11. Punto Fisso: iterazioni punto fisso, teorema delle contrazioni, convergenza locale, ordine di convergenza, Newton come iterazione di punto fisso

In questa lezione studieremo equazioni della forma

$$x = \phi(x), x \in I \subseteq \mathbb{R}$$

dove $\phi \in C(I)$ e I e' un intervallo chiuso (non necessariamente limitato) di $\mathbb{R}$ e la loro soluzione numerica tramite semplici iterazioni del tipo

$$x_{n+1} = \phi(x_n), n \geq 0, x_0 \in I$$

comunemente dette *iterazioni di punto fisso*.
In particolare, vedremo ipotesi che garantiscono la convergenza $x_n \to \xi, n \to \infty$ dove $\xi = \phi(\xi)$ e' detto punto fisso di $\phi \in I$ e studieremo l'ordine di convergenza dell'iterazione, scoprendo che il metodo di Newton puo' essere interpretato come iterazione di punto fisso.

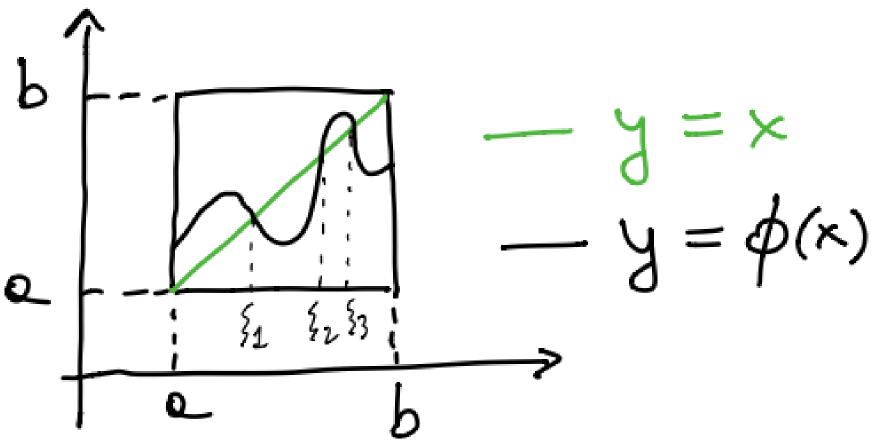
Enunciamo qui sotto un famoso teorema, detto *teorema delle contrazioni*.

> **Teorema** (esistenza e unicita' del punto fisso e convergenza delle iterazioni di punto fisso per una *contrazione*)
> Sia $\phi : I \subseteq \mathbb{R} \to \mathbb{R}$ una funzione derivabile nell'intervallo chiuso $I \subseteq \mathbb{R}$, tale che:
>
> 1. $\phi(I) \subseteq I$ cioe' l'immagine di I tramite ϕ, $\phi(I) = \{y : y = \phi(x), x \in I\}$, e' contenuta in I;
> 2. $\exists\ \Theta \in (0,1) : |\phi'(x)| < \Theta\ \forall\ x \in I \implies \exists! \xi \in I : \xi = \phi(\xi)$ (punto fisso) e $\forall\ x_0 \in I, \xi = lim_{n\to\infty} x_n$ dove $x_{n+1} = \phi(x_n), n \geq 0$.

Prima di dimostrare questo teorema (che puo' essere esteso ad ambiti molto piu' astratti, qui ci limitiamo a funzioni reali di variabile reale, come provato dal matematico polacco S. Banach nel 1919 facneodlo diventare uno dei risultati chiave dell'analisi matematica contemporanea), facciamo alcune osservazioni:

1. l'intervallo I e' assunto chiuso, ma puo' non essere limitato, cioe' $I = [a, b]$ con $-\infty < a < b < +\infty$ ma anche $I = [a, +\infty)$ oppure $I = (-\infty, b]$ o addirittura $I = \mathbb{R}$;
2. ϕ e' una contrazione (di I in se' stesso), cioe' contrae le distanze di un fattore $\Theta < 1$. Infatti per il teorema del valor medio $\forall\ x, y \in I$ vale $\phi(x) - \phi(y) = \phi'(z)(x - y), z \in \mathrm{int}(x, y) \subset I$ da cui $|\phi(x) - \phi(y)| = |\phi'(z)||x - y| \leq \Theta|x - y| < |x - y|$;
3. chiaramente la disuguaglianza appena provata, assunta direttamente come ipotesi, implica che ϕ e' continua in I, infatti $\forall\ x, \bar{x} \in I$ si ha che

$0 \leq |\phi(x) - \phi(\bar{x})| \leq \Theta|x - \bar{x}|$ e quindi per il teorema dei due carabinieri
$|\phi(x) - \phi(\bar{x})| \to 0, x \to \bar{x}$ che e' equivalente a dire che $\phi(x) \to \phi(\bar{x}), x \to \bar{x}$.

