CALCOLO NUMERICO

Francesca Mazzia

Dipartimento Interuniversitario di Matematica

Università di Bari

4. Algoritmi per la soluzione di sistemi lineari.

Sistemi triangolari inferiori

Sia L triangolare inferiore. La matrice è non singolare se gli elementi principali $l_{i,i} \neq 0$, per $i = 1, \cdots, n$.

Il sistema $Lx = b$ può scriversi:

$$
\begin{array}{ll}
l_{1,1}x_1 & = b_1 \\
l_{2,1}x_1 + l_{2,2}x_2 & = b_2 \\
l_{3,1}x_1 + l_{3,2}x_2 + l_{3,3}x_3 & = b_3 \\
\dots & \\
\dots & \\
l_{n,1}x_1 + l_{n,2}x_2 - \dots + l_{n,n}x_n & = b_n
\end{array}
$$

da cui si ottiene:

$$x_1 = \frac{b_1}{l_{1,1}}, \qquad x_2 = \frac{b_2 - l_{2,1}x_1}{l_{2,2}}$$

$$x_3 = \frac{b_3 - l_{3,1}x_1 - l_{3,2}x_2}{l_{3,3}}$$

in generale

$$x_i = \frac{b_i - \sum_{j=1}^{i-1} l_{i,j} x_j}{l_{i,i}} \qquad i = 1, 2, \ldots, n$$

Essendo $l_{i,i} \neq 0$ queste formule determinano in modo univoco x_i.

La procedura descritta si chiama algoritmo di sostituzione in avanti, in Matlab è descritta dal seguente algoritmo:

```
function b = soll(L,b)
n = length(b);
b(1) = b(1)/L(1,1);
for i=2:n
  b(i) = (b(i) - L(i,1:i-1)*b(1:i-1))/L(i,i);
end
```

L'istruzione matlab `L(i,1:i-1)*b(1:i-1)` esegue il prodotto riga per colonne fra il vettore `L(i,1:i-1)` e il vettore `b(1:i-1)`.

`L(i,1:i-1)` elementi

(`L(i,1)`, `L(i,2)`, $\cdots$, `L(i,i-1)`)

`b(1:i-1)` elementi

(`b(1)`, `b(2)`, $\cdots$, `b(i-1)`)

Costo computazionale:

moltiplicazioni :

`L(i,1:i-1)*b(1:i-1)`

richiede i-1 moltiplicazioni

viene eseguita per `i=2,n` sommando

$$\sum_{2}^{n}(i-1) = \sum_{1}^{n-1} i = \frac{n(n-1)}{2}$$

Il numero di addizioni è dello stesso tipo.

Sistemi triangolari superiori

$$
\begin{aligned}
u_{1,1}x_1 + u_{1,2}x_2 + \ldots + u_{1,n}x_n &= b_1 \\
u_{2,2}x_2 + \ldots + u_{2,n}x_n &= b_2 \\
\ldots & \\
\ldots & \\
u_{n,n}x_n &= b_n
\end{aligned}
$$

La matrice è non singolare se $u_{i,i} \neq 0$.

La soluzione di questo sistema è:

$$x_n = \frac{b_n}{u_{n,n}}, \qquad x_{n-1} = \frac{b_{n-1} - u_{n-1,n}x_n}{u_{n-1,n-1}}$$

$$x_{n-2} = \frac{b_{n-2} - u_{n-2,n-1}x_{n-1} - u_{n-2,n}x_n}{u_{n-2,n-2}}$$

In generale:

$$x_i = \frac{1}{u_{i,i}} \left(b_i - \sum_{j=i+1}^{n} u_{i,j} x_j \right) \qquad i = n, n-1, \ldots, 1$$

La procedura descritta si chiama algoritmo di sostituzione all'indietro, in Matlab è descritta dal seguente algoritmo:

```
function b = solu(U,b)
n = length(b);
b(n) = b(n)/U(n,n);
for i=n-1:-1:1
  b(i) = (b(i) - U(i,i+1:n)*b(i+1:n))/U(i,i);
end
```

Matrici di permutazione

$P_{r,s}$ matrici di permutazione elementari si ottengono dalla matrice I di dimensione $n \times n$ scambiando la r.ma riga con la s.ma.

