

Capitolo 1

Risoluzione di Sistemi lineari

1.1 Introduzione al Calcolo Numerico

Il Calcolo Numerico è una disciplina che fa parte di un ampio settore della Matematica Applicata che prende il nome di Analisi Numerica. Si tratta di una materia che è al confine tra la Matematica e l'Informatica poichè cerca di risolvere i consueti problemi matematici utilizzando però una via algoritmica. In pratica i problemi vengono risolti indicando un processo che, in un numero finito di passi, fornisca una soluzione numerica e soprattutto che sia implementabile su un elaboratore. I problemi matematici che saranno affrontati nelle pagine seguenti sono problemi di base: risoluzione di sistemi lineari, approssimazione delle radici di funzioni non lineari, approssimazione di funzioni e dati sperimentali, calcolo di integrali definiti. Tali algoritmi di base molto spesso non sono altro se non un piccolo ingranaggio nella risoluzione di problemi ben più complessi.

1.1.1 Elementi di Algebra Lineare

Sia \mathbb{R} l'insieme dei numeri reali. Generalmente si indica con $\mathbb{R}^{m \times n}$ l'insieme delle *matrici* ad elementi reali aventi m righe ed n colonne. Quindi una matrice è una *tabella a doppia entrata* di numeri reali. Per esempio:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

con $a_{ij} \in \mathbb{R}$ ed $m, n \in \mathbb{N}$. I numeri interi m ed n si dicono *dimensioni della matrice*, ovvero A si dice matrice di *dimensioni* $m \times n$ o di *ordine* $m \times n$. Se $m = n$ allora la matrice A si dice *quadrata* di dimensione n o di ordine n altrimenti si dice *rettangolare*. Se i e j sono numeri interi con $1 \leq i \leq m$ e $1 \leq j \leq n$ allora l'elemento della matrice A di dimensione $m \times n$ che si trova in posizione (i, j) viene indicato con a_{ij} . Gli elementi a_{ij} di una matrice quadrata A di ordine n tali che $i = j$ sono detti *elementi principali* o *diagonali* e formano la cosiddetta diagonale principale di A .

Assegnata una matrice $A \in \mathbb{R}^{m \times n}$ si definisce *matrice trasposta* di A la matrice $B = A^T \in \mathbb{R}^{n \times m}$ tale che

$$b_{ij} = a_{ji}, \quad i = 1, \dots, n, \quad j = 1, \dots, m.$$

Se accade che $A = A^T$ allora la matrice è detta *simmetrica*. Gli elementi di una matrice che si trovano al di sopra della diagonale principale sono detti *sopradiagonali*, mentre quelli che si trovano al di sotto della stessa diagonale principale sono detti *sottodiagonali*. Se una matrice ha tutti gli elementi sopradiagonali e sottodiagonali uguali a zero viene detta *matrice diagonale*. Se invece ha solo gli elementi sopradiagonali nulli allora viene detta *triangolo inferiore*. Se ha gli elementi sottodiagonali nulli allora è detta *triangolo superiore*.

Assegnate due matrici $A, B \in \mathbb{R}^{m \times n}$ si definisce *somma* di A e B , e si denota con $C = A + B$, la matrice $C \in \mathbb{R}^{m \times n}$ i cui elementi sono:

$$c_{ij} = a_{ij} + b_{ij} \quad i = 1, \dots, m, \quad j = 1, \dots, n.$$

In modo analogo si definisce la *differenza* tra matrici, infatti $D = A - B$ è la matrice avente elementi:

$$d_{ij} = a_{ij} - b_{ij} \quad i = 1, \dots, m, \quad j = 1, \dots, n.$$

Se $\alpha \in \mathbb{R}$ ed $A \in \mathbb{R}^{m \times n}$ allora la matrice $C = \alpha A$ è definita da:

$$c_{ij} = \alpha a_{ij}.$$

Se $A \in \mathbb{R}^{m \times p}$ e $B \in \mathbb{R}^{p \times n}$ si definisce *prodotto* di A per B la matrice $C \in \mathbb{R}^{m \times n}$ i cui elementi sono

$$c_{ij} = \sum_{k=1}^p a_{ik} b_{kj} \quad i = 1, \dots, m, \quad j = 1, \dots, n.$$

Si noti che affinché tale prodotto abbia senso è necessario che il numero delle colonne di A coincida con il numero delle righe di B . Quando ciò accade

le matrici si dicono *conformabili*, altrimenti si dicono *non conformabili*. Ad esempio nel nostro caso se $m \neq n$ allora il prodotto BA non ha senso. Ha sempre significato considerare i prodotti AB e BA se A e B sono matrici quadrate dello stesso ordine ($m = n$).

È facile verificare che il prodotto tra matrici gode della proprietà *associativa* ma in generale non di quella *commutativa*. Vale invece la seguente proprietà:

$$(AB)^T = B^T A^T.$$

Esempio 1.1.1.1 Siano A e B le seguenti matrici:

$$A = \begin{pmatrix} 3 & 1 & 0 \\ -1 & 2 & 1 \\ 3 & 1 & 1 \end{pmatrix}; \quad B = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix}.$$

Calcoliamo la matrice $C = AB$. L'elemento c_{ij} è uguale alla somma dei prodotti degli elementi della i -esima riga di A per la j -esima colonna di B .

$$\begin{aligned} c_{11} &= 3 \cdot 2 + 1 \cdot 0 + 0 \cdot 2 = 6 \\ c_{12} &= 3 \cdot 1 + 1 \cdot 1 + 0 \cdot 1 = 4 \\ c_{13} &= 3 \cdot (-1) + 1 \cdot 1 + 0 \cdot 1 = -2 \\ c_{21} &= -1 \cdot 2 + 2 \cdot 0 + 1 \cdot 2 = 0 \\ c_{22} &= -1 \cdot 1 + 2 \cdot 1 + 1 \cdot 1 = 2 \\ c_{23} &= -1 \cdot (-1) + 2 \cdot 1 + 1 \cdot 1 = 4 \\ c_{31} &= 3 \cdot 2 + 1 \cdot 0 + 1 \cdot 2 = 8 \\ c_{32} &= 3 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 = 5 \\ c_{33} &= 3 \cdot (-1) + 1 \cdot 1 + 1 \cdot 1 = -1. \end{aligned}$$

In definitiva

$$C = \begin{pmatrix} 6 & 4 & -2 \\ 0 & 2 & 4 \\ 8 & 5 & -1 \end{pmatrix}.$$

Calcolando il prodotto $D = BA$ si trova invece:

$$D = \begin{pmatrix} 2 & 3 & 0 \\ 2 & 3 & 2 \\ 8 & 5 & 2 \end{pmatrix}$$

da cui risulta evidente che $AB \neq BA$.

Siano $A, B \in \mathbb{R}^{n \times n}$ le seguenti matrici

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

e

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

dove $A_{11}, B_{11} \in \mathbb{R}^{p \times p}$, $A_{12}, B_{12} \in \mathbb{R}^{p \times (n-p)}$, $A_{21}, B_{21} \in \mathbb{R}^{(n-p) \times p}$ e infine $A_{22}, B_{22} \in \mathbb{R}^{(n-p) \times (n-p)}$, con $p < n$, rappresentano a loro volta matrici e non semplici elementi. Si dice cioè che A e B sono state suddivise a blocchi. Il prodotto AB può essere calcolato utilizzando tale decomposizione delle matrici:

$$AB = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}.$$

Si definisce *matrice identità di ordine n* la matrice quadrata diagonale I avente tutti gli elementi principali uguali a 1:

$$I = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

La matrice identità è l'elemento neutro per il prodotto, cioè se $A \in \mathbb{R}^{n \times n}$ si ha

$$AI = IA = A.$$

Definizione 1.1.1.1 Una matrice che si ottiene da I scambiando alcune righe (o colonne) viene detta matrice di permutazione.

Esempio 1.1.1.2 Sia P la seguente matrice di permutazione:

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

che è stata ottenuta da I_3 scambiando la prima riga con la terza. Consideriamo la seguente matrice A

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

e calcoliamo il prodotto PA :

$$PA = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix}.$$

La moltiplicazione a sinistra di una matrice di permutazione per A ha l'effetto di scambiare le righe di A esattamente nello stesso modo con cui erano state scambiate le righe dell'identità per ottenere P . Calcoliamo ora il prodotto AP :

$$AP = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 6 & 5 & 4 \\ 9 & 8 & 7 \end{pmatrix}.$$

Invece la moltiplicazione a destra di una matrice di permutazione per A ha l'effetto di scambiare le colonne di A .

Data una matrice $A \in \mathbb{R}^{m \times n}$, una matrice $B \in \mathbb{R}^{h \times k}$, $0 < h \leq m$, $0 < k \leq n$, è detta *sottomatrice* di A se è ottenuta da A eliminando $m - h$ righe ed $n - k$ colonne. Data una matrice $A \in \mathbb{R}^{m \times n}$, una sottomatrice quadrata B di ordine $k \leq n$ di A è detta *principale* se gli elementi principali di B sono anche gli elementi principali di A . Una sottomatrice B principale di ordine k di A è detta *principale di testa* se è formata dagli elementi a_{ij} , $i, j = 1, \dots, k$.

Definizione 1.1.1.2 Se $A \in \mathbb{R}^{n \times n}$ è una matrice di ordine 1, si definisce determinante di A il numero

$$\det A = a_{11}.$$

Se la matrice A è quadrata di ordine n allora fissata una qualsiasi riga (colonna) di A , diciamo la i -esima (j -esima) allora applicando la cosiddetta regola di Laplace il determinante di A è:

$$\det A = \sum_{j=1}^n a_{ij} (-1)^{i+j} \det A_{ij}$$

dove A_{ij} è la matrice che si ottiene da A cancellando la i -esima riga e la j -esima colonna.

