

Condizionamento del problema

$$\begin{aligned}x_1 + 2x_2 &= 3 \\ .499x_1 + 1.001x_2 &= 1.5\end{aligned}$$

La soluzione esatta è $x = (1, 1)^T$.

Perturbando la matrice dei coefficienti o il termine noto:

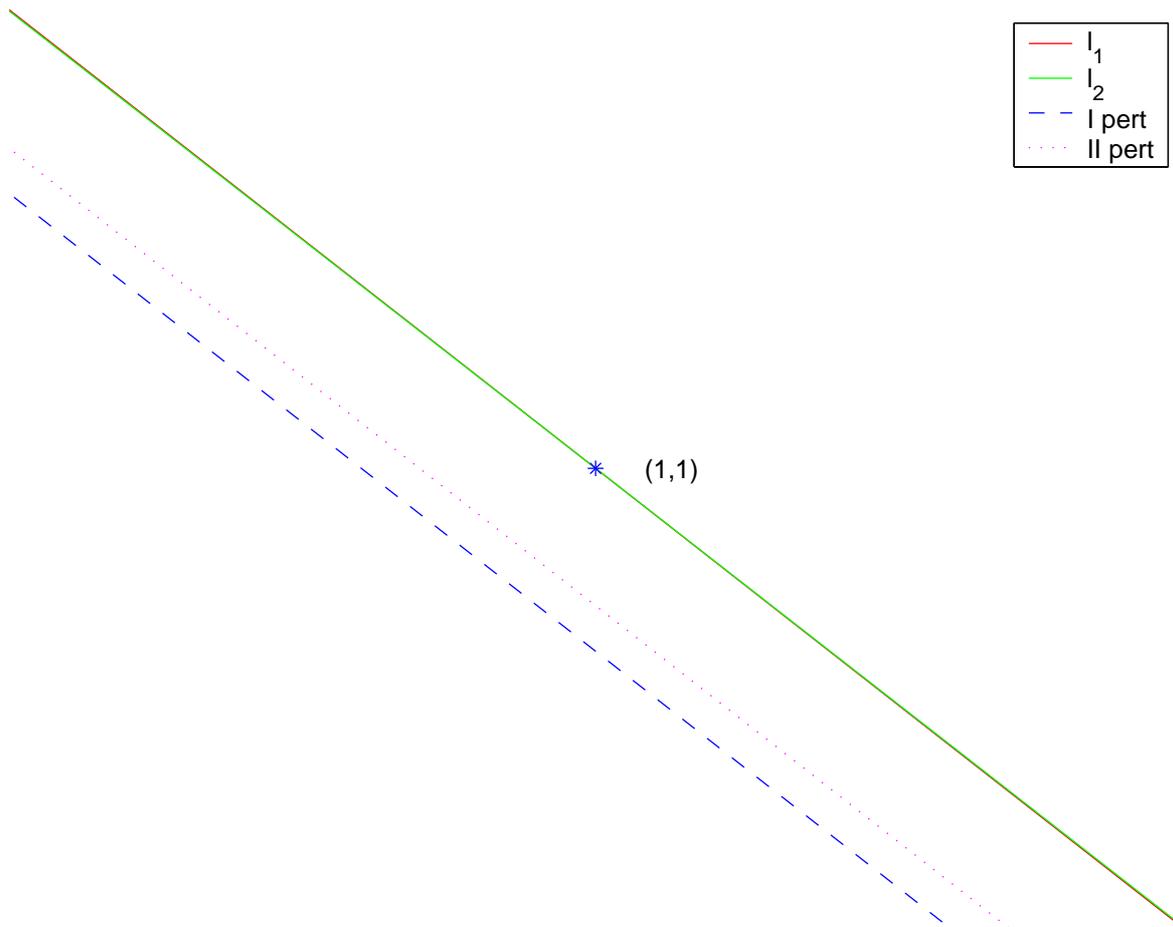
$$\begin{array}{ll}x_1 + 2x_2 = 3 & x_1 + 2x_2 = 3 \\ .5x_1 + 1.002x_2 = 1.5 & .499x_1 + 1.001x_2 = 1.4985\end{array}$$

la soluzione diventa:

$$(3, 0)^T \qquad (2, 0.5)^T$$

Perturbando di poco i dati iniziali, si trovano soluzioni diverse: si tratta di un **problema mal condizionato**.