$$
P_{r,s} = \begin{array}{c}
\begin{array}{ccccccccccc} & & r & & & & s & & & & \end{array} \\
\begin{pmatrix}
1 & 0 & & & & & & & & & \\
0 & \ddots & \ddots & & & & & & & & \\
 & \ddots & 1 & & & & & & & & \\
 & & & 0 & 0 & \dots & 0 & 1 & & & \\
 & & & 0 & 1 & & & 0 & & & \\
 & & & \vdots & & \ddots & & \vdots & & & \\
 & & & 0 & & & 1 & 0 & & & \\
 & & & 1 & 0 & \dots & 0 & 0 & & & \\
 & & & & & & & & 1 & \ddots & \\
 & & & & & & & & \ddots & \ddots & 0 \\
 & & & & & & & & & 0 & 1
\end{pmatrix}
\end{array}
\begin{array}{c}
\\ \\ \\ \\ r \\ \\ \\ \\ s \\ \\ \\ \\
\end{array}
$$

$P_{r,s}A$ scambia le righe r ed s

$AP_{r,s}$ scambia le colonne r ed s

Proprietà delle matrici di permutazione

Siano P, P_1, P_2 matrici di permutazione $n \times n$ e A una matrice $n \times n$ allora:

1. PA è come A con le righe permutate.

   AP è come A con le colonne permutate

2. $P^{-1} = P^T$

3. $\det(P) = \pm 1$

4. $P_1 P_2$ è un matrice di permutazione.

Fattorizzazione LU per risolvere $Ax = b$:

1. Fattorizzare A in modo che $PA = LU$ con

   P = matrice di permutazione

   L = matrice triangolare inferiore con elementi diagonali uguali a 1

   U = matrice triangolare superiore non singolare

2. Risolvere $P^{-1}z = b$, permutando le righe di b: $z = LUx = Pb$

3. risolvere$Ly = Pb$,$y = Ux$ con l'algoritmo di sostituzione in avanti

4. risolvere $Ux = L^{-1}(P^b)$ con l'algoritmo di sostituzione all'indietro : $x = U^{-1}(L^{-1}(Pb))$

La sottomatrice principale di A di ordine j è $A(1:j, 1:j)$.

Teorema. Le seguenti definizioni sono equivalenti:

1. Esiste un'unica matrice triangolare L con elementi principali uguali ad 1 a un'unica matrice trangolare superiore U non singolare tale che $A = LU$.

2. Tutte le sottomatrici principali di A sono non singolari

Dimostrazione: (1) implica (2)

Partizioniamo le matrici A, L e U

$$A = LU \equiv$$

$$\begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} = \begin{pmatrix} L_{1,1} & 0 \\ L_{2,1} & L_{2,2} \end{pmatrix} \begin{pmatrix} U_{1,1} & U_{1,2} \\ 0 & U_{2,2} \end{pmatrix}$$

$$= \begin{pmatrix} L_{1,1}U_{1,1} & L_{1,1}U_{1,2} \\ L_{2,1}U_{1,1} & L_{2,1}U_{1,2} + L_{2,2}U_{2,2} \end{pmatrix}$$

$A_{1,1}$ sottomatrice principale di dimensione j.

$\det(A_{1,1}) = \det(L_{1,1}U_{1,1}) = det(L_{1,1})\det(U_{1,1}) = 1 \cdot \prod_{k=1}^{j}(u_{k,k}) \neq 0$

proviamo che (2) implica (1) per induzione su n.

Per matrice di ordine 1 è vero: $a = 1 \cdot a$.

Supponiamo sia vero per $n-1$ e dimostriamolo per n.

Sia $\tilde{A}$ di ordine n, allora dobbiamo mostrare che esistono e sono unici :

L e U triangolari di ordine $n-1$,

s e t due vettori di lunghezza $n-1$ e η scalare tali che:

$$\tilde{A} = \begin{pmatrix} A & b \\ c^T & \delta \end{pmatrix} = \begin{pmatrix} L & 0 \\ s^T & 1 \end{pmatrix} \begin{pmatrix} U & t \\ 0 & \eta \end{pmatrix}$$

$$= \begin{pmatrix} LU & Lt \\ s^T U & s^T t + \eta \end{pmatrix}$$

Per induzione esistono L e U tali che $A = LU$. Siano $t = L^{-1}b$, $s^T = c^T U^{-1}$ e $\eta = \delta - s^T t$, essi sono unici. Gli elementi diagonali di U sono diversi da zero per induzione e $\eta \neq 0$ poichè $0 \neq det(\tilde{A}) = det(U)\eta$.

