

Μη γραμμικές εξισώσεις Τμήμα Ναυπηγών Μηχανικών

Δημήτρης Μητσούδης

Επίλυση μη γραμμικών εξισώσεων

Πώς θα λύσουμε εξισώσεις όπως η

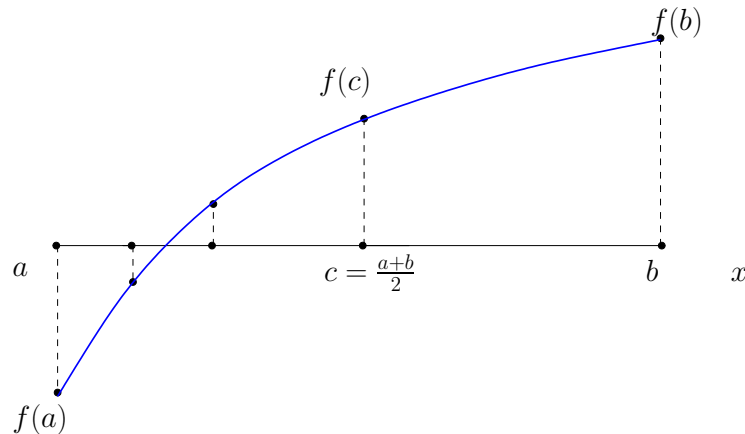
$$e^x - \sin(x) - 2 = 0;$$

Στόχος: να προσεγγίσουμε πραγματικές ρίζες ρ μιας εξίσωσης $f(x) = 0$, όπου η f μπορεί να είναι πολύπλοκη.

Μέθοδος διχοτόμησης

Θεώρημα (Bolzano)

Αν η f είναι συνεχής στο κλειστό διάστημα $[a, b]$ και $f(a)f(b) < 0$, τότε υπάρχει τουλάχιστον ένα $\rho \in [a, b]: f(\rho) = 0$.



Μέθοδος διχοτόμησης

Είσοδος: $a, b, \text{Tol}, N_{\max}, f$

υπολόγισε $f(a), f(b)$

Αν $f(a) \cdot f(b) > 0$ τότε **έξοδος**

μετρητής $\leftarrow 0$

while $((b - a)/2 > \text{Tol})$

υπολόγισε $c \leftarrow (a + b)/2, f(c)$

αν $f(c) = 0$

ρίζα = c , **έξοδος**

διαφορετικά (δηλ. αν $f(c) \neq 0$)

αν $f(a) \cdot f(c) < 0$

$b \leftarrow c, f(b) \leftarrow f(c)$

διαφορετικά

$a \leftarrow c, f(a) \leftarrow f(c)$

αν (μετρητής $\geq N_{\max}$)

τύπωσε 'Δεν επετεύχθη σύγκλιση σε $N_{\max}$ επαναλήψεις',

έξοδος

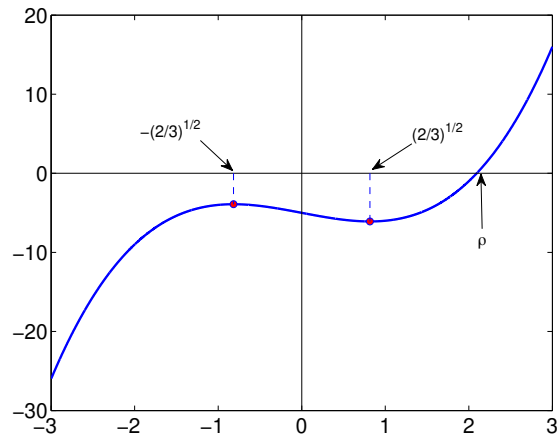
μετρητής $\leftarrow$ μετρητής + 1

ρίζα $\leftarrow (a + b)/2$, τύπωσε ρίζα

Παράδειγμα

$$f(x) = x^3 - 2x - 5 = 0$$

$$f(2) = -1 < 0, \quad f(3) = 16 > 0$$



$$1\text{η προσέγγιση: } x_1 = \frac{2+3}{2} = 2.5, \quad f(x_1) = f(2.5) = 5.625 > 0$$

$$2\text{η προσέγγιση: } x_2 = \frac{2+2.5}{2} = 2.25, \quad f(x_2) = f(2.25) = 1.890625 > 0$$

$$3\text{η προσέγγιση: } x_3 = \frac{2+2.25}{2} = 2.125, \quad f(x_3) = f(2.125) = 0.3457031 > 0$$

⋮

Επαναλ.	a	b	c	$b - a$	$f(c)$
1	2.0000000	3.0000000	2.5000000	1.0000000	5.6250000
2	2.0000000	2.5000000	2.2500000	0.5000000	1.8906250
3	2.0000000	2.2500000	2.1250000	0.2500000	0.3457031
4	2.0000000	2.1250000	2.0625000	0.1250000	-0.3513184
5	2.0625000	2.1250000	2.0937500	0.0625000	-0.0089417
6	2.0937500	2.1250000	2.1093750	0.0312500	0.1668358
7	2.0937500	2.1093750	2.1015625	0.0156250	0.0785623
⋮	⋮	⋮	⋮	⋮	⋮
16	2.0945435	2.0945740	2.0945587	0.0000305	0.0000807
17	2.0945435	2.0945587	2.0945511	0.0000153	-0.0000044
18	2.0945511	2.0945587	2.0945549	0.0000076	0.0000382
19	2.0945511	2.0945549	2.0945530	0.0000038	0.0000169
20	2.0945511	2.0945530	2.0945520	0.0000019	0.0000062

Πρόταση

Για τη μέθοδο της διχοτόμησης μπορεί εύκολα να αποδειχθεί και η ακόλουθη πρόταση:

Πρόταση

Έστω ότι η f είναι συνεχής στο $[a, b]$ και $f(a)f(b) < 0$. Αν (x_n) είναι η ακολουθία των προσεγγίσεων που παράγει η μέθοδος της διχοτόμησης (δηλ. η ακολουθία των μέσων των διαδοχικών διαστημάτων) τότε είτε $x_N = \rho$, για κάποιο N , είτε $x_n \rightarrow \rho$, $n \rightarrow \infty$, όπου ρ είναι ρίζα της εξίσωσης $f(x) = 0$.