> **Dimostrazione** (del teorema)
> Cominciamo dimostrando l'esistenza di un punto fisso, limitandoci al caso di $[a, b]$ limitato:
> in questo caso basta l'ipotesi *1.* e la continuita' di ϕ (non server che ϕ sia una contrazione).
> Siccome ϕ e' continua, tale e' $f(x) = x - \phi(x)$ se $a = \phi(a)$ oppure $b = \phi(b)$ allora a oppure b sono punto fisso.
> Se invece $a \neq \phi(a)$ e $b \neq \phi(b)$ siccome $a \leq \phi(x) \leq b \,\forall\, x \in [a, b]$ si ha $a - \phi(a) < 0$ e $b - \phi(b) > 0$ cioe' f e' continua e cambia segno agli estremi $\implies \exists\, \xi \in (a, b) : f(\xi) = 0$ cioe' $\exists\, \xi \in (a, b) : \xi = \phi(\xi)$.
> La continuita' non basta pero' a garantire l'unicita' del punto fisso, come si vede da questo disegno
>
> 
>
>
> Ma se ϕ e' una contrazione, l'unicita' e' assicurata.
> Infatti se $\exists \xi_1, \xi_2 \in I$ con $\xi_1 \neq \xi_2$ tali che $\xi_1 = \phi(\xi_1)$ e $\xi_2 = \phi(\xi_2)$ allora $|\xi_1 - \xi_2| = |\phi(\xi_1) - \phi(\xi_2)| \leq \Theta|\xi_1 - \xi_2|$ cioe' $\Theta \geq 1$ contro l'ipotesi che $\Theta < 1$.
> Resta da dimostrare che $\forall\, x_0 \in I$, definendo $x_{x+1} = \phi(x_n), n \geq 0$ si ha $lim_{n\to\infty} x_n = \xi$
> ora $e_{n+1} = |x_{n+1} - \xi| = |\phi(x_n) - \phi(\xi)| \leq \Theta|x_n - \xi| = \Theta e_n$ da cui
> $e_1 \leq \Theta e_0, e_2 \leq \Theta e_1 \leq \Theta^2 e_0, \ldots, e_n \leq \Theta^n e_0 \to 0, n \to \infty$ perche' $\Theta \in (0, 1)$.

E' il caso di fare subito alcune osservazioni importanti:

1. nel teorema delle contrazioni, la dimostrazione generale e' basata sul fatto che la successione $\{x_n\}$ e' di Cauchy e quindi convergente a uno $\xi \in I$ (perche' I e' chiuso) che 'e automaticamente punto fisso perche' per continuita' di ϕ,

$$\xi = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \phi(x_n) = \phi(lim_{n\to\infty} x_n) = \phi(\xi)$$

   Ma anche con la dimostrazione scritta sopra, si ottiene la **stima a priori** dell'errore $e_n \leq \Theta^n e_0$. Se Θ (ed e_0) sono noti questa permette a priori di stabilire il numero di iterazioni sufficiente ad ottenere ξ con una tolleranza $\epsilon > 0$, risolvendo la disuguaglianza

$$\Theta^n e_0 \le \epsilon \iff e^{nlog\Theta} \le e^{log(\epsilon/e_0)} \iff n \ge \frac{\log(\epsilon/e_0)}{\log(\Theta)} = -\frac{\log(\epsilon/e_0)}{|\log\Theta|} = \frac{\log(e_0/\epsilon)}{|\log\Theta|}$$

dato che $\Theta \in (0,1)$ e $\log\Theta < 0$. Notiamo che la convergenza e la stima ottenute valgono $\forall x_0 \in I$, cioe' le iterazioni di punto fisso che costruiscono (infinite) successioni diverse l'una dalll'altra al variare di x_0, in ogni caso forniscono successioni che convergono tutte allo *stesso limite* che e' *l'unico punto fisso* di ϕ in I, con una convergenza che e' almento lineare perche' $e_{n+1} \le \Theta e_n$;

