

Si considera il problema di determinare la soluzione dell'equazione

$$f(x) = 0$$

dove $f(x)$ è una funzione definita in un intervallo $[a, b]$, chiuso e limitato. Ogni valore $x^* \in [a, b]$ per cui $f(x^*) = 0$ si dice **radice dell'equazione** $f(x) = 0$ oppure **zero della funzione** $f(x)$.

Esempio. Data l'equazione $x^2 - 4 = 0$, essa ha una radice $x^* = 2$ nell'intervallo $[0, 4]$ e una radice $x^* = -2$ nell'intervallo $[-4, 0]$.

Se non si hanno a disposizione **strumenti analitici** (come nell'esempio), il grafico di $f(x)$ nell'intervallo considerato può fornire in modo rapido una **approssimazione** della soluzione cercata, ma **non** fornisce una **stima dell'accuratezza** dell'approssimazione determinata.

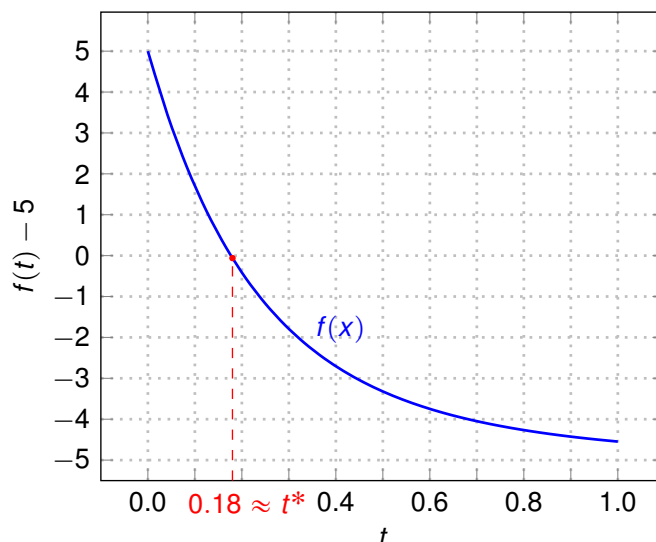
Risulta quindi necessario disporre di metodi numerici in grado di fornire soluzioni con una accuratezza prefissata.

Esempio

Supponiamo che una reazione chimica origini a un certo istante t una concentrazione di un particolare ione descritta dalla legge $f(t) = 7e^{-5t} + 3e^{-2t}$. Ci si chiede per quale istante t^* la concentrazione si dimezza rispetto a quella iniziale $f(0) = 10$, ossia quando $f(t^*) = 5$.

Si tratta di risolvere l'equazione $f(t) - 5 = 0$ in un intervallo $[0, c]$ dove c è un valore per cui $f(c) - 5 < 0$. Per esempio, per $c = 1$, $f(1) - 5 = -4.54$.

Dal grafico della funzione si deduce che $t^* \approx 0.18$. Vale che $f(0.18) - 5 = -0.0610$. L'accuratezza può essere non soddisfacente.



Metodo di bisezione

La tecnica più semplice per determinare lo zero di una funzione continua in $[a, b]$ che assume segno opposto agli estremi dell'intervallo, consiste nella dimostrazione costruttiva del seguente teorema (è un [corollario del Teorema del valor medio](#)).

Teorema degli zeri di una funzione continua

Sia $f(x)$ una funzione continua in $[a, b]$ chiuso e limitato e sia $f(a) \cdot f(b) < 0$. Allora esiste almeno un punto $x^* \in]a, b[$ tale che $f(x^*) = 0$.

Per determinare x^* si costruisce una successione di intervalli

$$\mathcal{I}_1 = [a_1, b_1], \quad \mathcal{I}_2 = [a_2, b_2], \quad \dots, \quad \mathcal{I}_k = [a_k, b_k], \quad \dots$$

tali che

$$\mathcal{I}_1 \supset \mathcal{I}_2 \supset \dots \supset \mathcal{I}_{k-1} \supset \mathcal{I}_k \supset \dots$$

con $f(a_k) \cdot f(b_k) < 0 \quad \forall k = 1, 2, \dots$, e con $a_1 = a$ e $b_1 = b$.

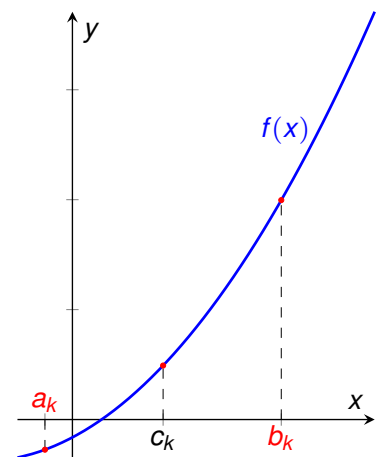
La costruzione degli intervalli avviene nel seguente modo. Si pone $a_1 = a$, $b_1 = b$. Siano $f(a_1) < 0$ e $f(b_1) > 0$ (in caso contrario si determinano le radici di $-f(x) = 0$). Al passo k si calcolano

$$c_k = \frac{a_k + b_k}{2} \quad \text{e} \quad f(c_k).$$

Metodo di bisezione

Se $f(c_k) = 0$, allora $c_k = x^*$ e abbiamo finito. Altrimenti si pone:

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



Pertanto è $f(a_{k+1}) < 0$, $f(b_{k+1}) > 0$ e $\mathcal{I}_{k+1} \subset \mathcal{I}_k$. Inoltre

$$b_{k+1} - a_{k+1} = \frac{1}{2}(b_k - a_k) = \frac{1}{2^2}(b_{k-1} - a_{k-1}) = \dots = \frac{1}{2^k}(b_1 - a_1)$$

Allora in \mathcal{I}_{k+1} giace x^* che è la radice di $f(x) = 0$ e $x^* = c_{k+1} \pm \varepsilon_{k+1}$, dove

$$|c_{k+1} - x^*| = \varepsilon_{k+1} \leq \frac{1}{2^{k+1}}(b - a)$$

Pertanto, fissata la soglia ε tale che $\varepsilon = \frac{1}{2^{k+1}}(b - a)$, il numero c_{k+1} è una approssimazione di x^* entro una tolleranza ε .

Col metodo di bisezione si determina una successione di approssimazioni $\{c_k\}$ di x^* tali che

$$|c_k - x^*| \leq \frac{1}{2^k} (b - a) \quad \forall k \geq 1$$

Per $k \rightarrow \infty$, $\{c_k\}$ converge a x^* con velocità di convergenza pari a quella della successione $\{2^{-k}\}$ (successione geometrica di ragione $1/2$).

Il metodo fornisce un valore maggiorante dell'errore. Da tale valore è possibile determinare il minimo numero di iterazioni k necessarie ad ottenere un errore assoluto che non superi una tolleranza ε prefissata. Infatti k è tale che

$$\frac{1}{2^k} (b - a) \leq \varepsilon$$

$$\Rightarrow 2^k \geq \frac{b - a}{\varepsilon} \Rightarrow k \geq \log_2 \left(\frac{b - a}{\varepsilon} \right) = \frac{\log_{10} \left(\frac{b - a}{\varepsilon} \right)}{\log_{10}(2)}$$

Occorrono allora almeno k passi e $k + 2$ valutazioni di funzione (1 per ogni iterazione più 2 per il passo iniziale).

Esempio

Si vuole risolvere $x^2 - 78.8 = 0$ in $[6, 12]$.

$$f(6) = 36 - 78.8 < 0$$

$$f(12) = 144 - 78.8 > 0$$

k	a_k	b_k	c_k	$f(c_k)$
1	6	12	9	2.2
2	6	9	7.5	-22.55
3	7.5	9	8.25	-10.7375
4	8.25	9	8.625	-4.409375
5	8.625	9	8.8125	-1.139844
6	8.8125	9	8.90625	0.5212891
7	8.8125	8.90625	8.859375	-0.3114746
8	8.859375	8.90625	8.882813	0.1043579

Il valore 8.882813 è una approssimazione della soluzione $\sqrt{78.8} \approx 8.876936408$ tale che

$$|8.882813 - x^*| \leq \frac{1}{2^8} 6 = 0.0234$$

L'errore assoluto è 0.00587.... Occorrono 10 valutazioni di funzione.

Esempio.

Determinare il numero di iterazioni necessarie a risolvere $x^3 + 4x^2 - 10 = 0$ con una tolleranza $\varepsilon = 10^{-5}$ in $[1, 2]$.

Deve essere

$$\frac{1}{2^k}(b-a) \leq 10^{-5}$$
$$2^k \geq \frac{b-a}{10^{-5}} \Rightarrow k \geq \log_2(10^5) \approx 16.6$$

Allora deve essere k almeno uguale a 17. Occorrono $17 + 2$ valutazioni di funzioni.

Osservazioni

- Se a e b sono valori molto vicini, operando in aritmetica finita il calcolo di c con la formula $(a+b)/2$ **può fornire un numero esterno all'intervallo $[a, b]$** .
Conviene calcolare il punto medio con la formula $c = a + (b-a)/2$.
Per esempio, se $a = 0.983$, $b = 0.986$, lavorando in aritmetica finita con tre cifre di precisione decimale e troncamento si ha:

$$\begin{aligned}\text{fl}(a+b) &= \text{fl}(0.983 + 0.986) = 0.196 \cdot 10^1 \\ \text{fl}((a+b)/2) &= \text{fl}(0.196 \cdot 10^1/2) = 0.980\end{aligned}$$

che è esterno all'intervallo $[0.983, 0.986]$. Invece

$$\begin{aligned}\text{fl}(b-a) &= \text{fl}(0.986 - 0.983) = 0.3 \cdot 10^{-2} \\ \text{fl}((b-a)/2) &= \text{fl}(0.3 \cdot 10^{-2}/2) = 0.150 \cdot 10^{-2} \\ \text{fl}(a + (b-a)/2) &= \text{fl}(0.983 + 0.150 \cdot 10^{-2}) \\ &= \text{fl}(0.983 + 0.0015) = 0.984\end{aligned}$$

che è interno all'intervallo $[0.983, 0.986]$.

