# Analisi Numerica Relazioni di Laboratorio

# Indrjo Dedej

Ultima revisione: 25 maggio 2023.

## Indice

1 Metodi diretti per sistemi 2 Metodi di ricerca di zeri lineari 1

Nota 0.1. Per la parte dei codici è stato usato GNU Octave al posto di Matlab.

**Nota 0.2.** Se A è il nome di una matrice, allora scriviamo  $A_{i,j}$  per indicare l'entrata di A alla i-esima riga e j-esima colonna.

## Metodi diretti per sistemi lineari

**Esercizio 1.1.** Scrivere una funzione tri\_solve che prenda in input una matrice triangolare A e un vettore b e calcoli la soluzione del sistema lineare Ax = b. Tale funzione deve gestire sia il caso triangolare superiore che triangolare inferiore (si usino eventualmente le funzioni istriu e istril).

Una matrice quadrata A di ordine n si dice triangolare superiore (rispettivamente, inferiore) qualora  $A_{i,j} = 0$  per ogni  $i, j \in \{1, ..., n\}$  con i > j (rispettivamente, i < j). Un sistema triangolare inferiore è quindi fatto quindi di equazioni

$$\sum_{j=1}^{i} A_{i,j} x_j = b_i \quad \text{per } i \in \{1, \dots, n\}.$$

Si verifica immediatamente che la soluzione  $x \coloneqq \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^T$  di questo sistema è tale che

$$x_{1} = \frac{b_{1}}{A_{1,1}}$$

$$x_{i} = \frac{1}{A_{i,i}} \left( b_{i} - \sum_{j=1}^{i-1} A_{i,j} x_{j} \right)$$

La funzione qui sotto prende una matrice triangolare inferiore A e un vettore b e restituisce la soluzione del sistema Ax = b.

```
function [x] = solve_lower_triangular(A,b)
n = length(b);
x(1) = b(1)/A(1,1);
for i = 2:n
x(i) = (b(i) - A(i,1:i-1) * x(1:i-1)')/A(i,i);
end
x = x';
end
```

Invece, la soluzione  $x := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^T$  di un sistema triangolare superiore

$$\sum_{j=n-i+1}^n A_{i,j} x_j = b_i \quad \text{per } i \in \left\{1, \dots, n\right\}$$

è data da

$$x_n = \frac{b_n}{A_{n,n}}$$

$$x_{n-i+1} = \frac{1}{A_{n-i+1,n-i+1}} \left( b_n - \sum_{j=n-i+2}^{i-1} A_{i,j} x_j \right)$$

La funzione che risolve Ax = b nel caso in cui A sia triangolare superiore è:

```
function [x] = solve_upper_triangular(A,b)

n = length(b);

x(n) = b(n)/A(n,n);

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

end
x = x';
end
```

Pertanto la funzione che risolve un sistema triangolare Ax = b è la seguente.

```
function [x] = solve_triangular(A,b)
if istril(A)
        [x] = solve_lower_triangular(A,b);
elseif istriu(A)
        [x] = solve_upper_triangular(A,b);
else
        disp("the matrix isn't triangular!");
        return;
end
end
```

**Esercizio 1.2.** Scrivere due funzioni, lu\_nopiv e lu\_piv, che calcolino la fattorizzazione LU con e senza pivoting di una matrice, rispettivamente.

L'algoritmo di eliminazione gaussiana consente, presa una matrice invertibile A di ordine n, di trasformare il sistema

$$Ax = b$$

in uno equivalente

$$Ux = b'$$

dove U è triangolare superiore. Il passaggio da A ad U avviene attraverso una matrice L triangolare inferiore con gli elementi sulla diagonale principale uguali ad 1:

$$A = LU$$
.

