Kapitola 5. SLAR - gradientní metody

Metody na řešení SLAR

- přímé (GEM, metoda LU-rozkladu) ✓
- iterační (Jacobiova m., Gauss-Seidelova m., metoda SOR) ✓
- gradientní

Motivace

Uvažujme kvadratickou funkci reálné proměnné x:

$$f(x) = \frac{1}{2}ax^2 - bx + c, \quad a > 0.$$

Nutná a postačující podmínka minima funkce $\big(f'(x) = 0\big)$ má tvar

$$ax = b.$$

To znamená, že místo řešení lineární rovnice můžeme řešit úlohu najít minimum konvexní kvadratické funkce $f(x)$ (obě úlohy mají stejná řešení).

Uvědomme si, že v případě funkce více proměnných je třeba splnit další podmínky kladené na matici soustavy $\mathbf{A}$, abychom zaručili konvexnost příslušné kvadratické funkce.

Uvažujeme soustavu (kde matice $\mathbf{A}$ je symetrická, pozitivně definitní)

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

Dále uvažujeme kvadratickou formu, tzv. **energetický funkcionál**

$$F(\mathbf{x}) = \frac{1}{2}\,\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{b}^T\mathbf{x}.$$

Platí

$$\operatorname{grad} F(\mathbf{x}) = \mathbf{A}\ \mathbf{x} - \mathbf{b}.$$

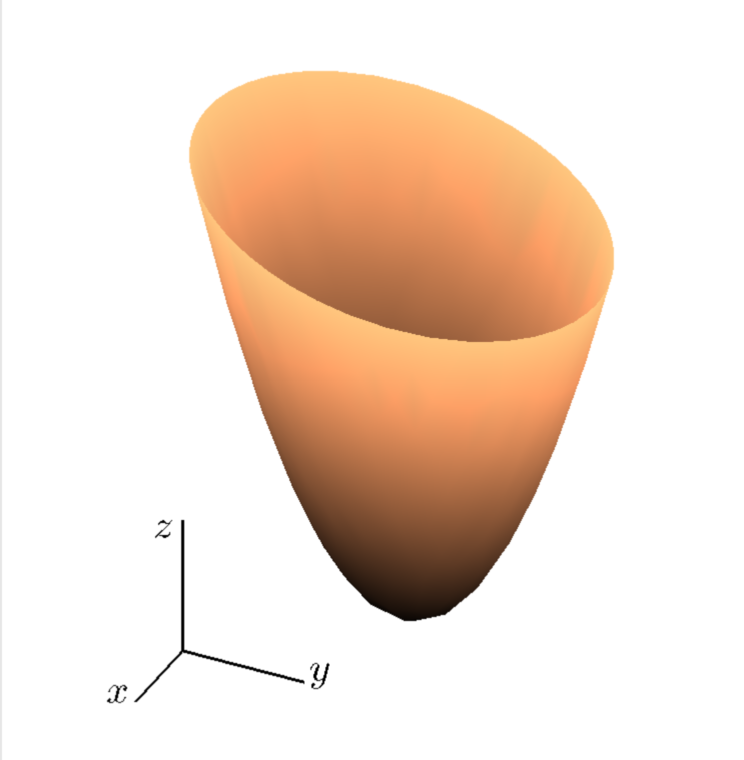
Funkce $F(\mathbf{x})$ je konvexní a kvadratická $\quad\Rightarrow\quad F(\mathbf{x})$ má globální minimum a pro bod minima $\tilde{\mathbf{x}}$ platí

$$\operatorname{grad} F(\tilde{\mathbf{x}}) = \mathbf{A}\tilde{\mathbf{x}} - \mathbf{b} = \mathbf{0}.$$

Bod minima $\tilde{\mathbf{x}}$ je tedy řešením soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Poznámka: Úlohy najít bod minima funkce F a řešit soustavu $\mathbf{A}\mathbf{x} = \mathbf{b}$ jsou ekvivalentní.

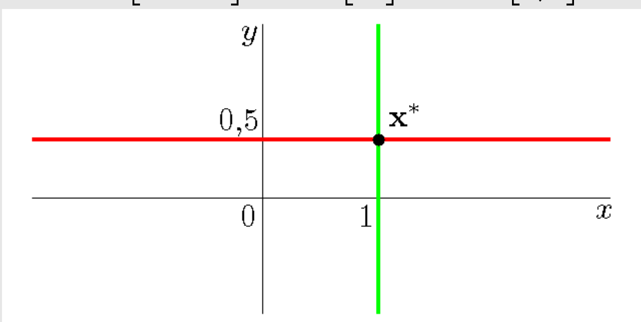
Poznámka: V případě soustavy 2 rovnic si lze udělat geometrickou představu, neboť pro $\mathbf{x} \in \mathbb{R}^2$ je grafem funkce $F(\mathbf{x})$ eliptický paraboloid, jehož vrstevnice jsou elipsy. Minima $F(\mathbf{x})$ se nabývá ve vrcholu paraboloidu.



Příklad 1

Uvažujme velmi jednoduchou soustavu $\mathbf{A}\mathbf{x} = \mathbf{b}$, kde

$$\mathbf{A} = \begin{bmatrix} 25 & 0 \\ 0 & 16 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 25 \\ 8 \end{bmatrix}, \quad \mathbf{x}^* = \begin{bmatrix} 1 \\ 0,5 \end{bmatrix}.$$



Odpovídající kvadratická funkce je

$$F(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{b}^T\mathbf{x} = \frac{1}{2}\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 25 & 0 \\ 0 & 16 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} 25 & 8 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{2}(25x^2 + 16y^2) - 25x - 8y.$$

Vrstevnice (hladiny):

$$\boxed{F(\mathbf{x}) = c}$$

$$\frac{1}{2}(25x^2 + 16y^2) - 25x - 8y = c$$

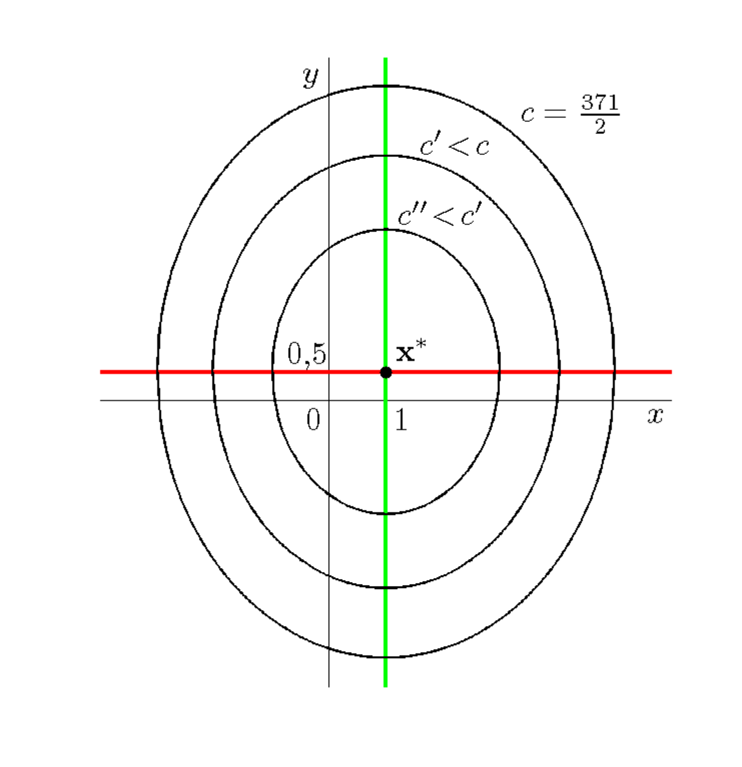
$$25x^2 + 16y^2 - 50x - 16y = 2c$$

$$25(x-1)^2 - 25 + 16(y - \frac{1}{2})^2 - 4 = 2c$$

$$25(x-1)^2 + 16(y - \frac{1}{2})^2 = 2c + 29$$

např. pro $c = \frac{371}{2}$:

$$\boxed{\frac{(x-1)^2}{4^2} + \frac{(y-\frac{1}{2})^2}{5^2} = \frac{2c+29}{400} = 1}$$

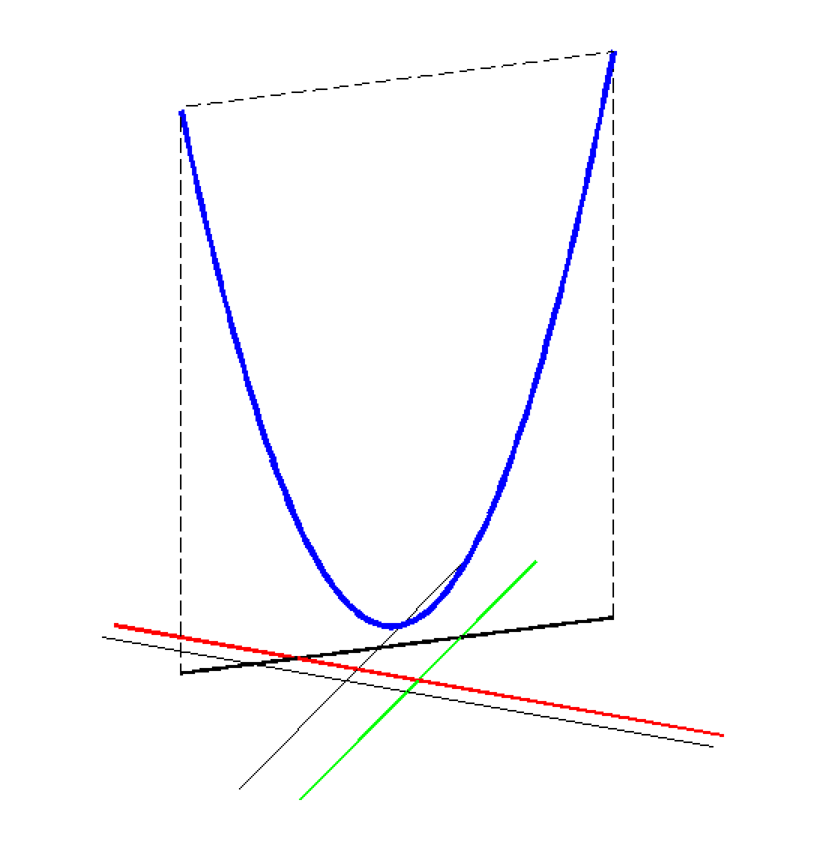


Řezy svislou rovinou $\quad y = px + q$

$$F(x) = \frac{1}{2}(25x^2 + 16y^2) - 25x - 8y = \frac{1}{2}(25x^2 + 16(px+q)^2) - 25x - 8(px+q) =$$

$$= \frac{1}{2}(25x^2 + 16(p^2x^2 + 2pqx + q^2)) - 25x - 8(px+q) =$$

$$= \underbrace{(\frac{25}{2} + \frac{16}{2}p^2)}_{>0} x^2 + (16pq - 25 - 8p)x + 8q^2 - 8q$$