- Pur essendo $f(a)$ e $f(b)$ rappresentabili sulla macchina, può essere che il loro prodotto, calcolato per verificare il segno, non sia rappresentabile sulla macchina. **È preferibile allora usare la funzione $\text{sgn}(x)$** per esaminare il segno di $f(a)$ e di $f(b)$, dove

$$\text{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$

- L'algoritmo di bisezione realizzato in aritmetica finita **può non avere fine** se si fissa una tolleranza troppo piccola. Se $a_k = 98.5$, $b_k = 98.6$ e $\varepsilon = 0.004$, in aritmetica con tre cifre decimali di precisione e troncamento il test $b_k - a_k < \varepsilon$ non è verificato. Infatti $c_k = a_k + (b_k - a_k)/2 = 98.5 + 0.1/2 = 98.55$, ma $\text{fl}(98.55) = 98.5$. Si genera quindi una **successione di iterati costanti**. Occorre dunque usare il test

$$|b_k - a_k| < \varepsilon + \epsilon \cdot \max\{|b|, |a|\} \quad \text{dove } \epsilon = \text{precisione di macchina}$$

(in modo che si eliminino inconvenienti dovuti al fatto di scegliere una tolleranza troppo piccola, inferiore relativamente rispetto alla precisione di macchina) e prevedere un massimo numero di iterazioni.

Codice Matlab

```
function [c, fc, it] = bisezione(fname, a, b, tol)
% INPUT:  fname (string/fhandle) - Function o function handle che calacola la funzione
%         a, b   (double)        - Estremi dell'intervallo
%         tol    (double)        - Tolleranza di approssimazione della soluzione
% OUTPUT: c      (double)        - Approssimazione della soluzione
%         fc     (double)        - Valore della funzione nell'approssimazione della soluzione
%         it     (integer)       - Numero di iterazioni eseguite
maxit = ceil( log2( (b - a)/tol ) ); it = 0;
fprintf('\nNumero massimo di passi necessari = %d\n', maxit);
fa = feval(fname, a); fb = feval(fname, b);
if ( sign(fa)*sign(fb) > 0 ), error('Intervallo non corretto');
elseif (fa == 0), c = a; fc = fa; return;
elseif (fb == 0), c = b; fc = fb; return;
else
    % Metodo di bisezione
    soglia = tol + eps*max( abs([a; b]) ); arresto = 0;
    while ( ~arresto )
        it = it + 1; c = a + (b - a)*0.5;
        fprintf('\n\tit = %d \tc = %g', it, c); % questa riga serve solo per la visualizz.
        fc = feval(fname, c);
        if ( fc == 0 ), break; end;
        if ( sign(fc)*sign(fa) > 0 )
            a = c; fa = fc;
        else
            b = c; fb = fc;
        end
        arresto = ( abs(b - a) < soglia ) || ( it == maxit );
    end
    fprintf('\n\n');
end
```

Il metodo di bisezione ha la caratteristica positiva di convergere a uno zero di $f(x)$ nella sola ipotesi di $f(x)$ continua con segno opposto agli estremi.

Tuttavia **la convergenza è lenta**: ad esempio, se $b - a = 1$ occorrono circa tre iterazioni per “guadagnare” un’ulteriore cifra decimale corretta nell’approssimazione della soluzione.

Pertanto spesso il metodo di bisezione viene usato solo per individuare l’intervallo entro cui cade lo zero di una funzione, come tecnica di **approssimazione iniziale**. Una volta individuato tale intervallo (entro cui in genere valgono ipotesi più forti), si raggiungono approssimazioni migliori con metodi a convergenza più veloce.

Metodo delle approssimazioni successive

Il problema di determinare lo zero di una funzione in genere **non** si risolve in un numero finito di passi.

Si tratta di generare un **procedimento iterativo**. Esso comporta:

- determinare una **approssimazione iniziale** x_0 alla soluzione x^* di $f(x) = 0$;
- determinare una **relazione funzionale** a partire da $f(x)$;
- generare, a partire da x_0 e valutando la relazione funzionale determinata, una successione $\{x_k\}_{k \in \mathbb{N}}$ di iterati fino ad ottenere la precisione desiderata o fino a che si raggiunge un massimo numero di iterazioni.

Occorre determinare le condizioni per cui $\{x_k\}$ converge per $k \rightarrow \infty$.

Se $\{x_k\}$ converge per ogni $x_0 \in [a, b]$ si parla di **convergenza globale**; se invece si ha convergenza solo per gli x_0 appartenenti a un opportuno intervallo che contiene x^* , allora si parla di **convergenza locale**.

Zeri di una funzione e problema di punto fisso

Il problema di cercare una radice di $f(x) = 0$ è strettamente connesso al problema di determinare la soluzione dell'equazione

$$x = g(x)$$

ossia il valore x^* per cui $x^* = g(x^*)$. Quando esiste, il punto x^* si dice **punto fisso** di $g(x)$.

Infatti, data $f(x)$ in $[a, b]$, sia $\phi(x)$ una funzione che non si annulla ed è limitata in $[a, b]$:

$$0 < |\phi(x)| < \infty \quad \forall x \in [a, b]$$

Definita $g(x) = x - \phi(x)f(x)$, il problema

$$f(x) = 0 \quad x \in [a, b]$$

è equivalente a trovare le soluzioni di

$$x = g(x) \quad x \in [a, b]$$

Zeri di una funzione e problema di punto fisso

Se infatti x^* è punto fisso di $g(x)$, ossia $x^* = g(x^*)$, essendo anche $g(x^*) = x^* - \phi(x^*)f(x^*)$ si ha che

$$f(x^*) = \frac{x^* - g(x^*)}{\phi(x^*)} = 0 \quad \text{perchè } \phi(x^*) \neq 0.$$

Viceversa se $f(x^*) = 0$, allora

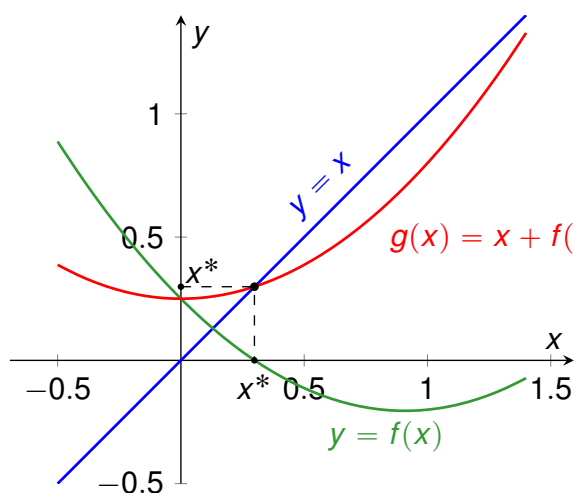
$$g(x^*) = x^* - \phi(x^*)f(x^*) = x^*$$

Pertanto **il problema di determinare uno zero di $f(x)$ in $[a, b]$ è ricondotto a quello di individuare il punto fisso di $g(x)$ in $[a, b]$.**

Geometricamente, determinare il punto fisso significa calcolare l'intersezione tra la retta $y = x$ (bisettrice del I e III quadrante) e la funzione $y = g(x)$.

$$\begin{cases} y = x \\ y = g(x) \end{cases}$$

Una possibile scelta di $\phi(x)$ è $\phi(x) = \pm 1$.



Teorema di esistenza e unicità del punto fisso

Condizioni **sufficienti** ad assicurare esistenza e unicità del punto fisso sono le seguenti.

Teorema

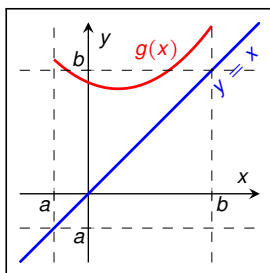
Sia $g(x)$ una funzione che soddisfa le seguenti ipotesi:

- ① $g(x)$ **assume valori in** $[a, b]$;
- ② $g(x)$ **è continua** in $[a, b]$;
- ③ sia L una costante $0 \leq L < 1$ tale che, per ogni $x, y \in [a, b]$, vale che

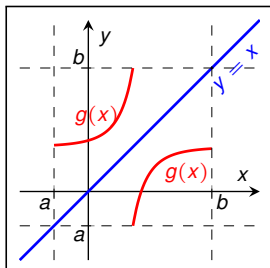
$$|g(x) - g(y)| \leq L|x - y|$$

(ossia $g(x)$ è una **contrazione** in $[a, b]$).

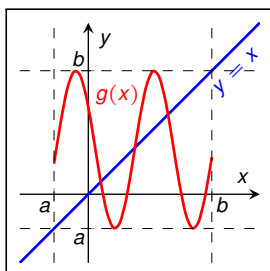
Allora esiste un unico punto fisso x^* di $g(x)$ in $[a, b]$.



Non soddisfa le ipotesi del Teorema:
 $\exists x \in [a, b]$ tale che $g(x) \notin [a, b]$



Non soddisfa le ipotesi del Teorema:
 $g(x)$ non è continua in $[a, b]$



Non soddisfa le ipotesi del Teorema:
 $g(x)$ oscilla troppo, ossia **non** è una contrazione in $[a, b]$

Dimostrazione del Teorema

Esistenza.

Sia $h(x) = x - g(x)$. Poiché $g(x) \in [a, b]$, $g(a) \geq a$ e $g(b) \leq b$. Se $g(a) = a$ o $g(b) = b$, allora si è già trovato il punto fisso. In caso contrario, è $g(a) > a$ e $g(b) < b$. Dunque,

$$h(a) = a - g(a) < 0$$

$$h(b) = b - g(b) > 0$$

e poiché $h(x)$ è continua, per il teorema degli zeri di una funzione continua, esiste $x^* \in]a, b[$ tale che $h(x^*) = 0$ e dunque $x^* = g(x^*)$.

Unicità.