Il determinante è pure uguale a

$$\det A = \sum_{i=1}^n a_{ij} (-1)^{i+j} \det A_{ij},$$

cioè il determinante è indipendente dall'indice di riga (o di colonna) fissato. Se A è la matrice di ordine 2

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

allora

$$\det A = a_{11}a_{22} - a_{21}a_{12}.$$

Il determinante ha le seguenti proprietà:

1. Se A è una matrice triangolare o diagonale allora

$$\det A = \prod_{i=1}^n a_{ii};$$

2. $\det I = 1$;
3. $\det A^T = \det A$;
4. $\det AB = \det A \det B$ (Regola di Binet);
5. se $\alpha \in \mathbb{R}$ allora $\det \alpha A = \alpha^n \det A$.
6. $\det A = 0$ se una riga (o una colonna) è combinazione lineare di due (o più) righe (o colonne) di A .
7. Se A è una matrice triangolare a blocchi

$$A = \begin{pmatrix} B & C \\ O & D \end{pmatrix}$$

con B e D matrici quadrate, allora

$$\det A = \det B \det D. \tag{1.1}$$

Una matrice A di ordine n si dice *non singolare* se il suo determinante è diverso da zero, in caso contrario viene detta *singolare*. Si definisce *inversa di* A la matrice A^{-1} tale che:

$$AA^{-1} = A^{-1}A = I$$

Per quello che riguarda il determinante della matrice inversa vale la seguente proprietà:

$$\det A^{-1} = \frac{1}{\det A}.$$

Vettori

Se $A \in \mathbb{R}^{m \times 1}$ (o $A \in \mathbb{R}^{1 \times m}$), la matrice si riduce ad una sola colonna (o una sola riga) e viene detta *vettore colonna* (o *riga*) *ad m elementi o componenti*. Solitamente il termine vettore viene associato a vettori colonna e l'insieme dei vettori ad m componenti viene indicato con \mathbb{R}^m . Per le operazioni tra vettori valgono le stesse regole viste per le matrici, cioè la somma e la differenza sono possibili tra vettori dello stesso tipo e con lo stesso numero di componenti. Se \mathbf{x} è un vettore colonna di m elementi allora \mathbf{x}^T è un vettore riga sempre di m elementi. Se $A \in \mathbb{R}^{m \times n}$ e $\mathbf{x} \in \mathbb{R}^n$ è possibile definire il prodotto matrice per vettore nel seguente modo:

$$\mathbf{y} = A\mathbf{x}, \quad y_i = \sum_{j=1}^n a_{ij}x_j, \quad i = 1, \dots, m$$

quindi $\mathbf{y} \in \mathbb{R}^m$. Non è possibile effettuare il prodotto $A\mathbf{x}^T$ perchè le dimensioni non sono compatibili.

Esempio 1.1.1.3 *Sia*

$$A = \begin{pmatrix} 5 & 1 & 0 \\ -1 & 1 & 2 \\ 5 & -5 & 1 \end{pmatrix}$$

e sia \mathbf{x} il vettore

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Calcoliamo il vettore prodotto $\mathbf{y} = A\mathbf{x}$:

$$\begin{aligned} y_1 &= 5 \cdot 1 + 1 \cdot 2 + 0 \cdot 3 = 7 \\ y_2 &= -1 \cdot 1 + 1 \cdot 2 + 2 \cdot 3 = 7 \\ y_3 &= 5 \cdot 1 - 5 \cdot 2 + 1 \cdot 3 = -2. \end{aligned}$$

Tra vettori sono consentite le seguenti operazioni:

1. *prodotto interno;*
2. *prodotto esterno.*

Il prodotto interno (o scalare), che viene spesso indicato come (\cdot, \cdot) , è definito nel seguente modo: siano $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, allora

$$(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y} = \sum_{i=1}^n x_i y_i = \alpha$$

e il risultato è un numero reale. Il prodotto scalare soddisfa le seguenti proprietà:

1. $\mathbf{x}^T \mathbf{x} \geq 0$ per ogni $\mathbf{x} \in \mathbb{R}^n$ e $(\mathbf{x}, \mathbf{x}) = 0$ se e solo se $\mathbf{x} = 0$;
2. $\mathbf{x}^T \mathbf{y} = \mathbf{y}^T \mathbf{x}$ per ogni $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$;
3. $(\alpha \mathbf{x})^T \mathbf{y} = \alpha(\mathbf{x}^T \mathbf{y})$ per ogni $\alpha \in \mathbb{R}$ e per ogni $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$;
4. $(\mathbf{x} + \mathbf{y})^T \mathbf{z} = \mathbf{x}^T \mathbf{z} + \mathbf{y}^T \mathbf{z}$ per ogni $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$.
5. se $\mathbf{x}^T \mathbf{y} = 0$ allora i due vettori si dicono *ortogonali*.

Se $\mathbf{x} \in \mathbb{R}^n$ e $\mathbf{y} \in \mathbb{R}^m$ allora il prodotto esterno viene definito nel seguente modo:

$$A = \mathbf{x}\mathbf{y}^T$$

e il risultato è una matrice di dimensione $n \times m$ i cui elementi sono:

$$a_{ij} = x_i y_j, \quad i = 1, \dots, n, \quad j = 1, \dots, m.$$

Esempio 1.1.1.4 Siano \mathbf{x} e \mathbf{y} i seguenti vettori:

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

e

$$\mathbf{y} = \begin{pmatrix} -1 \\ -2 \\ 4 \end{pmatrix}.$$

Calcoliamo prima il prodotto interno:

$$\mathbf{x}^T \mathbf{y} = 1 \cdot (-1) + 2 \cdot (-2) + 3 \cdot 4 = 7.$$

Osserviamo che tale operazione gode della proprietà commutativa, poichè $\mathbf{y}^T \mathbf{x} = 7$.

Per quello che riguarda il prodotto esterno, il risultato è la matrice

$$A = \mathbf{x}\mathbf{y}^T = \begin{pmatrix} -1 & -2 & 4 \\ -2 & -4 & 8 \\ -3 & -6 & 12 \end{pmatrix}.$$

Tale prodotto non gode della proprietà commutativa, infatti:

$$B = \mathbf{y}\mathbf{x}^T = \begin{pmatrix} -1 & -2 & -3 \\ -2 & -4 & -6 \\ 4 & 8 & 12 \end{pmatrix}.$$

Infatti $B \neq A$, anche se va osservato che $B = A^T$.

Norme Vettoriali

La funzione $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ si dice *norma* se per ogni vettore $\mathbf{x} \in \mathbb{R}^n$ $\|\mathbf{x}\|$ soddisfa:

1. $\|\mathbf{x}\| \geq 0$ e $\|\mathbf{x}\| = 0$ se e solo se $\mathbf{x} = 0$;
2. $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ per ogni $\alpha \in \mathbb{C}$;
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ per ogni $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ (*disuguaglianza triangolare*).

Tra le norme più utilizzate citiamo

$$\|\mathbf{x}\|_1 = \sum_{j=1}^n |x_j| \quad \text{norma 1}$$

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{j=1}^n |x_j|^2} \quad \text{norma 2 o norma euclidea}$$

$$\|\mathbf{x}\|_\infty = \max_{1 \leq j \leq n} |x_j| \quad \text{norma infinito.}$$

Norme su Matrici

Definizione 1.1.1.3 Una funzione $\|\cdot\| : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ tale che per ogni matrice $A \in \mathbb{R}^{n \times n}$, $\|A\|$ soddisfa:

1. $\|A\| \geq 0$ e $\|A\| = 0$ se e solo se $A = 0$;
2. $\|\alpha A\| = |\alpha|\|A\|$ per ogni $\alpha \in \mathbb{C}$;
3. $\|A + B\| \leq \|A\| + \|B\|$ per ogni $A, B \in \mathbb{C}^{n \times n}$;
4. $\|A \cdot B\| \leq \|A\| \cdot \|B\|$ per ogni $A, B \in \mathbb{C}^{n \times n}$;

si dice *norma di matrice*.

Definizione 1.1.1.4 Si dice che una norma di matrice è compatibile con una norma di vettore se per ogni matrice A e per ogni vettore \mathbf{x} risulta

$$\|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|.$$

Un modo per definire le norme di matrici compatibili con norme di vettori è il seguente. Sia \mathbf{x} un vettore diverso dal vettore nullo, allora considerata la norma del vettore $A\mathbf{x}$, $\|A\mathbf{x}\|$, definiamo come norma di A :

$$\|A\| = \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}. \quad (1.2)$$

$\|A\|$ è detta *norma naturale di A* oppure *norma di A indotta* dalla norma di vettore $\|\mathbf{x}\|$.

La (1.2) può anche scriversi:

$$\|A\| = \sup_{\|\mathbf{y}\|=1} \|A\mathbf{y}\|$$

anzi è possibile dimostrare che

$$\|A\| = \max_{\|\mathbf{y}\|=1} \|A\mathbf{y}\|.$$

Le norme matriciali indotte dalle norme vettoriali definite in precedenza sono

$$\begin{aligned} \|A\|_1 &= \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| & \|A\|_\infty &= \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \\ \|A\|_2 &= \sqrt{\rho(A^T A)}, \end{aligned}$$

dove $\rho(A^T A)$ è l'autovalore di modulo massimo della matrice $A^T A$, di cui ricordiamo la definizione.

