# Znanstveno računanje 2

## 5. dio vježbi
## Nelinearni sustavi jednadžbi

Nela Bosner

# Nelinearni sustavi jednadžbi: Newtonova metoda



- Neka je dana nelinearna funkcija $f : \mathbb{R}^n \to \mathbb{R}^n$.
- Želimo naći rješenje nelinearnog sustava od $n$ jednadžbi sa $n$ nepoznanica

$$f(x) = \begin{bmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_n(x_1, \ldots, x_n) \end{bmatrix} = 0.$$

- Ako pretpostavimo da je $x = \xi$ nultočka od $f$, da je $x_0$ aproksimacija od $\xi$, i da je $f$ diferencijabilna u nekoj okolini od $x_0$, tada možemo primijeniti Newtonovu metodu na ovaj sustav, tako da rješenje dobijemo iterativno

$$x_{i+1} = x_i - (Df(x_i))^{-1} f(x_i), \quad i = 0, 1, 2, \ldots.$$

- Newtonova metoda konvergira kvadratično, ali samo ako je početna aproksimacija $x_0$ dovoljno blizu traženom rješenju $\xi$.
- Kako bi se izbjegao problem uskog izbora početne aproksimacije koristi se **modificirana Newtonova metoda** koja je kombinacija
  - Newtonove metode
  - pretraživanje po pravcu — optimizacijska metoda
  - metode raspolavljanja

  i za koju se može dokazati globalna konvergencija za veliku klasu funkcija $f$.

- Definirajmo $h(x) = f(x)^T f(x) = \|f(x)\|_2^2$.
- Modifikacija algoritma se sastoji u uvođenju dodatnog parametra $\lambda$, i smjera traženja $s$ kako bismo definirali niz

$$x_{k+1} = x_k - \lambda_k s_k,$$

  gdje je u našem slučaju
  - $s_k = d_k = (Df(x_k))^{-1} f(x_k)$, smjer iz Newtonove metode
  - a $\lambda_k$ se odabire tako da minimizira funkciju $h_k(\lambda) = h(x_k - \lambda s_k)$,
  - da niz $\{h(x_k)\}$ bude strogo padajući, i da $x_k$ konvergira minimumu funkcije $h$.
- Budući da je $h(x) \geq 0$ za svaki $x$, slijedi

$$h(\xi) = 0 \Longleftrightarrow f(\xi) = 0.$$

## Algoritam (Modificirana Newtonova metoda)

$x_0$ *zadan;*
$k = 0$;
***while*** $\sim$*kriterij_zaustavljanja*
$\quad d_k = (Df(x_k))^{-1} f(x_k)$;
$\quad \gamma_k = \frac{1}{\kappa_2(Df(x_k))}$;
$\quad$ *Definiramo* $h(x) = f(x)^T f(x)$ *i* $h_k(\tau) = h(x_k - \tau d_k)$;
$\quad$ *Nađi najmanji cijeli broj* $j \geq 0$ *takav da je*
$\qquad h_k(2^{-j}) \leq h_k(0) - 2^{-j} \frac{\gamma_k}{4} \|d_k\|_2 \|Dh(x_k)\|_2$;
$\quad$ *Nađi* $i_{min} \in \{0, 1, \ldots, j\}$ *takav da je*
$\qquad h_k(2^{-i_{min}}) = \min_{i=0,\ldots,j} h_k(2^{-i})$;
$\quad \lambda_k = 2^{-i_{min}}$;
$\quad x_{k+1} = x_k - \lambda_k d_k$;
$\quad k = k + 1$;
***end***
$\xi \approx x_k$;

- Iz definicije je jasno da vrijedi

$$Dh(x) = 2f(x)^T Df(x).$$

- Pri izvršavanju danog algoritma, može se dogoditi da $\lambda_k$ bude jako mali, tako da korak koji se dodaje na $x_k$ bude skoro zanemariv. U tom slučaju dobro je staviti ograničenje na $\lambda_k$ odozdo, tako da se u slučaju $\lambda_k < 0.01$ postavi $\lambda_k = 0.01$.

# Zadaci



## Zadatak

*Napišite M-file funkciju* `mod_newton()` *koja implementira prethodno opisanu modificiranu Newtonovu metodu za sustav $f(x) = 0$. Funkcija neka ima ulazne parametre*

- *pokazivač na funkciju* $f$
- *pokazivač na funkciju* $df$ *koja implementira* $Df(x)$
- *početnu aproksimaciju* $x_0$

*Funkcija neka ispiše broj iteracija potreban za zadovoljavanje kriterija zaustavljanja i neka vraća*

- *rješenje* $x$, *takvo da je* $f(x) = 0$.

