

Capitolo 4

Equazioni e sistemi non lineari

4.1 Introduzione

Sia $f(x):\mathbb{R}\rightarrow\mathbb{R}$ una funzione continua almeno su un certo intervallo \mathcal{I} e si supponga che $f(x)$ non sia della forma $f(x) = a_1x + a_0$ con a_1 e a_0 costanti; la relazione

$$f(x) = 0 \tag{4.1}$$

si dice *equazione non lineare* nell'incognita x .

Il problema di determinare (se esistono) gli zeri di $f(x)$, ossia di trovare le eventuali radici dell'equazione (4.1), raramente può essere risolto con un metodo diretto, cioè effettuando sulla (4.1) un insieme finito di operazioni che conducano ad una espressione esplicita di ciascuna radice. Come esempio basti considerare il caso di una equazione con $f(x)$ polinomio di grado maggiore di 1, cioè il caso in cui la (4.1) sia della forma

$$a_mx^m + a_{m-1}x^{m-1} + \dots + a_1x + a_0 = 0, \quad (a_m \neq 0) \tag{4.2}$$

con m intero ≥ 2 .

La (4.2), che prende il nome di *equazione algebrica di grado m* , com'è noto possiede m radici nel campo complesso, ma queste, salvo casi speciali, si possono trovare con un metodo diretto soltanto per $m \leq 4$. In generale per calcolare numericamente una radice α di una equazione non lineare della forma (4.1) si ricorre ad un metodo iterativo, cioè all'applicazione ripetuta di una formula del tipo

$$x_{n+1} = \phi_n(x_n, x_{n-1}, \dots, x_{n-k+1}), \quad n = 0, 1, \dots, \quad k \geq 1, \tag{4.3}$$

dove ϕ_n si dice la *funzione di iterazione* del metodo. Tale funzione dipende non solo dagli argomenti $x_n, x_{n-1}, \dots, x_{n-k+1}$, ma anche dalla funzione $f(x)$ e la sua forma può variare al variare di n . Una opportuna scelta della funzione di iterazione e delle *approssimazioni iniziali* $x_0, x_{-1}, \dots, x_{-k+1}$ può far sì che la successione $\{x_n\}$ converga alla radice α . Il calcolo viene arrestato al verificarsi di qualche criterio di accettabilità prestabilito. Ad eccezione dei k valori iniziali, ogni altro termine di $\{x_n\}$ viene calcolato in funzione di k termini già noti: per questo motivo la (4.3) viene detta *metodo iterativo a k punti*. Se la forma di ϕ_n non varia al variare di n il metodo si dice *stazionario*.

Definizione 4.1.1 *Data una successione $\{x_n\}$ convergente ad un limite α , si ponga $e_n = x_n - \alpha$; se esistono due numeri reali $p \geq 1$ e $C \neq 0$ tali che sia*

$$\lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{|e_n|^p} = C, \quad (4.4)$$

si dice che la successione ha ordine di convergenza p e fattore di convergenza C .

Per $p = 1$ e $p = 2$ la convergenza si dice anche *lineare* e *quadratica* rispettivamente. Nel caso di $p = 1$ la convergenza ad α implica $C < 1$.

Dalla (4.4) segue che per n "abbastanza grande" si ha

$$\frac{|e_{n+1}|}{|e_n|^p} \simeq C. \quad (4.5)$$

Si dice che il metodo (4.3) è convergente di ordine p se tale è la successione da esso generata. Analogamente a quanto stabilito in 3.6 per i sistemi lineari, si può definire una velocità asintotica di convergenza, nel caso $p = 1$ (e $C < 1$), ponendo $V = -\text{Log}C$; per $p > 1$ la velocità di convergenza ha una espressione più complicata che dipende dai valori di p e di C e cresce al crescere di p e al decrescere di C .

Nei paragrafi che seguono verranno esposti alcuni metodi iterativi a due e ad un punto.

4.2 Metodo di bisezione

Il *metodo di bisezione* è il più semplice metodo iterativo per approssimare gli zeri reali di una funzione $f(x)$. In questo metodo ad ogni passo si

costruisce un intervallo contenente uno zero di $f(x)$ e si assume come approssimazione di tale zero l'ascissa del punto medio del detto intervallo.

Sia $f(x)$ continua sull'intervallo $[a, b]$ e poniamo $x_0 = a$, $x_1 = b$. Supposto che si abbia $f(x_0)f(x_1) < 0$, per la continuità di $f(x)$ si avrà almeno uno zero in $]x_0, x_1[$. Per semplicità supporremo che $]x_0, x_1[$ contenga un solo zero di $f(x)$. Il numero

$$x_2 = \frac{x_1 + x_0}{2},$$

cioè l'ascissa del punto medio di $]x_0, x_1[$, sarà certamente una approssimazione di α migliore di almeno una delle precedenti x_0 e x_1 . Se non si verifica $f(x_2) = 0$, si confronta il segno di $f(x_2)$ con quello di $f(x_1)$; se risulta $f(x_2)f(x_1) < 0$ allora $\alpha \in]x_2, x_1[$, nel caso contrario sarà $\alpha \in]x_0, x_2[$. Quindi la nuova approssimazione x_3 sarà data da

- a) $x_3 = (x_2 + x_1)/2$, se $f(x_2)f(x_1) < 0$,
 b) $x_3 = (x_2 + x_0)/2$, se $f(x_2)f(x_1) > 0$.

Indicando con \hat{x}_2 una variabile che può assumere i valori x_1 o x_0 , possiamo unificare i due casi nella sola formula:

$$x_3 = \frac{x_2 + \hat{x}_2}{2} \quad \text{dove } \hat{x}_2 = \begin{cases} x_1 & \text{se } f(x_2)f(x_1) < 0, \\ x_0 & \text{altrimenti.} \end{cases}$$

Per esempio, nel caso della Fig. 4.1 si verifica il caso b).

Ripetendo il procedimento, si determinano x_4, x_5, x_6, \dots secondo la formula generale

$$x_{n+1} = \frac{x_n + \hat{x}_n}{2}, \quad n = 1, 2, \dots; \quad (4.6)$$

dove per $n = 1$ è $\hat{x}_1 = x_0$ mentre per $n > 1$ si pone

$$\hat{x}_n = \begin{cases} x_{n-1} & \text{se } f(x_n)f(x_{n-1}) < 0, \\ \hat{x}_{n-1} & \text{altrimenti.} \end{cases}$$

Il metodo di bisezione è quindi un metodo iterativo a due punti non stazionario, infatti la funzione al secondo membro della (4.6) non è la stessa ad ogni passo.

Poiché ad ogni passo l'intervallo contenente α viene dimezzato, dopo n passi si ha una approssimazione x_{n+1} , tale che

$$|x_{n+1} - \alpha| \leq \frac{1}{2^n}(b - a), \quad (4.7)$$

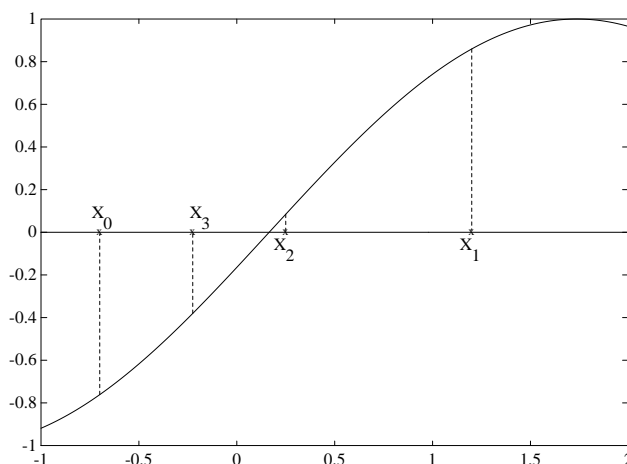


Figura 4.1: Metodo di bisezione.

da cui segue $\lim_{n \rightarrow \infty} |x_n - \alpha| = 0$, che prova la convergenza del metodo (4.6) alla radice α e fornisce anche una maggiorazione a priori dell'errore assoluto presente nell'iterata x_{n+1} . La (4.7) suggerisce come un possibile criterio di arresto la condizione $\frac{1}{2^n}(b-a) < \epsilon$, dove $\epsilon > 0$ è un numero opportunamente prefissato. La stessa condizione permette di conoscere a priori il numero n di iterazioni necessario per ridurre il modulo dell'errore assoluto al disotto di ϵ .

Il metodo (4.6) converge linearmente, infatti assumendo $|x_n - x_{n-1}|$ come stima di $|x_n - \alpha| = |e_n|$, si ha, per n abbastanza grande, l'uguaglianza approssimata della forma (4.5)

$$\frac{|e_{n+1}|}{|e_n|} \simeq \frac{|x_{n+1} - x_n|}{|x_n - x_{n-1}|} = \frac{1}{2};$$

poiché la convergenza è lenta, di solito questo metodo è usato per ottenere una prima approssimazione che consenta l'uso di altri metodi più efficienti.

4.3 Regola Falsi e metodo delle Secanti

Si abbiano le stesse condizioni iniziali del metodo di bisezione, ossia $x_0 < \alpha < x_1$ e $f(x_0)f(x_1) < 0$. Il *metodo di falsa posizione*, noto anche col nome di *regola falsi*, è un altro metodo iterativo a due punti, in generale non

stazionario, nel quale, partendo da $]x_0, x_1[$, ad ogni passo si costruisce un nuovo intervallo di estremi x_n e \hat{x}_n contenente uno zero di $f(x)$ e si assume come approssimazione x_{n+1} lo zero di una funzione lineare il cui grafico è la retta per i punti $A_n \equiv [x_n, f(x_n)]$ e $\hat{A}_n \equiv [\hat{x}_n, f(\hat{x}_n)]$. Si ha quindi lo schema

$$x_{n+1} = x_n - f(x_n) \frac{x_n - \hat{x}_n}{f(x_n) - f(\hat{x}_n)}, \quad n = 1, 2, \dots, \quad (4.8)$$

dove per $n = 1$ è $\hat{x}_1 = x_0$ mentre per $n > 1$ si pone, come nel metodo di bisezione,

$$\hat{x}_n = \begin{cases} x_{n-1} & \text{se } f(x_n)f(x_{n-1}) < 0, \\ \hat{x}_{n-1} & \text{altrimenti.} \end{cases}$$

Si noti che anche nel metodo di bisezione il numero x_{n+1} fornito dalla (4.6) può interpretarsi come lo zero di una funzione lineare il cui grafico è la retta passante per i punti $A_n \equiv [x_n, \text{sign}_0(f(x_n))]$ e $\hat{A}_n \equiv [\hat{x}_n, \text{sign}_0(f(\hat{x}_n))]$, dove si è posto

$$\text{sign}_0(z) := \begin{cases} \frac{z}{|z|} & \text{per } z \neq 0, \\ 0 & \text{per } z = 0. \end{cases}$$

Il fatto che nella regola falsi si utilizzino i valori di $f(x)$ anziché i soli segni, spiega perché, quando il metodo (4.8) converge, la convergenza è più rapida che nel metodo (4.6). L'interpretazione geometrica del metodo è illustrata nella Fig. 4.2.