Definizione 1.1.1.5 *Se $A \in \mathbb{R}^{n \times n}$ allora il numero λ si dice autovalore di A se esiste un vettore \mathbf{x} tale che:*

$$A\mathbf{x} = \lambda\mathbf{x}. \quad (1.3)$$

In generale anche se A è una matrice reale λ (e anche \mathbf{x}) può essere complesso. Inoltre se λ è un autovalore di A , ogni vettore non nullo \mathbf{x} che soddisfa la relazione (1.3) si dice *autovettore associato* a λ .

1.2 Metodi diretti per sistemi lineari

Assegnata una matrice quadrata A di ordine n , ed un vettore colonna di n elementi, \mathbf{b} , risolvere il sistema avente A come matrice dei coefficienti e \mathbf{b}

calcolabile. Infatti:

$$\begin{cases} x_n = \frac{b_n}{a_{nn}} \\ \\ x_i = \frac{b_i - \sum_{j=i+1}^n a_{ij}x_j}{a_{ii}} \quad i = n-1, \dots, 1. \end{cases} \quad (1.7)$$

Il metodo (1.7) prende il nome di *metodo di sostituzione all'indietro*, poichè il vettore \mathbf{x} viene calcolato partendo dall'ultima componente. Anche per il seguente sistema il vettore soluzione è calcolabile in modo analogo.

$$\begin{array}{rcccccc} a_{11}x_1 & & & & & = b_1 \\ a_{21}x_1 & +a_{22}x_2 & & & & = b_2 \\ \vdots & \vdots & \ddots & & & \vdots \\ a_{i1}x_1 & +a_{i2}x_2 & \dots & +a_{ii}x_i & & = b_i \\ \vdots & \vdots & & & \ddots & \vdots \\ a_{n1}x_1 & +a_{n2}x_2 & \dots & +a_{ni}x_i & \dots & +a_{nn}x_n = b_n \end{array} \quad (1.8)$$

In questo caso la matrice dei coefficienti è *triangolare inferiore* e la soluzione viene calcolata con il *metodo di sostituzione in avanti*:

$$\begin{cases} x_1 = \frac{b_1}{a_{11}} \\ \\ x_i = \frac{b_i - \sum_{j=1}^{i-1} a_{ij}x_j}{a_{ii}} \quad i = 2, \dots, n-1. \end{cases}$$

1.2.2 Metodo di Eliminazione di Gauss

L'idea di base del metodo di Gauss è appunto quella di operare delle opportune trasformazioni sul sistema originale $A\mathbf{x} = \mathbf{b}$, che non costino eccessivamente, in modo da ottenere un sistema equivalente, cioè un sistema che ammetta la stessa soluzione di quello di partenza, ma che sia facile da risolvere, per esempio uno avente come matrice dei coefficienti una matrice triangolare superiore. Prima di descrivere il metodo vediamo un esempio.

Supponiamo che il sistema da risolvere sia:

$$\begin{aligned} 2x_1 + x_2 + x_3 &= -1 \\ 6x_1 + 2x_2 + x_3 &= 1 \\ 4x_1 - 2x_2 + x_3 &= 2 \end{aligned}$$

La soluzione di un sistema lineare non cambia se un'equazione viene sostituita dalla combinazione lineare di due (o più) equazioni dello stesso sistema. L'idea alla base del metodo di Gauss è quella di ottenere un sistema lineare con matrice dei coefficienti triangolare superiore effettuando opportune combinazioni lineari tra le equazioni. Poniamo

$$A^{(1)} = \begin{pmatrix} 2 & 1 & 1 \\ 6 & 2 & 1 \\ 4 & -2 & 1 \end{pmatrix}, \quad \mathbf{b}^{(1)} = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$$

rispettivamente la matrice dei coefficienti e il vettore dei termini noti del sistema di partenza. Cerchiamo ora di determinare un sistema lineare equivalente a quello iniziale ma che abbia gli elementi sottodiagonali della prima colonna uguali a zero. Lasciamo inalterata la prima equazione. Poniamo

$$l_{21} = \frac{a_{21}}{a_{11}} = \frac{6}{2} = 3$$

e moltiplichiamo la prima equazione per l_{21} ottenendo:

$$6x_1 + 3x_2 + 3x_3 = -3$$

La nuova seconda equazione sarà la differenza tra la seconda equazione e la prima moltiplicata per l_{21} :

$$\begin{array}{r} 6x_1 + 2x_2 + x_3 = 1 \\ -6x_1 - 3x_2 - 3x_3 = 3 \\ \hline -x_2 - 2x_3 = 4 \quad \text{[Nuova seconda equazione].} \end{array}$$

Poniamo

$$l_{31} = \frac{a_{31}^{(1)}}{a_{11}^{(1)}} = \frac{4}{2} = 2$$

e moltiplichiamo la prima equazione per l_{31} ottenendo:

$$4x_1 + 2x_2 + 2x_3 = -2$$

La nuova terza equazione sarà la differenza tra la terza equazione e la prima moltiplicata per l_{31} :

$$\begin{array}{rcl} 4x_1 & -2x_2 & +x_3 & = 2 \\ -4x_1 & -2x_2 & -2x_3 & = 2 \\ \hline & -4x_2 & -x_3 & = 4 \quad [\text{Nuova terza equazione}]. \end{array}$$

Al secondo passo il sistema lineare è diventato:

$$\begin{array}{rcl} 2x_1 & +x_2 & +x_3 & = -1 \\ & -x_2 & -2x_3 & = 4 \\ & -4x_2 & -x_3 & = 4 \end{array}$$

La matrice dei coefficienti e il vettore dei termini noti sono diventati:

$$A^{(2)} = \begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & -4 & -1 \end{pmatrix}, \quad \mathbf{b}^{(2)} = \begin{pmatrix} -1 \\ 4 \\ 4 \end{pmatrix}.$$

Cerchiamo ora di azzerare gli elementi sottodiagonali della seconda colonna. Lasciamo inalterata le prime due equazioni del sistema. Poniamo

$$l_{32} = \frac{a_{32}^{(2)}}{a_{22}^{(2)}} = \frac{-4}{-1} = 4$$

e moltiplichiamo la seconda equazione per l_{32} ottenendo:

$$-4x_2 - 8x_3 = 16$$

La nuova terza equazione sarà la differenza tra la terza equazione e la seconda appena modificata

$$\begin{array}{rcl} -4x_2 & -x_3 & = 4 \\ 4x_2 & +8x_3 & = -16 \\ \hline & 7x_3 & = -12 \quad [\text{Nuova terza equazione}]. \end{array}$$

Abbiamo ottenuto un sistema triangolare superiore:

$$\begin{array}{rcl} 2x_1 & +x_2 & +x_3 & = -1 \\ & -x_2 & -2x_3 & = 4 \\ & & 7x_3 & = -12. \end{array}$$

La matrice dei coefficienti e il vettore dei termini noti sono diventati:

$$A^{(3)} = \begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 7 \end{pmatrix}, \quad \mathbf{b}^{(3)} = \begin{pmatrix} -1 \\ 4 \\ -12 \end{pmatrix}.$$

Vediamo ora di calcolare le formule che consentano di calcolare gli elementi della matrice dei coefficienti e del vettore dei termini noti ad ogni passo del metodo di Gauss. Abbiamo detto che $A^{(1)}$ e $\mathbf{b}^{(1)}$ sono assegnati inizialmente, ipotizziamo per il momento che $a_{11}^{(1)} \neq 0$. Calcoliamo ora gli stessi dati al passo 2 tenendo presente che:

1. La prima equazione del sistema resta invariata;
2. Gli elementi sottodiagonali della prima colonna di $A^{(2)}$ sono nulli;
3. La i -esima riga del sistema ($i \geq 2$) è ottenuta sottraendo dalla medesima riga la prima moltiplicata per $a_{i1}^{(1)}/a_{11}^{(1)}$.

Fissiamo quindi una riga i , $i \geq 2$, e calcoliamo gli elementi $a_{ij}^{(2)}$:

$$\begin{array}{cccccccccc} a_{i1}^{(1)} & a_{i2}^{(1)} & a_{i3}^{(1)} & \dots & a_{ij}^{(1)} & \dots & a_{in}^{(1)} & b_i^{(1)} & - & \\ \frac{a_{i1}^{(1)}}{a_{11}^{(1)}} \times & a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \dots & a_{1j}^{(1)} & \dots & a_{1n}^{(1)} & b_1^{(1)} & = & \\ \hline & 0 & a_{i2}^{(2)} & a_{i3}^{(2)} & \dots & a_{ij}^{(2)} & \dots & a_{in}^{(2)} & b_i^{(2)} & & \end{array}$$

dove

$$a_{ij}^{(2)} = a_{ij}^{(1)} - \frac{a_{i1}^{(1)}}{a_{11}^{(1)}} a_{1j}^{(1)}, \quad i, j = 2, \dots, n$$

e

$$b_i^{(2)} = b_i^{(1)} - \frac{a_{i1}^{(1)}}{a_{11}^{(1)}} b_1^{(1)}, \quad i = 2, \dots, n.$$

Se ipotizziamo che $a_{22}^{(2)} \neq 0$ possiamo calcolare gli elementi del sistema al passo 3 tenendo presente che:

1. La prime 2 equazioni del sistema restano invariate;
2. Gli elementi sottodiagonali della prime 2 colonna di $A^{(3)}$ sono nulli;

3. La i -esima riga del sistema ($i \geq 3$) è ottenuta sottraendo dalla medesima riga la seconda moltiplicata per $a_{i2}^{(2)}/a_{22}^{(2)}$.