Le entrate della matrice L sono ottenute durante l'algoritmo di Gauss. Qui sotto ci sono le funzioni che ad una matrice A restituisce la fattorizzazione LU, la prima senza e la seconda con pivotazione.

```
function [L U] = lu_nopiv(A)
    % l'ordine della matrice quadrata A
    n = rows(A);
    % All'inizio, la matrice triangolare inferiore è
    % l'identità, ma di volta in volta verrà aggiornata.
    L = eye(n);
    % Facciamo una copia di A, la quale verrà modificata
    % passo dopo passo. In realtà, non ce ne sarebbe il
    % bisogno, in quanto basta alterare A direttamente.
    U = A;
10
    for i = 1:n
11
      for j = i+1:n
12
13
        % Questi sono i coefficienti ottenuti durante
        % il MEG. La cosa importante è che U(i, i) = /= 0.
14
        % La pivotazione evita ciò, ma in questa funzione
15
        % non la consideriamo.
16
        L(j,i) = U(j,i)/U(i,i);
17
        % Le righe successive all'i-esima sono ottenute
18
        % sottraendo un opportuno multiplo dell'i-esima.
19
        U(j,:) = U(j,:)-L(j,i)*U(i,:);
21
22
    end
23
  end
```

La fattorizzazione LU con pivotazione coinvolge un'altra matrice:

$$PA = LU$$

dove *P* è ottenuta dall'identità permutandone le righe.

```
function [L U P] = lu_piv(A)
    n = rows(A);
    L = eye(n);
    U = A;
    P = eye(n);
      \% Ricerca del primo j per cui |A(j,i)| è massimo
      m = 0; k = i;
      for j = i:n
        if abs(A(j,i)) > m
10
          k = j; m = abs(A(j,i));
11
        end
12
      end
13
      L([i k],1:i-1) = L([k i],1:i-1);
      U([i k],:) = U([k i],:);
15
16
      P([i k],:) = P([k i],:);
      % Da qui in poi si procede come in lu_nopiv.
17
      for j = i+1:n
18
        % Fatta la pivotazione, si può proseguire come
19
```

```
L(j,i) = U(j,i)/U(i,i);

Le righe successive all'i-esima sono ottenute

sottraendo un opportuno multiplo dell'i-esima.

U(j,:) = U(j,:)-L(j,i)*U(i,:);

end

end

end

end
```

**Esercizio 1.3.** Scrivere una funzione che, presa in input una matrice reale A di ordine n non singolare, ne calcoli l'inversa usando la fattorizzazione LU.

In generale, l'inversa di A è una matrice  $X = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$  tale che

$$Ax_i = e_i \quad \text{per } i \in \{1, \dots, n\}.$$

Si tratta quindi di risolvere n sistemi lineari. Quando è disponibile la fattorizzazione LU di A, però si può seguire una strada che usa quanto implementato prima, ma che aumenta il numero di sistemi da risolvere. Sia quindi PA = LU, e quindi i sistemi diventano

$$LUx_i = Pe_i$$
.

Le soluzioni  $x_i$  si determinano risolvendo i sistemi triangolari

$$Ly_i = Pe_i$$
 e  $Ux_i = y_i$  per  $i \in \{1, ..., n\}$ .

```
function [X] = inverse_lu(A)

[L U P] = lu_piv(A);

n = rows(A);

for i=1:n

Y = solve_lower_triangular(L, P(:,i));

X(:,i) = solve_upper_triangular(U, Y);

end

end
```

**Esercizio 1.4.** Scrivere una funzione che calcola il determinante di una matrice A usando la fattorizzazione PA = LU. Per calcolare il determinate di P, si modifichi la funzione scritta al punto 2 in modo da contare il numero di scambi di righe effettuati.

Da PA = LU abbiamo che det  $P \det A = \det L \det U = \det U$ . Essendo P una permutazione delle righe dell'identità, si ha  $\det P = \pm 1$ . Inoltre si nota che PP = I. Quindi, in principio, la funzione qui sotto potrebbe andare bene.

```
function x = lu_det1(A)

[L U P] = lu_piv(A);

x = det(P)*det(U);
end
```

