

Capitolo 2

Interpolazione e Quadratura 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 *nodi*), x_j , cioè si pone

$$f(x_j) = \sum_{i=0}^n a_i \Phi_i(x_j) \quad j = 0, \dots, n. \quad (2.2)$$

Il processo di determinazione degli a_i attraverso la risoluzione del sistema (2.2) si chiama *metodo dei coefficienti indeterminati*. Il caso più studiato è quello dell'interpolazione polinomiale, in cui si pone:

$$\Phi_i(x) = x^i \quad i = 0, \dots, n$$

e perciò la funzione approssimante g assume la forma

$$\sum_{i=0}^n a_i x^i;$$

mentre le condizioni di coincidenza diventano

$$f(x_j) = \sum_{i=0}^n a_i x_j^i \quad j = 0, \dots, n. \quad (2.3)$$

Nell'ipotesi che i nodi x_j siano tutti distinti si può provare che la matrice dei coefficienti del sistema (2.3) è non singolare (*matrice di Vandermonde*) e pertanto il problema dell'interpolazione ammette sempre un'unica soluzione. Il polinomio interpolante fornito da (2.3) può essere ottenuto anche in modo esplicito. Va comunque notato che in generale è molto difficile ottenere una funzione di interpolazione in forma esplicita, il che rende necessario risolvere il sistema (2.3).

2.1.1 Il Polinomio Interpolante di Lagrange

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

$$L_n(x) = \sum_{k=0}^n l_{nk}(x) f(x_k) \quad (2.4)$$

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

$$l_{nk}(x_i) = \begin{cases} 0 & \text{se } k \neq i \\ 1 & \text{se } k = i \end{cases} \quad (2.5)$$

allora esse sono soddisfatte. In particolare la prima condizione di (2.5) 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 (2.5)

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

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:

$$L_n(x) = \sum_{k=0}^n \left(\prod_{i=0, i \neq k}^n \frac{x - x_i}{x_k - x_i} \right) f(x_k). \quad (2.6)$$

Il polinomio (2.6) 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*.

2.1.2 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

$$e(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \omega_{n+1}(x). \quad (2.7)$$

Esempio 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 $L_4(x)$:

$$L_4(x) = l_{4,0}(x)y_0 + l_{4,1}(x)y_1 + l_{4,2}(x)y_2 + l_{4,3}(x)y_3 + l_{4,4}(x)y_4 \quad (2.8)$$

e calcoliamo i 5 polinomi fondamentali di Lagrange:

$$l_{4,0}(x) = \frac{(x-0)(x-1)(x-3)(x-5)}{(-1-0)(-1-1)(-1-3)(-1-5)} =$$

$$= \frac{1}{48} x(x-1)(x-3)(x-5)$$

$$l_{4,1}(x) = \frac{(x+1)(x-1)(x-3)(x-5)}{(0+1)(0-1)(0-3)(0-5)} =$$

$$= -\frac{1}{15} (x+1)(x-1)(x-3)(x-5)$$

$$l_{4,2}(x) = \frac{(x+1)(x-0)(x-3)(x-5)}{(1+1)(1-0)(1-3)(1-5)} =$$

$$= \frac{1}{16} x(x+1)(x-3)(x-5)$$

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

Sostituendo in (2.8) 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.

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.

