Domácí úkol OPT 4

by a very good student

30. listopadu 2019

Proložení bodů kružnicí

Mějme m bodů v rovině, $\mathbf{a}_1,...,\mathbf{a}_m \in \mathbb{R}^2$. Chceme najít kružnici se středem $\mathbf{c} \in \mathbb{R}^2$ a poloměrem $r \geq 0$ takovou, že součet čtverců vzdáleností bodů od kružnice je nejmenší.

Označme jako $\mathbf{x} = (\mathbf{c}, r) \in \mathbb{R}^3$ vektor parametrů kružnice. Nechť dist (\mathbf{x}, \mathbf{a}) je orientovaná vzdálenost bodu \mathbf{a} od kružnice s parametry \mathbf{x} . Tedy $|\operatorname{dist}(\mathbf{x}, \mathbf{a})|$ je Eukleidovská vzdálenost bodu \mathbf{a} od kružnice, přičemž pro \mathbf{a} vně kružnice je dist $(\mathbf{x}, \mathbf{a}) > 0$ a pro \mathbf{a} uvnitř kružnice je dist $(\mathbf{x}, \mathbf{a}) < 0$. Chceme minimalizovat funkci

$$f(\mathbf{x}) = \sum_{i=1}^{m} \operatorname{dist}(\mathbf{x}, \mathbf{a}_i)^2$$

Funkce dist

Označili jsme parametry \mathbf{x} kružnice tak, že $\mathbf{c} = (x_1, x_2)$ a $r = x_3$. Nechť $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), ..., g_m(\mathbf{x}))^T \in \mathbb{R}^{m \times 1}$, a matice $\mathbf{A} = [\mathbf{a}_1, ..., \mathbf{a}_m] \in \mathbb{R}^{2 \times m}$. Funkci $g_i \in \mathbb{R}$, ze které se skládá funkce \mathbf{g} a která odpovídá funkci dist, můžeme vyjádřit jako

$$g_i(\mathbf{x}) = \operatorname{dist}(\mathbf{x}, \mathbf{a}_i) = \left\| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \mathbf{a}_i \right\| - x_3$$

Pro body na kružnici pak platí, že $g_i(\mathbf{x}) = 0$. Pro body uvnitř kružnice tato funkce nabývá hodnot 0 až $-x_3$, pro body vně kružnice je hodnota kladná. Nyní vypočteme derivaci funkce $g_i(\mathbf{x})$:

$$g_i(\mathbf{x})' = \begin{bmatrix} \frac{x_1 - a_{i1}}{\| x_1 \|_{x_2}^2 \| - \mathbf{a}_i} & \frac{x_2 - a_{i2}}{\| x_2 \|_{x_2}^2 \| - \mathbf{a}_i} \end{bmatrix} - 1 = \begin{bmatrix} \frac{\left(x_1 \|_{x_2}^2 - \mathbf{a}_i \|_{x_2}^2 \right) - 1}{\| x_2 \|_{x_2}^2 - \mathbf{a}_i} \end{bmatrix} \in \mathbb{R}^{1 \times 3}$$

Tedy derivace funkce $\mathbf{g}(\mathbf{x})$ je matice $\mathbb{R}^{m\times 3}$, složená v řádcích z těchto derivací g_i .

Gaussova-Newtonova iterační metoda

Nový bod \mathbf{x}' dostaneme z předchozího bodu \mathbf{x} iteračním krokem

$$\mathbf{x}' = \mathbf{x} - \mathbf{g}'(\mathbf{x})^{\dagger} \mathbf{g}(\mathbf{x}) = \mathbf{x} - (\mathbf{g}'(\mathbf{x})^T \mathbf{g}'(\mathbf{x}))^{-1} \mathbf{g}'(\mathbf{x})^T \mathbf{g}(\mathbf{x})$$

Levenbergova-Marquardtova iterační metoda

Tato iterační metoda upravuje iterační krok Gaussovy-Newtonovy metody přidáním regularizačního členu $\mu_k \mathbf{I}$.

$$\mathbf{x}' = \mathbf{x} - (\mathbf{g}'(\mathbf{x})^T \mathbf{g}'(\mathbf{x}) + \mu_k \mathbf{I})^{-1} \mathbf{g}'(\mathbf{x})^T \mathbf{g}(\mathbf{x})$$

Doplňující úkoly

1. Pro $m \ge 3$ a body $\mathbf{a}_1,...,\mathbf{a}_m$ v obecné poloze, zkuste nalézt co nejobecnější podmínky na body \mathbf{x} , ve kterých je funkce f diferencovatelná. Odpověď odůvodněte.