Problemi dell'algoritmo di fattorizzazione LU senza permutazioni:

Non può essere eseguita su matrici come:

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 & 5 \\ -1 & -2 & -7 \\ 1 & 1 & 3 \end{pmatrix}$$

Presenta problemi di instabilità numerica:

$$\begin{cases} 10^{-17}x_1 + x_2 & = 1 \\ \qquad\quad x_1 + x_2 & = 2 \end{cases}$$

Il problema è ben posto, ed ha come soluzione esatta $x \simeq (1,1)^T$. Calcolando la fattorizzazione LU ed operando in virgola mobile in base 2 in doppia precisione, si ottengono i fattori L ed U seguenti:

$$L = \begin{pmatrix} 1 & 0 \\ 10^{17} & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 10^{-17} & 1 \\ 0 & -10^{17} \end{pmatrix}$$

il prodotto LU non è A bensì la matrice seguente

$$\tilde{A} = \begin{pmatrix} 10^{-17} & 1 \\ 1 & 0 \end{pmatrix}$$

confrontando $\tilde{A}$ con A si vede che l' elemento a_{22} di A è stato cambiato, pertanto la soluzione numerica che è $\tilde{x} = (0,1)^T$ è totalmente sbagliata.

Teorema. Se A è non singolare, allora esistono due matrici di permutazione P_1 e P_2, una matrice triangolare inferiore L con elementi diagonali uguali a uno e una matrice triangolare superiore U tali che $P_1AP_2 = LU$. Solo una fra P_1 e P_2 è necessaria.

P_1A scambia le righe di A,

AP_2 scambia le colonne,

P_1AP_2 scambia righe e colonne di A.

Dimostrazione per induzione.

Per matrici di ordine 1 basta scegliere:

$P_1 = P_2 = L = 1$ e $U = A$.

Supponiamo sia vero per matrici di ordine $n-1$ e sia A di ordine n.

Essendo A non singolare, allora ha sicuramente un elemento diverso da zero.

Scegliamo P_1' e P_2' matrici di permutazione in modo da rendere l'elemento in posizione (1,1) di $P_1'AP_2'$ diverso da zero.

Scriviamo la fattorizzazione e risolviamo per le componenti incognite.

$$P_1'AP_2' = \begin{pmatrix} a_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ L_{2,1} & I \end{pmatrix} \begin{pmatrix} u_{1,1} & U_{1,2} \\ 0 & \tilde{A}_{2,2} \end{pmatrix}$$

$$\begin{pmatrix} a_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} = \begin{pmatrix} u_{1,1} & U_{1,2} \\ L_{2,1}u_{1,1} & L_{2,1}U_{1,2} + \tilde{A}_{2,2} \end{pmatrix}$$

$A_{2,2}$ e $\tilde{A}_{2,2}$ di ordine $n-1$, $L_{2,1}$ e $U_{1,2} \in \mathbb{R}^{n-1}$

Risolvendo per le componenti di questa fattorizzazione a blocchi otteniamo:

$u_{1,1} = a_{1,1} \neq 0$

$U_{1,2} = A_{1,2}$

$L_{2,1} = A_{2,1}/a_{1,1}$

$\tilde{A}_{2,2} = A_{2,2} - L_{2,1}U_{1,2}$

$\det(P_1' A P_2') = \pm\det(A) \neq 0$

$\det(P_1' A P_2') = 1 \cdot u_{1,1}\det(\tilde{A}_{2,2})$

$\det(\tilde{A}_{2,2}) \neq 0$

quindi per induzione esistono due matrici di permutazione $\tilde{P}_1$ e$\tilde{P}_2$ tali che $\tilde{P}_1\tilde{A}_{2,2}\tilde{P}_2 = \tilde{L}\tilde{U}$. Sostituendo si ha:

$$P_1'AP_2' = \begin{pmatrix} 1 & 0 \\ L_{2,1} & I \end{pmatrix} \begin{pmatrix} u_{1,1} & U_{1,2} \\ 0 & \tilde{P}_1^T\tilde{L}\tilde{U}\tilde{P}_2^T \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ L_{2,1} & I \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \tilde{P}_1^T\tilde{L} \end{pmatrix} \begin{pmatrix} u_{1,1} & U_{1,2} \\ 0 & \tilde{U}\tilde{P}_2^T \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ L_{2,1} & \tilde{P}_1^T\tilde{L} \end{pmatrix} \begin{pmatrix} u_{1,1} & U_{1,2}\tilde{P}_2 \\ 0 & \tilde{U} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \tilde{P}_2^T \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & \tilde{P}_1^T \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \tilde{P}_1L_{2,1} & \tilde{L} \end{pmatrix} \begin{pmatrix} u_{1,1} & U_{1,2}\tilde{P}_2 \\ 0 & \tilde{U} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \tilde{P}_2^T \end{pmatrix}$$

quindi

$$P_1AP_2 = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{P}_1 \end{pmatrix} P_1'AP_2' \begin{pmatrix} 1 & 0 \\ 0 & \tilde{P}_2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ \tilde{P}_1L_{2,1} & \tilde{L} \end{pmatrix} \begin{pmatrix} u_{1,1} & U_{1,2}\tilde{P}_2 \\ 0 & \tilde{U} \end{pmatrix}$$

