MAT–KOL (Banja Luka)
ISSN 0354–6969 (p), ISSN 1986–5228 (o)
Vol. XXII (1)(2016), 5–19
http://www.imvibl.org/dmbl/dmbl.htm

KVADRATNE INTERPOLACIJSKE METODE ZA JEDNODIMENZIONALNU BEZUVJETNU LOKALNU OPTIMIZACIJU [1]

Monika Zec[2], Kristian Sabo[3]

Sažetak. U radu je opisana klasa kvadratnih interpolacijskih metoda za jednodimenzionalnu lokalnu optimizaciju. U ovu se klasu metoda ubrajaju Newtonova metoda, Metoda dvije točke te Metoda tri točke. Za svaku od ovih metoda dana je geometrijska motivacija, izvod, odgovarajući algoritam te su navedeni rezultati o konvergenciji i brzini konvergencije. Spomenute metode su jednostavne te je za njihovo razumijevanje dovoljno osnovno znanje Diferencijalnog računa. U svrhu ilustracije efikasnosti metoda, dan je jedan numerički primjer izrađen u programskom paketu *Mathematica.*

Ključne riječi i fraze: Jednodimenzionalna bezuvjetna optimizacija, Interpolacija, Numerička procedura

Abstract. In this paper we describe the class of parabolic interpolation methods for one-dimensional local optimization. This class contains Newton method, Two points method and Three points method. For all of these methods, geometric motivation, derivation, corresponding algorithms and results about the convergence and convergence rate are given. Interpolation methods are very simple and for its understanding Differential Calculus knowledge is sufficient. In order to illustrate the methods, one numerical example is given that is made in *Mathematica* program package.

AMS Mathematics Subject Classification (2010): 65D05, 65K05

Key words and phrases: One-dimensional unconstrained optimization, Interpolation, Numerical procedure

---

[1] Ovaj rad je proizašao iz Diplomskog rada Monike Zec, studentice Odjela za matematiku Sveučilišta J. J. Strossmayera u Osijeku

[2] studentica Odjela za matematiku, Sveučilišta J.J. Strossmayera u Osijeku, Trg Ljudevita Gaja 6, 31000 Osijek, Republika Hrvatska, e-mail: mzec@mathos.hr

[3] Odjel za matematiku, Sveučilišta J.J. Strossmayera u Osijeku, Trg Ljudevita Gaja 6, 31000 Osijek, Republika Hrvatska, e-mail: ksabo@mathos.hr

1 Uvod

Mnogi problemi koji potječu iz različitih područja primjena svode se na traženje globalnog minimuma ili maksimuma funkcije $f : \mathbb{R}^n \to \mathbb{R}$ na skupu $\mathbb{R}^n$, odnosno točke iz $\mathbb{R}^n$ u kojoj se taj minimum, odnosno maksimum, postižu. Podsjetimo se kako za $x^\star \in \mathbb{R}^n$ kažemo da je točka globalnog minimuma funkcije f na $\mathbb{R}^n$ ako vrijedi $f(x^\star) \le f(x)$, za svaki $x \in \mathbb{R}^n$. U tom slučaju pišemo

$$x^\star = \operatorname*{argmin}_{x\in\mathbb{R}^n} f(x). \tag{1}$$

Vrijednost funkcije f u točki $x^\star$ zovemo globalni minimum funkcije te pišemo

$$f(x^\star) = \min_{x\in\mathbb{R}^n} f(x). \tag{2}$$

Slično, za $x^\star \in \mathbb{R}^n$ kažemo da je točka globalnog maksimuma funkcije f na $\mathbb{R}^n$ ako vrijedi $f(x^\star) \ge f(x)$, za svaki $x \in \mathbb{R}^n$ i pišemo

$$x^\star = \operatorname*{argmax}_{x\in\mathbb{R}^n} f(x). \tag{3}$$

Pri tome, vrijednost funkcije f u točki $x^\star$ zovemo globalni maksimum funkcije f i pišemo

$$f(x^\star) = \max_{x\in\mathbb{R}^n} f(x). \tag{4}$$

Problemi (1)-(2), odnosno (3)-(4), u literaturi se obično nazivaju i problemima bezuvjetne globalne optimizacije.

Kako je

$$\max_{x\in\mathbb{R}^n} f(x) = -\min_{x\in\mathbb{R}^n}(-f(x)),$$

bez smanjenja općenitost, umjesto oba problema (1)-(2), odnosno (3)-(4), dovoljno je analizirati samo problem traženja globalnog minimuma funkcije, odnosno točke u kojoj se taj minimum postiže. Nažalost, određivanje točke globalnog minimuma funkcije f, općenito je vrlo ozbiljan i zahtjevan posao i ponekad zahtijeva poznavanje svih točaka lokalnih minimuma funkcije f na $\mathbb{R}^n$. Sjetimo se kako za $x^\star \in \mathbb{R}^n$ kažemo da je točka lokalnog minimuma funkcije f na $\mathbb{R}^n$, ako postoji otvorena kugla $K(x^\star, \delta)$ sa središtem u točki $x^\star$ polumjera $\delta > 0$ takva da vrijedi $f(x^\star) \le f(x)$, za svaki $x \in K(x^\star, \delta)$. Pri tome vrijednost $f(x^\star)$ zovemo lokalni minimum funkcije f. Metode za traženje točke lokalnog minimuma nazivaju se lokalne optimizacijske metode.

