



**UNIVERSITÀ DEGLI STUDI DI CAGLIARI**  
**Facoltà di Ingegneria e Architettura**

Corso di laurea triennale in  
Ingegneria Elettrica ed Elettronica

**Problema ai minimi quadrati  
nella risoluzione di sistemi lineari  
indeterminati**

Relazione per il Seminario Avanzato di Matematica Applicata

**Studente**

Antonio Marongiu  
matr. n° 70/87/46164

**Docente**

Prof. G. Rodriguez

## Sistemi lineari $m \times n$ , con $m \neq n$ .

Un sistema lineare è un insieme finito di simultanee equazioni lineari aventi le stesse incognite  $x_1, \dots, x_n$

$$Ax = b$$

Una soluzione del sistema lineare è una  $n$ -upla  $(c_1, \dots, c_n) \in \mathbb{R}^n$  che sia soluzione di tutte le equazioni contemporaneamente. Un sistema lineare si dice risolubile (o compatibile) se ha almeno una soluzione.

Il sistema è costituito da  $m$  equazioni in  $n$  incognite ( $m, n$  qualsiasi). Il rango della matrice rettangolare  $A$  non può superare il più piccolo tra i numeri  $m$  ed  $n$ .

➤  $m > n$

Si può estrarre dalla matrice rettangolare  $A$  una sottomatrice quadrata di ordine  $n$ ; sia  $k$  il rango della matrice  $A$ .

$$\begin{cases} k(A) \neq k(A, \mathbf{b}) & \text{ sistema incompatibile } k(A, \mathbf{b}) > n \\ k(A) = k(A, \mathbf{b}) & \text{ sistema compatibile } \begin{cases} k(A) = n & \text{ soluzione unica} \\ k(A) < n & \infty^{n-k(A)} \text{ soluzioni} \end{cases} \end{cases}$$

Il problema potrebbe non avere soluzioni. Il sistema è detto sovradeterminato.

➤  $m < n$

Si può estrarre dalla matrice rettangolare  $A$  una sottomatrice quadrata di ordine  $m$ ; sia  $k$  il rango della matrice  $A$ . Poiché  $k \leq m < n$ , il sistema non potrà mai avere una ed una sola soluzione.

$$\begin{cases} k(A) \neq k(A, \mathbf{b}) & \text{ sistema incompatibile} \\ k(A) = k(A, \mathbf{b}) & \text{ sistema compatibile } \Rightarrow k(A) < n \quad \infty^{n-k(A)} \text{ soluzioni} \end{cases}$$

Il problema potrebbe avere infinite soluzioni. Il sistema è detto sottodeterminato.

Questi casi son esempi di problemi mal posti (soluzione inesistente o non unica). In particolare sia fatta attenzione sul primo caso, frequente nelle applicazioni: il numero di equazioni è ridondante e superiore al numero delle incognite.

Idealmente il numero di equazioni linearmente indipendenti è  $m - n$ , ma potrebbe non essere noto quali. Nel caso di dati affetti da errore è conveniente risolvere un numero sovrabbondante di equazioni per effettuare un processo di *noise reduction*.

Nel caso specifico di problema mal posto in un sistema sovradeterminato è ragionevole richiedere che sia minima la lunghezza euclidea del vettore residuo (detta anche varianza). Il *least squares problem*, o problema della regressione polinomiale, si esplica nella condizione di varianza minima

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2$$

Se la varianza minima è nulla, il sistema originale ammette una soluzione in senso classico (unica e dipendente con continuità dai dati). In caso contrario si ottiene la soluzione nel senso dei minimi quadrati.

### Metodo delle equazioni normali.

Data  $A \in \mathbb{R}^{m \times n}$  con  $m \geq n$ ,  $\mathbf{b} \in \mathbb{R}^m$ , si dice che  $\tilde{\mathbf{x}} \in \mathbb{R}^n$ , è soluzione nel senso dei minimi quadrati del sistema lineare  $A\mathbf{x} = \mathbf{b}$  se

$$\Phi(\tilde{\mathbf{x}}) = \|A\tilde{\mathbf{x}} - \mathbf{b}\|_2^2 \leq \min_{\mathbf{x} \in \mathbb{R}^n} \|A\mathbf{x} - \mathbf{b}\|_2^2 = \min_{\mathbf{x} \in \mathbb{R}^n} \Phi(\mathbf{x})$$

