

# Esame di Calcolo Numerico — 11 settembre 2023

Corso di Laurea in Ingegneria Chimica

Tempo a disposizione: 2 ore. È consentito consultare appunti e testi (cartacei).

**Esercizio 1 (15 punti)** Dati un vettore  $\mathbf{b} \in \mathbb{R}^n$  e una funzione  $f : \mathbb{R} \rightarrow \mathbb{R}$ , vogliamo costruire un'approssimazione di  $f$  della forma

$$r(x) = \sum_{j=1}^n \frac{w_j}{x + b_j}. \quad (1)$$

Nella (1) compaiono dei coefficienti incogniti  $\mathbf{w} = [w_1, w_2, \dots, w_n] \in \mathbb{R}^n$ , e vogliamo determinarli imponendo che valga l'uguaglianza

$$f(x_i) = r(x_i), \quad i = 1, 2, \dots, n, \quad (2)$$

per  $n$  reali distinti  $x_1, x_2, \dots, x_n$ .

1. Formulare il sistema di equazioni (2) come sistema lineare nell'incognita  $\mathbf{w}$ , scrivendone esplicitamente la matrice e il vettore dei termini noti (in funzione di  $\mathbf{b}, \mathbf{x}, f$ ). Qual è la dimensione di questa matrice?
2. Scrivere una `function w = ratinterp(b, x, y)` che calcoli il vettore dei coefficienti  $\mathbf{w}$  a partire dai vettori  $\mathbf{b}, \mathbf{x} = [x_1; x_2; \dots; x_n]$  e  $\mathbf{y} = [f(x_1); f(x_2); \dots; f(x_n)]$ . Riportarne sul foglio il codice.
3. Utilizzando la funzione scritta, calcolare (e riportare sul foglio) il vettore  $\mathbf{w}$  che risolve questo problema di approssimazione con  $f(x) = \exp(-x)$ ,

$$\mathbf{x} = \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}.$$

4. (\*) Per ognuno dei quattro valori  $z \in \{1, 3, 5, 7\}$ , calcolare (e riportare sul foglio) l'errore di interpolazione  $|r(z) - f(z)|$ . *Suggerimento: un modo possibile di farlo è scrivere una breve funzione che calcola il valore di  $r(z)$  in un punto  $z$  a partire da  $\mathbf{b}, \mathbf{w}, z$ .*
5. (\*) Due dei valori calcolati al punto precedente dovrebbero essere significativamente minori degli altri due; come mai?

**Esercizio 2 (15 punti)** Vogliamo risolvere un problema ai valori iniziali utilizzando il metodo di Runge-Kutta con tabella di Butcher

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \hline \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{array} \quad (3)$$

1. Scrivere esplicitamente le equazioni che legano  $y_{n+1}$  a  $y_n$ , per un'equazione differenziale generica  $y' = f(t, y)$ , usando tale metodo (3).
2. Scrivere una `function [T, Y] = rkeuler(N, lambda)` che, dato in ingresso il numero di intervalli  $N$  e un numero  $\lambda \in \mathbb{C}$ , risolve il problema test  $y' = \lambda y$ ,  $y(0) = 1$  sull'intervallo  $[0, 1]$  con il metodo (3), restituendo  $T, Y \in \mathbb{R}^{N+1}$  che contengono rispettivamente le approssimazioni  $t_n$  e  $y_n$  calcolate dal metodo ad ogni passo  $n = 0, 1, \dots, N$ . Riportare sul foglio il codice Matlab della funzione.
3. Per  $\lambda = -2.5$  e  $N \in \{50, 100, 200\}$ , riportare l'errore globale di discretizzazione  $E_N = \max_{n=1,2,\dots,N} \|y(t_n) - y_n\|_\infty$  tra la soluzione numerica calcolata e quella esatta. Cosa indicano i risultati ottenuti sull'ordine di convergenza del metodo?
4. Verificare (calcolandola) che la funzione di stabilità del metodo (3) è uguale a  $(1 + \frac{1}{2}q)^2$ . Qual è la regione di stabilità di questo metodo?

## Soluzioni

### Esercizio 1 (15 punti)

1. Scrivendo per esteso le (2), otteniamo il sistema lineare

$$\begin{bmatrix} \frac{1}{x_1+b_1} & \frac{1}{x_1+b_2} & \cdots & \frac{1}{x_1+b_n} \\ \frac{1}{x_2+b_1} & \frac{1}{x_2+b_2} & \cdots & \frac{1}{x_2+b_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{x_n+b_1} & \frac{1}{x_n+b_2} & \cdots & \frac{1}{x_n+b_n} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{bmatrix}.$$

2. Una soluzione possibile è la seguente.

```
function w = ratinterp(b, x, y)
n = length(b);
if not(length(x) == n) || not(length(y) == n)
    error('b, x, y devono avere la stessa lunghezza')
end

X = zeros(n, n);
for i = 1:n
    for j = 1:n
        X(i, j) = 1 / (x(i) + b(j));
    end
end

w = X \ y;
```

3. Una soluzione possibile è la seguente.

```
>> b = [1; 2; 3; 4]; x = b; y = exp(-x);
>> w = ratinterp(b, x, y)
w =
    -4.7871
    55.8314
   -117.6856
    67.8619
```