Geometricamente, si tratta delle equazioni di due rette l_1 , l_2 quasi parallele, di cui si vuole trovare l'intersezione; perturbando l_2 di poco si ottengono altri punti di intersezione, $(3, 0)$ o $(2, 0.5)$, che distano poco da punti di l_2 e appartengono a l_1 .



Il vettore residuo $r = b - Aw$ (ove w è la soluzione di una equazione perturbata e A e b sono matrice e termine noto del sistema non perturbato) è piccolo in entrambe i casi, pur essendo w significativamente diverso dalla soluzione esatta x .

$$\begin{array}{cc}
 w = (3, 0)^T & w = (2, 0.5)^T \\
 r = \begin{pmatrix} 0 \\ 0.003 \end{pmatrix} & r = \begin{pmatrix} 0 \\ 0.0015 \end{pmatrix}
 \end{array}$$

A causa degli errori di rappresentazione dei dati del problema e degli errori di arrotondamento nelle operazioni, un qualunque metodo numerico su calcolatore determina una soluzione approssimata w invece della soluzione esatta $x = A^{-1}b$.

Un criterio che si utilizza per decidere se w è una approssimazione accettabile consiste nel richiedere che la norma del residuo sia piccola.

$$r = b - Aw$$

Se $\|r\| = 0 \Rightarrow \|b - Aw\| = 0 \Rightarrow w \equiv x$.

Tale criterio non è sempre valido.

Infatti, riscrivendo $r = b - Aw$ come

$$Aw = b - r$$

w è soluzione esatta di un sistema in cui il termine noto è perturbato di $-r$. Anche se r ha elementi piccoli, se il problema è mal condizionato, w può essere molto diverso da x .

Si osservi che:

$$r = b - Aw = Ax - Aw = A(x - w)$$

L'errore assoluto è soluzione del sistema $r = Ay$. Per cui:

$$x - w = A^{-1}r$$

↓

$$\begin{aligned} \|x - w\| &\leq \|A^{-1}\| \|r\| \\ \frac{\|x - w\|}{\|x\|} &\leq \|A\| \|A^{-1}\| \frac{\|r\|}{\|b\|} \end{aligned}$$

ove l'ultima disuguaglianza segue da

$$Ax = b \Rightarrow \|b\| \leq \|A\| \|x\| \Rightarrow \frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|}$$

Pertanto dalla piccolezza del residuo non si può dedurre che l'errore assoluto o l'errore relativo siano piccoli, poichè le quantità $\|A^{-1}\|$ oppure $\|A\| \|A^{-1}\|$ forniscono la connessione tra residuo e accuratezza.

Solo se $\|A^{-1}\|$ o $\|A\| \|A^{-1}\|$ sono piccole, si può accettare w come soluzione di $Ax = b$ se $\|r\|$ è piccola.

Negli esempi sopra, con $A = \begin{pmatrix} 1 & 2 \\ .499 & 1.001 \end{pmatrix}$ si ha:

$$A^{-1} = \begin{pmatrix} 1.001 & -2 \\ -.499 & 1 \end{pmatrix} \frac{1}{1.001 - .998}$$

$$\|A\|_{\infty} = 3 \quad \|A^{-1}\|_{\infty} = \frac{3.001}{.003} = 1000.333$$

$$\|A\|_{\infty} \|A^{-1}\|_{\infty} \simeq 3001$$

$$w = (3, 0)^T$$

$$r = \begin{pmatrix} 0 \\ 0.003 \end{pmatrix}$$

$$\|x - w\|_{\infty} = 2$$

$$\|r\|_{\infty} = 3 \cdot 10^{-3}$$

$$\|A^{-1}\|_{\infty} \|r\|_{\infty} \simeq 3$$

$$w = (2, 0.5)^T$$

$$r = \begin{pmatrix} 0 \\ 0.0015 \end{pmatrix}$$

$$\|x - w\|_{\infty} = 1$$

$$\|r\|_{\infty} = 1.5 \cdot 10^{-3}$$

$$\|A^{-1}\|_{\infty} \|r\|_{\infty} \simeq 1.5$$

Si osservi anche che se si hanno due soluzioni approssimate $w^{(1)}$ e $w^{(2)}$ con residui $r^{(1)}$ e $r^{(2)}$ non è vero che se $\|r^{(1)}\| < \|r^{(2)}\|$ la prima soluzione sia più accurata della seconda. Infatti,

