Metoda největšího spádu

Teorie (velmi stručný výběr z přednášek)

Věta: Nechť A je symetrická, pozitivně definitní (*spd*) matice, b je vektor a $J(x)$ je kvadratický funkcionál: $J(x) = \frac{1}{2}x^T A\, x - x^T b\,.$
Pak $\quad A\bar{x} = b \;\Leftrightarrow\; J(\bar{x}) < J(x) \quad \forall\, x \neq \bar{x}.$

Tato věta říká, že místo řešení systému lineárních rovnic $Ax = b$ se *spd* maticí A můžeme hledat minimum kvadratického funkcionálu $J(x)$. K tomu lze využít *gradientní metody.* Asi nejnázornější gradientní metodou je *metoda největšího spádu.*

Metoda největšího spádu:

Hlavní myšlenka: Začneme v nějakém bodě x_0, najdeme směr, v němž hodnota $J(x)$ nejrychleji klesá, a pohybujeme se tímto směrem tak dlouho, dokud $J(x)$ klesá. V tomto novém bodě se zastavíme, znovu najděme směr největšího spádu hodnoty $J(x)$ a zopakujeme stejný postup k nalezeního dalšího bodu.

Poznámka: směr největšího poklesu hodnoty funkce je rovnoběžný s gradientem funkce, ale má opačnou orientaci. Gradient v libovolném bodě je kolmý k vrstevnici procházející tímto bodem. Viz Obr. 1.

Poznámka: Směr opačný ke gradientu $J(x)$ se rovná residuu $r = b - A\,x$ soustavy $Ax = b$.

Algoritmus:

Zvolte $x^{(0)}$. Pro $k = 0, 1, 2, \dots$ postupně počítejte

1. $r^{(k)} = b - A\,x^{(k)}$
2. $\alpha_k = (r^{(k)})^T\, r^{(k)} / (r^{(k)})^T A\; r^{(k)}$
3. $x^{(k+1)} = x^{(k)} + \alpha_k\, r^{(k)}$

dokud nenastane $||r^{(k)}|| < \epsilon$ pro nějakou malou, předem zvolenou hodnotu ϵ.

Příklad

Je dána soustava lineárních rovnic $Ax = b$, kde

$$A = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix}, \quad b = \begin{bmatrix} -1 \\ 7 \\ -7 \end{bmatrix}$$

a) Lze pro tuto soustavu použít metodu největšího spádu?
b) Pokud ano, spočítejte první tři iterace touto metodou; jako výchozí hodnotu zvolte $x^{(0)} = (0, 0, 0)^T$.

Řešení:

a) Nejdříve ověříme postačující podmínku pro konvergenci této metody. Potřebujeme zjistit, zda matice A je *spd*: A je symetrická, stačí tedy ověřit, že je i pozitivně definitní:

$$\det(3) = 3 > 0 \ , \ \det\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} = 8 > 0 \ , \ \det\begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix} = 20 > 0$$

Všehny hlavní minory jsou kladné, takže matice A je pozitivně definitní.

Závěr: metodu největšího spádu lze pro danou soustavu použít.

b) $\mathbf{k = 0}$:

1.

$$r^{(0)} = b - A\,x^{(0)} = \begin{bmatrix} -1 \\ 7 \\ -7 \end{bmatrix} - \begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 7 \\ -7 \end{bmatrix}$$

2.

$$(r^{(0)})^T\,r^{(0)} = \begin{bmatrix} -1 & 7 & -7 \end{bmatrix} \begin{bmatrix} -1 \\ 7 \\ -7 \end{bmatrix} = 1 + 49 + 49 = 99$$

$$(r^{(0)})^T A\,r^{(0)} = \begin{bmatrix} -1 & 7 & -7 \end{bmatrix} \begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 7 \\ -7 \end{bmatrix} = \begin{bmatrix} -1 & 7 & -7 \end{bmatrix} \begin{bmatrix} -17 \\ 29 \\ -29 \end{bmatrix} = 423$$

$$\alpha_0 = (r^{(0)})^T\,r^{(0)}/(r^{(0)})^T A\ r^{(0)} = 99/423 = 0.2340$$

3.

$$\mathbf{x^{(1)}} = x^{(0)} + \alpha_0\,r^{(0)} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + 0.2340 \begin{bmatrix} -1 \\ 7 \\ -7 \end{bmatrix} = \begin{bmatrix} -0.2340 \\ 1.6383 \\ -1.6383 \end{bmatrix}$$

$\mathbf{k = 1}$:

1.

$$r^{(1)} = b - A\,x^{(1)} = \begin{bmatrix} -1 \\ 7 \\ -7 \end{bmatrix} - \begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} -0.2340 \\ 1.6383 \\ -1.6383 \end{bmatrix} = \begin{bmatrix} 2.9787 \\ 0.2128 \\ -0.2128 \end{bmatrix}$$

2.