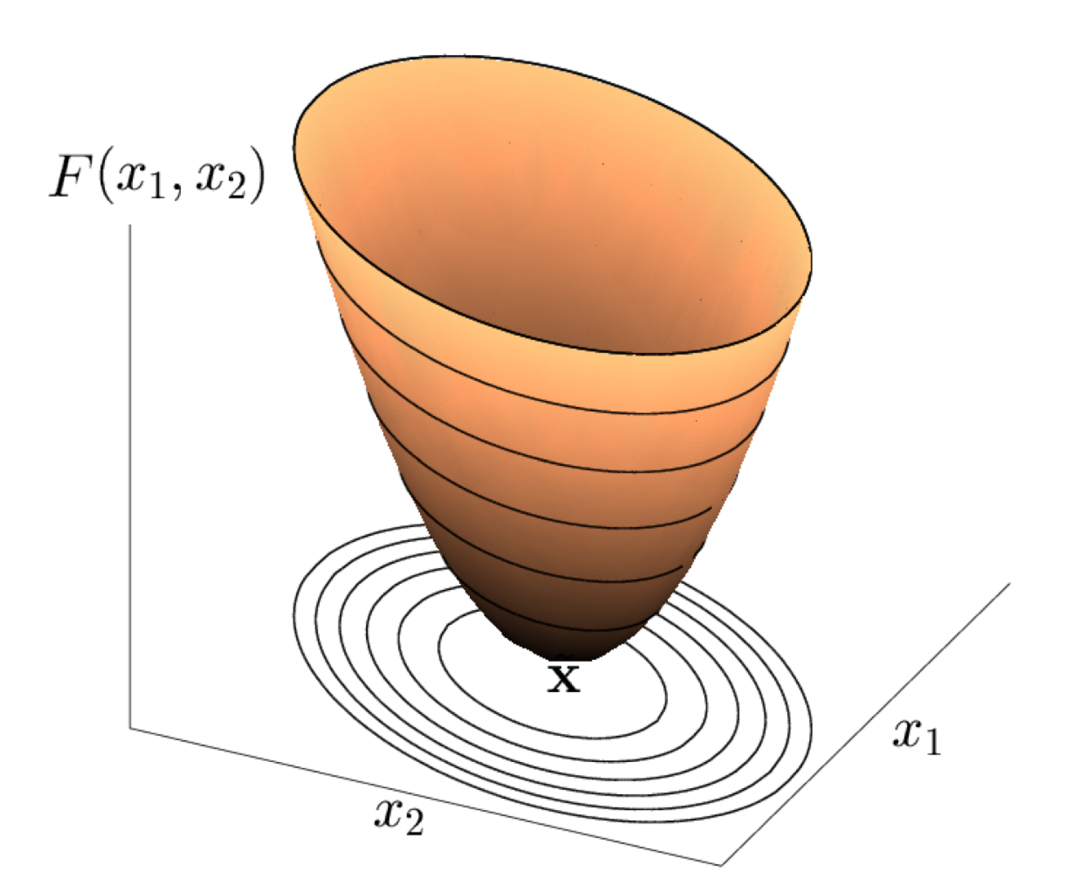


Princip

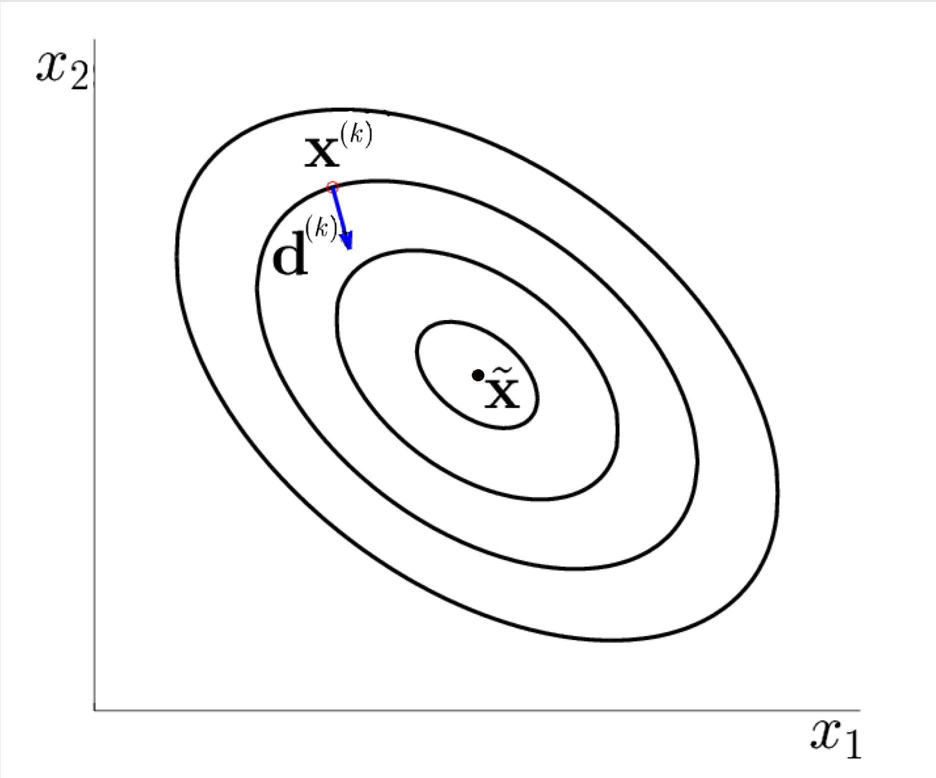
Stejně jako u každé iterační metody nejprve zvolíme počáteční aproximaci řešení $\mathbf{x}^{(0)}$.

Princip gradientních metod spočívá v tom, že zvolíme směr a v tomto směru se budeme chtít co nejvíce přiblížit k přesnému řešení. Gradientní metoda je tedy určena volbou směrů, ve kterých minimalizujeme funkci F.

Během jedné iterace se pohybujeme po povrchu grafu funkce $F(\mathbf{x})$ tak, abychom se dostali na nižší vrstevnici.



V případě soustavy dvou rovnic získáme promítnutím grafu funkce $F(\mathrm{x})$ do roviny proměnných x_1, x_2 systém soustředných elips - hladin (vrstevnic).



Metoda největšího spádu

Metodu největšího spádu získáme, pokud budeme za směrové vektory volit směry největšího spádu, tj. vektory

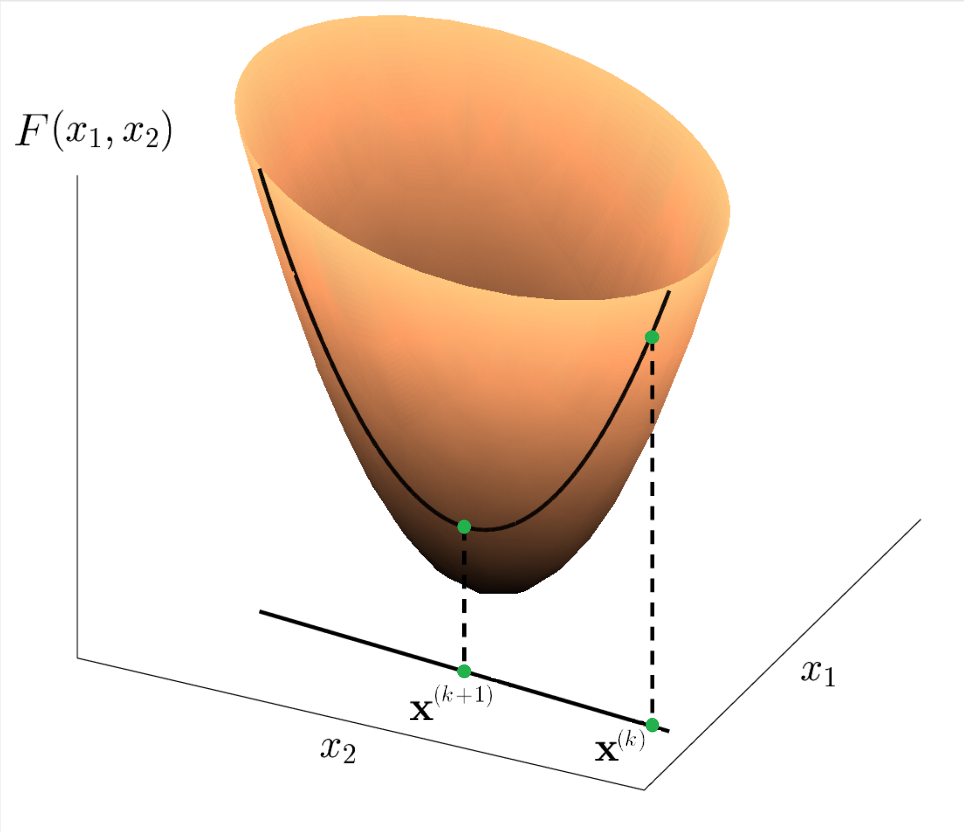
$$\boxed{\mathbf{d}^{(k)} = -\operatorname{grad} F(\mathbf{x}^{(k)}) = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k)}}.$$

Iterační formuli volíme ve tvaru

$$\boxed{\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + t^{(k)} \cdot \mathbf{d}^{(k)}},$$

v každém kroku metody určíme směr největšího spádu $\mathbf{d}^{(k)}$ a provedeme jednorozměrnou minimalizaci v tomto směru, tj.

$$\min_{t>0} F(\mathbf{x}^{(k)} + t\,\mathbf{d}^{(k)}).$$



Minimalizovanou funkci proměnné t označíme $\Psi(t)$.

Potom platí:

$$\underbrace{F(\mathbf{x}^{(k)} + t\,\mathbf{d}^{(k)})}_{\Psi(t)} = \frac{1}{2}(\mathbf{x}^{(k)} + t\,\mathbf{d}^{(k)})^T \mathbf{A}(\mathbf{x}^{(k)} + t\,\mathbf{d}^{(k)}) - \mathbf{b}^T(\mathbf{x}^{(k)} + t\,\mathbf{d}^{(k)}) =$$

$$= \frac{1}{2}\,\mathbf{x}^{(k)T}\mathbf{A}\mathbf{x}^{(k)} + \frac{1}{2}\,t\,\mathbf{x}^{(k)T}\mathbf{A}\mathbf{d}^{(k)} + \frac{1}{2}\,t\,\mathbf{d}^{(k)T}\mathbf{A}\mathbf{x}^{(k)} + \frac{1}{2}\,t^2\,\mathbf{d}^{(k)T}\mathbf{A}\mathbf{d}^{(k)} - \mathbf{b}^T\mathbf{x}^{(k)} - t\,\mathbf{b}^T\mathbf{d}^{(k)}$$

$$\frac{d\Psi(t)}{dt} = \frac{1}{2}\,\mathbf{x}^{(k)T}\mathbf{A}\mathbf{d}^{(k)} + \frac{1}{2}\,\mathbf{d}^{(k)T}\mathbf{A}\mathbf{x}^{(k)} + t\,\mathbf{d}^{(k)T}\mathbf{A}\mathbf{d}^{(k)} - \mathbf{b}^T\mathbf{d}^{(k)}$$

Poznámka: První 2 členy, tj. $\underline{\mathbf{x}^{(k)T}\mathbf{A}\mathbf{d}^{(k)}}$ a $\underline{\mathbf{d}^{(k)T}\mathbf{A}\mathbf{x}^{(k)}}$ jsou skaláry a jsou si pro symetrickou matici $\mathbf{A}$ rovny.

$$\mathbf{d}^{(k)T}\mathbf{A}\mathbf{x}^{(k)} = (\mathbf{d}^{(k)T}\mathbf{A}\mathbf{x}^{(k)})^T = (\mathbf{A}\mathbf{x}^{(k)})^T\mathbf{d}^{(k)} = \mathbf{x}^{(k)T}\mathbf{A}^T\mathbf{d}^{(k)}$$

$$\frac{d\Psi(t)}{dt} = t\,\mathbf{d}^{(k)^T}\mathbf{A}\mathbf{d}^{(k)} + \underbrace{\mathbf{x}^{(k)^T}\mathbf{A}\mathbf{d}^{(k)} - \mathbf{b}^T\mathbf{d}^{(k)}}_{\underbrace{\left(\mathbf{x}^{(k)^T}\mathbf{A} - \mathbf{b}^T\right)}_{-\mathbf{d}^{(k)^T}}\mathbf{d}^{(k)}}$$

$$\frac{d\Psi(t)}{dt} = t\,\mathbf{d}^{(k)^T}\mathbf{A}\mathbf{d}^{(k)} + \underbrace{\mathbf{x}^{(k)^T}\mathbf{A}\mathbf{d}^{(k)} - \mathbf{b}^T\mathbf{d}^{(k)}}_{\underbrace{\left(\mathbf{x}^{(k)^T}\mathbf{A} - \mathbf{b}^T\right)}_{-\mathbf{d}^{(k)^T}}\mathbf{d}^{(k)}}$$

$$\frac{d\Psi(t)}{dt} = t\,\mathbf{d}^{(k)^T}\mathbf{A}\mathbf{d}^{(k)} - \mathbf{d}^{(k)^T}\mathbf{d}^{(k)} = 0$$

$$\boxed{t^{(k)} = \frac{\mathbf{d}^{(k)^T}\mathbf{d}^{(k)}}{\mathbf{d}^{(k)^T}\mathbf{A}\,\mathbf{d}^{(k)}}}$$

Poznámka:

Pokud by matice $\mathbf{A}$ nesplňovala podmínku symetrie, jaký výsledek by nám dala metoda největšího spádu?

$$\text{grad}\,F(\mathbf{x}) = \mathbf{0}$$

$$\text{grad}\,F(\mathbf{x}) = \text{grad}\left(\frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{b}^T\mathbf{x}\right) = \frac{1}{2}(\mathbf{x}^T\mathbf{A})^T + \frac{1}{2}\mathbf{A}\mathbf{x} - \mathbf{b} = \frac{1}{2}\mathbf{A}^T\mathbf{x} + \frac{1}{2}\mathbf{A}\mathbf{x} - \mathbf{b} = \mathbf{0}$$

$$\Rightarrow \quad \boxed{\frac{1}{2}(\mathbf{A}^T + \mathbf{A})\mathbf{x} = \mathbf{b}}$$

Algoritmus **metody největšího spádu**

1. volba $\mathbf{x}^{(0)}$, ε
2. výpočet směru spádu $\mathbf{d}^{(k)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k)}$
3. výpočet koeficientu $t^{(k)} = \dfrac{\mathbf{d}^{(k)^T}\mathbf{d}^{(k)}}{\mathbf{d}^{(k)^T}\mathbf{A}\,\mathbf{d}^{(k)}}$
4. výpočet nové iterace $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + t^{(k)}\mathbf{d}^{(k)}$
5. $k = k + 1$ a zpět na 2) pokud $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| > \varepsilon$

Poznámka: Abychom ušetřili operace násobení matice a vektoru, určíme $\mathbf{d}^{(k+1)}$ takto:

$$\mathbf{d}^{(k+1)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k+1)} = \mathbf{b} - \mathbf{A}(\mathbf{x}^{(k)} + t^{(k)}\mathbf{d}^{(k)}) = \mathbf{d}^{(k)} - t^{(k)}\underbrace{\mathbf{A}\mathbf{d}^{(k)}}_{(*)}$$

$(*)$ toto se počítalo v kroku 3 v předchozí iteraci

Věta: Metoda největšího spádu konverguje (pro symetrickou, pozitivně definitní matici $\mathbf{A}$) pro libovolnou volbu počáteční aproximace $\mathbf{x}^{(0)}$ k přesnému řešení soustavy $\mathbf{Ax} = \mathbf{b}$.