$$w^{(1)} = (3, 0)^T$$

$$r^{(1)} = \begin{pmatrix} 0 \\ 0.003 \end{pmatrix}$$

$$\|r^{(1)}\|_{\infty} = 0.003$$

$$w^{(2)} = (0.4, 1.302)^T$$

$$r^{(2)} = \begin{pmatrix} -0.004 \\ -0.002902 \end{pmatrix}$$

$$\|r^{(2)}\|_{\infty} = 0.004$$

<

Ma

$$\|x - w^{(1)}\|_{\infty} = 2$$

>

$$\|x - w^{(2)}\|_{\infty} = 0.6$$

Teorema. Sia $A \in \mathbb{R}^{n \times n}$ non singolare. Dato il sistema $Ax = b$, nell'ipotesi di avere sulla matrice A una perturbazione δA tale che $\|\delta A\| < \frac{1}{\|A^{-1}\|}$, ove $\|\cdot\|$ è una norma naturale e $A + \delta A$ è non singolare, sia w la soluzione del sistema perturbato

$$(A + \delta A)w = b + \delta b$$

Allora

$$\frac{\|x - w\|}{\|x\|} \leq \frac{k(A)}{1 - k(A)\frac{\|\delta A\|}{\|A\|}} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right)$$

ove $k(A) = \|A\|\|A^{-1}\|$ si dice numero di condizione di A .

In particolare, se $\delta A = 0$ e $\delta b = -r$,

$$\frac{\|x - w\|}{\|x\|} \leq k(A) \frac{\|r\|}{\|b\|}$$

Se $\delta b = 0$,

$$\frac{\|x - w\|}{\|x\|} \leq \frac{k(A)}{1 - k(A)\frac{\|\delta A\|}{\|A\|}} \frac{\|\delta A\|}{\|A\|}$$

In una norma naturale, vale che $k(A) \geq 1$. Infatti

$$1 = \|I\| = \|AA^{-1}\| \leq \|A\|\|A^{-1}\| = k(A)$$

Se $k(A) \gg 1$, A è mal condizionata.

Se $k(A) \simeq 1$, A è ben condizionata.

$$k(I) = 1.$$

Se gli elementi di A sono normalizzati in modo che $\|A\| = 1$, un valore di $k(A)$ grande si riflette nell'enorme crescita di A^{-1} :

$$A = \frac{1}{2 + \epsilon} \begin{pmatrix} 1 & 1 \\ 1 & 1 + \epsilon \end{pmatrix} \quad \|A\|_{\infty} = 1$$

$$A^{-1} = \frac{2 + \epsilon}{\epsilon} \begin{pmatrix} 1 + \epsilon & -1 \\ -1 & 1 \end{pmatrix} \quad \|A^{-1}\|_{\infty} = \frac{(2 + \epsilon)^2}{\epsilon} = k(A)$$

Se $\epsilon = 10^{-k} \Rightarrow k(A) \simeq 10^k$.

La definizione di numero di condizione come

$$k(A) = \|A\| \|A^{-1}\|$$

è data per A non singolare. Si può estendere la definizione al caso di matrici qualunque.

Definizione. Sia $A \in \mathbb{R}^{m \times n}$. Si definisce numero di condizione di A rispetto a una norma naturale indotta da una norma vettoriale

$$k(A) = \frac{\max \|Ax\|_v}{\min \|Ax\|_v}$$

ove il minimo è fatto sui vettori $\|Ax\|_v \neq 0$. Il numero di condizione è il rapporto tra la perturbazione massima e la perturbazione minima che ogni vettore $x \in \mathbb{R}^n$ subisce per effetto della trasformazione lineare associata ad A .

Nel caso della norma euclidea, per definire il numero di condizione, introduciamo la definizione di valori singolari di A .

Definizione. Sia $A \in \mathbb{R}^{m \times n}$. Si dicono valori singolari di A le radici quadrate degli autovalori non nulli di $A^T A$:

$$\sigma_i = \sqrt{\lambda_i(A^T A)} \neq 0$$

Allora, rispetto alla norma euclidea, il numero di condizione si definisce come:

$$k_2(A) = \frac{\sqrt{\lambda_{max}(A^T A)}}{\sqrt{\lambda_{min}(A^T A)}}$$

ove $\lambda_{min}(A^T A) = \min_{\lambda \neq 0} \lambda(A^T A)$. Se $A \in \mathbb{R}^{n \times n}$ simmetrica, $k_2(A) = \frac{|\lambda(A)|_{max}}{|\lambda(A)|_{min}}$.