Supponiamo che ci siano due punti fissi x^* e y^* in $[a, b]$. Allora, se $x^* \neq y^*$, segue

$$|x^* - y^*| = |g(x^*) - g(y^*)| \leq L|x^* - y^*| < |x^* - y^*|$$

Ciò è assurdo.

Se $g(x)$ è derivabile in $[a, b]$ con $|g'(x)| \leq L < 1 \ \forall x \in [a, b]$, allora $g(x)$ è una contrazione.

Infatti per il teorema di Lagrange, esiste $\xi \in [a, b]$, tale che

$$\frac{g(x) - g(y)}{x - y} = g'(\xi) \quad \text{con } x < \xi < y$$

Allora

$$|g(x) - g(y)| = |g'(\xi)| |x - y| \leq L |x - y|$$

Il viceversa non è vero poiché $g(x)$ può non essere differenziabile.

Esempi

- ❶ $g(x) = x$ in $[0, 1]$ ha infiniti punti fissi: x è continua in $[0, 1]$ e $x \in [0, 1]$. Ma $g'(x) = 1$ per $x \in [0, 1]$.
- ❷ $g(x) = x - \sin(\pi x)$ in $[0, 1]$ ha due punti fissi: $x = 0$ e $x = 1$. La funzione $g(x)$ è continua in $[0, 1]$, ma $g(x)$ non appartiene sempre a $[0, 1]$ ($g(1/2) = -1/2$).
- ❸ $g(x) = (x^2 - 1)/3$ in $[-1, 1]$: $g(x)$ è continua e $g(x) \in [-1, 1]$. Inoltre $g'(x) = (2/3)x$ e dunque $|g'(x)| \leq 2/3 \ \forall x \in [-1, 1]$.
Dunque **esiste un unico punto fisso in $[-1, 1]$** . Esso deve soddisfare

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

Perciò $x^* = (3 - \sqrt{13})/2$.

In $[3, 4]$, $g(x)$ ha un altro punto fisso $x^* = (3 + \sqrt{13})/2$. Ma $g(x)$ non appartiene a $[3, 4]$ e $g'(4) = 8/3 > 1$. Quindi qui non sono soddisfatte le ipotesi del teorema. Infatti queste sono solo condizioni sufficienti.

Metodo iterativo

Data un'approssimazione iniziale x_0 di x^* , punto fisso di $g(x)$ in $[a, b]$, si genera una successione di iterati mediante il [metodo delle approssimazioni successive](#) (o [metodo del punto fisso](#), o [metodo di iterazione funzionale](#)):

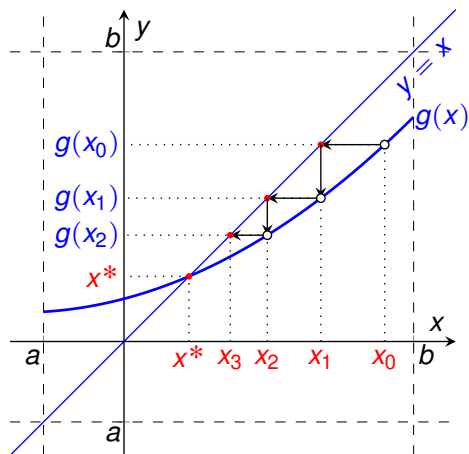
$$x_{k+1} = g(x_k)$$

Se $g(x)$ è continua e la successione $\{x_k\}$ degli iterati è convergente a un punto x^* per $k \rightarrow \infty$, allora x^* è punto fisso di $g(x)$. Infatti

$$x^* = \lim_{k \rightarrow \infty} x_{k+1} = \lim_{k \rightarrow \infty} g(x_k) = g\left(\lim_{k \rightarrow \infty} x_k\right) = g(x^*)$$

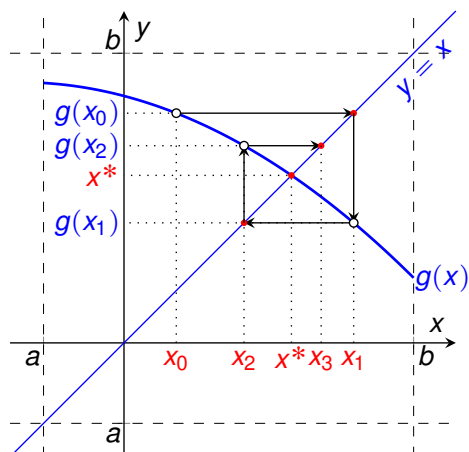
Dal punto di vista geometrico, il metodo dell'iterazione funzionale equivale alla costruzione di una **poligonale orientata** con lati orizzontali e verticali nel piano xy . Le figure seguenti aiutano a individuare le condizioni da imporre a $g(x)$ perché i vertici della poligonale convergano a $(x^*, g(x^*))$.

Metodo iterativo



Convergenza monotona:

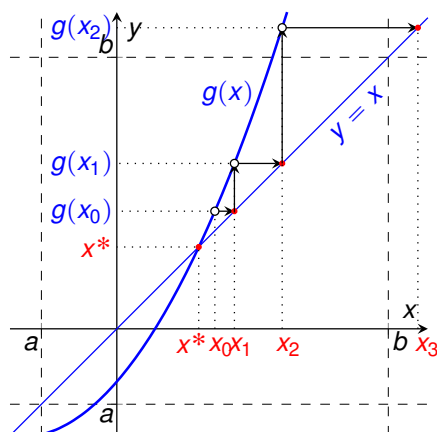
$\{x_k\}$ converge a x^* approssimando sempre per eccesso o sempre per difetto.



Convergenza alternata:

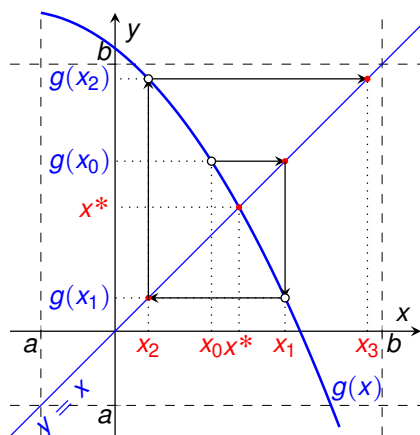
$\{x_k\}$ converge a x^* approssimando alternativamente per eccesso e per difetto.

Metodo iterativo: casi non convergenti



$$g'(x) > 1$$

Per $|g'(x)| > 1$ non c'è convergenza



$$g'(x) < -1$$

Convergenza globale del metodo delle approssimazioni successive

Teorema di convergenza globale

Sia $g(x)$ una funzione definita in un intervallo $[a, b]$ chiuso e limitato. Se

- $g(x) \in [a, b]$,
- $g(x)$ è continua in $[a, b]$,
- $g(x)$ è una contrazione in $[a, b]$,

allora per ogni $x_0 \in [a, b]$ la successione delle approssimazioni successive $\{x_k\}$ generate da $x_k = g(x_{k-1})$, $k = 1, 2, \dots$, converge per $k \rightarrow \infty$ all'unico punto fisso x^* di $g(x)$ in $[a, b]$.

Dimostrazione

Le ipotesi garantiscono l'esistenza e l'unicità del punto fisso x^* della funzione $g(x)$ in $[a, b]$. Poiché inoltre $x_k = g(x_{k-1})$, allora per ogni successione generata a partire da un punto $x_0 \in [a, b]$, l'iterato x_k è **ben definito**, ossia $x_k \in [a, b]$, poiché $g(x_{k-1}) \in [a, b]$.

Si consideri:

$$\begin{aligned} |x_k - x^*| &= |g(x_{k-1}) - g(x^*)| \leq L|x_{k-1} - x^*| = L|g(x_{k-2}) - g(x^*)| \\ &\leq L^2|x_{k-2} - x^*| = L^2|g(x_{k-3}) - g(x^*)| \\ &\vdots \\ &\leq L^k|x_0 - x^*| \end{aligned}$$

Poiché $0 \leq L < 1$, per $k \rightarrow \infty$ si ha $\lim_{k \rightarrow \infty} L^k = 0$ e dunque $\lim_{k \rightarrow \infty} x_k = x^*$. Si può trovare una maggiorazione dell'errore osservando che:

$$|x_0 - x^*| = |x_0 - x_1 + x_1 - x^*| \leq |x_0 - x_1| + L|x_0 - x^*|$$

Da ciò $|x_0 - x^*| \leq \frac{1}{1-L}|x_1 - x_0|$. Allora

$$|x_k - x^*| \leq \frac{L^k}{1-L}|x_1 - x_0|$$

La velocità di convergenza dipende da L . Quanto più L è vicino a 1, tanto più la convergenza è lenta e viceversa.