Si tratta dunque di minimizzare la norma euclidea del residuo. La soluzione può essere ottenuta imponendo che il gradiente<sup>1</sup> della funzione  $\Phi$  si annulli in  $\tilde{\mathbf{x}}$ . Essendo

$$\begin{aligned} \Phi(\mathbf{x}) &= (A\mathbf{x} - \mathbf{b})^T (A\mathbf{x} - \mathbf{b}) = \mathbf{x}^T A^T A \mathbf{x} - \mathbf{x}^T A^T \mathbf{b} - \mathbf{b}^T A \mathbf{x} + \mathbf{b}^T \mathbf{b} \\ &= \mathbf{x}^T A^T A \mathbf{x} - 2\mathbf{x}^T A^T \mathbf{b} + \mathbf{b}^T \mathbf{b} \end{aligned}$$

Si trova dunque

$$\nabla \Phi(\tilde{\mathbf{x}}) = 2A^T A \tilde{\mathbf{x}} - 2A^T \mathbf{b} = 0$$

Da cui si deduce che  $\tilde{\mathbf{x}}$  è la soluzione del sistema  $n \times n$

$$A^T A \tilde{\mathbf{x}} = A^T \mathbf{b}$$

Nota come *sistema normale* di equazioni. Se  $A$  è a rango pieno, tale sistema è non singolare. In tal caso, la soluzione nel senso dei minimi quadrati esiste ed è unica

$$\tilde{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}$$

Se  $A$  ha rango inferiore a  $n$ , la matrice  $A^T A$  risulta singolare, ma il sistema resta consistente in quanto il vettore  $A^T \mathbf{b}$  appartiene all'immagine di  $A^T$ , che coincide con l'immagine di  $A^T A$ . In questo caso, tra gli infiniti vettori soluzione, usualmente si assume come soluzione  $\tilde{\mathbf{x}}$  quella dotata di minima norma euclidea (soluzione normale). Per risolvere il sistema, nel caso a rango pieno, essendo  $A^T A$  simmetrica e definita positiva, si potrebbe usare la sua fattorizzazione di *Cholesky*:  $A^T A = R^T R$ , con  $R$  matrice  $n \times n$  triangolare superiore, il vettore  $\tilde{\mathbf{x}}$  può essere calcolato risolvendo in successione i due sistemi triangolari

$$\begin{cases} R^T \mathbf{y} = A^T \mathbf{b} \\ R \mathbf{x} = \mathbf{y} \end{cases}$$

Tuttavia, il sistema è solitamente mal condizionato; inoltre, a causa degli errori di arrotondamento, nel calcolo di  $A^T A$  possono andare perdute cifre significative con conseguente perdita della definita positività o della non singolarità della matrice. È quindi in generale più conveniente impiegare la fattorizzazione  $QR$ .

---

<sup>1</sup> Si riporta in breve la definizione di gradiente di una funzione. Si definisce **gradiente** un operatore del primo ordine che associa ad ogni campo scalare definito e differenziabile in  $\Omega$  il campo vettoriale le cui componenti sono le derivate parziali prime:

$$\nabla \varphi = \left( \frac{\partial \varphi}{\partial x_j} \right)_{1 \leq j \leq n}$$

### La decomposizione ai valori singolari (SVD) e la matrice pseudo-inversa.

Una qualunque matrice può essere ridotta in forma diagonale tramite pre e post-moltiplicazione per matrici unitarie.

Sia  $A \in \mathbb{R}^{m \times n}$ . Esistono due matrici unitarie  $U \in \mathbb{R}^{m \times m}$  e  $V \in \mathbb{R}^{n \times n}$  tali che

$$U^T A V = \Sigma = \text{diag}(\sigma_1, \dots, \sigma_p) \in \mathbb{R}^{m \times n}$$

Con  $p = \min(m, n)$  e  $\sigma_1 \geq \dots \geq \sigma_p \geq 0$ . Tale espressione è detta decomposizione in valori singolari (SVD) di  $A$  e i  $\sigma_i(A) = \sqrt{\lambda_i(A^*A)}$ ,  $i = 1, \dots, n$  sono detti valori singolari di  $A$ .

Supponiamo che il rango di  $A \in \mathbb{R}^{m \times n}$  sia pari a  $r$  e che  $A$  ammetta una SVD del tipo  $U^* A V = \Sigma$ .