A normalizzata ($\|A\|_2 = 1$) è mal condizionata se e solo se ha un piccolo valore singolare.

A simmetrica normalizzata è mal condizionata se ha un piccolo autovalore in modulo.

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

$$\sigma_{min}^2(A) \leq |\lambda_i(A)|^2 \leq \sigma_{max}^2(A)$$

Se A ha un piccolo autovalore in modulo è mal condizionata.

Se è mal condizionata non è detto che abbia un piccolo valore autovalore.

In genere una matrice è mal condizionata se è vicina alla singolarità.

Se A ha determinante piccolo non è detto che sia mal condizionata.

$$A = \text{diag}(1/2, 1/2, \dots, 1/2) \quad \det(A) = \frac{1}{2^n} \quad k_2(A) = 1$$

Se A è mal condizionata non è detto che abbia determinante piccolo.

$$T = \begin{pmatrix} 1 & -1 & \dots & -1 \\ & 1 & \dots & -1 \\ & & \ddots & -1 \\ & & & 1 \end{pmatrix} \quad \det(T) = 1 \quad \|T\|_{\infty} = n$$

$$T^{-1} = \begin{pmatrix} 1 & 2^0 & 2^1 & \dots & 2^{n-2} \\ & 1 & 2^0 & \dots & 2^{n-3} \\ & & 1 & 2^0 & \dots \\ & & & 1 & 2^0 \\ & & & & 1 \end{pmatrix}$$

$$\|T^{-1}\|_{\infty} = 2^{n-1} \quad k(T)_{\infty} = n2^{n-1}$$

Esempio. Matrici di Hilbert $H_n = [h_{ij}]$.

$$h_{ij} = \frac{1}{i+j-1} \quad i, j = 1, \dots, n$$

$$H_4 = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 1/3 & 1/4 & 1/5 & 1/6 \\ 1/4 & 1/5 & 1/6 & 1/7 \end{pmatrix}$$

In aritmetica finita, molti elementi vengono perturbati. In particolare, con $\beta = 10, t = 5$,

$$G_4 = \begin{pmatrix} 1 & 0.50000 & 0.33333 & 1/4 \\ 0.50000 & 0.33333 & 0.25000 & 0.20000 \\ 0.33333 & 0.25000 & 0.20000 & 0.16666 \\ 0.25000 & 0.20000 & 0.16666 & 0.14285 \end{pmatrix}$$

H_n^{-1} ha elemento

$$\bar{h}_{ij} = \frac{(-1)^{i+j}(n+i-1)!(n+j-1)!}{(i+j-1)!((i-1)!(j-1)!)^2(n-j)!(n-i)!}$$

$$H_4^{-1} = \begin{pmatrix} 16 & -120 & 240 & -140 \\ -120 & 1200 & -2700 & 1680 \\ 240 & -2700 & 6480 & -4200 \\ -140 & 1680 & -4200 & 2800 \end{pmatrix}$$

$$\|H_4\|_\infty \|H_4^{-1}\|_\infty = \frac{25}{12} \cdot 13620 = 28375 \simeq 2.8 \cdot 10^4$$

n	$k_2(H_n)$	$k_\infty(H_n)$
2	1.505	27
3	$5.241 \cdot 10^2$	$7.480 \cdot 10^2$
4	$1.551 \cdot 10^4$	$2.837 \cdot 10^4$
5	$4.766 \cdot 10^5$	$9.436 \cdot 10^5$
6	$1.495 \cdot 10^7$	$2.907 \cdot 10^7$
7	$4.754 \cdot 10^8$	$9.852 \cdot 10^8$
8	$1.526 \cdot 10^{10}$	$3.387 \cdot 10^{10}$
9	$4.932 \cdot 10^{11}$	$1.099 \cdot 10^{12}$
10	$1.603 \cdot 10^{13}$	$3.535 \cdot 10^{13}$

Il calcolo di $k(A)$ implica la valutazione di A^{-1} . Ma calcolare A^{-1} vuol dire risolvere $A\alpha_j = e_j$, $j = 1, \dots, n$, con il costo di risoluzione di n sistemi e gli stessi problemi connessi alla soluzione di $Ax = b$ in aritmetica finita.