Codice Matlab

```
function [x, it] = iterazione(gname, x0, tol, maxit)
% Metodo di iterazione funzionale (con visualizz. delle iterate)
% INPUT:  gname (string/fhandle) - Function per funz. d'iterazione
%         x0    (double)         - Iterato iniziale
%         tol    (double)         - Tolleranza per l'arresto
%         maxit  (integer)        - Numero massimo di iterazioni
% OUTPUT: x     (double)         - Approssimazione della soluzione
%         it     (integer)        - Numero di iterazioni eseguite
maxX = 1.0e100; % soglia per la divergenza della successione
it = 0; x = feval(gname, x0);
stop = ( abs(x - x0) <= tol + eps*abs(x0) || abs(x) >= maxX );
while ( ~stop )
    it = it + 1;
    x0 = x;
    x = feval(gname, x);
    fprintf('\nit = %d \tx = %g \tx0 = %g', it, x, x0);
    stop = ( abs(x - x0) <= tol + eps*abs(x0) || it == maxit ...
            || abs(x) >= maxX );
end
if ( it == maxit )
    fprintf('\nRaggiunto il limite massimo di iterazioni.\n');
elseif ( abs(x) >= maxX )
    fprintf('\nSuccessione numericamente divergente.\n');
end
end
```

Esempio

Si considerano vari modi di innescare un procedimento iterativo di punto fisso per poter trovare la soluzione della equazione

$$x^3 + 4x^2 - 10 = 0 \quad x \in [1, 2]$$

In tutti i casi si considera come punto iniziale $x_0 = 1.5$.

$$(a) \quad x = x - x^3 - 4x^2 + 10 = g_1(x);$$

$$(b) \quad x = \left(\frac{10}{x} - 4x \right)^{1/2} = g_2(x) \quad (\text{da } x^3 = 10 - 4x^2);$$

$$(c) \quad x = \frac{1}{2}(10 - x^3)^{1/2} = g_3(x) \quad (\text{da } x^2 = (10 - x^3)/4);$$

$$(d) \quad x = \left(\frac{10}{x+4} \right)^{1/2} = g_4(x) \quad (\text{da } x^3 + 4x^2 = 10);$$

$$(e) \quad x = x - \frac{x^3 + 4x^2 - 10}{3x^2 + 8x} = g_5(x) \quad (\text{da } x = x - f(x)/f'(x)).$$

Esempio

k	(a)	(b)	(c)	(d)	(e)
1	−0.875	0.8165	1.286953768	1.348399725	1.373333333
2	6.732	2.9969	1.402540804	1.367376372	1.365262015
3	−469.7	$(-8.65)^{1/2}$	1.345458374	1.364957015	1.365230014
4	$1.08 \cdot 10^8$	impossibile	1.375170253	1.365264748	1.365230013
⋮	diverge		⋮	⋮	
15			1.365223680	1.365230013	
⋮			⋮		
30			1.365230013		

Non tutte le scelte portano a un metodo convergente (caso (a)) o ben definito (caso (b)). Tuttavia anche se il metodo è convergente, ci sono funzioni che portano a una più rapida convergenza di altre (con il metodo di bisezione sono necessarie 27 valutazioni di funzione).

- (a) $g'_1(x) = 1 - 3x^2 - 8x$, dunque $|g'_1(x)| > 1$ per $x \in [1, 2]$.
- (b) $g_2(x)$ è definita per $-\sqrt{10}/2 \leq x \leq \sqrt{10}/2 \approx 1.581113$, dunque $g_2(x)$ non appartiene a $[1, 2]$.
- (c) $g'_3(x) = -\frac{3}{4} \frac{x^2}{(10 - x^3)^{1/2}}$, dunque $|g'_3(x)| \leq 2.1213$. Tuttavia, se ci si restringe a $[1, 1.5]$, $1.28 \leq g_3(x) \leq 1.5$ e $\max_{[1, 1.5]} |g'_3(x)| = |g'_3(1.5)| \approx 0.6556 < 1$.
- (d) $g'_4(x) = -\frac{1}{2} \sqrt{\frac{10}{x+4}} \frac{1}{x+4}$, per cui $|g'_4(x)| \leq 0.15$: la convergenza è più rapida che nel caso (c) perché la costante di contrattività è più piccola.
- (e) $g'_5(x) = \frac{6x^4 + 32x^3 + 32x^2 - 60x - 80}{(3x^2 + 8x)^2}$, quindi in $[1, 2]$ è $|g'_5(x)| \leq 0.6$.
Tuttavia, $g_5(x)$ è il metodo a convergenza più rapida: ciò dipende dal fatto che in $[1, 1.5]$, $|g'_5(x)| \leq 0.11$, invece $|g'_4(x)| \leq 0.123$. Da ciò segue la più veloce convergenza del caso (e) rispetto al caso (d).

Risultati di convergenza locale

Spesso è difficile verificare la condizione che per $x \in [a, b]$ si abbia $g(x) \in [a, b]$, condizione essenziale nel Teorema di convergenza globale. In tali casi, è utile un teorema di convergenza locale che assicura la convergenza di $\{x_k\}$ a un punto fisso x^* di $g(x)$ se x_0 è sufficientemente vicino a x^* .

È necessario sapere a priori che x^* è un punto fisso di $g(x)$, ossia che $g(x)$ ha un punto fisso.

Teorema di convergenza locale

Sia x^* un punto fisso di $g(x)$; si suppone che $g(x)$ sia continua e sia una contrazione per ogni $x \in [x^* - \rho, x^* + \rho] = \mathcal{I}_\rho$. Allora, per ogni $x_0 \in \mathcal{I}_\rho$, la successione degli $\{x_k\}$ è ben definita, ossia $x_k \in \mathcal{I}_\rho$, e converge a x^* per $k \rightarrow \infty$. Inoltre, x^* è l'unico punto fisso di $g(x)$ in \mathcal{I}_ρ .

Dimostrazione

Per ipotesi $x_0 \in \mathcal{I}_\rho$ e dunque

$$|x_0 - x^*| \leq \rho$$

Preso $x_1 = g(x_0)$, si ha

$$|x_1 - x^*| = |g(x_0) - g(x^*)| \leq L|x_0 - x^*| \leq L\rho < \rho$$

poiché $L < 1$. Supposto ora che $x_2, \dots, x_k \in \mathcal{I}_\rho$. Si dimostra che ciò vale anche per x_{k+1} :

$$|x_{k+1} - x^*| = |g(x_k) - g(x^*)| \leq L|x_k - x^*| \leq L\rho < \rho$$

da cui discende che la successione è ben definita. Inoltre,

$$|x_k - x^*| \leq L|x_{k-1} - x^*| \leq L^2|x_{k-2} - x^*| \leq \dots \leq L^k|x_0 - x^*| \leq L^k\rho$$

Per $k \rightarrow \infty$ si ha $\lim_{k \rightarrow \infty} L^k = 0$ e dunque x_k è convergente a x^* .

Supposto per assurdo che y^* sia un punto fisso di $g(x)$ in \mathcal{I}_ρ diverso da x^* , si ha che:

$$|x^* - y^*| = |g(x^*) - g(y^*)| \leq L|x^* - y^*| < |x^* - y^*|$$

Ciò è assurdo. Anche in questo caso si prova che:

$$|x_k - x^*| \leq \frac{L^k}{1-L}|x_1 - x_0|$$

Propagazione degli errori nel metodo delle approssimazioni successive

Poiché si opera in aritmetica finita anziché su numeri reali, è impossibile calcolare esattamente la funzione $g(x)$ per x assegnato. Di fatto, per x assegnato, si calcola una approssimazione di $g(x)$, data da

$$a(x) = g(x) + \delta(x)$$

dove $\delta(x)$ è l'errore commesso. Di solito è nota una maggiorazione di tale errore: $|\delta(x)| \leq \delta$. Pertanto, **operando in aritmetica finita**, il metodo delle approssimazioni successive diventa:

$$w_{k+1} = a(w_k) = g(w_k) + \delta_k \quad k = 0, 1, 2, \dots$$

dove w_k è il k -esimo iterato ottenuto operando con **numeri finiti** e $|\delta_k| \leq \delta$. **In generale la successione $\{w_k\}$ non converge**. Tuttavia, è possibile determinare, sotto opportune condizioni, una approssimazione di x^* tanto più accurata quanto più δ è piccolo.

Teorema

Sia x^* un punto fisso di $g(x)$ e supponiamo che, in un intervallo $\mathcal{I}_\rho = [x^* - \rho, x^* + \rho]$, $g(x)$ sia continua e sia contrattiva: allora, per ogni $w_0 \in \mathcal{I}_{\rho_0} = [x^* - \rho_0, x^* + \rho_0]$, con $\rho_0 = \rho - \frac{\delta}{1-L}$ dove $\delta \geq |\delta_k|$, la successione dei w_k , ottenuta da $w_k = g(w_{k-1}) + \delta_{k-1}$ è tale che:

$$|w_k - x^*| \leq \frac{\delta}{1-L} + L^k \left(\rho_0 - \frac{\delta}{1-L} \right)$$

e i w_k appartengono a $\mathcal{I}_\rho \forall k$.

Nella disuguaglianza, il primo termine è indipendente da k e può essere grande se L è prossimo a 1, mentre il secondo termine tende a 0 per $k \rightarrow \infty$. Pertanto, non si ha più convergenza della successione degli iterati a x^* .

Per quanto k sia preso grande, la differenza tra due iterati successivi non può essere più piccola di $\frac{2\delta}{1-L}$, a causa degli errori di arrotondamento nel calcolo di $g(x)$.

Criteri di arresto per il metodo delle approssimazioni successive

Occorre determinare un criterio per vedere se l'approssimazione ottenuta è punto fisso di $g(x)$, ossia se $x - g(x) = \phi(x)f(x) = 0$. Si ritiene che x_k sia una approssimazione accettabile se contemporaneamente:

$$|f(x_k)| \leq \tau_1 \quad \text{e} \quad |x_k - x_{k-1}| \leq \tau_2$$

oppure

$$\frac{|f(x_k)|}{f_{\max}} \leq \sigma_1 \quad \text{e} \quad \frac{|x_k - x_{k-1}|}{|x_k|} \leq \sigma_2$$

dove $\tau_1, \tau_2, \sigma_1, \sigma_2$ sono tolleranze assegnate e $f_{\max} = \max_{x \in \mathcal{I}_\rho} |f(x)|$.

La scelta delle tolleranze è cruciale. Può essere che $|f(x_k)|$ sia piccola pur non essendo x_k accettabile e, viceversa, può essere che $|f(x_k)|$ sia grande pur essendo x_k accettabile.

La tolleranza τ_2 deve essere inoltre maggiore o uguale a $\frac{2\delta}{1-L}$, poiché questo termine, che tiene conto degli errori di arrotondamento, non converge a 0 per $k \rightarrow \infty$. D'altra parte, la successione delle differenze $\{x_k - x_{k-1}\}$ può convergere a 0, pur essendo le due successioni divergenti. Se non si conosce nulla di $f(x)$, conviene applicare i test relativi.