Ako funkciju f, umjesto na $\mathbb{R}^n$, promatramo na nekoj omeđenoj i zatvorenoj domeni $\mathcal{D} \subset \mathbb{R}^n$, u cilju traženja točke globalnog minimuma mogu se koristiti neke poznate globalne optimizacijske metode. Ovdje ćemo spomenuti dvije takve metode. Prva metoda poznata pod nazivom `DIRECT` metoda za globalnu optimizaciju ([2]), koristi se ako je funkcija f Lipshitz-neprekidna na $\mathcal{D}$, te ako $\mathcal{D}$ ima specijalni oblik tzv. hiperkvadra. U općem slučaju, mogu se koristiti globalni genetički algoritmi te odgovarajuće modifikacije (vidjeti primjerice [4]).

U ovom radu govorit ćemo o jednoj klasi lokalnih jednodimenzionalnih optimizacijskih metoda. Kao što je već ranije navedeno, takve metode traže točku lokalnog minimuma funkcije f na $\mathbb{R}$ i općenito ne postoji jamstvo da je dobiveno rješenje ujedno točka globalnog minimuma. Treba spomenuti da neke specijalne funkcije, kao što su konveksne funkcije, imaju svojstvo da točka lokalnog minimuma, koja je određena nekom lokalnom metodom, ujedno predstavlja točku globalnog minimuma te funkcije (vidjeti [1]).

Ako je funkcija derivabilna na $\mathbb{R}$, sukladno poznatom Fermatovom teoremu, problem traženja lokalnih minimuma svodi se na rješavanje jednadžbe $f'(x) = 0$. Ponekad je ovu jednadžbu moguće riješiti egzaktno, te se na taj način, uz primjenu dovoljnih uvjeta ekstrema, mogu identificirati točke lokalnih minimuma. U suprotnom, nužno je koristiti neku numeričku iterativnu metodu, koja se obično zasniva na konstrukciji niza realnih brojeva (x_k), koji uz određene uvjete konvergira prema točki lokalnog minimuma $x^\star$ funkcije f. Pri tome x_k zovemo $k-$ta aproksimacija, a $e_k = x^\star - x_k$ pogreška $k-$te aproksimacije. Broj iteracija koje je potrebno provesti kako bi se zadovoljila neka unaprijed zadana točnost, ovisi o metodi. S ciljem postavljanja nekog kriterija na osnovi kojeg je moguće uspoređivanje efikasnosti metoda, navodimo sljedeću definiciju.

Definicija 1.1 *Neka niz (x_k), dobiven nekom iterativnom metodom, konvergira prema $x^\star \in \mathbb{R}$ i neka je $e_k = x^\star - x_k$ pogreška k-te aproksimacije. Kažemo da metoda ima brzinu konvergencije r, ako postoje dvije pozitivne konstante $A, r \in \mathbb{R}_+$, takve da vrijedi*

$$\lim_{k\to\infty} \frac{|e_{k+1}|}{|e_k|^r} = A. \tag{5}$$

Uočimo, ako je x_k blizu točke $x^\star$, onda će metodi s većom brzinom konvergencije trebati manji broj iteracija kako bi se postigla zadana točnost.

U ovom radu govorimo o klasi numeričkih kvadratnih interpolacijskih metoda koje se mogu koristiti za traženje točke lokalnog minimuma funkcije. Najčešće korištena metoda ove klase je poznata Newtonova metoda. Newtonovu metodu je smisleno koristiti ako funkcija ima neprekidnu drugu derivaciju. Ako funkcija nije dva puta derivabilna, Newtonova metoda postaje beskorisna. U tom slučaju mogu se koristiti neke druge metode ove klase, kojima nije potrebna druga derivacija funkcije, nego se zasnivaju na poznavanju prve derivacije i/ili vrijednosti funkcije. Očekivano, smanjivanjem uvjeta na funkciju, smanjuje se brzina konvergencije te posljedično metoda sporije konvergira.

U udžbenicima iz numeričke analize, obično su iskazani i dokazani teoremi koji se odnose na brzinu konvergencije Newtonove metode. Cilj ovog rada je ilustrirati metodologiju kojom se dokazuje brzina konvergencije i za neke druge metode ove klase, koje nisu tako česte u literaturi.

Članak je organiziran na sljedeći način. U drugom odjeljku opisane su kvadratne interpolacijske metode, u koju se ubrajaju Newtonova metoda, Metoda dvije točke te Metoda tri točke. Za svaku od metoda napravljen je kratki izvod, dani su odgovarajući algoritmi i rezultati vezani uz konvergenciju. Za

Metodu dvije točke dan je i potpuni dokaz o konvergenciji te brzini konvergencije. U trećem odjeljku izrađen je jedan jednostavan numerički primjer s kojim su ilustrirani prethodno opisani teorijski rezultati.

2 Interpolacijske metode

Pretpostavimo da je zadana funkcija $f : \mathbb{R} \to \mathbb{R}$ koja u točki $x^\star \in \mathbb{R}$ postiže lokalni minimum. Ako je funkcija f neprekidno derivabilna, onda je $f'(x^\star) = 0$. Kao što je u Uvodu već spomenuto, ako jednadžbu $f'(x) = 0$ ne možemo riješiti egzaktno, u cilju određivanja točke $x^\star$ potrebno je koristiti neku numeričku proceduru. Problem se komplicira ako funkcija nije derivabilna te u tom slučaju za traženje točke lokalnog minimuma treba koristiti neku metodu koja ne zahtijeva poznavanje derivacije. U nastavku opisat ćemo jednu klasu metoda za numeričko određivanje točke lokalnog minimuma, poznatu kao kvadratne interpolacijske metode. Zajedničko svojstvo kvadratnih interpolacijskih metoda je da funkciju f interpoliraju kvadratnom funkcijom u jednoj, dvije ili tri točke te umjesto traženja točke lokalnog minimuma funkcije f, traže globalni minimum kvadratne funkcije te se postupak uz određena pravila iterativno nastavlja. U slučaju interpolacije u jednoj točki dobivamo najčešće korištenu Newtonovu metodu ili metodu jedne točke, dok u slučajevima interpolacije u dvije odosno tri točke, dobivamo Metodu dvije točke, odnosno Metodu tri točke.

