SVD*

- **Norma matriciale euclidea**

la norma 2 è il valore singolare più grande

↳ tipicamente sono i valori della DIAGONALE PRINCIPALE

$$||A||_2 = \max_{x \neq 0} \frac{||Ax||_2}{||x||_2} = \sigma_{max}.$$

- **Numero di condizionamento in norma euclidea**

$$\mu_2(A) = ||A||_2||A^{-1}||_2 = \frac{\sigma_{max}}{\sigma_{min}}$$

Tale definizione può essere estesa anche a matrici rettangolari di qualsiasi forma. Se $rango(A) < \min(m, n)$ allora $\sigma_{min} = 0$ e $\mu_2(A) = \infty$.

- **Calcolo del rango**

Il rango di una matrice A è uguale al numero di valori singolari diversi da zero. In pratica il rango può non essere ben determinato quando alcuni valori singolari sono molto piccoli. Solitamente si fissa un valore di soglia e si parla di *rango numerico* della matrice.

- **Calcolo pseudoinversa**

$$A^\dagger = V\Sigma^\dagger U^T.$$

Altri utilizzi della *SVD* : basi ortonormali

- Le colonne di U corrispondenti ai valori singolari $\sigma_i \neq 0$ formano una base ortonormale di $span(A)$.
- Le colonne di U corrispondenti ai valori singolari $\sigma_i = 0$ formano una base ortonormale di $span(A)^{\perp}$.
- Le colonne di V corrispondenti ai valori singolari $\sigma_i = 0$ formano una base ortonormale di $null(A) = (x : Ax = 0)$.
- Le colonne di V corrispondenti ai valori singolari $\sigma_i \neq 0$ formano una base ortonormale di $null(A)^{\perp}$.

Altri utilizzi della *SVD* : approssimazione di matrici

Se conosciamo la DECOMPOSIZIONE a val. SING di A

La SVD della matrice $A \in \mathbb{R}^{m \times n}$ permette di esprimerla

$$A = \sum_{i=1}^{n} \sigma_i u^{(i)} v^{(i)T}$$

dove $u^{(i)} \in \mathbb{R}^m$ e $v^{(i)} \in \mathbb{R}^n$ sono le colonne della matrice U e V rispettivamente. Le matrici $u^{(i)} v^{(i)T}$ hanno rango uno.

Considerando $k < n$ valori singolari (ordinati in ordine crescente) otteniamo un'approssimazione $\tilde{A}$ della matrice A di rango k che verifica

$$||\tilde{A} - A||_2 = \sigma_{k+1}$$

$$||\tilde{A} - A||_2 = \min_{\{rango(B)=k\}} ||B - A||_2$$

Approssimazione di matrici: compressione e miglioramento di immagini digitali

→ matrici utilizzate per la RAPPRESENTAZIONE di un'IMMAGINE

Nel trattamento delle immagini digitali, utilizzando k valori singolari si può ottenere un'approssimazione molto buona della matrice originale A che rappresenta l'immagine.
La memorizzazione della matrice $\tilde{A}$ richiede $k(m+n+1)$ locazioni di memoria contro le mn richieste dalla matrice A con un notevole risparmio quando k molto più piccolo di m o n.

Oltre alla compressione di immagini, l'approssimazione di basso rango $\tilde{A}$ permette di migliorare l'immagine, eliminando i termini corrispondenti ai più piccoli valori singolari che possono essere responsabili di rumore o effetto sfocato.

→ es. nell'analisi del feto con ECG, elimina le immagini che danno "rumore" nello sfondo

Altri utilizzi della *SVD* : problema ai minimi quadrati completo

Nei problemi ai minimi quadrati standard, che nascono per esempio nel *fitting* di dati, si assume implicitamente che gli elementi di A siano noti esattamente, mentre gli elementi di b siano soggetti ad errori. Si cerca infatti di minimizzare $||y-b||_2$ con il vincolo $y \in span(A)$.