Επιπλέον, ισχύει

$$|x_n - \rho| \leq \frac{b-a}{2^n}, \quad n = 1, 2, \dots \quad (1)$$

Απόδειξη

Θέτουμε $a_1 = a$, $b_1 = b$.

Συμβολίζουμε $[a_i, b_i]$, $i = 1, 2, \dots$, τα διαδοχικά διαστήματα που κατασκευάζει η μέθοδος της διχοτόμησης, και c_i το μέσον του $[a_i, b_i]$, $i = 1, 2, \dots$

Προφανώς, $[a_{i+1}, b_{i+1}] \subset [a_i, b_i]$.

Σε κάθε $[a_i, b_i]$ υπάρχει μια ρίζα, και αν δεν τύχει να συμβεί $c_N = \rho$ για κάποιο N , τότε θα υπάρχουν άπειρα τέτοια διαστήματα που θα περιέχουν τη ρίζα ρ .

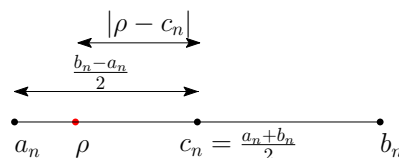
$$\begin{aligned} \text{Ισχύει: } b_n - a_n &= \frac{b_{n-1} - a_{n-1}}{2} = \frac{b_{n-2} - a_{n-2}}{2^2} = \dots = \frac{b_1 - a_1}{2^{n-1}} \\ \Rightarrow b_n - a_n &= \frac{b-a}{2^{n-1}}. \end{aligned}$$

Επίσης

$$\rho \in [a_n, b_n] \quad \text{και} \quad c_n = \frac{1}{2}(a_n + b_n).$$

Άρα

$$|\rho - c_n| \leq \frac{b_n - a_n}{2} = \frac{b-a}{2^n}.$$



Εκτίμηση του σφάλματος

Η (1):

$$|x_n - \rho| \leq \frac{b-a}{2^n}, \quad n = 1, 2, \dots,$$

αποτελεί μια εκτίμηση του σφάλματος, δηλ. μας δίνει πληροφορία για το πόσο απέχει η n -οστή προσέγγιση x_n από τη ρίζα ρ .

Από την (1) μπορούμε να υπολογίσουμε **εκ των προτέρων** το πλήθος των βημάτων που απαιτείται για να προσεγγίσουμε τη ρίζα με σφάλμα μικρότερο κάποιου επιθυμητού επιπέδου.

Παράδειγμα

Θεωρούμε και πάλι την εξίσωση

$$f(x) = x^3 - 2x - 5 = 0, \quad [a, b] = [2, 3],$$

και έστω ότι θέλουμε να υπολογίσουμε τη ρίζα με σφάλμα μικρότερο του 10^{-5} .

Αρκεί, λόγω της (1), να έχουμε

$$\begin{aligned} \frac{3-2}{2^n} < 10^{-5} &\Leftrightarrow 2^n > 10^5 \Leftrightarrow \log_{10} 2^n > \log_{10} 10^5 \\ &\Leftrightarrow n \log_{10} 2 > 5 \Leftrightarrow n > \frac{5}{\log_{10} 2} \approx 16.6 \end{aligned}$$

Άρα αρκεί να εκτελέσουμε 17 βήματα του αλγόριθμου της διχοτόμησης.

Παράδειγμα (Süli & Mayers)

Θεωρήστε τη συνάρτηση

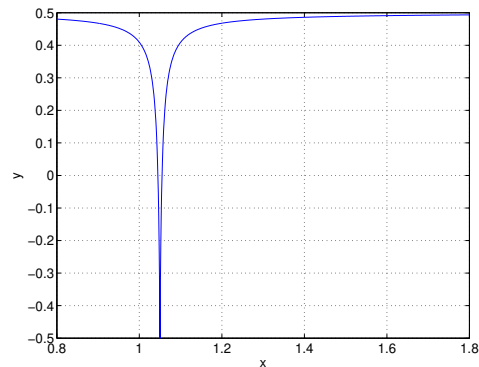
$$f(x) = \frac{1}{2} - \frac{1}{1 + M|x - 1.05|}.$$

Προφανώς η $f(x) = 0$ έχει δύο ρίζες, τις $\rho_{1,2} = 1.05 \pm \frac{1}{M}$.

Απόσταση των δύο ριζών $= 2/M$, οπότε αν το M είναι μεγάλο η απόσταση γίνεται πολύ μικρή.

Γραφική παράσταση στο διάστημα $[0.8, 1.8]$ για $M = 200$.

```
>> x = 0.8:0.001:1.8;  
>> y = 0.5 - 1./(1 +  
200*abs(x-1.05));  
>> plot(x,y)  
>> grid on  
>> xlabel('x')  
>> ylabel('y')
```



Ο εντοπισμός διαστήματος $[a, b]$ στα άκρα του οποίου οι τιμές της f να είναι ετερόσημες δεν είναι αριθμητικά εύκολος. (Πειραματιστείτε με το MATLAB για $M = 2000, 20000, 200000$.)

