Αριθμητική

Επίλυση μη-γραμμικών εξισώσεων (και συστημάτων)

Θα βρούμε ακολουθίες ~~ακολουθίες~~ αριθμών $(x_n)=\{x_1,x_2,x_3,\dots\}$ που συγκλίνουν στη ρίζα r της εξίσωσης $f(x)=0$. Προσεγγίζουμε τη ρίζα από κάποιον όρο x_N, για "μεγάλο" N. Η ταχύτητα σύγκλισης εξαρτάται από την $f(x)$ και τη χρησιμοποιούμενη μέθοδο. Γενικά συμφέρει να αρχίσουμε την επαναληπτική διαδικασία με αρχική τιμή που βρίσκεται κοντά στη ρίζα r.

π.χ. Αν $f(x)=f_1(x)-f_2(x)$ τότε $f(x)=0 \Leftrightarrow f_1(x)=f_2(x)$, άρα σχεδιάζουμε τις $f_1(x), f_2(x)$ και τα σημεία τομής τους δίνουν τα r (έστω κατά προσέγγιση να τα βρούμε)

Ένας αριθμός λέγεται αλγεβρικός αν είναι ρίζα πολυωνύμου με ακέραιους συντελεστές. Ο βαθμός του πολυωνύμου λέγεται βαθμός του αλγεβρικού αριθμού. Η αντίστοιχη πολυωνυμική εξίσωση λέγεται ελαχιστική (του αλγεβρικού αριθμού) αν δεν υπάρχει άλλο πολυώνυμο μικρότερου βαθμού που να έχει τον αλγεβρικό αριθμό ρίζα.

(οι αλγεβρικοί αριθμοί είναι αριθμήσιμοι ή απαριθμητοί-countable, όπως και οι ρητοί)

π.χ. η εξίσωση $x^2-2=0$ είναι η ελαχιστική εξίσωση του αλγεβρικού αριθμού $\sqrt{2}$ του οποίου ο βαθμός είναι 2.

π.χ. η εξίσωση $x^2-x-1=0$ είναι η ελαχιστική εξίσωση του αλγεβρικού αριθμού $(\sqrt{5}+1)/2$ του οποίου ο βαθμός είναι 2.

Υπερβατικός (transcendental) αριθμός είναι το αντίθετο του αλγεβρικού αριθμού, δηλ. δεν είναι ρίζα κανενός πολυωνύμου με ακέραιους συντελεστές.

π.χ. e ← Hermite, π ← Lindemann, $\xi=\sum_{n=1}^{\infty}10^{-n!}$ ← Liouville, $2^{\sqrt{2}}$ ← Hilbert, 0.123456789101112131415... ← Champernowne

(οι πραγματικοί αριθμοί δεν είναι αριθμήσιμοι, δηλ. κυρίως αποτελούνται από υπερβατικούς αριθμούς).

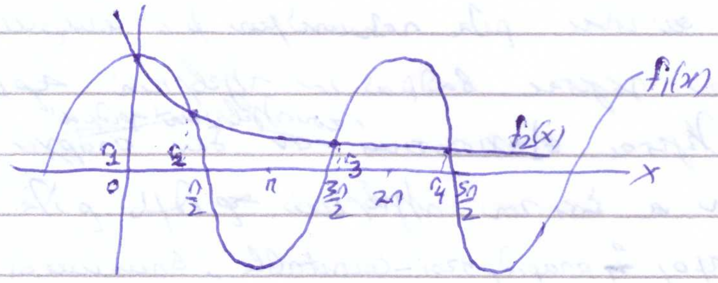
Υπερβατικές εξισώσεις είναι αυτές που περιέχουν υπερβατικές συναρτήσεις, δηλ. τριγωνομετρικές, λογαριθμικές, εκθετικές κλπ., π.χ. $f(x)=\sin x+e^{-x^2}+\ln x-1$, και εν γένει δεν μπορούν να λυθούν αναλυτικά.

Στα υπολογιστικά μαθηματικά ενδιαφερόμαστε να λύσουμε αριθμητικά αλγεβρικές και υπερβατικές εξισώσεις.

Μέθοδοι ολικής σύγκλισης όταν δεν απαιτείται η αρχική τιμή να είναι κοντά στη ρίζα.

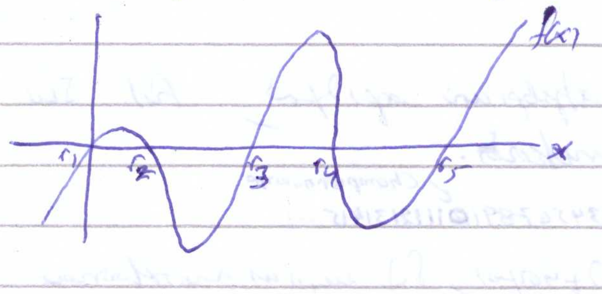
Μέθοδοι τοπικής σύγκλισης το αντίθετο.

$\longrightarrow \quad f(x) = \cos x - e^{-0.5x} = f_1(x) - f_2(x)$



Άρα εύκολα βλέπουμε ότι $r_1 = 0$, $0 < r_2 < \frac{\pi}{2}$, $r_3 > \frac{3\pi}{2}$, $r_4 < \frac{5\pi}{2}$, ... άπειρες ρίζες κατά προσέγγιση $\frac{(2m+1)\pi}{2}$, $m = 3, 4, \dots$

Η $f(x)$ είναι συνεχής και παραγωγίσιμη αλλά και εδώ δεν έχουμε πληροφορίες για το πού ακριβώς είναι οι ρίζες



Αν $a_n \to 0$, $x_n \to x^*$ και $|x_n - x^*| \le c|a_n|$ για n μεγάλα, λέμε ότι η x_n συγκλίνει στο x^* με ρυθμό σύγκλισης $O(a_n)$

Αν $x_n \to x^*$ και $\dfrac{|x_{n+1} - x^*|}{|x_n - x^*|^{\alpha}} \xrightarrow[n\to\infty]{} c$ λέμε ότι η x_n συγκλίνει με τάξη α στο x^* με ασυμπτωτική σταθερά σφάλματος c. (αν το α είναι μεγαλύτερο η ακολουθία συγκλίνει ταχύτερα). (το c εδώ είναι πιο σημαντικό για τη σύγκλιση).

Για $\alpha=1$ λέμε ότι η x_n συγκλίνει γραμμικά
Για $\alpha=2$ λέμε ότι η x_n συγκλίνει τετραγωνικά

Μια επαναληπτική διαδικασία $x_{n+1} = g(x_n)$ λέμε ότι είναι τάξης α αν η x_n συγκλίνει στη λύση $r = g(r)$ με τάξη α.

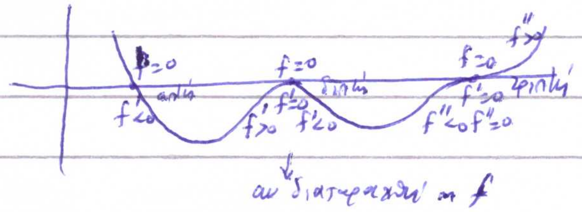
Η ρίζα r της εξίσωσης $f(x)=0$ λέγεται ότι έχει πολλαπλότητα m αν $f(x) = (x-r)^m q(x)$, $\lim_{x\to r} q(x) \neq 0$.

π.χ. για την $f(x) = x^6 - 14x^5 + 80x^4 - 238x^3 + 387x^2 - 324x + 108$
$= (x-1)(x-2)^2(x-3)^3$

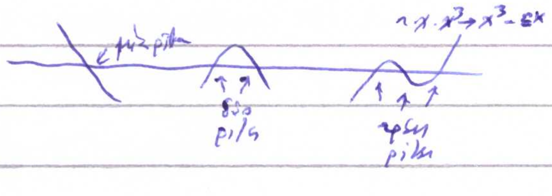
$f(x) = (x-r_1)^1 q(x)$, $r_1 = 1$ ← απλή, $q(x) = (x-2)^2(x-3)^3$
$f(x) = (x-r_2)^2 q(x)$, $r_2 = 2$ ← διπλή, $q(x) = (x-1)(x-3)^3$
$f(x) = (x-r_3)^3 q(x)$, $r_3 = 3$ ← τριπλή, $q(x) = (x-1)(x-2)^2$

Η $f \in C^m[a,b]$ έχει μια ρίζα πολλαπλότητας m αν $f(r) = f'(r) = \dots = f^{(m-1)}(r) = 0$, $f^{(m)}(r) \neq 0$.

π.χ.



αν διαταραχθεί η f



Η ικανοποίηση του θεωρήματος του Bolzano αποτελεί ικανό (όχι αναγκαίο) κριτήριο ύπαρξης ρίζας στο αντίστοιχο διάστημα, δηλ. αν $f(a)f(b)<0$, f συνεχής, τότε $\exists\, r \in (a,b)$ (το πλήθος των ριζών είναι περιττό), με $f(r)=0$

Αν $f(a)f(b)>0$ τότε ή $\nexists$ ρίζα ή ο αριθμός ριζών είναι άρτιος.