2.1 Newtonova metoda–metoda jedne točke

Ovdje ćemo pretpostaviti da je funkcija f dva puta neprekidno derivabilna. Neka je zadana točka $x_0 \in \mathbb{R}$, za koju pretpostavljamo da leži dovoljno blizu točke $x^\star$ koju želimo odrediti. Funkciju f u okolini točke x_0 aproksimirat ćemo kvadratnom funkcijom $g : \mathbb{R} \to \mathbb{R}$, $g(x) = \alpha x^2 + \beta x + \gamma$, pri čemu koeficijente α, β, γ određujemo iz sljedećih uvjeta (Slika 1):

$$
\begin{aligned}
g(x_0) &= \alpha x_0^2 + \beta x_0 + \gamma = f(x_0) =: f_0, \\
g'(x_0) &= 2\alpha x_0 + \beta = f'(x_0) =: f_0', \\
g''(x_0) &= 2\alpha = f''(x_0) =: f_0''.
\end{aligned}
\tag{6}
$$

Rješavanjem sustava (6) dobivamo da su:

$$\alpha = \frac{f_0''}{2}, \quad \beta = f_0' - f_0'' x_0, \quad \gamma = f_0 - \alpha x_0^2 - \beta x_0.$$

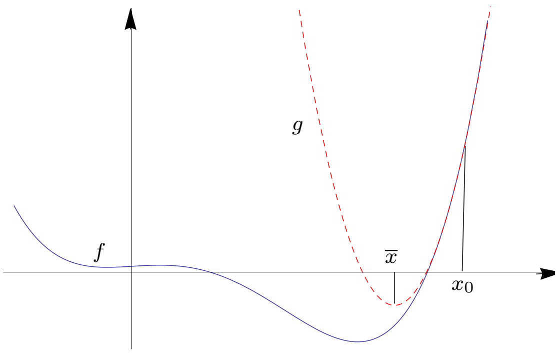
Umjesto traženja točke lokalnog minimuma funkcije f, potražit ćemo točku u kojoj se postiže globalni minimum njezine aproksimacije g. Kako je g kvadratna funkcija, slijedi da se njezin globalni minimum postiže u točki

$$\overline{x} = -\frac{\beta}{2\alpha} = x_0 - \frac{f_0'}{f_0''},$$

odakle dobivamo rekurzivnu formulu, na kojoj se temelji poznata Newtonova metoda (vidjeti također Algoritam 1):

$$x_{k+1} = x_k - \frac{f'_k}{f''_k}, \quad k = 0, 1, \ldots, \tag{7}$$

pri čemu su $f'_k := f'(x_k)$ te $f''_k := f''(x_k)$, $k = 0, 1, \ldots$.



Slika 1: Graf funkcije f i kvadratne interpolacijske funkcije g

U udžbenicima iz numeričke analize mogu se pronaći teoremi o konvergenciji Newtonove metode te njezinoj brzini konvergencije. S obzirom to da je riječ o standardnim rezultatima, sljedeći teorem navodimo bez dokaza. Iskaz i dokaz ovog teorema s nešto blažim uvjetima na funkciju f, može se naći u [6].

Teorem 2.1 *Neka je $f : \mathbb{R} \to \mathbb{R}$ tri puta neprekidno diferencijabilna. Ako je $x^\star \in \mathbb{R}$ točka takva da je $f'(x^\star) = 0$ te $f''(x^\star) \neq 0$, onda postoji $\delta > 0$ takav da za svaki $x_0 \in \langle x^\star - \delta, x^\star + \delta \rangle$ niz (x_k) zadan rekurzivno sa (7) konvergira prema $x^\star$ pri čemu je brzina konvergencije metode $r = 2$.*

Algorithm 1 Metoda jedne točke

Input: f, x_0, ε

$f'_0 = f'(x_0)$, $f''_0 = f''(x_0)$

$\bar{x} = x_0 - \frac{f'_0}{f''_0}$

while $|f'(\bar{x})| < \varepsilon$ (ili dok nije zadovoljen neki drugi kriterij zaustavljanja) **do**

  $x_0 = \bar{x}$

  $f'_0 = f'(x_0)$, $f''_0 = f''(x_0)$

  $\bar{x} = x_0 - \frac{f'_0}{f''_0}$

end while

Output: $\bar{x}$

Primijetimo da je u svrhu primjene rekurzivne formule (7) dovoljno da funkcija ima drugu derivaciju. Međutim, kako bi se dokazao prethodno navedeni teorem o konvergenciji te brzini konvergencije, postavljen je zahtjev da je funkcija tri puta neprekidno derivabilna. Situacija u kojoj su za primjenu metode potrebni blaži uvjeti, dok su za bilo kakav teorijski rezultat potrebni puno jači uvjeti, vrlo je česta u numeričkoj analizi.

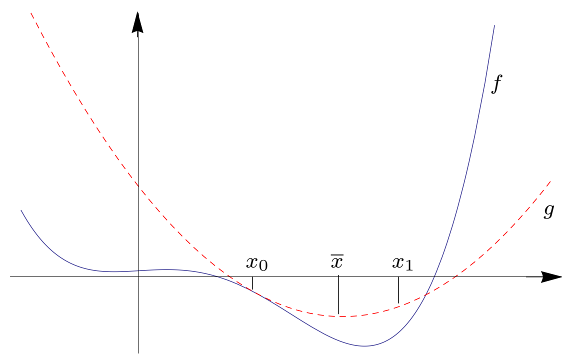
Sukladno Teoremu 2.1 brzina Newtonove metode je $r = 2$. Za takvu metodu kažemo da ima kvadratnu brzinu konvergencije. U sljedećim pododjeljcima analizirat ćemo metode koje se ne temelje na računanju druge derivacije, već koriste samo prvu derivaciju ili samo vrijednost funkcije. Smanjenje zahjeva na funkciju, očekivano će rezultirati smanjenjem brzine konvergencije odgovarajuće metode.