Teorema 4.3.1 *Se $f(x) \in C^2(\mathcal{I})$ e se il metodo regola falsi converge ad uno zero α di $f(x)$ con $\alpha \in \mathcal{I}$ tale che $f'(\alpha) \neq 0$, $f''(\alpha) \neq 0$, la convergenza è di ordine $p = 1$.*

Una variante importante della regola falsi è il *metodo delle secanti*, in cui sono richieste due approssimazioni iniziali senza alcun'altra condizione e senza la necessità di controllare il segno di $f(x)$ ad ogni passo.

In questo metodo il calcolo dell'approssimazione x_{n+1} utilizza le informazioni precedenti, x_n e x_{n-1} , secondo la formula

$$x_{n+1} = x_n - f(x_n) \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}, \quad n = 1, 2, \dots; \quad (4.9)$$

il numero x_{n+1} è individuato geometricamente (cfr. Fig. 4.3) dal punto in cui la secante passante per i punti $A_n \equiv [x_n, f(x_n)]$ e $A_{n-1} \equiv [x_{n-1}, f(x_{n-1})]$ incontra l'asse delle ascisse.

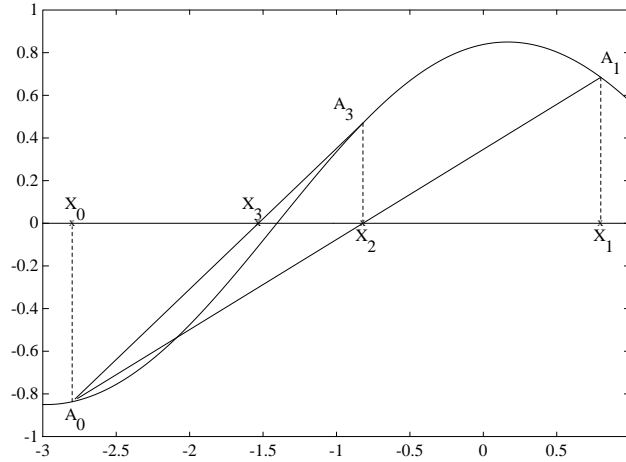


Figura 4.2: Regula falsi.

La (4.9) rappresenta un metodo iterativo stazionario a due punti, ma, per $n \geq 2$, ad ogni passo il calcolo di x_{n+1} richiede la sola valutazione di $f(x_n)$.

La convergenza del metodo è garantita se le approssimazioni x_0 e x_1 si scelgono "abbastanza vicine" alla radice α ; vale il seguente teorema.

Teorema 4.3.2 *Se $f(x) \in C^2(\mathcal{I})$ e se il metodo delle secanti converge ad uno zero α di $f(x)$ con $\alpha \in \mathcal{I}$ e tale che $f'(\alpha) \neq 0$, $f''(\alpha) \neq 0$, allora l'ordine della convergenza è $p = (1 + \sqrt{5})/2 \simeq 1.618$.*

4.4 Metodi iterativi a un punto

In questo paragrafo si esporranno alcune proprietà generali dei metodi iterativi stazionari a un punto.

Data l'equazione $f(x) = 0$, si può sempre costruire una funzione $\phi(x)$ tale che l'equazione data sia equivalente alla seguente

$$x = \phi(x). \quad (4.10)$$

Basta infatti porre, ad esempio,

$$\phi(x) = x - g(x)f(x),$$

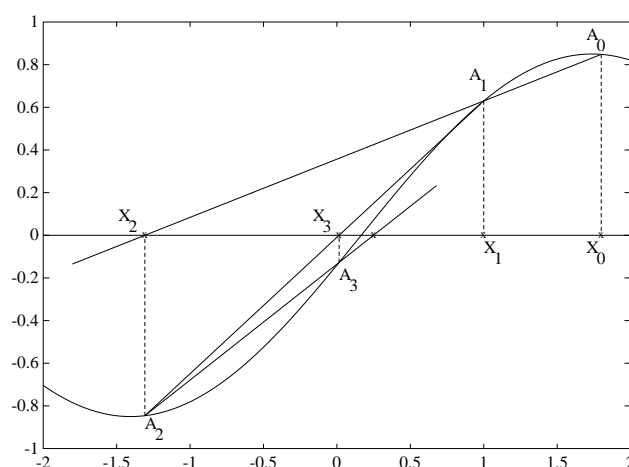


Figura 4.3: Metodo delle secanti.

dove $g(x)$ è un'arbitraria funzione continua che si suppone diversa da zero nei punti di una regione contenente gli zeri di $f(x)$. Se α è uno zero di $f(x)$ si ha anche $\alpha = \phi(\alpha)$ e viceversa; α si dice un *punto fisso* della trasformazione (4.10) o, semplicemente, di $\phi(x)$.

Per ogni scelta della funzione $\phi(x)$ si può considerare un metodo iterativo stazionario ad un punto della forma

$$x_{n+1} = \phi(x_n), \quad n = 0, 1, 2, \dots, \quad (4.11)$$

e il problema di approssimare uno zero α di $f(x)$ si riduce a quello di costruire mediante la (4.11) una successione convergente ad α , punto fisso di $\phi(x)$.

Dal teorema seguente risulta una condizione sufficiente per la *convergenza locale* del metodo (4.11), cioè una convergenza assicurata dalla scelta di x_0 in un opportuno intorno di α .

Teorema 4.4.1 *Sia α un punto fisso di $\phi(x)$ interno ad un intervallo \mathcal{I} sul quale $\phi(x)$ sia derivabile con continuità e si supponga che esistano due numeri positivi ρ e K con $K < 1$, tali che $\forall x \in [\alpha - \rho, \alpha + \rho] \subset \mathcal{I}$ si verifichi la condizione*

$$|\phi'(x)| \leq K; \quad (4.12)$$

allora per il metodo (4.11) valgono le seguenti proposizioni:

1. se $x_0 \in]\alpha - \rho, \alpha + \rho[$ allora è anche $x_n \in]\alpha - \rho, \alpha + \rho[$ per $n = 1, 2, \dots$;
2. la successione $\{x_n\}$, con $x_0 \in]\alpha - \rho, \alpha + \rho[$, converge e si ha $\lim_{n \rightarrow \infty} x_n = \alpha$;
3. α è l'unico punto fisso di $\phi(x)$ in $[\alpha - \rho, \alpha + \rho]$.

DIMOSTRAZIONE. La proposizione 1 si dimostra per induzione, cioè, scelto $x_0 \in]\alpha - \rho, \alpha + \rho[$, si ammette per ipotesi che sia, per un certo n , $x_n \in]\alpha - \rho, \alpha + \rho[$, ossia $|x_n - \alpha| < \rho$ e si deduce che deve essere $x_{n+1} \in]\alpha - \rho, \alpha + \rho[$, ovvero $|x_{n+1} - \alpha| < \rho$. Infatti, facendo uso della (4.11) e del teorema del valor medio, si ha

$$x_{n+1} - \alpha = \phi(x_n) - \phi(\alpha) = \phi'(\xi)(x_n - \alpha),$$

dove ξ è compreso tra x_n e α ; dall'ipotesi fatta su x_n e da quelle del teorema segue poi

$$|x_{n+1} - \alpha| = |\phi'(\xi)| |x_n - \alpha| \leq K |x_n - \alpha| < \rho. \quad (4.13)$$

La proposizione 2 segue dall'ipotesi $0 < K < 1$ e dalla disuguaglianza

$$|x_{n+1} - \alpha| \leq K^{n+1} |x_0 - \alpha|$$

che si ottiene dalla (4.13).

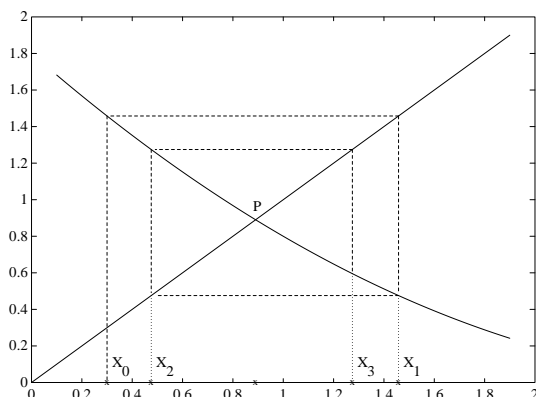
Infine la proposizione 3 risulta per assurdo, infatti se in $]\alpha - \rho, \alpha + \rho[$ esistesse un altro punto fisso $\alpha' \neq \alpha$, si avrebbe

$$|\alpha - \alpha'| = |\phi(\alpha) - \phi(\alpha')| = |\phi'(\xi)| |\alpha - \alpha'| \leq K |\alpha - \alpha'| < |\alpha - \alpha'|.$$

□

Il Teorema 4.4.1 trova applicazione in molti casi concreti in cui qualche zero di $f(x)$ sia stato già localizzato (per esempio col metodo di bisezione) in un intervallo su cui valga la (4.12). Nelle Figure 4.4 e 4.5 sono riportati i grafici delle funzioni $y = x$ e $y = \phi(x)$ che si incontrano nel punto P di ascissa α . In esse è data l'interpretazione geometrica di un metodo della forma (4.11) con la condizione (4.12) verificata in due modi diversi.

In generale l'ordine di convergenza di un metodo iterativo è un numero reale ≥ 1 (vedi per esempio il metodo delle secanti). Per i metodi iterativi stazionari ad un punto vale però il teorema seguente.

Figura 4.4: $-1 < \phi'(x) < 0$.