Důkaz:

Konvergenci dokážeme v normě $||.||_{\mathbf{A}} = \sqrt{\mathbf{x}^T\mathbf{Ax}}$ (tzv. energetická norma).

$||.||_{\mathbf{A}}$ je s euklidovskou normou $||.||_2$ ekvivalentní, tj. z toho již plyne i konvergence v $||.||_2 = \sqrt{\mathbf{x}^T\mathbf{x}}$

(Definice: X ... lineární prostor, $||.||_1$ a $||.||_2$ ... normy na X;

$||.||_1$ a $||.||_2$ jsou ekvivalentní, existují-li čísla $c, C > 0$: $\forall \mathbf{x} \in X \quad c||\mathbf{x}||_1 \leq ||\mathbf{x}||_2 \leq C||\mathbf{x}||_1$)

tj. má platit

$$c^2\mathbf{x}^T\mathbf{x} \quad \leq \quad \mathbf{x}^T\mathbf{Ax} \quad \leq \quad C^2\mathbf{x}^T\mathbf{x}$$

$$\mathbf{x}^T(c^2\mathbf{I})\mathbf{x} \quad \leq \quad \mathbf{x}^T\mathbf{Ax} \quad \leq \quad \mathbf{x}^T(C^2\mathbf{I})\mathbf{x}$$

platí pro $c = |\lambda_{min}|$, $C = |\lambda_{max}|$

$\mathbf{x}^*$ ... přesné řešení $\mathbf{Ax} = \mathbf{b}$

$\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}^*$ ... chyba k-té iterace

Odvoďme nejprve vztah pro energetickou normu chyby k-té iterace.

$$F(\mathbf{x}^{(k)}) - F(\mathbf{x}^*) = F(\mathbf{x}^* + \mathbf{e}^{(k)}) - F(\mathbf{x}^*) = \ \dots \qquad (*)$$

Obecně pro 2 body $\mathbf{x}$, $\mathbf{x} + t\mathbf{d}$ platí:

$$\mathbf{F}(\mathbf{x} + \mathbf{td}) - \mathbf{F}(\mathbf{x}) = \frac{1}{2}(\mathbf{x} + t\mathbf{d})^T\mathbf{A}(\mathbf{x} + t\mathbf{d}) - \mathbf{b}^T(\mathbf{x} + t\mathbf{d}) - \frac{1}{2}\mathbf{x}^T\mathbf{Ax} + \mathbf{b}^T\mathbf{x} =$$

$$= \underline{t\mathbf{x}^T\mathbf{Ad}} + \frac{1}{2}t^2\mathbf{d}^T\mathbf{Ad}\underline{- t\mathbf{b}^T\mathbf{d}} =$$

$$= \underline{\mathbf{td}^{\mathbf{T}}(\mathbf{Ax} - \mathbf{b})} + \frac{\mathbf{1}}{\mathbf{2}}\mathbf{t}^{\mathbf{2}}\mathbf{d}^{\mathbf{T}}\mathbf{Ad}$$

Pro náš případ $\mathbf{x} = \mathbf{x}^*$, $t = 1$, $\mathbf{d} = \mathbf{e}^{(k)}$

$$\dots = F(\mathbf{x}^* + \mathbf{e}^{(k)}) - F(\mathbf{x}^*) = \frac{1}{2}\mathbf{e}^{(k)^T}\mathbf{Ae}^{(k)} = ||\mathbf{e}^{(k)}||^2_{\mathbf{A}} \qquad (**)$$

$(*) + (**) \quad \Rightarrow \quad \boxed{F(\mathbf{x}^{(k)}) - F(\mathbf{x}^*) = \tfrac{1}{2}\mathbf{e}^{(k)^T}\mathbf{Ae}^{(k)}} \qquad (***)$

$(***) \quad \Rightarrow \quad F(\mathbf{x}^{(k+1)}) - F(\mathbf{x}^*) = \tfrac{1}{2}\mathbf{e}^{(k+1)^T}\mathbf{Ae}^{(k+1)} \qquad (****)$

Odečtením $(****)$ a $(***)$ dostaneme

$$\boxed{F(\mathbf{x}^{(k+1)}) - F(\mathbf{x}^{(k)}) = \frac{1}{2}\mathbf{e}^{(k+1)^T}\mathbf{A}\mathbf{e}^{(k+1)} - \frac{1}{2}\mathbf{e}^{(k)^T}\mathbf{A}\mathbf{e}^{(k)}}, \qquad (\clubsuit)$$

kde iterace $\mathbf{x}^{(k+1)}$ je vypočtena metodou největšího spádu, tj.

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + t^{(k)}\mathbf{r}^{(k)}.$$

Pro výraz na levé straně opět použijeme **zvýrazněný vztah** pro hodnoty $\mathbf{x} = \mathbf{x}^{(k)}$, $t = t^{(k)}$, $\mathbf{d} = \mathbf{r}^{(k)}$

$$F(\mathbf{x}^{(k)} + t^{(k)}\mathbf{r}^{(k)}) - F(\mathbf{x}^{(k)}) = t^{(k)}\mathbf{r}^{(k)^T}\big(\underbrace{\mathbf{A}\mathbf{x}^{(k)} - \mathbf{b}}_{-\mathbf{r}^{(k)}}\big) + \frac{1}{2}t^{(k)^2}\mathbf{r}^{(k)^T}\mathbf{A}\mathbf{r}^{(k)} =$$

$\Big($ $t^{(k)}$ jsme počítali podle vztahu $\boxed{t^{(k)} = \dfrac{\mathbf{r}^{(k)^T}\mathbf{r}^{(k)}}{\mathbf{r}^{(k)^T}\mathbf{A}\mathbf{r}^{(k)}}}$ $\Big)$

$$= -\frac{(\mathbf{r}^{(k)^T}\mathbf{r}^{(k)})^2}{\mathbf{r}^{(k)^T}\mathbf{A}\mathbf{r}^{(k)}} + \frac{1}{2}\frac{(\mathbf{r}^{(k)^T}\mathbf{r}^{(k)})^2}{(\mathbf{r}^{(k)^T}\mathbf{A}\mathbf{r}^{(k)})^2}\mathbf{r}^{(k)^T}\mathbf{A}\mathbf{r}^{(k)} = -\frac{1}{2}\frac{(\mathbf{r}^{(k)^T}\mathbf{r}^{(k)})^2}{\mathbf{r}^{(k)^T}\mathbf{A}\mathbf{r}^{(k)}} \qquad (\spadesuit)$$

Porovnáním ($\spadesuit$) a ($\clubsuit$) dostaneme

$$\boxed{-\frac{1}{2}\frac{(\mathbf{r}^{(k)^T}\mathbf{r}^{(k)})^2}{\mathbf{r}^{(k)^T}\mathbf{A}\mathbf{r}^{(k)}} = \frac{1}{2}\mathbf{e}^{(k+1)^T}\mathbf{A}\mathbf{e}^{(k+1)} - \frac{1}{2}\mathbf{e}^{(k)^T}\mathbf{A}\mathbf{e}^{(k)}}. \qquad (\heartsuit)$$

Nyní poslední rovnici
 a) vynásobíme 2
 b) poslední člen převedeme na druhou stranu
 c) a vydělíme jím rovnici

Dostaneme

$$\boxed{\frac{\mathbf{e}^{(k+1)^T}\mathbf{A}\mathbf{e}^{(k+1)}}{\mathbf{e}^{(k)^T}\mathbf{A}\mathbf{e}^{(k)}} = 1 - \frac{(\mathbf{r}^{(k)^T}\mathbf{r}^{(k)})^2}{\mathbf{r}^{(k)^T}\mathbf{A}\mathbf{r}^{(k)}\ \mathbf{e}^{(k)^T}\underbrace{\mathbf{A}\mathbf{e}^{(k)}}_{=\mathbf{r}^{(k)}}}} \qquad (\diamondsuit)$$

Platí $\boxed{\mathbf{A}\mathbf{e}^{(k)} = \mathbf{r}^{(k)}}$, protože $\mathbf{A}\mathbf{x}^* - \mathbf{b} = \mathbf{0}$ a $\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k)}$

$$\mathbf{r}^{(k)} = \mathbf{b}\ \underbrace{-\mathbf{A}\mathbf{x}^{(k)} + \mathbf{A}\mathbf{x}^*}_{\mathbf{A}(\underbrace{\mathbf{x}^* - \mathbf{x}^{(k)}}_{\mathbf{e}^{(k)}})} - \mathbf{b}$$

Dále z $\boxed{\mathbf{A}\mathbf{e}^{(k)} = \mathbf{r}^{(k)}}$ plyne $\mathbf{e}^{(k)} = \mathbf{A}^{-1}\mathbf{r}^{(k)}$ a tedy odhad

$$||\mathbf{e}^{(k)}|| \le ||\mathbf{A}^{-1}|| \cdot ||\mathbf{r}^{(k)}||$$

Pro ($\diamondsuit$) dostáváme odhad:

$$\frac{\mathbf{e}^{(k+1)^T}\mathbf{A}\mathbf{e}^{(k+1)}}{\mathbf{e}^{(k)^T}\mathbf{A}\mathbf{e}^{(k)}} \le 1 - \frac{||\mathbf{r}^{(k)}||^4}{||\mathbf{A}|| \cdot ||\mathbf{r}^{(k)}||^2 \cdot ||\mathbf{A}^{-1}|| \cdot ||\mathbf{r}^{(k)}||^2} = 1 - \frac{1}{||\mathbf{A}|| \cdot ||\mathbf{A}^{-1}||} = q < 1$$

Tj.

$$\boxed{\mathbf{e}^{(k+1)^T}\mathbf{A}\mathbf{e}^{(k+1)} \le q\ \mathbf{e}^{(k)^T}\mathbf{A}\mathbf{e}^{(k)} \quad \forall k}. \qquad (\blacksquare)$$

$$\mathbf{e}^{(k+1)^T}\mathbf{A}\mathbf{e}^{(k+1)} \le q\ \mathbf{e}^{(k)^T}\mathbf{A}\mathbf{e}^{(k)} \quad \forall k$$

Důkaz:

Jde o to odhadnout q v (■).

$$\|\mathbf{e}^{(k+1)}\|_{\mathbf{A}}^2 \;\leq\; q\,\|\mathbf{e}^{(k)}\|_{\mathbf{A}}^2 \qquad \Rightarrow \qquad \|\mathbf{e}^{(k)}\|_{\mathbf{A}}^2 \;\leq\; q^k\,\|\mathbf{e}^{(0)}\|_{\mathbf{A}}^2$$

$$q = 1 - \frac{1}{\underbrace{\|\mathbf{A}\|}_{=\lambda_{max}} \cdot \underbrace{\|\mathbf{A}^{-1}\|}_{=\frac{1}{\lambda_{min}}}} = 1 - \frac{1}{\varkappa(\mathbf{A})} = \boxed{\frac{\varkappa(\mathbf{A})-1}{\varkappa(\mathbf{A})} \leq \sqrt{\frac{\varkappa(\mathbf{A})-1}{\varkappa(\mathbf{A})+1}}}$$