2. si puo' facilmente ottenre una **stima a posteriori** dell'errore che spesso e' piu' precisa della stima a priori: basta infatti scrivere

$$x_{n+1} - \xi = x_{n+1} - x_n + x_n - \xi$$

ma

$$x_{n+1} - \xi = \phi(x_n) - \phi(\xi) = \phi'(z_n)(x_n - \xi), \quad z_n \in \text{int}(x_n, \xi)$$

da cui

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

ovvero

$$(1 - \phi'(z_n))(x_n - \xi) = x_n - x_{n+1}$$
$$\implies \frac{|x_{n+1} - x_n|}{1 - \phi'(z_n)} = e_n \le \underbrace{\frac{|x_{n+1} - x_n|}{1 - \Theta}}_{\text{prima stima a posteriori}}$$

perche' $|\phi'(z_n)| \le \Theta < 1$

$$\implies |1 - \phi'(z_n)| \ge |1 - |\phi'(z_n)|| \ge 1 - \Theta > 0 \quad \text{e} \quad \frac{1}{|1 - \phi'(z_n)|} \le \frac{1}{1 - \Theta}$$

In pratica abbiamo fatto vedere che l'errore e' stimato dallo $\textit{step} = |x_{n+1} - x_n|$ a meno del fattore $\frac{1}{1-\Theta} = \textit{peso}$.
Se Θ e' piccolo lo step diventa da solo una buona stima dell'errore perche' il perso e' ≈ 1; invece se Θ e' vicino ad 1 lo step va corretto per evitare una possibile sottostima, ma se $\phi \in C^1(I)$, siccome $z_n \to \xi$ allora $\phi'(z_n) \to \phi'(\xi), n \to \infty$, quindi una stima a posteriori migliore e' tendenzialmente la stima empirica (almeno per n abbastanza grande)

$$e_n = \frac{|x_{n+1} - x_n|}{1 - \phi'(z_n)} \approx \underbrace{\frac{x_{n+1} - x_n}{1 - \phi'(x_n)}}_{\text{seconda stima a posteriori}}$$

Esempio

Consideriamo l'equazione di punto fisso

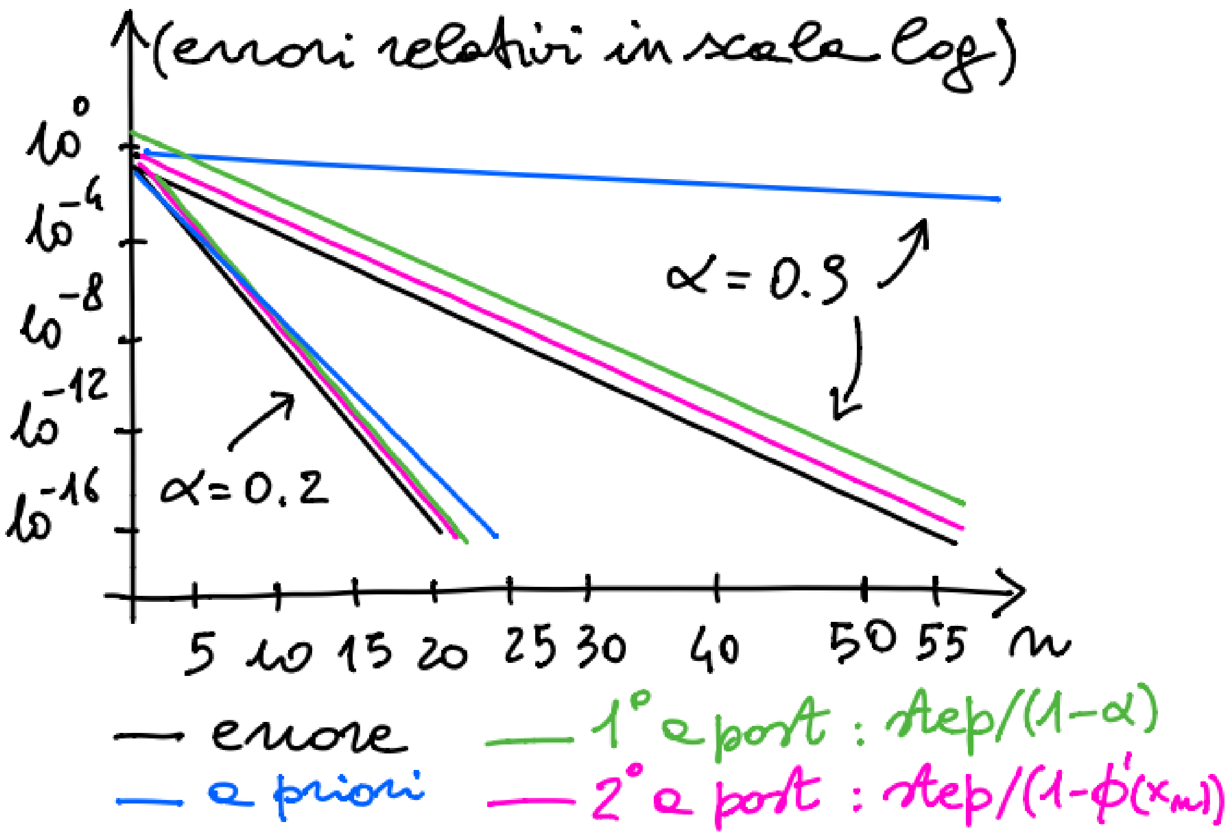
$$x = \phi(x) = e^{-\alpha x}, \alpha > 0$$