Teorema 4.4.2 *Un metodo iterativo ad un punto, la cui funzione di iterazione $\phi(x)$ sia sufficientemente derivabile, ha ordine di convergenza uguale ad un numero intero positivo p . Precisamente se il metodo (4.11) converge ad α , la convergenza è di ordine p allora e solo che si abbia*

$$\phi(\alpha) = \alpha, \quad \phi^{(i)}(\alpha) = 0 \quad \text{per } 1 \leq i < p, \quad \phi^{(p)}(\alpha) \neq 0. \quad (4.14)$$

DIMOSTRAZIONE. Usando la (4.11) e la formula di Taylor si ha in generale

$$\begin{aligned} x_{n+1} - \alpha &= \phi(x_n) - \phi(\alpha) = \phi'(\alpha)(x_n - \alpha) + \phi''(\alpha) \frac{(x_n - \alpha)^2}{2!} + \dots \\ &+ \dots + \phi^{(p-1)}(\alpha) \frac{(x_n - \alpha)^{p-1}}{(p-1)!} + \phi^{(p)}(\xi) \frac{(x_n - \alpha)^p}{p!} \end{aligned}$$

dove ξ è compreso tra x_n e α ; quindi, se valgono le (4.14), si ha

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = \frac{|\phi^{(p)}(\alpha)|}{p!} \neq 0, \quad (4.15)$$

cioè l'ordine di convergenza è p ed il fattore di convergenza è $C = \frac{|\phi^{(p)}(\alpha)|}{p!}$.

Viceversa se l'ordine è p , sia $\phi^{(i)}(\alpha)$ la prima derivata non nulla nel precedente sviluppo di Taylor intorno al punto α ; se fosse $i \neq p$, per il ragionamento diretto anche l'ordine sarebbe $i \neq p$ contro l'ipotesi. Quindi deve essere $i = p$ cioè valgono le (4.14). \square

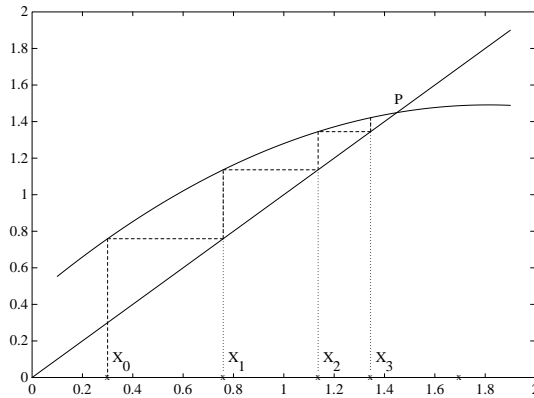


Figura 4.5: $0 < \phi'(x) < 1$.

È importante notare che i Teoremi 4.4.1, 4.4.2 e quelli di 4.3 valgono per la successione teorica $\{x_n\}$ definita dal processo iterativo in assenza di errori di arrotondamento. Nella pratica però si introduce al passo n -esimo un errore δ_n dovuto agli arrotondamenti nel calcolo di ϕ e agli errori dei passi precedenti, per cui, nel caso specifico del processo (4.11), la successione che si ottiene è in realtà

$$\tilde{x}_{n+1} = \phi(\tilde{x}_n) + \delta_n, \quad n = 0, 1, \dots; \quad (4.16)$$

per tale successione $\{\tilde{x}_n\}$, anche quando fossero verificate le ipotesi del Teorema 4.4.1, non è più garantita la convergenza ad α .

Tuttavia si può dimostrare che nelle ipotesi del teorema di convergenza e verificandosi la condizione $|\delta_n| \leq \delta$, $n = 0, 1, \dots$, i termini della successione (4.16) finiscono col cadere in un intorno di α la cui ampiezza dipende linearmente da δ e inoltre, se K non è troppo vicino a 1 e se si opera con una precisione tale da rendere trascurabile il numero $\delta/(1-K)$, l'errore $|\tilde{x}_{n+1} - \alpha|$ è dell'ordine di $|\tilde{x}_{n+1} - \tilde{x}_n|$. Quindi come criterio di arresto dell'algoritmo iterativo si può assumere la condizione $|\tilde{x}_{n+1} - \tilde{x}_n| \leq \epsilon$ oppure $\left| \frac{\tilde{x}_{n+1} - \tilde{x}_n}{\min(|\tilde{x}_n|, |\tilde{x}_{n+1}|)} \right| \leq \epsilon$ a seconda che si voglia limitare l'errore assoluto o quello relativo.

Si abbia una successione $\{x_n\}$ convergente linearmente; in base alla (4.5) si ha quindi, per $p = 1$ e per n "sufficientemente" grande,

$$x_{n+1} - \alpha \simeq C(x_n - \alpha),$$

$$x_{n+2} - \alpha \simeq C(x_{n+1} - \alpha),$$

da cui

$$\alpha \simeq \frac{x_n x_{n+2} - x_{n+1}^2}{x_{n+2} - 2x_{n+1} + x_n}. \quad (4.17)$$

Si ha perciò una approssimazione di α costruita con tre termini successivi x_n, x_{n+1}, x_{n+2} . Il secondo membro della (4.17) si può considerare come il termine z_n di una nuova successione, che, per evitare errori di cancellazione, si pone, di solito, nella forma

$$z_n = x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n}. \quad (4.18)$$

Introducendo i simboli $\Delta_n := x_{n+1} - x_n$ e $\Delta_n^2 := x_{n+2} - 2x_{n+1} + x_n$ ($\Delta_n^2 = \Delta_{n+1} - \Delta_n$) la successione $\{z_n\}$ può scriversi

$$z_n = x_n - \frac{(\Delta_n)^2}{\Delta_n^2}, \quad n = 0, 1, \dots \quad (4.19)$$

Lo schema di calcolo definito dalla (4.19) va sotto il nome di *processo* Δ^2 di *Aitken* ed ha lo scopo di accelerare la convergenza di una successione che converga linearmente. Vale infatti il seguente teorema.

Teorema 4.4.3 *Se una successione $\{x_n\}$ converge ad α linearmente, allora la successione $\{z_n\}$ converge allo stesso limite α più rapidamente di $\{x_n\}$, ossia si ha*

$$\lim_{n \rightarrow \infty} \frac{z_n - \alpha}{x_n - \alpha} = 0.$$

Nella Fig. 4.6 è data una interpretazione geometrica facilmente verificabile del processo di Aitken, nel caso di una successione generata con iterazioni del tipo $x_{n+1} = \phi(x_n)$: z_n è l'ascissa della intersezione della retta per i punti $P_n \equiv [x_n, \phi(x_n)]$ e $P_{n+1} \equiv [x_{n+1}, \phi(x_{n+1})]$ con la retta di equazione $y = x$.

Sulla base della (4.17), supponendola valida anche nel caso di una successione che converga non linearmente, è possibile, partendo da un metodo iterativo $x_{n+1} = \phi(x_n)$ di un dato ordine, costruirne un altro di ordine più elevato. Infatti, esprimendo il membro destro della (4.18) in funzione della sola x_n , si ha

$$F(x_n) = x_n - \frac{(\phi(x_n) - x_n)^2}{\phi(\phi(x_n)) - 2\phi(x_n) + x_n}.$$

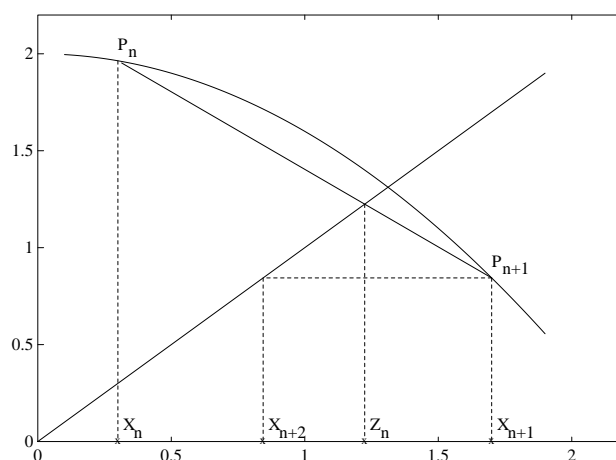


Figura 4.6: Processo di Aitken.

In generale $F(x)$ ha gli stessi punti fissi di $\phi(x)$.

Il metodo iterativo

$$x_{n+1} = F(x_n), \quad n = 0, 1, \dots, \quad (4.20)$$

è noto come *metodo di Steffensen* e per esso si può dimostrare il teorema seguente.

Teorema 4.4.4 *Sia $\phi(x)$ la funzione di iterazione di un metodo di ordine p per approssimare un suo punto fisso α . Per $p > 1$ il corrispondente metodo di Steffensen (4.20) per approssimare α ha ordine $2p - 1$, mentre per $p = 1$ e nella ipotesi $\phi'(\alpha) \neq 1$, il metodo (4.20) ha ordine almeno 2.*

4.5 Metodo di Newton

Il più importante fra i metodi ad un punto è il *metodo di Newton*. Tale metodo si può applicare per approssimare uno zero α di $f(x)$ se, in tutto un intorno di α , $f(x)$ è derivabile con continuità. In tal caso, assumendo la funzione di iterazione della forma

$$\phi(x) = x - \frac{f(x)}{f'(x)} \quad (4.21)$$

si ha il metodo

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, \dots \quad (4.22)$$

L'iterata x_{n+1} è individuata dal punto d'incontro dell'asse delle ascisse con la tangente alla curva $y = f(x)$ nel punto $A_n \equiv [x_n, f(x_n)]$ (cfr. Fig. 4.7); per questo si usa anche la denominazione di *metodo delle tangenti*.

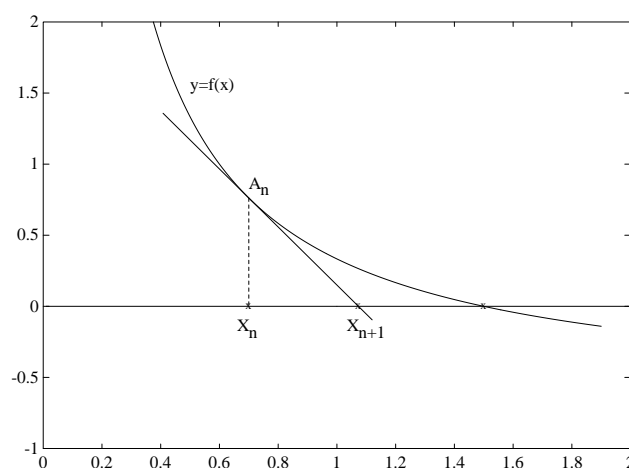


Figura 4.7: Metodo di Newton.