Allora si usa una stima di $\|A^{-1}\|$ ottenuta nel seguente modo:

- si calcola una soluzione approssimata w_0 ;
- si calcola il residuo in doppia precisione $r_0 = b - Aw_0 = A(x - w_0)$;

Pertanto $x - w_0$ è soluzione di

$$Ay = r_0$$

La soluzione calcolata è una approssimazione di $y = A^{-1}r_0$:

$$y_0 = A^{-1}r_0$$

$$\|y_0\| \simeq \|x - w_0\| \leq \|A^{-1}\| \|r_0\|$$

Pertanto $\frac{\|y_0\|}{\|r_0\|}$ è una sottostima di $\|A^{-1}\|$. Si può anche provare che

$$\|r_0\| \simeq 10^{-t} \|A\| \|w_0\|$$

ove t è il numero di cifre dell'aritmetica usata.

$$\|y_0\| \simeq \|x - w_0\| \leq \|A^{-1}\| \|r_0\| \simeq \|A^{-1}\| \|A\| 10^{-t} \|w_0\|$$

Da cui,

$$k(A) \simeq 10^t \frac{\|y_0\|}{\|w_0\|}$$

E' opportuno calcolare r_0 in doppia precisione.

Esempio.

$$\begin{aligned}x_1 + 2x_2 &= 3 \\ .499x_1 + 1.001x_2 &= 1.5\end{aligned}$$

Si assume $t = 3$.

$$w_0 = (0, 1.5)^T$$

$$r_0 = \begin{pmatrix} 3 - 0 - 2 \cdot 1.5 \\ 1.5 - .499 \cdot 0 - 1.001 \cdot 1.5 \end{pmatrix} = \begin{pmatrix} 0 \\ -1.5 \cdot 10^{-3} \end{pmatrix}$$

$$\|r_0\|_\infty = 1.5 \cdot 10^{-3}$$

Si risolve $Ay = r_0 \Rightarrow y_0 = \begin{pmatrix} 1.5 \\ -0.75 \end{pmatrix}$, $\|y_0\|_\infty = 1.5$.

Si osservi che $\frac{\|y_0\|_\infty}{\|r_0\|_\infty} = 10^3$ che è una sottostima di $\|A^{-1}\|_\infty = 1000.333$.

Inoltre $\|r_0\|_\infty \simeq 10^{-3} \|A\|_\infty \|w_0\|_\infty = 10^{-3} \cdot 3 \cdot 1.5 = 4.5 \cdot 10^{-3}$. Da ciò

$$3001 \simeq k_\infty(A) \simeq 10^3 \frac{\|y_0\|_\infty}{\|w_0\|_\infty} = \frac{1.5 \cdot 10^3}{1.5} = 1000$$

.

Raffinamento iterativo della soluzione

Si è osservato che, dato $Ax = b$, con un metodo numerico si calcola una approssimazione w_0 della soluzione.

- Soluzione di $Ax = b \Rightarrow w_0$
- $r_0 = b - Aw_0 = A(x - w_0)$, calcolato in doppia precisione
- Soluzione di $Ay = r_0 \Rightarrow y_0$; calcolo di $w_1 = w_0 + y_0$
- $r_1 = b - Aw_1 = A(x - w_1)$, calcolato in doppia precisione
- Soluzione di $Ay = r_1 \Rightarrow y_1$; calcolo di $w_2 = w_1 + y_1$
- ...

Per sistemi molto mal condizionati il metodo fallisce o converge molto lentamente. Ciò capita se $k(A)$ è molto grande rispetto alla precisione di macchina. E' opportuno controllare ad ogni passo se si verifica che:

$$\frac{\|y_i\|}{\|w_{i+1}\|} < \frac{1}{2} \frac{\|y_{i-1}\|}{\|w_i\|}$$

In questo caso il metodo converge lentamente.

Algoritmo. Sia N il massimo numero di iterazioni previste e TOL una tolleranza prefissata ($TOL = u\|A\|$):

```
begin   $r = b;$    $x = 0;$ 
       $k = 0;$ 
      while   $k \leq N$  do
      begin  risolvere  $Ay = r;$ 
             $xx = x + y;$ 
             $r = b - A * xx;$  D.P.
            if  $\|x - xx\|_{\infty} < TOL$  then
                stampare  $xx;$ 
            else
                 $k = k + 1;$ 
                 $x = xx$ 
            end;
      end;
end;
```

Stabilità dei metodi diretti

In aritmetica finita, i fattori di A o di PA sono affetti da errore. Detti \mathcal{L} e \mathcal{R} i fattori calcolati, essi possono essere ritenuti fattori esatti di una matrice perturbata mediante una matrice δA :

$$PA + \delta A = \mathcal{L}\mathcal{R} = (L + \delta L)(R + \delta R) = LR + L\delta R + \delta LR + \delta L\delta R$$

$$\Rightarrow \delta A = L\delta R + \delta LR + \delta L\delta R$$

δA è piccolo (e dunque \mathcal{L} e \mathcal{R} sono accettabili) se gli elementi di L e di R non sono troppo grandi rispetto ad A . Si cercano algoritmi che mantengano L e R limitati. Tali algoritmi si dicono **stabili**.