Πλεονεκτήματα

- Απλή μέθοδος που μπορεί να χρησιμοποιηθεί υπό πολύ γενικές συνθήκες για την f . Απαιτείται μόνο η συνέχεια της f και αλλαγή προσήμου σε μια περιοχή μιας ρίζας.
- Συγκλίνει πάντα και απαιτεί μόνον έναν υπολογισμό της f ανά βήμα.
- Μπορούμε να προσδιορίσουμε εκ των προτέρων το πλήθος των βημάτων που απαιτείται για να προσεγγίσουμε τη ρίζα με δεδομένη ακρίβεια.

Μειονεκτήματα

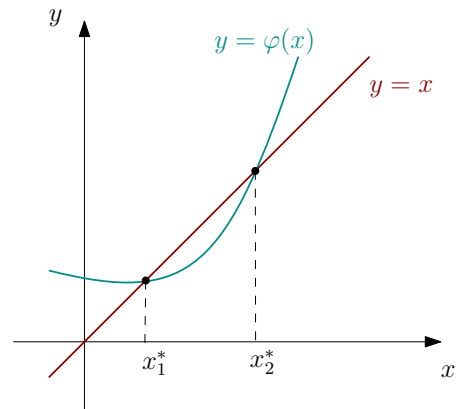
- Η μέθοδος συγκλίνει πολύ αργά σε σύγκριση με άλλες μεθόδους. Αυτό άλλωστε φαίνεται από την εκτίμηση (1), όπου υποδεικνύει ότι το σφάλμα σε κάθε βήμα απλώς υποδιπλασιάζεται.
- Έτσι το συνολικό κόστος της μεθόδου, που είναι ανάλογο του αριθμού των υπολογισμών της f είναι υψηλό.

Σταθερό σημείο συνάρτησης

Ορισμός

Ένα σημείο x^* του πεδίου ορισμού μιας συνάρτησης φ λέγεται **σταθερό σημείο** της φ , αν $\varphi(x^*) = x^*$.

Γεωμετρικά: Βρες τα σημεία τομής του γραφήματος της $y = \varphi(x)$ με την ευθεία $y = x$ (δηλ. τη διχοτόμο του α' και γ' τεταρτημορίου).



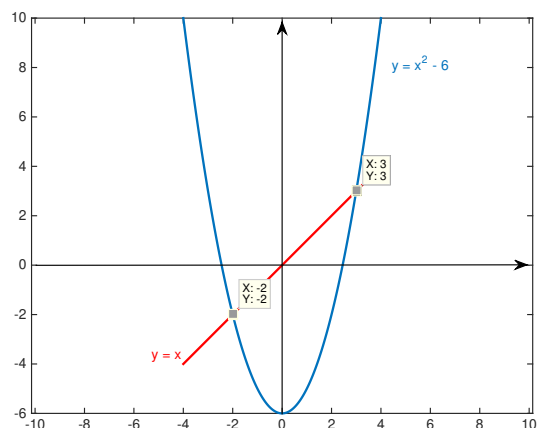
Παράδειγμα

Βρείτε τα σταθερά σημεία της συνάρτησης $\varphi(x) = x^2 - 6$.

Ένα x^* είναι σταθερό σημείο της φ αν:

$$\varphi(x^*) = x^* \Leftrightarrow (x^*)^2 - 6 = x^* \Leftrightarrow (x^*)^2 - x^* - 6 = 0 \Leftrightarrow (x^* - 3)(x^* + 2) = 0.$$

Επομένως τα σταθερά σημεία της φ είναι: $x^* = -2$ και $x^* = 3$.



Πρόταση

Αν

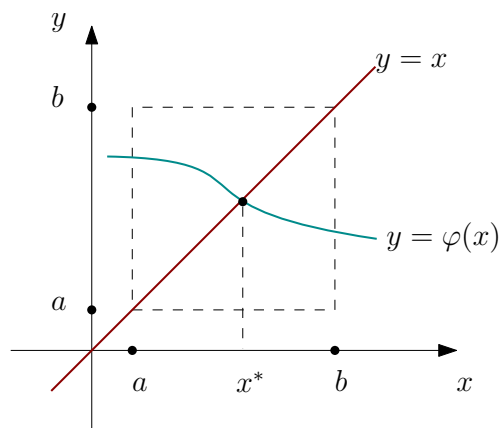
- φ είναι συνεχής στο $[a, b]$, και
- $\varphi(x) \in [a, b]$ για κάθε $x \in [a, b]$,

τότε η φ έχει τουλάχιστον ένα σταθερό σημείο στο διάστημα $[a, b]$.

Αν, επιπλέον, η φ είναι παραγωγίσιμη στο (a, b) και υπάρχει μια θετική σταθερά $L < 1$ τέτοια ώστε

$$|\varphi'(x)| \leq L \quad \forall x \in (a, b),$$

τότε η φ έχει ακριβώς ένα σταθερό σημείο στο διάστημα $[a, b]$.



Απόδειξη – (i) ύπαρξη

Αν $\varphi(a) = a$ ή $\varphi(b) = b$, τότε η φ έχει σταθερό σημείο σε ένα από τα άκρα του διαστήματος.

Αν όχι, τότε αναγκαστικά $\varphi(a) > a$ και $\varphi(b) < b$.

Θέτουμε $g(x) = \varphi(x) - x$.

