

RISOLUZIONE DI EQUAZIONI NON LINEARI

Introduzione

Data una $f \in C[a, b]$ affrontiamo il problema di determinare eventuali $x \in [a, b]$ tali che : $f(x) = 0$.

E' questo il problema della ricerca degli zeri che, se non esiste un metodo per determinare le radici in modo immediato, si basa su metodi iterativi che danno luogo a successioni che si spera convergano alla soluzione.

Tali metodi possono essere chiusi (bracketing methods) che richiedono due valori iniziali all'interno dei quali deve trovarsi la soluzione, oppure aperti (open methods).

Mentre i metodi chiusi convergono sicuramente alla radice, quelli aperti possono divergere ma, se convergono, questi ultimi sono più veloci di quelli chiusi.

I metodi chiusi che esamineremo sono: **bisezione**, **falsa posizione**. I metodi aperti: **Newton** e **secanti**.

Trattiamo per il momento solo radici semplici di funzioni continue.

1. Metodo di bisezione

Sia $f \in C[a, b]$ e $f(a) \cdot f(b) < 0$. Per il teorema di esistenza degli zeri $\exists s \in]a, b[: f(s) = 0$.

Valutiamo $f(x)$ per $x_m = \frac{a+b}{2}$. Sia $f(x_m) \neq 0$ e tale che $f(a) \cdot f(x_m) < 0$.

Poniamo: $x_1 = \frac{a+b}{2}$, $x_0 = a$

Deve $\exists s \in [x_0, x_1] : f(s) = 0$. Poniamo $x_2 = \frac{x_0 + x_1}{2}$

Iterando tale procedimento n volte, la radice, se non ancora trovata, è contenuta in un intervallo di

lunghezza $\frac{b-a}{2^n}$. L'errore commesso nel considerare x_n come radice è :

$$|e_n| \leq h^n (b-a), \quad h = \frac{1}{2}$$

Un metodo si dice di ordine p se, detta α la radice ed $\{x_n\}$ le iterate, si ha:

$$|x_{n+1} - \alpha| \leq c|x_n - \alpha|^p, \quad n \geq 0, \quad c \neq f(p)$$

se $p = 1$ e $c < 1$: $|x_{n+1} - \alpha| \leq c|x_n - \alpha|$ e la convergenza si dice lineare.

Il metodo di bisezione è generalmente lineare poiché

$$|e_{n+1}| \leq h|e_n|$$

anche se ci possono essere dei casi particolari in cui non si ha una riduzione monotona dell'errore.

In un metodo iterativo, per decidere quando terminare i calcoli si può stimare l'errore relativo:

$$|E_r| = \left| \frac{x_{att} - x_{prec}}{x_{att}} \right|$$

dove x_{att} e x_{prec} sono le stime al passo attuale e a quello precedente.

Se $E_r < \text{accuratezza} \Rightarrow \text{STOP}$.

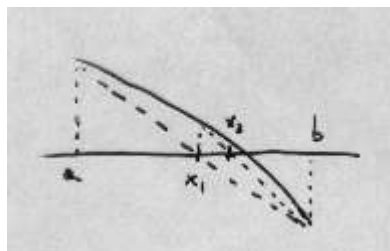
Lo svantaggio principale del metodo di bisezione è la lenta convergenza.

Un metodo più veloce è dato dal metodo della falsa posizione.

2. Metodo della falsa posizione (regula falsi)

Il metodo di bisezione divide a metà l'intervallo in cui è contenuta la radice senza tenere conto delle informazioni che possono dare i valori della f . Se ad esempio $f(x_a)$ è più vicina a zero di $f(x_b)$ è più probabile che x_a sia più vicina alla radice di x_b .

In tale metodo si sostituisce la f con una retta che determina una falsa posizione della radice.



La retta che unisce a e b è data da:

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

per $y = 0$ si ha x_1 :

$$x_1 = \frac{a \cdot f(b) - b \cdot f(a)}{f(b) - f(a)}$$

Si sostituisce a oppure b con x_1 , a seconda se è a o b ad avere la f con lo stesso segno di $f(x_1)$

Il rate di convergenza del metodo è lineare.

Esaminiamo ora i metodi aperti che hanno origine dai metodi di punto fisso.