Sia A fattorizzabile e normalizzata in modo che $\max_{i,j} |a_{i,j}| = 1$.

$$A = LR \quad (PA = LR)$$

Se esistono costanti positive a e b indipendenti dagli elementi di A e dall'ordine di A tali che

$$|l_{ij}| \leq a \quad |r_{ij}| \leq b$$

la fattorizzazione LR si dice stabile in senso forte.

Se a e b dipendono dall'ordine di A , allora la fattorizzazione di A si dice stabile in senso debole.

- Sia A simmetrica definita positiva.

$$A = LL^T \Rightarrow a_{ii} = \sum_{j=1}^i l_{ij}^2 \Rightarrow l_{ij}^2 \leq a_{ii} \leq \max_{r,s} |a_{rs}|$$

$$\frac{1}{\max_{r,s} |a_{rs}|} A = \frac{L}{\sqrt{\max_{r,s} |a_{rs}|}} \frac{L^T}{\sqrt{\max_{r,s} |a_{rs}|}} = HH^T$$

ove $H = \frac{L}{\sqrt{\max_{r,s} |a_{rs}|}}$.

$$|h_{ij}| = \frac{|l_{ij}|}{\sqrt{\max_{r,s} |a_{rs}|}} \leq 1$$

Poichè $a = b = 1$, la fattorizzazione di Cholesky è stabile in senso forte.

- Algoritmo di eliminazione di Gauss con pivoting parziale: si dimostra:

$$|l_{ij}| \leq 1 \quad |r_{ij}| \leq 2^{n-1} \max_{r,s} |a_{rs}|$$

Infatti per la scelta del perno come elemento di modulo massimo sulla colonna k a partire dalla posizione diagonale k , i moltiplicatori sono in modulo minori o

uguali a 1 ($|m_{ij}| \leq 1 \Rightarrow |l_{ij}| \leq 1 \Rightarrow a = 1$).

$$a_{ij}^{(2)} = a_{ij} - m_{i1}a_{1j} \Rightarrow |a_{ij}^{(2)}| \leq 2\max_{r,s}|a_{rs}|$$

$$a_{ij}^{(3)} = a_{ij}^{(2)} - m_{i2}a_{2j}^{(2)} \Rightarrow |a_{ij}^{(3)}| \leq 2\max_{r,s}|a_{rs}^{(2)}|$$

$$\leq 2^2\max_{r,s}|a_{rs}|$$

...

$$a_{ij}^{(n)} = a_{ij}^{(n-1)} - m_{in-1}a_{n-1j}^{(2)} \Rightarrow |a_{ij}^{(n)}| \leq 2^{n-1}\max_{r,s}|a_{rs}|$$

Per cui $b = 2^{n-1}$. Esistono matrici per cui tale limite è raggiunto.

$$a_{ij} = \begin{cases} 1 & i = j \\ -1 & j < i \\ 1 & j = n \\ 0 & \text{altrove} \end{cases} \quad r_{ij} = \begin{cases} 1 & i = j \\ 2^{i-1} & j = n \\ 0 & \text{altrove} \end{cases}$$

L coincide con il triangolo inferiore di A e $r_{nn} = 2^{n-1}$. Se r_{nn} viene alterato, $\tilde{r}_{nn} = 2^{n-1} - \epsilon$, $L\tilde{R}$ è fattorizzazione esatta di una matrice che differisce dalla A solo in $\tilde{a}_{nn} = 1 - \epsilon$. Se $\epsilon = 0.5$, \tilde{R} è circa uguale a R . Ma in tal modo \tilde{R} diventa fattore di una matrice molto perturbata rispetto ad A .