Ordine di convergenza

Un metodo iterativo che genera una successione di iterati $\{x_k\}$ convergenti a x^* , si dice che è **di ordine p** o che ha **velocità di convergenza pari a p** se esistono due costanti C e p tali che:

$$\lim_{k \rightarrow \infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*|^p} = C \quad \text{con } p > 1$$

oppure

$$\lim_{k \rightarrow \infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*|} = C \quad 0 < C < 1 \quad (p = 1)$$

Se $p = 1$ e $0 < C < 1$, il metodo si dice **lineare**. Se invece è

$\lim_{k \rightarrow \infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*|} = 0$, il metodo è **superlineare**, mentre se il limite vale 1, il metodo si dice **sublineare**. C si dice **costante asintotica di errore**. Vale che

$$e_{k+1} = |x_{k+1} - x^*| \approx C|x_k - x^*|^p = Ce_k^p$$

o anche

$$e_{k+1} = (C + \delta_k)e_k^p$$

con $\lim_{k \rightarrow \infty} \delta_k = 0$.

Esempio. Siano $p = 2$, $C = 2$. Per $e_k = 0.1$, se il metodo è di ordine 2 si ha

$$e_{k+1} \approx Ce_k^2 = 2 \cdot 10^{-2} = 0.02$$

$$e_{k+2} \approx Ce_{k+1}^2 = 2 \cdot 4 \cdot 10^{-4} = 8 \cdot 10^{-4} = 0.0008$$

$$e_{k+3} \approx Ce_{k+2}^2 = 2 \cdot 64 \cdot 10^{-8} = 1.28 \cdot 10^{-6} = 0.00000128$$

Come trovare metodi con velocità di convergenza maggiore di 1?

Teorema

Sia x^* un punto fisso di $g(x)$. Supponiamo che $g(x)$ sia di classe C^p in un opportuno intervallo $\mathcal{I}_\rho = [x^* - \rho, x^* + \rho]$ con $g^{(i)}(x^*) = 0$, per $i = 1, 2, \dots, p-1$, e $g^{(p)}(x^*) \neq 0$. Se $p = 1$, assumiamo $|g'(x^*)| < 1$.

Allora per ogni $x_0 \in \mathcal{I}_\rho$ opportuno, il metodo di iterazione funzionale è un metodo convergente di ordine p .

Dim. Poiché $g(x) \in C^p$ ed è $g'(x^*) = 0$ oppure $|g'(x^*)| < 1$ per $p = 1$, allora esiste un intervallo $\mathcal{I}_\rho = [x^* - \rho, x^* + \rho]$ per cui $|g'(x)| \leq L < 1$ per $x \in \mathcal{I}_\rho$ e inoltre $g(x) \in \mathcal{I}_\rho$. Infatti, $x_0 \in \mathcal{I}_\rho$ e vale che

$$|x_1 - x^*| = |g(x_0) - g(x^*)| = |g'(\xi)||x_0 - x^*| < |x_0 - x^*| < \rho$$

In generale

$$|x_k - x^*| = |g(x_{k-1}) - g(x^*)| \leq L|x_{k-1} - x^*| \leq \dots \leq L^k|x_0 - x^*| < \rho$$

Dunque $x_k \in \mathcal{I}_\rho$ e, come per il teorema di convergenza locale, $x_k \rightarrow x^*$ per $k \rightarrow \infty$.

Come trovare metodi con velocità di convergenza maggiore di 1?

Se si considera lo sviluppo in serie di Taylor di punto iniziale x^* , si ha:

$$\begin{aligned}x_{k+1} &= g(x_k) = g(x^*) + g'(x^*)(x_k - x^*) + \dots \\&\quad \dots + \frac{g^{(p-1)}(x^*)}{(p-1)!}(x_k - x^*)^{p-1} + \frac{g^{(p)}(\xi_k)}{p!}(x_k - x^*)^p \\&= x^* + \frac{g^{(p)}(\xi_k)}{p!}(x_k - x^*)^p\end{aligned}$$

con $\xi_k \in]x_k, x^*[\subset \mathcal{I}_\rho$. Pertanto si ha

$$x_{k+1} - x^* = \frac{g^{(p)}(\xi_k)}{p!}(x_k - x^*)^p$$

Da

$$\lim_{k \rightarrow \infty} g^{(p)}(\xi_k) = g^{(p)}\left(\lim_{k \rightarrow \infty} (\xi_k)\right) = g^{(p)}(x^*)$$

segue

$$\lim_{k \rightarrow \infty} \frac{x_{k+1} - x^*}{(x_k - x^*)^p} = \frac{g^{(p)}(x^*)}{p!}$$

Se $p = 1$, $|g'(x^*)| < 1$. $\frac{g^{(p)}(x^*)}{p!}$ si dice **costante asintotica di errore**.

Conseguenze

- 1 Se x^* è punto fisso di $g(x)$ e $g \in C^1$, con $g'(x^*) \neq 0$ ed è $|g'(x^*)| < 1$, allora esiste un intorno $\mathcal{I}_\rho = [x^* - \rho, x^* + \rho]$ per cui $|g'(x)| < 1$ per $x \in \mathcal{I}_\rho$. In tale intervallo, per ogni $x_0 \in \mathcal{I}_\rho$, il metodo iterativo **converge al punto fisso in modo lineare**.
- 2 Se x^* è punto fisso di $g(x)$ e $g \in C^2$, con $g'(x^*) = 0$ e $g''(x^*) \neq 0$, allora esiste un intorno $\mathcal{I}_\rho = [x^* - \rho, x^* + \rho]$ tale che, per ogni $x_0 \in \mathcal{I}_\rho$, il metodo iterativo **converge al punto fisso con velocità di convergenza quadratica** e vale che

$$\lim_{k \rightarrow \infty} \frac{|x_{k+1} - x^*|}{(x_k - x^*)^2} = \frac{|g''(x^*)|}{2}$$

o equivalentemente

$$|x_{k+1} - x^*| = \frac{|g''(\xi_k)|}{2}(x_k - x^*)^2$$

con $\xi_k \in \mathcal{I}_\rho$.

Se M è un valore maggiorante del valore assoluto della derivata seconda di $g(x)$ in \mathcal{I}_ρ , allora

$$|e_{k+1}| \leq \frac{M}{2} e_k^2$$

$\frac{M}{2}$ è la costante che maggiora $\frac{e_{k+1}}{e_k^2}$ il cui valore asintotico è $\frac{|g''(x^*)|}{2}$.

Definizione

Si dice che x^* è uno zero di $f(x)$ di **molteplicità** m se $f(x) = (x - x^*)^m q(x)$, $x \neq x^*$, e $\lim_{x \rightarrow x^*} q(x) \neq 0$. Se $m = 1$, x^* è uno zero semplice.

Teorema

Sia $f \in C^1([a, b])$, dove $[a, b]$ contiene uno zero x^* di $f(x)$:
il punto x^* è uno **zero semplice** $\Leftrightarrow f(x^*) = 0, f'(x^*) \neq 0$.

Dim. Se x^* è uno zero semplice, allora $f(x^*) = 0$ e $f(x) = (x - x^*)q(x)$, con $\lim_{x \rightarrow x^*} q(x) \neq 0$. Si ha:

$$f'(x) = (x - x^*)q'(x) + q(x)$$

Allora

$$\lim_{x \rightarrow x^*} f'(x) = f'(x^*) = \lim_{x \rightarrow x^*} q(x) \neq 0.$$

Se, viceversa, $f(x^*) = 0$ ed $f'(x^*) \neq 0$, allora

$$f(x) = f(x^*) + (x - x^*)f'(\xi) = (x - x^*)f'(\xi)$$

con ξ compreso tra x e x^* . Posto $q(x) = f'(\xi)$, si ha $\lim_{x \rightarrow x^*} f'(\xi) = f'(\lim_{x \rightarrow x^*} \xi) = f'(x^*) \neq 0$.

Teorema

Sia $f \in C^m([a, b])$, dove $[a, b]$ contiene uno zero x^* di $f(x)$. Il punto x^* è uno **zero di molteplicità** $m \Leftrightarrow f^{(i)}(x^*) = 0$ per $i = 0, 1, \dots, m-1$, ed è $f^{(m)}(x^*) \neq 0$.

Metodi a convergenza superlineare

Data l'equazione $f(x) = 0$, si può determinare x^* , soluzione di $f(x) = 0$, come punto fisso di

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

con $\phi(x) \neq 0$ per ogni x nell'intervallo in cui si cerca la soluzione.

Per esaminare la velocità di convergenza, si consideri $g'(x)$:

$$g'(x) = 1 - \phi(x)f'(x) - \phi'(x)f(x)$$

In x^* è $g'(x^*) = 1 - \phi(x^*)f'(x^*)$.

Il metodo iterativo ha velocità di convergenza lineare se $g'(x^*) \neq 0$, ossia se $\phi(x^*) \neq 1/f'(x^*)$, supposto che $f'(x^*) \neq 0$ (x^* deve essere zero semplice di $f(x)$). Se $\phi(x)$ è costante, con $\phi(x) = m \neq 1/f'(x^*)$, il metodo è lineare.