*Kao kriterij zaustavljanja uzmite da je reletivna razlika komponenata između dvije uzastopne aproksimacija manja od mašinskog epsilona:*

$$\max_{i=1,\dots,n} \frac{|x_{k+1}(i) - x_k(i)|}{|x_k(i)|} < \varepsilon_M.$$

## Zadatak

*Svoju funkciju* `mod_newton()` *testirajte na sustavu*

$$f(x) = \begin{bmatrix} x_1^2 + x_2^2 - 2 \\ e^{x_1 - 1} + x_2^3 - 2 \end{bmatrix} = 0,$$

*za kojeg je točno rješenje jednako*

$$\xi = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

- *Usporedite rezultate sa običnom Newtonovom metodom koja ima isti kriterij zaustavljanja.*
- *Za početne točke uzmite*
  1. $x_0 = [1.5, 2]^T$,
  2. $x_0 = [0.5, 0.4]^T$.

## Zadatak (nastavak)

*Što primjećujete kod rješavanja ovog problema ovim dvijema različitim varijantama Newtonove metode?*

- *Za* $x_0 = [1.5, 2]^T$ *modificirana Newtonova metoda čak konvergira jednako kao i obična Newtonova metode, međutim,*
- *za* $x_0 = [0.5, 0.4]^T$ *obična Newtonova metoda ima problema sa izvođenjem zbog skoro singularnog diferencjala* $Df(x_i)$ *i divergira, dok modificirana Newtonova metoda uspijeva izračunati rješenje, i to je njena prednost.*

# Primjeri iz primjene: Grijanje ploče sa zaštitnim slojem



## Primjer

- *Zaštitni sloj ploče izložen je zračenju topline iz grijača.*
- *Temperaturu zaštitnog sloja možemo dobiti iz procesa prijenosa topline konvekcijom i zračenjem.*
- *Ako se zračenje tretira pod određenim uvjetima, tada, uz uvjet da je temperatura okoline jednaka* $298^\circ K$*, dobivamo sljedeći nelinearni sustav jednadžbi za nepoznanice* $J_g$, $T_g$, $J_{zs}$ *i* $T_{zs}$

$$
\begin{aligned}
5.67 \cdot 10^{-8} T_{zs}^4 + 17.41 T_{zs} - J_{zs} &= 5188.18 \\
J_{zs} - 0.71 J_g + 7.46 T_{zs} &= 2352.71 \\
5.67 \cdot 10^{-8} T_g^4 + 1.865 T_g - J_g &= 2250 \\
J_g - 0.71 J_{zs} + 7.46 T_g &= 11093
\end{aligned}
$$

## Primjer (nastavak)

*gdje su*

- $J_g$ *i* $J_{zs}$ *potpuno emitirano i reflektirano zračenje sa površine grijalice odnosno zaštitnog sloja,*
- $T_g$ *i* $T_{zs}$ *temperature grijalice i zaštitnog sloja izražene u* $^\circ K$.

## Zadatak

*Riješite gornji primjer pomoću modificirane Newtonove metode, pri čemu su početne vrijednosti:*

| $J_g$ | $T_g$ | $J_{zs}$ | $T_{zs}$ |
|---|---|---|---|
| 8000 | 298 | 5000 | 298 |

# Računanje minimuma funkcije: gradijentna metoda



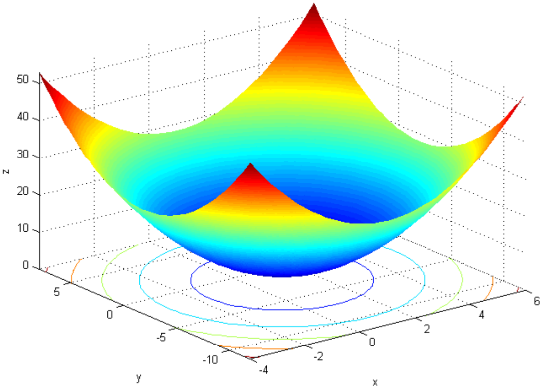
$$\boxed{f \searrow \min}$$

- Imamo: $f : \mathbb{R}^n \longrightarrow \mathbb{R}$.
- Tražimo: $x^* = \arg\min_{x\in\mathbb{R}^n} f(x)$.
- Ono što ćemo naći: lokalni minimum.