La matrice  $A^\dagger = V \Sigma^\dagger U^*$  è detta pseudo-inversa di Moore-Penrose o inversa generalizzata, solo se  $m \geq n$  e  $A$  è a rango pieno; essendo

$$\Sigma^\dagger = \text{diag}\left(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}, 0, \dots, 0\right)$$

### Risoluzione QR.

Si dice che una matrice  $A \in \mathbb{R}^{m \times n}$ , con  $m \geq n$ , ammette una fattorizzazione  $QR$  se esistono una matrice  $Q \in \mathbb{R}^{m \times m}$  ortogonale ed una matrice triangolare superiore  $R \in \mathbb{R}^{m \times n}$  tali che  $A = QR$ .

La fattorizzazione  $QR$  esiste ed è unica, avendo  $A$  rango pieno.

$$\|Ax - b\|^2 = \|QRx - b\|^2 = \|Q(Rx - Q^T b)\|^2 = \|Rx - c\|^2$$

Avendo posto  $c = Q^T b$  e ricordando che una matrice ortogonale ha norma-2 unitaria. La matrice  $R$  ha la forma

$$R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$$

Con  $R_1$  triangolare superiore, quadrata e, se  $A$  è a rango pieno, non singolare. Partizionando in maniera coerente anche  $c$

$$c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, c_1 \in \mathbb{R}^n, c_2 \in \mathbb{R}^{m-n}$$

Si ha

$$\|Ax - b\|^2 = \|Rx - c\|^2 = \left\| \begin{bmatrix} R_1 x \\ 0 \end{bmatrix} - \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \right\|^2 = \|R_1 x - c_1\|^2 + \|c_2\|^2$$

Se  $\det(R_1) \neq 0$ , il sistema

$$R_1 x = c_1$$

Ammette una e una sola soluzione, per la quale si ha  $\|R_1 x - c_1\|^2 = 0$ . In corrispondenza a tale soluzione

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2 = \|c_2\|$$

Perciò l'eventuale annullarsi del vettore  $\mathbf{c}_2$  significa che  $\mathbf{x}$  è la soluzione classica del sistema iniziale  $A\mathbf{x} = \mathbf{b}$ . In caso contrario, essa è la soluzione nel senso dei minimi quadrati, e la norma di  $\mathbf{c}_2$  fornisce la misura della varianza minima. Rispetto all'approccio basato sulle equazioni normali questo metodo ha il vantaggio di operare direttamente sulla matrice  $A$ , il cui condizionamento è pari alla radice quadrata di quello di  $A^T A$ .

Sia  $A \in \mathbb{R}^{m \times n}$ , con  $m \geq n$  una matrice di rango pieno. Allora l'unica soluzione nel senso dei minimi quadrati è data da

$$\tilde{\mathbf{x}} = \tilde{R}^{-1} \tilde{Q}^T \mathbf{b}$$

Essendo  $\tilde{R} \in \mathbb{R}^{n \times n}$  e  $\tilde{Q} \in \mathbb{R}^{m \times n}$  le matrici definite dalla fattorizzazione  $QR$ . Inoltre il minimo di  $\Phi$  vale

$$\Phi(\tilde{\mathbf{x}}) = \sum_{i=n+1}^m [(Q^T \mathbf{b})_i]^2$$

Infatti

$$\|R\mathbf{x} - Q^T \mathbf{b}\|_2^2 = \|\tilde{R}\mathbf{x} - \tilde{Q}^T \mathbf{b}\|_2^2 + \sum_{i=n+1}^m [(Q^T \mathbf{b})_i]^2$$

Ed il minimo è quindi raggiunto per

$$\mathbf{x} = \tilde{\mathbf{x}} = \tilde{R}^{-1} \tilde{Q}^T \mathbf{b}$$

Se  $A$  non ha rango pieno tali tecniche di risoluzione si dimostrano inappropriate: se  $\tilde{\mathbf{x}}$  è una soluzione nel senso dei minimi quadrati, anche  $\tilde{\mathbf{x}} + \mathbf{z}$  con  $\mathbf{z} \in \ker(A)$  è soluzione. Si deve perciò introdurre un ulteriore vincolo per fissare un'unica soluzione. Ciò viene ottenuto richiedendo che  $\tilde{\mathbf{x}}$  abbia norma euclidea minima, di modo che il problema precedente diventa

Trovare  $\tilde{\mathbf{x}}$  tale che  $\|\tilde{\mathbf{x}}\|_2$  sia minima e

$$\|A\tilde{\mathbf{x}} - \mathbf{b}\|_2^2 \leq \min_{\mathbf{x} \in \mathbb{R}^n} \|A\mathbf{x} - \mathbf{b}\|_2^2$$

Questo problema è consistente con il caso precedente qualora  $A$  abbia rango pieno, in quanto in tal caso  $\min_{\mathbf{x} \in \mathbb{R}^n} \Phi(\mathbf{x})$  ammette un'unica soluzione che avrà necessariamente norma euclidea minima.