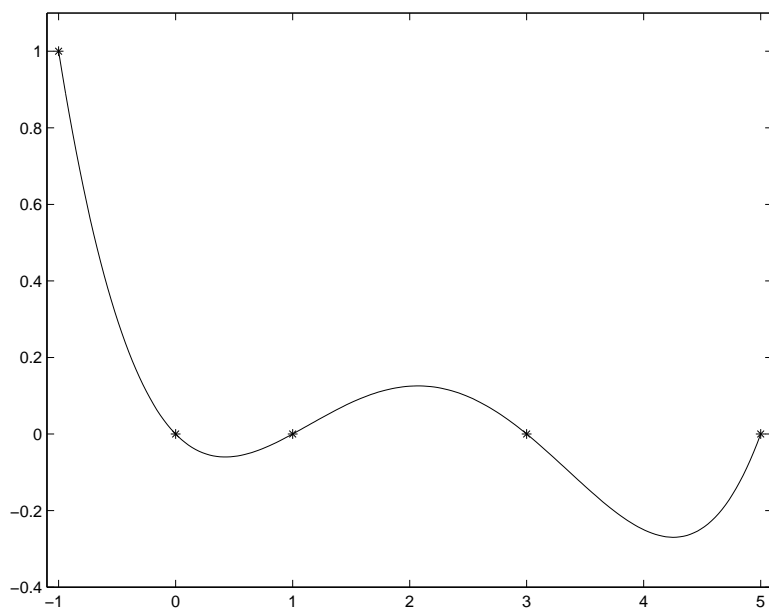


Figura 2.1: Grafico del polinomio $l_{40}(x)$.

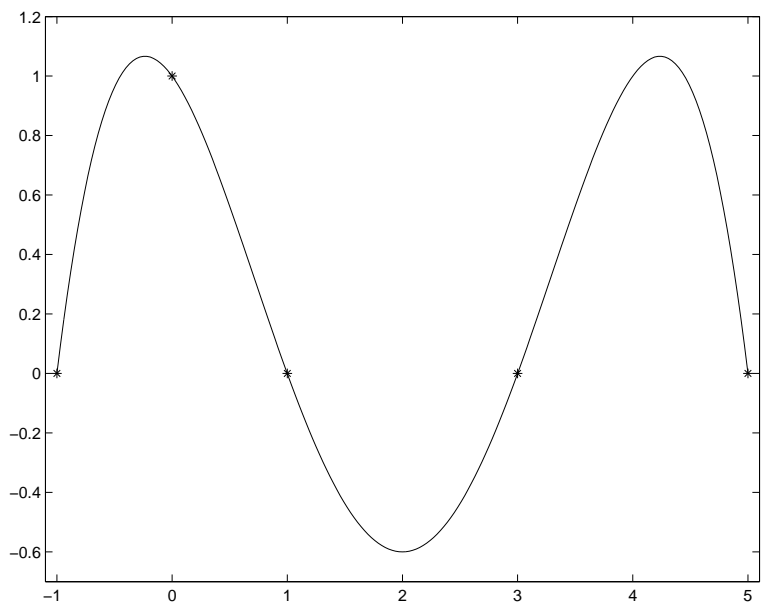


Figura 2.2: Grafico del polinomio $l_{41}(x)$.

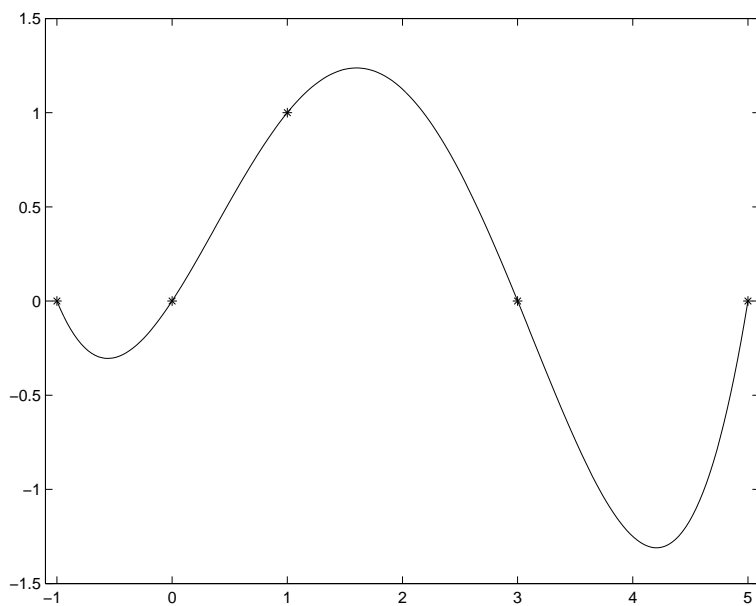


Figura 2.3: Grafico del polinomio $l_{42}(x)$.