Η g είναι συνεχής στο $[a, b]$, με

$$g(a) = \varphi(a) - a > 0 \quad \text{και} \quad g(b) = \varphi(b) - b < 0.$$

Άρα από Θ. Bolzano θα υπάρχει τουλάχιστον ένα $x^* \in (a, b)$: $g(x^*) = 0$.

Το x^* είναι ένα σταθερό σημείο της φ αφού

$$0 = g(x^*) = \varphi(x^*) - x^* \Rightarrow \varphi(x^*) = x^*.$$

Απόδειξη – (ii) μοναδικότητα

Ας υποθέσουμε, επιπλέον, ότι $|\varphi'(x)| \leq L < 1$ και ότι στο $[a, b]$ υπάρχουν δύο σταθερά σημεία x^*, z^* της φ .

Αν $x^* \neq z^*$, τότε από το Θεώρημα Μέσης Τιμής $\exists$ τουλάχιστον ένα ξ μεταξύ των x^* και z^* , τέτοιο ώστε

$$\frac{\varphi(x^*) - \varphi(z^*)}{x^* - z^*} = \varphi'(\xi) \Rightarrow |\varphi(x^*) - \varphi(z^*)| = |\varphi'(\xi)| |x^* - z^*|.$$

Επομένως,

$$|x^* - z^*| = |\varphi(x^*) - \varphi(z^*)| = |\varphi'(\xi)| |x^* - z^*| \leq L |x^* - z^*| < |x^* - z^*|,$$

ΑΤΟΠΟ!

Το άτοπο αυτό οφείλεται στη μόνη υπόθεση που κάναμε ότι $x^* \neq z^*$.

Άρα, αναγκαστικά $x^* = z^*$.

Παράδειγμα

Δείξτε ότι η συνάρτηση $\varphi(x) = (x^2 - 4)/5$ έχει μοναδικό σταθερό σημείο στο διάστημα $[-2, 2]$.

Η μέγιστη και η ελάχιστη τιμή της $\varphi(x)$, για $x \in [-2, 2]$, θα λαμβάνεται είτε στα άκρα του διαστήματος ή στα σημεία μηδενισμού της παραγώγου της.

Προφανώς $\varphi'(x) = 2x/5$, και η φ είναι συνεχής και παραγωγίσιμη στο $[-2, 2]$. Άρα η μέγιστη και η ελάχιστη τιμή της $\varphi(x)$, για $x \in [-2, 2]$, θα λαμβάνεται ή στο $x = -2$ ή στο $x = 0$ είτε στο $x = 2$. Όμως

$$\varphi(-2) = 0, \quad \varphi(0) = -\frac{4}{5}, \quad \varphi(2) = 0.$$

$\therefore$ η ελάχιστη τιμή της $\varphi(x)$, για $x \in [-2, 2]$ είναι $-4/5$ και η μέγιστη είναι 0.

$$\Rightarrow \forall x \in [-2, 2] : -\frac{4}{5} \leq \varphi(x) \leq 0.$$

Επιπλέον,

$$|\varphi'(x)| = \left| \frac{2x}{5} \right| \leq \frac{4}{5}, \quad \text{για κάθε } x \in (-2, 2).$$

Άρα η φ ικανοποιεί όλες τις υποθέσεις του Θεωρήματος, επομένως έχει μοναδικό σταθερό σημείο στο $[-2, 2]$.

Παρατήρηση

Για την προηγούμενη συνάρτηση μπορούμε να υπολογίσουμε το σταθερό σημείο της φ αναλυτικά:

$$x^* = \varphi(x^*) = \frac{(x^*)^2 - 4}{5} \Leftrightarrow (x^*)^2 - 5x^* - 4 = 0.$$

Άρα

$$x^* = \frac{5 \pm \sqrt{41}}{2},$$

και έτσι το σταθερό σημείο της φ που ανήκει στο $[-2, 2]$ είναι το $x^* = \frac{5 - \sqrt{41}}{2} \approx -0.702$.

Βεβαίως η φ έχει και μοναδικό σταθερό σημείο στο διάστημα $[5, 6]$ το $x^* = \frac{5 + \sqrt{41}}{2} \approx 5.702$.

Όμως $\varphi(6) = 6.4 > 6$ και $\varphi'(6) = 2.4 > 1 \Rightarrow$ η φ ΔΕΝ ικανοποιεί τις υποθέσεις του Θεωρήματος στο διάστημα $[5, 6]$.

Οι υποθέσεις του Θεωρήματος είναι ικανές, αλλά όχι όμως και αναγκαίες, για την ύπαρξη μοναδικού σταθερού σημείου.

Επαναληπτικές μέθοδοι

Πολλές φορές αντί να λύσω την εξίσωση

$$f(x) = 0,$$

είναι προτιμότερο να λύσω ένα ισοδύναμο πρόβλημα της μορφής

$$\varphi(x) = x.$$

(Δηλαδή, αντί να ψάχνω ρίζες της f , να ψάχνω για σταθερά σημεία της φ .)

Φυσικά, μπορώ να βρώ πολλές φ που να έχουν ως σταθερό σημείο μια ρίζα της f .