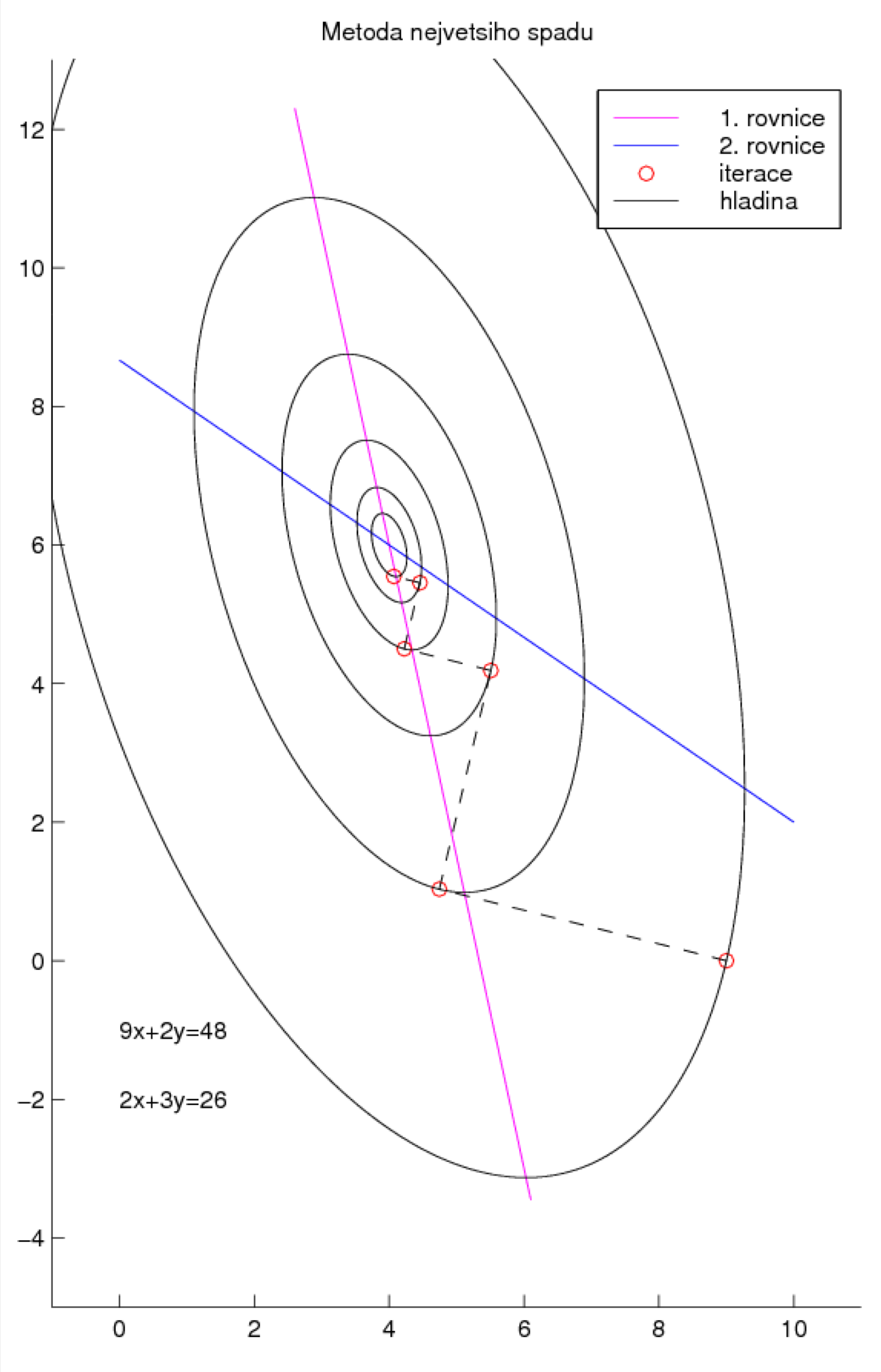
Důkaz poslední nerovnosti:

$$\frac{(\varkappa(\mathbf{A})-1)^2}{\varkappa^2(\mathbf{A})} \leq \frac{\varkappa(\mathbf{A})-1}{\varkappa(\mathbf{A})+1} \qquad \Big/ : (\varkappa(\mathbf{A})-1)$$

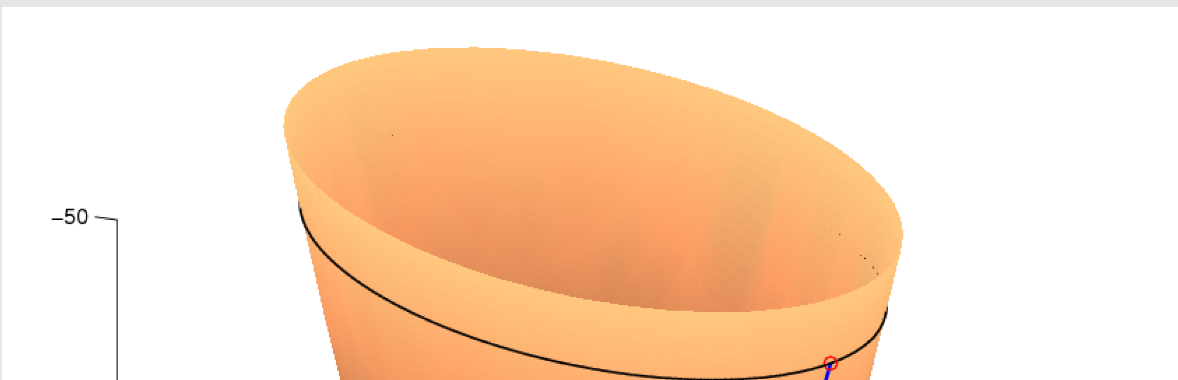
$$\frac{\varkappa(\mathbf{A})-1}{\varkappa^2(\mathbf{A})} \leq \frac{1}{\varkappa(\mathbf{A})+1} \qquad \Big/ \cdot \varkappa^2(\mathbf{A})(\varkappa(\mathbf{A})+1)$$

$$\varkappa^2(\mathbf{A}) - 1 \leq \varkappa^2(\mathbf{A}) \qquad \ldots\; OK$$

Geometrický význam metody největšího spádu



k	$x^{(k)}$	$y^{(k)}$
0	9.0000	0
1	4.7425	1.0321
2	5.5081	4.1902
3	4.2240	4.5015
4	4.4549	5.4541
5	4.0676	5.5480



Vlastnost reziduí

Všimněme si faktu, že vždy po sobě jdoucí iterace směru spádu, tj. $\mathbf{d}^{(k)}$ a $\mathbf{d}^{(k+1)}$ jsou na sebe kolmé.

Cvičení: Ukažte, že platí $\boxed{\mathbf{d}^{(k)^T}\,\mathbf{d}^{(k+1)} = 0}$.

$$
\begin{aligned}
\mathbf{d}^{(k)^T}\,\mathbf{d}^{(k+1)} &= \mathbf{d}^{(k)^T}\,(\mathbf{b} - \mathbf{A}\mathbf{x}^{(k+1)}) = \\
&= \mathbf{d}^{(k)^T}\,(\mathbf{b} - \mathbf{A}(\mathbf{x}^{(k)} + t^{(k)}\mathbf{d}^{(k)})) = \\
&= \mathbf{d}^{(k)^T}\,(\mathbf{b} - \mathbf{A}\mathbf{x}^{(k)} - t^{(k)}\mathbf{A}\,\mathbf{d}^{(k)}) = \\
&= \mathbf{d}^{(k)^T}\,(\mathbf{d}^{(k)} - t^{(k)}\mathbf{A}\,\mathbf{d}^{(k)}) = \\
&= \mathbf{d}^{(k)^T}\mathbf{d}^{(k)} - t^{(k)}\mathbf{d}^{(k)^T}\mathbf{A}\,\mathbf{d}^{(k)} = \\
&= \mathbf{d}^{(k)^T}\mathbf{d}^{(k)} - \frac{\mathbf{d}^{(k)^T}\,\mathbf{d}^{(k)}}{\mathbf{d}^{(k)^T}\mathbf{A}\,\mathbf{d}^{(k)}}\,\mathbf{d}^{(k)^T}\mathbf{A}\,\mathbf{d}^{(k)} = \\
&= \mathbf{d}^{(k)^T}\mathbf{d}^{(k)} - \mathbf{d}^{(k)^T}\mathbf{d}^{(k)} = 0
\end{aligned}
$$

Poznámka:

V případě, že budou hladiny (elipsy) „velmi protáhlé“, bude obecně metoda největšího spádu konvergovat velmi pomalu, nastane tzv. **cik-cak efekt**.

Na druhou stranu, pokud budou hladiny (elipsy) „skoro kružnice“, bude metoda největšího spádu konvergovat velmi rychle.

Nevýhodu cik-cak efektu odstraní nová metoda, tzv. **metoda sdružených gradientů**, která využívá důmyslnější volby směrů minimalizace, a sice tak, aby se neopakovali, jak k tomu docházelo u metody největšího spádu.

Příklad 1 - pokračování

Uvažovali jsme jednoduchou soustavu $\mathbf{A}\mathbf{x} = \mathbf{b}$, kde

$$
\mathbf{A} = \begin{bmatrix} 25 & 0 \\ 0 & 16 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 25 \\ 8 \end{bmatrix}, \quad \mathbf{x}^* = \begin{bmatrix} 1 \\ 0,5 \end{bmatrix}.
$$

Jedna z vrstevnic měla tvar

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

poměr poloos:

$$
\sqrt{\lambda_2} = \sqrt{16} \to 4 : 5 \leftarrow \sqrt{25} = \sqrt{\lambda_1}
$$

$$
\boxed{\sqrt{\lambda_2} : \sqrt{\lambda_1}}
$$

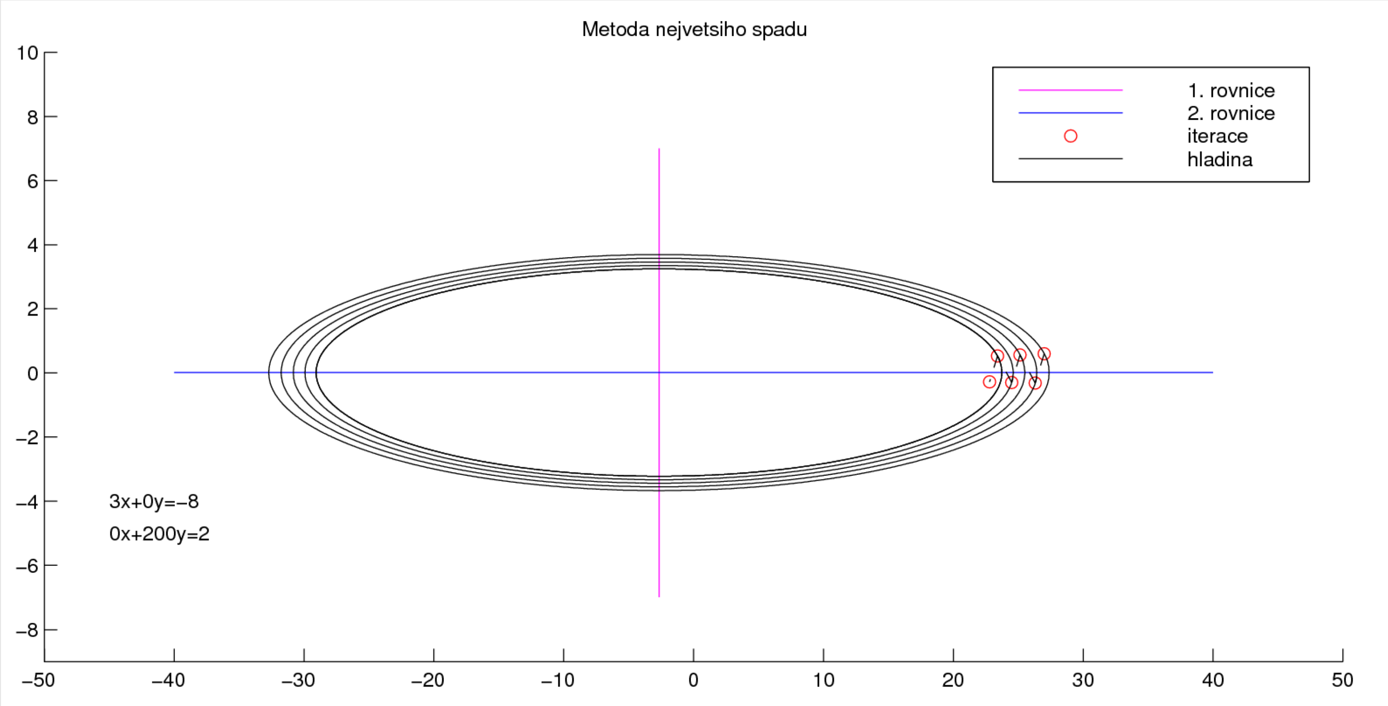
Poznámka:

- Pro případ $\lambda_2 \gg \lambda_1$ získáme protáhlé elipsy
- Pro případ $\lambda_2 \approx \lambda_1$ získáme skoro kružnice

Příklad 2

Pomocí metody největšího spádu řešte soustavu $\mathbf{A}\mathbf{x} = \mathbf{b}$, kde

$$\mathbf{A} = \begin{bmatrix} 3 & 0 \\ 0 & 200 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -8 \\ 2 \end{bmatrix}, \quad \text{počáteční iterace } \mathbf{x}^{(0)} = \begin{bmatrix} 27 \\ 0{,}6 \end{bmatrix}.$$



k	$x^{(k)}$	$y^{(k)}$
0	27.000000	0.600000
1	26.307758	-0.317804
2	25.139940	0.563008
3	24.491101	-0.297251
4	23.396504	0.528335
5	22.788346	-0.277987
⋮	⋮	⋮
250	-2.657606	0.010180

vlastní čísla matice $\mathbf{A}$:

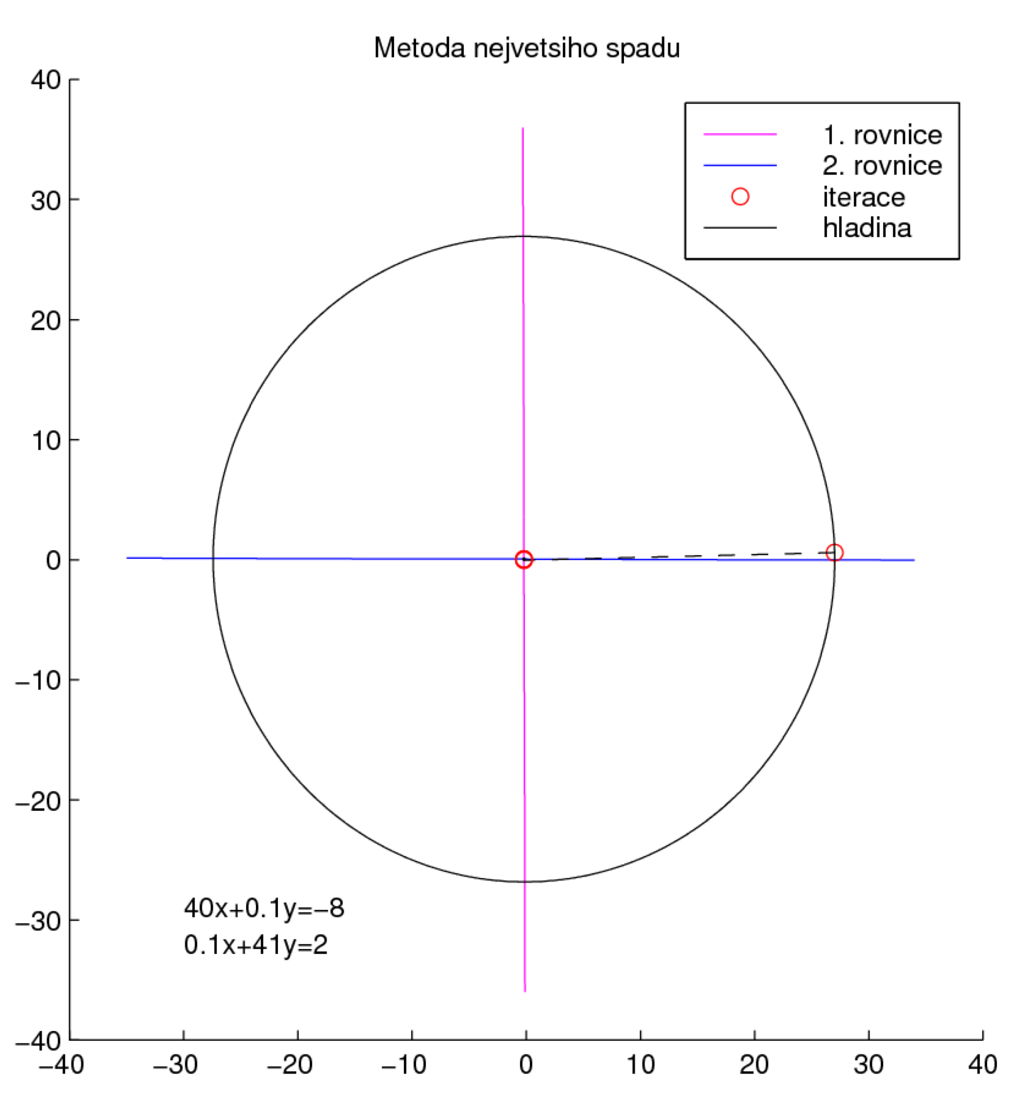
$$\lambda_1 = 3 \quad \text{a} \quad \lambda_2 = 200$$

přesné řešení soustavy je $\quad \mathbf{x}^* = \begin{bmatrix} -\frac{8}{3} \\ \frac{1}{100} \end{bmatrix}$

Příklad 3

Pomocí metody největšího spádu řešte soustavu $\mathbf{A}\mathbf{x} = \mathbf{b}$, kde

$$\mathbf{A} = \begin{bmatrix} 40 & 0{,}1 \\ 0{,}1 & 41 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -8 \\ 2 \end{bmatrix}, \quad \text{počáteční iterace } \mathbf{x}^{(0)} = \begin{bmatrix} 27 \\ 0{,}6 \end{bmatrix}.$$



k	$x^{(k)}$	$y^{(k)}$
0	27.000000	0.600000
1	-0.197972	-0.032418
2	-0.199872	0.049274
3	-0.200123	0.049268

vlastní čísla matice $\mathbf{A}$:

$$\lambda_1 \doteq 39{,}99 \quad \text{a} \quad \lambda_2 \doteq 41{,}01$$

přesné řešení soustavy $\mathbf{x}^*$ se s $\mathbf{x}^{(3)}$ shoduje na 6 desetiných míst

Poznámky k rychlosti konvergence:

$$\boxed{\|\mathbf{x}^{(k)} - \mathbf{x}^*\|_{\mathbf{A}} = \left(\frac{\varkappa(\mathbf{A}) - 1}{\varkappa(\mathbf{A}) + 1}\right)^k \|\mathbf{x}^{(0)} - \mathbf{x}^*\|_{\mathbf{A}}}$$

- Je-li $\varkappa(\mathbf{A}) \gg 1$, tj. $\lambda_{max} \gg \lambda_{min}$, pak metoda největšího spádu konverguje pomalu

$$\frac{\varkappa(\mathbf{A}) - 1}{\varkappa(\mathbf{A}) + 1} = \frac{\varkappa(A) - 1 + 1 - 1}{\varkappa(A) + 1} = 1 - \frac{2}{\underbrace{\varkappa(A) + 1}_{\to\infty \text{ pro } \varkappa(A)\to\infty}} \lessapprox 1$$

- Je-li $\varkappa(\mathbf{A}) \gtrapprox 1$, tj. $\lambda_{max} \approx \lambda_{min}$, pak metoda největšího spádu konverguje rychle

$$\frac{\varkappa(\mathbf{A})-1}{\varkappa(\mathbf{A})+1} = 1 - \frac{2}{\underbrace{\varkappa(\mathbf{A})+1}_{\approx 2}} \approx 0$$

- Pokud jsou vrstevnice sféry (v $\mathbb{R}^2$ kružnice), potom metoda největšího spádu nalezne řešení (přesné) v jednom kroku.

Poznámka:

Směr, ve kterém provádíme minimalizaci v rámci jednoho kroku metody, můžeme volit i jinak než směr největšího spádu.

Obecně označme používané směry $\mathbf{s}^{(k)}$.

Novou iteraci hledáme ve tvaru

$$\boxed{\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + t\,\mathbf{s}^{(k)}}$$

Koeficient t získáme z jednorozměrné minimalizace

$$\boxed{\min_{t>0} \underbrace{F(\mathbf{x}^{(k)} + t\,\mathbf{s}^{(k)})}_{\Phi(t)}}$$

$$\Phi(t) = \frac{1}{2}(\mathbf{x}^{(k)} + t\,\mathbf{s}^{(k)})^T\mathbf{A}(\mathbf{x}^{(k)} + t\,\mathbf{s}^{(k)}) - \mathbf{b}^T(\mathbf{x}^{(k)} + t\,\mathbf{s}^{(k)})$$

$$\frac{d\Phi(t)}{dt} = t\mathbf{s}^{(k)^T}\mathbf{A}\mathbf{s}^{(k)} + \underbrace{\mathbf{x}^{(k)^T}\mathbf{A}\mathbf{s}^{(k)} - \mathbf{b}^T\mathbf{s}^{(k)}}_{\underbrace{(\mathbf{x}^{(k)^T}\mathbf{A} - \mathbf{b}^T)}_{\underbrace{(\mathbf{A}\mathbf{x}^{(k)} - \mathbf{b})^T}_{= -\mathbf{r}^{(k)}\ \text{(reziduum)}}}\ \mathbf{s}^{(k)}} = 0$$

$$\boxed{t^{(k)} = \frac{\mathbf{r}^{(k)^T}\mathbf{s}^{(k)}}{\mathbf{s}^{(k)^T}\mathbf{A}\mathbf{s}^{(k)}}}$$

Volíme-li za vektory $\mathbf{s}^{(k)}$ postupně jednotkové vektory souřadných os, získáme Gauss-Seidelovu metodu !!!

Pokud na vektor $\mathbf{x}$ aplikujeme 1 iteraci gradientní metody se směrovým vektorem $\mathbf{e}_i = [0, \ldots, 0, 1, 0, \ldots 0]^T$ (na i-té pozici 1, jinak 0), dostaneme:

$$\mathbf{x} := \mathbf{x} + \frac{\mathbf{r}^T\mathbf{e}_i}{\mathbf{e}_i^T\mathbf{A}\mathbf{e}_i}\,\mathbf{e}_i. \qquad (\bullet)$$

Platí: $\mathbf{e}_i^T\mathbf{A}$ ... i-tý řádek matice $\mathbf{A}$

$\mathbf{e}_i^T\mathbf{A}\mathbf{e}_i$ ... diagonální prvek a_{ii} matice $\mathbf{A}$

$\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}$ ... reziduum

$\mathbf{r}^T\mathbf{e}_i = r_i$ ... i-tá složka vektoru $\mathbf{r}$

$r_i = b_i - \sum_{j=1}^{n} a_{ij}x_j$

Vztah $(\bullet)$ zvětší i-tou složku vektoru $\mathbf{x}$ o hodnotu r_i, tj.

$$x_i := x_i + \underbrace{\frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{n} a_{ij}x_j\right)}_{=r_i},$$

$$x_i := \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^{n} a_{ij}x_j\right).$$

Script v MATLABu

```
function [vysledky_gs,vysledky_gm]=gs_gm(A,b,x0,iteraci);

%************************************************************
%           Porovnani Gauss-Seidelovy metody a
%         gradientni metody, kde za smery volime
%            jednotkove vektory souradnych os
%************************************************************

n=size(A,1);

%************************************************************
%  Gauss-Seidelova metoda
%************************************************************

x=x0;
vysledky_gs=x';
D=diag(diag(A));  L=tril(A)-D;  U=triu(A)-D;
H=-(L+D)\U;
g=(L+D)\b;

for i=1:iteraci
  x=H*x+g;
  vysledky_gs=[vysledky_gs;x'];
end

%************************************************************
%  Gradientni metoda
%************************************************************

x=x0;
vysledky_gm=x';

for i=1:iteraci
   for j=1:n
     s=zeros(n,1);
     s(j)=1;
     r=-A*x+b;
     t=(r'*s)/(s'*A*s);
     x=x+t*s;
   end;
   vysledky_gm=[vysledky_gm;x'];
end
```

Úvaha

Při vhodné volbě směrových vektorů $\mathbf{s}^{(k)}$ je možné dojít do přesného řešení za konečný počet kroků $\leq n$. Musí existovat n vektorů $\mathbf{s}^{(k)}$ tak, že

$$\boxed{\mathbf{x}^* - \mathbf{x}^{(0)} = \sum_{k=1}^{n} t^{(k)} \mathbf{s}^{(k)}} \qquad (*)$$

Jak volit směry $\mathbf{s}^{(k)}$?