Fissiamo quindi una riga i , $i \geq 3$, e calcoliamo gli elementi $a_{ij}^{(3)}$:

$$\begin{array}{cccccccccc}
 0 & a_{i2}^{(2)} & a_{i3}^{(2)} & \dots & a_{ij}^{(2)} & \dots & a_{in}^{(2)} & b_i^{(2)} & - \\
 \frac{a_{i2}^{(2)}}{a_{22}^{(2)}} \times & 0 & a_{22}^{(2)} & a_{23}^{(2)} & \dots & a_{2j}^{(2)} & \dots & a_{2n}^{(2)} & b_2^{(2)} & = \\
 \hline
 0 & 0 & a_{i3}^{(3)} & \dots & a_{ij}^{(3)} & \dots & a_{in}^{(3)} & b_i^{(3)} & &
 \end{array}$$

dove

$$a_{ij}^{(3)} = a_{ij}^{(2)} - \frac{a_{i2}^{(2)}}{a_{22}^{(2)}} a_{2j}^{(2)}, \quad i, j = 3, \dots, n$$

e

$$b_i^{(3)} = b_i^{(2)} - \frac{a_{i2}^{(2)}}{a_{22}^{(2)}} b_2^{(2)}, \quad i = 3, \dots, n.$$

Avendo ricavato esplicitamente le formule per i primi due passi del metodo di Gauss è semplice ricavare quelle per un generico passo k . La matrice $A^{(k)}$ ha gli elementi sottodiagonali delle prime $k - 1$ colonne uguali a zero, e, supposto $a_{kk}^{(k)} \neq 0$, gli elementi di $A^{(k+1)}$ e di $\mathbf{b}^{(k+1)}$ sono quindi:

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} a_{kj}^{(k)}, \quad i, j = k + 1, \dots, n \quad (1.9)$$

e

$$b_i^{(k+1)} = b_i^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} b_k^{(k)}, \quad i = k + 1, \dots, n. \quad (1.10)$$

Il valore di k varia da 1 (matrice dei coefficienti e vettori dei termini noti iniziali) fino a $n - 1$, infatti la matrice $A^{(n)}$ avrà gli elementi sottodiagonali delle prime $n - 1$ colonne uguali a zero.

Tutto il discorso fatto finora va bene se gli elementi $a_{kk}^{(k)}$ sono diversi da zero per ogni k , prima di affrontare come modificare il metodo di Gauss qualora tale situazione non si verifichi consideriamo una formulazione alternativa dello stesso metodo. Nel descrivere il metodo di Gauss si è fatta l'implicita ipotesi (vedere formule (1.9) e (1.10)) che i cosiddetti *elementi pivotali* $a_{kk}^{(k)}$

siano non nulli per $k = 1, 2, \dots, n - 1$. Tale ipotesi non è tuttavia molto limitante in quanto la non singolarità di A garantisce che nella k -esima colonna ci sia sicuramente un elemento al di sotto della diagonale diverso da zero e quindi è possibile, con un opportuno scambio di righe in $A^{(k)}$, trovare un elemento pivotale non nullo. Infatti scambiare due righe in $A^{(k)}$ significa sostanzialmente scambiare due equazioni nel sistema $A^{(k)}\mathbf{x} = \mathbf{b}^{(k)}$ e ciò non altera la natura del sistema stesso. Vediamo di provare questa proprietà. Supponiamo che la matrice A sia non singolare, consideriamo la matrice $A^{(k)}$ e supponiamo $a_{kk}^{(k)} = 0$. Se tutti gli elementi $a_{ik}^{(k)}$ per $i = k + 1, \dots, n$, fossero nulli allora la matrice $A^{(k)}$ avrebbe la seguente struttura:

$$A^{(k)} = \begin{pmatrix} a_{11}^{(1)} & \cdots & a_{1,k-1}^{(1)} & a_{1k}^{(1)} & a_{1,k+1}^{(1)} & \cdots & a_{1n}^{(1)} \\ & \ddots & \vdots & \vdots & \vdots & & \vdots \\ & & a_{k-1,k-1}^{(k-1)} & a_{k-1,k}^{(k-1)} & a_{k-1,k+1}^{(k-1)} & \cdots & a_{k-1,n}^{(k-1)} \\ & & & 0 & a_{k,k+1}^{(k)} & & a_{kn}^{(k)} \\ & 0 & & \vdots & \vdots & & \vdots \\ & & & 0 & a_{n,k+1}^{(k)} & \cdots & a_{nn}^{(k)} \end{pmatrix}$$

Partizioniamo $A^{(k)}$ nel seguente modo

$$A^{(k)} = \begin{pmatrix} A_{11}^{(k)} & A_{12}^{(k)} \\ 0 & A_{22}^{(k)} \end{pmatrix}$$

con $A_{11}^{(k)}$ matrice di ordine $k - 1$. Poichè

$$\det A^{(k)} = \det A_{11}^{(k)} \det A_{22}^{(k)}$$

si ha $\det A^{(k)} = 0$ (infatti la matrice $A_{22}^{(k)}$ è singolare), ma questo contrasta con la proprietà vista in precedenza:

$$\det A^{(k)} = \det A \neq 0$$

quindi possiamo concludere che se $a_{kk}^{(k)} = 0$ e $\det A \neq 0$ deve necessariamente esistere un elemento $a_{ik}^{(k)} \neq 0$, con $i \in \{k + 1, k + 2, \dots, n\}$. L'esistenza di questo elemento diverso da zero garantisce che il metodo di Gauss possa essere applicato anche se $a_{kk}^{(k)} = 0$. Infatti in una simile eventualità si può applicare la cosiddetta strategia di *Pivoting parziale*:

1. si determina l'elemento $a_{rk}^{(k)}$ tale che

$$|a_{rk}^{(k)}| = \max_{k \leq i \leq n} |a_{ik}^{(k)}|;$$

2. si effettua lo scambio tra la r -esima e la k -esima equazione.

Capitolo 2

Interpolazione di Funzioni

2.1 Introduzione

Nel campo del Calcolo Numerico si possono incontrare diversi casi nei quali è richiesta l'approssimazione di una funzione (o di una grandezza incognita):
1) non è nota l'espressione analitica della funzione $f(x)$ ma si conosce il valore che assume in un insieme finito di punti x_1, x_2, \dots, x_n . Si potrebbe pensare anche che tali valori siano delle misure di una grandezza fisica incognita valutate in differenti istanti di tempo.

2) Si conosce l'espressione analitica della funzione $f(x)$ ma è così complicata dal punto di vista computazionale che è più conveniente cercare un'espressione semplice partendo dal valore che essa assume in un insieme finito di punti.

In questo capitolo analizzeremo un particolare tipo di approssimazione di funzioni cioè la cosiddetta interpolazione che richiede che la funzione approssimante assume in determinate ascisse esattamente lo stesso valore di $f(x)$.

In entrambi i casi appena citati è noto, date certe informazioni supplementari, che la funzione approssimante va ricercata della forma:

$$f(x) \simeq g(x; a_0, a_1, \dots, a_n). \quad (2.1)$$

Se i parametri a_0, a_1, \dots, a_n sono definiti dalla condizione di coincidenza di f e g nei punti x_0, x_1, \dots, x_n , allora tale procedimento di approssimazione si chiama appunto *Interpolazione*. Invece se $x \notin [\min_i x_i, \max_i x_i]$ allora si parla di *Estrapolazione*.

Tra i procedimenti di interpolazione il più usato è quello in cui si cerca la

funzione g in (2.1) nella forma

$$g(x; a_0, a_1, \dots, a_n) = \sum_{i=0}^n a_i \Phi_i(x)$$

dove $\Phi_i(x)$, per $i = 0, \dots, n$, sono funzioni fissate e i valori di a_i , $i = 0, \dots, n$, sono determinati in base alle condizioni di coincidenza di f con la funzione approssimante nei punti di interpolazione (detti anche *nodì*), x_j , cioè si pone

$$f(x_j) = \sum_{i=0}^n a_i \Phi_i(x_j) \quad j = 0, \dots, n.$$

Vedremo nel successivo paragrafo di dare una risposta al nostro problema nel caso in cui si cerchino le funzioni $\Phi_i(x)$ di tipo polinomiale.

2.2 Il Polinomio Interpolante di Lagrange

Al fine di dare una forma esplicita al polinomio interpolante, scriviamo il candidato polinomio nella seguente forma:

$$L_n(x) = \sum_{k=0}^n l_{nk}(x) f(x_k) \quad (2.2)$$

dove gli $l_{nk}(x)$ sono per il momento generici polinomi di grado n . Imponendo le condizioni di interpolazione

$$L_n(x_i) = f(x_i) \quad i = 0, \dots, n$$

deve essere, per ogni i :

$$L_n(x_i) = \sum_{k=0}^n l_{nk}(x_i) f(x_k) = f(x_i)$$

ed è evidente che se

$$l_{nk}(x_i) = \begin{cases} 0 & \text{se } k \neq i \\ 1 & \text{se } k = i \end{cases} \quad (2.3)$$

allora esse sono soddisfatte. In particolare la prima condizione di (2.3) indica che $l_{nk}(x)$ si annulla negli n nodi $x_0, x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n$ e quindi deve avere la seguente struttura:

$$l_{nk}(x) = c_k \prod_{i=0, i \neq k}^n (x - x_i)$$

mentre imponendo la seconda condizione di (2.3)

$$l_{nk}(x_k) = c_k \prod_{i=0, i \neq k}^n (x_k - x_i) = 1$$

si trova immediatamente:

$$c_k = \frac{1}{\prod_{i=0, i \neq k}^n (x_k - x_i)}.$$

In definitiva il polinomio interpolante ha la seguente forma:

$$L_n(x) = \sum_{k=0}^n \left(\prod_{i=0, i \neq k}^n \frac{x - x_i}{x_k - x_i} \right) f(x_k). \quad (2.4)$$

Il polinomio (2.4) prende il nome di *Polinomio di Lagrange* mentre i polinomi:

$$l_{nk}(x) = \prod_{i=0, i \neq k}^n \frac{x - x_i}{x_k - x_i}; \quad k = 0, 1, \dots, n$$

si chiamano *Polinomi Fondamentali di Lagrange*.