Το πλήθος N των απλών ριζών της $f(x)=0$ στο (a,b) είναι

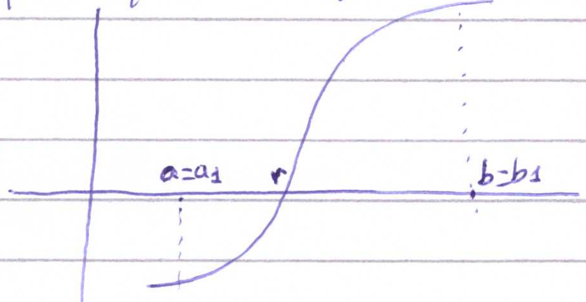
$$N = -\frac{\gamma}{\pi}\int_a^b \frac{f(x)f''(x) - f'(x)^2}{f^2(x) + \gamma^2 f'(x)^2}\,dx + \frac{1}{\pi}\arctan\left(\frac{\gamma\,[f(a)f'(b) - f(b)f'(a)]}{f(a)f(b) + \gamma^2 f'(a)f'(b)}\right), \quad \gamma > 0 \text{ μικρός π.χ. } \gamma \simeq 1$$

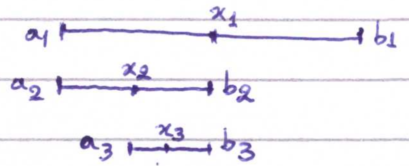
Για ρίζες μεγαλύτερης πολλαπλότητας υπάρχουν αντίστοιχοι τύποι.

Μέθοδος Διχοτόμησης

~~Μέθοδος ολικής~~ σύγκλιση $f(a)f(b)<0$ $\quad$ f συνεχής (υποθέσεις)

Είναι μέθοδος ολικής σύγκλισης ~~σύγκλιση για οποιοδήποτε~~ διότι δεν χρειάζεται αρχική εκτίμηση της ρίζας, συγκλίνει πάντα, ο αριθμός των επαναλήψεων για να προσεγγιστεί η ρίζα με δοσμένη ακρίβεια είναι προκαθορισμένος, απαιτεί έναν υπολογισμό της f ανά βήμα, δυστυχώς η σύγκλιση είναι αργή (μεγάλο κόστος), στην πράξη χρησιμοποιείται για έναν χονδρικό υπολογισμό της ρίζας και μετά καταφεύγουμε σε άλλη μέθοδο. Επίσης αν υπάρχουν περισσότερες της μίας ρίζες, η μέθοδος προσεγγίζει κάποια, άγνωστο εκ των προτέρων. Η μέθοδος δεν μπορεί να εφαρμοστεί όταν η ρίζα έχει πολλαπλότητα άρτιο αριθμό.





$x_1 = \frac{a_1+b_1}{2}$, $f(a_1)f(x_1)<0 \rightarrow [a_2,b_2]=[a_1,x_1]$

$x_2 = \frac{a_2+b_2}{2}$, $f(x_2)f(b_2)<0 \rightarrow [a_3,b_3]=[x_2,b_2]$

$x_3 = \frac{a_3+b_3}{2}$, ...

Δημιουργείται μια ακολουθία διαστημάτων $[a_1,b_1]\supset[a_2,b_2]\supset[a_3,b_3]\supset\dots\supset[a_n,b_n]\supset\dots$ (εκτός αν τύχει το r να είναι το μέσο κάποιου διαστήματος, οπότε σταματά η ακολουθία), που όλα περιέχουν τη ρίζα r, δηλ.

$a\le a_1\le a_2\le\dots\le a_{n-1}\le a_n\le x_n\le b_n\le b_{n-1}\le\dots\le b_2\le b_1\le b$, $\quad x_n=\frac{a_n+b_n}{2}$

και $f(a_n)f(b_n)<0$

Η ακολουθία (a_n) των αριστερών άκρων των διαστημάτων είναι αύξουσα και φραγμένη, ενώ η (b_n) των δεξιών άκρων είναι φθίνουσα και φραγμένη και ~~άρα~~ υπάρχει

$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = x^*$ $\quad$ $(\operatorname{sgn} f(a_n)=\operatorname{sgn} f(a),\ \operatorname{sgn} f(b_n)=\operatorname{sgn} f(b))$

Το κάθε επόμενο διάστημα έχει μήκος το μισό του προηγούμενου του, δηλ. $b_2-a_2=\frac{b_1-a_1}{2}$, $b_3-a_3=\frac{b_2-a_2}{2}=\frac{b_1-a_1}{2^2}$, ... άρα $b_{n+1}-a_{n+1}=\frac{b-a}{2^n}$.

f συνεχής, άρα $f(a_n)f(b_n)<0 \Rightarrow \lim_{n\to\infty}[f(a_n)f(b_n)]\le 0 \Rightarrow \lim_{n\to\infty} f(a_n)\cdot\lim_{n\to\infty} f(b_n)\le 0 \Rightarrow f(\lim_{n\to\infty} a_n)\cdot f(\lim_{n\to\infty} b_n)\le 0$

$\Rightarrow f(x^*)\cdot f(x^*)\le 0 \Rightarrow f(x^*)^2\le 0 \Rightarrow f(x^*)=0 \Rightarrow x^*=r$.

Δηλαδή η ακολουθία (x_n) (όπως και οι $(a_n),(b_n)$) συγκλίνει στη ρίζα r.

Ισχύει $|x_n - r| \le \frac{b-a}{2^n}$

διότι $|x_n - r| = \left|\frac{a_n+b_n}{2} - r\right| = \left|\frac{a_n+b_n-2r}{2}\right| = \left|\frac{b_n-a_n-2(r-a_n)}{2}\right| \overset{a_n\le r}{\le} \left|\frac{b_n-a_n}{2}\right| = \frac{b_n-a_n}{2} = \frac{b-a}{2^n}$

~~Προφανώς~~ $\cancel{|x_n - r| = \left|\frac{a_n+b_n}{2} - r\right| \le \frac{b_n-a_n}{2}}$

Αν $N \ge \frac{\log_{10}((b-a)\varepsilon^{-1})}{\log_{10}2}$ $\left(\Leftrightarrow N = \lceil \log_2((b-a)\varepsilon^{-1})\rceil\right)$ τότε $|x_N - r| \le \varepsilon$

Πράγματι αν $N \ge \frac{\log((b-a)\varepsilon^{-1})}{\log 2} \Rightarrow N\log 2 \ge \log((b-a)\varepsilon^{-1}) \Rightarrow$

$\log 2^N \ge \log((b-a)\varepsilon^{-1}) \Rightarrow 2^N \ge (b-a)\varepsilon^{-1} \Rightarrow \frac{b-a}{2^N} \le \varepsilon$, Αλλά

$|x_N - r| \le \frac{b-a}{2^N} \Rightarrow |x_N - r| \le \varepsilon$

Για $\varepsilon = 10^{-\alpha} \Rightarrow N \ge \frac{\log_{10}(b-a)+\alpha}{\log 2}$

~~Άλλο~~ κριτήριο τερματισμού της διαδικασίας μπορεί να είναι $\frac{|x_{n+1}-x_n|}{|x_{n+1}|} < \varepsilon$.

Το $|x_{n+1} - x_n| < \varepsilon$ δεν είναι καλό κριτήριο γιατί η $x_n = \sum_{i=1}^{n} \frac{1}{i}$ αποκλίνει ενώ $\lim_{n\to\infty}(x_{n+1}-x_n) = 0$

Επίσης το $|f(x_n)| < \varepsilon$ δεν είναι καλό κριτήριο γιατί μπορεί το x_n να είναι πολύ μακριά από το r και το $f(x_n)$ να είναι πολύ μικρό, π.χ. $f(x) = (x^5-1)^8$, όταν $x_n = \frac{1}{2}$ είναι $x_n \ne r = 1$, αλλά $f(x_n) = 9.1\times 10^{-13} \lll 1$

Ισχύει $\frac{|x_{n+1}-r|}{|x_n - r|} \le \frac{1}{2}$ (??? proof), άρα η μέθοδος συγκλίνει γραμμικά και σε κάθε επανάληψη βρίσκει ένα δυαδικό ψηφίο της λύσης.

π.χ. Να βρεθεί η $\sqrt[3]{7}$ με ακρίβεια $\varepsilon = 10^{-3}$

$f(x) = x^3 - 7 = 0$

$a=0,\ b=3,\ f(0)f(3)<0$

Για $\varepsilon=10^{-3}$, $N=\left\lceil \frac{\log((b-a)\varepsilon^{-1})}{\log 2}\right\rceil \simeq \lceil 11.550746\rceil = 12$,

άρα χρειάζονται 12 επαναλήψεις.

Διαλέγοντας τον κανόνα έχουμε $x_{12}=\frac{a_{12}+b_{12}}{2}=$

$=\frac{1.911621+1.913086}{2}=1.912354$

Η ακριβής $r=\sqrt[3]{7}=1.912931$

Το σφάλμα $|x_{12}-r|=|1.912354-1.912931|=0.577\times10^{-3}<10^{-3}$ σφάλμα.

α.χ. $f(x)=x^3+x^2-3x-3=0$, $[1,2]$

Για 3 επαναλήψεις $x_3=?$

Πόσες επαναλήψεις χρειάζονται για σφάλμα $\varepsilon=10^{-6}$

$f(1)f(2)=(-4)\times3=-12<0 \Rightarrow \exists r\in[1,2]=[a_1,b_1]$

$x_1=\frac{1+2}{2}=1.5$

Στο $[1,1.5]$, $f(1)f(1.5)=(-4)\times(-1.875)>0$

Στο $[1.5,2]=[a_2,b_2]$, $x_2=\frac{a_2+b_2}{2}=\frac{1.5+2}{2}=1.75$

Στο $[1.5,1.75]=[a_3,b_3]$, $f(a_3)f(b_3)=f(1.5)f(1.75)=(-1.875)\times(0.1718)<0$,

$x_3=\frac{a_3+b_3}{2}=\frac{1.5+1.75}{2}=1.625 \simeq r=1.732051$

Για $\alpha=6$, $N\geq\frac{\log(b-a)+\alpha}{\log2}=\frac{6}{\log2}=19.9315\Rightarrow N=20$, Άρα χρειάζονται

20 επαναλήψεις για σφάλμα $\varepsilon\leq10^{-6}$.