La convergenza è quadratica se

$$\phi(x^*) = \frac{1}{f'(x^*)}$$

con $f'(x^*) \neq 0$. Allora, o si pone $\phi(x) = 1/f'(x^*)$ costante (ma x^* è incognito), oppure si pone

$$\phi(x) = \frac{1}{f'(x)}$$

ottenendo un **metodo a convergenza quadratica** dato da

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

Il metodo è detto **metodo di Newton** (o di Newton-Raphson). Vale che

$$g'(x^*) = 1 - \frac{(f'(x^*))^2 - f(x^*)f''(x^*)}{(f'(x^*))^2} = \frac{f(x^*)f''(x^*)}{(f'(x^*))^2} = 0$$

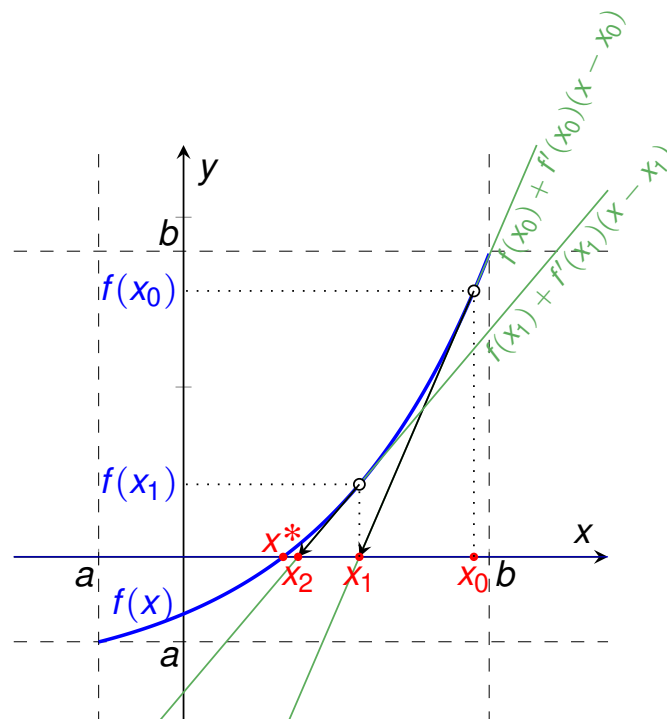
$$g''(x^*) = \frac{(f'(x^*)f''(x^*) + f(x^*)f'''(x^*))(f'(x^*))^2 - 2f(x^*)(f''(x^*))^2 f'(x^*)}{(f'(x^*))^4} = \frac{f''(x^*)}{f'(x^*)}$$

Allora se $f(x^*) = 0$, $f'(x^*) \neq 0$, $f''(x^*) \neq 0$, **il metodo di Newton ha convergenza quadratica con costante asintotica di errore $\frac{f''(x^*)}{2f'(x^*)}$.**

Metodo di Newton

Il metodo di Newton si dice anche **metodo delle tangenti**, perché geometricamente x_{k+1} è il punto d'intersezione tra $y = 0$ (asse delle ascisse) e la retta tangente a $f(x)$ in $(x_k, f(x_k))$:

$$y = f(x_k) + f'(x_k)(x - x_k)$$



Riassumiamo tutte le condizioni per la **convergenza locale** del metodo di Newton.

Sia x^* uno zero di $f(x)$. Sia $f(x)$ continua insieme alle sue derivate prima, seconda e terza (continuità di g, g', g'') in un intorno del punto fisso. Sia $f'(x) \neq 0$ per x in un opportuno intorno di x^* e sia $f''(x^*) \neq 0$ (perché $f(x)/f'(x)$ deve essere definita e deve essere $g''(x^*) \neq 0$).

Allora, per ogni $x_0 \in [x^* - \rho, x^* + \rho]$, la successione generata dal metodo di Newton converge a x^* in modo quadratico.

Esempi

1 $\sin(x) - (x/2)^2 = 0$ in $[1, 2]$. $f'(x) = \cos(x) - x/2$, $f''(x) = -\sin(x) - 1/2$

$$x_{k+1} = x_k - \frac{\sin(x_k) - \left(\frac{x_k}{2}\right)^2}{\cos(x_k) - \frac{x_k}{2}}$$

k	x_k	$f(x_k)$	$f'(x_k)$	$-f(x_k)/f'(x_k)$
0	1.5	0.434995	-0.67926	0.64039
1	2.14039	-0.303197	-1.60948	-0.18838
2	1.95201	-0.024372	-1.34805	-0.01808
3	1.93393	-0.000233	-1.32217	-0.00018
4	1.93375	0.000005		

Costante asintotica di errore: $g''(x^*)/2 = \frac{f''(x^*)}{2f'(x^*)} \approx 0.54$.

2 $x^2 - \gamma = 0$

$$x_{k+1} = x_k - \frac{x_k^2 - \gamma}{2x_k} = \frac{1}{2} \left(x_k + \frac{\gamma}{x_k} \right)$$

Per $\gamma = 2$ e $[a, b] = [1, 2]$ si ha:

k	x_k
0	1.5
1	1.41666666
2	1.41421568
3	1.414213561
4	1.414213562

La costante asintotica di errore vale $f''(x^*)/(2f'(x^*)) = 1/(2\sqrt{\gamma})$. Se γ è piccolo, la convergenza può essere lenta.

La complessità del metodo è pari a una valutazione di funzione e una valutazione della derivata prima per ciascun passo; [se la complessità di \$f'\$ è analoga a quella di \$f\$](#) , si dice che il metodo richiede due valutazioni di funzioni per passo.

Codice Matlab

```
function [x, it] = newton(fname, fpname, x0, tolx, tolf, maxit)
% INPUT:  fname (string o fhandle) - Function della funzione
%         fpname (string o fhandle) - Function della derivata prima
%         x0     (double)         - Stima iniziale
%         tolx   (double)         - Distanza minima fra iterati successivi
%         tolf   (double)         - Soglia verso zero dei valori di f(x)
%         maxit  (integer)        - Numero massimo di interazioni
% OUTPUT: x     (double)         - Approssimazione della soluzione
%         it     (integer)        - Numero di iterazioni eseguite
tolfp = min( tolf, 10*eps );
% Metodo di Newton
x = x0; fx = feval(fname, x); it = 0;
stop = ( abs(fx) < tolf );
while ( ~stop )
    it = it + 1;
    fpx = feval(fpname, x);
    if (abs(fpx) < tolf), error('|f''(xk)| troppo piccolo.');

```

Teorema (convergenza globale del metodo di Newton)

Sia $f \in C^2([a, b])$ con $f'(x) \neq 0 \forall x \in [a, b]$. Sia inoltre verificato **uno** dei seguenti quattro gruppi di condizioni:

$$\begin{array}{ll} \left\{ \begin{array}{l} f(a) < 0, f(b) > 0 \\ f''(x) \leq 0 \\ |f(b)| \leq (b-a)|f'(b)| \end{array} \right. & \text{oppure} \quad \left\{ \begin{array}{l} f(a) < 0, f(b) > 0 \\ f''(x) \geq 0 \\ |f(a)| \leq (b-a)|f'(a)| \end{array} \right. \\ \text{oppure} \quad \left\{ \begin{array}{l} f(a) > 0, f(b) < 0 \\ f''(x) \geq 0 \\ |f(b)| \leq (b-a)|f'(b)| \end{array} \right. & \text{oppure} \quad \left\{ \begin{array}{l} f(a) > 0, f(b) < 0 \\ f''(x) \leq 0 \\ |f(a)| \leq (b-a)|f'(a)| \end{array} \right. \end{array}$$

Allora il metodo di Newton genera una successione di iterati convergenti all'unica soluzione di $f(x) = 0$ appartenente ad $[a, b]$, a partire da un **qualunque** $x_0 \in [a, b]$.

Esempio

Data l'equazione $f(x) = x^3 - x - 1 = 0$, verificare che il metodo di Newton sia globalmente convergente in $[1, 2]$.

Si ha:

- $f(1) = -1 < 0, f(2) = 5 > 0$;
- $f'(x) = 3x^2 - 1$; $f'(x) = 0$ per $x = \pm 1/\sqrt{3}$, entrambi esterni all'intervallo $[1, 2]$, dunque $f'(x) \neq 0$ in $[1, 2]$;
- $f''(x) = 6x > 0$ nell'intervallo $[1, 2]$;
- $(b-a)|f'(1)| = 2 > 1 = |f(1)|$.

Essendo verificate tutte le condizioni del Teorema, il metodo di Newton è globalmente convergente per l'equazione data nell'intervallo $[1, 2]$ ($x^* \approx 1.32472$):

$x_0 = 1.0$		$x_0 = 1.52$		$x_0 = 1.9$		$x_0 = 0.8^*$	
k	x_k	k	x_k	k	x_k	k	x_k
1	1.5	1	1.35278	1	1.49725	1	2.2
2	1.34783	2	1.32542	2	1.34718	2	1.64911
3	1.32520	3	1.32472	3	1.32517	3	1.39267
4	1.32472			4	1.32472	4	1.32866
						5	1.32473
						6	1.32472

* **N.B.:** si ha convergenza, ma questo caso **non** rientra fra le ipotesi del Teorema perché $x_0 = 0.8 \notin [1, 2]$.

Metodo di Newton per zeri multipli

Se x^* è uno **zero semplice** di $f(x)$, il metodo di Newton converge **quadraticamente**.

Se x^* è uno **zero di molteplicità m** , allora, poiché $f(x) = (x - x^*)^m q(x)$, con $\lim_{x \rightarrow x^*} q(x) \neq 0$, si ha:

$$f'(x) = m(x - x^*)^{m-1} q(x) + (x - x^*)^m q'(x) = (x - x^*)^{m-1} (mq(x) + (x - x^*)q'(x))$$

$$\begin{aligned} f''(x) &= m(m-1)(x - x^*)^{m-2} q(x) + m(x - x^*)^{m-1} q'(x) \\ &\quad + m(x - x^*)^{m-1} q'(x) + (x - x^*)^m q''(x) \\ &= (x - x^*)^{m-2} (m(m-1)q(x) + 2mq'(x)(x - x^*) + (x - x^*)^2 q''(x)) \end{aligned}$$

$$\frac{f(x)}{f'(x)} = (x - x^*) \frac{q(x)}{mq(x) + (x - x^*)q'(x)}$$

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

$$\begin{aligned} g'(x) &= 1 - \frac{(f'(x^*))^2 - f(x)f''(x)}{(f'(x^*))^2} = \frac{f(x)f''(x)}{(f'(x^*))^2} \\ &= \frac{q(x) (m(m-1)q(x) + 2mq'(x)(x - x^*) + (x - x^*)^2 q''(x))}{(mq(x) + (x - x^*)q'(x))^2} \end{aligned}$$

$$g'(x^*) = \frac{(m^2 - m)q^2(x^*)}{m^2 q^2(x^*)} = 1 - \frac{1}{m} \begin{cases} \neq 0 \\ \neq 1 \end{cases} \quad \forall m > 1$$

Metodo di Newton per zeri multipli

Dunque $0 < g'(x^*) < 1$, da cui discende che il metodo di Newton diventa convergente solo linearmente. Tuttavia se si modifica l'iterazione nel seguente modo (**metodo di Newton modificato**):

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

la convergenza ritorna almeno quadratica:

$$\begin{aligned} g'(x) &= 1 - m \frac{(f'(x))^2 - f(x)f''(x)}{(f'(x))^2} \\ &= 1 - m + m \frac{f(x)f''(x)}{(f'(x))^2} \\ g'(x^*) &= 1 - m + m \frac{f(x^*)f''(x^*)}{(f'(x^*))^2} = 1 - m + m \left(1 - \frac{1}{m}\right) = 0 \end{aligned}$$

Un altro modo per ottenere un metodo a **convergenza quadratica** consiste nell'osservare che

$$\frac{f(x)}{f'(x)} = (x - x^*) \frac{q(x)}{mq(x) + (x - x^*)q'(x)}$$

ha in x^* uno zero di molteplicità 1. Dunque si applica il metodo di Newton a $f(x)/f'(x)$ (**metodo di Halley**):

$$x_{k+1} = x_k - \underbrace{\frac{f(x_k)/f'(x_k)}{\left((f'(x_k))^2 - f(x_k)f''(x_k)\right) / (f'(x_k))^2}}_{g(x_k)}$$

Se $g(x)$ è sufficientemente regolare in un opportuno intervallo che contiene x^* , allora si ha **convergenza quadratica** al punto fisso x^* .