Sulla convergenza e l'ordine del metodo di Newton, vale il seguente teorema.

Teorema 4.5.1 *Sia $f(x) \in C^3([a, b])$, $a < \alpha < b$, $f(\alpha) = 0$, $f'(\alpha) \neq 0$, allora valgono le proposizioni:*

1. *esiste un numero $\rho > 0$ tale che per ogni $x_0 \in [\alpha - \rho, \alpha + \rho]$ il metodo (4.22) converge;*
2. *la convergenza è di ordine $p \geq 2$;*
3. *se $p = 2$ il fattore di convergenza è $C = \frac{1}{2} \left| \frac{f''(\alpha)}{f'(\alpha)} \right|$.*

DIMOSTRAZIONE. Dalla (4.21) segue

$$\phi'(x) = \frac{f(x)f''(x)}{f'^2(x)} \quad (4.23)$$

da cui

$$\phi'(\alpha) = 0; \quad (4.24)$$

quindi, fissato un numero positivo $K < 1$, esiste un numero $\rho > 0$ tale che per ogni $x \in [\alpha - \rho, \alpha + \rho]$ si abbia $|\phi'(x)| \leq K$ e perciò vale il teorema di convergenza 4.4.1.

Per dimostrare l'asserto 2, dalla (4.23) si ricava

$$\phi''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)};$$

ne segue che se $f''(\alpha) \neq 0$ è anche $\phi''(\alpha) \neq 0$; questa insieme alla (4.24), garantisce per il Teorema 4.4.2, che l'ordine è $p = 2$. Se invece si ha $f''(\alpha) = 0$ è anche $\phi''(\alpha) = 0$ e quindi si ha $p > 2$.

Se $p = 2$, l'asserto 3 segue dalla (4.15) essendo $\phi'(\alpha) = 0$ e $\phi''(\alpha) \neq 0$. \square

Notiamo che la convergenza di cui si parla nella proposizione 1 del teorema precedente è di tipo locale, cioè si verifica se si sceglie x_0 in un intorno di α di raggio ρ abbastanza piccolo. A un tale intorno si può pervenire, ad esempio, col metodo di bisezione.

Vi sono però casi speciali in cui si può avere una *convergenza globale*, cioè che si verifica per qualunque scelta di x_0 in un dato intervallo limitato $[a, b]$. Ciò è asserto nel teorema seguente.

Teorema 4.5.2 *Sia $f(x) \in C^2([a, b])$ con $f(a)f(b) < 0$ e si supponga*

1. $f'(x) \neq 0, \forall x \in [a, b]$,
2. $f''(x) \geq 0$, oppure $f''(x) \leq 0, \forall x \in [a, b]$,
3. $\left| \frac{f(a)}{f'(a)} \right| < b - a, \left| \frac{f(b)}{f'(b)} \right| < b - a$,

allora $f(x)$ ha un solo zero $\alpha \in [a, b]$ ed il metodo di Newton converge ad α per ogni $x_0 \in [a, b]$.

Osservazione 4.5.1 Si noti che ponendo nella (4.22) $x_n = a$ oppure $x_n = b$ si ottiene rispettivamente

$$|x_{n+1} - a| = \left| \frac{f(a)}{f'(a)} \right| \quad \text{e} \quad |x_{n+1} - b| = \left| \frac{f(b)}{f'(b)} \right|,$$

per cui le condizioni dell'ipotesi 3 equivalgono a $|x_{n+1} - a| < b - a, |x_{n+1} - b| < b - a$, cioè le tangenti negli estremi $A \equiv [a, f(a)]$ e $B \equiv [b, f(b)]$ intersecano l'asse delle ascisse all'interno di $[a, b]$.

Sotto le ipotesi 1 e 2 del Teorema 4.5.2 cadono in particolare le funzioni monotone convesse o concave su $[a, b]$. Una importante applicazione si ha nella speciale equazione

$$f(x) = x^m - c = 0, \quad m \in \mathbb{R}, \quad m \geq 2, \quad c > 0, \quad (4.25)$$

che possiede la soluzione positiva $\alpha = \sqrt[m]{c}$, radice m -esima aritmetica di c .

Poiché per $x > 0$ si ha $f'(x) > 0$ e $f''(x) > 0$, per applicare il Teorema 4.5.2 basta constatare che si possono assegnare intervalli $[a, b]$ con $a < \alpha < b$, tali da soddisfare le condizioni 3. Se $a > 0$ è tale che sia $f(a) < 0$, si avrà $a < \alpha$. Ogni numero $b > a + \left| \frac{f(a)}{f'(a)} \right| = a + \frac{a^m - c}{ma^{m-1}}$ verifica la prima delle condizioni 3; inoltre, con questa scelta, b risulta a destra dell'intersezione della tangente in $A \equiv [a, f(a)]$ con l'asse delle ascisse, quindi è $b > \alpha$ e si ha

$$\begin{aligned} \left| \frac{f(b)}{f'(b)} \right| &= \frac{b^m - \alpha^m}{mb^{m-1}} \\ &= (b - \alpha) \frac{b^{m-1} + \alpha b^{m-2} + \dots + \alpha^{m-1}}{mb^{m-1}} \\ &< b - \alpha < b - a. \end{aligned}$$

Pertanto un tale numero b verifica anche la seconda condizione 3. Ne segue che ogni numero positivo può considerarsi interno a un intervallo $[a, b]$ che verifichi le condizioni 3 e perciò il metodo di Newton applicato alla (4.25) converge ad α , comunque si scelga una approssimazione iniziale $x_0 > 0$. Il metodo fornisce la formula

$$x_{n+1} = \frac{1}{m} [(m-1)x_n + cx_n^{1-m}] \quad n = 0, 1, \dots$$

Per $m = 2$ si ha l'algoritmo $x_{n+1} = \frac{1}{2} \left(x_n + \frac{c}{x_n} \right)$ per estrarre la radice quadrata di c , notevole per il suo basso costo computazionale.

Il Teorema 4.5.1 vale nell'ipotesi $f'(\alpha) \neq 0$, cioè se α è uno zero semplice di $f(x)$. Supponiamo ora che α sia per $f(x)$ uno zero di molteplicità $s \geq 1$; in tal caso si può scrivere

$$f(x) = g(x)(x - \alpha)^s, \quad \text{con } g(x) = \frac{f(x)}{(x - \alpha)^s} \quad \text{e } g(\alpha) = \lim_{x \rightarrow \alpha} g(x) \neq 0.$$

Quindi per il metodo di Newton si ha

$$\phi(x) = x - \frac{f(x)}{f'(x)} = x - \frac{g(x)(x - \alpha)}{sg(x) + g'(x)(x - \alpha)}$$

da cui segue, con facili calcoli,

$$\phi'(\alpha) = 1 - \frac{1}{s}. \quad (4.26)$$

Perciò, per $s > 1$, il metodo convergere linearmente con fattore di convergenza $C = 1 - \frac{1}{s}$.

Se si sostituisce la funzione $f(x)$ con $F(x) = f(x)/f'(x)$ che ha gli stessi zeri di $f(x)$ ma tutti semplici, il metodo di Newton, applicato all'equazione $F(x) = 0$ fornisce ancora una convergenza almeno quadratica verso la radice α , essendo ora $F'(\alpha) \neq 0$. Il costo computazionale è però maggiore di quello richiesto dall'equazione $f(x) = 0$, poiché, in effetti, si usa la formula iterativa

$$x_{n+1} = x_n - \frac{f(x_n)f'(x_n)}{f'^2(x_n) - f(x_n)f''(x_n)}, \quad n = 0, 1, \dots \quad (4.27)$$

Quando si conosce la molteplicità della radice α si può usare una formula meno costosa della (4.27), ottenuta modificando il metodo di Newton in modo che la convergenza sia almeno quadratica anche per una radice di molteplicità $s > 1$. Scrivendo infatti

$$x_{n+1} = x_n - s \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \dots, \quad (4.28)$$

si definisce un metodo per il quale si ha

$$\phi(x) = x - s \frac{f(x)}{f'(x)},$$

da cui segue, tenuto conto della (4.26),

$$\phi'(\alpha) = 0.$$

Nell'uso del metodo (4.22) può accadere che i valori di $f'(x_n)$ varino molto lentamente al variare di n . In tal caso può essere utile una variante del metodo di Newton in cui lo stesso valore di $f'(x_n)$ viene utilizzato per un certo numero di passi successivi, dopo i quali viene ricalcolato; in altri termini

$f'(x_n)$ non si calcola ad ogni passo ma solo ogni volta che n coincide con un elemento di un prefissato sottoinsieme dei numeri naturali, lasciando $f'(x_n)$ uguale all'ultimo valore calcolato se n non appartiene a tale sottoinsieme. Si noti che se il valore di $f'(x_n)$ viene mantenuto costante per ogni n , la convergenza del metodo dipende dal valore di tale costante ed è comunque lineare. Anche il metodo delle secanti può considerarsi ottenuto dal metodo di Newton approssimando $f'(x_n)$ con il rapporto incrementale $\frac{f(x_n)-f(x_{n-1})}{x_n-x_{n-1}}$.

4.6 Metodi iterativi in \mathbb{R}^n

La teoria dei metodi iterativi precedentemente esposta può essere estesa al caso in cui, nella (4.1), sia $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$, cioè al caso di un sistema di n equazioni (non lineari) in altrettante incognite

$$\begin{aligned} f_1(x_1, \dots, x_n) &= 0 \\ f_2(x_1, \dots, x_n) &= 0 \\ &\vdots \\ f_n(x_1, \dots, x_n) &= 0. \end{aligned} \tag{4.29}$$

In analogia con quanto esposto in 4.4, il sistema (4.29) si può scrivere in una forma equivalente che consente di approssimare una soluzione $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$ come punto fisso di una opportuna funzione di iterazione $\phi(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ data da

$$\phi(x) = x - G(x)f(x), \tag{4.30}$$

dove $\phi = (\phi_1, \phi_2, \dots, \phi_n)^T$, $x = (x_1, x_2, \dots, x_n)^T$, $f = (f_1, f_2, \dots, f_n)^T$ e $G(x)$ è una matrice $n \times n$ non singolare in un dominio $D \subseteq \mathbb{R}^n$ contenente α .