2.2.1 Il Resto del Polinomio di Lagrange

Assumiamo che la funzione interpolata $f(x)$ sia di classe $\mathcal{C}^{n+1}([a, b])$ e valutiamo l'errore che si commette nel sostituire $f(x)$ con $L_n(x)$ in un punto $x \neq x_i$. Supponiamo che l'intervallo $[a, b]$ sia tale da contenere sia i nodi x_i che l'ulteriore punto x . Sia dunque

$$e(x) = f(x) - L_n(x)$$

l'errore (o resto) commesso nell'interpolazione della funzione $f(x)$. Poichè

$$e(x_i) = f(x_i) - L_n(x_i) = 0 \quad i = 0, \dots, n$$

è facile congetturare per $e(x)$ la seguente espressione:

$$e(x) = c(x)\omega_{n+1}(x)$$

dove

$$\omega_{n+1}(x) = \prod_{i=0}^n (x - x_i)$$

è il cosiddetto *polinomio nodale* mentre $c(x)$ è una funzione da determinare. Definiamo ora la funzione

$$\Phi(t; x) = f(t) - L_n(t) - c(x)\omega_{n+1}(t)$$

dove t è una variabile ed x è un valore fissato. Calcoliamo la funzione $\Phi(t; x)$ nei nodi x_i :

$$\Phi(x_i; x) = f(x_i) - L_n(x_i) - c(x)\omega_{n+1}(x_i) = 0$$

e anche nel punto x :

$$\Phi(x; x) = f(x) - L_n(x) - c(x)\omega_{n+1}(x) = e(x) - c(x)\omega_{n+1}(x) = 0$$

pertanto la funzione $\Phi(t; x)$ (che è derivabile con continuità $n+1$ volte poichè $f(x)$ è di classe \mathcal{C}^{n+1}) ammette almeno $n+2$ zeri distinti. Applicando il teorema di Rolle segue che $\Phi'(t; x)$ ammette almeno $n+1$ zeri distinti. Riapplicando lo stesso teorema segue che $\Phi''(t; x)$ ammette almeno n zeri distinti. Così proseguendo segue che

$$\exists \xi_x \in [a, b] \ni' \Phi^{(n+1)}(\xi_x; x) = 0.$$

Calcoliamo ora la derivata di ordine $n+1$ della funzione $\Phi(t; x)$, osservando innanzitutto che la derivata di tale ordine del polinomio $L_n(x)$ è identicamente nulla. Pertanto

$$\Phi^{(n+1)}(t; x) = f^{(n+1)}(t) - c(x) \frac{d^{n+1}}{dt^{n+1}} \omega_{n+1}(t).$$

Calcoliamo la derivata di ordine $n+1$ del polinomio nodale. Osserviamo innanzitutto che

$$\omega_{n+1}(t) = \prod_{i=0}^n (t - x_i) = t^{n+1} + p_n(t)$$

dove $p_n(t)$ è un polinomio di grado al più n . Quindi

$$\frac{d^{n+1}}{dt^{n+1}}\omega_{n+1}(t) = \frac{d^{n+1}}{dt^{n+1}}t^{n+1}.$$

Poichè

$$\frac{d}{dt}t^{n+1} = (n+1)t^n$$

e

$$\frac{d^2}{dt^2}t^{n+1} = (n+1)nt^{n-1}$$

è facile dedurre che

$$\frac{d^{n+1}}{dt^{n+1}}t^{n+1} = \frac{d^{n+1}}{dt^{n+1}}\omega_{n+1}(t) = (n+1)!.$$

Pertanto

$$\Phi^{(n+1)}(t; x) = f^{(n+1)}(t) - c(x)(n+1)!$$

e quindi

$$\Phi^{(n+1)}(\xi_x; x) = f^{(n+1)}(\xi_x) - c(x)(n+1)! = 0$$

cioè

$$c(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!}$$

e in definitiva

$$e(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!}\omega_{n+1}(x). \quad (2.5)$$

Esempio 2.2.1.1 *Supponiamo di voler calcolare il polinomio interpolante di Lagrange passante per i punti $(-1, -1)$, $(0, 1)$, $(1, -1)$, $(3, 2)$ e $(5, 6)$. Il grado di tale polinomio è 4, quindi definiamo i nodi*

$$x_0 = -1, \quad x_1 = 0, \quad x_2 = 1, \quad x_3 = 3, \quad x_4 = 5,$$

cui corrispondono le ordinate che indichiamo con y_i , $i = 0, \dots, 4$:

$$y_0 = -1, \quad y_1 = 1, \quad y_2 = -1, \quad y_3 = 2, \quad y_4 = 6.$$

Scriviamo ora l'espressione del polinomio $L_4(x)$:

$$L_4(x) = l_{4,0}(x)y_0 + l_{4,1}(x)y_1 + l_{4,2}(x)y_2 + l_{4,3}(x)y_3 + l_{4,4}(x)y_4 \quad (2.6)$$

e calcoliamo i 5 polinomi fondamentali di Lagrange:

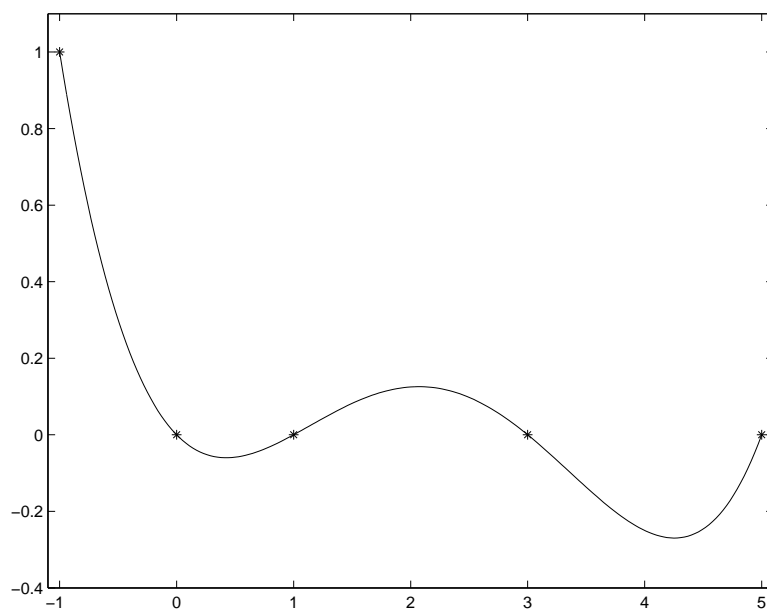
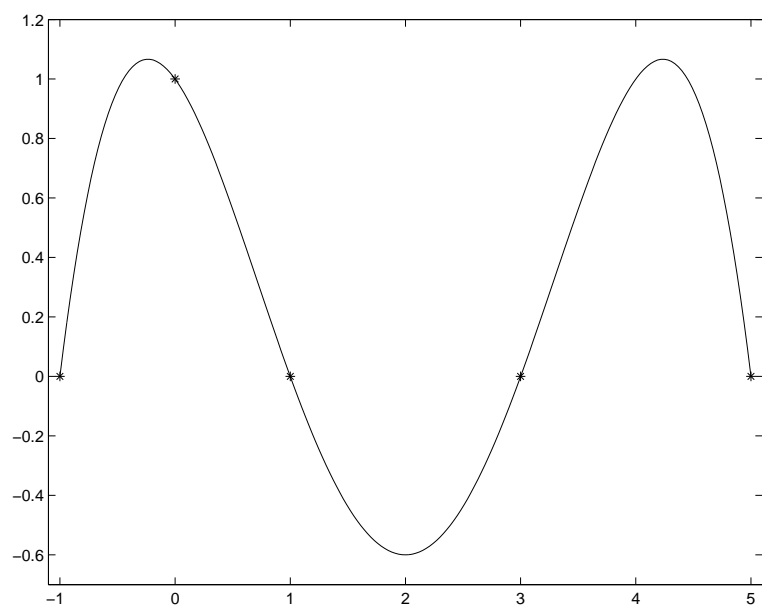
$$\begin{aligned}
 l_{4,0}(x) &= \frac{(x-0)(x-1)(x-3)(x-5)}{(-1-0)(-1-1)(-1-3)(-1-5)} = \\
 &= \frac{1}{48} x(x-1)(x-3)(x-5) \\
 l_{4,1}(x) &= \frac{(x+1)(x-1)(x-3)(x-5)}{(0+1)(0-1)(0-3)(0-5)} = \\
 &= -\frac{1}{15}(x+1)(x-1)(x-3)(x-5) \\
 l_{4,2}(x) &= \frac{(x+1)(x-0)(x-3)(x-5)}{(1+1)(1-0)(1-3)(1-5)} = \\
 &= \frac{1}{16}x(x+1)(x-3)(x-5) \\
 l_{4,3}(x) &= \frac{(x+1)(x-0)(x-1)(x-5)}{(3+1)(3-0)(3-1)(3-5)} = \\
 &= -\frac{1}{48}x(x+1)(x-1)(x-5) \\
 l_{4,4}(x) &= \frac{(x+1)(x-0)(x-1)(x-3)}{(5+1)(5-0)(5-1)(5-3)} = \\
 &= \frac{1}{240}x(x+1)(x-1)(x-3)
 \end{aligned}$$