Lo strumento che consente di calcolare la soluzione di

$$\|A\tilde{\mathbf{x}} - \mathbf{b}\|_2^2 \leq \min_{\mathbf{x} \in \mathbb{R}^n} \|A\mathbf{x} - \mathbf{b}\|_2^2$$

è la decomposizione in valori singolari (SVD). Vale infatti il seguente risultato

Sia  $A \in \mathbb{R}^{m \times n}$ . Allora l'unica soluzione di  $\|A\tilde{\mathbf{x}} - \mathbf{b}\|_2^2 \leq \min_{\mathbf{x} \in \mathbb{R}^n} \|A\mathbf{x} - \mathbf{b}\|_2^2$  è

$$\tilde{\mathbf{x}} = A^\dagger \mathbf{b}$$

Essendo  $A^\dagger$  la pseudo-inversa di  $A$ .

Utilizzando la SVD di  $A$ ,  $A = U\Sigma V^T$ , il problema ai minimi quadrati è equivalente a cercare  $\mathbf{w} = V^T \mathbf{x}$  tale che  $\mathbf{w}$  abbia norma euclidea minima e

$$\|\Sigma \mathbf{w} - U^T \mathbf{b}\|_2^2 \leq \|\Sigma \mathbf{y} - U^T \mathbf{b}\|_2^2, \quad \forall \mathbf{y} \in \mathbb{R}^n$$

Se  $r$  è il numero di valori singolari  $\sigma_i$  non nulli di  $A$ , allora

$$\|\Sigma \mathbf{w} - U^T \mathbf{b}\|_2^2 = \sum_{i=1}^r (\sigma_i \omega_i - (U^T \mathbf{b})_i)^2 + \sum_{i=r+1}^m ((U^T \mathbf{b})_i)^2$$

Che è minimo se  $\omega_i = \frac{(U^T \mathbf{b})_i}{\sigma_i}$  per  $i = 1, \dots, r$ . Inoltre è evidente che fra tutti i vettori  $\mathbf{w}$  di  $\mathbb{R}^n$  che hanno le prime  $r$  componenti fissate, quello che rende minima la norma euclidea ha le restanti  $n - r$  componenti nulle. Dunque, il vettore soluzione da cercare è  $\tilde{\mathbf{w}} = \Sigma^\dagger U^T \mathbf{b}$  ossia  $\tilde{\mathbf{x}} = V \Sigma^\dagger U^T \mathbf{b} = A^\dagger \mathbf{b}$ .

Per quanto riguarda la stabilità del problema, ci limitiamo a fare notare che se il rango di  $A$  non è pieno, la soluzione  $\tilde{\mathbf{x}}$  non è necessariamente una funzione continua dei dati e piccoli cambiamenti su di essi possono produrne di grandi su  $\tilde{\mathbf{x}}$ .

Nel caso di sistemi sotto determinati, in cui  $m < n$ , se  $A$  ha rango massimo si può ancora utilizzare la fattorizzazione  $QR$ . In particolare, se si opera sulla matrice traspota si ottiene la soluzione del sistema con norma euclidea minima. Se invece la matrice non ha rango massimo, si ricorre nuovamente alla  $SVD$ . Nel caso in cui  $m = n$ , si possono evidentemente ancora usare la  $SVD$  o la fattorizzazione  $QR$  in alternativa al metodo di Gauss per la risoluzione del sistema lineare  $A\mathbf{x} = \mathbf{b}$ . Anche se questi algoritmi hanno un costo di gran lunga superiore a quello richiesto dal metodo di Gauss, essi risultano preferibili ad esso quanto il sistema sia mal condizionato e prossimo ad essere singolare.