- Zkusíme takto: nechť $\mathbf{s}^{(k)}$ tvoří bázi (ortogonální) n-rozměrného euklidovského prostoru, potom vynásobením $(*)$ skalárně s $\mathbf{s}^{(k)}$ a úpravou získáme

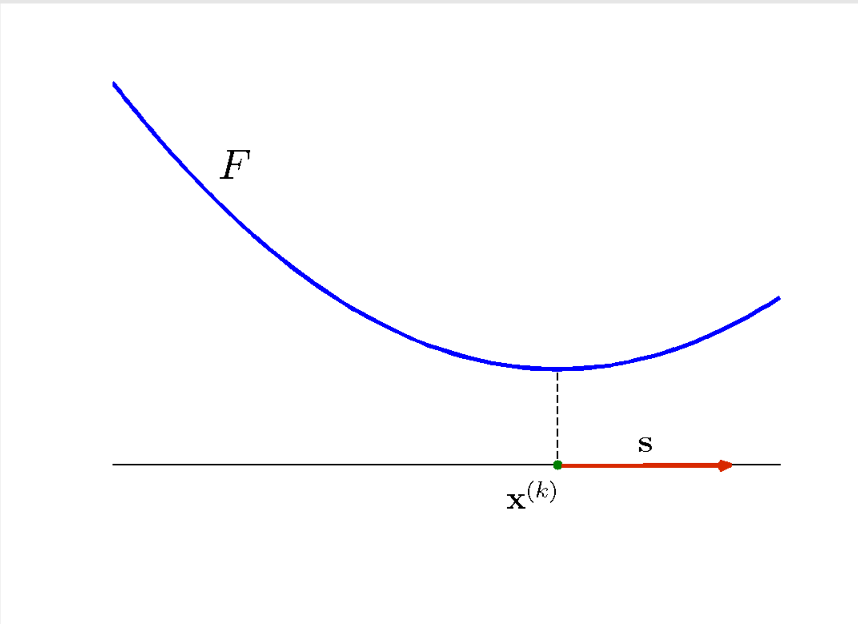
$$\mathbf{s}^{(k)^T}(\mathbf{x}^* - \mathbf{x}^{(0)}) = t^{(k)} \mathbf{s}^{(k)^T} \mathbf{s}^{(k)}$$

$$t^{(k)} = \frac{\mathbf{s}^{(k)^T}(\mathbf{x}^* - \mathbf{x}^{(0)})}{\mathbf{s}^{(k)^T}\mathbf{s}^{(k)}} \qquad \ldots \text{ nešikovné !, obsahuje přesné řešení}$$

- je třeba zvolit lepší strategii volby vektorů $\mathbf{s}^{(k)}$

Definice $\mathbf{x}^{(k)}$ je **optimální vzhledem ke směru** $\mathbf{s} \neq \mathbf{0}$, jestliže

$$\boxed{F(\mathbf{x}^{(k)}) \leq F(\mathbf{x}^{(k)} + t\mathbf{s}) \quad \forall t \in \mathbb{R}} \qquad (\star)$$



Poznámka: Je-li $\mathbf{x}^{(k)}$ optimální vzhledem k libovolnému směru z vektorového prostoru V, říkáme, že je $\mathbf{x}^{(k)}$ **optimální vzhledem k** V.

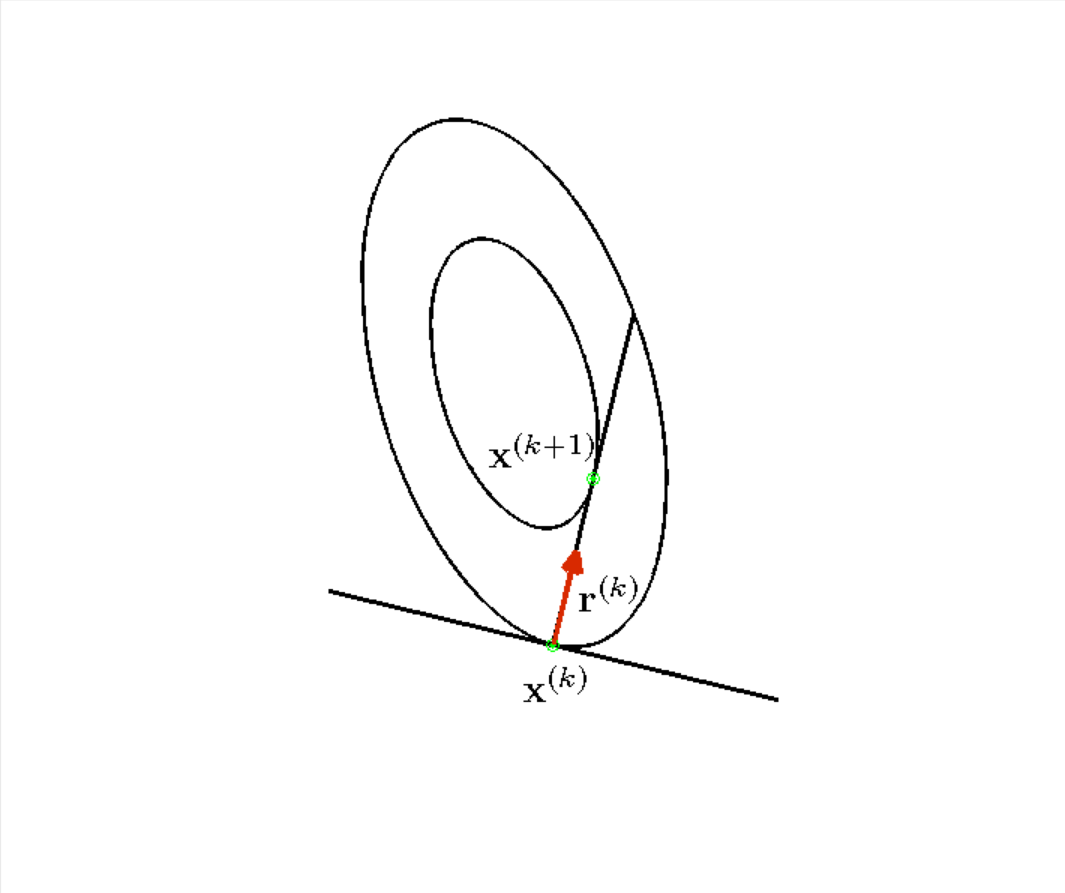
Podle $(\star)$ se minima nabývá pro $t = 0$, tzn. že derivace F podle t je v minimu ($t = 0$) rovna 0:

$$\frac{\partial F}{\partial t}\left(\mathbf{x}^{(k)} + t\,\mathbf{s}\right) = t\,\mathbf{s}^T\mathbf{A}\mathbf{s} + \mathbf{s}^T\left(\mathbf{A}\mathbf{x}^{(k)} - \mathbf{b}\right)$$

$$\frac{\partial F}{\partial t}\left(\mathbf{x}^{(k)}\right) = \mathbf{s}^T \underbrace{\left(\mathbf{A}\mathbf{x}^{(k)} - \mathbf{b}\right)}_{=\mathbf{r}^{(k)}} = 0$$

$$\boxed{\mathbf{s} \perp \mathbf{r}^{(k)}}$$

Poznámka: Iterace $\mathbf{x}^{(k+1)}$ metody největšího spádu je optimální vzhledem k reziduu $\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k)}$ ... směry, ve kterých minimalizujeme.



Naším cílem je, aby se i v dalších iteracích zachovávala optimalita k již použitým směrům.

To pro metodu největšího spádu bohužel neplatí.

Např. pro soustavu ve $2D$ jsme ukazovali, že směry největšího spádu (reziduí) jsou na sebe kolmé,

$$\text{tj.} \quad \mathbf{r}^{(k)} \perp \mathbf{r}^{(k+1)} \quad \text{a} \quad \mathbf{r}^{(k+1)} \perp \mathbf{r}^{(k+2)} \quad \Rightarrow \quad \boxed{\mathbf{r}^{(k)} \parallel \mathbf{r}^{(k+2)}} \;!!!$$

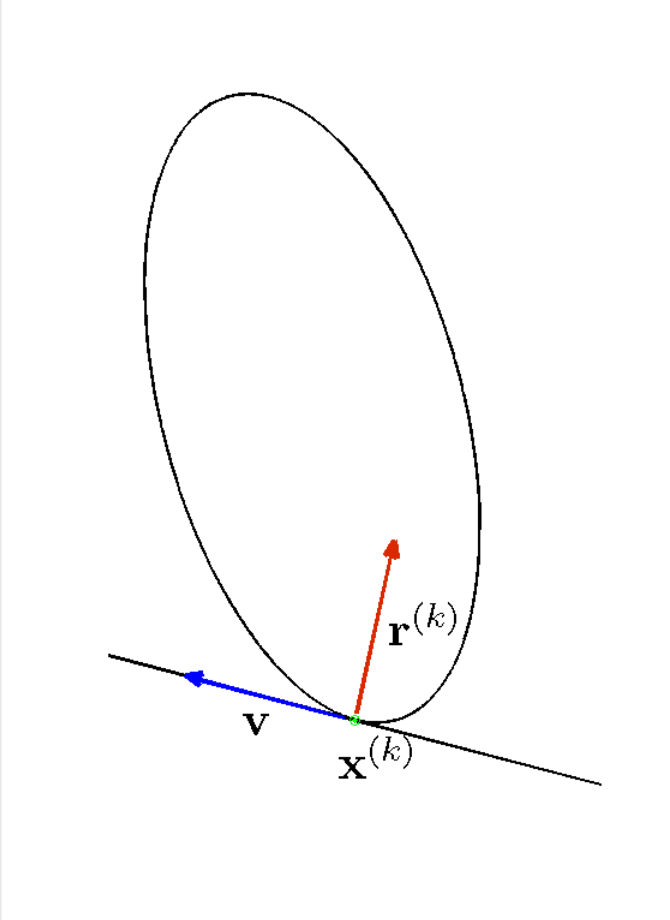
$\Rightarrow$ $\mathbf{x}^{(k+2)}$ je optimální vzhledem k $\mathbf{r}^{(k+1)}$, ale již není optimální vzhledem k $\mathbf{r}^{(k)}$

Existují směry, které udržují optimalitu k předchozím?

Nechť

$$\boxed{\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{s}}.$$

Předpokládejme, že $\mathbf{x}^{(k)}$ ke optimální vzhledem k $\mathbf{v}$ (tj. $\mathbf{r}^{(k)} \perp \mathbf{v}$).



Chceme-li, aby bylo i $\mathbf{x}^{(k+1)}$ optimální vzhledem k $\mathbf{v}$, (tj. $\mathbf{r}^{(k+1)} \perp \mathbf{v}$), musí platit:

$$0 = \mathbf{v}^T\mathbf{r}^{(k+1)} = \mathbf{v}^T(\mathbf{b} - \mathbf{A}\mathbf{x}^{(k+1)}) = \mathbf{v}^T\left(\mathbf{b} - \mathbf{A}(\mathbf{x}^{(k)} + \mathbf{s})\right) = \mathbf{v}^T\left(\underbrace{\mathbf{b} - \mathbf{A}\mathbf{x}^{(k)}}_{\mathbf{r}^{(k)}} - \mathbf{A}\mathbf{s}\right) = \mathbf{v}^T\left(\mathbf{r}^{(k)} - \mathbf{A}\mathbf{s}\right) = -\mathbf{v}^T\mathbf{A}\mathbf{s}$$

Závěr

Chceme-li zachovat optimalitu vzhledem ke všem použitým směrům, musí tyto směry splňovat podmínky tzv. A-**ortogonality**, tj. pro 2 různé směry $\mathbf{s}$ a $\mathbf{v}$ musí platit:

$$\boxed{\mathbf{v}^T\mathbf{A}\mathbf{s} = \mathbf{0}}.$$

Poznámka: Vektorům které jsou A-**ortogonální** se také říká A-**sdružené**.

Metoda sdružených gradientů

Za směry, ve kterých minimalizujeme, budeme brát A-ortogonální vektory $\mathbf{s}^{(k)}$.

Platí tedy:

$$\boxed{{\mathbf{s}^{(k)}}^T\mathbf{A}\,\mathbf{s}^{(l)} = 0, \quad k \neq l}.$$

Chceme, aby platilo:

$$ {\mathbf{s}^{(k)}}^T\mathbf{A}\cdot / \qquad \boxed{\mathbf{x}^* - \mathbf{x}^{(0)} = \sum_{k=1}^{n} t^{(k)}\mathbf{s}^{(k)}} \qquad (*)$$

$$\mathbf{s}^{(k)T} \underbrace{\mathbf{A}(\mathbf{x}^* - \mathbf{x}^{(0)})} = t^{(k)} \mathbf{s}^{(k)T} \mathbf{A}\, \mathbf{s}^{(k)}$$

$$\mathbf{A}\mathbf{x}^* - \mathbf{A}\mathbf{x}^{(0)} = \underbrace{\mathbf{A}\mathbf{x}^* - \mathbf{b}}_{=\mathbf{0}} \underbrace{-\mathbf{A}\mathbf{x}^{(0)} + \mathbf{b}}_{\mathbf{r}^{(0)}}$$

$$\boxed{t^{(k)} = \frac{\mathbf{s}^{(k)T}\mathbf{r}^{(0)}}{\mathbf{s}^{(k)T}\mathbf{A}\mathbf{s}^{(k)}}}$$

Strategie volby směrů

- Máme-li ortogonální bázi $\mathbb{R}^n$, lze z ní procesem A-ortogonalizace získat A-ortogonální bázi.
- Za ortogonální bázi budeme volit reziduové vektory.

  Aby proces ortogonalizace vedl k cíli, musíme zaručit, že reziduové vektory tvoří bázi.

  Ortogonalitu ukážeme vzápětí; může se stát, že se některé reziduum anuluje.