Π.χ.

$$f(x) = 0 \Leftrightarrow x + f(x) = x \quad \text{επομένως παίρνουμε ως} \quad \varphi(x) = x + f(x).$$

ή

$$f(x) = 0 \Leftrightarrow -3f(x) = 0 \Leftrightarrow x - 3f(x) = x \quad \text{άρα παίρνουμε ως} \quad \varphi(x) = x - 3f(x).$$

Ξεκινάμε με μια αρχική προσέγγιση x_0 και κατασκευάζουμε μια ακολουθία προσεγγίσεων $(x_n)_{n=1}^{\infty}$ από τη σχέση:

$$x_n = \varphi(x_{n-1}), \quad n = 1, 2, \dots$$

Αν αυτή η ακολουθία συγκλίνει στο x^* και η φ είναι συνεχής, τότε

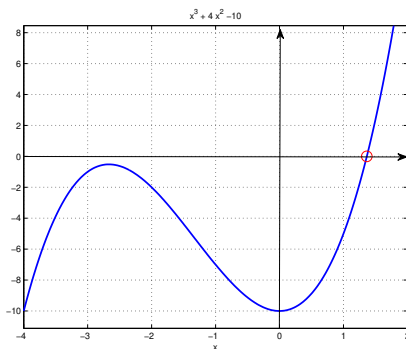
$$x^* = \lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \varphi(x_{n-1}) = \varphi\left(\lim_{n \rightarrow \infty} x_{n-1}\right) = \varphi(x^*),$$

επομένως το όριο της ακολουθίας αυτής είναι ένα σταθερό σημείο της φ .

Παράδειγμα

Να λυθεί η εξίσωση:

$$x^3 + 4x^2 - 10 = 0$$



Μοναδική ρίζα στο $[1, 2]$.

Η εύρεση ρίζας της $f(x) = x^3 + 4x^2 - 10$ μπορεί να διατυπωθεί ισοδύναμα ως εύρεση σταθερού σημείου για καθεμιά από τις ακόλουθες συναρτήσεις

$$(a) \quad x = \varphi_1(x) = x - x^3 - 4x^2 + 10$$

$$(b) \quad x = \varphi_2(x) = \sqrt{\frac{10}{x} - 4x}$$

$$(c) \quad x = \varphi_3(x) = \frac{1}{2}\sqrt{10 - x^3}$$

$$(d) \quad x = \varphi_4(x) = \sqrt{\frac{10}{4 + x}}$$

$$(e) \quad x = \varphi_5(x) = x - \frac{x^3 + 4x^2 - 10}{3x^2 + 8x}$$

Η επαναληπτική μέθοδος $x_{n+1} = \varphi_i(x_n)$, $n = 0, 1, 2, \dots$, με αρχική προσέγγιση $x_0 = 1.5$ δίνει σε καθεμιά από τις πέντε περιπτώσεις:

n	(a)	(b)	(c)	(d)	(e)
0	1.5	1.5	1.5	1.5	1.5
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	1.03×10^8		1.375170253	1.365264748	1.365230013
5			1.360094193	1.365225594	
6			1.367846968	1.365230576	
7			1.363887004	1.365229942	
8			1.365916734	1.365230022	
9			1.364878217	1.365230012	
10			1.365410062	1.365230014	
15			1.365223680	1.365230013	
20			1.365230236		
25			1.365230006		
30			1.365230013		

Ερώτηση:

Πώς μπορούμε να κατασκευάσουμε μια ακολουθία η οποία να συγκλίνει «αξιόπιστα» και όσο το δυνατόν ταχύτερα στη λύση ενός δεδομένου προβλήματος εύρεσης ρίζας μιας συνάρτησης;

Θεώρημα

Έστω φ μια συνεχής συνάρτηση στο $[a, b]$ τέτοια ώστε $\varphi(x) \in [a, b]$ για κάθε $x \in [a, b]$. Έστω, επιπλέον, ότι η φ είναι παραγωγίσιμη στο (a, b) και ότι υπάρχει μια θετική σταθερά $L < 1$ τέτοια ώστε

$$|\varphi'(x)| \leq L, \quad \text{για κάθε } x \in (a, b).$$

Τότε για κάθε $x_0 \in [a, b]$ η ακολουθία $(x_n)_{n=0}^\infty$ με

$$x_n = \varphi(x_{n-1}), \quad n = 1, 2, \dots,$$

συγκλίνει στο μοναδικό σταθερό σημείο x^* της φ στο $[a, b]$.

Επιπλέον, για τα σφάλματα $x_n - x^*$ ισχύουν οι εκτιμήσεις

$$|x_n - x^*| \leq L^n |x_0 - x^*| \leq L^n \max\{x_0 - a, b - x_0\}, \quad (2)$$

$$|x_n - x^*| \leq \frac{L^n}{1 - L} |x_1 - x_0|, \quad (3)$$

$$|x_n - x^*| \leq \frac{L}{1 - L} |x_n - x_{n-1}|. \quad (4)$$

Απόδειξη

Από το προηγούμενο Θεώρημα $\Rightarrow \exists$ μοναδικό $x^* \in [a, b] : \varphi(x^*) = x^*$.

Επειδή $\varphi([a, b]) \subset [a, b]$, η ακολουθία $(x_n)_{n=0}^\infty$ θα ορίζεται για κάθε $n \geq 1$, και τα $x_n \in [a, b]$ για όλα τα n .

Τότε $\forall n \geq 1$:

$$|x_n - x^*| = |\varphi(x_{n-1}) - \varphi(x^*)| \stackrel{\Theta\text{MT}}{=} |\varphi'(\xi_n)| |x_{n-1} - x^*| \leq L |x_{n-1} - x^*|,$$

όπου το $\xi_n \in (a, b)$.