### Condizionamento di un problema dei minimi quadrati.

Sia  $A \in \mathbb{R}^{m \times n}$ , con rango  $n < m$ . Si può assumere come numero di condizionamento di  $A$  il numero

$$k_2(A) = \frac{\sigma_1}{\sigma_n}$$

Sia  $\mathbf{x}$  la soluzione del problema

$$\min_{\mathbf{y}} \|A\mathbf{y} - \mathbf{b}\|_2$$

E sia  $\mathbf{r} = A\mathbf{x} - \mathbf{b}$  il residuo. Poniamo  $\sin \vartheta = \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}$ . Sia  $\mathbf{x} + \Delta\mathbf{x}$  la soluzione di

$$\min_{\mathbf{y}} \|(A + \delta A)\mathbf{y} - \mathbf{b}\|_2$$

e supponiamo che  $\varepsilon = \frac{\|\delta A\|}{\|A\|}$  sia piccolo. Allora

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \varepsilon [\tan \vartheta k_2^2(A) + k_2(A)] + O(\varepsilon^2)$$

$$\frac{\|\Delta\mathbf{r}\|}{\|\mathbf{b}\|} \leq 2\varepsilon k_2(A) + O(\varepsilon^2)$$

### Matrici di Householder.

Le matrici di *Householder* sono matrici elementari del tipo

$$H(\mathbf{w}, \mathbf{w}) = I - 2\mathbf{w}\mathbf{w}^T$$

ove  $\mathbf{w} \in \mathbb{R}^n$  è un vettore di lunghezza unitaria:  $\|\mathbf{w}\|_2 = 1$ . Si tratta di matrici ortogonali, simmetriche e involutorie. Vengono anche indicate come matrici di riflessione o riflettori elementari, in quanto rappresentano la simmetria rispetto all'iperpiano passante per l'origine e ortogonale al vettore  $\mathbf{w}$ .

Se si pre-moltiplica una matrice  $A = (\mathbf{a}_1, \dots, \mathbf{a}_n)$  per  $H$ , si ottiene una matrice  $A'$  le cui colonne sono trasformate nel seguente modo

$$\mathbf{a}'_k = (I - 2\mathbf{w}\mathbf{w}^T)\mathbf{a}_k = \mathbf{a}_k - 2(\mathbf{w}^T \mathbf{a}_k)\mathbf{w}$$

Allo stesso modo nella post-moltiplicazione sono le righe ad essere trasformate indipendentemente.

Dal punto di vista numerico le matrici di *Householder* sono importanti per la seguente proprietà.

Siano  $\mathbf{x}$  e  $\mathbf{y}$  due vettori qualunque di  $\mathbb{R}^n$ , linearmente indipendenti con  $\|\mathbf{y}\|_2 = 1$ . Si può allora determinare un vettore  $\mathbf{w} \in \mathbb{R}^n$ , con  $\|\mathbf{w}\|_2 = 1$ , e uno scalare  $\alpha$  tali che se  $H = I - 2\mathbf{w}\mathbf{w}^T$ , allora  $H\mathbf{x} = \alpha\mathbf{y}$ .

Dimostrazione.

Poiché  $H$  è ortogonale, si ha

$$\|\mathbf{x}\|_2 = \|H\mathbf{x}\|_2 = \|\alpha\mathbf{y}\|_2 = |\alpha|$$

Da cui si hanno due possibili scelte per  $\alpha$ :  $\|\mathbf{x}\|$  o  $-\|\mathbf{x}\|$ . L'uguaglianza  $H\mathbf{x} = \alpha\mathbf{y}$  si scrive

$$\mathbf{x} - 2(\mathbf{w}^T \mathbf{x})\mathbf{w} = \alpha\mathbf{y}$$

Oppure

$$\mathbf{x} - \alpha\mathbf{y} = 2\lambda\mathbf{w}$$

con  $\lambda = \mathbf{w}^T \mathbf{x}$ .

Per calcolare  $\lambda$  si fa il prodotto scalare dell'uguaglianza precedente per  $\mathbf{x}$ :

$$\mathbf{x}^T \mathbf{x} - \alpha \mathbf{x}^T \mathbf{y} = 2\lambda \mathbf{x}^T \mathbf{w} = 2\lambda^2$$

Poiché i vettori  $\mathbf{x}, \mathbf{y}$  non sono linearmente dipendenti si ha:

$$|\mathbf{x}^T \mathbf{y}| < \|\mathbf{x}\| \|\mathbf{y}\| = |\alpha|$$

E l'uguaglianza precedente permette di determinare  $\lambda \neq 0$ , poiché  $\mathbf{x}^T \mathbf{x} - \alpha \mathbf{x}^T \mathbf{y} = \alpha^2 - \alpha \mathbf{x}^T \mathbf{y} > 0$ . Il vettore  $\mathbf{w}$  è allora definito da  $\mathbf{w} = \frac{\mathbf{x} - \alpha\mathbf{y}}{2\lambda}$ . La matrice  $H$  non dipende dal segno di  $\lambda$ .