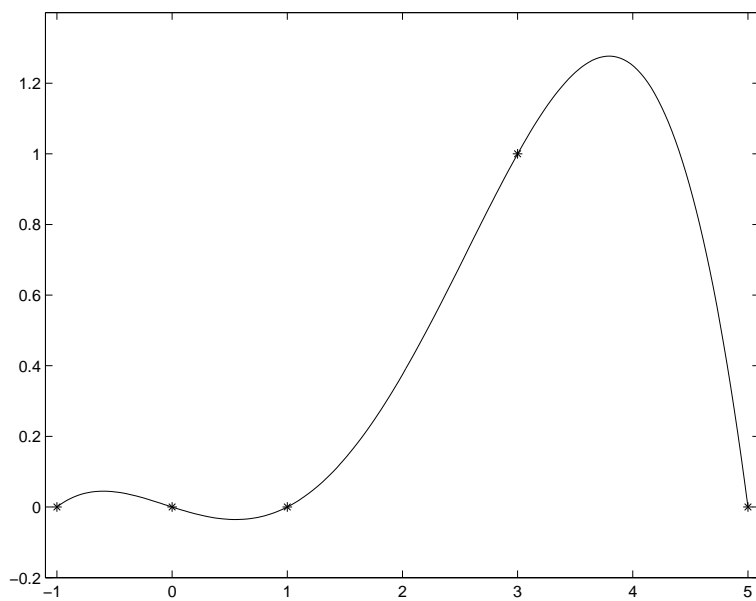


Figura 2.4: Grafico del polinomio $l_{43}(x)$.

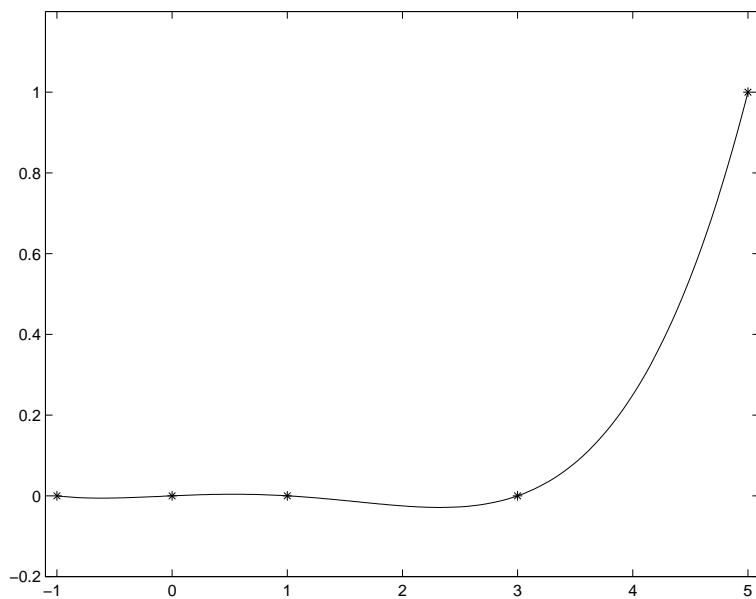


Figura 2.5: Grafico del polinomio $l_{44}(x)$.

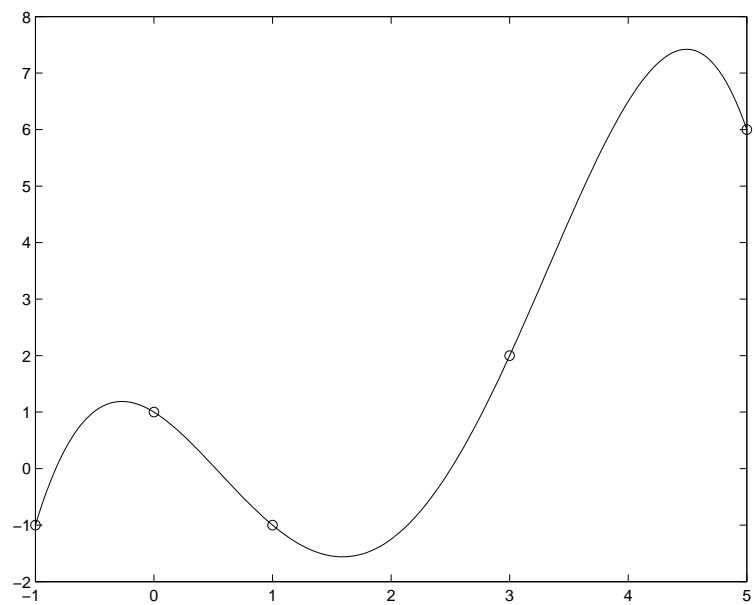


Figura 2.6: Grafico del polinomio interpolante di Lagrange $L_4(x)$.