2.2 Metoda dvije točke

2.2.1 Metoda I

Neka su $x_0, x_1 \in \mathbb{R}$, $x_0 < x_1$ dvije različite točke za koje su poznate vrijednosti $f(x_0)$, $f'(x_0)$ i $f'(x_1)$, pri čemu dodatno pretpostavljamo da je $f'(x_0) \neq f'(x_1)$. U tom slučaju možemo potražiti kvadratnu interpolacijsku funkciju $g : \mathbb{R} \to \mathbb{R}$, oblika $g(x) = \alpha x^2 + \beta x + \gamma$, $\alpha, \beta, \gamma \in \mathbb{R}$, sa sljedećim svojstvima (vidi Sliku 2):

$$\begin{aligned} g(x_0) &= \alpha x_0^2 + \beta x_0 + \gamma = f(x_0) =: f_0, \\ g'(x_0) &= 2\alpha x_0 + \beta = f'(x_0) =: f_0', \\ g'(x_1) &= 2\alpha x_1 + \beta = f'(x_1) =: f_1'. \end{aligned} \tag{8}$$



Slika 2: Graf funkcije f i kvadratne interpolacijske funkcije g

Rješavanjem sustava linearnih jednadžbi (8) dobivamo

$$\alpha = \frac{f_1' - f_0'}{2(x_1 - x_0)}, \quad \beta = \frac{x_1 f_0' - x_0 f_1'}{x_1 - x_0}, \quad \gamma = f_0 - \alpha x_0^2 - \beta x_0.$$

Kvadratna funkcija g postiže globalni minimum u točki čija apscisa glasi

$$\overline{x} = -\frac{\beta}{2\alpha} = x_0 - \frac{x_0 - x_1}{f_0' - f_1'} f_0'.$$

Općenito, dobivamo rekurzivnu formulu

$$x_{k+1} = x_{k-1} - \frac{x_{k-1} - x_k}{f'_{k-1} - f'_k} f'_{k-1}, \quad k = 1, 2, \ldots, \tag{9}$$

pri čemu su $f'_k := f'(x_k)$, $k = 0, 1, \ldots$.

Prije no što navedemo potpuni algoritam, u sljedećem Teoremu odgovaramo na pitanje konvergencije iterativne procedure kao i odgovarajuće brzine konvergencije. Pri tome najviše koristimo [8] i [9]. Kao što ćemo vidjeti u cilju dokaza teorema potrebno je postaviti dodatne uvjete na funkciju f.

Teorem 2.2 *Neka je $f : \mathbb{R} \to \mathbb{R}$ tri puta neprekidno diferencijabilna. Ako je $x^\star \in \mathbb{R}$ takav da je $f'(x^\star) = 0$ te $f''(x^\star) \neq 0$ i $f'''(x^\star) \neq 0$, onda postoji $\delta > 0$ takav da za svaki $x_0, x_1 \in \langle x^\star - \delta, x^\star + \delta\rangle$, $x_0 \neq x_1$, niz (x_k) zadan rekurzivno s (9) konvergira prema $x^\star$ pri čemu je brzina konvergencije metode $r = \frac{1+\sqrt{5}}{2}$.*

Dokaz. Najprije uočimo da vrijedi $L(x_{k+1}) = 0$, gdje je

$$L(x) = \frac{x - x_k}{x_{k-1} - x_k} f'_{k-1} + \frac{x - x_{k-1}}{x_k - x_{k-1}} f'_k,$$

interpolacijski polinom prvog stupnja čiji graf prolazi točkama (x_{k-1}, f'_{k-1}) te (x_k, f'_k).

Bez smanjenja općenitosti možemo pretpostaviti da je $x_{k-1} < x_k$, jer ako to nije ispunjeno, možemo ih međusobno zamijeniti. Prema ocjeni pogreške za interpolacijski polinom prvog stupnja funkcije f', postoji $\xi \in \langle x_{k-1}, x_k\rangle$ (vidjeti primjerice [6]) takav da je

$$f'(x) - L(x) = \frac{1}{2} f^{'''}(\xi)(x - x_{k-1})(x - x_k).$$

Specijalno, za $x = x_{k+1}$, zbog $L(x_{k+1}) = 0$, dobivamo

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

Uvrštavanjem izraza (9) u prethodnu formulu slijedi

$$f'(x_{k+1}) = f'_{k+1} = \frac{1}{2} f^{'''}(\xi) f'_{k-1} f'_k \frac{(x_k - x_{k-1})^2}{(f'_k - f'_{k-1})^2}. \tag{10}$$

Prema Lagrangeovom teoremu srednje vrijednosti (vidi primjerice [3]) primijenjenom na funkciju f' na $[x_{k-1}, x_k]$ postoji $\kappa_0 \in \langle x_{k-1}, x_k\rangle$ takav da je

$$f'_k - f'_{k-1} = f^{''}(\kappa_0)(x_k - x_{k-1}).$$

Analogno, za funkciju f' na $[x_i, x^\star]$, $i = k-1, k, k+1$ (pri čemu smo bez smanjenja općenitosti pretpostavili da je $x_i < x^\star$), postoje brojevi $\xi_i \in \langle x_i, x^\star\rangle$, $i = k-1, k, k+1$ takvi da je

$$f'_i = f'_i - f'(x^\star) = (x_i - x^\star) f''(\xi_i), \quad i = k-1, k, k+1. \tag{11}$$