3. Metodi di punto fisso

Data la funzione $f(x)$ di cui si vuole trovare, per $x \in [a, b]$, s tale che: $f(s) = 0$, costruiamo una funzione ausiliaria $g(x)$ tale che: $s = g(s)$ quando $f(s) = 0$.

Tale costruzione non è unica. Per esempio se:

$$f(x) = x^3 - 13x + 18 \quad \text{allora}$$

$$1) \quad g(x) = \frac{x^3 + 18}{13}$$

$$2) \quad g(x) = (13x - 18)^{1/3}$$

$$3) \quad g(x) = x^3 - 12x + 18$$

$$4) \quad g(x) = \frac{13x - 18}{x^2}$$

$$5) \quad g(x) = \sqrt{\frac{13x - 18}{x}}$$

Il problema di trovare s tale che : $s = g(s)$ è noto come problema del punto fisso ed s è detto punto fisso di $g(x)$.

TEOREMA DEL PUNTO FISSO

Sia $g([a,b]) \subseteq [a,b]$ ed inoltre

$$\forall x, y \in [a,b]: |g(x) - g(y)| \leq k|x - y| \text{ con } k < 1$$

$$\Rightarrow 1) \exists \alpha \in [a,b]: \alpha = g(\alpha)$$

$$2) \forall x_0 \in [a,b], x_{n+1} = g(x_n) \Rightarrow \lim_{n \rightarrow \infty} x_n = \alpha$$

$$3) |\alpha - x_n| \leq \frac{k^n}{1-k} (x_1 - x_0)$$

DIM.

1) Supponiamo $g \in C^1[a,b]$ allora dall'ipotesi: $|g'(x)| \leq k$

Poniamo: $\psi(x) = x - g(x)$

Si ha: $\psi(a) \leq 0$ poichè $g(a) \geq a$

$\psi(b) \geq 0$ poichè $g(b) \leq b$ poichè $g([a,b]) \subseteq [a,b]$

Se è valido l' = allora a oppure b è punto fisso.

Se non è valido l' = si applica il teorema di esistenza degli zeri.

Dimostriamo ora l'unicità del punto fisso.

Supponiamo per assurdo che $\exists \beta, \alpha \neq \beta : \alpha = g(\alpha)$

$$\beta = g(\beta)$$

$$|\alpha - \beta| = |g(\alpha) - g(\beta)| \text{ per Lagrange: } \leq |\alpha - \beta| \cdot |g'(\xi)| \leq k|\alpha - \beta| < |\alpha - \beta| \text{ poichè } k < 1$$

ma ciò è assurdo.

$$2) |\alpha - x_n| = |g(\alpha) - g(x_{n-1})| \leq \text{Lagrange} \leq |\alpha - x_{n-1}| \cdot |g'(\xi)| \leq k|\alpha - x_{n-1}| = k|g(\alpha) - g(x_{n-2})| \leq$$

$$\leq k^2 |\alpha - x_{n-2}| \leq \dots \leq k^n |\alpha - x_0|$$

Poichè x_0 è fissato $|\alpha - x_0|$ è una costante : $0 < |\alpha - x_0| < k^n$

Poichè $k < 1$, $\lim_{n \rightarrow \infty} k^n = 0$ e per il teorema del confronto: $\lim_{n \rightarrow \infty} |\alpha - x_n| = 0$, $\lim_{n \rightarrow \infty} x_n = \alpha$

$$\begin{aligned} 3) \quad |\alpha - x_n| &= |\alpha - g(x_n) + g(x_n) - x_n| \leq |\alpha - g(x_n)| + |g(x_n) - x_n| = |g(\alpha) - g(x_n)| + |g(x_n) - x_n| \leq \\ &\leq |g'(\xi)| |\alpha - x_n| + |g(x_n) - x_n| \end{aligned}$$

La prima si ha per Lagrange; la seconda, ricordando che $x_n = g(x_{n-1})$, dà:

$$\begin{aligned} |g(x_n) - x_n| &= |g(x_n) - g(x_{n-1})| \leq k |x_n - x_{n-1}| = k |g(x_n) - g(x_{n-1})| \leq \\ &\leq k^2 |x_{n-1} - x_{n-2}| \leq \dots \leq k^n |x_1 - x_0| \end{aligned}$$