se $\alpha < 1, |\phi'(x)| = |-\alpha e^{-\alpha x}| \le \alpha < 1$ e quindi ϕ contrae le distanze, quindi l'ipotesi 2. del teorema delle contrazioni e' soddisfatta con $\Theta = \alpha$; d'altra parte, $0 < \phi(x) \le 1, \forall x \in [0, +\infty)$, quindi anche l'ipotesi 1. e' soddisfatta con $I = [0, +\infty)$.
In realta' siccome $\phi'(x) = -\alpha e^{-\alpha x} < 0$, ϕ e' strettamente descrescente (e positiva), quinid visto che $\phi(0) = 1$ e $0 < \phi(1) = 1 - e^{-\alpha} < 1$ si ha $0 < \phi(x) \le 1, \forall x \in [0, 1]$ cioe' ϕ e' anche una contrazione di $[a, b] = [0, 1]$ in se stesso.

Allora $\exists!$ ξ punto fisso di $\phi \in [0, 1]$: prendiamo $x_0 = \dfrac{1}{2} \implies e_0 = |x_0 - \xi| \le \dfrac{1}{2}$ e la successione $x_{n+1} = \phi(x_n),\ n \ge 0$ converge a ξ con la stima a priori dell'errore $e_n \le \dfrac{\alpha^n}{2}$.
Nel grafico qui sotto mostriamo l'errore effettivo, la stima a priori e la stima a posteriori dello step (non pesato) e dello step pesato da $\dfrac{1}{1-\alpha}$ e da $\dfrac{1}{1-\phi'(x_n)}$ con

$$\alpha = 0.2 \implies \mathrm{fl}(\xi) = 0.8445798\ldots$$
$$\alpha = 0.9 \implies \mathrm{fl}(\xi) = 0.5887032\ldots$$

(questi valori sono stati calcolati col metodo di Newton)



Si vede chiaramente che la convergenza e' lineare, che la stima a priori e' una sovrastima, molto distante dall'errore per $\alpha = 0.9$: infatti
$\dfrac{e_{n+1}}{e_n} = |\phi'(z_n)| \to |\phi'(\xi)|, n \to \infty$ cioe' $|\phi'(\xi)| = \alpha e^{-\alpha \xi} = L$ e' la costante asintotica, il parametro che effettivamente regola la velocita' di convergenza (α e' solo una stima) perche' per n abbastanza grande $e_{n+1} \approx L e_n$.
Qui per $\alpha = 0.9$ si ha $L \approx 0.53$ che e' ben minore di α, mentre per $\alpha = 0.2$ si ha $L \approx 0.17$ che e' poco minore di α.

In effetti la seconda stima a posteriori, $e_n \approx \frac{|x_{n+1}-x_n|}{1-\phi'(x_n)}$, e' una stima aderente dell'errore ma e' shiftata in avanti di 1, fenomeno che abbiamo gia' visto nella stima con lo step nel motodo di Newton, perche' per stimare e_n bisogna essere al passo $n+1$ ($step$=$|x_{n+1}-x_n|$).