Tuttavia questa implementazione richiede inutilmente alla macchina di calcolare L. Inoltre di P si sa che è I con le righe permutate, quindi det P è 1 se c'è stato un numero pari di scambi, -1 altrimenti. In breve, det  $P=(-1)^{\delta}$ , dove  $\delta$  è il numero di scambi di righe effettuati. Una versione più economica quindi è:

```
function [d] = lu_det2(A)
n = rows(A); U = A;
delta = 0;
```

```
for i = 1:n
      m = 0; k = i;
      for j = i:n
        if abs(A(j,i)) > m
          k = j; m = abs(A(j,i));
      U([i \ k],:) = U([k \ i],:);
11
      if i != k
12
        delta = delta+1;
13
14
       for j = i+1:n
15
        m = U(j,i)/U(i,i);
16
        U(j,:) = U(j,:)-m*U(i,:);
17
18
19
    % A questo punto U è una matrice triangolare: il suo
20
    % determinante è dato dal prodotto delle entrate
21
    % della diagonale principale.
22
    det_U = U(1,1);
23
    if n \ge 2
      for i = 2:n
25
        det_U = det_U * U(i,i);
26
27
    d = (-1)^delta*det_U;
```

Esercizio 1.5. Scrivere due funzioni sgs e mgs che calcolino la fattorizzazione QR ridotta di una matrice, usando rispettivamente il metodo di Gram-Schmidt e il metodo di Gram-Schmidt modificato.

Siano  $a_1, ..., a_n \in \mathbb{R}^m$  linearmente indipendenti. L'algoritmo di Gram-Schmidt dà n vettori  $q_1, \dots, q_n \in \mathbb{R}^m$  tra loro ortogonali:

$$\begin{aligned} q_1 &:= x_1 \\ q_i &:= x_i - \sum_{j=1}^{i-1} \frac{q_j^T x_i}{\|q_j\|^2} q_j = x_i - \sum_{j=1}^{i-1} \left[ \left( \frac{q_j}{\|q_j\|} \right)^T x_i \right] \frac{q_j}{\|q_j\|} \quad \text{per } i \in \{2, \dots, n\} \end{aligned}$$

Se introduciamo la matrice R triangolare superiore attraverso

$$R_{j,i} \coloneqq \begin{cases} \|q_i\| & \text{se } i = j \\ \left(\frac{q_j}{\|q_j\|}\right)^T x_i & \text{se } j \le i - 1 \\ 0 & \text{altrimenti} \end{cases}$$

possiamo scrivere

$$X = QR$$
,

dove  $Q \coloneqq \begin{bmatrix} \frac{q_1}{\|q_1\|} & \cdots & \frac{q_n}{\|q_n\|} \end{bmatrix}$ . Ecco l'implementazione di questo algoritmo.

```
function [Q R] = sgs(A)
  [m n] = size(A);
   Q = A;
  R = zeros(n, n);
```

```
for i = 1:n
    for j = 1:i-1
        R(j,i) = Q(:,j)'*A(:,i);
        Q(:,i) = Q(:,i)-R(j,i)*Q(:,j);
end
        R(i,i) = norm(Q(:,i));
        Q(:,i) = Q(:,i)/R(i,i);
end
a end
```

Per ragioni numeriche, è meglio però l'implementazione che segue:

```
function [Q R] = mgs(A)

[m n] = size(A);

R = zeros(n, n);

for i = 1:n

R(i,i) = norm(A(:,i));

Q(:,i) = A(:,i)/R(i,i);

for j = i+1:n

R(i,j) = Q(:,i)'*A(:,j);

A(:,j) = A(:,j)-R(i,j)*Q(:,i);

end
end
end
```

**Esercizio 1.6.** Sia m = 50 e n = 12. Sia  $f(t) = \cos(4t)$  e si definiscano i punti  $t_i = \frac{i-1}{m-1}$  per i = 1, ..., m e b il vettore di componenti  $b_i = f(t_i)$ . Sia A la matrice del problema dei minimi quadrati che si ottiene approssimando con un polinomio di grado n - 1 la sequenza dei punti  $(t_i, b_i)$ . Si risolva e visualizzi (con tutti e sedici i decimali) la soluzione del problema calcolata con i seguenti metodi:

- 1. formazione e soluzione del sistema di equazioni normali usando il comando \ di MATLAB
- 2. fattorizzazione QR utilizzando sgs
- 3. fattorizzazione QR utilizzando mgs
- 4. fattorizzazione QR utilizzando qr (funzione built-in di MATLAB che utilizza il metodo di Householder).