Επαγωγικά:

$$\begin{aligned} |x_n - x^*| &\leq L |x_{n-1} - x^*| \leq L^2 |x_{n-2} - x^*| \leq \cdots \leq L^n |x_0 - x^*| \\ \Rightarrow |x_n - x^*| &\leq L^n |x_0 - x^*|. \end{aligned} \quad (5)$$

Επειδή $0 < L < 1 \Rightarrow L^n \rightarrow 0$. Παίρνουμε όρια στην (5):

$$\lim_{n \rightarrow \infty} |x_n - x^*| \leq \underbrace{\lim_{n \rightarrow \infty} L^n}_{=0} |x_0 - x^*| = 0 \Rightarrow x_n \rightarrow x^*.$$

Επειδή το $x^* \in [a, b]$ η εκτίμηση (2) προκύπτει άμεσα από την ανισότητα (5):

$$|x_n - x^*| \leq L^n |x_0 - x^*| \leq L^n \max\{x_0 - a, b - x_0\}.$$

Ακολουθώντας τα ίδια βήματα με αυτά που κάναμε στην πορεία της απόδειξης της (5) έχουμε:

$$|x_{n+1} - x_n| = |\varphi(x_n) - \varphi(x_{n-1})| \leq L |x_n - x_{n-1}| \leq \cdots \leq L^n |x_1 - x_0|.$$

Επομένως, αν $m > n \geq 1$,

$$\begin{aligned} |x_m - x_n| &= |x_m - x_{m-1} + x_{m-1} - x_{m-2} + \cdots + x_{n+1} - x_n| \\ &\leq |x_m - x_{m-1}| + |x_{m-1} - x_{m-2}| + \cdots + |x_{n+1} - x_n| \\ &\leq L^{m-1} |x_1 - x_0| + L^{m-2} |x_1 - x_0| + \cdots + L^n |x_1 - x_0| \\ &= L^n (1 + L + L^2 + \cdots + L^{m-n-1}) |x_1 - x_0|. \end{aligned}$$

Επειδή $\lim_{m \rightarrow \infty} x_m = x^*$ έχουμε

$$\begin{aligned} |x^* - x_n| &= \lim_{m \rightarrow \infty} |x_m - x_n| \leq \lim_{m \rightarrow \infty} \left(L^n (1 + L + L^2 + \cdots + L^{m-n-1}) |x_1 - x_0| \right) \\ &\leq L^n \left(\lim_{m \rightarrow \infty} \frac{1 - L^{m-n}}{1 - L} \right) |x_1 - x_0| \\ &\stackrel{L \leq 1}{\leq} L^n \frac{1}{1 - L} |x_1 - x_0| = \frac{L^n}{1 - L} |x_1 - x_0|. \end{aligned}$$

Τέλος, για να αποδείξουμε την (4) θέτουμε

$$y_0 := x_{n-1} \quad \text{και} \quad y_1 := \varphi(x_{n-1}) = x_n.$$

Τότε από την (3) έχουμε

$$|y_1 - x^*| \leq \frac{L^1}{1-L} |y_1 - y_0|,$$

η οποία αποδεικνύει την (4).

Παρατηρήσεις

- Η εκτίμηση (2) δεν έχει μεγάλη πρακτική σημασία, αφού περιέχει το σταθερό σημείο x^* το οποίο μας είναι άγνωστο.
- Η φύση των εκτιμήσεων (3) και (4) είναι αρκετά διαφορετική. Η (3) μας δίνει πληροφορία για την x_n χωρίς να χρησιμοποιεί την ίδια τη x_n , **εκ των προτέρων (a priori) εκτίμηση**.
Αντίθετα, στην (4) χρησιμοποιείται η προσέγγιση x_n για την οποία δίνει πληροφορία, **εκ των υστέρων (a posteriori) εκτίμηση**.
- Σε όλες τις εκτιμήσεις (2)–(4) η ταχύτητα με την οποία η $(x_n)_{n=0}^\infty$ συγκλίνει στο x^* εξαρτάται από τον παράγοντα L^n . Όσο μικρότερη είναι η τιμή του L τόσο μεγαλύτερη θα είναι και η ταχύτητα σύγκλισης.
- Αν στο Θεώρημα η σταθερά $L = 1$, τότε η ακολουθία $(x_n)_{n=0}^\infty$, $x_n = \varphi(x_{n-1})$, $n = 1, 2, \dots$, $x_0 \in [a, b]$ δεν συγκλίνει κατ' ανάγκην.
Π.χ., θεωρήστε την $\varphi(x) = -x$, για $x \in [-1, 1]$. (Προφανώς $|\varphi'(x)| = 1$.) Για οποιοδήποτε x_0 πάρουμε στο $[-1, 1]$ με $x_0 \neq 0$, θα προκύψει η ακολουθία $x_0, -x_0, x_0, -x_0, \dots$, η οποία προφανώς δεν συγκλίνει.

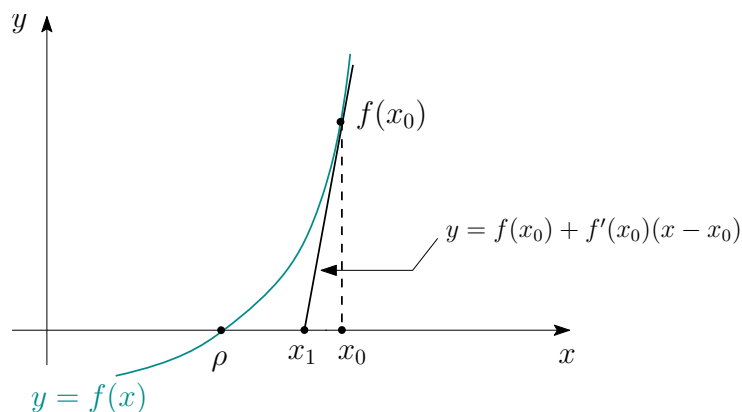
Ερώτημα: Πώς μπορούμε να κατασκευάσουμε μια ακολουθία η οποία να συγκλίνει «αξιόπιστα» και όσο το δυνατόν ταχύτερα στη λύση ενός δεδομένου προβλήματος εύρεσης ρίζας μιας συνάρτησης;

Απάντηση: Μετατρέπουμε το πρόβλημα εύρεσης ρίζας μιας συνάρτησης σε ένα ισοδύναμο πρόβλημα εύρεσης σταθερού σημείου μιας άλλης συνάρτησης που ικανοποιεί τις υποθέσεις του Θεωρήματος, και έχει παράγωγο που κατ' απόλυτη τιμή είναι όσο το δυνατόν μικρότερη κοντά στο σταθερό σημείο.