Si considerano quindi metodi iterativi della forma

$$x^{(k+1)} = \phi(x^{(k)}), \quad k = 0, 1, \dots, \tag{4.31}$$

per i quali si definisce l'ordine di convergenza come in 4.1 cambiando nella Definizione 4.1.1 i valori assoluti in norme di vettori.

Se esistono continue le derivate prime delle funzioni ϕ_i , introducendo la matrice jacobiana

$$\Phi(x) = \begin{pmatrix} \frac{\partial \phi_1}{\partial x_1} & \dots & \frac{\partial \phi_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi_n}{\partial x_1} & \dots & \frac{\partial \phi_n}{\partial x_n} \end{pmatrix}, \tag{4.32}$$

si può generalizzare il Teorema 4.4.1 e dimostrare il seguente teorema di convergenza locale.

Teorema 4.6.1 *Se α è un punto fisso di $\phi(x)$, condizione sufficiente per la convergenza ad α del metodo (4.31) è che esistano due numeri positivi K e ρ , con $K < 1$, tali che si abbia*

$$\|\Phi(x)\| \leq K, \quad \forall x \in D_\rho = \{x \mid \|x - \alpha\| \leq \rho\}; \quad (4.33)$$

purché $x^{(0)}$ sia scelto in D_ρ ; in tal caso α è l'unico punto fisso di ϕ in D_ρ .

Il Teorema 4.4.2 non può essere direttamente esteso in \mathbb{R}^n ; tuttavia se esistono continue le derivate seconde delle ϕ_i , si può dimostrare che, se il metodo (4.31) converge ad un punto fisso α , condizione sufficiente perché converga linearmente è che sia

$$\Phi(\alpha) \neq \mathbf{O}.$$

Mentre, se $\Phi(\alpha) = \mathbf{O}$, la convergenza è almeno quadratica ed è esattamente quadratica se risulta non nulla almeno una delle matrici hessiane

$$H_i(\alpha) = \left(\begin{array}{cccc} \frac{\partial^2 \phi_i}{\partial x_1^2} & \frac{\partial^2 \phi_i}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 \phi_i}{\partial x_1 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 \phi_i}{\partial x_n \partial x_1} & \frac{\partial^2 \phi_i}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 \phi_i}{\partial x_n^2} \end{array} \right)_{x=\alpha}, \quad i = 1, \dots, n.$$

Per estendere il metodo di Newton al sistema non lineare (4.29), supponiamo che le funzioni f_i siano derivabili con continuità rispetto a ciascuna variabile e che la matrice jacobiana

$$J(x) = \left(\begin{array}{ccc} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{array} \right)$$

sia non singolare in un dominio contenente nel suo interno una soluzione α del sistema (4.29).

Specializzando la funzione di iterazione (4.30) nella forma

$$\phi(x) = x - J^{-1}(x)f(x),$$

si ha il sistema $x = x - J^{-1}(x)f(x)$ equivalente a (4.29), da cui il metodo iterativo di Newton o di *Newton-Raphson*

$$x^{(k+1)} = x^{(k)} - J^{-1}(x^{(k)})f(x^{(k)}), \quad k = 0, 1, \dots$$

Nell'uso pratico del metodo l'iterata $x^{(k+1)}$ si ricava dalla soluzione del sistema lineare

$$J(x^{(k)})d^{(k)} = -f(x^{(k)}), \quad k = 0, 1, \dots, \quad (4.34)$$

dove $d^{(k)} = x^{(k+1)} - x^{(k)}$.

Sull'ordine e sulla convergenza locale del metodo si ha il seguente teorema.

Teorema 4.6.2 *Sia $\alpha \in D$ soluzione di (4.29) e le funzioni f_i siano della classe $C^3(D)$ e tali che $J(x)$ sia non singolare in D ; allora l'ordine di convergenza del metodo di Newton è almeno $p = 2$ e si ha convergenza per ogni scelta di $x^{(0)}$ in un opportuno dominio contenente α .*

DIMOSTRAZIONE. Tenuto conto della forma della funzione di iterazione ϕ , la colonna j -esima della matrice (4.32) risulta:

$$\begin{aligned} \frac{\partial \phi}{\partial x_j} &= \frac{\partial x}{\partial x_j} - \frac{\partial}{\partial x_j} [J^{-1}(x)f(x)] \\ &= \frac{\partial x}{\partial x_j} - J^{-1}(x) \frac{\partial f(x)}{\partial x_j} - \frac{\partial J^{-1}(x)}{\partial x_j} f(x) \\ &= -\frac{\partial J^{-1}(x)}{\partial x_j} f(x); \end{aligned}$$

infatti i vettori $\frac{\partial x}{\partial x_j}$ e $J^{-1}(x) \frac{\partial f(x)}{\partial x_j}$ coincidono con la j -esima colonna della matrice identica. Ne segue che per $x = \alpha$ si ha

$$\left(\frac{\partial \phi}{\partial x_j} \right)_{x=\alpha} = - \left(\frac{\partial J^{-1}(x)}{\partial x_j} \right)_{x=\alpha} f(\alpha) = 0, \quad j = 1, 2, \dots, n,$$

quindi è

$$\Phi(\alpha) = \mathbf{O}, \quad (4.35)$$

cioè l'ordine di convergenza del metodo è almeno $p = 2$.

Dalla continuità di $\Phi(x)$ e dalla (4.35) segue infine, per un assegnato $K < 1$, l'esistenza di un dominio $D_\rho = \{x \mid \|x - \alpha\| \leq \rho\} \subset D$ in cui vale la condizione (4.33) e quindi il metodo converge $\forall x^{(0)} \in D_\rho$. \square

Nel caso di una lenta variazione della matrice jacobiana $J(x)$, si può ricorrere anche ora ad un *metodo di Newton semplificato* della forma

$$J(x^{(0)})d^{(k)} = -f(x^{(k)}), \quad k = 0, 1, \dots, \quad (4.36)$$

dove la matrice jacobiana è valutata una sola volta in corrispondenza ad una buona approssimazione iniziale $x^{(0)}$. Se il metodo (4.36) converge, la convergenza è in generale lineare.

Un modo di evitare il calcolo dell'intera matrice $J(x)$ ad ogni passo, consiste nel considerare l' i -esima equazione del sistema (4.29) come una equazione nella sola incognita x_i ed applicare a ciascuna equazione del sistema il metodo di Newton per equazioni in una incognita. Supposto che l'ordinamento delle equazioni sia tale che, in un dominio contenente α , si abbia

$$\frac{\partial f_i(x_1, \dots, x_n)}{\partial x_i} \neq 0, \quad i = 1, 2, \dots, n,$$

si ottiene

$$x_i^{(k+1)} = x_i^{(k)} - \frac{f_i(x_1^{(k)}, \dots, x_n^{(k)})}{\frac{\partial f_i(x_1^{(k)}, \dots, x_n^{(k)})}{\partial x_i}}, \quad i = 1, \dots, n; \quad k = 0, 1, \dots \quad (4.37)$$

Questo metodo, detto *metodo non lineare di Jacobi-Newton*, è un metodo iterativo ad un punto e può esprimersi nella forma matriciale

$$x^{(k+1)} = x^{(k)} - D^{-1}(x^{(k)})f(x^{(k)}), \quad k = 0, 1, \dots,$$

dove $D = \text{diag}(\frac{\partial f_1}{\partial x_1}, \frac{\partial f_2}{\partial x_2}, \dots, \frac{\partial f_n}{\partial x_n})$; la convergenza, lineare, è garantita se vale la condizione (4.33), tenendo conto che nel metodo in esame si ha $\phi(x) = x - D^{-1}(x)f(x)$.

Una variante del metodo (4.37) si ottiene introducendo la tecnica iterativa di Gauss-Seidel, della forma

$$x_i^{(k+1)} = x_i^{(k)} - \frac{f_i(x_1^{(k+1)}, \dots, x_{i-1}^{(k+1)}, x_i^{(k)}, \dots, x_n^{(k)})}{\frac{\partial f_i(x_1^{(k+1)}, \dots, x_{i-1}^{(k+1)}, x_i^{(k)}, \dots, x_n^{(k)})}{\partial x_i}}, \quad (4.38)$$

$$i = 1, 2, \dots, n, \quad k = 0, 1, \dots,$$

dove si sono utilizzate, per il calcolo di $x_i^{(k+1)}$, le prime $i - 1$ componenti già note dell'iterata $x^{(k+1)}$ in corso di calcolo.

Il metodo (4.38) è noto come *metodo non lineare di Gauss-Seidel*.

Un modo di accelerare la convergenza lineare di questo metodo è quello di introdurre un fattore di rilassamento $\omega \in]0, 2[$ ponendo

$$x_i^{(k+1)} = x_i^{(k)} - \omega F_i^{(k+1,k)}, \quad i = 1, 2, \dots, n, \quad k = 0, 1, \dots, \quad (4.39)$$

dove $F_i^{(k+1,k)}$ indica la frazione che compare in (4.38). Alla (4.39), con $\omega > 1$ si dà il nome di *metodo di sovrarilassamento di Newton*.

Una classe di metodi, mutuati dal metodo di Newton-Raphson e che possono essere interpretati come una estensione del metodo delle secanti in \mathbb{R}^n , si ottiene sostituendo in (4.34) la matrice $J(x)$ con una sua approssimazione discreta; si ha così l'evidente vantaggio di non dovere calcolare le n^2 derivate parziali costituenti $J(x)$.

Per approssimare $J(x^{(k)})$ si può introdurre la matrice $\Delta(x^{(k)})$, reale di ordine n , i cui elementi sono dati da

$$\left(\Delta(x^{(k)})\right)_{ij} = \frac{f_i(x^{(k)} + h_j^{(k)} e^{(j)}) - f_i(x^{(k)})}{h_j^{(k)}},$$

$$i, j = 1, 2, \dots, n, \quad k = 0, 1, \dots,$$

essendo $h_j^{(k)}$ opportuni scalari, infinitesimi al crescere di k ; ad esempio si può assumere $h_j^{(k)} = f_j(x^{(k)})$ oppure, più semplicemente, $h_j^{(k)} = x_j^{(k-1)} - x_j^{(k)}$. In generale questi metodi presentano un ordine di convergenza $p \leq 2$.