Ισχύει $x_{n+1}=x_n+\frac{h}{2^{n+1}}\operatorname{sgn}f(x_n)$, $n=0,1,2,\ldots$, $h=(b-a)\operatorname{sgn}f(x_0)$, $x_0=a$ (? ? proof)

$x_{n+1}=x_n-\frac{h}{2^{n+1}}\operatorname{sgn}f(x_n)$.

Άρα η μέθοδος της διχοτόμησης τη μόνη πληροφορία που χρειάζεται είναι τα πρόσημα της f.

π.χ. $n=0$, $x_1=x_0+\frac{b-a}{2}\operatorname{sgn}f(x_0)\operatorname{sgn}f(x_0)=\frac{a+b}{2}$ (ok)

$n=1$, $x_2=x_1+\frac{b-a}{2^2}\operatorname{sgn}f(x_0)\operatorname{sgn}f(x_1)=\cancel{\frac{a+b}{2}}\cancel{\frac{b-a}{4}}=$

$=b_2-\frac{1}{2}(b_2-a_2)=\frac{b_2+a_2}{2}$ (ok) ..

Επαναληπτική μέθοδος σταθερού σημείου

$f(x)=0$, f συνεχής στο $[a,b]$

$\Leftrightarrow x=g(x)$, g συνεχής (η γραφή αυτή δεν είναι μοναδική)

π.χ. $f(x) = x^2-2x-3 \Leftrightarrow x = x^2-x-3 = g_1(x)$

$\Leftrightarrow x(x-2)-3=0 \Leftrightarrow x = \frac{3}{x-2} = g_2(x)$

$\Leftrightarrow x = \frac{x^2-3}{2} = g_3(x)$

$\Leftrightarrow x = \pm\sqrt{2x+3} = g_4(x)$

$\Leftrightarrow \ldots\ldots$

π.χ. $f(x) = x^3-x^2+1=0 \Leftrightarrow x = x^3-x^2+x+1 = g_1(x)$

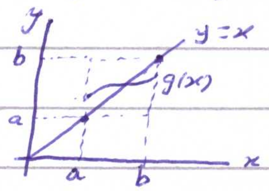
$\Leftrightarrow x = x - \frac{x^3-x^2+1}{x^2+4} = g_2(x)$

$\Leftrightarrow x = \frac{1}{x-x^2} = g_3(x)$

$\Leftrightarrow x = \sqrt{\frac{1}{1-x}} = g_4(x)$

$\Leftrightarrow \ldots\ldots$

Πάντα θα υποθέτουμε ότι η g είναι τέτοια που $g(x)\in[a,b]$, $\forall x\in[a,b]$. Έτσι θα εξασφαλίζεται η ύπαρξη ρίζας r, (σταθερό σημείο της g) $r=g(r) \Leftrightarrow f(r)=0$ και το πρόβλημα προσδιορισμού ρίζας θα ανάγεται σε πρόβλημα προσδιορισμού σταθερού σημείου}.



π.χ. $g(x) = x^2+x-1 \Rightarrow g(-1)=-1$, $g(1)=1 \Rightarrow -1, 1$ σταθερά σημεία της g

Δηλαδή, αν $g(x)\in[a,b]$, $\forall x\in[a,b] \Rightarrow \exists\, r\in[a,b]$ με $r=g(r)$

Πράγματι, αν $g(a)=a$ ή $g(b)=b$ υπάρχει σταθερό σημείο της g. Αφού $g(x)\in[a,b]$ απομένει να εξετάσουμε την περίπτωση με $g(a)>a$, $g(b)<b$. Αν $\varphi(x)=g(x)-x$, τότε $\varphi(a)=g(a)-a>0$, $\varphi(b)=g(b)-b<0$. Άρα $\exists\, r\in(a,b)$ με $\varphi(r)=0 \Leftrightarrow g(r)=r$.

Η παραπάνω συνθήκη είναι ικανή, αλλά όχι αναγκαία, π.χ. $g: [-1,1]\to[0,2]$, $g(x)=2x^2$ έχει $g(0)=0$, $g(\frac{1}{2})=\frac{1}{2}$

π.χ. Η $g(x) = 2^{-x}$, $x\in[0,1]$ έχει $g'(x)=-\ln 2\cdot 2^{-x}<0 \Rightarrow g$ γνήσια φθίνουσα $g(0)=1$, $g(1)=1/2 \Rightarrow g([0,1]) = [\frac{1}{2},1] \subset [0,1] \Rightarrow \exists$ σταθερό σημείο της g στο $[0,1]$ (και μάλιστα μοναδικό).

$g(x)\in[a,b],\ \forall x\in[a,b]$, $\quad |g(x)-g(y)| \le L\,|x-y|,\ \forall x,y\in[a,b]$ (συνθήκη Lipschitz), $0<L<1$
(g συστολή)

$\Rightarrow \exists!\ r\in[a,b],\ r=g(r)$

Πράγματι, $g(x)\in[a,b] \Rightarrow \exists\, r\in[a,b]$ με $r=g(r)$. Έστω $r_1\neq r_2 \in[a,b]$ με $r_1=g(r_1)$, $r_2=g(r_2)$. Τότε $|r_1-r_2| = |g(r_1)-g(r_2)| \le L\,|r_1-r_2| < |r_1-r_2|$ άτοπο, άρα $r_1=r_2$, δηλ. η ρίζα είναι μοναδική.

Αν $g\in C^1[a,b]$, δηλ. η g είναι παραγωγίσιμη στο $[a,b]$, ικανοποιεί τη συνθήκη Lipschitz με $L=\max\limits_{a\le x\le b}|g'(x)|$.

Πράγματι, θεώρημα μέσης τιμής $\exists\,\xi:\ \dfrac{g(x)-g(y)}{x-y}=g'(\xi) \Rightarrow$

$g(x)-g(y)=g'(\xi)(x-y) \Rightarrow |g(x)-g(y)| = |g'(\xi)|\,|x-y|$. Αν $L=\max\limits_{a\le x\le b}|g'(x)|$

τότε $|g'(\xi)|\le L$, άρα $|g(x)-g(y)| \le L\,|x-y|$

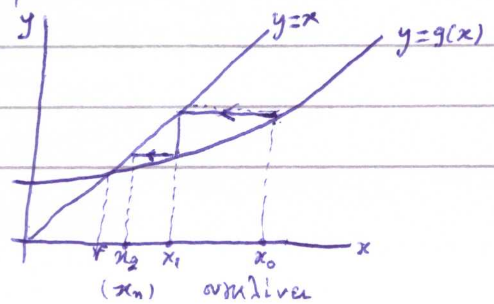
Αν $g\in C^1(a,b)$ τότε δεν ικανοποιείται κατ' ανάγκη η συνθήκη Lipschitz, π.χ. $g(x)=\sqrt{x}$ στο $(0,1)$.

π.χ. $g(x)=2^{-x}$, $x\in[0,1]$. Η g είναι συστολή διότι
$g'(x)=-\ln 2\cdot 2^{-x}$, άρα $L=\max\limits_{0\le x\le 1}|g'(x)| = \max\limits_{0\le x\le 1}|\ln 2\cdot 2^{-x}| = \ln 2 \simeq 0.693<1$

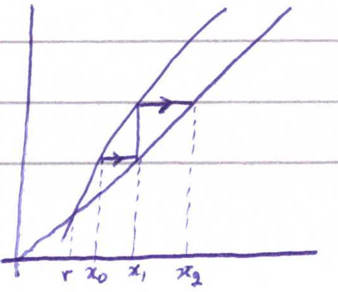
Αν ορίσουμε την επαναληπτική διαδικασία $x_{n+1}=g(x_n)$, $n=0,1,2,\ldots$ (fixed point iteration, functional iteration) με κάποιο αρχικό x_0, αν η (x_n) συγκλίνει στο $r=\lim\limits_{n\to\infty}x_n$, τότε θα είναι $r=g(r)$.

Πράγματι, αφού g συνεχής, $r=\lim\limits_{n\to\infty}x_n = \lim\limits_{n\to\infty} g(x_{n-1}) = g\left(\lim\limits_{n\to\infty}x_{n-1}\right) = g(r)$

Γεωμετρικά η επαναληπτική διαδικασία έχει ως εξής, όπου άλλοτε η (x_n) συγκλίνει, άλλοτε αποκλίνει.



(x_n) συγκλίνει



(x_n) αποκλίνει

Θεώρημα Αν $g(x) \in [a,b]$, $\forall x \in [a,b]$ και $|g(x)-g(y)| \le L|x-y|$, $0 \le L < 1$, $\forall x \in [a,b]$ τότε η (x_n) με $x_{n+1} = g(x_n)$ συγκλίνει στο $r = g(r)$ για οποιαδήποτε αρχική τιμή x_0 και μάλιστα

$$|x_n - r| \le L^n \max\{x_0 - a, b - x_0\}$$

$$|x_n - r| \le \frac{L^n}{1-L}|x_1 - x_0| \quad \text{(Εκ των προτέρων εκτίμηση – δεν χρησιμοποιείται η } x_n)$$

$$|x_n - r| \le \frac{L}{1-L}|x_n - x_{n-1}| \quad \text{(Εκ των υστέρων εκτίμηση – χρησιμοποιείται η } x_n)$$

(όταν το $L \simeq 1$ η σύγκλιση μπορεί να είναι αργή)

Πράγματι, αφού $g(x) \in [a,b]$, $|g(x)-g(y)| \le L|x-y|$, $0 \le L < 1$, άρα $\exists!\, r = g(r)$.