Nelle applicazioni si ha interesse a porre  $H$  sotto la forma  $I - \frac{1}{\beta} \mathbf{v}\mathbf{v}^T$ , con

$\mathbf{v} = 2\lambda\mathbf{w} = \mathbf{x} - \alpha\mathbf{y}$  e  $\beta = 2\lambda^2 = \alpha^2 - \alpha(\mathbf{x}^T \mathbf{y})$ , che evita di dover estrarre una radice quadrata per calcolare  $\lambda$ . In generale non si ha bisogno di calcolare  $H$  esplicitamente, ma solamente di poter calcolare il trasformato di un generico vettore  $\mathbf{z}$ , cioè:

$$H\mathbf{z} = \mathbf{z} - \frac{\mathbf{v}\mathbf{v}^T \mathbf{z}}{\beta}$$

Tale risultato è ottenuto calcolando dapprima  $\gamma = \frac{\mathbf{v}^T \mathbf{z}}{\beta}$  e poi  $H\mathbf{z} = \mathbf{z} - \gamma\mathbf{v}$ .

### Metodo di Householder.

Sia  $A$  una matrice quadrata scritta in termini di vettori colonna

$$A = A^{(1)} = [\mathbf{a}_1^{(1)} \dots \mathbf{a}_n^{(1)}]$$

È possibile costruire la matrice elementare di *Householder*  $H_1$  tale che

$$H_1 \mathbf{a}_1^{(1)} = k_1 \mathbf{e}_1$$

Si genera in tal modo la prossima matrice della successione

$$A^{(2)} = H_1 A^{(1)} = \begin{bmatrix} \mathbf{a}_1^{(2)} & \dots & \mathbf{a}_n^{(2)} \end{bmatrix} = \begin{bmatrix} k_1 & \mathbf{v}_1^T \\ \mathbf{0} & \hat{A}^{(2)} \end{bmatrix}$$

Successivamente, si consideri la sottomatrice  $n - 1$

$$\hat{A}^{(2)} = [\hat{\mathbf{a}}_2^{(2)} \dots \hat{\mathbf{a}}_n^{(2)}]$$

E si costruisca il riflettore elementare di dimensione  $n - 1$  tale che

$$\hat{H}_2 \hat{\mathbf{a}}_2^{(2)} = k_2 \mathbf{e}_1$$

A tal punto, si orli la matrice  $\hat{H}_2$  con una riga e una colonna della matrice identità al fine di raggiungere la dimensione  $n$

$$H_2 = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \hat{H}_2 \end{bmatrix}$$

E si moltiplichi a sinistra per  $A^{(2)}$

$$A^{(3)} = H_2 A^{(2)} = \begin{bmatrix} k_1 & \mathbf{v}_1^T \\ \mathbf{0} & \widehat{H}_2 A^{(2)} \end{bmatrix}$$

E così via. Se  $A$  è quadrata, l'algoritmo termina al passo  $n - 1$  con la matrice triangolare superiore

$$R = A^{(n)} = H_{n-1} A^{(n-1)} = H_{n-1} \dots H_1 A^{(1)} = Q^T A$$

La matrice ortogonale

$$Q = H_1 \dots H_{n-1}$$

essendo prodotto di matrici ortogonali implica

$$A = QR$$

che costituisce la fattorizzazione  $QR$  di  $A$  secondo *A. Householder*.

## Esperimento su Matlab.

In questo esperimento su Matlab è stato implementato un sistema sovra-determinato di equazioni lineari ( $m > n$ ); più precisamente con  $m = 200$  ed  $n \in [1,87]$  attraverso un ciclo `for`.

```
A=rand(m,n);  
x=rand(n,1);  
b=A*x;
```

La prima parte effettua la risoluzione del sistema in senso classico di cui sono stati calcolati il numero di condizionamento, l'errore assoluto (ovviamente nullo) e la condizione di varianza minima.

```
N=norm(A*x-b);  
M=min(N);  
C(n)=cond(A)  
E(n)=norm(x-x)
```

Successivamente è stata introdotta una “piccola” perturbazione sui dati ( $1 \cdot 10^{-3}$ )

```
b1=b+(1e-03)*rand(m,1);
```

ed è stato risolto il sistema mediante la fattorizzazione di Cholesky e QR.