4.7 Zeri di polinomi

L'equazione (4.1) sia algebrica di grado $m \geq 2$, cioè della forma

$$P(x) = a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0 = 0 \quad (a_m \neq 0) \quad (4.40)$$

con i coefficienti a_i reali. All'equazione (4.40) sono applicabili tutti i metodi visti nei paragrafi precedenti, ma le particolari proprietà della classe dei polinomi forniscono ulteriori procedimenti per la localizzazione e l'approssimazione delle radici. Uno strumento utile a tale scopo è la *successione di Sturm*.

Definizione 4.7.1 *Si dice successione di Sturm per il polinomio $P(x)$ con $a_i \in \mathbb{R}$, una successione di polinomi reali*

$$P_0(x) := P(x), \quad P_1(x), \dots, \quad P_k(x), \quad (k \leq m) \quad (4.41)$$

per cui valgono le seguenti proprietà:

1. $P_k(x)$ non ha zeri reali,
2. $\alpha \in \mathbb{R}$, $P_r(\alpha) = 0$ per $1 \leq r \leq k-1$ implica $P_{r-1}(\alpha)P_{r+1}(\alpha) < 0$,
3. $\alpha \in \mathbb{R}$, $P_0(\alpha) = 0$ implica $P'_0(\alpha)P_1(\alpha) > 0$.

Dalla proprietà 3 della definizione segue che gli zeri reali di $P(x)$ sono tutti semplici.

Teorema 4.7.1 (di Sturm) *Se la (4.41) è una successione di Sturm il numero degli zeri del polinomio $P(x)$ nell'intervallo $a < x \leq b$ è dato da $V(a) - V(b)$, dove $V(x)$ è il numero delle variazioni di segno presenti nella successione dei valori non nulli assunti dalla (4.41) nel punto x .*

DIMOSTRAZIONE. Al variare di x la funzione $V(x)$ può cambiare di valore soltanto se qualche polinomio $P_i(x)$ della successione (4.41) cambia di segno e pertanto si annulla in un punto $x = \alpha$. Per la proprietà 2 non possono annullarsi in $x = \alpha$ due polinomi consecutivi, inoltre, se $P_i(\alpha) = 0$, non può essere $i = k$ per la proprietà 1. Quindi deve essere $0 \leq i \leq k-1$.

Distinguiamo i due casi $i > 0$ e $i = 0$.

Nel primo caso, tenuto conto della proprietà 2 e della continuità, i polinomi $P_{i-1}(x)$, $P_i(x)$ e $P_{i+1}(x)$, calcolati nel punto $x = \alpha$ e in punti sufficientemente vicini $x = \alpha - \delta$ e $x = \alpha + \delta$ ($\delta > 0$), presentano una sola delle possibili distribuzioni di segni illustrate nei due quadri seguenti:

	$\alpha - \delta$	α	$\alpha + \delta$		$\alpha - \delta$	α	$\alpha + \delta$
P_{i-1}	+	+	+		-	-	-
P_i	\pm	0	\pm		\pm	0	\pm
P_{i+1}	-	-	-		+	+	+

Quindi si ha comunque $V(\alpha - \delta) = V(\alpha) = V(\alpha + \delta)$ e il valore di $V(x)$ non cambia quando x attraversa uno zero di $P_i(x)$.

Nel secondo caso, tenendo conto della proprietà 3, si hanno le due possibilità:

	$\alpha - \delta$	α	$\alpha + \delta$		$\alpha - \delta$	α	$\alpha + \delta$
P_0	+	0	-		-	0	+
P_1	-	-	-		+	+	+

Perciò se α è radice di $P(x) = 0$ nella successione (4.41) si ha

$$V(\alpha - \delta) - V(\alpha + \delta) = 1.$$

In particolare se $\alpha = a$, dalle ultime due colonne di ciascuno dei quadri relativi a P_0 e P_1 , si ha $V(a) - V(a + \delta) = 0$, mentre se $\alpha = b$, dalle prime due colonne si ha $V(b - \delta) - V(b) = 1$.

Quindi il valore di $V(x)$ si abbassa di una unità ogni volta che x attraversa crescendo uno zero di $P(x)$, eccettuato il caso in cui tale zero coincida con $x = a$.

Da quanto sopra resta dimostrata la tesi. \square

Corollario 4.7.1 *Se nella (4.41) si ha $k = m$ e se inoltre tutti i polinomi hanno i coefficienti dei termini di grado massimo dello stesso segno, allora l'equazione (4.40) ha m radici reali e distinte e viceversa.*

Per costruire effettivamente una successione di Sturm per il polinomio (4.40), si pone

$$\begin{aligned} P_0(x) &:= P(x), & P_1(x) &:= P'(x), \\ P_{r-1}(x) &= P_r(x)Q_r(x) - P_{r+1}(x), & r &= 1, 2, \dots, \end{aligned} \quad (4.42)$$

dove $Q_r(x)$ e $-P_{r+1}(x)$ sono rispettivamente quoziente e resto della divisione $P_{r-1}(x)/P_r(x)$, $r = 1, 2, \dots$.

Il processo (4.42) ha termine, poiché il grado dei polinomi decresce al crescere dell'indice e perciò per un certo $k \leq m$ risulta

$$P_{k-1}(x) = P_k(x)Q_k(x).$$

Si riconosce nel processo (4.42) il noto *algoritmo di Euclide* che fornisce il massimo comune divisore di $P_0(x)$ e $P_1(x)$, cioè si ha

$$P_k(x) = \text{M.C.D. } \{P(x), P'(x)\}. \quad (4.43)$$

Dalle (4.42) e (4.43) segue che, nel caso in cui $P(x)$ e $P'(x)$ non abbiano zeri reali in comune, i polinomi $P_0(x), P_1(x), \dots, P_k(x)$ formano una successione di Sturm per $P(x)$. Infatti, nell'ipotesi fatta e per la (4.43), $P_k(x)$ non ha zeri reali (proprietà 1).

Se $P_r(\alpha) = 0$ con $1 \leq r \leq k - 1$, allora si ha $P_{r-1}(\alpha) = -P_{r+1}(\alpha) \neq 0$ come segue dalla (4.42) e quindi $P_{r-1}(\alpha)P_{r+1}(\alpha) < 0$ (proprietà 2).

Infine, dalla definizione di $P_0(x)$ e di $P_1(x)$, se $P_0(\alpha) = 0$, segue $P'_0(\alpha)P_1(\alpha) = (P'_0(\alpha))^2 > 0$ (proprietà 3).

Se $P(x)$ e $P'(x)$ hanno zeri reali in comune, questi, per la (4.43), sono tutti zeri di $P_k(x)$ con la stessa molteplicità che hanno come zeri di $P'(x)$ e la (4.42) non fornisce una successione di Sturm.

In tal caso si può verificare che la successione

$$\frac{P_0(x)}{P_k(x)}, \frac{P_1(x)}{P_k(x)}, \dots, \frac{P_k(x)}{P_k(x)}, \quad (4.44)$$

è una successione di Sturm per il polinomio $P(x)/P_k(x)$ che ha tanti zeri semplici quanti sono gli zeri distinti di $P(x)$.

La differenza $V(a) - V(b)$ valutata per la successione (4.44) fornisce allora il numero delle radici reali e distinte (indipendentemente dalla loro molteplicità) dell'equazione $P(x) = 0$ sull'intervallo $a < x \leq b$.

Una successione di Sturm può essere usata per individuare un intervallo $[a, b]$ contenente una sola radice reale α di una equazione algebrica e quindi, con successivi dimezzamenti dell'intervallo come nel metodo di bisezione, si può approssimare α con qualsivoglia accuratezza anche se piuttosto lentamente.

L'approssimazione di uno zero di $P(x)$, che sia stato separato per esempio con una successione di Sturm, può essere fatta usando uno qualunque dei metodi esposti precedentemente. In particolare, si può utilizzare il metodo di Newton

$$x_{n+1} = x_n - \frac{P(x_n)}{P'(x_n)},$$

che richiede ad ogni passo il calcolo di $P(x_n)$ e $P'(x_n)$. Il calcolo di un polinomio in un punto $x = c$ può farsi agevolmente; infatti, posto

$$P(x) = (x - c)Q(x) + R, \quad (4.45)$$

si ha $P(c) = R$ e l'espressione di R insieme a quella dei coefficienti del quoziente $Q(x)$ si ottiene in funzione dei coefficienti a_i di $P(x)$ usando l'*algoritmo di Ruffini-Horner*, cioè ponendo

$$Q(x) = b_{m-1}x^{m-1} + b_{m-2}x^{m-2} + \dots + b_0, \quad R = b_{-1},$$

e calcolando successivamente (in base al principio di identità applicato alla (4.45))

$$\begin{aligned} b_{m-1} &= a_m, \\ b_{m-2} &= b_{m-1}c + a_{m-1}, \\ &\dots \quad \dots \\ b_i &= b_{i+1}c + a_{i+1}, \\ &\dots \quad \dots \\ b_{-1} &= b_0c + a_0 = R. \end{aligned} \quad (4.46)$$

L'algoritmo (4.46) ottimizza il calcolo di $P(c) = R$ riducendolo ad m moltiplicazioni ed altrettante addizioni ed equivale a calcolare per $x = c$ il polinomio $P(x)$ scritto nella forma

$$((\dots((a_m x + a_{m-1})x + a_{m-2})x + a_{m-3})x + \dots + a_1)x + a_0.$$

Con lo stesso algoritmo si può calcolare anche $P'(x)$ per $x = c$, avendosi dalla (4.45)

$$P'(x) = (x - c)Q'(x) + Q(x), \quad P'(c) = Q(c).$$

Un'operazione possibile nelle equazioni algebriche è la cosiddetta *deflazione* che consiste nell'utilizzare una radice α , già trovata, per abbassare il grado dell'equazione, in base al fatto che le rimanenti radici di $P(x) = 0$ sono quelle dell'equazione di grado $m - 1$

$$S_1(x) = \frac{P(x)}{x - \alpha} = 0.$$

Di questa equazione si può calcolare una nuova radice per poi ripetere una deflazione e così via, fino ad ottenere, almeno in linea di principio, tutte le radici.

Tale procedimento è però sconsigliabile in pratica, infatti quand'anche si conoscesse α esattamente, il calcolo dei coefficienti di $S_1(x)$ introduce inevitabilmente degli errori che possono privare di attendibilità i risultati successivi.