Allora:

$$|\alpha - x_n| \leq k |\alpha - x_n| + k^n |x_1 - x_0| \Rightarrow$$

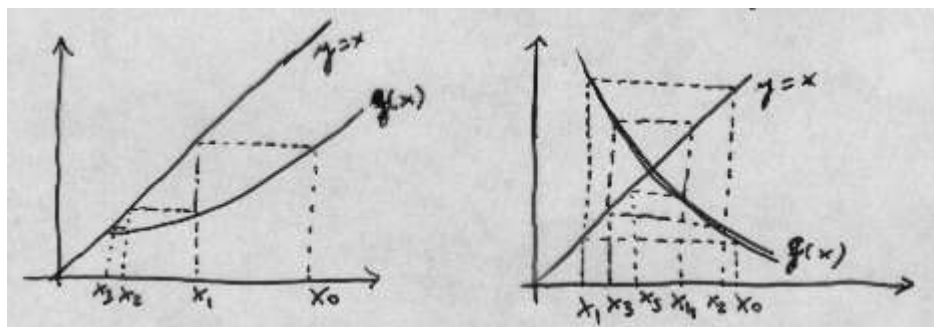
$$|\alpha - x_n| - k |\alpha - x_n| \leq k^n |x_1 - x_0| \Rightarrow$$

$$(1 - k) |\alpha - x_n| \leq k^n |x_1 - x_0| \Rightarrow |\alpha - x_n| \leq \frac{k^n}{(1 - k)} |x_1 - x_0| . \quad \bullet$$

Una funzione g definita da: $g([a, b]) \subseteq [a, b]$, $|g(x) - g(y)| \leq k|x - y|$ è detta **contrazione**.

SIGNIFICATO GEOMETRICO DEL TEOREMA DI PUNTO FISSO

Ogni radice reale di $x = g(x)$ è l'ascissa di un punto di intersezione di $y = g(x)$ con la retta $y = x$.



La forma a gradini o a spirale si ha se $g'(x) > 0$ oppure $g'(x) < 0$.

Algoritmo del punto fisso: $x_{n+1} = g(x_n)$

VELOCITA' DI CONVERGENZA DEI METODI DI PUNTO FISSO

$g \in C^{k+1}[a, b]$, $x^* = g(x^*)$, $e_n = x_n - x^*$, $g(x_n) = x_{n+1}$

$$e_{n+1} = x_{n+1} - x^* = g(x_n) - g(x^*) = g'(x^*)(x_n - x^*) + g''(x^*) \frac{(x_n - x^*)^2}{2} + \dots + g^{(k)}(x^*) \frac{(x_n - x^*)^k}{k!} + O(k+1)$$

Se: $g'(x^*) \neq 0$ si ha:

$$e_{n+1} = g'(\xi_n) e_n \quad \xi_n \in]x_n, x^*[$$

ed essendo g' continua in x^* : $\lim_{n \rightarrow \infty} \frac{e_{n+1}}{e_n} = g'(x^*)$

Tale rate di convergenza è detto lineare.

Se $g'(x^*) = 0$, $g''(x^*) \neq 0$ $\lim_{n \rightarrow \infty} \frac{e_{n+1}}{e_n^2} = g''(x^*)/2$ e il rate di convergenza è quadratico.

Più derivate si annullano in x^* , più veloce è il metodo. Il termine $\frac{g^{(k+1)}(x^*)}{(k+1)!}$ è detto fattore asintotico di convergenza.

Ricaviamo il numero r di iterazioni da eseguire perché l'errore sia $< 10^{-m} e_{n-1}$