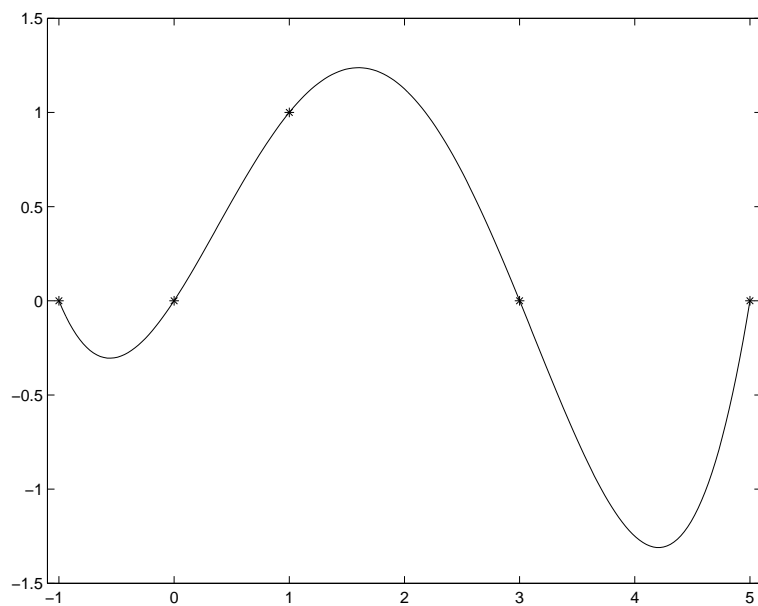
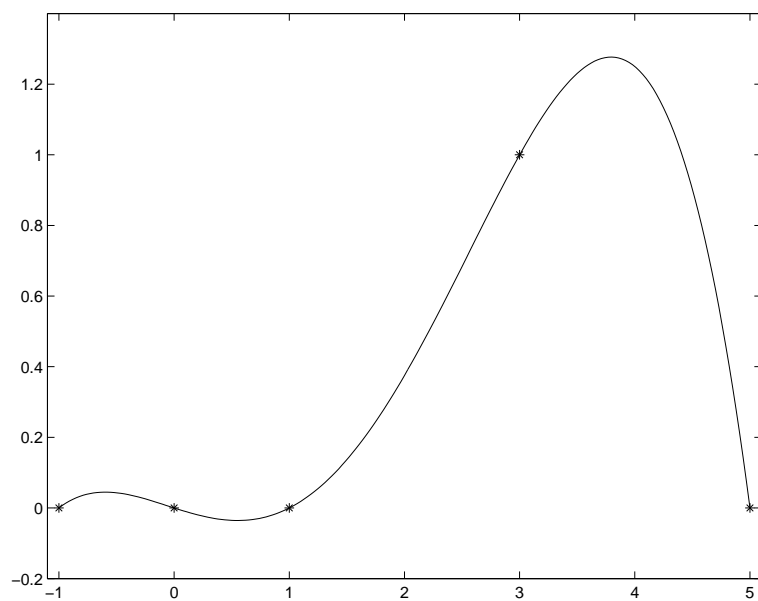
Sostituendo in (2.6) il valore della funzione nei nodi si ottiene l'espressione finale del polinomio interpolante:

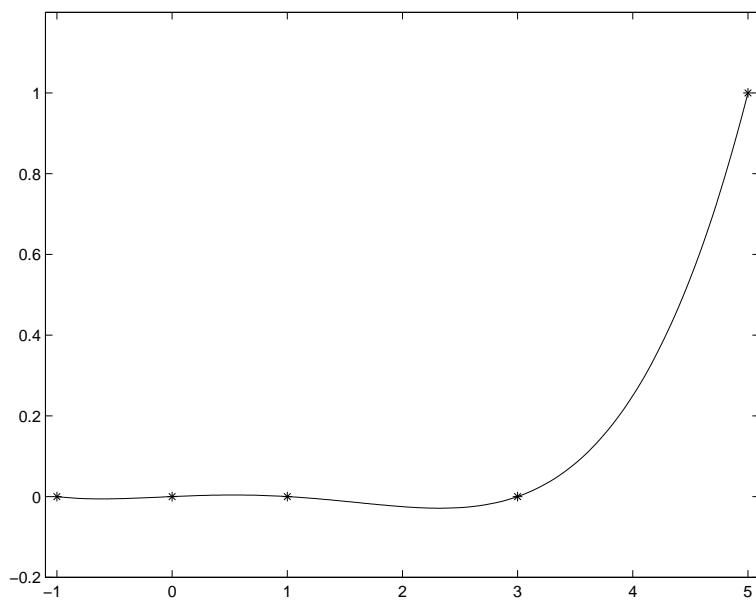
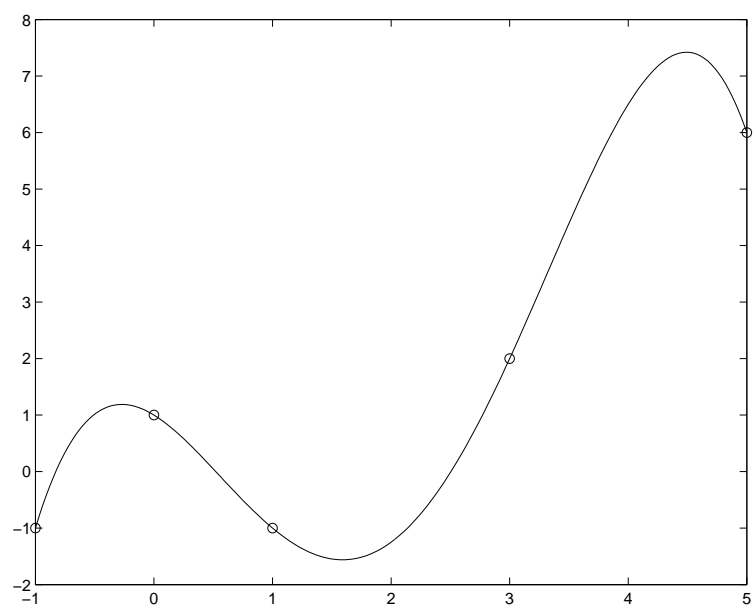
$$L_4(x) = -l_{4,0}(x) + l_{4,1}(x) - l_{4,2}(x) + 2l_{4,3}(x) + 6l_{4,4}(x).$$

Se vogliamo calcolare il valore approssimato della funzione $f(x)$ in un'ascissa diversa dai nodi, per esempio $x = 2$ allora dobbiamo calcolare il valore del polinomio interpolante $L_4(2)$.

Nelle figure 2.1-2.5 sono riportati i grafici dei cinque polinomi fondamentali di Lagrange: gli asterischi evidenziano il valore assunto da tali polinomi nei nodi di interpolazione. Nella figura 2.6 è tracciato il grafico del polinomio interpolante di Lagrange, i cerchi evidenziano ancora una volta i punti di interpolazione.

Figura 2.1: Grafico del polinomio $l_{40}(x)$.Figura 2.2: Grafico del polinomio $l_{41}(x)$.

Figura 2.3: Grafico del polinomio $l_{42}(x)$.Figura 2.4: Grafico del polinomio $l_{43}(x)$.

Figura 2.5: Grafico del polinomio $l_{44}(x)$.Figura 2.6: Grafico del polinomio interpolante di Lagrange $L_4(x)$.

Il polinomio interpolante di Lagrange ha il pregio di fornire un'espressione esplicita del polinomio stesso ma ha un grosso svantaggio. Infatti se, una volta calcolato $L_n(x)$, supponiamo di venire a conoscenza di un ulteriore valore $(x_{n+1}, f(x_{n+1}))$ si devono ripetere tutti i calcoli. Questa motivazione ci spinge a cercare una diversa rappresentazione del polinomio interpolante. Vedremo che una soluzione a questo problema ci verrà fornita dal polinomio interpolante di Newton. Prima di affrontare tale argomento dimostriamo l'unicità del polinomio interpolante di grado n e successivamente definiamo le cosiddette differenze divise.

Supponiamo per assurdo che esistano due diversi polinomi $L_n(x)$ e $P_n(x)$, entrambi di grado n ed entrambi interpolanti la funzione $f(x)$ nei nodi x_0, \dots, x_n ,

$$L_n(x_i) = P_n(x_i) = f(x_i) \quad i = 0, \dots, n.$$

Definiamo il polinomio differenza, che sarà di grado al più n :

$$D_n(x) = L_n(x) - P_n(x)$$

e calcoliamo il suo valore nei nodi:

$$D_n(x_i) = L_n(x_i) - P_n(x_i) = f(x_i) - f(x_i) = 0.$$

Il polinomio $D_n(x)$ ammette quindi $n + 1$ zeri ma questo contrasta con il teorema fondamentale dell'algebra che stabilisce che un polinomio di grado n ammette al più n zeri. Pertanto

$$D_n(x) \equiv 0.$$

Capitolo 3

Equazioni non Lineari

3.1 Introduzione

Le radici di un'equazione non lineare $f(x) = 0$ non possono, in generale, essere espresse esplicitamente e anche se ciò è possibile spesso l'espressione si presenta in forma talmente complicata da essere praticamente inutilizzabile. Di conseguenza per poter risolvere equazioni di questo tipo siamo obbligati ad utilizzare metodi numerici che sono, in generale, di tipo iterativo, cioè partendo da una (o più) approssimazioni della radice, producono una successione x_0, x_1, x_2, \dots , convergente alla radice. Per alcuni di questi metodi per ottenere la convergenza è sufficiente conoscere un intervallo $[a, b]$ che contiene la soluzione, altri metodi richiedono invece la conoscenza di una buona approssimazione iniziale. Talvolta è opportuno utilizzare in maniera combinata due metodi, uno del primo tipo e uno del secondo.