Επίσης η (x_n) έχει $|x_n - r| = |g(x_{n-1}) - g(r)| \le L|x_{n-1} - r|$

$\Rightarrow |x_n - r| \le L|x_{n-1} - r|$

Συνεχίζοντας $|x_{n-1} - r| = \cancel{|g(x_{n-2}) - g(r)|} \le L|x_{n-2} - r|$, άρα $|x_n - r| \le L^2 |x_{n-2} - r|$,

άρα τελικά $|x_n - r| \le L^n |x_0 - r| \le L^n \max\{x_0 - a, b - x_0\}$

Αλλά $L < 1 \Rightarrow L^n \xrightarrow[n\to\infty]{} 0 \Rightarrow |x_n - r| \xrightarrow[n\to\infty]{} 0 \Rightarrow x_n \xrightarrow[n\to\infty]{} r$

Επίσης $|x_{n+1} - x_n| = |g(x_n) - g(x_{n-1})| \le L|x_n - x_{n-1}| \le \dots \le L^n |x_1 - x_0| \Rightarrow |x_{n+1} - x_n| \le L^n |x_1 - x_0|$

Άρα, $\forall k$, $|x_{n+k} - x_n| = |(x_{n+k} - x_{n+k-1}) + (x_{n+k-1} - x_{n+k-2}) + \dots + (x_{n+1} - x_n)|$

$\le |x_{n+k} - x_{n+k-1}| + |x_{n+k-1} - x_{n+k-2}| + \dots + |x_{n+1} - x_n|$

$\le L^{n+k-1}|x_1 - x_0| + L^{n+k-2}|x_1 - x_0| + \dots + L^n |x_1 - x_0|$

$= L^n (1 + L + \dots + L^{k-1})|x_1 - x_0| = L^n \frac{1 - L^k}{1-L}|x_1 - x_0| \le \frac{L^n}{1-L}|x_1 - x_0|$

$\Rightarrow (x_n)$ Cauchy ακολουθία στο $[a,b]$ $\Rightarrow (x_n)$ συγκλίνει στο $r \in [a,b]$ με $r = g(r)$.

Αν $\varphi(x) = |x - x_n|$ τότε $|r - x_n| = \varphi(r) = \varphi(\lim\limits_{k\to\infty} x_{n+k}) = \lim\limits_{k\to\infty} \varphi(x_{n+k}) =$

$= \lim\limits_{k\to\infty} |x_{n+k} - x_n| \le \frac{L^n}{1-L}|x_1 - x_0|$

Επίσης αν $y_0 = x_{n-1}$, $y_1 = g(y_0) = g(x_{n-1}) = x_n$, αφού

$|y_n - r| \le \frac{L^n}{1-L}|y_1 - y_0| \overset{n=1}{\Rightarrow} |y_1 - r| \le \frac{L^1}{1-L}|y_1 - y_0| \Rightarrow |x_n - r| \le \frac{L}{1-L}|x_n - x_{n-1}|$

Η παραπάνω 3η εκτίμηση σφάλματος είναι καλύτερη από τη 2η διότι

$\frac{L}{1-L}|x_n - x_{n-1}| \le \frac{L}{1-L} L^{n-1} |x_1 - x_0| = \frac{L^n}{1-L}|x_1 - x_0|$

π.χ. η $g(x) = 2^{-x}$, $x \in [0,1]$ είναι συστολή, άρα η $x_{n+1} = 2^{-x_n}$ συγκλίνει $\forall x_0 \in [0,1]$ στο $r = g(r)$

Αν $|x_{N+1} - x_N| < \varepsilon \Rightarrow |x_{N+1} - r| \le \frac{L\varepsilon}{1-L}$ (έτσι εξασφαλίζουμε ένα άνω φράγμα στο σφάλμα $|x_{N+1} - r|$)

διότι η ακολουθία $\delta_n = |x_{n+1} - x_n|$ έχει $\delta_n = |x_{n+1} - x_n| \le L|x_n - x_{n-1}| = L\delta_{n-1}$, άρα η δ_n φθίνει μονότονα.
Αν N ο μικρότερος φυσικός με $|x_{N+1} - x_N| < \varepsilon$ τότε $|x_{N+1} - r| \le \frac{L}{1-L}|x_{N+1} - x_N| \le \frac{L}{1-L}\varepsilon$

Αν $g \in C^1[a,b]$ με $g'(r) \neq 0$ τότε $\lim\limits_{n\to\infty} \dfrac{x_{n+1}-r}{x_n-r} = g'(r)$ και $\varepsilon_{n+1} \simeq g'(r)\,\varepsilon_n$, $\varepsilon_n = x_n - r$

~~$\varepsilon_{n+1} \simeq g'(r)\varepsilon_n$, $\varepsilon_n = x_n - r$~~

Δηλ. $\varepsilon_{n+1} = x_{n+1} - r = g(x_n) - g(r) = g'(\xi_n)(x_n - r)$.

$\Rightarrow \lim\limits_{n\to\infty} \dfrac{x_{n+1}-r}{x_n-r} = \lim\limits_{n\to\infty} g'(\xi_n) = g'(\lim\limits_{n\to\infty}\xi_n)$

Αλλά αφού $x_n \underset{n\to\infty}{\longrightarrow} r \Rightarrow \xi_n \underset{n\to\infty}{\longrightarrow} r$, άρα $\lim\limits_{n\to\infty} \dfrac{x_{n+1}-r}{x_n-r} = g'(r) = \lim\limits_{n\to\infty} \dfrac{\varepsilon_{n+1}}{\varepsilon_n}$

(*)

π.χ. $f(x) = x^2 - 2x - 3 = 0 \Leftrightarrow x = \pm\sqrt{2x+3} = g(x)$, $x_0 = 4$, $x \in [1,4]$

$g'(x) = \dfrac{1}{\sqrt{2x+3}} \Rightarrow |g'(x)| = \dfrac{1}{\sqrt{2x+3}} < 1$, $\forall x \in [1,4]$, άρα η (x_n) συγκλίνει

$x_0 = 4$, $x_1 = g(x_0) = g(4) = 3.316$, $x_2 = 3.104, \ldots$ $x_n = 3$ μονότονη σύγκλιση

π.χ. $f(x) = x^2 - 2x - 3 = 0 \Leftrightarrow x = \dfrac{3}{x-2} = g(x)$, $x_0 = 4$, $x \in (-\infty, 1]$

$g'(x) = -\dfrac{3}{(x-2)^2} \Rightarrow |g'(x)| = \dfrac{3}{(x-2)^2} < 1$, $\forall x \in (-\infty, 1]$, άρα (x_n) συγκλίνει

$x_0 = 4$, $x_1 = g(x_0) = g(4) = 1.5$, $x_2 = -6, \ldots, x_n = -1$ παλινδρομική σύγκλιση

π.χ. $f(x) = x^2 - 2x - 3 = 0 \Leftrightarrow x = \dfrac{x^2-3}{2} = g(x)$, $x_0 = 4$, $x \in [1,4]$

$g'(x) = x \Rightarrow |g'(x)| = |x| > 1$, άρα (x_n) αποκλίνει

$x_0 = 4$, $x_1 = 6.5$, $x_2 = 19.6$, $x_3 = 191.0, \ldots$

π.χ. $f(x) = x^3 - x^2 + 1 = 0 \Leftrightarrow x = x - \dfrac{x^3 - x^2 + 1}{x^2 + 4} = g(x)$, $x_0 = 0.5$, $\epsilon = 10^{-5}$

Η $x_{n+1} = g(x_n)$ συγκλίνει με $x_{16} = -0.75488$ ενώ $r = -0.754877666624 = g(r)$

Είναι $g'(r) = 0.28 < 1$

π.χ. $f(x) = x^3 - x^2 + 1 = 0 \Leftrightarrow x = x^3 - x^2 + x + 1 = g(x)$, $x_0 = 0.5$, $\epsilon = 10^{-5}$

Η $x_{n+1} = g(x_n)$ αποκλίνει αφού $g'(x) = 3x^2 - 2x + 1 \Rightarrow g'(r) \simeq 4.22 > 1$

π.χ. $f(x) = x^3 + 2x^2 + 10x - 20 = 0 \Leftrightarrow x = \dfrac{20}{x^2 + 2x + 10} = g(x)$, $x_0 = 1$ ακρίβεια 10 δεκαδικών ψηφίων

$x_{n+1} = g(x_n)$, $x_0 = 1 \Rightarrow x_1 = g(x_0) = g(1) = \dfrac{20}{1+2+10} = \dfrac{20}{13} \simeq 1.538461538 5, \ldots$

$x_{28} = 1.3688081078$

π.χ. $f(x) = x^3 - 2x - 5 = 0 \Leftrightarrow x = \dfrac{x^3 - 5}{2} = g_1(x)$ όχι σωστή σε διάστημα $[a,b]$ που να ικανοποιούνται υποθέσεις σταθερού σημείου

g συγκλίνει για $x < \sqrt{2/3}$ $(|g_1'(x)| < 1)$

$f(x)=0 \Leftrightarrow x=\frac{x^3-(2-\lambda)x-5}{\lambda}=g_2(x),\ \lambda\neq 0$

Στο $[2,3]$, $g_2(x)\in[2,3]$ και g συστολή στο $[2,3]$ για $\lambda\leq -16$

Για $\lambda=-19$, $x_0=3$ για $|x_{N+1}-x_N|<\varepsilon=10^{-13} \Rightarrow n\geq N=31$

$x_{31}=2.0945514815425$

Λέμε ότι η τάξη σύγκλισης της (x_n) είναι (τουλάχιστον) $p\geq 1$ αν $\exists\, C>0: \frac{|x_{n+1}-r|}{|x_n-r|^p}\leq C,\ \forall n\in\mathbb{N},\quad r=\lim_{n\to\infty}x_n$

Τότε η ακολουθία $\frac{|x_{n+1}-r|}{|x_n-r|^p}$ είναι φραγμένη.