**Primjer**

*Tražimo minimum funkcije* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$

$$f(x, y) = x^2 + \frac{1}{4}y^2 - 2x + y + 5$$




Slika: Graf funkcije $f(x,y) = x^2 + \frac{1}{4}y^2 - 2x + y + 5$.

# Svojstva funkcije





**Teorem**

*Ako je $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ neprekidno diferencijabilna funkcija ($f$ je klase $C^1$) i ako ona poprima lokalni minimum u $x^*$, tada $\nabla f(x^*) = 0$.*

**Teorem**

*Ako je $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ klase $C^2$ i ako ona poprima lokalni minimum u $x^*$, tada $\nabla f(x^*) = 0$ i Hesseova matrica $\nabla^2 f(x^*)$ je pozitivno semidefinitna.*

### Teorem

*Ako je $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ klase $C^2$, te $\nabla f(x^*) = 0$ i Hesseova matrica $\nabla^2 f(x^*)$ je pozitivno definitna, tada funkcija $f$ poprima striktni minimum u $x^*$.*

### Teorem

*Ako je $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ konveksna funkcija, tada je svaki lokalni minimum ujedno i globalni minimum.*

## Primjer

*Funkcija* $f(x,y) = x^2 + \frac{1}{4}y^2 - 2x + y + 5$ *je klase* $C^2$*:*

$$\nabla f(x,y) = \begin{bmatrix} 2x-2 \\ \frac{1}{2}y+1 \end{bmatrix},$$
$$\nabla^2 f(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

*Štoviše,*

- *Hesseova matrica* $\nabla^2 f(x,y)$ *je pozitivno definitna za svaki* $x$*,*
- $f$ *je konveksna funkcija,*
- *iz uvjeta* $\nabla f(x^*,y^*) = 0$ *slijedi da se globalni minimum postiže u* $(1,-2)$*.*

# Skica metode



- Rješavamo problem $\min_{x \in \mathbb{R}^n} f(x)$,
- imamo početnu aproksimaciju $x_0 \in \mathbb{R}^n$.

Odredit ćemo niz točaka $x_0, x_1, x_2, \ldots$ za koje je

$$f(x_0) > f(x_1) > f(x_2) > \cdots$$

**Algoritam (Skica iterativnog postupka)**

1. *odaberi* $x_0$, $i = 0$;
2. *dok je* $\nabla f(x_i)! = \Theta$
3.     *nađi* $x_{i+1}$ *takav da je* $f(x_{i+1}) < f(x_i)$;
4.     $i = i + 1$;
5. *kraj petlje.*

- Ako je niz $\{f(x_i)\}_i$ odozdo ograničen, tada postoji
$$\lim_{i\to\infty} f(x_i).$$
- Međutim, $\lim_{i\to\infty} x_i$ ne mora postajati. Npr.

$$f(x) = e^{-x} \qquad\qquad x_i = i$$
$$\lim_{i\to\infty} f(x_i) = 0 \qquad\qquad \lim_{i\to\infty} x_i \text{ ne postoji.}$$

Da bi postojao $\lim_{i\to\infty} x_i$ nužno nam je dodatno svojstvo:

**Svojstvo**

*Ako postoji $x_0 \in \mathbb{R}^n$ takav da je skup*
$$\{x \mid f(x) \le f(x_0)\}$$
*kompaktan, tada $\lim_{i\to\infty} x_i$ postoji.*

# Pretraživanje po pravcu



$$\boxed{\text{Kako odabrati } x_{i+1}\text{?}}$$

- Svesti višedimenzionalan problem u jednodimenzionalan: **pretraživanje po pravcu.**
- Osnovna ideja pretraživanja po pravcu je:
  1. odaberi smjer $s_i$,
  2. odaberi korak $\lambda_i$,
  3. definiraj $x_{i+1} = x_i + \lambda_i s_i$, takav da je $f(x_{i+1}) < f(x_i)$.