I calcoli precedenti generano quattro liste di 12 coefficienti. Prendendo come riferimento la soluzione con l'ultimo metodo, si cancellino in ogni lista le cifre decimali che appaiono errate (affette da arrori di arrotondamento). Quali metodi si mostrano più instabili? Non è necessario spiegare le osservazioni fatte

In generale, presi m punti  $(x_1, y_1)$ , ... e  $(x_m, y_m)$  di  $\mathbb{R}^2$ , si cerca un polinomio  $p_* \in \mathbb{P}_n$ , con n < m - 1, tale che

$$\sum_{i=1}^{m} (p_{*}(x_{i}) - y_{i})^{2} \leq \sum_{i=1}^{m} (p(x_{i}) - y_{i})^{2} \quad \text{per ogni } p \in \mathbb{P}_{n}.$$

Per ricondurci al problema dei minimi quadrati, basta osservare che, se

$$p(x) = c_1 + c_2 x + c_3 x^3 + \dots + c_{n+1} x^n$$

allora

$$\sum_{i=1}^{m} (p(x_i) - y_i)^2 = ||Ac - y||_2^2$$

dove

$$A := \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_i & x_i^2 & \cdots & x_i^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^m \end{bmatrix} \in \mathbb{R}^{m \times (n+1)}.$$

Per trovare quindi il  $p_* \in \mathbb{P}_n$  che minimizza  $\sum_{i=0}^n (p(x_i) - y_i)^2$  è equivalente a trovare il  $c \in \mathbb{R}^{n+1}$  che minimizza  $||Ac - y||_2^2$ . Cioè si deve risolvere l'equazione

$$A^T A c = A^T \gamma$$

rispetto a c.

Diamo una funzione che presa una lista di ascisse  $x \in \mathbb{R}^m$  e un numero naturale n produce la matrice A di sopra:

```
function [A] = vandermonde(n, x)
% !!! Qui x è un vettore *colonna* di R^m
for i = 0:n
A(:,i+1) = x.^i;
end
end
```

Diamo a anche le seguenti funzioni che risolvono il problema nei modi richiesti nell'ordine.

```
function [c] = minquad(n, x, y)
% !!! Qui, x e y sono vettori *colonna* di R^m
A = vandermonde(n, x');
% Dobbiamo risolvere il problema A'*Ac = A'y in c.
% Qui, lo facciamo con l'operatore \ di Matlab.
c = (A'*A)\(A'*y);
end
```

Se usiamo la decomposizione A = QR dell'esercizio precedente, allora

$$R^T \underbrace{Q^T Q}_{=I} Rc = R^T Q^T y$$

da cui, perché R è triangolare,

$$Rc = Q^T y$$
.

Le funzioni che seguono si basano sulla risoluzione di questa equazione. Abbiamo già dato negli esercizi precedenti le funzioni che risolvono sistemi triangolari superiori.

```
function [c] = minquad_sgs(n, x, y)
    % !!! Qui, x e y sono vettori *colonna* di R^m
    A = vandermonde(n, x);
    % Dobbiamo risolvere il problema A'*Ac = A'y in c.
    [Q R] = sgs(A);
    c = solve_upper_triangular(R, Q'*y);
end

function [c] = minquad_mgs(n, x, y)
    % !!! Qui, x e y sono vettori *colonna* di R^m
    A = vandermonde(n, x');
    % Dobbiamo risolvere il problema A'*Ac = A'y in c.
    [Q R] = mgs(A);
    c = solve_upper_triangular(R, Q'*y);
end
```

```
function [c] = minquad_qr(n, x, y)
A = vandermonde(n, x');

% Dobbiamo risolvere il problema A'*Ac = A'y in c.
[Q R] = sgs(A);
% Qui Q è una matrice m*m, mentre R è m*(n+1).
% Ma R è trapeizoidale superiore, quindi
Q = Q(:,1:n+1);
R = R(1:n+1,:);
c = (inverse_lu(R) * Q') * y;
end
```