Για $p=1$, γραμμική σύγκλιση

$p=2$, τετραγωνική σύγκλιση

$p=3$, κυβική σύγκλιση

Όσο μεγαλύτερο το p τόσο ταχύτερη σύγκλιση της (x_n) στο r.

Αν η (x_n) συγκλίνει κατά τάξη p τότε συγκλίνει και κατά τάξη q, $1\leq q<p$, διότι για $|x_n-r|<1 \Rightarrow |x_n-r|^q>|x_n-r|^p \Rightarrow \frac{1}{|x_n-r|^q}<\frac{1}{|x_n-r|^p} \Rightarrow$

$\frac{|x_{n+1}-r|}{|x_n-r|^q}<\frac{|x_{n+1}-r|}{|x_n-r|^p}\leq C$

Αν $\lim_{n\to\infty}\frac{x_{n+1}-r}{(x_n-r)^p}=a\neq 0 \Rightarrow$ η τάξη σύγκλισης της (x_n) είναι p

Η τάξη σύγκλισης της (x_n) είναι τουλάχιστον p. Αν η τάξη σύγκλισης ήταν $p+\varepsilon$, $\varepsilon>0$ τότε

$\frac{|x_{n+1}-r|}{|x_n-r|^{p+\varepsilon}}\leq C \Rightarrow \frac{|x_{n+1}-r|}{|x_n-r|^p}\leq C|x_n-r|^\varepsilon$, άρα $\lim_{n\to\infty}\frac{|x_{n+1}-r|}{|x_n-r|^p}\leq C\lim_{n\to\infty}|x_n-r|^\varepsilon$

$\Rightarrow |a|\leq 0$

(*) $g\in C^1[a,b],\ g'(r)\neq 0 \Rightarrow$ τάξη σύγκλισης της (x_n) $p=1$

Για να πάρουμε ακολουθία με τάξη σύγκλισης $p>1$ πρέπει $g'(r)=0$ (π.χ. μέθοδος Newton)

Μέθοδος Newton - Raphson

Η μέθοδος αυτή συγκλίνει, υπό προϋποθέσεις, τοπικά, δηλαδή για αρχική προσέγγιση x_0 αρκετά κοντά στο r. Η ταχύτητα σύγκλισης εξαρτάται από το x_0 και μπορεί να επιταχύνεται ή ακόμα μπορεί και να αποκλίνει. Απαιτεί δύο συναρτησιακούς υπολογισμούς, της f και της f' σε κάθε επανάληψη.

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

Αν $g(x)=x-\dfrac{f(x)}{f'(x)}$ τότε $f(x)=0 \Leftrightarrow x=g(x)$

Άρα $x_{n+1}=g(x_n) \Rightarrow \underline{x_{n+1}=x_n-\dfrac{f(x_n)}{f'(x_n)}}$ ορίζει την ακολουθία (x_n)

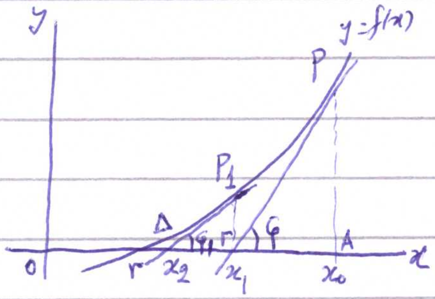
Αλλιώς, αν x_0 είναι κοντά στη ρίζα r τότε
$f(x_0+\delta)=0 \Rightarrow f(x_0)+f'(x_0)\delta+\ldots=0 \Rightarrow \delta=-\dfrac{f(x_0)}{f'(x_0)}$, άρα $x_1=x_0+\delta=$
$=x_0-\dfrac{f(x_0)}{f'(x_0)}$ είναι μία καλύτερη προσέγγιση της ρίζας. Όμοια $x_2=x_1-\dfrac{f(x_1)}{f'(x_1)}$,
$\ldots, \; x_{n+1}=x_n-\dfrac{f(x_n)}{f'(x_n)}$

Αλλιώς η εφαπτόμενη στο $(x_0,f(x_0))$ είναι $y=f(x_0)+f'(x_0)(x-x_0)$, άρα για $y=0$ προκύπτει το x_1, δηλ. $0=f(x_0)+f'(x_0)(x_1-x_0)=0 \Rightarrow x_1=x_0-\dfrac{f(x_0)}{f'(x_0)}$
κλπ..

Ακόμα $0=f(r)=f(x_n)+(r-x_n)f'(x_n)+\frac{1}{2}f''(\xi_n)(r-x_n)^2$, ξ_n μεταξύ x_n και r
$\Rightarrow r=x_n-\dfrac{f(x_n)}{f'(x_n)}-\dfrac{1}{2}(r-x_n)^2\dfrac{f''(\xi_n)}{f'(\xi_n)}$

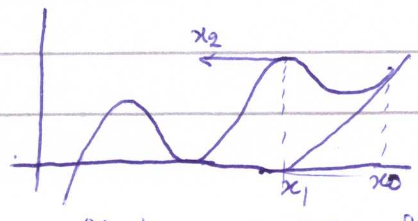
Άρα $x_{n+1}=x_n-\dfrac{f(x_n)}{f'(x_n)}$ καλύτερη προσέγγιση της r από το x_n

Γεωμετρικά

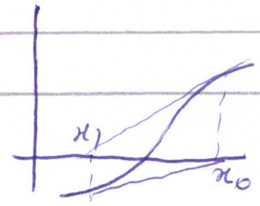


$$x_1=0\Gamma=0A-\Gamma A=x_0-\frac{PA}{\tan\varphi}=x_0-\frac{f(x_0)}{f'(x_0)}$$

$$x_2=0\Delta=0\Gamma-\Delta\Gamma=x_1-\frac{P_1\Gamma}{\tan\varphi}=x_1-\frac{f(x_1)}{f'(x_1)}$$



απομάκρυνση, αφού $f'(x_1)=0$



f'' όχι σταθερό πρόσημο στην περιοχή σύγκλισης, ανακύκλωση π.χ. $f(x)=x^3-8=0$, $x_0=1$, $\varepsilon=|x_{N+1}-x_N|=0.02$

$f'(x)=3x^2$

$x_{n+1}=x_n-\dfrac{f(x_n)}{f'(x_n)}=x_n-\dfrac{x_n^3-8}{3x_n^2}$, $n=0,1,2,\dots$

$n=0 \rightsquigarrow x_1=x_0-\dfrac{x_0^3-8}{3x_0^2}=1.5-\dfrac{-4.625}{6.75}=2.185$, $|x_1-x_0|=0.685>\varepsilon$

$n=1 \rightsquigarrow x_2=x_1-\dfrac{x_1^3-8}{3x_1^2}=2.185-\dfrac{2.434}{14.325}=2.015$, $|x_2-x_1|=0.1699>\varepsilon$

$n=2 \rightsquigarrow x_3=x_2-\dfrac{x_2^2-8}{3x_2^2}=2.015-\dfrac{0.1844}{12.1837}=2.000115$, $|x_3-x_2|=0.0151<\varepsilon=0.02$

$N_{min}=2$, ακριβής λύση $r=2$

π.χ. $f(x)=3\sin x+4x-5$, $x_0=1$, $\varepsilon=10^{-4}$

$f'(x)=3\cos x+4$

$x_{n+1}=x_n-\dfrac{3\sin x_n+4x_n-5}{3\cos x_n+4}$

$n=0$, $x_1=1-\dfrac{1.5244}{5.6209}=0.72879$, $|x_1-x_0|=0.271>\varepsilon$

$n=1$, $x_2=0.72879-\dfrac{0.0869}{6.2379}=0.742726$, $|x_2-x_1|=0.0139>\varepsilon$

$n=2$, $x_3=0.7427-\dfrac{0.000194}{6.20988}=0.742758$, $|x_3-x_2|=0.00003<\varepsilon=10^{-4}$

$\underline{g'(r)=0}$

δηλ. $g'(x)=1-\dfrac{f'}{f'}+\dfrac{ff''}{f'^2}=\dfrac{ff''}{f'^2}\Rightarrow g'(r)=\dfrac{f(r)f''(r)}{f'(r)^2}=0$

π.χ. $f(x)=x^3-2x-5=0$, ακρίβεια 15 δεκαδικών ψηφίων;

$x_0=3$: $x_1=x_0-\dfrac{f(x_0)}{f'(x_0)}=x_0-\dfrac{x_0^3-2x_0-5}{3x_0^2-2}=2.36$