4. Scrivendo la funzione

```
function r = calcola(b, w, z)
r = 0;
n = length(b);

for j = 1:n
    r = r + w(j) / (z + b(j));
end
```

possiamo calcolare

```
>> abs(calcola(b, w, 1) - exp(-1))
ans =
    4.4409e-16
>> abs(calcola(b, w, 3) - exp(-3))
ans =
    1.2143e-15
>> abs(calcola(b, w, 5) - exp(-5))
```

```
ans =
    8.3201e-04
>> abs(calcola(b, w, 7) - exp(-7))
ans =
    0.0049
```

5. I primi due valori sarebbero uguali a zero se facessimo tutti i calcoli in aritmetica esatta, perché tra le condizioni (2) stiamo imponendo che  $f(1) = r(1)$  e  $f(3) = r(3)$ . Quindi il piccolo valore che osserviamo è solo un errore algoritmico dell'ordine della precisione di macchina. Gli altri valori invece sono più grandi, perché in generale  $f(5) \neq r(5)$ ,  $f(7) \neq r(7)$ : le due funzioni  $f(x)$  e  $r(x)$  non hanno nessun motivo per essere uguali al di fuori dei punti appartenenti al vettore  $\mathbf{x}$ .

## Esercizio 2 (15 punti)

1. Usando direttamente la definizione dei metodi di Runge–Kutta, le equazioni risultano

$$\kappa_1 = f(t_n, y_n) \quad (4a)$$

$$\kappa_2 = f(t_n + \frac{1}{2}h, y_n + \frac{1}{2}h\kappa_1) \quad (4b)$$

$$y_{n+1} = y_n + \frac{1}{2}h\kappa_1 + \frac{1}{2}h\kappa_2 \quad (4c)$$

2. Una soluzione possibile è la seguente.

```
function [T, Y] = rkeuler(N, lambda)

h = 1 / N;
T = 0:h:1;
Y = zeros(1, N+1);
Y(1) = 1;
for n = 1:N
    k1 = lambda * Y(n);
    k2 = lambda * (Y(n) + h/2*k1);
    Y(n+1) = Y(n) + h/2*k1 + h/2*k2;
end
```

3. Come abbiamo visto a lezione, la soluzione esatta del problema test è  $y(t) = \exp(\lambda t)$ . Quindi abbiamo

```
>> [T, Y] = rkeuler(50, -2.5);
>> E50 = max(abs(Y - exp(-2.5*T)))
E50 =
    0.0046
>> [T, Y] = rkeuler(100, -2.5);
>> E100 = max(abs(Y - exp(-2.5*T)))
E100 =
    0.0023
>> [T, Y] = rkeuler(200, -2.5);
>> E200 = max(abs(Y - exp(-2.5*T)))
E200 =
    0.0012
>> E50 / E100, E100 / E200
ans =
    2.0106
ans =
    2.0052
```

I rapporti calcolati sono vicini a 2, il che indica che il metodo ha ordine di convergenza 1.

4. Per calcolare la funzione di stabilità, applichiamo il metodo al problema test, ottenendo

$$\begin{aligned}\kappa_1 &= \lambda y_n \\ \kappa_2 &= \lambda(y_n + \frac{1}{2}h\kappa_1) = \lambda(1 + \frac{1}{2}h\lambda)y_n \\ y_{n+1} &= y_n + \frac{1}{2}h\kappa_1 + \frac{1}{2}h\kappa_2 = y_n(1 + \frac{1}{2}h\lambda + \frac{1}{2}h\lambda(1 + \frac{1}{2}h\lambda)).\end{aligned}$$

Sostituendo  $h\lambda = q$ , otteniamo che  $y_n$  nell'ultimo termine è moltiplicato per

$$R(q) = 1 + \frac{1}{2}q + \frac{1}{2}q + \frac{1}{4}q^2 = 1 + q + \frac{1}{4}q^2 = (1 + \frac{1}{2}q)^2.$$

La regione di stabilità è l'insieme dei punti per cui  $|R(q)| < 1$ ; questo succede quando  $|1 + \frac{1}{2}q| < 1$ , o equivalentemente  $|2 + q| < 2$ . Dobbiamo quindi determinare la regione dei punti  $q$  nel piano complesso che distano meno di 2 dal punto  $-2$ , che è l'interno di un cerchio di centro  $-2$  e raggio 2 (quindi tangente all'origine).

*Nota:* definendo  $y_{n+1/2} = y_n + \frac{1}{2}hf(t_n, y_n)$ , possiamo manipolare le (4) per ottenere

$$y_{n+1/2} = y_n + \frac{1}{2}hf(t_n, y_n), \quad y_{n+1} = y_{n+1/2} + \frac{1}{2}hf(t_n + \frac{1}{2}h, y_{n+1/2}).$$

Questo vuol dire che effettuare un passo del metodo (3) è equivalente a effettuare in sequenza due passi del metodo di Eulero esplicito con lunghezza del passo  $\frac{1}{2}h$  ognuno; quindi il comportamento del metodo (3) è lo stesso del metodo di Eulero esplicito con passo  $\frac{1}{2}h$ , in particolare per quanto riguarda l'ordine di convergenza e la regione di assoluta stabilità.