$$(r^{(1)})^T\,r^{(1)} = \begin{bmatrix} 2.9787 & 0.2128 & -0.2128 \end{bmatrix} \begin{bmatrix} 2.9787 \\ 0.2128 \\ -0.2128 \end{bmatrix} = 8.9633$$

$$(r^{(1)})^T A\, r^{(1)} = \begin{bmatrix} 2.9787 & 0.2128 & -0.2128 \end{bmatrix} \begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 2.9787 \\ 0.2128 \\ -0.2128 \end{bmatrix} =$$

$$= \begin{bmatrix} 2.9787 & 0.2128 & -0.2128 \end{bmatrix} \begin{bmatrix} 8.5106 \\ -2.1277 \\ 2.1277 \end{bmatrix} = 24.4455$$

$$\alpha_1 = (r^{(1)})^T\, r^{(1)} / (r^{(1)})^T A\; r^{(1)} = 8.9633/24.4455 = 0.3667$$

3.

$$\mathbf{x^{(2)}} = x^{(1)} + \alpha_1\, r^{(1)} = \begin{bmatrix} -0.2340 \\ 1.6383 \\ -1.6383 \end{bmatrix} + 0.3667 \begin{bmatrix} 2.9787 \\ 0.2128 \\ -0.2128 \end{bmatrix} = \begin{bmatrix} 0.8582 \\ 1.7163 \\ -1.7163 \end{bmatrix}$$

k = 2:

1.

$$r^{(2)} = b - A\, x^{(2)} = \begin{bmatrix} -1 \\ 7 \\ -7 \end{bmatrix} - \begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 0.8582 \\ 1.7163 \\ -1.7163 \end{bmatrix} = \begin{bmatrix} -0.1418 \\ 0.9929 \\ -0.9929 \end{bmatrix}$$

2.

$$(r^{(2)})^T\, r^{(2)} = \begin{bmatrix} -0.1418 & 0.9929 & -0.9929 \end{bmatrix} \begin{bmatrix} -0.1418 \\ 0.9929 \\ -0.9929 \end{bmatrix} = 1.9919$$

$$(r^{(2)})^T A\, r^{(2)} = \begin{bmatrix} -0.1418 & 0.9929 & -0.9929 \end{bmatrix} \begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} -0.1418 \\ 0.9929 \\ -0.9929 \end{bmatrix} =$$

$$= \begin{bmatrix} -0.1418 & 0.9929 & -0.9929 \end{bmatrix} \begin{bmatrix} -2.4113 \\ 4.1135 \\ -4.1135 \end{bmatrix} = 8.5106$$

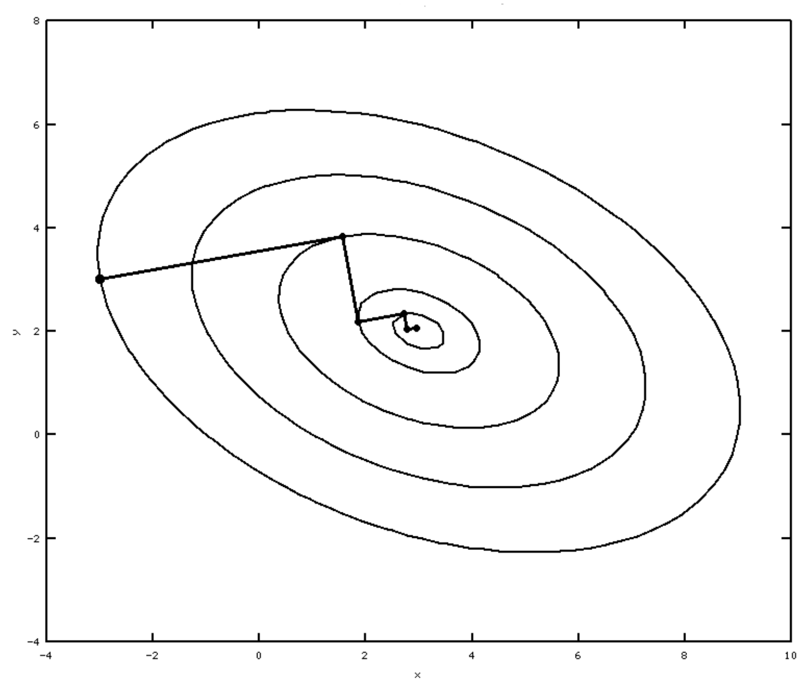
$$\alpha_2 = (r^{(2)})^T\, r^{(2)} / (r^{(2)})^T A\; r^{(2)} = 1.9919/8.5106 = 0.2340$$

3.

$$\mathbf{x^{(3)}} = x^{(2)} + \alpha_2\, r^{(2)} = \begin{bmatrix} 0.8582 \\ 1.7163 \\ -1.7163 \end{bmatrix} + 0.2340 \begin{bmatrix} -0.1418 \\ 0.9929 \\ -0.9929 \end{bmatrix} = \begin{bmatrix} 0.8250 \\ 1.9487 \\ -1.9487 \end{bmatrix}$$

$$r^{(3)} = b - A\, x^{(3)} = \begin{bmatrix} -1 \\ 7 \\ -7 \end{bmatrix} - \begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 0.8250 \\ 1.9487 \\ -1.9487 \end{bmatrix} = \begin{bmatrix} 0.4225 \\ 0.0302 \\ -0.0302 \end{bmatrix}$$

Konvergence je dost pomalá - přesné řešení je $\bar{x} = (1,\ 2,\ -2)^T$.



Obr. 1: Elipsy představují vrstevnice kvadratické funkce dvou proměnných. Lomená čára začínající na vnější elipse je cesta směrem k minimu kvadratické funkce spočítaná metodou největšího spádu.