$|x_{N+1}-x_N|<10^{-15}\Rightarrow N_{min}=6$, $x_6=2.094551481542327$

(ταχύτατη σύγκλιση) $x_7=$ ίδιο

αφού $x_0=3$ 'κοντά' στο r

$x_0=0$: $x_1=-\dfrac{5}{2}=-2.5$

$|x_{N+1}-x_N|<10^{-15}\Rightarrow N_{min}=19\Rightarrow x_{19}=2.094551481542327$

$x_{20}=$ ίδιο

(αργή σύγκλιση αφού $x_0=0$ πιο μακριά από το r)

π.χ. $\sqrt{A}$, $A>0$

$f(x)=x^2-A=0$, $f'(x)=2x$

$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}=x_n-\frac{x_n^2-A}{2x_n}=\frac{x_n^2+A}{2x_n}=\frac{1}{2}\left(x_n+\frac{A}{x_n}\right)$ τύπος Ήρωνα εύρεσης $\sqrt{A}$

π.χ. $\sqrt[3]{7}$ ακρίβεια 6 δεκαδικών ψηφίων

$f(x)=x^3-7=0$, $f'(x)=3x^2$, $x_0=3$

$x_1=x_0-\frac{f(x_0)}{f'(x_0)}=x_0-\frac{x_0^3-7}{3x_0^2}=2.259259$

$|x_{N+1}-x_N|<10^{-6} \Rightarrow N_{min}=5 \Rightarrow x_5=1.912931$, $f(x_5)=0.000000$

π.χ. $f(x)=x-2^{-x}=0$, $r=0.64118\ldots$

$f'(x)>0$, $\forall x\in\mathbb{R}$, r απλή ρίζα

$x_0=0.5$, $|x_{N+1}-x_N|<10^{-13} \Rightarrow N_{min}=3$, $x_3=0.64118574450496$

$x_4= \ldots 499$

πολύ γρήγορη σύγκλιση σε σχέση με τον τύπο $x_{n+1}=2^{-x_n}$

π.χ. $f(x)=x^3-2x-5=0$, $r=2.09455\ldots$

$x_0=3$, $|x_{N+1}-x_N|<10^{-13} \Rightarrow N_{min}=5$, $x_5=2.0945514815423$

x_6 = ίδιο

πολύ γρήγορη σύγκλιση σε σχέση με τη διχοτόμηση ή σε σχέση με μια απλή επαναληπτική μέθοδο.

$x_0=-2$, $|x_{N+1}-x_N|<5\times10^{-6} \Rightarrow N=7$, $x_7=2.09455$

$x_0=0$, $|x_{N+1}-x_N|<5\times10^{-6} \Rightarrow N=18$, $x_{18}=2.09455$

δηλαδή, όταν το x_0 δεν είναι κοντά στο r, αργεί η σύγκλιση, μπορεί και να αποκλίνει.

Θεώρημα $f(r)=0$, $f'(r)\neq0$ (r απλή ρίζα της f)

f δύο φορές συνεχώς παραγωγίσιμη σε μία περιοχή του r.

Τότε η (x_n) με $x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}$ συγκλίνει στο r.

και $\lim_{n\to\infty}\frac{x_{n+1}-r}{(x_n-r)^2}=\frac{f''(r)}{2f'(r)}$, δηλ. τουλάχιστον τετραγωνική σύγκλιση

Αν $f''(r)\neq0 \Rightarrow$ τάξη σύγκλισης $=2$.

Παρατήρηση, η $g(x) = x - \frac{f(x)}{f'(x)}$ έχει $g'(r)=0 \Rightarrow \exists$ κλειστό διάστημα I με μέσον r ώστε $\max_{x\in I} |g'(x)| = L < 1$

$\forall x \in I$, $|g(x)-r| = |g(x)-g(r)| = |g'(\xi)|\,|x-r| \le L|x-r| < |x-r| \Rightarrow g(x) \in I$
$\Rightarrow g$ συστολή

Άρα η (x_n), $x_{n+1} = g(x_n)$ συγκλίνει στο $r = g(r)$.

$$f(x_n) = f(r) + f'(r)(x_n - r) + \frac{1}{2} f''(\xi_{n_1})(x_n - r)^2$$
$$f'(x_n) = f'(r) + f''(\xi_{n_2})(x_n - r)$$

ξ_{n_1}, ξ_{n_2} μεταξύ x_n, r.

$$x_{n+1} = x_n - \frac{f'(r)(x_n - r) + \frac{1}{2} f''(\xi_{n_1})(x_n-r)^2}{f'(r) + f''(\xi_{n_2})(x_n - r)}$$

$$\Rightarrow x_{n+1} - r = (x_n - r) - (x_n - r)\,\frac{f'(r) + \frac{1}{2} f''(\xi_{n_1})(x_n - r)}{f'(r) + f''(\xi_{n_2})(x_n - r)}$$
$$= (x_n - r)^2\,\frac{f''(\xi_{n_2}) - \frac{1}{2} f''(\xi_{n_1})}{f'(r) + f''(\xi_{n_2})(x_n - r)}$$

$$\Rightarrow \lim_{n\to\infty} \frac{x_{n+1} - r}{(x_n - r)^2} = \frac{f''(r) - \frac{1}{2} f''(r)}{f'(r)} = \frac{f''(r)}{2f'(r)}$$

Άρα η σύγκλιση είναι τετραγωνική εφόσον το x_0 βρεθεί στο κλειστό διάστημα I που η g είναι συστολή. Όμως το I δεν είναι εκ των προτέρων γνωστό και μπορεί να είναι πολύ μικρό. Έτσι χρησιμοποιούμε τη μέθοδο αυτή αφού χρησιμοποιήσουμε πρώτα μια άλλη μέθοδο π.χ. διχοτόμηση και έτσι επιταχύνουμε μετά τη σύγκλιση με το Newton.

Αντίστοιχα μπορούμε να γράψουμε

$$g(x) = g(r) + g'(r)(x - r) + \frac{1}{2} g''(\xi)(x-r)^2, \quad \xi \text{ μεταξύ } x \text{ και } r$$
$$\Rightarrow g(x) - r = \frac{1}{2} g''(\xi)(x-r)^2$$
$$x_{n+1} - r = g(x_n) - r = \frac{1}{2} g''(\xi_n)(x_n - r)^2$$
$$\Rightarrow |x_{n+1} - r| = \frac{1}{2}|g''(\xi_n)|\,|x_n - r|^2 \le \frac{1}{2} L\, |x_n - r|^2$$
$$\Rightarrow \frac{|x_{n+1} - r|}{|x_n - r|^2} \le \frac{1}{2} L = c \Rightarrow \text{σύγκλιση τουλάχιστον τάξης 2.}$$

Αν $f(r)=f'(r)=\dots=f^{(m-1)}(r)=0$, $f^{(m)}(r)\neq 0$ (r ρίζα πολλαπλότητας m)

$m\neq 1 \Rightarrow \lim_{n\to\infty} \dfrac{x_{n+1}-r}{x_n-r} = 1-\dfrac{1}{m}$ (Γραμμική σύγκλιση)

Πράγματι, $f(x_n)=f(r)+f'(r)(x_n-r)+\dots+\dfrac{f^{(m-1)}(r)}{(m-1)!}(x_n-r)^{m-1}+\dfrac{f^{(m)}(\xi_{n_1})}{m!}(x_n-r)^m$

$= \dfrac{f^{(m)}(\xi_{n_1})}{m!}(x_n-r)^m$, ξ_{n_1} μεταξύ x_n, r

$f'(x_n)=\dfrac{f^{(m)}(\xi_{n_2})}{(m-1)!}(x_n-r)^{m-1}$, ξ_{n_2} μεταξύ x_n, r

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)} \Rightarrow x_{n+1}-r = x_n-r-\frac{f^{(m)}(\xi_{n_1})}{m!}\frac{(x_n-r)^m}{(x_n-r)^{m-1}}\frac{(m-1)!}{f^{(m)}(\xi_{n_2})}$$

$$=(x_n-r)\left(1-\frac{f^{(m)}(\xi_{n_1})}{m\,f^{(m)}(\xi_{n_2})}\right)$$

$$\Rightarrow \lim_{n\to\infty}\frac{x_{n+1}-r}{x_n-r}=1-\frac{f^{(m)}(r)}{m\,f^{(m)}(r)}=1-\frac{1}{m}$$

π.χ. $f(x)=x^2=0 \Rightarrow r=0$ $\quad f'(x)=2x$, $f'(r)=0$, $f''(r)\neq 0$, $m=2$

$$x_{n+1}=x_n-\frac{x_n^2}{2x_n}=\frac{1}{2}x_n \Rightarrow x_n\to 0, \ \forall x_0$$

και $\dfrac{x_{n+1}-r}{x_n-r}=\dfrac{x_{n+1}}{x_n}=\dfrac{1}{2}=1-\dfrac{1}{2}$

1ης τάξης σύγκλιση και όχι τετραγωνική λόγω πολ/ρίζας.