Quando sia A che b sono soggetti ad incertezza, si parla di **problema ai minimi quadrati completo** e si cerca il sistema compatibile, i.e. $y \in span(\bar{A})$, tale che $[\bar{A}\ y]$ sia il più vicino a $[A\ b]$.
Osserviamo che

- $[\bar{A}\ y]$ definisce un sistema compatibile sse il suo rango è al più n;
- una soluzione $\bar{x}$ verifica

$$[\bar{A}\ y][\begin{array}{c} x \\ -1 \end{array}] = 0$$

e quindi $[x^T, -1]^T \in null([\bar{A}\ y])$.

Problema ai minimi quadrati completo, continua

Calcolate U, Σ, V tali che

$$[A\ b] = U \Sigma V^T,$$

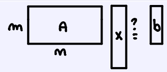
abbiamo che

- la matrice $[\bar{A}\ y]$ costruita usando usando i primi n valori singolari di $[A\ b]$ è un'approssimazione di rango n
- una soluzione $[\bar{x}^T, -1]$ risulta proporzionale a v_{n+1}

  Pertanto

$$\bar{x} = -\frac{1}{v_{n+1,n+1}} \begin{pmatrix} v_{n+1,1} \\ v_{n+1,2} \\ \cdots \\ v_{n+1,n} \end{pmatrix}$$

Soluzione di sistemi lineari sotto-determinati



ho TANTI GRADI DI LIBERTÀ
POCHI VINCOLI $\Rightarrow \infty$ soluz.

Tale sistema è caratterizzato da meno equazioni (m) che incognite (n) ed ha **infinite** soluzioni.

Ci proponiamo di trovare la soluzione $\bar{x}$ che ha norma due minima.

↳ $\min \|x\|_2$

Supponiamo che A abbia rango m. Vale

↳ m colonne LIN. INDIP.

Theorem

Ogni soluzione del sistema sotto-determinato si esprime

applico A → $Ax = \underbrace{AA^T(AA^T)^{-1}}_{I} b + Ay$

$$\underbrace{x = A^T(AA^T)^{-1}b}_{\text{questa è la soluz. di } \|\cdot\|_2 \text{ MINIMA}} + y \Rightarrow$$

→ $Ax = b + \underbrace{Ay}$

↳ se lo imponiamo $=0$, segue che x è sol.

dove $\bar{x} = A^T(AA^T)^{-1}b$ è una particolare soluzione del sistema e $Ay = 0$. Inoltre

$\Rightarrow$ ovviamente il calcolo dell'inversa è da EVITARE

$$||\bar{x}||_2 \leq ||x||_2, \quad \forall x : Ax = b$$

Oss: La matrice $A^T(AA^T)^{-1}$ è la **pseudoinversa** di A Inoltre si può dimostrare che $y = (I - A^T(AA^T)^{-1}A)z$, z qualsiasi.

Soluzione di sistemi lineari sotto-determinati, continua

Il calcolo della soluzione di norma minima $\bar{x}$ si riconduce quindi ai seguenti passi computazionali

devo calcolare $\bar{x} = A^T\underbrace{(AA^T)^{-1}b}_{z}$,

$z = (AA^T)^{-1}b$ sse $\underbrace{(AA^T)}_{M}z = b$

uso questo trucco quando c'è un INVERSA, riformulo il problema in modo da ottenere un sist. lineare

- calcolo di $M = AA^T$
- risoluzione di $Mz = b$
- calcolo di $\bar{x} = A^T z$

$M = A \cdot A^T = \boxed{A} \cdot \boxed{A^T} = \boxed{M}$, poi $\boxed{M} \cdot \boxed{z} = \boxed{b} \Rightarrow$ trovo z

infine $\boxed{A^T} \cdot \boxed{z} = \boxed{\bar{x}}$ e trovo $\bar{x}$

M è **simmetrica e definita positiva** quindi devo calcolare solo $\frac{m(m+1)}{2}$ elementi ed il sistema di equazioni di dimensione m può essere risolto con l'algoritmo di Cholesky.

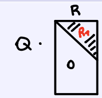
la metà degli elementi

Gli aspetti numericamente delicati sono legati al calcolo di M dove gli errori di arrotondamento possono far perdere alla matrice la proprietà di essere definita positiva. Esempio:

$$A = \begin{pmatrix} 1 & \epsilon & 0 \\ 1 & 0 & \epsilon \end{pmatrix} \rightsquigarrow AA^T = \begin{pmatrix} 1+\epsilon^2 & 1 \\ 1 & 1+\epsilon^2 \end{pmatrix}$$

che risulta singolare se $\epsilon^2 < u$.