Η μέθοδος του Νεύτωνα

Είναι η πιο γνωστή επαναληπτική μέθοδος για την επίλυση μη γραμμικών εξισώσεων.

Γεωμετρική ιδέα: Ξεκινάμε από ένα σημείο x_0 , βρίσκουμε το σημείο $(x_0, f(x_0))$, και φέρνουμε την εφαπτομένη στο γράφημα της f στο σημείο $(x_0, f(x_0))$. Η επόμενη προσέγγιση x_1 θα είναι το σημείο τομής της εφαπτομένης με τον άξονα των x .



Άρα

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

Συνεχίζουμε με τον ίδιο τρόπο.

Η **μέθοδος του Νεύτωνα** είναι η επαναληπτική μέθοδος που δίνεται από την αναδρομική σχέση

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

υπό την προϋπόθεση ότι $f'(x_n) \neq 0$ για κάθε n , και ότι η x_0 είναι μια αρχική προσέγγιση της ρίζας ρ .

Θεώρημα

Έστω ότι οι f, f', f'' είναι συνεχείς και ρ είναι μια απλή ρίζα της εξίσωσης $f(x) = 0$ (δηλαδή $f(\rho) = 0$ και $f'(\rho) \neq 0$). Τότε υπάρχει ένα κλειστό διάστημα με μέσον το ρ , τέτοιο ώστε για κάθε αρχική τιμή x_0 στο διάστημα I η ακολουθία (x_n) της μεθόδου του Νεύτωνα συγκλίνει στο ρ . Επιπλέον, ισχύει ότι

$$|x_{n+1} - \rho| \leq C(x_n - \rho)^2. \quad (6)$$

Παράδειγμα

$$f(x) = x^3 + 4x^2 - 10 = 0.$$

$$x_0 = 3$$

n	x_n	$f(x_n)$
0	3.000000000	53.000000000
1	1.960784314	12.917279176
2	1.486238507	2.118579113
3	1.371823522	0.109233476
4	1.365251224	0.000350265
5	1.365230014	0.000000004
6	1.365230013	0.000000000

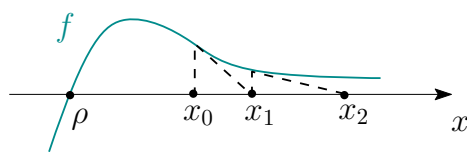
$$x_0 = -3$$

n	x_n	$f(x_n)$
0	-3.000000000	-1.000000000
1	-2.666666667	-0.518518519
2	Inf	Inf

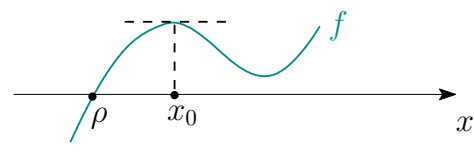
$$x_0 = -1$$

n	x_n	$f(x_n)$
0	-1.000000000	-7.000000000
1	-2.400000000	-0.784000000
2	-2.808333333	-0.601639468
$\vdots$	$\vdots$	$\vdots$
17	4.432109964	155.636984582
18	2.783197825	42.543940045
19	1.848251512	9.977823969
20	1.449682410	1.452938592
21	1.368522615	0.054459849
22	1.365235316	0.000087556
23	1.365230013	0.000000000
24	1.365230013	0.000000000

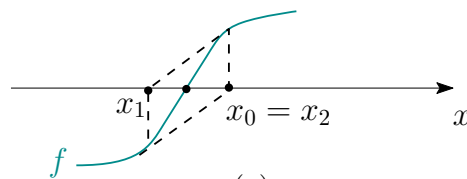
Παραδείγματα αποτυχίας της μεθόδου του Νεύτωνα λόγω κακής αρχικής επιλογής (x_0)



(a)



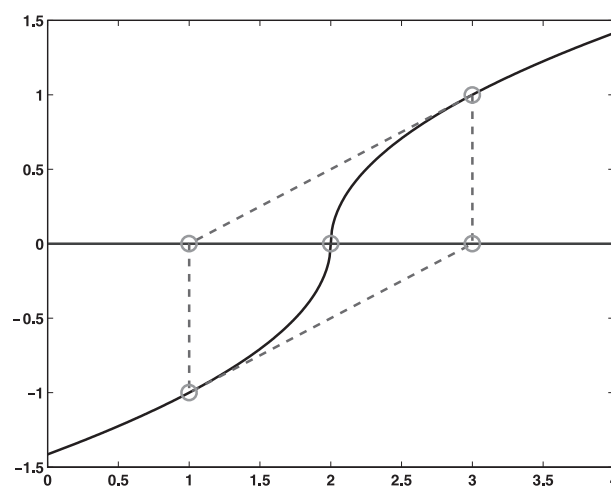
(b)



(c)

(Ακραίο) παράδειγμα αποτυχίας της μεθόδου του Νεύτωνα

$$f(x) = \begin{cases} \sqrt{x-2}, & x \geq 2 \\ -\sqrt{2-x}, & x < 2 \end{cases}$$



Πλεονεκτήματα

- Η μέθοδος του Νεύτωνα συγκλίνει ταχύτατα, υπό τις προϋποθέσεις ότι η αρχική προσέγγιση είναι αρκετά κοντά στη ρίζα ρ , ότι η ρίζα είναι απλή (δηλ. $f(\rho) = 0$, $f'(\rho) \neq 0$), και ότι η f είναι δύο φορές συνεχώς παραγωγίσιμη σε μια περιοχή της ρ .