Per evitare queste difficoltà si può ricorrere alla cosiddetta *deflazione implicita* in cui ogni radice già calcolata viene utilizzata per calcolarne un'altra, operando sempre sul polinomio originale $P(x)$, mediante una modifica del metodo di Newton.

Siano $\alpha_1, \alpha_2, \dots, \alpha_m$ gli zeri di $P(x)$; partendo dalla nota decomposizione

$$P(x) = a_m(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_m),$$

e supposto di conoscere le radici $\alpha_1, \alpha_2, \dots, \alpha_r$ ($1 \leq r < m$), sia $S_r(x)$ il polinomio di grado $m - r$ che si otterrebbe operando senza errori r deflazioni; sarà allora

$$S_r(x) = \frac{P(x)}{(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_r)}.$$

Da questa segue che si può applicare il metodo di Newton all'equazione $S_r(x) = 0$, senza conoscere esplicitamente il polinomio $S_r(x)$, ottenendo

$$x_{n+1} = x_n - \frac{1}{\frac{P'(x_n)}{P(x_n)} - \left(\frac{1}{x_n - \alpha_1} + \cdots + \frac{1}{x_n - \alpha_r} \right)}. \quad (4.47)$$

La (4.47) prende il nome di *metodo di Newton con deflazione implicita*.

Il metodo di Newton può essere usato anche nel campo complesso.

Esistono anche metodi efficienti che forniscono simultaneamente tutte le radici reali o complesse di una equazione algebrica.

Uno dei più usati consiste nel considerare le radici dell'equazione $P(x) = 0$ come autovalori della matrice di Frobenius associata a $P(x)$ (cfr. Capitolo 2) e nell'applicare per il calcolo degli autovalori il metodo QR descritto nel Capitolo 5.

Un altro metodo, che consente di approssimare tutte le radici di una equazione algebrica come soluzione di uno speciale sistema non lineare, è quello esposto nell'Esempio 6.8.3.

4.8 Complementi ed esempi

I metodi iterativi possono essere analizzati in base alla loro *efficienza*. Anche se questo concetto non ha una definizione formale unica, esso è legato essenzialmente all'ordine p del metodo (e quindi alla sua capacità di ridurre l'errore ad ogni iterazione) ed al numero v di valutazioni di funzione per ogni iterazione (e quindi alla sua complessità computazionale). Molto semplicemente, ad esempio, l'efficienza E può essere definita da

$$E = \frac{p}{v}.$$

Limitandosi al caso di radici semplici, per il metodo delle secanti (4.9) si ha $E \simeq 1.618$ avendosi $v = 1$ per ogni iterazione (tranne che per la prima), mentre risulta $E = 1$ sia per la regola falsi (4.8) sia per il metodo di Newton (4.22). Analogamente si ha $E = 1$ per il metodo di Steffensen

$$x_{n+1} = x_n - \frac{(f(x_n))^2}{f(x_n) - f(x_n - f(x_n))} \quad (4.48)$$

ottenuto dalla (4.20) ove si ponga $\phi(x) = x - f(x)$: infatti, come si verifica facilmente, per esso è $p = 2$ e $v = 2$.

Il vantaggio del metodo delle secanti rispetto ad altri metodi, ed in particolare rispetto al metodo di Newton, può risultare decisivo nel caso di funzioni complicate e di non immediata derivabilità.

Esempio 4.8.1 L'equazione

$$f(x) = x \left(\arccos \frac{1}{1+x} \right)^{\sin(\log(2x+1))} - \frac{1}{\pi} = 0$$

ha una radice in $]0, 1[$.

Assumendo $x_0 = 0.6$ e $x_1 = 0.4$, il metodo delle secanti produce con 6 valutazioni di $f(x)$, $x_5 = 0.30212622627176\dots$ dove le 14 cifre riportate sono esatte.

Il metodo di Newton, applicato con $x_0 = 0.4$, fornisce lo stesso risultato con l'iterazione x_4 : sono però necessarie otto valutazioni di funzione di cui quattro di $f'(x)$, funzione, questa, dotata di una maggiore complessità computazionale. \square

4.8.1 Radici multiple

Il metodo di Newton e le sue varianti (metodo delle secanti, metodo di Steffensen (4.48)) mostrano una riduzione dell'ordine di convergenza, e quindi dell'efficienza, in presenza di radici multiple.

Tuttavia, nota la molteplicità s di una radice α di $f(x) = 0$, è possibile, generalizzando la (4.28), costruire schemi iterativi aventi ordine di convergenza superiore al primo.

Esempio 4.8.2 Si consideri la funzione di iterazione

$$\phi(x) = x - \frac{2f^{(s-1)}(x)}{2f^{(s)}(x) + h(x)f^{(s+1)}(x)}$$

dove si intende $f^{(0)}(x) = f(x)$ e dove, per definizione di molteplicità di una radice, si ha $f^{(r)}(\alpha) = 0$, $r = 0, 1, \dots, s-1$, $f^{(s)}(\alpha) \neq 0$. La funzione $h(x)$ può essere determinata in modo che il metodo $x_{n+1} = \phi(x_n)$, $n = 0, 1, \dots$, abbia un ordine di convergenza superiore al primo.

Poiché risulta

$$\phi'(\alpha) = 1 - \frac{2f^{(s)}(\alpha)}{2f^{(s)}(\alpha) + h(\alpha)f^{(s+1)}(\alpha)},$$

il metodo ha ordine di convergenza almeno due se $\phi'(\alpha) = 0$, ossia se $h(\alpha) = 0$. Derivando ancora viene

$$\phi''(\alpha) = \frac{f^{(s)}(\alpha) [2f^{(s+1)}(\alpha) + h'(\alpha)f^{(s+1)}(\alpha)] - f^{(s+1)}(\alpha)f^{(s)}(\alpha)}{[f^{(s)}(\alpha)]^2}.$$

Pertanto l'ordine di convergenza è almeno tre se $\phi''(\alpha) = 0$, ossia se $h'(\alpha) = -1$.

In particolare se si assume $h(x) = -f^{(s-1)}(x)/f^{(s)}(x)$, le precedenti condizioni sono soddisfatte e poiché, come si constata facilmente, risulta $\phi'''(\alpha) \neq 0$, il metodo ha ordine di convergenza esattamente $p = 3$.

Per esempio, applicando il metodo alla funzione $f(x) = (x^3 + x - 1)^2$ che ha uno zero di molteplicità due in $]0, 1[$, ponendo $s = 2$ e $x_0 = 1$ si ottiene:

$$\begin{aligned} x_1 &= 0.73170731 \dots, \\ x_2 &= 0.68269610 \dots, \\ x_3 &= 0.68232780 \dots \end{aligned}$$

(in x_3 le otto cifre decimali sono esatte). L'algoritmo di Newton (4.22), che in questo caso perde la convergenza quadratica, con $x_0 = 1$ fornisce otto cifre decimali esatte in x_{27} .

Applicando il metodo alla funzione $f(x) = x^3 + x - 1$ con $s = 1$ e $x_0 = 1$, si ottiene:

$$\begin{aligned} x_1 &= 0.69230769 \dots, \\ x_2 &= 0.68232811 \dots, \\ x_3 &= 0.68232780 \dots \end{aligned}$$

In questo caso, per il Teorema 4.4.2, il fattore di convergenza è

$$C = \frac{1}{6} |\phi'''(\alpha)| = \frac{6\alpha^2 - 1}{3\alpha^2 + 1} \simeq 0.3122,$$

dove si è approssimato α con $0.68232780 \simeq x_3$.

Alternativamente C può essere stimato dalle iterazioni fatte, in base alla (4.5),

$$C \simeq \frac{|x_3 - x_2|}{|x_3 - x_1|^3} \simeq 0.3119.$$

Per questo metodo con $h(x) = -f^{(s-1)}(x)/f^{(s)}(x)$ risulta $E = 1$, indipendentemente da s . Tuttavia, a parte la necessaria conoscenza a priori di s , può costituire un inconveniente l'uso delle derivate successive di $f(x)$. \square

4.8.2 Il caso delle equazioni algebriche

L'equazione algebrica (4.2) è dotata di importanti relazioni fra i coefficienti, reali o complessi, e le radici α_i , $i = 1, 2, \dots, m$. Fra queste si riportano le due seguenti.

Posto

$$s_i = \sum_{1 \leq k_1 < k_2 < \dots < k_i \leq m} \alpha_{k_1} \alpha_{k_2} \cdots \alpha_{k_i}, \quad i = 1, 2, \dots, m,$$

dove in particolare è $s_1 = \alpha_1 + \alpha_2 + \dots + \alpha_m$ e $s_m = \alpha_1 \alpha_2 \cdots \alpha_m$,

$$S_k = \sum_{i=1}^m \alpha_i^k, \quad k = 1, 2, \dots,$$

si possono dimostrare le seguenti relazioni:

$$s_i = (-1)^i \frac{a_{m-i}}{a_m}, \quad i = 1, 2, \dots, m; \quad (4.49)$$

$$\sum_{i=1}^k a_{m-i+1} S_{k-i+1} = -k a_{m-k}, \quad k = 1, 2, \dots, m. \quad (4.50)$$

Le (4.50) sono dette *formule di Girard-Newton*.

Nel caso di equazioni a coefficienti interi, dalla (4.49) con $i = m$, segue facilmente che le eventuali radici razionali, della forma p/q , con $q \neq 0$, p e q interi e primi fra loro, sono da ricercarsi nell'insieme

$$\left\{ \frac{\pm \text{divisori interi di } a_0}{\pm \text{divisori interi di } a_m} \right\}.$$

Le (4.49) e (4.50) consentono, inoltre, di ricondurre la risoluzione di una particolare classe di sistemi algebrici non lineari alla risoluzione di una equazione algebrica.