2.2 Formule di Quadratura di Tipo Interpolatorio

Siano assegnati due valori a, b , con $a < b$, ed una funzione f integrabile sull'intervallo (a, b) . Il problema che ci poniamo è quello di costruire degli algoritmi numerici che ci permettano di valutare, con errore misurabile, il numero

$$I(f) = \int_a^b f(x) dx.$$

Diversi sono i motivi che possono portare alla richiesta di un algoritmo numerico per questi problemi.

Per esempio pur essendo in grado di calcolare una primitiva della funzione f , questa risulta così complicata da preferire un approccio di tipo numerico. Non è da trascurare poi il fatto che il coinvolgimento di funzioni, elementari e non, nella primitiva e la loro valutazione negli estremi a e b comporta comunque un'approssimazione dei risultati. Un'altra eventualità è che f sia nota solo in un numero finito di punti o comunque può essere valutata in ogni valore dell'argomento solo attraverso una routine. In questi casi l'approccio analitico non è neanche da prendere in considerazione.

Supponiamo dunque di conoscere la funzione $f(x)$ nei punti distinti x_0, x_1, \dots, x_n prefissati o scelti da noi, ed esaminiamo la costruzione di formule del tipo

$$\sum_{k=0}^n w_k f(x_k) \tag{2.9}$$

che approssimi realizzare $I(f)$.

Formule di tipo (2.9) si dicono *di quadratura*, i numeri reali x_0, x_1, \dots, x_n e w_0, \dots, w_n si chiamano rispettivamente *nodi* e *pesi* della formula di quadratura.

Il modo più semplice ed immediato per costruire formule di tipo (2.9) è quello di sostituire la funzione integranda $f(x)$ con il polinomio di Lagrange $L_n(x)$ interpolante $f(x)$ nei nodi $x_i, i = 0, \dots, n$. Posto infatti

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

dove $e(x)$ è la funzione errore, abbiamo:

$$\begin{aligned} \int_a^b f(x)dx &= \int_a^b [L_n(x) + e(x)]dx = \\ &= \int_a^b L_n(x)dx + \int_a^b e(x)dx = \\ &= \int_a^b \sum_{k=0}^n l_{nk}(x)f(x_k)dx + \int_a^b e(x)dx = \\ &= \sum_{k=0}^n \left(\int_a^b l_{nk}(x)dx \right) f(x_k) + \int_a^b e(x)dx. \end{aligned}$$

Ponendo

$$w_k = \int_a^b l_{nk}(x)dx \quad k = 0, 1, \dots, n \quad (2.10)$$

e

$$R_{n+1}(f) = \int_a^b e(x)dx \quad (2.11)$$

otteniamo

$$I(f) \simeq \sum_{k=0}^n w_k f(x_k)$$

con un errore stabilito dalla relazione (2.11). Le formule di quadratura con pesi definiti dalle formule (2.10) si dicono *interpolatorie*. La quantità $R_{n+1}(f)$ prende il nome di *Resto della formula di quadratura*. Un utile concetto per misurare il grado di accuratezza con cui una formula di quadratura, interpolatoria o meno, approssima un integrale è il seguente.

Definizione 2.2.1 *Una formula di quadratura ha grado di precisione q se fornisce il valore esatto dell'integrale quando la funzione integranda è un qualunque polinomio di grado al più q ed inoltre esiste un polinomio di grado $q + 1$ tale che l'errore è diverso da zero.*

È evidente da questa definizione che ogni formula di tipo interpolatorio con nodi x_0, x_1, \dots, x_n ha grado di precisione almeno n .