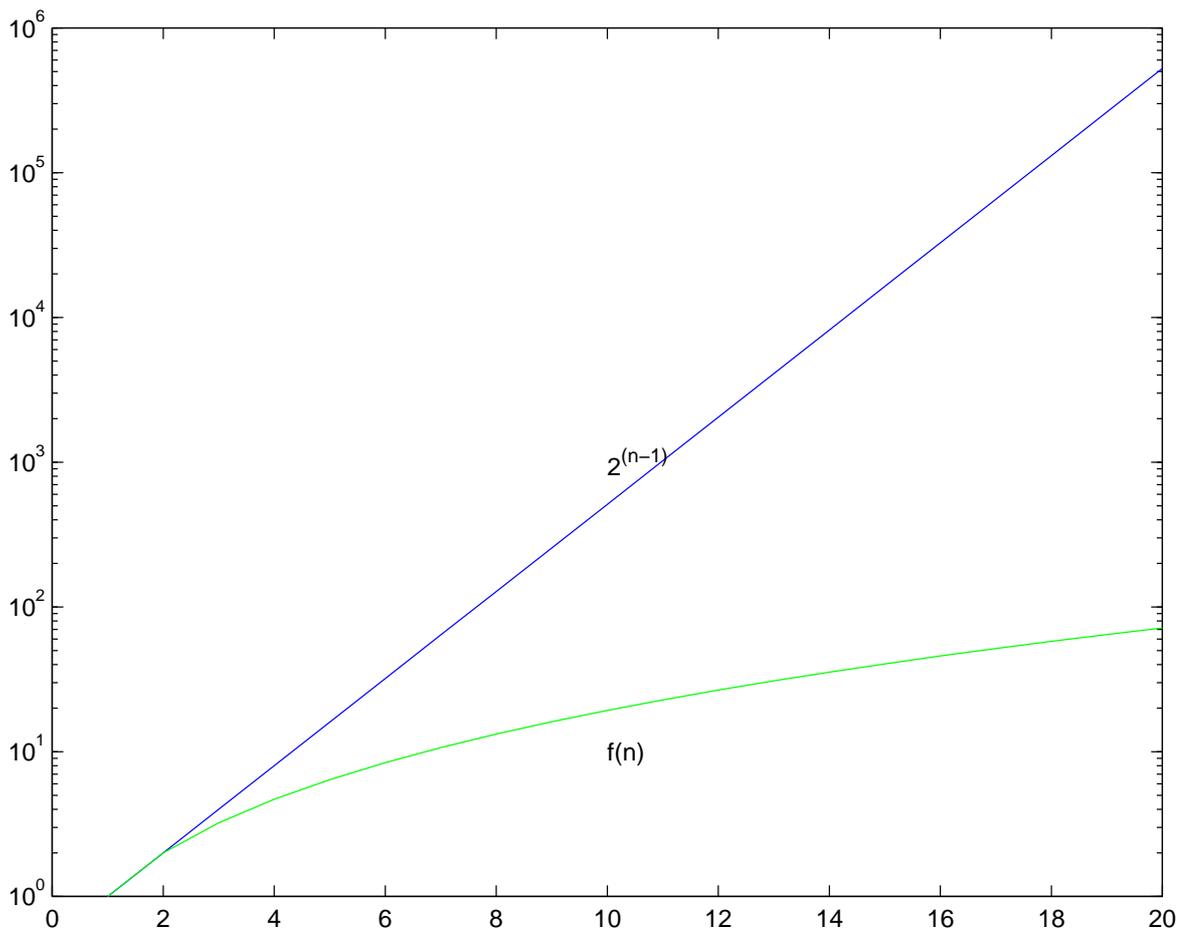
- Nel caso di pivoting totale, si dimostra che:

$$|l_{ij}| \leq 1 \quad |r_{ij}| \leq f(n)\max_{r,s}|a_{rs}|$$

$$f(n) = \sqrt{n} \sqrt{1.2.3^{1/2}4^{1/3} \dots n^{1/(n-1)}}$$

Non si conoscono matrici per cui vale l'uguaglianza. Per $n \leq 4$, si dimostra che $f(n) = n$.

n	$f(n)$ G.P.T	2^{n-1} G.P.P
10	19	2^9
20	67	2^{19}
50	530	2^{49}
100	3300	2^{99}



- Nel caso di matrici di Hessemberg, l'uso del pivoting parziale produce $a = 1$ e $b = n$

- Nel caso di matrici tridiagonali, diagonali dominanti, $a = 1, b = 2$.
- Per la fattorizzazione QR , vale che:

$$\max_{i,j} |q_{ij}| = 1/n \|Q\|_T \leq \|Q\|_2 = 1$$

Da $Q^T A = R$, segue per ogni j :

$$\begin{aligned} \max_i |r_{ij}| &= \|r_{*j}\|_\infty \leq \|r_{*j}\|_2 = \|Q^T a_{*j}\|_2 \\ &\leq \|Q^T\|_2 \sqrt{n} \|a_{*j}\|_\infty = \sqrt{n} \max_i |a_{ij}| \end{aligned}$$

In questo caso $a = 1$ e $b = \sqrt{n}$. La fattorizzazione è stabile in senso debole.