Dopo aver discusso questo esempio semplice ma significativo, vediamo che anche per le iterazioni punto fisso vale un risultato di *convergenza locale* (mentre la formulazione generale del teorema delle contrazioni ha carattere *globale*, visto che $x_0 \in I$ e' arbitrario e per la convergenza non e' importante che x_0 sia vicino a ξ).

> **Teorema** (convergenza locale delle iterazioni di punto fisso)
> Sia ξ punto fisso di $\phi \in C^1(I_\delta(\xi))$ dove $I_\delta(\xi)=[\xi-\delta,\xi+\delta], \delta>0$ e sia $|\phi'(\xi)|<1$ allora
>
> $$\exists\delta' \leq \delta : x_{n+1}=\phi(x_n), n \geq 0$$
>
> converge a $\xi, \forall x_0 \in I_{\delta'}(\xi)$.

E' chiaro il carattere *locale* di questo risultato, che fornisce condizioni sufficienti per la convergenza delle iterazioni di punto fisso purche' x_0 *sia abbastanza vicino a* ξ.

Come per tutti i metodi iterativi, e' importante capire quale sia l'*ordine di convergenza* delle iterazioni di punto fisso.
Abbiamo gia' osservato che nel caso di una contrazione l'ordine e' almento $p=1$ perche' vale $e_{n+1} \leq \Theta e_n$ con $\Theta \in (0,1)$.
D'altra parte, nell'esempio svolto prima abbiamo fatto vedere che l'ordine e' esattamente $p=1$ se $\phi'(\xi) \neq 0$ con costante asintotica $L=|\phi'(\xi)|$. Diamo ora una caratterizzazione completa col seguente teorema.

> **Teorema** (ordine di convergenza delle iterazioni di punto fisso)
> Sia ξ punto fisso di $\phi \in C^p(I), p \geq 1$, con I intervallo di $\mathbb{R}$ e supponiamo di essere in ipotesi che garantiscano la convergenza a ξ di $x_{n+1}=\phi(x_n), n \geq 0, x_0 \in I$ (ad esempio le ipotesi del teorema delle contrazioni). Allora:
>
> 1. $\{x_n\}$ ha ordine esattamente $p=1 \iff 0<|\phi'(\xi)|<1$;
> 2. $\{x_n\}$ ha ordine esattamente $p>1 \iff \phi^{(j)}(\xi)=0, 1 \leq j \leq p-1$ e $\phi^{(p)}(\xi) \neq 0$.

> **Dimostrazione**
> 1. si dimostra subito visto che $e_{n+1}=|\phi'(z_n)|e_n, z_n \in \text{int}(\xi, x_n)$ per il teorema del valor medio, quindi $\lim_{n\to\infty} \frac{e_{n+1}}{e_n}=|\phi'(\lim z_)|=|\phi'(\xi)|$.
> Per il 2. utilizziamo la formula di Taylor di grado $p-1$ centrata in ξ con resto p-esimo in forma di Lagrange
>
> $$x_{n+1}=\phi(x_n)=\phi(\xi)+\phi'(\xi)(x_n-\xi)+\frac{\phi''(\xi)}{2}(x_n-\xi)^2+\cdots$$
> $$\cdots+\frac{\phi^{(p-1)}(\xi)}{(p-1)!}(x_n-\xi)^{p-1}+\frac{\phi^{(p)}(u_n)}{p!}(x_n-\xi)^p, u_n \int \text{int}(\xi,x_n)$$