Ako je $f''(\xi_{k+1}) = 0$, onda je $f'(x_{k+1}) = 0$ pa niz (x_k), konvergira prema nul-točki funkcije f'. U suprotnom, zbog (10) i (11) imamo

$$\begin{aligned} x_{k+1} - x^\star &= \frac{f'_{k+1}}{f''(\xi_{k+1})} \\ &= \frac{1}{2} f'''(\xi) f'_{k-1} f'_k \frac{(x_k - x_{k-1})^2}{(f'_k - f'_{k-1})^2 f''(\xi_{k+1})} \\ &= \frac{1}{2} \frac{f'''(\xi) f''(\xi_k) f''(\xi_{k-1})}{f''(\xi_{k+1})(f''(\tau_0))^2} (x_k - x^\star)(x_{k-1} - x^\star). \end{aligned} \tag{12}$$

Neka su x_0 i x_1 početne aproksimacije. Označimo s $e_i = x_i - x^\star$, $i \in \mathbb{N}$. Iz (12) slijedi

$$e_2 = \frac{1}{2} \frac{f'''(\xi) f''(\xi_1) f''(\tau_0)}{f''(\xi_2)(f''(\tau_0))^2} e_1 e_0 =: \eta(e_0, e_1) e_0 e_1,$$

gdje je

$$\eta(e_{k-1}, e_k) = \frac{1}{2} \frac{f'''(\xi) f''(\xi_k) f''(\xi_{k-1})}{f''(\xi_{k+1})(f''(\tau_0))^2}. \tag{13}$$

Uočimo da je $\eta(0,0) = \frac{f'''(x^\star)}{2f''(x^\star)}$. Ako je $C \in \langle 0, 1\rangle$ unaprijed zadani broj, onda uvijek možemo odrediti dovoljno maleni broj $\delta > 0$ sa svojstvom da za sve $u, v \in \mathbb{R}$, $|u|, |v| < \delta$ vrijedi

$$|v\, \eta(u, v)| \le C < 1.$$

Odaberimo x_0, x_1 tako da je $|e_0|, |e_1| \le \delta$. Onda je

$$|e_2| = |\eta(e_0, e_1) e_0 e_1| \le C|e_1| < |e_1| \le \delta$$

odakle slijedi

$$|e_1 \eta(e_1, e_2)| \le C < 1$$

Induktivno dobivamo

$$|e_k| = |\eta(e_{k-2}, e_{k-1}) e_{k-2} e_{k-1}| \le C|e_{k-1}| \le C^2 |e_{k-2}| \le \cdots \le C^{k-1} |e_1|.$$

Zbog $C < 1$ za $k \to \infty$ slijedi da je $\lim\limits_{k\to\infty} e_k = 0$. Konačno, to znači da smo pokazali kako postoji $\delta > 0$ takav da za sve $x_0, x_1 \in \langle x^\star - \delta, x^\star + \delta\rangle$ niz (x_k) konvergira prema $x^\star$.

Preostaje pokazati da je brzina konvergencije metode $r = \frac{1+\sqrt{5}}{2}$, odnosno da postoji $A > 0$ takav da je

$$\lim_{k\to\infty} \frac{|e_{k+1}|}{|e_k|^r} = A.$$

Uočimo da iz (12) slijedi

$$e_{k+1} = \eta(e_{k-1}, e_k) e_{k-1} e_k \tag{14}$$

odakle prelaskom na limes zbog (13) dobivamo

$$\lim_{k\to\infty} \frac{e_{k+1}}{e_{k-1}e_k} = \frac{f'''(x^\star)}{2f''(x^\star)}. \tag{15}$$

Ako označimo s $s_k = \frac{|e_{k+1}|}{|e_k|^r}$, $k \in \mathbb{N}$, onda je

$$|e_k| = s_{k-1}|e_{k-1}|^r$$

te

$$|e_{k+1}| = s_k|e_k|^r = s_k s_{k-1}^r |e_{k-1}|^{r^2} \tag{16}$$

Iz (14) te (16) dobivamo

$$|\eta(e_{k-1}, e_k)| = \frac{|e_{k+1}|}{|e_{k-1}||e_k|} = \frac{s_k s_{k-1}^r |e_{k-1}|^{r^2}}{s_{k-1}|e_{k-1}|^{r+1}} = s_k s_{k-1}^{r-1} |e_{k-1}|^{r^2-r-1}.$$

Kako je $r^2 - r - 1 = 0$, slijedi $|e_{k-1}|^{r^2-r-1} = 1$ te je prema tome

$$|\eta(e_{k-1}, e_k)| = s_k s_{k-1}^{r-1}.$$

Neka je $y_k = \log|\eta(e_{k-1}, e_k)|$ te $\varepsilon_k = \log s_k$, onda je

$$\varepsilon_k = y_k - (r-1)\varepsilon_{k-1}. \tag{17}$$

Budući da niz $(\eta(e_{k-1}, e_k))$ konvergira prema broju koji je različit od 0, sukladno uvjetima teorema, slijedi da je niz (y_k) konvergentan te postoji $y^\star = \lim_{k\to\infty} y_k$. Pokažmo da je (ε_k) konvergentan te da je $\lim_{k\to\infty} \varepsilon_k = \frac{y^\star}{r}$. U tu svrhu najprije ćemo pokazati da je niz $(\varepsilon_k - \varepsilon_{k-1})$ konvergentan, te da mu je limes jednak nuli. Kako je (y_k) konvergentan, za proizvoljni $\mu > 0$ postoji $m \in \mathbb{N}$ takav da za sve $k \geq m$ vrijedi

$$|y_{k+1} - y_k| < \frac{(2-r)\mu}{2}. \tag{18}$$

Nadalje iz (17) slijedi da je

$$|\varepsilon_{k+1} - \varepsilon_k| \leq |y_{k+1} - y_k| + (r-1)|\varepsilon_k - \varepsilon_{k-1}|,$$