Η μέθοδος Newton-Raphson εφαρμόζεται και για την εύρεση μιγαδικών ριζών

π.χ. $z^3-1=0$ ακρίβεια 5 δεκαδικών ψηφίων

$r_1=1$, $r_{2,3}=e^{\pm\frac{2\pi i}{3}}=\cos\frac{2\pi}{3}\pm i\sin\frac{2\pi}{3}=-\frac{1}{2}\pm i\frac{\sqrt{3}}{2}$

Για $z_0=-0.81+i\,0.001 \Rightarrow z_{n+1}=z_n-\dfrac{f(z_n)}{f'(z_n)}=z_n-\dfrac{z_n^3-1}{3z_n^2}$

$$\Rightarrow z_1=z_0-\frac{z_0^3-1}{3z_0^2}=-0.03195+i\,0.00192$$

$|z_{N+1}-z_N|<10^{-5} \Rightarrow N_{min}=20$, $z_{20}=1.00000+i\,0.00000$ δηλ. προσεγγίζει την r_1

Για $z_0=-0.81+i\,0.03 \Rightarrow z_1=z_0-\dfrac{z_0^3-1}{3z_0^2}=-0.03403+i\,0.05753$

$|z_{N+1}-z_N|<10^{-5} \Rightarrow N_{min}=16$, $z_{16}=-0.50000+i\,0.86603$ δηλ. προσεγγίζει την r_2

Για $z_0=-0.81-i\,0.04 \Rightarrow z_1=-0.03565-i\,0.07660$

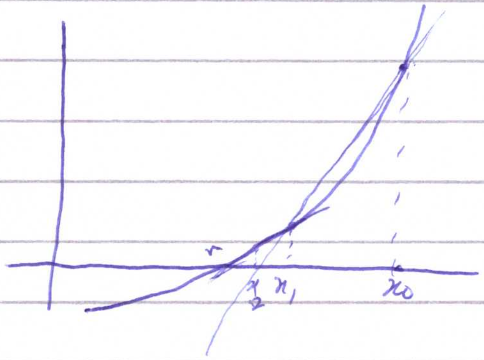
$|z_{N+1}-z_N|<10^{-5} \Rightarrow N_{min}=15$, $z_{15}=-0.50000-i\,0.86603$ δηλ. προσεγγίζει την r_3.

Μέθοδος της τέμνουσας (secant)

Στη μέθοδο Newton-Raphson μπορεί να είναι δύσκολος ο υπολογισμός της παραγώγου ή μπορεί να μην υπάρχει παράγωγος. Προσεγγίζουμε $f'(x_n) \simeq \dfrac{f(x_n)-f(x_{n-1})}{x_n - x_{n-1}}$, άρα

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

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



Είναι μέθοδος 2 βημάτων, χρειάζονται οι 2 προηγούμενες προσεγγίσεις.
Η μέθοδος πάντα συγκλίνει?, συγκλίνει πάντα βραδύτερα από τη Newton-Raphson, δεν μπορεί να εφαρμοστεί αν η ρίζα r έχει πολλαπλότητα άρτιο αριθμό. Στην πράξη εφαρμόζεται συχνότερα από τη Newton-Raphson επειδή απαιτεί τη γνώση της f και όχι της f'.

π.χ. $x^3 - 8 = 0$, $x_0 = 3$, $x_1 = 1.5$ $\quad \varepsilon = 0.02$

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{(x_n^3-8)-(x_{n-1}^3-8)}\,(x_n^3-8) = x_n - \frac{x_n^3 - 8}{x_n^2 + x_n x_{n-1} + x_{n-1}^2}, \quad n=1,2,\ldots$$

$n=1 \rightsquigarrow x_2 = 1.793651$, $|x_2 - x_1| = 0.2936509 > \varepsilon$

$n=2 \rightsquigarrow x_3 = 2.066952$, $|x_3 - x_2| = 0.2733011 > \varepsilon$

$n=3 \rightsquigarrow x_4 = 1.992769$, $|x_4 - x_3| = 0.0741832 > \varepsilon$

$n=4 \rightsquigarrow x_5 = 1.999763$, $|x_5 - x_4| = 0.00699389 < \varepsilon = 0.02$

<u>Θεώρημα</u> $f\in C^2(a,b)$, $f(r)=0$, $r\in(a,b)$, $f'(r)\neq 0$, $f''(r)\neq 0$

Τότε η (x_n) με $x_{n+1} = x_n - \dfrac{f(x_n)}{\frac{f(x_n)-f(x_{n-1})}{x_n-x_{n-1}}}$ συγκλίνει στο r με τάξη σύγκλισης $p=\frac{1+\sqrt5}{2}\simeq 1.62$.

<u>Απόδειξη</u>: Γράφουμε, όπου $p(x)\equiv f(x_n)+\dfrac{f(x_n)-f(x_{n-1})}{x_n-x_{n-1}}(x-x_n)$

$\Rightarrow p(x_{n+1})=0$, $p(x_n)=f(x_n)$, $p(x_{n-1})=f(x_{n-1})$

Αν $\varphi(t)\equiv f(t)-p(t)-\dfrac{f(x)-p(x)}{(x-x_n)(x-x_{n-1})}(t-x_n)(t-x_{n-1})$

$\Rightarrow \varphi(x_n)=\varphi(x_{n-1})=\varphi(x)=0$

Από θ. Rolle, η φ' έχει δύο τουλάχιστον μηδενικά και άρα η φ'' έχει 1 τουλάχιστον ρίζα ξ_n, άρα

$$0=\varphi''(\xi_n)=f''(\xi_n)-2\frac{f(x)-p(x)}{(x-x_n)(x-x_{n-1})}$$

$\Rightarrow f(x)-p(x)=\frac12 f''(\xi_n)(x-x_n)(x-x_{n-1})$

$\Rightarrow f(r)-p(r)=\frac12 f''(\xi_n)(r-x_n)(r-x_{n-1})$

$\Rightarrow p(r)-p(x_{n+1})=-\frac12 f''(\xi_n)(r-x_n)(r-x_{n-1})$

$\Rightarrow (r-x_{n+1})\dfrac{f(x_n)-f(x_{n-1})}{x_n-x_{n-1}}=-\frac12 f''(\xi_n)(r-x_n)(r-x_{n-1})$

Από το θ. μέσης τιμής $\dfrac{f(x_n)-f(x_{n-1})}{x_n-x_{n-1}}=f'(\zeta_n)$, ζ_n μεταξύ x_n και x_{n-1}

Αφού $f'(r)\neq 0$ επιλέγοντας x_n, x_{n-1} αρκετά κοντά στο r έχουμε $f'(\zeta_n)\neq 0$

$\Rightarrow \varepsilon_{n+1}=-\frac12\dfrac{f''(\xi_n)}{f'(\zeta_n)}\varepsilon_n\varepsilon_{n-1}$, $\varepsilon_n\equiv r-x_n$

Αν $0<\varepsilon<1$, $I=[r-\varepsilon,r+\varepsilon]$ και $M=\frac{\varepsilon}{2}\max\limits_{x,y\in I}\left|\dfrac{f''(x)}{f'(y)}\right|<1$ τότε

$\varepsilon_n\xrightarrow[n\to\infty]{}0 \Rightarrow x_n\xrightarrow[n\to\infty]{} r$

και $|\varepsilon_{n+1}|=C_n|\varepsilon_n||\varepsilon_{n-1}|$, $C_n=\frac12\left|\dfrac{f''(\xi_n)}{f'(\zeta_n)}\right|\xrightarrow[n\to\infty]{}\frac12\left|\dfrac{f''(r)}{f'(r)}\right|\equiv C$

$\overset{p>1}{\Rightarrow}\dfrac{|\varepsilon_{n+1}|}{|\varepsilon_n|^p}=C_n|\varepsilon_n|^{1-p}|\varepsilon_{n-1}|=C_n\left(\dfrac{|\varepsilon_n|}{|\varepsilon_{n-1}|^p}\right)^r \Rightarrow r=1-p,\ r=-\frac1p \Rightarrow$

$-\frac{1}{p} = 1-p \Rightarrow p^2 - p - 1 = 0 \Rightarrow p = \frac{1+\sqrt{5}}{2}$

$y_{n+1} = C_n y_n^{r}$, $\quad y_n = \frac{|\varepsilon_n|}{|\varepsilon_{n-1}|^p}$

$$y_{n+1} = C_n (C_{n-1} y_{n-1}^{r})^{r} = C_n C_{n-1}^{r} y_{n-1}^{r^2} = C_n C_{n-1}^{r} (C_{n-2} y_{n-2}^{r})^{r^2} =$$
$$= C_n C_{n-1}^{r} C_{n-2}^{r^2} y_{n-2}^{r^3} = \dots = C_n C_{n-1}^{r} C_{n-2}^{r^2} \cdots C_0^{r^n} y_0^{r^{n+1}}$$

Αν $c_n = \frac{C_n}{C} \xrightarrow[n\to\infty]{} 1$ τότε

$$y_{n+1} = C^{1+r+\dots+r^n} c_n c_{n-1}^{r} c_{n-2}^{r^2} \cdots c_0^{r^n} y_0^{r^{n+1}}$$

Όμως $|r|<1 \Rightarrow r^n \xrightarrow[n\to\infty]{} 0 \Rightarrow y_0^{r^{n+1}} \to 1$ και

$$C^{1+r+\dots+r^n} = C^{\frac{1-r^{n+1}}{1-r}} \xrightarrow[n\to\infty]{} C^{\frac{1}{1-r}} = C^{\frac{1}{p}}$$

Επίσης $|\ln(c_n c_{n-1}^{r} \cdots c_0^{r^n})| = |\ln c_n + r\ln c_{n-1} + \dots + r^n \ln c_0|$
$\le |\ln c_n| + |r|\,|\ln c_{n-1}| + \dots + |r|^n |\ln c_0|$