> Dimostriamo prima "$\Longleftarrow$" (condizione sufficiente): se $\phi^{(j)}(\xi) = 0, 1 \le j \le p-1$ e $\phi^{(p)}(\xi) \neq 0$, da Taylor $x_{n+1} - \xi = \frac{\phi^{(p)}(u_n)}{p!}(x_n - \xi)^p$ e passando ai moduli
>
> $$\frac{e_{n+1}}{e_n^p} = \frac{|\phi^{(p)}(u_n)|}{p!} \to \frac{\phi^{(p)}(\xi)}{p!} \neq 0, n \to \infty$$
>
> perche' $u_n \to \xi, n \to \infty$ e $\phi^{(p)}$ e' continua quindi $\lim \phi^{(p)}(u_n) = \phi^{(p)}(\lim u_n) = \phi^{(p)}(\xi)$ ovvero $\{x_n\}$ ha ordine esattamente p.
> Per dimostrare "$\Longrightarrow$" (condizione necessaria) supponiamo per assurdo che $\{x_n\}$ abbia ordine esattamente p ma che $\exists j < p$ tale che $\phi^{(j)} \neq 0$.
> Prendiamo $k = \min\{j < p : \phi^{(j)}(\xi) \neq 0\}$ e scriviamo $\frac{e_{n+1}}{e_n^p} = \frac{e_{n+1}}{e_n^k} e_n^{k-p}$. Ora per ipotesi $\frac{e_{n+1}}{e_n^p} \to L \neq 0$, d'altra parte, con lo stesso ragionamento usato per la dimostrazione della condizione sufficiente, tramite la formula di Taylor si avrebbe
>
> $$\frac{e_{n+1}}{e_n^k} \to \frac{|\phi^{(k)}(\xi)|}{k!} = L' \neq 0$$
>
> ma allora $\frac{e_{n+1}}{e_n^p} = \underbrace{\frac{e_{n+1}}{e_n^k}}_{L'} \underbrace{e_n^{k-p}}_{\infty}$ perche' $k - p < 0$ e $e_n \to 0$ cioe' alla fine $\frac{e_{n+1}}{e_n^p} \to \infty, n \to \infty$, contraddicendo l'ipotesi che abbia limite finito.

Ribadiamo che la condizione data e' *necessaria* e *sufficiente*, cioe' fornisce una caratterizzazione completa di quando le iterazioni di punto fisso hanno ordine $p \ge 1$.
In particolare (e questo ci servira' tra poco) le iterazioni di punto fisso possono avere convergenza *quadratica*, *cubica*, etc...

Per concludere il capitolo mostriamo infatti che il metodo di Newton si puo' re-interpretare come iterazione di punto fisso e che in base a quanto visto sopra si vede subito che ammette convergenza locale e che la convergenza e' quadratica per zeri semplici.
Infatti l'iterazione di Newton

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, n \ge 0$$

e' di tipo punto fisso con $\phi(x) = x - \frac{f(x)}{f'(x)}$ e se $f \in C^k(I)$ con $f'(x) \neq 0, \forall x \in I$ intervallo, allora $\phi \in C^{k-1}(I)$.
E' evidente che ξ e' zero (semplice) di $f \iff \xi$ e' punto fisso di ϕ.
Inoltre, il teorema di convergenza locale per Newton (zeri semplici) discende immediatamente dal teorema di convergenza locale per le iterazioni di punto fisso se $f \in C^2(I_\delta(\xi))$, infatti

$$\phi'(x) = \frac{d}{dx}(x - \frac{f}{f'}) = 1 - \underbrace{\frac{(f')^2 - ff''}{(f')^2}}_{\text{derivata del rapporto}} = \frac{ff''}{(f')^2}$$

quindi $f(\xi) = 0 \implies \phi'(\xi) = 0$ e $|\phi'(\xi)| = 0 < 1$, allora $\exists \delta' \leq \delta$ tale che l'iterazione di Newton converge come iterazione di punto fisso $\forall x_0 \in I_{\delta'}(\xi)$.

D'altra parte, sempre interfpretando Newton come iterazione di punto fisso, e' immediato che la convergenza per zeri semplici e' almeno quadratica perche' $\phi'(\xi) = 0$ ed e' esattamente quadratica se $f''(\xi) \neq 0$, utilizzando la caratterizzazione vista sopra (che pero' richiede $\phi \in C^2$ e quindi $f \in C^3$).
Infatti

$$\phi'' = \frac{d}{dx}(\frac{ff''}{(f')^2}) = \frac{(f'f'' + ff''')(f')^2 - 2f'f''(ff'')}{(f')^4}$$

da cui $\phi''(\xi) = \dfrac{f''(\xi)}{f'(\xi)} \neq 0$ perche' $f''(\xi) \neq 0$ (e $f'(\xi) \neq 0$).

Non e' difficile vedere che interpretando Newton come iterazione di punto fisso, nel caso di zero multiplo l'ordine di convergenza diventa $p = 1$ con costante asintotica $|\phi'(\xi)| = 1 - \dfrac{1}{m}$ dove m e' la molteplicita' di ξ (il numero di derivate successive che si annullano in ξ) e che se m e' nota l'iterazione

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

torna di ordine almeno $p = 2$ perche' $\phi'_m = \lim_{x \to \xi} \phi'_m(x) = 0$.