Prima di analizzare alcuni metodi per l'approssimazione delle radici dell'equazione $f(x) = 0$ diamo la definizione di molteplicità di una radice.

Definizione 3.1.1.1 *Sia $f \in C^r([a, b])$ per un intero $r > 0$. Una radice α di $f(x)$ si dice di molteplicità r se*

$$\lim_{x \rightarrow \alpha} \frac{f(x)}{(x - \alpha)^r} = c \quad c \neq 0, c \neq \pm\infty. \quad (3.1)$$

Una definizione alternativa di molteplicità è che se la funzione $f(x)$ si annulla in α insieme a tutte le derivate fino all'ordine $r - 1$ allora la radice ha molteplicità r .

3.2 Metodo di Bisezione

Sia $f : [a, b] \rightarrow \mathbb{R}$, $f \in \mathcal{C}([a, b])$, e sia $f(a)f(b) < 0$. Sotto tali ipotesi esiste sicuramente almeno un punto nell'intervallo $[a, b]$ in cui f si annulla. L'idea alla base del *Metodo di Bisezione* consiste nel costruire una successione di intervalli $\{I_k\}_{k=0}^{\infty}$, con $I_0 = [a_0, b_0] \equiv [a, b]$, tali che:

1. $I_{k+1} \subset I_k$;
2. $\alpha \in I_k, \forall k \geq 0$;
3. l'ampiezza di I_k tende a zero per $k \rightarrow +\infty$.

La successione degli I_k viene costruita nel seguente modo. Innanzitutto si pone

$$I_0 = [a_0, b_0] = [a, b]$$

e si calcola il punto medio

$$c_1 = \frac{a_0 + b_0}{2}.$$

Se $f(c_1) = 0$ allora $\alpha = c_1$, altrimenti si pone:

$$I_1 = [a_1, b_1] \equiv \begin{cases} a_1 = a_0 & b_1 = c_1 & \text{se } f(a_0)f(c_1) < 0 \\ a_1 = c_1 & b_1 = b_0 & \text{se } f(a_0)f(c_1) > 0. \end{cases}$$

Ora, a partire da $I_1 = [a_1, b_1]$, si ripete la stessa procedura. In generale al passo k si calcola

$$c_{k+1} = \frac{a_k + b_k}{2}.$$

Se $f(c_{k+1}) = 0$ allora $\alpha = c_{k+1}$, altrimenti si pone:

$$I_{k+1} = [a_{k+1}, b_{k+1}] \equiv \begin{cases} a_{k+1} = a_k & b_{k+1} = c_k & \text{se } f(a_k)f(c_{k+1}) < 0 \\ a_{k+1} = c_{k+1} & b_{k+1} = b_k & \text{se } f(a_k)f(c_{k+1}) > 0. \end{cases}$$

La successione di intervalli I_k così costruita soddisfa automaticamente le condizioni 1) e 2). Per quanto riguarda la 3) abbiamo:

$$b_k - a_k = \frac{b_{k-1} - a_{k-1}}{2} = \frac{b_0 - a_0}{2^k}$$

e dunque l'ampiezza di I_k tende a zero quando $k \rightarrow +\infty$. La continuità di f garantisce poi la convergenza delle successioni $\{a_k\}_{k=0}^{\infty}$ e $\{b_k\}_{k=0}^{\infty}$ ad una

radice di $f(x)$.

Infatti, per costruzione, $\{b_k\}_{k=0}^{\infty}$ è una successione monotona decrescente, limitata inferiormente e dunque convergente, mentre $\{a_k\}_{k=0}^{\infty}$ è una successione monotona crescente, limitata superiormente e dunque convergente. Allora da:

$$\begin{aligned} \lim_{k \rightarrow +\infty} (b_k - a_k) &= 0 \\ \lim_{k \rightarrow +\infty} a_k &= \xi_1 \quad \lim_{k \rightarrow +\infty} b_k = \xi_2 \end{aligned}$$

segue

$$\xi_1 = \xi_2 = \xi.$$

Ancora poichè $f(a_k)f(b_k) \leq 0$ per ogni k , si ha

$$\lim_{k \rightarrow +\infty} f(a_k)f(b_k) = f\left(\lim_{k \rightarrow +\infty} a_k\right) f\left(\lim_{k \rightarrow +\infty} b_k\right) \leq 0 \Leftrightarrow f^2(\xi) \leq 0$$

e quest'ultima ha senso solo se $f(\xi) = 0$. Quindi in definitiva $\alpha = \xi$.

Generalmente costruendo le successioni $\{a_k\}$ e $\{b_k\}$ accade che la condizione $f(c_k) = 0$, per un certo valore k , non si verifica mai a causa degli errori di arrotondamento. Quindi è necessario stabilire un opportuno criterio di stop che ci permetta di fermare la procedura quando riteniamo di aver raggiunto una precisione soddisfacente. Per esempio si può imporre:

$$b_k - a_k \leq \varepsilon \tag{3.2}$$

dove ε è una prefissata tolleranza. La (3.2) determina anche un limite per il numero di iterate infatti:

$$\frac{b_0 - a_0}{2^k} \leq \varepsilon \quad \Rightarrow \quad k > \log_2 \left(\frac{b_0 - a_0}{\varepsilon} \right).$$

Poichè $b_k - \alpha \leq b_k - a_k$, il criterio (3.2) garantisce che α è approssimata da c_{k+1} con un errore assoluto minore di ε . Se $0 \notin [a, b]$ si può usare come criterio di stop

$$\frac{b_k - a_k}{\min(|a_k|, |b_k|)} \leq \varepsilon \tag{3.3}$$

che garantisce che α è approssimata da c_{k+1} con un errore relativo minore di ε . Un ulteriore criterio di stop è fornito dal test:

$$|f(c_k)| \leq \varepsilon. \tag{3.4}$$

È comunque buona norma utilizzare due criteri di stop insieme, per esempio (3.2) e (3.4) oppure (3.3) e (3.4).

3.3 Metodi di Iterazione Funzionale

Il metodo di bisezione può essere applicato ad una vastissima classe di funzioni, in quanto per la sua costruzione è richiesta solo la continuità della funzione. Tuttavia ha lo svantaggio di risultare piuttosto lento, infatti ad ogni passo si guadagna in precisione una cifra binaria. Poichè $10^{-1} \simeq 2^{-3.3}$ per passare da t a $t + 1$ cifre decimali significative sono mediamente necessarie 3.3 iterate. Inoltre la velocità di convergenza è indipendente dalla funzione $f(x)$ poichè il metodo utilizza esclusivamente il segno assunto dalla funzione in determinati punti e non il suo valore. Il metodo delle bisezioni può essere comunque utilizzato con profitto per determinare delle buone approssimazioni della radice α che possono essere utilizzate dai metodi iterativi che ci apprestiamo a trattare.

Infatti richiedendo alla funzione $f(x)$ supplementari condizioni di regolarità è possibile individuare una vasta classe di metodi che forniscono le stesse approssimazioni del metodo di bisezione utilizzando però un numero di iterate molto minore. In generale questi metodi sono del tipo:

$$x_{k+1} = g(x_k) \quad k = 0, 1, 2, \dots \quad (3.5)$$

con x_0 assegnato valore iniziale e forniscono un'approssimazione della soluzione dell'equazione

$$x = g(x). \quad (3.6)$$

Ogni punto α tale che $\alpha = g(\alpha)$ si dice *punto fisso* o *punto unito* di g .

Per poter applicare uno schema del tipo (3.5) all'equazione $f(x) = 0$, bisogna prima trasformare questa nella forma (3.6). Ad esempio se $[a, b]$ è l'intervallo di definizione di f ed $h(x)$ è una qualunque funzione tale che $h(x) \neq 0$, per ogni $x \in [a, b]$, si può porre:

$$g(x) = x - \frac{f(x)}{h(x)}. \quad (3.7)$$

Ovviamente ogni punto fisso di g è uno zero di f e viceversa.

Teorema 3.3.1.1 *Sia $g \in \mathcal{C}([a, b])$ e assumiamo che la successione $\{x_k\}$ generata da (3.5) sia contenuta in $[a, b]$. Allora se tale successione converge, il limite è il punto fisso di g .*

Dimostrazione.

$$\alpha = \lim_{k \rightarrow +\infty} x_{k+1} = \lim_{k \rightarrow +\infty} g(x_k) = g\left(\lim_{k \rightarrow +\infty} x_k\right) = g(\alpha). \quad \square$$

Teorema 3.3.1.2 *Sia α punto fisso di g e $g \in \mathcal{C}^1([\alpha - \rho, \alpha + \rho])$, per qualche $\rho > 0$. Scelto x_0 tale che*

$$|x_0 - \alpha| \leq \rho$$

per la successione $\{x_k\}_{k=0}^{\infty}$ generata da (3.5) si ha che se $|g'(x)| < 1$, per $|x - \alpha| \leq \rho$, allora $|x_k - \alpha| \leq \rho$, per ogni k , e la successione $\{x_k\}$ converge a α .