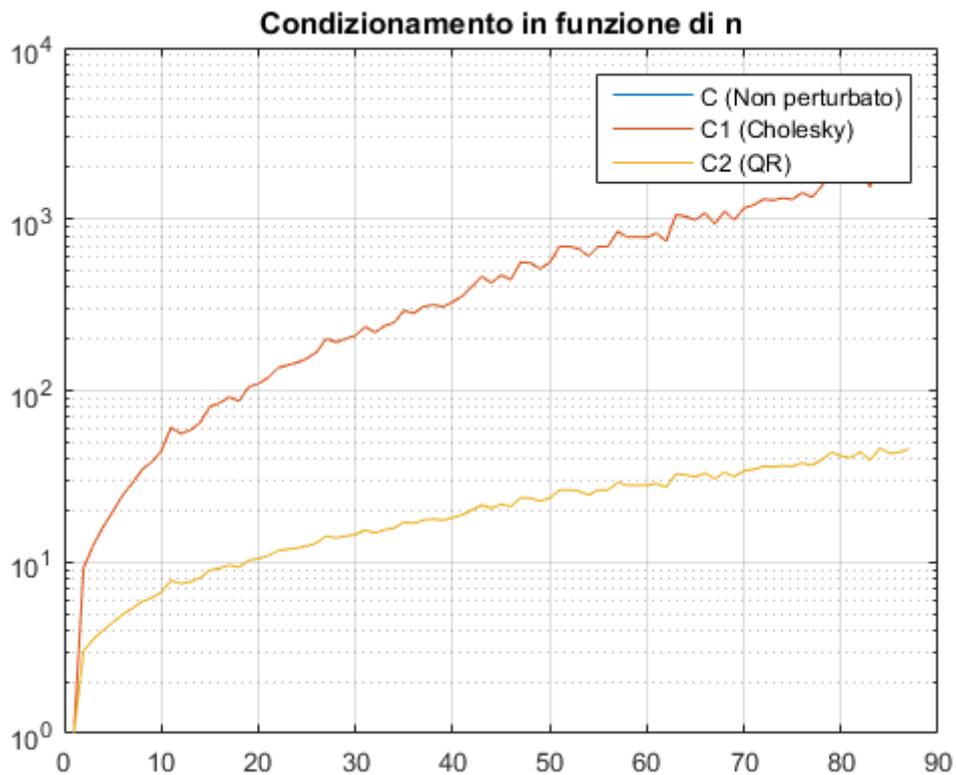
```
%% Cholesky factorization  
R1=chol(A'*A);  
y=R1'\(A'*b1);  
x1=R1\y;  
N1=norm(A*x-b1);  
M1=min(N1);  
C1(n)=cond(A'*A)  
E1(n)=norm(x-x1)  
  
%% QR factorization  
[Q,R]=qr(A);  
t=Q'*b1;  
x2=R\t;  
N2=norm(Q*R*x-b1);  
M2=min(N2);  
C2(n)=C(n)+tand(asin(N/norm(b1,2)))*(C(n))^2  
E2(n)=norm(x-x2)
```

Anche in questo caso sono stati calcolati: l'errore assoluto, il numero di condizionamento e la condizione di varianza minima; in particolar modo, nella fattorizzazione QR, il numero di condizionamento è stato determinato tramite la formula ai minimi quadrati.