Possiamo scegliere $P_2' = I$ e P_1' in modo tale che $a_{1,1}$ sia il più grande elemento in valore assoluto nella colonna, il che implica che $L_{2,1} = A_{2,1}/a_{1,1}$ ha elementi limitati da 1 in valore assoluto. In generale al passo i della fattorizzazione, ordiniamo le righe da i a n in modo da portare l'elemento più grande in valore assoluto in posizione i, i. Questa tecnica si chiama fattorizzazione con Pivot Parziale.

Possiamo scegliere P_2' e P_1' in modo tale che $a_{1,1}$ sia il più grande elemento in valore assoluto di tutta la matrice. In generale al passo i della fattorizzazione, ordiniamo le righe e le colonne da i a n in modo da portare l'elemento più grande in valore assoluto in posizione i, i. Questa tecnica si chiama fattorizzazione con Pivot Totale.

```
function [L,U] = fattlu(A)
[n,m] = size(A);


for k=1:n-1
    if abs(A(k,k)) <= realmin
      error(' Un elemento pivotale e'' piccolo')
    else
      L(k+1:n,k)= A(k+1:n,k)/A(k,k);
    end
    A(k+1:n,k+1:n)=A(k+1:n,k+1:n) - ....
    L(k+1:n,k)*A(k,k+1:n);
end
U = triu(A);
end
```

La funzione Matlab che esegue la fattorizzazione LU con pivot è **lu**. Essa può essere richiamata con l'istruzione:

```
>> [L,U,P]=lu(A)
```

Supponiamo che $\det(A) = 0$ e U abbia l'ultimo elemento diagonale nullo.

Quindi l'ultima riga di U è un vettore nullo che si ottiene moltiplicando l'ultima riga di L^{-1} per le righe di A, cioè:

$$l_{n,1}a_1^T + \ldots + l_{n,n-1}a_{n-1}^T + l_{n,n}a_n^T = 0$$

con $l_{n,i}$ non tutti nulli e ciò implica la lineare dipendenza delle righe di A.

L'annullarsi del determinante di una matrice quadrata è condizione necessaria e sufficiente affinchè i vettori riga siano linearmente dipendenti.

Lo stesso discorso può farsi con i vettori colonna.

Numero di operazioni

1) Calcolo dei coefficienti $m_{i,j}$. Questi sono $n-1$ per la prima colonna, $n-2$ per la seconda, fino ad 1 per la penultima. In totale abbiamo:
$$\sum_{k=1}^{n-1}(n-k) = \frac{n(n-1)}{2}$$
divisioni.

2) Calcolo degli $a_{i,j}^{(k)}$. Questi sono $(n-1)^2$ al primo passo $(k=2)$, $(n-2)^2$ al secondo passo fino ad 1 all'ultimo passo $(k=n-1)$. In totale il numero degli $a_{i,j}^{(k)}$ calcolati sono:
$$\sum_{k=1}^{n-1}(n-k)^2 = \frac{n(n-1)(2n-1)}{6} \sim \frac{n^3}{3}.$$

Riassumendo:

moltiplicazioni e divisioni:$\sim n^3/3$ sottrazioni: $\sim n^3/3$.

Sia $Ax = b$ e $(A + \delta A)\widehat{x} = b + \delta b$, $\widehat{x} = x + \delta x$

Per studiare il comportamento di δx in funzione della perturbazione sui dati di input, utilizziamo il lemma seguente.

Lemma. Sia $\|\cdot\|$ una norma matriciale. Allora $\|B\| < 1$ implica che $I - B$ è invertibile e $\|(I - B)^{-1}\| \leq 1/(1 - \|B\|)$.