# Odabir smjera

Za neki smjer $s$ definirajmo funkciju $g : \mathbb{R} \longrightarrow \mathbb{R}$

$$g(\lambda) = f(x + \lambda s).$$

Svojstva funkcije $g$:

- $g(0) = f(x)$,
- ako je $f$ klase $C^1$ tada je i $g$ klase $C^1$,
- ako je $s$ smjer u kojem funkcija $f$ lokalno pada u okolini točke $x$, tada i funkcija $g$ pada u 0:

$$\begin{aligned} f(x + \lambda s) &< f(x), \qquad \lambda \in \mathbb{R}, \lambda > 0, \text{mali} \\ g'(0) &< 0. \end{aligned}$$

Vrijedi:

$$g'(\lambda) = \langle \nabla f(x + \lambda s), s \rangle,$$

pa iz uvjeta $g'(0) < 0$ slijedi

$$g'(0) = \langle \nabla f(x), s \rangle < 0.$$

### Definicija

*Za funkciju $f$ i točku $x$ definiramo skup*

$$D(x) = \{s \mid \langle \nabla f(x), s \rangle < 0\}.$$

*Ovaj skup nazivamo* ***skup smjerova silaska*** *a njegove elemente* ***smjerovi silaska****.*

# Odabir koraka

- veličina koraka je bitna za konvergenciju metode.
- Uvjet $f(x_{i+1}) < f(x_i)$ za odabir točke $x_{i+1}$ nije dovoljan. Npr.

$$f(x) = x^2 \qquad x_0 = 2, x_1 = 1.5, \ldots, x_i = 1 + \frac{1}{i+1},$$

$$\lim_{i\to\infty} f(x_i) = 1 \qquad \lim_{i\to\infty} x_i = 1,$$

ali niz $\{f(x_i)\}_i$ ne konvergira ka minimumu funkcije $f$.

**Maksimalno spuštanje:**

$$\lambda_i = \arg\min_{\lambda>0} f(x_i + \lambda s_i).$$

- Maksimalno spuštanje po pravcu je problem minimizacije funkcije $g : \mathbb{R} \to \mathbb{R}$, $g(\lambda) = f(x_i + \lambda s_i)$.
- Taj minimum računamo tako da tražimo nultočku funkcije

$$g'(\lambda) = \langle \nabla f(x_i + \lambda s_i), s_i \rangle = \nabla f(x_i + \lambda s_i)^T s_i.$$

- Izračunati ga možemo pomoću obične Newtonove metoda sa $\lambda_0 = 0$.
- Za Newtonovu metodu nam još treba i

$$g''(\lambda) = s_i^T \nabla^2 f(x_i + \lambda s_i) s_i.$$

# Gradijentna metoda (metoda najbržeg silaska)



**Ideja**

*Odaberi smjer u kojem funkcija f najbrže pada.*

**Koji je to smjer?**

Podsjetimo se: $g'(0) = \langle \nabla f(x), s \rangle$.

- Riješimo problem
$$\min_{\|s\|=1} \langle \nabla f(x_i), s \rangle.$$
- Neka je $\theta$ kut između $\nabla f(x_i)$ i $s$,
- budući da je $\|s\| = 1$, tada vrijedi
$$\langle \nabla f(x_i), s \rangle = \|\nabla f(x_i)\| \cos \theta.$$
- Minimum se postiže za $\theta = \pi$, odakle je $\cos \theta = -1$, i traženi smjer je do na skalirajući faktor jednak
$$s_i = -\nabla f(x_i).$$

# Algoritam gradijentne metode





## Algoritam

1. *odaberi* $x_0$, $i = 0$;
2. *dok je* $\nabla f(x_i)! = \Theta$
3. $\quad s_i = -\nabla f(x_i)$;
4. $\quad \lambda_i = \arg\min_{\lambda>0} f(x_i + \lambda s_i)$;
5. $\quad x_{i+1} = x_i + \lambda_i s_i$;
6. $\quad i = i + 1$;
7. *kraj petlje.*

# Konvergencija gradijentne metode

**Teorem**

*Neka je* $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ *i neka je* $x_0 \in \mathbb{R}^n$ *izabran da vrijedi*

- $K := \{x \mid f(x) \leq f(x_0)\}$ *je kompaktan,*
- $f$ *je neprekidno diferencijabilna na nekom otvorenom skupu koji sadrži* $K$*.*

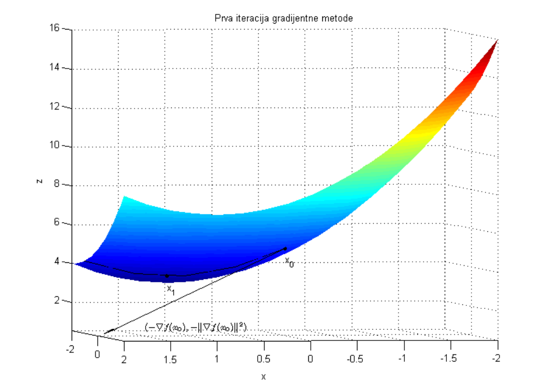
*Tada za svaki niz* $\{x_i\}_i$ *definiran gradijentnom metodom vrijedi*