$\exists \varepsilon'>0, N: |\ln c_n| < \varepsilon', \forall n>N$, άρα

$$|\ln(c_n c_{n-1}^{r} \cdots c_0^{r^n})| \le \varepsilon'(1+|r|+|r|^2+\dots+|r|^{n-N}) + |r|^{n-N+1} \max_i |\ln c_i|$$
$$\Rightarrow \ln(c_n c_{n-1}^{r} \cdots c_0^{r^n}) \xrightarrow[n\to\infty]{} 0 \Rightarrow c_n c_{n-1}^{r} \cdots c_0^{r^n} \xrightarrow[n\to\infty]{} 1$$

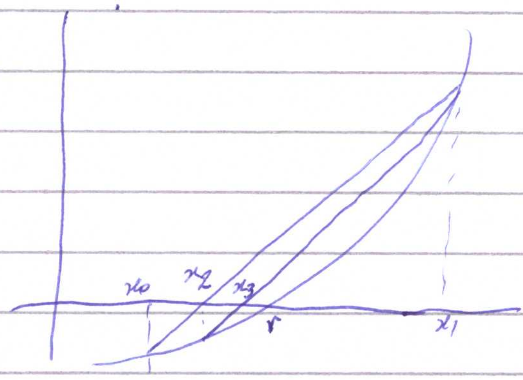
Τελικά $y_{n+1} \xrightarrow[n\to\infty]{} C^{\frac{1}{p}} \cdot 1 \cdot 1 \neq 0 \Rightarrow \lim_{n\to\infty} \frac{|\varepsilon_n|}{|\varepsilon_{n-1}|^p} = C^{\frac{1}{p}} \Rightarrow \lim_{n\to\infty} \frac{|x_n - r|}{|x_{n-1} - r|^p} = C^{\frac{1}{p}}$

$\Rightarrow$ τάξη σύγκλισης της (x_n) είναι p.

Μέθοδος εσφαλμένης θέσης

Είναι παραλλαγή της μεθόδου της τέμνουσας όπου τώρα η ρίζα είναι πάντοτε μεταξύ δύο διαδοχικών προσεγγίσεων.

Επιλέγουμε τις αρχικές προσεγγίσεις x_0, x_1 ώστε $f(x_0)f(x_1)<0$



Εν συνεχεία αν $f(x_1)f(x_2)<0$ προχωράμε (αλλιώς $f(x_0)f(x_2)<0$ και προχωράμε). Έτσι η ρίζα r εγκλωβίζεται όπως και στη μέθοδο της διχοτόμησης.

Η μέθοδος αυτή συγκλίνει γρηγορότερα από τη μέθοδο της διχοτόμησης.

π.χ. $f(x)=\cos x - x=0$, $r=0.739085133215161\ldots$

ακρίβεια 8 δεκαδικών ψηφίων

$x_0=0.5$ N-R

$x_0=0.5$, $x_1=3$ τέμνουσα, εσφαλμένη θέση

N-R: $x_3=0.73908513$

τέμνουσα: $x_5=0.73908513$

εσφαλμένη θέση: $x_8=0.73908513$

Μέθοδοι Illinois και Pegasus

Παραλλαγή της μεθόδου εσφαλμένης θέσης

$$x_{n+2} = \frac{\alpha f(x_{n-1})}{\alpha f(x_{n-1}) - f(x_{n+1})} x_{n+1} + \frac{f(x_{n+1})}{f(x_{n+1}) - \alpha f(x_{n-1})} x_{n-1} ,$$

$\alpha = \frac{1}{2}$ για Illinois

$$\alpha = \frac{f(x_i)}{f(x_i) + f(x_{i+1})} \quad \text{Pegasus}$$

Πολύ καλές όταν τα δύο σημεία που χρησιμοποιούν η f έχει αντίθετα πρόσημα. Μειονέκτημα ότι απαιτούν μεγάλη ακρίβεια πλησιάζοντας τη ρίζα.

Μέθοδος Ridders

$$x_{n+1} = x_n + (x_n - x_{n-2}) \frac{\operatorname{sgn}(x_{n-2} - x_{n-1}) f(x_n)}{\sqrt{f(x_n)^2 - f(x_{n-2}) f(x_{n-1})}}$$

Το x_{n+1} πάντα βρίσκεται στο διάστημα (x_{n-2}, x_{n-1}).

Μέθοδος Aitken (μέθοδος επιτάχυνσης Aitken)

Έστω γενική επαναληπτική μέθοδος έστω $\varepsilon_{n+1} \simeq g'(r)\,\varepsilon_n$

$$\Rightarrow \quad x_{n+1} - r \simeq g'(r)(x_n - r)$$
$$x_{n+2} - r \simeq g'(r)(x_{n+1} - r)$$

$$\Rightarrow \quad \frac{x_{n+1}-r}{x_{n+2}-r} \simeq \frac{x_n - r}{x_{n+1}-r} \Rightarrow (x_{n+1}-r)^2 \simeq (x_{n+2}-r)(x_n - r)$$

$$\Rightarrow x_{n+1}^2 - 2x_{n+1}r + \cancel{r^2} \simeq x_{n+2}x_n - (x_n + x_{n+2})r + \cancel{r^2}$$

$$\Rightarrow (x_{n+2} - 2x_{n+1} + x_n)r \simeq x_{n+2}x_n - x_{n+1}^2$$

$$\Rightarrow r \simeq \frac{x_{n+2}x_n - x_{n+1}^2}{x_{n+2} - 2x_{n+1} + x_n} = \frac{x_n^2 + x_n x_{n+2} - 2x_n x_{n+1} + 2x_n x_{n+1} - x_n^2 - x_{n+1}^2}{x_{n+2} - 2x_{n+1} + x_n}$$

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

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

$$\Rightarrow \quad y_n = x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n} = x_n - \frac{(g(x_n) - x_n)^2}{g(g(x_n)) - 2g(x_n) + x_n}$$

νέα ακολουθία να συγκλίνει γρηγορότερα στο r από ότι η (x_n)

Χρησιμοποιεί 3 διαδοχικές προσεγγίσεις της γενικής επαναληπτικής μεθόδου

Αν η (x_n) συγκλίνει γραμμικά στο r και για αρκετά μεγάλα n είναι $(x_n - r)(x_{n+1} - r) > 0$, τότε η ακολουθία (y_n) συγκλίνει στο r ταχύτερα από την (x_n) με την έννοια ότι $\lim_{n\to\infty} \frac{y_n - r}{x_n - r} = 0$.

$x_0,\ x_1 = g(x_0),\ x_2 = g(x_1) \Rightarrow y_0$

$x_1,\ x_2,\ x_3 \Rightarrow y_1$

π.χ. $x^3 + 2x^2 + 10x - 20 = 0 \Rightarrow x = \frac{20}{x^2+2x+10} = g(x) \Rightarrow x_{28} = 1.3688081078$

Με Aitken $x_{12} = 1.3687861026$, $x_{13} = 1.3688178744$, $x_{14} = 1.3688037731$

$$\Rightarrow y_{12} = x_{12} - \frac{(x_{13} - x_{12})^2}{x_{14} - 2x_{13} + x_{12}} = 1.3688081078 = x_{28}$$, άρα πιο γρήγορη σύγκλιση.

Μέθοδος Steffensen

$$x_{n+1} = x_n - \frac{(g(x_n)-x_n)^2}{g(g(x_n)) - 2g(x_n) + x_n}, \quad n=0,1,2,\dots \qquad x_0 = \text{δεδομένο}$$

Η μέθοδος αυτή όταν εφαρμόζεται σε μία γραμμικά συγκλίνουσα ακολουθία που δημιουργείται από τη γενική επαναληπτική μέθοδο, μπορεί να μετατρέψει τη σύγκλιση σε τετραγωνική, χωρίς να απαιτεί υπολογισμό της g' (όπως η N-R)

Π.χ. $x^3 - x^2 + 1 = 0 \Leftrightarrow x = x - \dfrac{x^3 - x^2 + 1}{x^2+4} = g(x)$, $x_0 = -0.5$, $\varepsilon = 10^{-10}$.

$x_0 = -0.5 \Rightarrow x_1 = g(x_0) = -0.6470588235$

$x_2 = g(x_1) = -0.7173061864$

$$x_3 = x_0 - \frac{(x_1 - x_0)^2}{x_2 - 2x_1 + x_0} = -0.7815504013$$

$x_4 = g(x_3) = -0.7624196543$

$x_5 = g(x_4) = -0.7570792583$

$$x_6 = x_3 - \frac{(x_4 - x_3)^2}{x_5 - 2x_4 + x_3} = -0.7550111580$$

$\vdots$

$x_{12} = -0.7548776662$ π.χ. $r = -0.754877666246\,69$

Ενώ η γενική επαναληπτική μέθοδος με $x = x - \dfrac{x^3 - x^2 + 1}{x^2+4} = g(x)$

δίνει $x_{19} = -0.7548776662$, δηλ. συγκλίνει πιο γρήγορα η Steffensen.

Π.χ. $x^3+2x^2+10x-20=0 \Leftrightarrow x=\dfrac{20}{x^2+2x+10}=g(x)$, $x_0=1$, $\varepsilon=10^{-15}$

$x_1=g(x_0)=1.538461538461539$

$x_2=g(x_1)=1.295019157088123$

$x_3=x_0-\dfrac{(x_1-x_0)^2}{x_2-2x_1+x_0}=1.370813882687234$

$x_4=g(x_3)=1.367918090298850$

$x_5=g(x_4)=1.369203162587276$

$x_9=\ldots=1.368808107821373$

$x_{10}=g(x_9)=1.368808107821373$, άρα η ακρίβεια επιτεύχθη στο x_9.