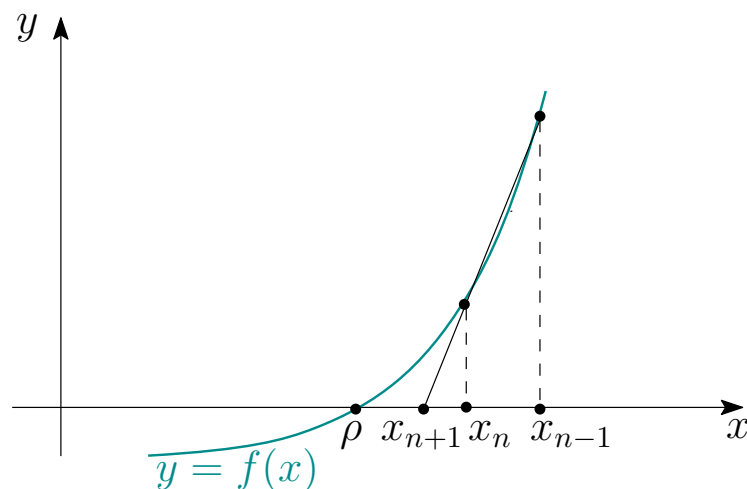
Μειονεκτήματα

- Απαιτείται η γνώση τόσο της f όσο και της f' , και υπάρχουν περιπτώσεις όπου η f' δεν είναι γνωστή ή είναι δύσκολο να υπολογιστεί.
- Σε κάθε βήμα της μεθόδου απαιτείται ο υπολογισμός των τιμών δύο συναρτήσεων: της f και της f' .

Μέθοδος της τέμνουσας

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

με αρχικές προσεγγίσεις x_0 και x_1 .



Συνδυασμοί μεθόδων (υβριδικά σχήματα)

Πολλές φορές για να βρούμε έναν αξιόπιστο και αποτελεσματικό αλγόριθμο για τον υπολογισμό ριζών της εξίσωσης $f(x) = 0$ χρησιμοποιούμε ένα συνδυασμό από αριθμητικές μεθόδους, καταφεύγουμε δηλαδή στις λεγόμενες **υβριδικές μεθόδους**.

Π.χ., μπορούμε να χρησιμοποιήσουμε αρχικά τη μέθοδο της διχοτόμησης κάνοντας μερικές επαναλήψεις για να εντοπίσουμε «χονδρικά» τη ρίζα, και στη συνέχεια να χρησιμοποιήσουμε μια πολύ ταχύτερη μέθοδο, όπως η μέθοδος του Νεύτωνα ή της τέμνουσας ώστε να προσεγγίσουμε ταχύτερα τη ρίζα.

Ο αλγόριθμος zeroin

Χρησιμοποιεί ένα συνδυασμό τριών μεθόδων: διχοτόμησης, τέμνουσας και αντίστροφης τετραγωνικής παρεμβολής. (Η τελευταία μέθοδος έχει τάξη σύγκλισης περίπου 1.84, και απαιτεί τρεις αρχικές τιμές.)

Ο αλγόριθμος zeroin υλοποιείται στο MATLAB μέσω της συνάρτησης **fzero**. Η **fzero** έχει πολλές δυνατότητες: Μπορεί να λάβει ως είσοδο μια τιμή και να αναζητήσει ένα διάστημα στα άκρα του οποίου η f έχει τιμές ετερόσημες. Οι τιμές που επιστρέφει κάθε φορά η f ελέγχονται για NaN, απειρισμούς και μιγαδικές τιμές. Ο χρήστης μπορεί να αλλάξει το προκαθορισμένο επίπεδο ακρίβειας στην προσέγγιση της ρίζας, καθώς και να πληροφορηθεί κατά την έξοδο για το πλήθος των υπολογισμών των τιμών της συνάρτησης, τον αριθμό των επαναλήψεων, κ.ο.κ.

Η συνάρτηση **fzero** (στην πιο απλή της κλήση) παίρνει δύο ορίσματα: Το πρώτο είναι η συνάρτηση f της οποίας ψάχνουμε ρίζα, και το δεύτερο είναι είτε ένα σημείο κοντά στο οποίο αναζητούμε τη ρίζα ή ένα διάστημα μέσα στο οποίο αναζητούμε τη ρίζα.

Σημειώστε ότι μια συνάρτηση περνάει ως όρισμα μιας άλλης συνάρτησης στο MATLAB με δύο τρόπους:

- (α) Ως function handle, ή
- (β) ως ανώνυμη συνάρτηση (anonymous function).

Παραδείγματα

Θέλουμε να υπολογίσουμε τη ρίζα της $f(x) = x^3 - 2x - 5 = 0$ με αρχική εκτίμηση το 3:

```
>> format long;  
>> f = @(x) x.^3 - 2*x - 5;  
>> y = fzero(f,3)  
y =  
    2.094551481542327
```

Υπολογίζουμε μια ρίζα της $\cos(x)$ στο διάστημα $[1, 2]$ γραφοντας:

```
>> format long;  
>> x = fzero(@cos, [1 2])  
x =  
    1.570796326794897
```