Poniamo: $E_0 = e_{n-1}$, $e_{n-1+r} = E_r$, $\frac{|e_n|}{|e_{n-1}|^p} = C$

Si ha: $E_r \approx |E_{r-1}|^p C \approx C^{1+p} |E_{r-2}|^{p^2} \approx \dots \approx C^{1+p+\dots+p^{r-1}} |E_0|^{p^r}$

$$\text{cioè: } |E_r| \approx \begin{cases} C^r |E_0| & p = 1 \\ C^{\frac{p^r-1}{p-1}} |E_0|^{p^r} & p > 1 \end{cases}$$

$$|E_r| = 10^{-m} |E_0| = C^r |E_0| \quad p = 1$$

$$r \geq \frac{m}{-\lg C}$$

$$|E_r| = 10^{-m} |E_0| = C^{\frac{p^r-1}{p-1}} |E_0|^{p^r} \quad p > 1$$

$$p^r \geq 1 + \frac{m}{-\lg \left(C^{\frac{1}{p-1}} |E_0|^{p^r} \right)}$$

Dal confronto per $p = 1$ e $p > 1$ si ha che, se sono richieste 130 iterazioni nel metodo lineare per avere una data accuratezza, sono necessarie solo 7 iterazioni nel metodo quadratico.

4. Metodo di Newton o delle tangenti.

Sia x_0 un punto arbitrario di $[a, b]$ in cui è definita la $f(x)$ di cui si vuole determinare la radice.

Consideriamo la tangente alla $f(x)$ nel punto $(x_0, f(x_0))$

$$y - f(x_0) = f'(x_0)(x - x_0) \text{ e sia } f'(x_0) \neq 0.$$

x_1 sia dato dalla intersezione di tale retta con \bar{x}

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

ripetendo iterativamente si ha:

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}}$$

con $f'(x_n) \neq 0$. Se la radice è semplice il metodo è quadratico.

TEOREMA

$$f \in C^2[a, b], \exists \delta > 0, f'(x) \neq 0 \forall x \in]x^* - \delta, x^* + \delta[, f(x^*) = 0$$

$\Rightarrow \exists \varepsilon > 0, \{x_n\} \rightarrow x^*$ quadraticamente per $|x_0 - x^*| < \varepsilon$ ovvero:

$$\lim_{n \rightarrow \infty} \frac{e_{n+1}}{e_n^2} = c, \quad c = -\frac{f''(x^*)}{2f'(x^*)}$$

Dimostrazione.

Espandiamo la f attorno ad x^* :

$$f(x^*) = f(x_n) + (x^* - x_n)f'(x_n) + (x^* - x_n)^2 \frac{f''(\xi_n)}{2} \quad \xi_n \in]x^* - x_n, x^* + x_n[$$

Poiché: $e_n = x^* - x_n$, $f(x^*) = 0$ si ha: $0 = f(x_n) + e_n f'(x_n) + e_n^2 \frac{f''(\xi_n)}{2}$.

Da cui: $\frac{f(x_n)}{f'(x_n)} = -e_n - e_n^2 \frac{f''(\xi_n)}{2f'(x_n)}$.

Poiché il metodo è: $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ sostituendo, si ha:

$$x_{n+1} - x_n = e_n + e_n^2 \frac{f''(\xi_n)}{2f'(x_n)} = x_{n+1} - x^* + x^* - x_n = -e_{n+1} + e_n$$

$$\Rightarrow e_{n+1} = -e_n^2 \frac{f''(\xi_n)}{2f'(x_n)} \leq -e_n^2 \frac{\max|f''(\xi_n)|}{2 \min|f'(x_n)|} = -Me_n^2$$

dove: $M = \frac{\max|f''(\xi_n)|}{2 \min|f'(x_n)|}$

Pertanto: $|e_{n+1}| \leq M|e_n|^2$ da cui la convergenza quadratica.

La convergenza è garantita se: $M|e_0| \leq 1$

Infatti: $M|e_{n+1}| \leq M^2 e_n^2 \Rightarrow M|e_n| \leq M^2 e_{n-1}^2$

↑ ↓

$$|e_n| \leq Me_{n-1}^2 \quad M^2 e_n^2 \leq M^4 e_{n-1}^4$$

$M|e_{n+1}| \leq M^4 e_{n-1}^4 \leq \dots \leq (M|e_0|)^{2^{n+1}}$ da cui la convergenza se $M|e_0| < 1$.