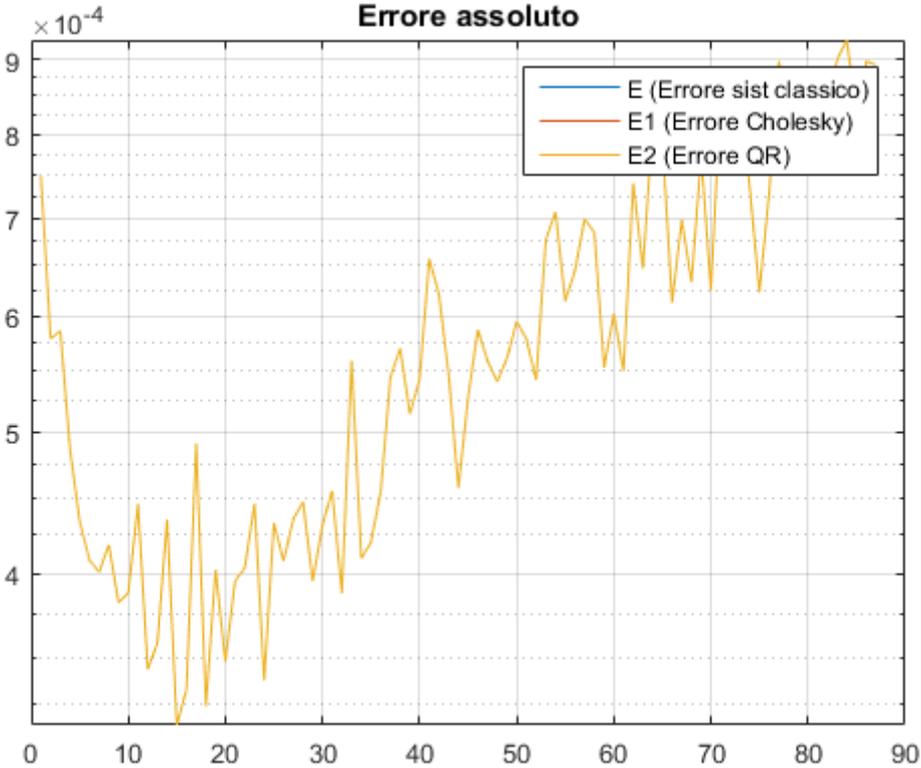
Infine, è stata implementata una matrice mal-condizionata, del tipo di Hilbert, e si è notato come sia sufficiente una “piccola” dimensione affinché il sistema risultasse irrisolvibile.

```
H=hilb(n);  
x3=rand(n,1)  
h=H*x3;  
  
N3=norm(H*x-h);  
M3=min(N3);  
C3=cond(H)  
E3=abs(x-x3)
```

Al termine sono stati eseguiti due grafici in scala semilogaritmica: nel primo grafico, relativo ai condizionamenti del sistema, si nota che due di loro ( $C$  e  $C2$ ) coincidono per il fatto che il sistema è compatibile, e quindi il termine noto appartiene all'immagine di  $A$ .



Nel secondo grafico sono stati rappresentati gli errori assoluti del sistema e, anche in questo caso, due di essi coincidono (*E1 ed E2*).



Di seguito è riportato l'intero codice implementato:

```

clc; clear all;
m=200; nmax=87;
C= zeros (nmax,1);
C1= zeros (nmax,1);
C2= zeros (nmax,1);
C3=zeros (nmax,1);
E= zeros (nmax,1);
E1= zeros (nmax,1);
E2= zeros (nmax,1);
%% Metodi di risoluzione
for n=1:nmax;
    %% Caso di soluzione in senso classico
    %%Si introducano A e x di un sistema sovradeterminato%
    %%Si consideri il caso di soluzione in senso classico%
    A=rand(m,n);
    x=rand(n,1);
    b=A*x; %b è il vettore dei termini noti%
    %M rappresenta la condizione di varianza N minima%
    N=norm(A*x-b);
    M=min(N);
    %C indica il condizionamento di A%
    C(n)=cond(A)
    %E indica l'errore assoluto%
    E(n)=norm(x-x)
    %% Caso "lievemente" perturbato
    %%Sia introdotta una "piccola" perturbazione sui dati%
    b1=b+(1e-03)*rand(m,1);
    %% Cholesky factorization
    R1=chol(A'*A);
    y=R1'\(A'*b1);
    x1=R1\y;
    N1=norm(A*x-b1);
    M1=min(N1);
    C1(n)=cond(A'*A)
    E1(n)=norm(x-x1)
    %% QR factorization
    [Q,R]=qr(A);
    t=Q'*b1;
    x2=R\t;
    N2=norm(Q*R*x-b1);
    M2=min(N2);
    C2(n)=C(n)+tand(asin(N/norm(b1,2)))*(C(n))^2
    E2(n)=norm(x-x2)
end
%% Si introduca una matrice di Hilbert.
%%Si noti come sia sufficiente una "piccola" dimensione %
% affinché il sistema sia irrisolvibile.%
H=hilb(n);
x3=rand(n,1);
h=H*x3; %h è il vettore dei termini noti%
N3=norm(H*x-h);
M3=min(N3);
C3=cond(H)
E3=norm(x-x3)
%% Grafici
figure(1)
semilogy([C C1 C2]);
grid;
title('Condizionamento in funzione di n');
legend('C (Non perturbato)', 'C1 (Cholesky)', 'C2 (QR)');
figure(2)

```

```
semilogy([E E1 E2]);  
grid;  
title('Errore assoluto');  
legend('E (Errore sist classico)', 'E1 (Errore Cholesky)', 'E2 (Errore QR)');
```