Esempio 4.8.3 Si consideri il sistema

$$\sum_{i=0}^2 b_i x_i^k = m_k, \quad k = 0, 1, 2, 3, 4, 5,$$

nelle incognite $b_0, b_1, b_2, x_0, x_1, x_2$. Tale sistema è formalmente analogo al sistema (7.4) con $n = 2$. Dalle prime tre equazioni si ricavano b_0, b_1, b_2 , (implicate linearmente) in funzione di $x_0, x_1, x_2, m_0, m_1, m_2$, ottenendo

$$b_0 = \frac{m_0 x_1 x_2 - m_1 (x_1 + x_2) + m_2}{(x_1 - x_0)(x_2 - x_0)},$$

$$b_1 = \frac{m_0 x_0 x_2 - m_1 (x_0 + x_2) + m_2}{(x_0 - x_1)(x_2 - x_1)},$$

$$b_2 = \frac{m_0 x_0 x_1 - m_1 (x_0 + x_1) + m_2}{(x_0 - x_2)(x_1 - x_2)}.$$

Sostituendo nelle ultime tre equazioni e ponendo $s_1 = x_0 + x_1 + x_2$, $s_2 = x_0 x_1 + x_0 x_2 + x_1 x_2$, $s_3 = x_0 x_1 x_2$, si ottiene il sistema lineare in s_1, s_2, s_3 ,

$$\begin{pmatrix} m_2 & -m_1 & m_0 \\ m_3 & -m_2 & m_1 \\ m_4 & -m_3 & m_2 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} m_3 \\ m_4 \\ m_5 \end{pmatrix}.$$

Ponendo $m_k = \frac{b^{k+1} - a^{k+1}}{k+1}$, si può verificare che la matrice dei coefficienti è non singolare, perciò, risolvendo, si trova

$$s_1 = 3 \left(\frac{b+a}{2} \right), \quad s_2 = 3 \left(\frac{b+a}{2} \right)^2 - \frac{3}{5} \left(\frac{b-a}{2} \right)^2,$$

$$s_3 = \left[\left(\frac{b+a}{2} \right)^2 - \frac{3}{5} \left(\frac{b-a}{2} \right)^2 \right] \frac{b+a}{2}.$$

Evidentemente, per le (4.49), i numeri x_0, x_1, x_2 , sono le radici dell'equazione algebrica

$$x^3 - s_1 x^2 + s_2 x - s_3 = 0.$$

Cercando le soluzioni di questa equazione fra i divisori di s_3 , si trova che $(b+a)/2$ è radice; ne segue quindi

$$x_0 = \frac{b+a}{2} - \frac{b-a}{2} \sqrt{\frac{3}{5}}, \quad x_1 = \frac{b+a}{2}, \quad x_2 = \frac{b+a}{2} + \frac{b-a}{2} \sqrt{\frac{3}{5}}.$$

Con tali valori si ricava, poi, facilmente

$$b_0 = b_2 = \frac{5}{18}(b-a), \quad b_1 = \frac{4}{9}(b-a).$$

Si noti che con $a = -1$ e $b = 1$ si ottengono i nodi e i pesi della formula di quadratura di Gauss-Legendre a tre punti (cfr. 7.1 e 7.4).

Un esempio di applicazione delle (4.50) si ha se b_0, b_1, b_2 , sono arbitrariamente prefissati uguali. Dalla prima equazione del sistema si ottiene $b_0 = b_1 = b_2 = m_0/3$, mentre per il calcolo di x_0, x_1, x_2 , possono usarsi le successive tre equazioni. Con $a = -1$ e $b = 1$ si ha:

$$\begin{aligned}x_0 + x_1 + x_2 &= 0 \\x_0^2 + x_1^2 + x_2^2 &= 1 \\x_0^3 + x_1^3 + x_2^3 &= 0.\end{aligned}$$

Questo sistema ha per soluzioni tutte le permutazioni dei tre numeri che sono radici del polinomio algebrico $a_3x^3 + a_2x^2 + a_1x + a_0$, i cui coefficienti possono ricavarsi con le formule di Girard-Newton. Ponendo nelle (4.50) $k = 3$, e, come suggerito dal sistema, $S_1 = 0$, $S_2 = 1$, $S_3 = 0$, si ottiene

$$\begin{aligned}a_0 &= 0 \\a_3 + 2a_1 &= 0 \\a_2 + 3a_0 &= 0.\end{aligned}$$

Fissando arbitrariamente uno dei coefficienti a_i , ad esempio $a_3 = 1$, si ha l'equazione

$$x^3 - \frac{1}{2}x = 0$$

e quindi $x_0 = -x_2 = \frac{\sqrt{2}}{2}$, $x_1 = 0$. □

4.8.3 Una particolare successione di Sturm

Si consideri la matrice tridiagonale hermitiana

$$T = \begin{pmatrix} a_1 & \bar{b}_2 & & \\ b_2 & \ddots & \ddots & \\ & \ddots & \ddots & \bar{b}_n \\ & & b_n & a_n \end{pmatrix}$$

con $b_i \neq 0$, $i = 2, 3, \dots, n$. Il suo polinomio caratteristico $P(\lambda) = \det(A - \lambda I)$ può essere calcolato per ricorrenza come visto in 2.11.5 e si ha

$$P(\lambda) := P_n(\lambda) = (a_n - \lambda)P_{n-1}(\lambda) - |b_n|^2 P_{n-2}(\lambda)$$

con $P_0(\lambda) = 1$ e $P_1(\lambda) = a_1 - \lambda$.

Al riguardo vale il seguente teorema.

Teorema 4.8.1 *Nelle ipotesi fatte su T , la successione*

$$(-1)^n P_n(\lambda), (-1)^{n-1} P_{n-1}(\lambda), \dots, -P_1(\lambda), P_0(\lambda), \quad (4.51)$$

è una successione di Sturm per $P(\lambda)$ e gli zeri di $P(\lambda)$ sono distinti.

Esempio 4.8.4 Si considera la matrice di ordine n

$$T = \begin{pmatrix} 0 & 1/2 & & \\ 1/2 & \ddots & \ddots & \\ & \ddots & \ddots & 1/2 \\ & & 1/2 & 0 \end{pmatrix}.$$

Si verifica facilmente che, per la successione (4.51), risulta $V(-1) - V(1) = n$. Inoltre se n è pari si ha $V(-1) - V(0) = V(0) - V(1) = n/2$, mentre se n è dispari si ha $V(-1) - V(0) = (n+1)/2$ e $V(0) - V(1) = (n-1)/2$: in questo secondo caso $\lambda = 0$ è uno zero di $P(\lambda)$. In entrambi i casi si verifica che $\lambda = 1$ non è uno zero di $P(\lambda)$; ne segue che $|\lambda_i| < 1$, $i = 1, 2, \dots, n$. Questo risultato si trova anche applicando il Teorema 2.8.1 alla matrice T e osservando che $\lambda = 1$ e $\lambda = -1$ non possono essere autovalori in quanto le matrici $T - I$ e $T + I$ sono a predominanza diagonale debole con grafo orientato fortemente connesso e quindi non singolari (cfr. Corollario 2.11.2).

Si può poi dimostrare che gli autovalori di T sono

$$\lambda_i = \cos\left(\frac{i\pi}{n+1}\right), \quad i = 1, 2, \dots, n.$$

□

4.8.4 Il teorema di Newton-Kantorovich

Per il metodo di Newton

$$x^{(k+1)} = x^{(k)} - J^{-1}(x^{(k)})f(x^{(k)}), \quad k = 0, 1, \dots,$$

sussiste il seguente teorema.

Teorema 4.8.2 (di Newton-Kantorovich) Sia $x^{(0)} \in \mathbb{R}^n$ e $f(x) \in C^2(D)$ dove $D = \{x \mid \|x - x^{(0)}\| \leq 2b\} \subset \mathbb{R}^n$. Se risulta:

$$\begin{aligned} \|J^{-1}(x^{(0)})\| &\leq a; \\ \|J^{-1}(x^{(0)})f(x^{(0)})\| &\leq b; \\ \sum_{r=1}^n \left| \frac{\partial^2 f_i}{\partial x_j \partial x_r} \right| &\leq \frac{c}{n}, \quad \forall x \in D, \quad 1 \leq i, j \leq n; \\ abc &\leq \frac{1}{2}; \end{aligned}$$

allora si ha:

$$\begin{aligned} x^{(k)} &\in D, \quad k = 1, 2, \dots; \\ \lim_{k \rightarrow \infty} x^{(k)} &= \alpha, \quad \text{essendo } \alpha \text{ radice di } f(x) = 0; \\ \alpha &\text{ è unica in } D; \\ \|\alpha - x^{(k)}\| &\leq \frac{b(2abc)^{2^k-1}}{2^{k-1}}. \end{aligned}$$

Per quanto non di agevole applicazione, il teorema è importante perché fornisce un legame tra la convergenza e la scelta della stima iniziale $x^{(0)}$ oltreché un risultato di esistenza e unicità della soluzione di $f(x) = 0$ (si noti che nel Teorema 4.6.2 l'esistenza della soluzione è ammessa per ipotesi).

Esempio 4.8.5 Si consideri in \mathbb{R}^3 il sistema

$$f(x) = \begin{pmatrix} 2x_1 + \cos x_2 + \cos x_3 - 1.9 \\ \cos x_1 + 2x_2 + \cos x_3 - 1.8 \\ \cos x_1 + \cos x_2 + 2x_3 - 1.7 \end{pmatrix} = 0.$$

La matrice jacobiana è

$$J(x) = \begin{pmatrix} 2 & -\sin x_2 & -\sin x_3 \\ -\sin x_1 & 2 & -\sin x_3 \\ -\sin x_1 & -\sin x_2 & 2 \end{pmatrix}.$$

Si assume come stima iniziale $x^{(0)} = 0$. Segue, con semplici calcoli,

$$\|J^{-1}(0)\| = \frac{1}{2} = a, \quad \|J^{-1}(0)f(0)\| = \frac{3}{20} = b,$$

dove si è usata la norma ∞ .

Inoltre, per ogni x , si ha $\sum_{r=1}^n |\partial^2 f_i / \partial x_j \partial x_r| \leq 1$, $1 \leq i, j \leq 3$, da cui $c = 3$.

Poiché $abc = 9/40 < \frac{1}{2}$ il metodo di Newton è convergente.

Si noti che risulta $D = \{x \mid \|x\| \leq \frac{3}{10}\}$, per cui i punti di coordinate $\pm \frac{\pi}{2}$ non appartengono a D ; ne viene che $J(x)$ è, per $x \in D$, diagonalmente dominante in senso forte e pertanto la soluzione del sistema (4.34) può essere approssimata, per ogni k , con il metodo iterativo di Jacobi o di Gauss-Seidel.

Il risultato

$$x_4 = \begin{pmatrix} -0.04236905027717\dots \\ -0.09413400044134\dots \\ -0.14733761588854\dots \end{pmatrix}$$

è stato ottenuto con $x^{(0)} = 0$, usando il metodo di Jacobi.