Numericamente, per x_k sufficientemente vicino a x^* , se entrambi $f'(x^*)$ e $f(x^*)$ sono piccoli, si possono avere problemi di cancellazione.

Altra derivazione del metodo di Newton

Il metodo di Newton si può derivare considerando lo sviluppo in serie di Taylor a partire da x_k e calcolandolo in x^* , trascurando i termini di ordine superiore a 1.

$$\begin{aligned} 0 = f(x^*) &= f(x_k) + f'(x_k)(x^* - x_k) + \frac{f''(\xi)}{2}(x^* - x_k)^2 \\ \Rightarrow x^* &\approx x_k - \frac{f(x_k)}{f'(x_k)} \end{aligned}$$

Se si sostituisce la funzione $f(x)$ con il polinomio di grado 1 tale che $p(x_k) = f(x_k)$ e $p'(x_k) = f'(x_k)$ (**modello lineare di $f(x)$**), si ottiene il metodo di Newton.

Se si sostituisce la funzione $f(x)$ con il polinomio di grado 2, anziché di grado 1, tale che il polinomio coincide con $f(x)$ fino alla derivata seconda, si ottiene la seguente **iterazione funzionale**:

$$x_{k+1} = x_k - \frac{f'(x_k) \pm \sqrt{(f'(x_k))^2 - 2f(x_k)f''(x_k)}}{f''(x_k)}$$

Per uno zero semplice, il metodo ha ordine di convergenza 3. Tuttavia queste iterazioni funzionali comportano **alta complessità** poiché richiedono ad ogni passo il calcolo della derivata seconda, oltre quello della funzione e della derivata prima.

Anziché usare polinomi di Taylor per approssimare $f(x)$ e calcolarne uno zero, si possono usare polinomi di interpolazione. Da tale idea nascono i **metodi di interpolazione** per la determinazione degli zeri di una funzione.

Siano $a_0 \equiv a$, $b_0 \equiv b$, tali che $f(a) \cdot f(b) < 0$, con $f \in C([a, b])$. Per ogni $k = 0, 1, 2, \dots$, si costruisce a partire da $[a_k, b_k]$ con $f(a_k) \cdot f(b_k) < 0$ una successione di iterati dati da:

$$c_k = b_k - \frac{f(b_k)}{\frac{f(b_k) - f(a_k)}{b_k - a_k}} = b_k - \frac{f(b_k)}{f(b_k) - f(a_k)}(b_k - a_k)$$

Geometricamente c_k è l'intersezione tra l'asse delle ascisse e la retta che unisce i punti di coordinate $(a_k, f(a_k))$ e $(b_k, f(b_k))$:

$$\begin{cases} y = 0 \\ y = f(b_k) + \frac{f(b_k) - f(a_k)}{b_k - a_k}(x - b_k) \end{cases}$$

Si calcola c_k e se $f(c_k) = 0$, allora $x^* = c_k$, altrimenti si pone

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

La regola falsi è convergente e l'ordine di convergenza è lineare.

Non è detto che abbia maggiore efficienza del metodo di bisezione. Ciò dipende dalla particolare scelta della funzione. Si veda ad esempio il comportamento di $f(x) = \sin(x) - 0.9$ nell'intervallo $[0, 2]$.

Metodo delle secanti

Nel **metodo delle secanti**, x_{k+1} è intersezione tra $y = 0$ e il polinomio di grado 1 che interpola $f(x)$ nei punti $(x_{k-1}, f(x_{k-1}))$ e $(x_k, f(x_k))$:

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

Tuttavia, non è detto che $f(x_k)$ e $f(x_{k-1})$ siano di segno opposto e questo, in aritmetica finita, può dare problemi di cancellazione che non si presentano con il metodo della regola falsi.

Si ha bisogno di due approssimazioni iniziali, ma di una sola valutazione di funzione per ogni passo dopo il primo.

La convergenza (locale) è ottenibile solo se x_0 e x_1 appartengono a un intervallo sufficientemente piccolo intorno a x^* , zero di $f(x)$.

```
function [x, it] = secanti(fname, x0, x1, tolx, tolf, maxit)
% Metodo delle secanti
it = 0;
if ( abs(x1 - x0) < tolx ), x = (x0 + x1) / 2; return; end
fx0 = feval(fname, x0); if ( abs(fx0) < tolf ), x = x0; return; end
fx1 = feval(fname, x1); if ( abs(fx1) < tolf ), x = x1; return; end
stop = 0;
while ( ~stop )
    it = it + 1;
    t = fx0 / fx1;
    d = (x1 - x0) / (1 - t);
    x = x1 - d;
    fx = feval(fname, x);
    stop = ( (abs(fx) < tolf && abs(d) < tolx + eps*abs(x1)) ...
        || fx == 0 || it == maxit );
    fx0 = fx1;
    fx1 = fx;
    x0 = x1;
    x1 = x;
end
if ( it >= maxit )
    fprintf('\nRaggiunto il massimo numero di iterazioni.\n');
end
```

Convergenza locale del metodo delle secanti

$$x_{k+1} = x_k - \frac{f(x_k)}{\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}} = x_k - \frac{f(x_k)}{f[x_k, x_{k-1}]}$$

Sottraendo x^* ad ambo i membri e ricordando che $f(x^*) = 0$, si ottiene

$$\begin{aligned} x_{k+1} - x^* &= x_k - x^* - \frac{f(x_k) - f(x^*)}{f[x_k, x_{k-1}]} = (x_k - x^*) \left(1 - \frac{\frac{f(x_k) - f(x^*)}{x_k - x^*}}{f[x_k, x_{k-1}]} \right) \\ &= (x_k - x^*) \left(1 - \frac{f[x_k, x^*]}{f[x_k, x_{k-1}]} \right) = (x_k - x^*) \frac{f[x_k, x_{k-1}] - f[x_k, x^*]}{f[x_k, x_{k-1}]} \\ &= (x_k - x^*) \frac{f[x_k, x_{k-1}] - f[x_k, x^*]}{x_{k-1} - x^*} \cdot \frac{x_{k-1} - x^*}{f[x_k, x_{k-1}]} \\ &= (x_k - x^*)(x_{k-1} - x^*) \frac{f[x_k, x_{k-1}, x^*]}{f[x_k, x_{k-1}]} \end{aligned}$$

Convergenza locale del metodo delle secanti

Se $f \in C^2$ in un intervallo che contiene x^* , x_k e x_{k-1} , allora

$$\frac{f[x_k, x_{k-1}, x^*]}{f[x_k, x_{k-1}]} = \frac{f''(\xi)}{2f'(\eta)}$$

con $\xi \in \mathcal{I}(x_k, x_{k-1}, x^*)$, $\eta \in \mathcal{I}(x_k, x_{k-1})$.

Se $f'(x^*) \neq 0$, esiste un intervallo in cui $f'(x) \neq 0$. Sia $\mathcal{I}_\rho = [x^* - \rho, x^* + \rho]$ tale intervallo: poiché f' ed f'' sono funzioni continue, esiste $M > 0$ tale che

$$\left| \frac{f''(x)}{2f'(y)} \right| \leq M$$

per tutti gli $x, y \in \mathcal{I}_\rho$. Allora

$$|x_{k+1} - x^*| \leq M \cdot |x_k - x^*| \cdot |x_{k-1} - x^*|$$

Posto $e_k = M|x_k - x^*|$, si ha che

$$e_{k+1} \leq e_k e_{k-1}$$

Siano ora x_0 e x_1 tali che e_0 ed e_1 siano minori del minimo tra 1 e ρM :

$$e_0 < \min\{1, \rho M\} \quad e_1 < \min\{1, \rho M\}$$

Convergenza locale del metodo delle secanti

Posto $K = \max\{e_0, e_1^{1/p}\} < 1$, con p tale che $p^2 = p + 1$, si ha

$$e_0 \leq K$$

$$e_1 \leq K^p$$

$$e_2 \leq e_1 e_0 \leq K^{p+1} = K^{p^2}$$

In generale, se si suppone $e_i \leq K^{p^i}$, si ha

$$e_{i+1} \leq e_i e_{i-1} \leq K^{p^i} K^{p^{i-1}} = K^{p^{i-1}(p+1)} = K^{p^{i-1}p^2} = K^{p^{i+1}}$$

Allora il metodo delle secanti, per $x_0, x_1 \in \mathcal{I}_\rho$, genera una successione di iterati tale che

$$e_{i+1} \leq K^{p^{i+1}}$$

ossia, per $i \rightarrow \infty$, la quantità e_{i+1} (che è un multiplo del modulo dell'errore al passo $(i+1)$ -esimo) converge a 0 come $(K^{p^i})^p = (K^p)^{p^i}$.