2.3 Formule di Newton-Cotes

Suddividiamo l'intervallo $[a, b]$ in n intervallini di ampiezza h , con

$$h = \frac{b - a}{n}$$

e definiamo i nodi

$$x_i = a + ih \quad i = 0, 1, \dots, n.$$

La formula di quadratura interpolatoria costruita su tali nodi, cioè

$$\int_a^b f(x) dx = \sum_{i=0}^n w_i f(x_i) + R_{n+1}(f)$$

è detta *Formula di Newton-Cotes*.

Una proprietà di cui godono i pesi delle formule di Newton-Cotes è la cosiddetta *proprietà di simmetria*. Infatti poichè i nodi sono a due a due simmetrici rispetto al punto medio c dell'intervallo $[a, b]$, cioè $c = (x_i + x_{n-i})/2$, per ogni i , tale proprietà si ripercuote sui pesi che infatti sono a due a due uguali, cioè $w_i = w_{n-i}$, per ogni i .

Vediamo ora due esempi di formule di Newton-Cotes.

2.3.1 Formula dei Trapezi

Siano $x_0 = a$, $x_1 = b$ e $h = b - a$.

$$\begin{aligned} T_2 &= w_0 f(x_0) + w_1 f(x_1) \\ w_0 &= \int_a^b l_{1,0}(x) dx = \int_a^b \frac{x - x_1}{x_0 - x_1} dx = \int_a^b \frac{x - b}{a - b} dx = \\ &= \frac{1}{a - b} [(x - b)^2]_{x=a}^{x=b} = \frac{h}{2}. \end{aligned}$$

Poichè i nodi scelti sono simmetrici rispetto al punto medio $c = (a + b)/2$ è

$$w_1 = w_0 = \frac{h}{2}.$$

Otteniamo dunque la formula

$$T_2 = \frac{h}{2} [f(a) + f(b)].$$

che viene detta *Formula dei Trapezi*. Per quanto concerne il resto abbiamo

$$R_2(f) = \frac{1}{2} \int_a^b (x-a)(x-b)f''(\xi_x)dx.$$

Prima di vedere come tale espressione può essere manipolata dimostriamo il seguente teorema che è noto come *teorema della media generalizzato*.

Teorema 2.3.1 *Siano $f, g : [a, b] \rightarrow \mathbb{R}$, funzioni continue con $g(x)$ a segno costante e $g(x) \neq 0$ per ogni $x \in]a, b[$. Allora*

$$\int_a^b f(x)g(x)dx = f(\xi) \int_a^b g(x)dx \quad \xi \in [a, b]. \quad \square$$

Poichè la funzione $(x-a)(x-b)$ è a segno costante segue:

$$R_2(f) = \frac{1}{2} f''(\eta) \int_a^b (x-a)(x-b)dx$$

posto $x = a + ht$ otteniamo

$$R_2(f) = \frac{1}{2} f''(\eta) h^3 \int_0^1 t(t-1)dt = -\frac{1}{12} h^3 f''(\eta).$$

2.3.2 Formula di Simpson

Siano $x_0 = a$, $x_2 = b$ mentre poniamo $x_1 = c$, punto medio dell'intervallo $[a, b]$. Allora

$$S_3 = w_0 f(a) + w_1 f(c) + w_2 f(b).$$

Posto

$$h = \frac{b-a}{2}$$

abbiamo

$$w_0 = \int_a^b l_{2,0}(x)dx = \int_a^b \frac{(x-c)(x-b)}{(a-c)(a-b)} dx.$$

Effettuando il cambio di variabile $x = c + ht$ è facile calcolare quest'ultimo integrale, infatti

$$x = a \Rightarrow a = c + ht \Rightarrow a - c = ht \Rightarrow -h = ht \Rightarrow t = -1$$

e

$$x = b \Rightarrow b = c + ht \Rightarrow b - c = ht \Rightarrow h = ht \Rightarrow t = 1.$$