Soluzione di sistemi lineari sotto-determinati, continua



e $R^T = [R_1^T\ 0]$

Un approccio alternativo è il seguente. Data $A^T = QR$ allora

↳ $A = R^T Q^T$

$$Ax = b \rightsquigarrow R^T \underbrace{Q^T x}_{y} = b$$

Sia

$$y = Q^T x = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \begin{array}{l} \leftarrow \mathbb{R}^m \\ \leftarrow \mathbb{R}^{n-m} \end{array},$$

allora

$R^T \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = [R_1^T y_1, 0]$ e deve essere $= b$

$$R^T y = (R_1^T, 0^T) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = b \rightsquigarrow R_1^T y_1 = b.$$

↳ questo è un SISTEMA TRIANGOLARE, trovo y_1

La soluzione di norma minima si ottiene ponendo $y_2 = 0$ e calcolando

$$\bar{x} = Qy = (Q_1, Q_2) \begin{pmatrix} y_1 \\ 0 \end{pmatrix} \rightsquigarrow \bar{x} = Q_1 y_1.$$

Soluzione di sistemi lineari sotto-determinati, continua

Il calcolo della soluzione di minima norma $\bar{x}$ si riconduce quindi ai seguenti passi computazionali

- calcolo di $A^T = QR = (Q_1, Q_2)\begin{pmatrix} R_1 \\ 0 \end{pmatrix}$
- risoluzione di $R_1^T y_1 = b$
- calcolo di $\bar{x} = Q_1 y_1$

Il costo di tale algoritmo è $m^2 n - m^3/3$ flop

Un'applicazione dei sistemi lineari sotto-determinati

Un' importante applicazione nasce nell'ambito dell'ottimizzazione con vincoli. Supponiamo di dover risolvere

funz. di 2 var.

$$\min_{x_1+x_2=1} F(x_1, x_2).$$

Il problema può essere semplificato ponendo $x_2 = 1 - x_1$ e risolvendo il problema

funz. di 1 var.

$$\min_{x_1} F(x_1, 1 - x_1)$$

che non ha più vincoli ed ha un'unica variabile x_1.
Con la notazione matriciale, il vincolo si scrive come $Ax = b$ con $A = (1, 1)$ e $b = 1$ e la generica soluzione del sistema sotto-determinato $Ax = b$ come

$$x = (0, 1)^T + p(1, -1)^T, \ \ p \in \mathbb{R},$$

dove $\bar{x} = (0, 1)^T$ è una particolare soluzione $Ax = b$ e $y = p(1, -1)^T, p \in \mathbb{R}$ risolve $Ay = 0$. Il problema iniziale viene riformulato

$$\min_{p} F(p, 1 - p).$$

È importante trovare le soluzioni y di $Ay = 0$.

Applicazione, continua

In generale un problema di dimensione n si scrive

$$\min_{Ax=b} F(x)$$

con $A \in \mathbb{R}^{m\times n}$, $m < n$. Data $A^T = QR$ risulta

$$AQ = (AQ_1,\ AQ_2) = R^T = (R_1^T,\ 0)$$

da cui

$$AQ_2 = 0.$$

Data una soluzione particolare $\bar{x}$ di $Ax = b$ il problema di ottimizzazione iniziale si riscrive

$$\min_p F(\bar{x} + Q_2 p),$$

con $p \in \mathbb{R}^{n-m}$. La nuova riformulazione ha quindi $n - m$ variabili e non ha più vincoli.

Sistemi lineari sotto-determinati, esercizio

Su un corpo di massa $m = 1$, inizialmente in posizione di riposo, agiscono nella stessa direzione ed in sequenza delle forze f_i per $(i-1)s \leq t \leq is, i = 1, 2, 3, 4$ in modo da portarlo a fermarsi dopo $4s$ alla distanza di un metro dal punto iniziale. Scrivi il sistema sotto-determinato da risolvere per determinare le forze f_i, utilizzando le leggi del moto uniformemente accelerato. Calcola la soluzione (f_1, f_2, f_3, f_4) di norma minima.