Per stimare la convergenza al passo n-esimo basta considerare la differenza tra due iterate successive. Infatti, per il teorema di Lagrange:

$$f(x_n) - f(x^*) = f'(\xi_n)(x_n - x^*)$$

Ma: $f(x^*) = 0 \Rightarrow f(x_n) = f'(\xi_n)(x_n - x^*)$

$$e_n = x^* - x_n = -\frac{f(x_n)}{f'(\xi_n)} \approx -\frac{f(x_n)}{f'(x_n)} = x_{n+1} - x_n$$

5. Il metodo di Newton come metodo di punto fisso.

Cerchiamo $x^* : f(x^*) = 0$

e risolviamo tale problema come ricerca del punto fisso di $g(x) : x^* = g(x^*)$

Scelta “naturale” di $g(x)$: $g(x) = x + c \cdot f(x)$ con $c \neq 0$.

Abbiamo visto che l'iterazione di punto fisso: $x_{n+1} = g(x_n)$

è veloce se si annullano le derivate di $g(x)$.

Si ha: $g'(x^*) = 1 + c \cdot f'(x^*)$

Allora: $g'(x^*) \neq 0$ a meno che: $c = -\frac{1}{f'(x^*)}$

Però, non conoscendo x^* non conosciamo c .

Poniamo quindi:

$$g(x) = x + h(x) \cdot f(x) .$$

Poiché $f(x^*) = 0$ si ha: $g'(x^*) = 1 + h'(x^*) \cdot f(x^*) + h(x^*) \cdot f'(x^*) = 1 + h(x^*) \cdot f'(x^*) \equiv 0$

$$\Rightarrow h(x) = -\frac{1}{f'(x)}$$

$$\Rightarrow x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad \text{che è il metodo di Newton}$$

6. Radici Multiple

Se la radice di $f(x)$ non è semplice, la convergenza del metodo di Newton non è più quadratica ma lineare.

Ponendo però:

$$x_{n+1} = x_n - p \frac{f(x_n)}{f'(x_n)}$$

si può far vedere che il metodo è quadratico. È però di scarsa utilità perché solo raramente si conosce la molteplicità di una radice.

7. Metodo delle secanti

Una difficoltà del metodo di Newton è relativa al calcolo della $f'(x)$. Sostituendo la tangente con la secante si ha un metodo più lento ma meno oneroso:

$$x_{n+1} = x_n - f(x_n) \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}$$

da cui:

$$x_{n+1} = \frac{f(x_n)x_{n-1} - f(x_{n-1})x_n}{f(x_n) - f(x_{n-1})}$$

Poiché x_{n+1} dipende esplicitamente da x_n ed x_{n-1} non è una iterazione di punto fisso e non è un metodo chiuso poiché, sebbene siano richiesti due valori iniziali, non è richiesto un cambio di segno della $f(x)$. Tale metodo può quindi divergere.

Mostriamo che la convergenza è superlineare.

Poniamo: $e_n = \xi - x_n$ dove $\xi : f(\xi) = 0$. Vediamo se $e_n \rightarrow 0$.

Si ha, dalle differenze divise:

$$f(x) = f(x_0) + (x - x_0)f[x_0, x_1] + (x - x_0)(x - x_1)f[x_0, x_1, x]$$

Poniamo: $x = \xi$, $x_0 = x_n$, $x_1 = x_{n-1}$

$$f(\xi) = f(x_n) + (\xi - x_n)f[x_n, x_{n-1}] + (\xi - x_n)(\xi - x_{n-1})f[x_n, x_{n-1}, \xi] = 0$$

Dividiamo per $f[x_n, x_{n-1}]$:

$$f(\xi) = \frac{f(x_n)}{f[x_n, x_{n-1}]} + (\xi - x_n) + \frac{(\xi - x_n)(\xi - x_{n-1})}{f[x_n, x_{n-1}]} f[x_n, x_{n-1}, \xi] = 0$$

Ma: $x_n - x_{n+1} = \frac{f(x_n)}{f[x_n, x_{n-1}]}$ e quindi:

$$f(\xi) = x_n - x_{n+1} + \xi - x_n + (\xi - x_n)(\xi - x_{n-1}) \frac{f[x_n, x_{n-1}, \xi]}{f[x_n, x_{n-1}]} = 0$$