odakle za $p \in \mathbb{N}$, $p \geq 3$ induktivno dobivamo

$$\begin{aligned}
&|\varepsilon_{m+p+1} - \varepsilon_{p+1}| \leq |y_{m+p+1} - y_{m+p}| + (r-1)|\varepsilon_{m+p} - \varepsilon_{m+p-1}| \\
&\leq |y_{m+p+1} - y_{m+p}| + (r-1)|y_{m+p} - y_{m+p-1}| + (r-1)^2|\varepsilon_{m+p-1} - \varepsilon_{m+p-2}| \\
&\leq |y_{m+p+1} - y_{m+p}| + (r-1)|y_{m+p} - y_{m+p-1}| + \cdots + (r-1)^p|y_{m+1} - y_m| \\
&+ (r-1)^{p+1}|\varepsilon_m - \varepsilon_{m-1}|.
\end{aligned}$$

Iz (18) slijedi

$$
\begin{aligned}
&|\varepsilon_{m+p+1} - \varepsilon_{p+1}| \leq (1 + (r-1) + (r-1)^2 + \cdots + \\
&(r-1)^p)\frac{(2-r)\mu}{2} + (r-1)^{p+1}|\varepsilon_m - \varepsilon_{m-1}| \\
&< (1 + (r-1) + (r-1)^2 + \cdots + \\
&(r-1)^p + \cdots)\frac{(2-r)\mu}{2} + (r-1)^{p+1}|\varepsilon_m - \varepsilon_{m-1}| \\
&= \frac{1}{2-r}\frac{(2-r)\mu}{2} + (r-1)^{p+1}|\varepsilon_m - \varepsilon_{m-1}| = \frac{\mu}{2} + (r-1)^{p+1}|\varepsilon_m - \varepsilon_{m-1}|.
\end{aligned}
$$

Budući da je $\lim_{k\to\infty}(r-1)^k = 0$, postoji prirodni broj p_0 takav da za sve $p \geq p_0$ vrijedi

$$(r-1)^{p+1}|\varepsilon_m - \varepsilon_{m-1}| < \frac{\mu}{2}.$$

Konačno to znači da je $|\varepsilon_{k+1} - \varepsilon_k| < \mu$ za svaki $k \geq m + p_0$, odnosno da je $(\varepsilon_{k+1} - \varepsilon_k)$ konvergentan te da konvergira prema nuli.

Definirajmo nadalje novi niz (t_k) na sljedeći način

$$t_k = r\,\varepsilon_k = \varepsilon_k + (r-1)\varepsilon_k.$$

Iz definicije niza (t_k) i (17) slijedi

$$t_k = \varepsilon_k + y_{k+1} - \varepsilon_{k+1}.$$

Budući da niz $(\varepsilon_{k+1} - \varepsilon_k)$ konvergira prema nuli, slijedi da niz (t_k) također konvergira te da je $\lim_{k\to\infty} t_k = y^\star$. Kako je $\varepsilon_k = \frac{t_k}{r}$, slijedi da je (ε_k) konvergentan te da mu je limes jednak $\frac{y^\star}{r}$.

To znači da je

$$\lim_{k\to\infty}\frac{|e_{k+1}|}{|e_k|^r} = \lim_{k\to\infty} s_k = e^{\frac{y^\star}{r}} > 0,$$

čime smo pokazali da je brzina konvergencije ove metode $r = \frac{1+\sqrt{5}}{2}$.

□

Iz iterativne procedure (9) proizlazi Algoritam 2. Uočimo da smo u Algoritmu 2 naveli kriterij zaustavljanja, prema kojem Algoritam staje kada je derivacija funkcije f u trenutnoj aproksimaciji manja od unaprijed zadanog broja $\varepsilon > 0$. U literaturi postoje različiti drugi kriteriji zaustavljanja. Budući da zahtijevaju posebnu analizu i obrazloženja ovdje ih nećemo razmatrati.

2.2.2 Metoda II

U ovom pododjeljku kratko ćemo opisati još jednu metodu dvije točke. Slično kao i prvoj metodi pretpostavimo da su $x_0, x_1 \in \mathbb{R}$, $x_0 < x_1$ dvije različite točke za koje su poznate vrijednosti $f(x_0)$, $f(x_1)$ te $f'(x_0)$, pri čemu je $f(x_0) \neq f(x_1)$.

Algorithm 2 **Metoda dvije točke**

Input: f, x_0, x_1, ε

1: $f_0' = f'(x_0)$, $f_1' = f'(x_1)$
2: $\bar{x} = x_0 - \frac{x_0 - x_1}{f_0' - f_1'} f_0'$
3: **while** $|f'(\bar{x})| < \varepsilon$ (ili dok nije zadovoljen neki drugi kriterij zaustavljanja) **do**
4: $\quad x_0 = x_1$, $x_1 = \bar{x}$
5: $\quad f_0' = f'(x_0)$, $f_1' = f'(x_1)$
6: $\quad \bar{x} = x_0 - \frac{x_0 - x_1}{f_0' - f_1'} f_0'$
7: **end while**

Output: $\bar{x}$

U tom slučaju tražimo kvadratnu interpolacijsku funkciju $g : \mathbb{R} \to \mathbb{R}$, $g(x) = \alpha x^2 + \beta x + \gamma$ sa sljedećim svojstvima

$$\begin{aligned} g(x_0) &= \alpha x_0^2 + \beta x_0 + \gamma = f(x_0) =: f_0, \\ g(x_1) &= \alpha x_1^2 + \beta x_1 + \gamma = f(x_1) =: f_1, \\ g'(x_0) &= 2\alpha x_0 + \beta = f'(x_0) =: f_0'. \end{aligned} \tag{19}$$