Inoltre $a - c = -h$ e $a - b = -2h$ mentre

$$x - c = c + ht - c = ht, \quad x - b = c + ht - b = c - b + ht = -h + ht = h(t - 1),$$

ed il differenziale $dx = hdt$ cosicchè

$$\begin{aligned} w_0 &= \int_a^b \frac{(x - c)(x - b)}{(a - c)(a - b)} = \int_{-1}^1 \frac{hth(t - 1)}{(-h)(-2h)} hdt = \\ &= \frac{h}{2} \int_{-1}^1 (t^2 - t) dt = \frac{h}{2} \int_{-1}^1 t^2 dt = \frac{h}{2} \left[\frac{t^3}{3} \right]_{-1}^1 = \frac{h}{3}. \end{aligned}$$

Per la simmetria è anche

$$w_2 = w_0 = \frac{h}{3}$$

mentre possiamo calcolare w_1 senza ricorrere alla definizione. Infatti possiamo notare che la formula deve fornire il valore esatto dell'integrale quando la funzione è costante nell'intervallo $[a, b]$, quindi possiamo imporre che, prendendo $f(x) = 1$ in $[a, b]$, sia

$$\int_a^b dx = b - a = \frac{h}{3}(f(a) + f(b)) + w_1 f(c)$$

da cui segue

$$w_1 = b - a - \frac{2}{3}h = 2h - \frac{2}{3}h = \frac{4}{3}h.$$

Dunque

$$S_3 = \frac{h}{3} [f(a) + 4f(c) + f(b)].$$

Questa formula prende il nome di *Formula di Simpson*. Per quanto riguarda l'errore si può dimostrare, e qui ne omettiamo la prova, che vale la seguente relazione

$$R_3(f) = -h^5 \frac{f^{(4)}(\sigma)}{90} \quad \eta, \sigma \in (a, b),$$

che assicura che la formula ha grado di precisione 3.

2.4 Formule di Quadratura Composte

Come abbiamo già avuto modo di vedere le formule di quadratura interpolatorie vengono costruite approssimando su tutto l'intervallo di integrazione la funzione integranda con un unico polinomio, quello interpolante la funzione sui nodi scelti. Per formule convergenti la precisione desiderata si ottiene prendendo n sufficientemente grande. In tal modo comunque, per ogni fissato n , bisogna costruire la corrispondente formula di quadratura. Una strategia alternativa che ha il pregio di evitare la costruzione di una nuova formula di quadratura, e che spesso produce risultati più apprezzabili, è quella delle *formule composte*. L'idea delle formule composte è la seguente.

Scelta una formula di quadratura l'intervallo di integrazione (a, b) viene suddiviso in N intervallini di ampiezza h ,

$$h = \frac{b - a}{N} \quad (2.12)$$

sicchè

$$\int_a^b f(x)dx = \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} f(x)dx$$

dove i punti x_i sono:

$$x_i = a + ih \quad i = 0, \dots, N \quad (2.13)$$

quindi la formula di quadratura viene applicata ad ognuno degli intervallini $[x_i, x_{i+1}]$. Il grado di precisione della formula di quadratura composta coincide con il grado di precisione della formula da cui deriva. Descriviamo ora la *Formula dei Trapezi Composta*.

2.4.1 Formula dei Trapezi Composta

Per quanto visto in precedenza suddividiamo l'intervallo $[a, b]$ in N sottointervalli, ognuno di ampiezza data da h , come in (2.12), e con i nodi x_i definiti in (2.13). Applichiamo quindi in ciascuno degli N intervalli $[x_i, x_{i+1}]$ la formula

dei trapezi. Otteniamo

$$\begin{aligned} \int_a^b f(x)dx &= \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} f(x)dx = \\ &= \sum_{i=0}^{N-1} \left[\frac{h}{2} (f(x_i) + f(x_{i+1})) - \frac{1}{12} h^3 f''(\eta_i) \right] \quad \eta_i \in (x_i, x_{i+1}). \end{aligned}$$