Dimostrazione. Sia

$$\lambda = \max_{|x - \alpha| \leq \rho} |g'(x)| < 1.$$

Proviamo per induzione che tutti gli elementi della successione $\{x_k\}$ sono contenuti nell'intervallo di centro α e ampiezza 2ρ . Per $k = 0$ si ha banalmente $x_0 \in [\alpha - \rho, \alpha + \rho]$. Assumiamo che $|x_k - \alpha| \leq \rho$ e dimostriamolo per $k + 1$.

$$|x_{k+1} - \alpha| = |g(x_k) - g(\alpha)| = |g'(\xi_k)| |x_k - \alpha|$$

dove $|\xi_k - \alpha| < |x_k - \alpha| \leq \rho$. Pertanto

$$|x_{k+1} - \alpha| \leq \lambda |x_k - \alpha| < |x_k - \alpha| \leq \rho.$$

Proviamo ora che:

$$\lim_{k \rightarrow +\infty} x_k = \alpha.$$

Da $|x_{k+1} - \alpha| \leq \lambda |x_k - \alpha|$ segue

$$|x_{k+1} - \alpha| \leq \lambda^{k+1} |x_0 - \alpha|.$$

Conseguentemente qualunque sia x_0 si ha:

$$\lim_{k \rightarrow +\infty} |x_k - \alpha| = 0 \Leftrightarrow \lim_{k \rightarrow +\infty} x_k = \alpha. \quad \square$$

Definizione 3.3.1.2 *Un metodo iterativo del tipo (3.5) si dice localmente convergente ad una soluzione α del problema $f(x) = 0$ se esiste un intervallo $[a, b]$ contenente α tale che, per ogni $x_0 \in [a, b]$, la successione generata da (3.5) converge a α .*

Una volta determinata una condizione sufficiente per la convergenza della successione $\{x_k\}$ ad un punto fisso di $g(x)$ si deve essere sicuri che tale punto fisso è unico. Infatti se, oltre ad α esistesse anche $\beta \in [a, b]$ tale che $\beta = g(\beta)$, con $\alpha \neq \beta$, allora

$$|\alpha - \beta| = |g(\alpha) - g(\beta)| = |g'(\xi)| |\alpha - \beta|$$

con $\xi \in [a, b]$. Poichè $|g'(\xi)| < 1$ si ha:

$$|\alpha - \beta| < |\alpha - \beta|$$

e ciò è assurdo.

Come abbiamo già visto nel caso del metodo delle bisezioni anche per metodi di iterazione funzionale è necessario definire dei criteri di arresto per il calcolo delle iterazioni. I criteri di arresto per un processo iterativo basato sullo schema (3.5) sono i seguenti:

$$\begin{aligned} |x_{k+1} - x_k| &\leq \varepsilon \\ \frac{|x_{k+1} - x_k|}{\min(|x_{k+1}|, |x_k|)} &\leq \varepsilon \quad |x_{k+1}|, |x_k| \neq 0 \\ |f(x_k)| &\leq \varepsilon \end{aligned}$$

dove ε è una prefissata tolleranza dipendente dalla precisione con cui si vuole approssimare una soluzione.

Per confrontare differenti metodi iterativi che approssimano la stessa radice α di $f(x) = 0$, si può considerare la velocità con cui tali successioni convergono verso α . Lo studio della velocità di convergenza passa attraverso il concetto di ordine del metodo.

Definizione 3.3.1.3 Sia $\{x_k\}_{k=0}^{\infty}$ una successione convergente ad α e tale che $x_k \neq \alpha$, per ogni k . Se esiste un numero reale $p \geq 1$ tale che

$$\lim_{k \rightarrow +\infty} \frac{|x_{k+1} - \alpha|}{|x_k - \alpha|^p} = \gamma \quad \text{con} \quad \begin{cases} 0 < \gamma \leq 1 & \text{se } p = 1 \\ \gamma > 0 & \text{se } p > 1 \end{cases} \quad (3.8)$$

allora si dice che la successione ha ordine di convergenza p . La costante γ prende il nome di costante asintotica di convergenza.

In particolare se $p = 1$ e $0 < \gamma < 1$ allora la convergenza si dice *lineare*, se $p = 1$ e $\gamma = 1$ allora la convergenza si dice *sublineare*, mentre se $p > 1$ allora la convergenza si dice *superlineare*.

Osservazione. La relazione (3.8) implica che esiste una costante positiva β ($\beta \simeq \gamma$) tale che, per k sufficientemente grande:

$$|x_{k+1} - \alpha| \leq \beta |x_k - \alpha|^p \quad (3.9)$$

La (3.9) indica che la riduzione di errore assoluto ad ogni passo è tanto maggiore quanto più alto è l'ordine di convergenza e, a parità di ordine, quanto più piccola è la costante asintotica di convergenza. I metodi aventi convergenza lineare sono caratterizzati dal fatto che $g'(\alpha) \neq 0$ mentre il seguente teorema stabilisce una relazione tra la funzione $g(x)$ e l'ordine di convergenza di un metodo nel caso sia superlineare.

Teorema 3.3.1.3 (Enunciato). *Sia α punto fisso di $g \in \mathcal{C}^p([a, b])$, $p \geq 2$, intero. Scelto $x_0 \in [a, b]$ la successione $\{x_k\}$ generata dal metodo (3.5) converge a α con ordine p se e soltanto se*

$$g'(\alpha) = g''(\alpha) = \dots = g^{(p-1)}(\alpha) = 0, \quad g^{(p)}(\alpha) \neq 0. \quad \square$$

Definizione 3.3.1.4 *Un metodo iterativo convergente ad α si dice di ordine p (di ordine almeno p) se tutte le successioni ottenute al variare del punto iniziale in un opportuno intorno di α convergono con ordine di convergenza p (almeno p).*

3.3.2 Metodo di Newton-Raphson

Nell'ipotesi che f sia derivabile ed ammetta derivata prima continua allora un altro procedimento per l'approssimazione dello zero della funzione $f(x)$ è il *metodo di Newton-Raphson*, noto anche come *metodo delle tangenti*. Nella Figura 3.1 è riportata l'interpretazione geometrica di tale metodo. A partire dall'approssimazione x_0 si considera la retta tangente la funzione f passante per il punto P_0 di coordinate $(x_0, f(x_0))$. Si calcola l'ascissa x_1 del punto di intersezione tra tale retta tangente e l'asse delle x e si ripete il procedimento a partire dal punto P_1 di coordinate $(x_1, f(x_1))$.

È facile vedere che il metodo definisce il seguente processo iterativo:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad k = 0, 1, 2, \dots \quad (3.10)$$

che equivale, scegliendo in (3.7) $h(x) = f'(x)$, al metodo di iterazione funzionale in cui la funzione $g(x)$ è

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

Per esaminare l'ordine di convergenza del metodo in questione consideriamo prima il α è radice semplice dell'equazione $f(x) = 0$. Supponiamo che

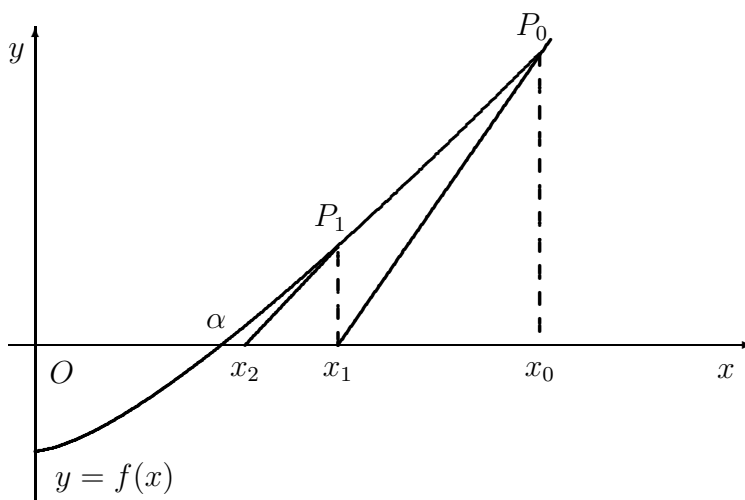


Figura 3.1: Metodo di Newton-Raphson

per ipotesi $f'(x) \neq 0$, per $x \in [a, b]$, dove $[a, b]$ è un opportuno intervallo contenente α . Calcoliamo quindi la derivata prima di $g(x)$:

$$g'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}.$$

Poichè α è semplice risulta $f'(\alpha) \neq 0$ e quindi:

$$g'(\alpha) = \frac{f(\alpha)f''(\alpha)}{[f'(\alpha)]^2} = 0$$

esiste quindi un intorno di α nel quale $|g'(x)| < 1$, per ogni x , e per il teorema (3.3.1.2) comunque si sceglie un punto iniziale appartenente a tale intorno il metodo di Newton-Raphson risulta convergente. Il metodo ha ordine almeno 2 se la derivata prima è diversa da zero in α . Per calcolare esattamente l'ordine di convergenza bisogna calcolare $g''(\alpha)$. Infatti se la funzione $f \in \mathcal{C}^{(3)}([a, b])$ allora

$$g''(x) = \frac{[f'(x)f''(x) + f'''(x)f(x)]f'(x) - 2f'(x)f(x)f''(x)}{[f'(x)]^3}$$

e quindi

$$g''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)}$$

cosicchè se $f''(\alpha) \neq 0$ il metodo ha ordine 2 altrimenti ha addirittura ordine 3.

Tuttavia se la radice α ha molteplicità $r > 1$ l'ordine di convergenza del metodo non è più 2. Se x_0 è sufficientemente vicino ad α è $|g'(x)| < 1$ e quindi per il teorema 3.3.1.2 il metodo è ancora convergente ma l'ordine di convergenza è 1.

Osservazione. Quando la derivata prima della funzione $f(x)$ si mantiene all'incirca costante allora si può porre

$$M = f'(x)$$

e applicare, al posto dello schema di Newton-Raphson, la formula

$$x_{k+1} = x_k - \frac{f(x_k)}{M}. \quad (3.11)$$

La (3.11) definisce un metodo che viene detto *metodo di Newton semplificato* oppure *metodo della direzione costante* in quanto geometricamente equivale all'applicazione del metodo di Newton in cui anzichè prendere la retta tangente la curva f si considera la retta avente coefficiente angolare uguale a M .