Rješavanjem sustava (19) dobivamo da je

$$\begin{aligned} \alpha &= \tfrac{f_0 - f_1 - f_0'(x_0 - x_1)}{-(x_0 - x_1)^2}, \\ \beta &= f_0' + 2x_0 \tfrac{f_0 - f_1 - f_0'(x_0 - x_1)}{(x_0 - x_1)^2} \\ \gamma &= f(x_0) - \alpha x_0^2 - \beta x_0, \end{aligned}$$

odakle slijedi rekurzivna formula:

$$x_{k+1} = x_{k-1} - \frac{1}{2} \frac{(x_k - x_{k-1}) f_k'}{f_k' - \frac{f_k - f_{k-1}}{x_k - x_{k-1}}}, \quad k = 1, 2 \ldots \tag{20}$$

Za metodu koja se zasniva na iterativnoj proceduri (20) po uzoru na dokaz Teorema 2.2, može se pokazati analogna tvrdnja.

3 Metoda tri točke

Neka je $f : \mathbb{R} \to \mathbb{R}$ neprekidna funkcija koja općenito nije derivabilna. Kao i dosada s x^* označimo nepoznatu točku u kojoj se postiže lokalni minimum funkcije f. Odaberimo tri međusobno različite početne aproksimacije x_0, x_1 i x_2 i to tako da su točke $(x_0, f(x_0)), (x_1, f(x_1))$ te $(x_2, f(x_2))$ nekolinearne te da vrijedi

$$\min_{i=0,1,2} x_i \leq x^* \leq \max_{i=0,1,2} x_i.$$

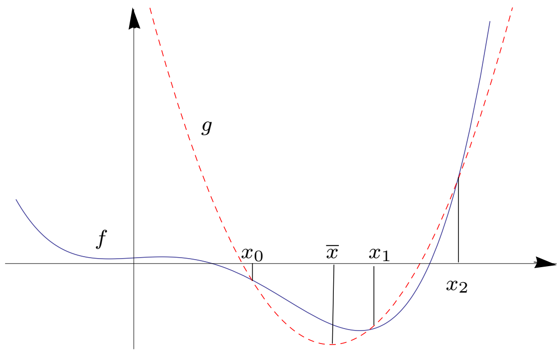
U tom slučaju postoji jedinstveni interpolacijski polinom g drugog stupnja koji prolazi kroz te tri točke te glasi

$$g(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}f_0 + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}f_1 + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}f_2,$$

gdje su $f_k := f(x_k), k = 0,1,2$. Slično kao i kod prethodnih metoda, umjesto točke u kojoj se postiže lokalni minimum funkcije f tražimo točku u kojoj se postiže globalni minimum kvadratne funkcije g. Nije teško vidjeti da iz $g'(x) = 0$ dobivamo

$$\overline{x} = \frac{1}{2}(x_0+x_1) + \frac{1}{2}\frac{(f_0-f_1)(x_1-x_2)(x_2-x_0)}{(x_1-x_2)f_0+(x_2-x_0)f_1+(x_0-x_1)f_2}. \tag{21}$$

Iz formule (21) proizlazi metoda kod koje polazeći od trojke točaka x_0, x_1 i x_2, dobivamo novu točku $\overline{x}$. U svakoj od tako dobivenih četvorki točaka izračunamo vrijednost funckije f te među njima biramo one tri točke $\tilde{x}_0, \tilde{x}_1, \tilde{x}_2$ u kojima je vrijednost funkcije najmanja. Postupak nastavljamo sve dok se ne zadovolji neka unaprijed zadana točnost. Potpuna procedura opisana je u Algoritmu 3.



Slika 3: Graf funkcije f i kvadratne interpolacijske funkcije g

Vrijedi sljedeći teorem, čiji iskaz i dokaz se mogu naći primjerice u [9].

Teorem 3.1 *Neka je $f : \mathbb{R} \to \mathbb{R}$ četiri puta neprekidno diferencijabilna. Ako je $x^\star \in \mathbb{R}$ takav da je $f'(x^\star) = 0$ te $f''(x^\star) \neq 0$, $f'''(x^\star) \neq 0$, $f^{iv}(x^\star) \neq 0$ onda postoji $\delta > 0$ takav da za svaki $x_0, x_1, x_2 \in \langle x^\star - \delta, x^\star + \delta\rangle$, $x_i \neq x_j$, $i \neq j$, Algoritam 3 konvergira prema $x^\star$ pri čemu je brzina konvergencije metode $r = 1.32$.*

Analogno kao i dosada u svrhu dokaza teorema potrebno je postaviti znatno jači zahtjev na funkciju f, od onog koji je potreban za provođenje Algoritma 3.

Algorithm 3 **Metoda tri točke**

Input: f, x_0, x_1, x_2 tako da je $\min\{x_0, x_1, x_2\} \leq x^* \leq \max\{x_0, x_1, x_2\}$, ε

1: $f_0 = f(x_0)$, $f_1 = f(x_1)$, $f_2 = f(x_2)$

2: $\bar{x} = \frac{1}{2}(x_0 + x_1) + \frac{1}{2}\frac{(f_0 - f_1)(x_1 - x_2)(x_2 - x_0)}{(x_1 - x_2)f_0 + (x_2 - x_0)f_1 + (x_0 - x_1)f_2}$

3: **if** $(\overline{x} - x_0)(\overline{x} - x_2) \geq 0$ **then**

4:     **goto** Korak 8

5: **else**

6:     **goto** Korak 9

7: **end if**

8: Konstruirati novu trojku $\{x_0, x_1, x_2\}$ iz x_0, x_1, x_2 i $\overline{x}$ za koje je vrijednost funkcije f najmanja i **goto** Korak 1