Testiamo le funzioni appena implementate. Prima di tutto, diamo i vettori di  $\mathbb{R}^{50}$ :

```
octave> format long
octave> t = (@(i) (i-1)/49)(1:50)';
octave> b = (@(x) cos(4*x))(t);
```

Chiediamo ora di risolvere il problema ai minimi quadrati, usando un polinomi di grado 11:

```
octave> c1 = minquad(11, t, b)
c1 =

9.999999998598327e-01
-1.637969801790591e-07
-7.999989743148679e+00
-2.072197377938371e-04
1.066866439686697e+01
-1.069780529188339e-02
-5.655104676152303e+00
-6.193169964808305e-02
1.679168741518508e+00
1.575953725551536e-02
-3.779623750302716e-01
8.865738750316085e-02
```

```
loctave> c2 = minquad_sgs(11, t, b)
c2 =

1.000010858217736e+00
-1.597441622479323e-03
-7.961790735605575e+00
-3.563609846100917e-01
1.232237295279629e+01
-4.198742022992519e+00
1.143539843615144e-01
-3.712350976158632e+00
1.556815307967554e+00
1.317363348240178e+00
-7.841381000343972e-01
5.040691898039995e-02
```

```
octave> c3 = minquad_mgs(11, t, b)
c3 =

9.99999993467416e-01
2.936344856152573e-08
-7.999997687339061e+00
```

```
-8.218757375288988e-05
1.066765586643305e+01
-5.947011668467894e-03
-5.669068083108868e+00
-3.556294157169759e-02
1.647176449303515e+00
3.985971841029823e-02
-3.882198705396149e-01
9.054209982423345e-02
```

```
octave> c4 = minquad_qr(11, t, b)

c4 =

1.000000000996612e+00

-4.227432497394545e-07

-7.999981235680764e+00

-3.187632846675115e-04

1.066943079626799e+01

-1.382028918305878e-02

-5.647075624787249e+00

-7.531602913513780e-02

1.693606968736276e+00

6.032105768099427e-03

-3.742417020548601e-01

8.804057586530689e-02
```

## 2 Metodi di ricerca di zeri

## Esercizio 2.1. Scrivere una funzione

```
[c,fc,iter]=bisezione(f,a,b,tol,itmax)
```

metodo di bisezione per la ricerca di una funzione f nell'intervallo [a, b]. Qui, tol è la tolleranza sull'approssimazione della radice e itmax è il numero massimo di iterazioni disposte a fare. Per quanto riguarda l'output, c è l'approssimazione trovata, fc è f(c) e iter è il numero di iterazioni fatte effettivamente per arrivare all'approssimazione c. Testare la funzione per la ricerca di una delle radici di

$$f(x) \coloneqq e^x - x^2 - \sin x - 1$$

sull'intervallo [-2, 2].

L'implementazione qui sotto usa il "teorema degli zeri" e l'algoritmo di bisezione.

```
function [c, fc, iter] = bisezione(f, a, b, tol, itmax)
   % INPUT:
              : funzione di cui cercare le radici
       a, b : estremi delll'intervallo di ricerca
   %
       tol : sulla soluzione trovata
   %
       itmax : numero massimo di iterazioni disposti
   %
               a compiere per approssimare uno zero
   % OUTPUT:
            : approssimazione trovata
   %
       С
   %
       fc
              : la funzione f valutata in c
10
       iter : numero iterazioni effettuate
```

```
iter = 0;
12
    while abs(a-b) > tol && iter < itmax
13
      % Nel caso in cui uno dei due estremi è
14
      % uno zero, ci fermiamo direttamente.
15
      if f(a) == 0
16
17
        c = a; return;
      elseif f(b) == 0
19
        c = b; return;
      else
20
        % Altrimenti, suddividiamo l'intervallo [a, b]
21
        % in due intervalli di uguale lunghezza, [a, c]
22
        % e [c, b], dove c è il punto medio di [a, c], e
23
        % si seleziona quello che ha uno zero.
24
         c = (a+b)/2;
         if f(a)*f(c) < 0
26
           b = c; % lo zero sta nell'intervallo [a, c]
27
28
           a = c; % lo zero sta nell'intervallo [c, b]
29
30
31
      end
32
    iter = iter+1;
33
    end
    fc = f(c);
34
35 end
```