  Potom ovšem iterační proces končí - dosáhli jsme přesného řešení.
- Provádíme tedy současně 2 procesy!
  - iterační proces
  - proces A-ortogonalizace
- Vektory reziduí budeme značit $\mathbf{r}^{(k)}$, získané sdružené směry označíme $\mathbf{s}^{(k)}$
  - pro zadané $\mathbf{x}^{(0)}$ určíme $\mathbf{r}^{(0)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(0)}$
  - $\mathbf{s}^{(0)}$ položíme rovno $\mathbf{r}^{(0)}$
  - určíme $\mathbf{x}^{(1)}$ optimální vzhledem k $\mathbf{s}^{(0)}$
  - určíme $\mathbf{r}^{(1)}$
  - $\mathbf{s}^{(1)}$ určujeme z $\mathbf{r}^{(1)}$ tak, aby $\mathbf{s}^{(1)T}\mathbf{A}\mathbf{s}^{(0)} = 0$
  - atd.

Proces A-ortogonalizace

$$\boxed{\mathbf{s}^{(k)} = \mathbf{r}^{(k)} + \sum_{i=1}^{k-1} \beta_{ki}\, \mathbf{s}^{(i)}} \qquad (\bullet\bullet)$$

(Při určení $\mathbf{s}^{(k)}$ vyjdeme z $\mathbf{r}^{(k)}$. Přičítáme násobky předchozích $\mathbf{s}^{(i)}$ tak, abychom zaručili A-ortogonalitu.)

Koeficienty β_{ki} volíme tak, aby

$$\boxed{\mathbf{s}^{(k)}\mathbf{A}\mathbf{s}^{(i)} = 0, \quad (i < k)}$$

($\bullet\bullet$) vynásobíme $\quad \mathbf{s}^{(i)T}\mathbf{A} \cdot \Big/$

$$\underbrace{\mathbf{s}^{(i)T}\mathbf{A}\mathbf{s}^{(k)}}_{=\,0} = \mathbf{s}^{(i)T}\mathbf{A}\mathbf{r}^{(k)} + \beta_{ki}\, \mathbf{s}^{(i)T}\mathbf{A}\mathbf{s}^{(i)}$$

$$\Rightarrow \quad \boxed{\beta_{ki} = -\frac{\mathbf{s}^{(i)T}\mathbf{A}\mathbf{r}^{(k)}}{\mathbf{s}^{(i)T}\mathbf{A}\mathbf{s}^{(i)}}}$$

Z vlastností A-ortogonality vyplývá řada skutečností.

Věta 1 Platí

$$\boxed{\begin{array}{rcll} \mathbf{r}^{(k)^T}\cdot\mathbf{s}^{(j)} &=& \mathbf{r}^{(0)^T}\cdot\mathbf{s}^{(j)} & \qquad k\le j \\ \mathbf{r}^{(k)^T}\cdot\mathbf{s}^{(j)} &=& 0 & \qquad k>j \end{array}}$$

Důkaz:

$$-\mathbf{b}\cdot\Big/ \quad \mathbf{A}\cdot\Big/ \qquad \mathbf{x}^{(k+1)}=\mathbf{x}^{(k)}+t^{(k)}\mathbf{s}^{(k)}$$

$$\Rightarrow \qquad -\mathbf{r}^{(k+1)}=-\mathbf{r}^{(k)}+t^{(k)}\mathbf{A}\,\mathbf{s}^{(k)}$$

$$\Rightarrow \qquad \mathbf{r}^{(k)}=\mathbf{r}^{(0)}-\sum_{j=1}^{k-1}t^{(j)}\mathbf{A}\,\mathbf{s}^{(j)}$$

vynásobíme skalárně s $\mathbf{s}^{(j)}$

$$\boxed{\mathbf{r}^{(k)^T}\mathbf{s}^{(j)}=\mathbf{r}^{(0)^T}\mathbf{s}^{(j)}-\underbrace{t^{(j)}}_{(*)}\ \mathbf{s}^{(j)^T}\mathbf{A}\,\mathbf{s}^{(j)}}$$

$(*)$ počítáme (viz dříve) takto

$$t^{(j)}=\frac{\mathbf{r}^{(0)^T}\mathbf{s}^{(j)}}{\mathbf{s}^{(j)^T}\mathbf{A}\mathbf{s}^{(j)}}$$

□

Důkaz:

vztah $(\bullet\bullet)$ vynásobíme skalárně s $\mathbf{A}\mathbf{s}^{(j)}$

$$\mathbf{s}^{(k)}=\mathbf{r}^{(k)}+\sum_{i=1}^{k-1}\beta_{ki}\,\mathbf{s}^{(i)} \qquad (\bullet\bullet)$$

$$\underbrace{\mathbf{s}^{(j)^T}\mathbf{A}\,\mathbf{s}^{(k)}}_{\substack{=0\ (k<j)\\ \neq 0\ (k=j)}}=\mathbf{s}^{(j)^T}\mathbf{A}\,\mathbf{r}^{(k)}+\sum_{i=1}^{k-1}\beta_{ki}\ \underbrace{\mathbf{s}^{(j)^T}\mathbf{A}\,\mathbf{s}^{(i)}}_{=0\ (k\le j)}$$

□

Důkaz:

Úplnou matematickou indukcí ukážeme, že $\boxed{\mathbf{r}^{(j)^T}\mathbf{r}^{(k)}=0 \quad \text{pro} \quad j>k}$.

Platí

$$\mathbf{r}^{(k+1)}=\mathbf{b}-\mathbf{A}\mathbf{x}^{(k+1)}=\mathbf{b}-\mathbf{A}(\mathbf{x}^{(k)}+t^{(k)}\mathbf{r}^{(k)})=\mathbf{r}^{(k)}-t^{(k)}\mathbf{A}\mathbf{r}^{(k)}$$

1. $\underline{\underline{j=1}} \Rightarrow \quad k=0$

$$\mathbf{r}^{(1)^T}\mathbf{r}^{(0)}=\left(\mathbf{r}^{(0)}+t^{(0)}\mathbf{A}\,\mathbf{s}^{(0)}\right)^T\mathbf{r}^{(0)}=\mathbf{r}^{(0)^T}\mathbf{r}^{(0)}+t^{(0)}\mathbf{s}^{(0)^T}\mathbf{A}\,\mathbf{r}^{(0)}=\mathbf{r}^{(0)^T}\mathbf{r}^{(0)}-\frac{\mathbf{r}^{(0)^T}\mathbf{s}^{(0)}}{\mathbf{s}^{(0)^T}\mathbf{A}\,\mathbf{s}^{(0)}}\ \mathbf{s}^{(0)^T}\mathbf{A}\,\mathbf{r}^{(0)}=0$$

$(\mathbf{r}^{(0)}=\mathbf{s}^{(0)})$

2. a) $\forall k < j$ platí: $\boxed{\mathbf{r}^{(j+1)^T}\mathbf{r}^{(k)} = 0}$

$$\mathbf{r}^{(j+1)^T}\mathbf{r}^{(k)} = \left(\mathbf{r}^{(j)} + t^{(j)}\mathbf{A}\,\mathbf{s}^{(j)}\right)^T\mathbf{r}^{(k)} = \underbrace{\mathbf{r}^{(j)^T}\mathbf{r}^{(k)}}_{=0\ \textbf{(předpoklad)}} + t^{(j)}\underbrace{\mathbf{s}^{(j)^T}\mathbf{A}\,\mathbf{r}^{(k)}}_{=0\ \textbf{(Věta 2)}}$$

b) $\boxed{\mathbf{r}^{(j+1)^T}\mathbf{r}^{(j)} = 0}$

$$\mathbf{r}^{(j+1)^T}\mathbf{r}^{(j)} = \left(\mathbf{r}^{(j)} + t^{(j)}\mathbf{A}\,\mathbf{s}^{(j)}\right)^T\mathbf{r}^{(j)} = \mathbf{r}^{(j)^T}\mathbf{r}^{(j)} + t^{(j)}\mathbf{s}^{(j)^T}\mathbf{A}\,\mathbf{r}^{(j)} = \underbrace{\mathbf{r}^{(j)^T}\mathbf{r}^{(j)} - \frac{\mathbf{r}^{(0)^T}\mathbf{s}^{(j)}}{\mathbf{s}^{(j)^T}\mathbf{A}\,\mathbf{s}^{(j)}}\mathbf{s}^{(j)^T}\mathbf{A}\,\mathbf{r}^{(j)}}$$

$$\mathbf{s}^{(j)^T}\mathbf{A}\,\mathbf{s}^{(j)} = \mathbf{s}^{(j)^T}\mathbf{A}\,\mathbf{r}^{(j)} \qquad \textbf{(Věta 2}$$

$$\mathbf{r}^{(j)^T}\mathbf{r}^{(j)} = \mathbf{r}^{(0)^T}\mathbf{s}^{(j)} \qquad \textbf{(Věta 1}$$

□

Důkaz: Platí:

$$\beta_{ki} = -\frac{\mathbf{s}^{(i)^T}\mathbf{A}\mathbf{r}^{(k)}}{\mathbf{s}^{(i)^T}\mathbf{A}\,\mathbf{s}^{(i)}}$$

Pro čitatel platí:

$$\mathbf{s}^{(i)^T}\mathbf{A}\,\mathbf{r}^{(k)} = \mathbf{r}^{(k)^T}\mathbf{A}\,\mathbf{s}^{(\mathbf{i})} = \mathbf{r}^{(k)^T}\frac{1}{\mathbf{t}^{(\mathbf{i})}}\left(\mathbf{r}^{(\mathbf{i+1})} - \mathbf{r}^{(\mathbf{i})}\right),$$

kde se použil vztah

$$\mathbf{r}^{(\mathbf{i+1})} = \mathbf{r}^{(\mathbf{i})} + \mathbf{t}^{(\mathbf{i})}\mathbf{A}\,\mathbf{s}^{(\mathbf{i})}, \quad t^{(i)} \neq 0 \text{ pro } \mathbf{r}^{(i)} \neq \mathbf{0}$$

Platí

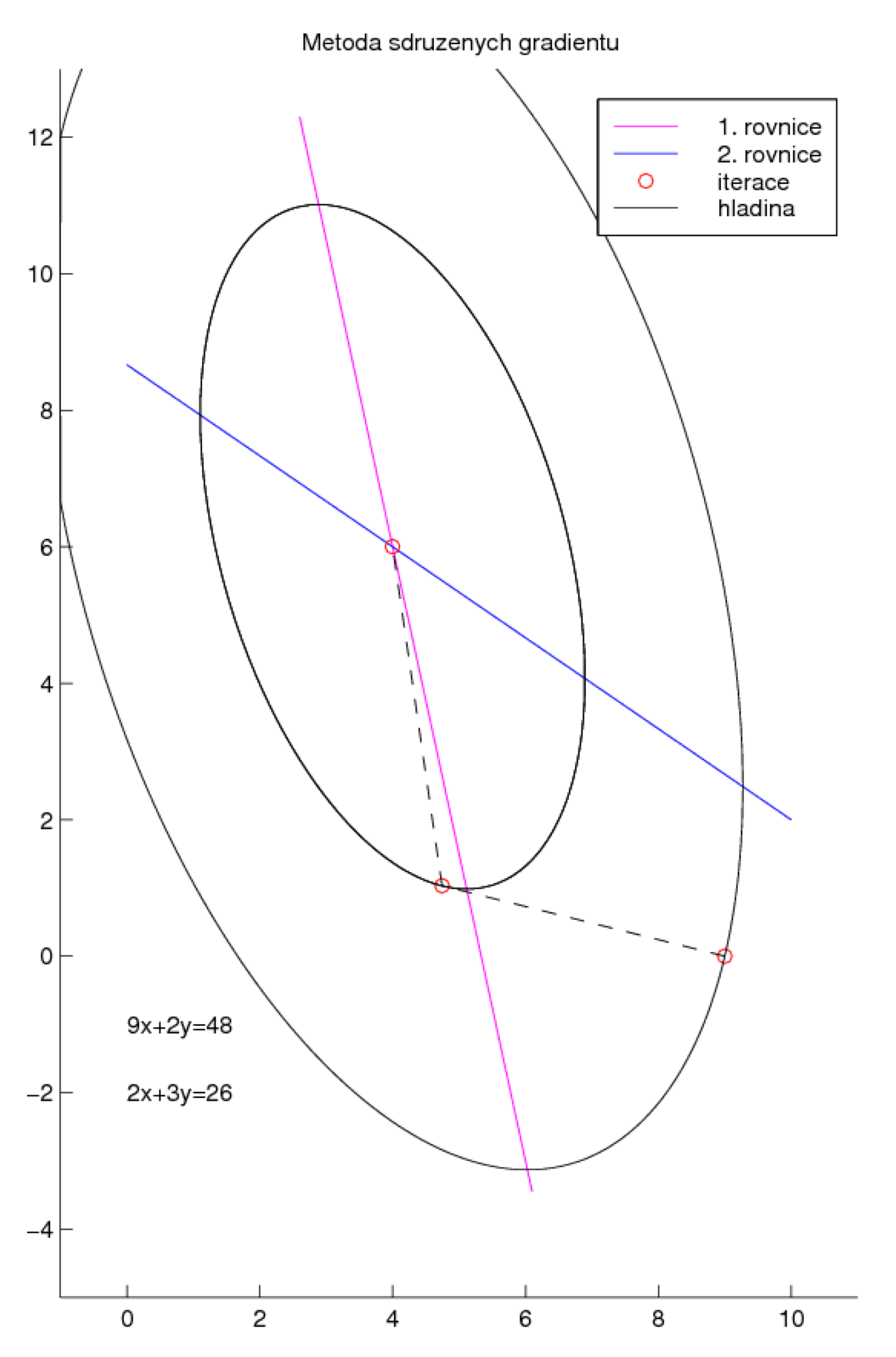
- pro $i < k-1$: čitatel $\beta_{ki} = \mathbf{r}^{(k)^T}\left(\mathbf{r}^{(i+1)} - \mathbf{r}^{(i)}\right)\frac{1}{t^{(i)}} = 0$ **(Tvrzení)**
- pro $i = k-1$: čitatel $\beta_{k,k-1} = \mathbf{r}^{(k)^T}\left(\mathbf{r}^{(k)} - \mathbf{r}^{(k-1)}\right)\frac{1}{t^{(k-1)}} = \mathbf{r}^{(k)^T}\mathbf{r}^{(k)}\frac{1}{t^{(k-1)}} \neq 0$ (pro $\mathbf{r}^{(k)} \neq \mathbf{0}$)