9: **if** $|\overline{x} - x_1| < \varepsilon$ **then**

10:     **goto** Korak 13

11: **else**

12:     **goto** Korak 8

13: **end if**

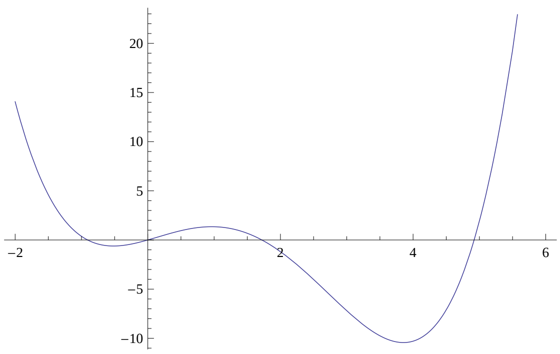
Output: $\bar{x}$

4 Numerički primjer

U ovom odjeljku s jednim jednostavnim numeričkim primjerom ilustrirat ćemo efikasnost prethodno navedenih metoda te potvrditi teorijske rezultate koji se odnose na brzinu konvergencije pojedine metode. Promatrajmo u tu svrhu funkciju $f : \mathbb{R} \to \mathbb{R}$, zadanu formulom

$$f(x) = 2 - \frac{2}{25}x + \frac{61}{100}x^2 - \frac{43}{30}x^3 + \frac{1}{4}x^4.$$

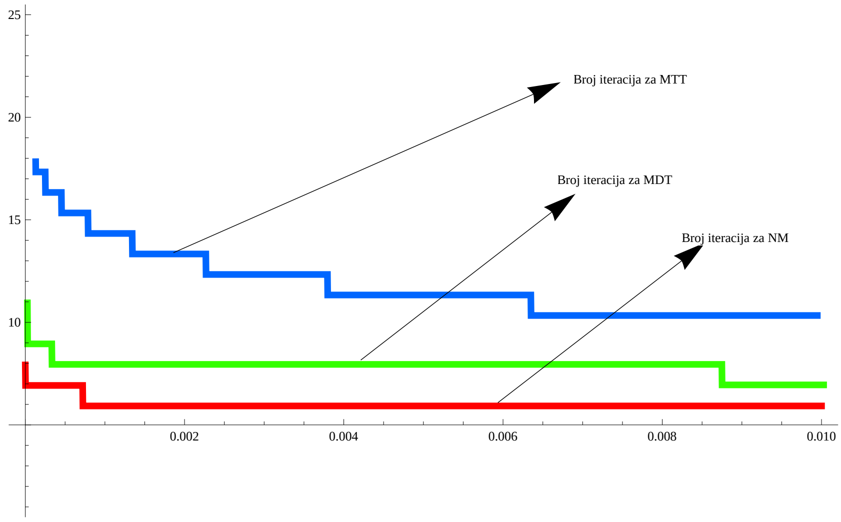
Graf funkcije f prikazan je na Slici 4.



Slika 4: Graf funkcije f

Nije teško vidjeti da f u $x^\star = 4$ postiže lokani minimum, koji je ujedno i globalni minimum funkcije. Svaku od prehodno opisanih metoda primjenjujemo na funkciju f uz početne aproksimacije $x_0 = 1.5, x_1 = 3$ i $x_2 = 6$. Kriterij zaustavljanja definiramo tako da algoritam staje ako je $|x_k - 4| < \varepsilon$, gdje je $\varepsilon > 0$ unaprijed zadana točnost. Pri tome za različite $\varepsilon \in [10^{-8}, 0.01]$ određujemo broj iteracija koji je potreban da bi se odgovarajući algoritam zaustavio. Početne aproksimacije odabrane su tako da sve metode konvergiraju.

Na Slici 5 prikazan je broj iteracija za Metodu jedne točke-Newtonovu metodu `(NM)`, Metodu dvije točke `(MDT)` i Metodu tri točke `(MTT)` u ovisnosti o parametru ε. Kako je i očekivano sa Slike 5 vidimo da `NM` treba najmanje iteracija, `MDT` nešto više, dok `MTT` treba značajno najviše iteracija da bi se zadovoljila zadana točnost.



Slika 5: Broj iteracija za Newtonovu metodu `(NM)`, Metodu dvije točke `(MDT)` i Metodu tri točke `(MTT)`

Literatura

[1] M. Alić, G. Nogo, *Optimizacija: Uvod u teoriju nužnih i dovoljnih uvjeta ekstrema*, Odjel za matematiku, Sveučilište u Osijeku, Osijek, 2004

[2] D. R. Jones, C. D. Perttunen, B. E. Stuckman, *Lipschitzian optimization without the Lipschitz constant*, JOTA **79**(1993), 157–181

[3] S. Kurepa, *Matematika analiza 2, funkcije jedne varijable*, Tehnika knjiga, Zagreb, 1987.

[4] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evoluti on Programs*, Springer, New York, 1996

[5] K. V. Price, R. M. Storn, J. A. Lampinen, *Differential Evolution. A Practical Approach to Global Optimization*, Springer-Verlag Berlin Heidelberg, 2005

[6] R. Scitovski, *Numerička matematika*, Sveučilište u Osijeku, Odjel za matematiku, 2014.

[7] R. Scitovski, N. Truhar, Z. Tomljanović, *Metode optimizacije*, Sveučilište Josipa Jurja Strossmayera u Osijeku, Odjel za matematiku., Osijek, 2014.

[8] G. W. Stewart, *On the convergence of multipoint iterations*, Numerical Mathematics, **68**(1994), 143-147

[9] W. Sun, Y. Yuan, *Optimization theory and methods, Nonlinear programming*, Springer-Verlag, New York, 2006.

Primljeno u redakciju Časopisa 06.09.2015; Revidirana verzija 03.11.2015;
Dostupno na Internetu 09.11.2015.