Testiamo il metodo di bisezione su f:

```
>> f = @(x) exp(x)-x.^2-sin(x)-1;
>> df = @(x) exp(x)-2*x-cos(x);
>> [c, fc, it] = bisezione(f, 1, 1.5, 1e-12,100)
c = 1.2797
fc = 6.6813e-13
it = 39
```

#### Esercizio 2.2. Scrivere una funzione

```
[c,fc,iter]=newton(f,df,x0,tol,itmax)
```

che implementa il metodo di Newton-Raphson per la ricerca di uno zero di una funzione f con derivata df. Qui, x0 è il punto da cui parte il metodo, tol è la tolleranza sull'approssimazione della radice e itmax è il numero massimo di iterazioni disposte a fare. Per quanto riguarda l'output, c è l'approssimazione trovata, fc è f(c) e iter è il numero di iterazioni fatte effettivamente per arrivare all'approssimazione c. Testare la funzione per la ricerca di una delle radici di

$$f(x) \coloneqq e^x - x^2 - \sin x - 1$$

sull'intervallo [-2, 2].

Richiamiamo che il metodo di Newton-Raphson si scrive come

$$\begin{cases} x_0 \in \mathbb{R} \text{ scelto} \\ x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \text{ per } n \in \mathbb{N} \end{cases}$$

Osserviamo che se dobbiamo controllare l'incremento, ci dobbiamo arrestare appena che il valore assoluto di  $\frac{f(x_n)}{f'(x_n)}$  scende al di sotto di una certa tolleranza.

```
function [a, fa, iter] = newton(f, df, a, tol, itmax)
    % INPUT:
              : funzione di cui cercare le radici
    %
       f
    %
        df
              : derivata prima della funzione f
              : punto da cui far partire il metodo
              : sulla soluzione trovata
    %
        itmax : numero massimo di iterazioni disposti
                 a fare per l'approssimazione
    %
    % OUTPUT:
    % a : approssimazione trovata
10
           : la funzione f valutata in a
11
    % iter : numero massimo di iterazioni disposti
12
               a fare per arrivare ad a
13
   iter = 0;
    delta = tol+1; % per innescare il loop
15
    while iter < itmax && abs(delta) > tol
16
      iter = iter+1;
17
      delta = f(a)/df(a);
18
     a = a-delta;
19
    end
20
    fa = f(a);
21
22 end
```

Testiamo il metodo di bisezione su f:

```
>> f = @(x) exp(x)-x.^2-sin(x)-1;

>> df = @(x) exp(x)-2*x-cos(x);

>> [c, fc, it] = newton(f, df, 1.5, 1e-6,100)

c = 1.2797

fc = 2.2204e-16

it = 5
```

# Esercizio 2.3. Testo

```
function [errs] = newton_errors(f, df, a, m, tol, itmax)
    % INPUT:
    %
        f
              : funzione di cui cercare le radici
    %
        df
              : derivata prima della funzione f
              : punto da cui far partire il metodo
    %
        а
              : molteplicità della radice attesa
    %
       tol : sulla soluzione trovata
    %
        itmax : numero massimo di iterazioni disposti
    %
                 a fare per l'approssimazione
    % OUTPUT:
10
    % errs : lista degli errori commessi da un passo
11
    %
                al successivo
    iter = 0;
13
    delta = tol+1; % per innescare il loop
14
    while iter < itmax && abs(delta) > tol
15
      iter = iter+1;
16
      delta = m * f(a)/df(a);
17
      a = a-delta;
18
      errs(iter) = abs(delta);
19
    end
  end
```