Analisi all'indietro dell'errore nella risoluzione di un sistema

Partendo da numeri finiti e supponendo che le operazioni di macchina non alterino la scelta del pivot, si dimostra che la fattorizzazione di Gauss ottenuta in aritmetica finita è la fattorizzazione esatta di

$$(PA + \delta A) = \mathcal{L}\mathcal{R}$$

$$\|\delta A\|_{\infty} \leq u \cdot n^2 \cdot \max_{i,j} |r_{ij}|$$

Le soluzioni calcolate \tilde{y} e \tilde{x} dei sistemi $\mathcal{L}y = Pb$ e $\mathcal{R}x = y$ sono soluzioni esatte di

$$\begin{cases} (\mathcal{L} + \delta\mathcal{L})\tilde{y} = Pb & \|\delta\mathcal{L}\|_{\infty} \leq 1.01u \frac{n(n+1)}{2} \max |l_{ij}| \\ (\mathcal{R} + \delta\mathcal{R})\tilde{x} = \tilde{y} & \|\delta\mathcal{R}\|_{\infty} \leq 1.01u \frac{n(n+1)}{2} \max |r_{ij}| \end{cases}$$

Allora \tilde{x} è soluzione esatta di

$$(PA + E)\tilde{x} = Pb$$

$$(\mathcal{L} + \delta\mathcal{L})(\mathcal{R} + \delta\mathcal{R})\tilde{x} = Pb$$

$$(\mathcal{L}\mathcal{R} + \mathcal{L}\delta\mathcal{R} + \delta\mathcal{L}\mathcal{R} + \delta\mathcal{L}\delta\mathcal{R})\tilde{x} = Pb$$

Poichè $PA + \delta A = \mathcal{L}\mathcal{R}$, segue

$$E = \delta A + \mathcal{L}\delta\mathcal{R} + \delta\mathcal{L}\mathcal{R} + \delta\mathcal{L}\delta\mathcal{R}$$

$$\begin{aligned}
\|E\|_\infty &\leq n^2 u \max|r_{ij}| + \\
&+ n \cdot 1.01 u \frac{n(n+1)}{2} \max|r_{ij}| + \\
&+ 1.01 u \frac{n(n+1)}{2} n \max|r_{ij}| + \\
&+ (1.01)^2 u^2 \left(\frac{n(n+1)}{2} \right)^2 \max|r_{ij}| \\
&\leq u \max|r_{ij}| (n^2 + 1.01(n^3 + n^2) + 1.02 n^2) \\
&= 1.01 u \max|r_{ij}| (3n^2 + n^3)
\end{aligned}$$

posto $u \frac{(n+1)^2}{4} < 1$.

Conclusione: a partire da numeri finiti, la soluzione calcolata è soluzione esatta di

$$(PA + E)\tilde{x} = Pb$$

con $\|E\|_\infty \leq 1.01 u(3n^2 + n^3) f(n) \max|a_{ij}|$ ove

$$f(n) = \begin{cases} 1 & \text{A simm. def. pos.} \\ n & \text{A di Hessemberg con Piv. p.} \\ 2^{n-1} & \text{Gauss piv. p.} \\ \sqrt{n} \sqrt{1 \cdot 2 \cdot 3^{1/2} \dots n^{1/(n-1)}} & \text{Gauss piv. t.} \\ 2 & \text{matrici tridiagonali diag. dom.} \\ \sqrt{n} & \text{QR} \end{cases}$$

Nella pratica $\|E\|_\infty \leq u n \|A\|_\infty$. Si conclude che:

$$\frac{\|x - \tilde{x}\|_\infty}{\|x\|_\infty} \leq \frac{k_\infty(A)}{1 - k_\infty(A) \frac{\|E\|_\infty}{\|A\|_\infty}} \frac{\|E\|_\infty}{\|A\|_\infty}$$