Scrivendo diversamente la stessa espressione

$$\begin{aligned} \int_a^b f(x)dx &= \frac{h}{2} (f(x_0) + f(x_N)) + h \sum_{i=1}^{N-1} f(x_i) - \frac{1}{12} h^3 \sum_{i=0}^{N-1} f''(\eta_i) = \\ &= \frac{h}{2} (f(x_0) + f(x_N)) + h \sum_{i=1}^{N-1} f(x_i) - \frac{1}{12} h^3 N f''(\eta) \end{aligned}$$

dove $\eta \in (a, b)$. L'esistenza di tale punto η è garantito dal cosiddetto *Teorema della media nel discreto* applicato a $f''(x)$, che stabilisce che se $g(x)$ è una funzione continua in un intervallo $[a, b]$ e $\eta_i \in [a, b]$ $i = 1, N$, sono N punti distinti, allora esiste un punto $\eta \in (a, b)$ tale che

$$\sum_{i=1}^N g(\eta_i) = N g(\eta).$$

Dunque la formula dei trapezi composta è data da:

$$T_C(h) = \frac{h}{2} (f(x_0) + f(x_N)) + h \sum_{i=1}^{N-1} f(x_i)$$

con resto

$$R_T = -\frac{1}{12} h^3 N f''(\eta) = -\frac{1}{12} \frac{(b-a)^3}{N^3} N f''(\eta) = -\frac{1}{12} \frac{(b-a)^3}{N^2} f''(\eta).$$

Quest'ultima formula talvolta può essere utile per ottenere a priori una suddivisione dell'intervallo $[a, b]$ in un numero di intervalli che permetta un errore non superiore ad una prefissata tolleranza. Infatti

$$|R_T| \leq \frac{1}{12} \frac{(b-a)^3}{N^2} M, \quad M = \max_{x \in [a, b]} |f''(x)|.$$

Imponendo che $|R_T| \leq \varepsilon$, precisione prefissata, segue

$$N_\varepsilon \geq \sqrt{\frac{(b-a)^3 M}{12\varepsilon}}. \quad (2.14)$$

Tuttavia questo numero spesso risulta una stima eccessiva a causa della maggiorazione della derivata seconda tramite M .

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

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

3.1.1 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).

3.1.2 Il metodo della falsa posizione

Una variante del metodo delle bisezioni è appunto il metodo della falsa posizione. Partendo sempre da una funzione $f(x)$ continua in un intervallo $[a, b]$ tale che $f(a)f(b) < 0$, in questo caso si approssima la radice considerando l'intersezione della retta passante per i punti $(a, f(a))$ e $(b, f(b))$ con l'asse x . L'equazione della retta è

$$y = f(b) + \frac{f(b) - f(a)}{b - a}(x - b)$$

pertanto il punto c_1 , sua intersezione con l'asse x , è:

$$c_1 = b - f(b) \frac{b - a}{f(b) - f(a)}.$$

Si testa a questo punto l'appartenenza della radice α ad uno dei due intervalli $[a, c_1]$ e $[c_1, b]$ e si procede esattamente come nel caso del metodo delle bisezioni, ponendo

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

Ad un generico passo k si calcola

$$c_k = b_{k-1} - f(b_{k-1}) \frac{b_{k-1} - a_{k-1}}{f(b_{k-1}) - f(a_{k-1})}$$

e si pone

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

Anche per questo metodo è possibile dimostrare la convergenza nella sola ipotesi di continuità della funzione $f(x)$.

3.2 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 f 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:

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

dove x_0 è un'assegnato valore iniziale e forniscono un'approssimazione delle soluzioni 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 3.2.1 *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 3.2.2 *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 3.2.1 *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 3.2.2 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 3.2.3 (*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 3.2.3 *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).*

3.2.1 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 alla 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:

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

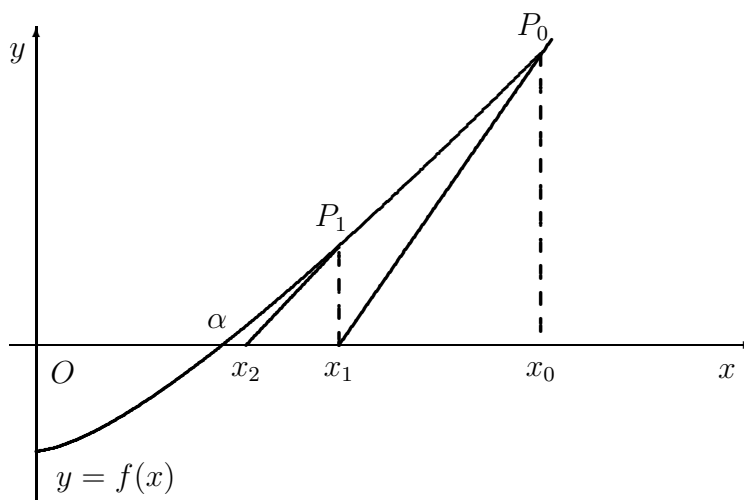


Figura 3.1: Metodo di Newton-Raphson

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

$$g(x) = x - \frac{f(x)}{f'(x)} \quad (3.11)$$

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

$$g'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}. \quad (3.12)$$

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 (3.2.2) comunque si sceglie un punto iniziale appartenente a tale intorno il metodo di Newton-Raphson risulta convergente.

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 3.2.2 il metodo è ancora convergente ma l'ordine di convergenza è 1. *Osservazione.* Se applicando ripetutamente la formula di Newton-Raphson accade che la derivata prima della funzione $f(x)$ si mantiene sensibilmente costante allora si può porre

$$M = f'(x)$$

e applicare la formula

$$x_{k+1} = x_k - \frac{f(x_k)}{M} \quad (3.13)$$

anzichè la (3.10). La (3.13) 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 .

3.2.2 I Metodi delle Secanti e delle Secanti a due punti

Il metodo delle secanti è definito dalla relazione

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - c)}{f(x_k) - f(c)}$$

dove $c \in [a, b]$. Il significato geometrico di tale metodo è il seguente: ad un generico passo k si considera la retta congiungente i punti di coordinate $(x_k, f(x_k))$ e $(c, f(c))$ e si pone x_{k+1} pari al punto di intersezione di tale retta con l'asse x . Dalla formula si evince che la funzione iteratrice del metodo è

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

In base alla teoria vista nei paragrafi precedenti il metodo ha ordine di convergenze 1 se $g'(\alpha) \neq 0$. Può avere ordine di convergenza almeno 1 se $g'(\alpha) = 0$. Tale eventualità si verifica se la tangente alla curva in α ha lo stesso coefficiente angolare della retta congiungente i punti $(\alpha, 0)$ e $(c, f(c))$.

Poichè il metodo delle secanti ha lo svantaggio di avere, solitamente, convergenza lineare mentre il metodo di Newton-Raphson, pur avendo convergenza quadratica, ha lo svantaggio di richiedere, ad ogni passo, due valutazioni di funzioni: $f(x_k)$ ed $f'(x_k)$, quindi se il costo computazionale di $f'(x_k)$ è molto

più elevato rispetto a quello di $f(x_k)$ può essere più conveniente l'uso di metodi che necessitano solo del calcolo del valore della funzione $f(x)$. Un metodo di questo tipo è **il metodo delle secanti a due punti**, che consiste nel considerare, alla k -esima iterata, la retta passante per i punti di coordinate $(x_{k-1}, f(x_{k-1}))$ e $(x_k, f(x_k))$ e definire x_{k+1} come il punto d'intersezione di tale retta con l'asse x . Ne risulta il seguente metodo:

$$x_{k+1} = x_k - f(x_k) \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}. \quad (3.14)$$

Poichè il metodo delle secanti a due punti non rientra tra quelli di iterazione funzionale, per ricavare l'ordine del metodo non si possono applicare i teoremi enunciati in precedenza, pertanto si deve procedere in un modo più che sperimentale. Infatti in questo caso si può vedere che l'ordine non è neanche un numero intero ma irrazionale:

$$p = \frac{1 + \sqrt{5}}{2}.$$

Il metodo delle secanti a due punti ha comunque una convergenza di tipo superlineare, come il metodo di Newton-Raphson, ma ha il vantaggio di richiedere solo una valutazione funzionale ad ogni passo (tranne al primo passo) e di non richiedere il calcolo della derivata prima.