Dimostrazione. Sia λ autovalore di B con x autovettore, allora $(I - B)$ ha l'autovalore $1 - \lambda$ con x autovettore:

$$(I - B)x = x - Bx = x - \lambda x = (1 - \lambda)x$$

quindi $(I - B)$ è non singolare, infatti gli autovalori sono tutti diversi da zero e quindi il determinante è diverso da zero, essendo uguale al prodotto degli autovalori.

$$I = (I - B)(I - B)^{-1} = (I - B)^{-1} - B(I - B)^{-1}$$
$$I + B(I - B)^{-1} = (I - B)^{-1}$$
$$\|(I - B)^{-1}\| \leq \|I\| + \|B\|\|(I - B)^{-1}\|$$
$$\|(I - B)^{-1}\| - \|B\|\|(I - B)^{-1}\| \leq 1$$
$$\|(I - B)^{-1}\| \leq 1/(1 - \|B\|)$$

Troviamo una maggiorazione relativa per $\|\delta x\|$

$(A + \delta A)(x + \delta x) - Ax = b + \delta b - b$
$\delta Ax + A\delta x + \delta A\delta x = \delta b$
$\delta Ax + (A + \delta A)\delta x = \delta b$
$(A + \delta A)\delta x = \delta b - \delta Ax$
$(I + A^{-1}\delta A)\delta x = A^{-1}(\delta b - \delta Ax)$
Se $\|A^{-1}\|\|\delta A\| < 1$ possiamo applicare il lemma e

$$\delta x = (I + A^{-1}\delta A)^{-1}A^{-1}(\delta b - \delta Ax)$$

passando alle norme

$$\|\delta x\| \leq \|(I+A^{-1}\delta A)^{-1}\|\|A^{-1}\|(\|\delta b\|+\|\delta A\|\|x\|)$$

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

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

$\|b\| = \|Ax\| \leq \|A\|\|x\|$

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

La quantità $k(A) = \|A\|\|A^{-1}\|$ è il numero di condizione della matrice A

Residuo : $r = A\widehat{x} - b$

$r = 0$ se $\widehat{x} = x$

$A\delta x = A\widehat{x} - Ax = A\widehat{x} - b$

$\delta x = A^{-1} r$

$\|\delta x\| \leq \|A^{-1}\| \|r\|$

Matrice di Hilbert

La matrice di Hilbert rappresenta un classico esempio di matrice *mal condizionata* reale, nel senso che ha origine da modelli reali.

Gli elementi della matrice di Hilbert sono dati dalla espressione seguente:

$$a_{ij}^{(n)} = \frac{1}{i+j-1}, \qquad i,j = 1,...,n$$

La tabella riporta i valori del numero di condizione di A, in norma 2 e in norma ∞, per diversi valori di n.

n	$\kappa_2(A^{(n)})$	$\kappa_\infty(A^{(n)})$
2	1.505	27
3	$5.241\ 10^{2}$	$7.480\ 10^{2}$
4	$1.551\ 10^{4}$	$2.837\ 10^{4}$
5	$4.766\ 10^{5}$	$9.436\ 10^{5}$
6	$1.495\ 10^{7}$	$2.907\ 10^{7}$
7	$4.754\ 10^{8}$	$9.852\ 10^{8}$
8	$1.526\ 10^{10}$	$3.387\ 10^{10}$
9	$4.932\ 10^{11}$	$1.099\ 10^{12}$
10	$1.603\ 10^{13}$	$3.535\ 10^{13}$

Si consideri la matrice A e il vettore b definiti da:

$$A = \begin{bmatrix} 1.2969 & 0.8648 \\ 0.2161 & 0.1441 \end{bmatrix}, b = \begin{bmatrix} 0.8642 \\ 0.1440 \end{bmatrix}.$$

La soluzione esatta del sistema $Ax = b$ è $x = [2, -2]^T$.

Sia $\widehat{x} = [0.9911, -0.4870]^T$, e sia $r = A\widehat{x} - b$, quindi $r = [-10^{-8}, 10^{-8}]^T$.

Questo valore di r fa intendere che $\widehat{x}$ è una buona approssimazione della soluzione esatta anche se in realtà non lo è.

La matrice inversa di A è data da

$$A^{-1} = -10^{-8} \begin{bmatrix} 0.8648 & -0.1441 \\ -1.2969 & 0.2161 \end{bmatrix}$$

per cui $\kappa_\infty(A) = 3\ 10^8$, quindi A è mal condizionata.

Il residuo non è una buona indicazione di precisione della soluzione di un sistema quando la matrice è mal condizionata. Infatti, vale la disuguaglianza

$$\frac{\|\delta x\|}{\|x\|} \leq \kappa(A) \frac{\|r\|}{\|b\|}$$

che si ottiene da

$$\|\delta x\| \leq \|A^{-1}\| \|r\|$$

dividendo per $\|x\|$ e sfruttando la disuguaglianza $\|b\| \leq \|A\| \cdot \|x\|$.