Si dimostra che l'ordine di convergenza del metodo delle secanti è p e, poiché p è la radice di $t^2 - t - 1 = 0$, **la velocità di convergenza è frazionaria**, data da $p = (1 + \sqrt{5})/2 \approx 1.618 \dots$ (sezione aurea).

Un passo del metodo delle secanti richiede una **sola valutazione di funzione**. Dunque si può assumere che **due passi del metodo delle secanti abbiano la stessa complessità del metodo di Newton**.

Considerando **coppie di passi**, per quanto riguarda la convergenza, si ha:

$$e_{i+2} \leq K^{p^{i+2}} = (K^{p^i})^{p^2} \leq (K^{p^i})^{p+1} = (K^{p+1})^{p^i}$$

perchè $p^2 = p + 1$. Pertanto, si ha un procedimento con velocità di convergenza $p + 1 = 2.618\dots$, **localmente più veloce di quello di Newton**.

Generalizzazione del metodo delle secanti

Dati $r + 1$ valori approssimati $x_k, x_{k-1}, \dots, x_{k-r}$, si determina x_{k+1} come intersezione tra $y = 0$ e il **polinomio interpolante di grado r** tale che

$$f(x_{k-i}) = p(x_{k-i}) \quad \forall i = 0, 1, \dots, r$$

Per $r = 1$, si ottiene il **metodo delle secanti**.

Per $r = 2$, si ottiene il **metodo di Muller**: dati x_k, x_{k-1}, x_{k-2} , si determina

$$p(x) = a(x - x_k)^2 + b(x - x_k) + c$$

tale che

$$p(x_k) = f(x_k) = c$$

$$p(x_{k-1}) = f(x_{k-1}) = a(x_{k-1} - x_k)^2 + b(x_{k-1} - x_k) + c$$

$$p(x_{k-2}) = f(x_{k-2}) = a(x_{k-2} - x_k)^2 + b(x_{k-2} - x_k) + c$$

Pertanto

$$c = f(x_k)$$

$$b = \frac{(x_{k-2} - x_k)^2 (f(x_{k-1}) - f(x_k)) - (x_{k-1} - x_k)^2 (f(x_{k-2}) - f(x_k))}{(x_{k-2} - x_k)(x_{k-2} - x_{k-1})(x_{k-1} - x_k)}$$

$$a = \frac{(x_{k-1} - x_k)(f(x_{k-2}) - f(x_k)) - (x_{k-2} - x_k)(f(x_{k-1}) - f(x_k))}{(x_{k-2} - x_k)(x_{k-2} - x_{k-1})(x_{k-1} - x_k)}$$

$$x_{k+1} = x_k - \frac{2c}{b \pm \sqrt{b^2 - 4ac}}$$

Il segno è scelto in modo che $x_{k+1} - x_k$ sia più piccolo possibile, ossia la quantità al denominatore più grande possibile:

$$x_{k+1} = x_k - \frac{2c}{b + \operatorname{sgn}(b)\sqrt{b^2 - 4ac}}$$

Se $f \in C^3$ e $f'(x^*) \neq 0$, in prossimità di x^* si ha **convergenza di ordine $p = 1.84$** , dove p è radice di $p^3 = p^2 + p + 1$.

Metodi dell'interpolazione inversa

Il metodo delle secanti possiede un'altra generalizzazione.

Se $f(x)$ nell'intorno di x^* è **biunivoca**, allora essa è invertibile, ossia esiste una funzione $g(y)$ tale che $g(f(x)) = x$.

Pertanto vale che $g(f(x^*)) = g(0) = x^*$.

Allora a partire da $f(x_k), \dots, f(x_{k-r})$ si può costruire un **polinomio interpolante $q(y)$** di grado r tale che $q(f(x_{k-i})) = x_{k-i} \forall i = 0, \dots, r$. Si considera come nuovo iterato x_{k+1} il valore di $q(0)$.

Questi metodi vengono detti **metodi dell'interpolazione inversa**.

Per $r = 1$ si riottiene il metodo delle secanti.

Per $r = 2$ si ha il **metodo dell'interpolazione quadratica inversa**: si tratta di costruire un polinomio $q(y)$ di grado 2

$$q(y) = \alpha_1 y^2 + \alpha_2 y + \alpha_3$$

che passa per i punti $(f(x_k), x_k), (f(x_{k-1}), x_{k-1}), (f(x_{k-2}), x_{k-2})$, e di considerare come nuovo iterato il valore $x_{k+1} = q(0) = \alpha_3$. Il procedimento viene poi reiterato.

La funzione **fzero** di Matlab è un **polialgoritmo** che usa al proprio interno il metodo di bisezione, il metodo delle secanti e il metodo dell'interpolazione quadratica inversa. La procedura sceglie automaticamente quale metodo applicare in modo da garantire convergenza globale.

Positura del problema di determinare lo zero di una funzione $f(x)$

Sia x^* una radice di $f(x) = 0$. Si vuole studiare cosa succede se si **perturba** la funzione $f(x)$ di una piccola quantità e si determina lo zero x_ε^* di $f_\varepsilon(x)$ anziché di $f(x)$.

Se $f_\varepsilon(x) = f(x) + \varepsilon h(x)$ con ε abbastanza piccolo, si può valutare $x_\varepsilon^* - x^*$, supponendo che $f(x)$ e $h(x)$ siano funzioni sufficientemente regolari.

Se x_ε^* è uno zero semplice e si sviluppa $f_\varepsilon(x)$ in serie di Taylor con punto iniziale x^* , si ha:

$$\begin{aligned} 0 &= f_\varepsilon(x_\varepsilon^*) = f_\varepsilon(x^*) + f'_\varepsilon(x^*)(x_\varepsilon^* - x^*) + \frac{1}{2}f''_\varepsilon(\xi)(x_\varepsilon^* - x^*)^2 \\ 0 &= f(x^*) + \varepsilon h(x^*) + f'(x^*)(x_\varepsilon^* - x^*) + \varepsilon h'(x^*)(x_\varepsilon^* - x^*) \\ &\quad + \frac{1}{2}f''(\xi)(x_\varepsilon^* - x^*)^2 + \frac{1}{2}\varepsilon h''(\xi)(x_\varepsilon^* - x^*)^2 \end{aligned}$$

con $\xi \in]x_\varepsilon, x_\varepsilon^*[$. **Tralasciando perturbazioni del secondo ordine**, si ha:

$$0 \approx \varepsilon h(x^*) - f'(x^*)(x^* - x_\varepsilon^*) \quad \Rightarrow \quad |x^* - x_\varepsilon^*| \approx \left| \frac{\varepsilon h(x^*)}{f'(x^*)} \right|$$

Positura del problema di determinare lo zero di una funzione $f(x)$

La perturbazione nel risultato è pari a quella del dato amplificata di $1/|f'(x^*)|$. Se $|f'(x^*)|$ è piccolo, il problema è mal condizionato; se $|f'(x^*)|$ è grande, il problema è ben condizionato (zero ben separato).

Se x_ε^* è zero di molteplicità m , si ha:

$$\begin{aligned} 0 &= f_\varepsilon(x_\varepsilon^*) = f_\varepsilon(x^*) + f'_\varepsilon(x^*)(x_\varepsilon^* - x^*) + \dots + \frac{f_\varepsilon^{(m-1)}(x^*)}{(m-1)!}(x_\varepsilon^* - x^*)^{m-1} \\ &\quad + \frac{f_\varepsilon^{(m)}(x^*)}{m!}(x_\varepsilon^* - x^*)^m + \frac{f_\varepsilon^{(m+1)}(\xi)}{(m+1)!}(x_\varepsilon^* - x^*)^{m+1} \\ 0 &= f(x^*) + \varepsilon h(x^*) + \sum_{j=1}^m \left(\frac{f^{(j)}(x^*)}{j!}(x_\varepsilon^* - x^*)^j + \varepsilon \frac{h^{(j)}(x^*)}{j!}(x_\varepsilon^* - x^*)^j \right) \\ &\quad + \frac{f^{(m)}(x^*)}{m!}(x_\varepsilon^* - x^*)^m + \varepsilon \frac{h^{(m)}(x^*)}{m!}(x_\varepsilon^* - x^*)^m \\ &\quad + \frac{f^{(m+1)}(\xi)}{(m+1)!}(x_\varepsilon^* - x^*)^{m+1} + \varepsilon \frac{h^{(m+1)}(\xi)}{(m+1)!}(x_\varepsilon^* - x^*)^{m+1} \end{aligned}$$

con $\xi \in]x_\varepsilon, x_\varepsilon^*[$. **Tralasciando perturbazioni del tipo $\varepsilon(x_\varepsilon^* - x^*)^j$, $j = 1, \dots, m+1$** , è:

$$0 \approx \varepsilon h(x^*) + \frac{f^{(m)}(x^*)}{m!}(x_\varepsilon^* - x^*)^m \quad \Rightarrow \quad |x^* - x_\varepsilon^*| \approx \left| \frac{m! \varepsilon h(x^*)}{f^{(m)}(x^*)} \right|^{1/m}$$

Anche se ε è piccolo, $\varepsilon^{1/m}$ può essere grande. Il problema è sempre mal condizionato.

Esempio

$$p(x) = (x - 20)(x - 19) \cdots (x - 1) = x^{20} - 210x^{19} + \dots = \sum_{j=0}^{20} a_j x^j$$

Se si perturba $a_{19} = -210$ di una quantità ε abbastanza piccola, si ha che

$$|x_\varepsilon^* - x^*| \approx \left| \frac{\varepsilon a_{19}(x^*)^{19}}{p'(x^*)} \right| = \frac{\varepsilon 210 (x^*)^{19}}{\prod_{\substack{i=1, \dots, 20 \\ i \neq x^*}} (x^* - i)}$$

Se $x^* = 20$, si ha che:

$$|x_\varepsilon^* - 20| \approx \frac{\varepsilon 210 (20)^{19}}{19!} = \varepsilon 0.9 \cdot 10^{10}$$

La perturbazione finale è **estremamente grande** anche se gli zeri sono estremamente separati.