□

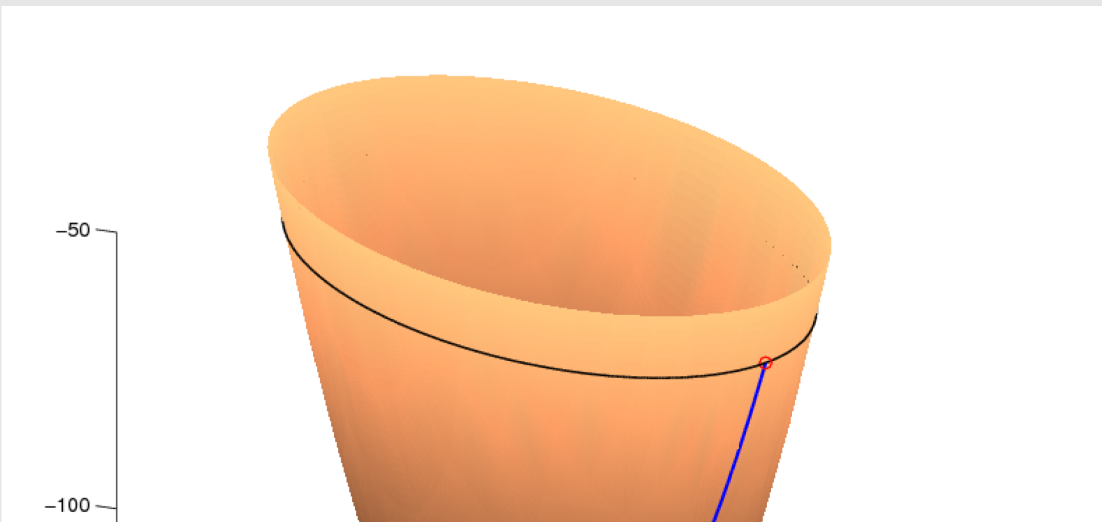
Algoritmus (**Metoda sdružených gradientů**)

1. $\mathbf{x}^{(0)}$, ε
2. $\mathbf{r}^{(0)} = \mathbf{b} - \mathbf{A}\,\mathbf{x}^{(0)}$, $\mathbf{s}^{(0)} = \mathbf{r}^{(0)}$
3. $t^{(k)} = \dfrac{\mathbf{s}^{(k)^T}\mathbf{r}^{(k)}}{\mathbf{s}^{(k)^T}\mathbf{A}\,\mathbf{s}^{(\mathbf{k})}}$
4. $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + t^{(k)}\mathbf{s}^{(k)}$
5. $\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} + t^{(k)}\mathbf{A}\,\mathbf{s}^{(\mathbf{k})}$
6. $\beta_k = -\dfrac{\mathbf{s}^{(k)^T}\mathbf{A}\,\mathbf{r}^{(k+1)}}{\mathbf{s}^{(k)^T}\mathbf{A}\,\mathbf{s}^{(\mathbf{k})}}$
7. $\mathbf{s}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta_k\mathbf{s}^{(k)}$
8. If $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| < \varepsilon$ then konec, else $\to$ add 3

Geometrický význam metody sdružených gradientů



k	$x^{(k)}$	$y^{(k)}$
0	9.0000	0
1	4.7425	1.0321
2	4.0000	6.0000



Poznámka: Gradientní metody patří mezi **nestacionární metody**.

např. pro metodu největšího spádu platí

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + t^{(k)}\mathbf{d}^{(k)} = \mathbf{x}^{(k)} + t^{(k)}\left(\mathbf{b} - \mathbf{A}\,\mathbf{x}^{(k)}\right) = \underbrace{\left(\mathbf{I} - t^{(k)}\mathbf{A}\right)}_{\mathbf{H}^{(k)}}\mathbf{x}^{(k)} + \underbrace{t^{(k)}\mathbf{b}}_{\mathbf{g}^{(k)}}$$

V každém kroku se mění matice $\mathbf{H}^{(k)}$.

Platí-li $||\mathbf{H}^{(k)}|| \to 0$ (pro $k \to \infty$), dostaneme metody se superlineární rychlostí konvergence.

Věta Nechť $\mathbf{A}$ je symetrická pozitivně definitní. Potom metoda sdružených gradientů konverguje nejvýše po n krocích. Navíc chyba k-té iterace ($k < n$) je ortogonální na směry $\mathbf{s}^{(j)}$, $j = 0, 1, \ldots, k-1$ a platí:

$$\boxed{||\mathbf{x}^{(k)} - \mathbf{x}^{*}||_{\mathbf{A}} \leq \frac{2C^k}{1 + C^{2k}}||\mathbf{x}^{(0)} - \mathbf{x}^{*}||_{\mathbf{A}}},$$

kde $C = \dfrac{\sqrt{\varkappa(\mathbf{A})} - 1}{\sqrt{\varkappa(\mathbf{A})} + 1}$, $\varkappa(\mathbf{A}) = \dfrac{\lambda_{max}}{\lambda_{min}}$.

Poznámka: V metodě největšího spádu vystupuje ve vztahu pro chybu k-té iterace koeficient

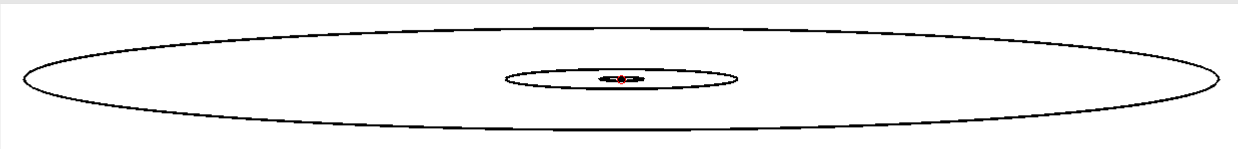
$$\left(\frac{\varkappa(\mathbf{A}) - 1}{\varkappa(\mathbf{A}) + 1}\right)^k.$$

Je zřejmé, že na rychlost konvergence má vliv číslo $\varkappa(\mathbf{A})$, tj. λ_{max} a λ_{min}. Čím blíže je λ_{max} a λ_{min}, tím rychleji metody konvergují.

Příklad 4

Řešte soustavu $\boxed{\mathbf{A}\mathbf{x} = \mathbf{b}}$, kde

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 10000 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 10000 \end{bmatrix}, \quad \text{přesné řešení } \mathbf{x}^{*} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$



- poměr poloos elips je $\sqrt{10000} : \sqrt{1} = 100 : 1$!!!
- $\varkappa(\mathbf{A}) = \frac{10000}{1} = 10000 \quad \Rightarrow \quad$ pomalá konvergence!

Vezměme si matici $\mathbf{P}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{10000} \end{bmatrix}$ $(\det(\mathbf{P}^{-1}) \neq 0)$ a řešme soustavu

$$\boxed{\mathbf{P}^{-1}\mathbf{A}\mathbf{x} = \mathbf{P}^{-1}\mathbf{b}}$$

$$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} = \begin{bmatrix}1\\ 1\end{bmatrix}$$

- $\varkappa(\mathbf{P}^{-1}\,\mathbf{A}) = \frac{1}{1} = 1 \quad \Rightarrow$ rychlá konvergence (1. iterace). Mluvíme o tzv. **předpodmiňování**.

Poznámka:

Chceme-li i novou (předpodmíněnou) soustavu řešit metodou sdružených gradientů, musí být její matice symetrická pozitivně definitní.

Místo matice $\mathbf{P}^{-1}\mathbf{A}$ vezmeme matici (podobnou $\mathbf{A}$) $\mathbf{P}^{-\frac{1}{2}}\mathbf{A}\mathbf{P}^{-\frac{1}{2}}$ ($\mathbf{P}$ ... symetrická pozitivně definitní) a řešíme soustavu

$$\boxed{\widetilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}}$$

$$\widetilde{\mathbf{A}} = \mathbf{P}^{-\frac{1}{2}}\mathbf{A}\mathbf{P}^{-\frac{1}{2}}\,, \quad \tilde{\mathbf{x}} = \mathbf{P}^{\frac{1}{2}}\mathbf{x}, \quad \tilde{\mathbf{b}} = \mathbf{P}^{-\frac{1}{2}}\mathbf{b}$$

Jak volit matice předpodmínění $\mathbf{P}$?

... řada možností, např. $\mathbf{P} = \operatorname{diag}(\mathbf{A})$

Příklad 5

Porovnejte vlastní čísla zadané matice $\mathbf{A}$ a matice získané pomocí diagonálního předpodmínění.

$$\mathbf{A} = \begin{bmatrix}1000000 & 200 & 30 & 0\\ 200 & 10000 & 40 & 0\\ 30 & 40 & 100 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix}1000000 & 0 & 0 & 0\\ 0 & 10000 & 0 & 0\\ 0 & 0 & 100 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

$$\mathbf{P}^{-\frac{1}{2}} = \begin{bmatrix}\frac{1}{\sqrt{1000000}} & 0 & 0 & 0\\ 0 & \frac{1}{\sqrt{10000}} & 0 & 0\\ 0 & 0 & \frac{1}{\sqrt{100}} & 0\\ 0 & 0 & 0 & \frac{1}{\sqrt{1}}\end{bmatrix} = \begin{bmatrix}\frac{1}{1000} & 0 & 0 & 0\\ 0 & \frac{1}{100} & 0 & 0\\ 0 & 0 & \frac{1}{10} & 0\\ 0 & 0 & 0 & \frac{1}{1}\end{bmatrix}$$

$$\widetilde{\mathbf{A}} = \mathbf{P}^{-\frac{1}{2}}\mathbf{A}\mathbf{P}^{-\frac{1}{2}} = \begin{bmatrix}\frac{1}{1000} & 0 & 0 & 0\\ 0 & \frac{1}{100} & 0 & 0\\ 0 & 0 & \frac{1}{10} & 0\\ 0 & 0 & 0 & \frac{1}{1}\end{bmatrix}\begin{bmatrix}1000000 & 200 & 30 & 0\\ 200 & 10000 & 40 & 0\\ 30 & 40 & 100 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\frac{1}{1000} & 0 & 0 & 0\\ 0 & \frac{1}{100} & 0 & 0\\ 0 & 0 & \frac{1}{10} & 0\\ 0 & 0 & 0 & \frac{1}{1}\end{bmatrix} =$$

$$= \begin{bmatrix}1 & 0.002 & 0.003 & 0\\ 0.002 & 1 & 0.04 & 0\\ 0.003 & 0.04 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

vlastní čísla matice $\mathbf{A}$:

$$\lambda_1 = 1; \quad \lambda_2 \doteq 99{,}837534233; \quad \lambda_3 \doteq 10000{,}121161153; \quad \lambda_4 \doteq 1000000{,}041304614$$

$$\boxed{\varkappa(\mathbf{A}) \doteq \frac{1000000{,}041304614}{1} = 1000000{,}041304614}$$

vlastní čísla matice $\widetilde{\mathbf{A}} = \mathbf{P}^{-\frac{1}{2}}\mathbf{A}\mathbf{P}^{-\frac{1}{2}}$:

$$\widetilde{\lambda}_1 \doteq 0{,}9599874549379999; \quad \widetilde{\lambda}_2 \doteq 0{,}999702401514713; \quad \widetilde{\lambda}_3 = 1; \quad \widetilde{\lambda}_4 \doteq 1{,}040310143547287$$

$$\boxed{\varkappa(\widetilde{\mathbf{A}}) \doteq \frac{1{,}040310143547287}{0{,}9599874549379999} \doteq 1{,}083670560689229}$$