- $x_i \in K$ *za sve* $i = 0, 1, 2, \ldots$ *i* $\{x_i\}_i$ *ima barem jedno gomilište* $\bar{x}$ *u* $K$*.*
- *Svako gomilište niza* $\{x_i\}_i$ *je stacionarna točka od* $f$*; to jest* $\nabla f(\bar{x}) = 0$*.*

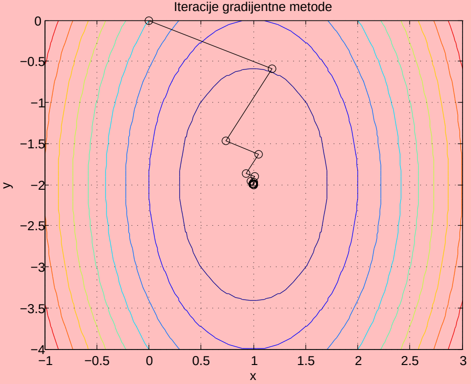
## Primjer

*Izvršavamo gradijentnu metodu za*

$$f(x,y) = x^2 + \frac{1}{4}y^2 - 2x + y + 5, \quad i \quad x_0 = (0,0).$$



## Primjer




Slika: Iteracije gradijentne metode za funkciju $f(x, y) = x^2 + \frac{1}{4}y^2 - 2x + y + 5$ i $x_0 = (0, 0)$.

# Zadaci



## Zadatak

*Napišite M-file funkciju* `gradijentna_metoda()` *koja implementira gradijentnu metodu za traženje minimuma funkcije $f$. Funkcija neka ima ulazne parametre*

- *pokazivač na funkciju gradf koja predstavlja $\nabla f$*
- *pokazivač na funkciju hessf koji predstavlja $\nabla^2 f$*
- *početnu aproksimaciju minimuma $x_0$.*

*Kriterij zaustavljanja iteracija gradijentne metode neka je*

$$\|\nabla f(x_i)\|_2 \leq 10^{-10}.$$

*Implementirajte i običnu Newtonovu metodu za funkciju $g'(\lambda)$, pri čemu je njen kriterij zaustavljanja dan sa*

$$\frac{|\lambda_j - \lambda_{j-1}|}{\lambda_{j-1}} \leq 10^{-5}.$$

### Zadatak (nastavak)

*Funkcija neka vraća*

- *x koji je minum funkcije f*
- *broj iteracija i potrebnih za dostizanje tražene točnosti.*

# Primjeri iz primjene: Optimalni smještaj tvornice

## Primjer

- *Pretpostavimo da se proizvođači sirovina potrebnih za neku tvornicu nalaze na koordinatama* $(x_1, y_1)$, $(x_2, y_2)$ *i* $(x_3, y_3)$, *i da se trgovine koje prodaju proizvode te tvornice nalaze na koordinatama* $(x_4, y_4)$ *i* $(x_5, y_5)$.
- *Pretpostavimo da su troškovi po kilometru transporta od ili do gornjih lokacije dani sa* $c_i$, $i = 1, \ldots, 5$ *(ovisno o vrsti transporta).*
- *Problem optimalnog smještaja tvornice se tada svodi na pronalaženje lokacije tvornice za koju bi ukupni troškovi transporta od proizvođača do tvornice, i od tvornice do trgovina bili minimalni.*

## Primjer (nastavak)

- *Dakle, minimiziramo funkciju ukupnih troškova transporta*

$$f(x, y) = \sum_{i=1}^{5} c_i \sqrt{(x - x_i)^2 + (y - y_i)^2}.$$

## Zadatak

*Primijenite svoju funkciju* `gradijentna_metoda()` *na ovaj primjer, pri čemu su zadani sljedeći parametri:*

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $x_i$ | 43 | 13 | 115 | 119 | 33 |
| $y_i$ | 167 | 29 | 119 | 4 | 17 |
| $c_i$ | 12 | 10 | 14 | 9 | 19 |

## Primjer (nastavak)

- *Sami izračunajte i napišite funkcije koje implementiraju* $\nabla f$ *i* $\nabla^2 f$.
- *Pomoću MATLAB-ove funkcije* `contour()` *nacrtajte nivo-skupove funkcije* $f$ *na kvadratu* $[-100, 100] \times [-100, 100]$*, i iz dobivenog grafa odredite početnu aproksimaciju* $x_0$.
- *Za konačno rješenje provjerite* $\|\nabla f(x)\|_2$.