Se $f \in C^2[a, b]$ poiché: $f[x_0, x_1, \dots, x_n, x] = \frac{f''(\xi)}{n!}$

$$f[x_n, x_{n-1}, \xi] = \frac{f''(u)}{2!}$$

$$f[x_n, x_{n-1}] = f'(v)$$

Per cui:

$$e_{n+1} = -e_n e_{n-1} \frac{f''(u)}{2f'(v)}$$

Ponendo: $M = \frac{\max f''(x)}{2 \min f'(x)}$ si ha:

$$|e_{n+1}| \leq M |e_n| |e_{n-1}|$$

Poniamo: $d_n = M e_n$

$$M e_{n+1} \leq M^2 e_n e_{n-1} \Rightarrow d_{n+1} \leq d_n d_{n-1}$$

$$d_2 \leq d_0 d_1$$

Poniamo: $d = \max(d_0, d_1)$

$$d_2 \leq d^2, d_3 \leq d_2 d_1 \leq d^3, \dots, d_n \leq d^{2n} \text{ con:}$$

$$\alpha_{n+1} = \alpha_n + \alpha_{n-1}, \alpha_0 = 1, \alpha_1 = 1 \quad (1)$$

Tale relazione dà la successione di Fibonacci

Poniamo: $\alpha_n = r^n, (1) \Rightarrow r^{n+1} = r^n + r^{n-1} \Rightarrow r^2 - r - 1 = 0$

$$r = \frac{1 \pm \sqrt{5}}{2}, r_1 = \frac{1 + \sqrt{5}}{2}, r_2 = \frac{1 - \sqrt{5}}{2}$$

La soluzione generale è: $\alpha_n = c_1 (r_1)^n + c_2 (r_2)^n$

dove c_1 e c_2 si ricavano da:

$$c_1 + c_2 = 1 \equiv \alpha_0$$

$$c_1 r_1 + c_2 r_2 = 1 \equiv \alpha_1$$

$$\Rightarrow c_1 = \frac{1 + \sqrt{5}}{2\sqrt{5}}, c_2 = \frac{-1 - \sqrt{5}}{2\sqrt{5}}$$

Ovvero: $\alpha_n = \frac{1+\sqrt{5}}{2\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2\sqrt{5}}\right) \left(\frac{1-\sqrt{5}}{2}\right)^n$

Poiché $\left(\frac{1-\sqrt{5}}{2}\right)^n \rightarrow 0$ il rate di convergenza è determinato da r_1 . Inoltre:

$$M|e_n| < d^{\alpha_n}$$

Se $d < 1$ il metodo è convergente e ciò si ha se: $M|e_0| < 1$, $M|e_1| < 1$. Si può far vedere che r_1 è l'ordine di convergenza del metodo e poiché: $1 < r_1 < 2$ La convergenza è superlineare.

DIFFERENZA TRA IL METODO DI NEWTON E LE SECANTI

Il metodo di Newton necessita il calcolo della $f'(x)$. D'altronde il metodo delle secanti ha il rate di convergenza inferiore. Vediamo quando conviene l'uno o l'altro metodo.

Valutiamo il numero delle iterate per ottenere una certa precisione.

$$|e_n| \approx \frac{1}{M} (M|e_0|)^{2^n} \quad \text{Newton}$$

$$|e_n| \approx \frac{1}{M} (M|e_0|)^{r^n} \quad \text{secanti}$$

perché il metodo abbia accuratezza $\varepsilon : |e_n| \leq \varepsilon$

da cui: $n_N \geq \frac{k}{e_n^2}$

$$n_S \geq \frac{k}{e_n^r} \quad \text{con } k = \ln \frac{\ln \varepsilon^r}{\ln(M|e_0|)}$$

Sia m il tempo di calcolo di $f(x)$ ed m^* quello di $f'(x)$ ($t \neq 1$). Il tempo minimo per avere accuratezza ε è dato da :

$$T_N = (m + mt)n_N = (1+t) \frac{mk}{e_n^2}$$

$$T_S = mn_S = \frac{mk}{e_n^r}$$

$$\frac{T_S}{T_N} = \frac{e_n^2}{(1+t)e_n^r}, \text{ e quindi } T_S < T_N \text{ se:}$$

$$e_n^2 < (1+t)e_n^r \text{ ovvero:}$$

$$t > \frac{e_n^2}{e_n^r} - 1 \approx 0.44$$

Pertanto il metodo delle secanti è più veloce del metodo di Newton se la frazione di tempo del calcolo della f' è maggiore del 44% del tempo di calcolo della f .