Funkci f upravíme a pak zderivujeme

$$f(\mathbf{x}) = \sum_{i=1}^{m} g_i(\mathbf{x})^2 = \|\mathbf{g}(\mathbf{x})\|^2 = \mathbf{g}(\mathbf{x})^T \mathbf{g}(\mathbf{x})$$

$$f'(\mathbf{x}) = 2 \|\mathbf{g}(\mathbf{x})\| \, \frac{\mathbf{g}(\mathbf{x})^T}{\|\mathbf{g}(\mathbf{x})\|} \mathbf{g'}(\mathbf{x}) = 2 \mathbf{g}(\mathbf{x})^T \mathbf{g'}(\mathbf{x})$$

Aby tento výraz dával smysl, musí existovat derivace $g_i(\mathbf{x})$. Tj. musí platit

$$\left\| \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \mathbf{a}_i \right\| \neq 0$$

pro každý bod \mathbf{a}_i , neboli střed kružnice nesmí ležet v žádném z bodů \mathbf{a}_i . Pro všechna ostatní \mathbf{x} je funkce f diferencovatelná.

2. Může se zdát, že algoritmy na nelineární nejmenší čtverce bez omezení nejde použít, protože máme omezení $r \geq 0$. Vadí to? Co se stane, budeme-li toto omezení ignorovat? Můžou algoritmy konvergovat k řešení se záporným r? Odpovědi odůvodněte.

Vynechání omezení ničemu nevadí. Pro r < 0 bude funkce g_i nabývat hodnot $\geq r$. Funkce f potom nebude nabývat minima pro body ležící co nejblíže kružnici (tam se hodnoty sumy čtverců blíží nule, ostatní jsou kladné), ale bude nabývat minima r^2 pro body blížící se středu kružnice. Tedy první iterací budeme pouze umistovat střed kružnice mezi body \mathbf{a}_i .

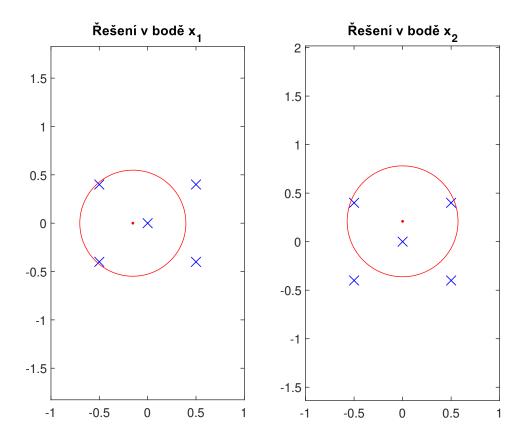
Dalšími iteracemi se ovšem poloměr x_3 změní na kladný a algoritmus najde správné řešení, protože při první iteraci je hodnota gradientu funkce f ve třetí složce záporná (třetí prvek v \mathbf{g}' je -1 a hodnota \mathbf{g} musí být pro $x_3 < 0$ kladná).

3. Může mít pro nějakou množinu $\mathbf{a}_1,...,\mathbf{a}_m$ funkce f více lokálních minim s různou funkční hodnotou?

Ano, funkce f může mít více lokálních minim s různou funkční hodnotou. Např. pro body

$$\mathbf{a}_1 = \begin{bmatrix} -0.5 \\ 0.4 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 0.5 \\ 0.4 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} 0.5 \\ -0.4 \end{bmatrix}, \mathbf{a}_4 = \begin{bmatrix} -0.5 \\ -0.4 \end{bmatrix}, \mathbf{a}_5 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

zvolíme dva různé počáteční body pro konvergenci, konkrétně $\mathbf{x}_{01} = (-0.5, 0, 0.3)^T$ a $\mathbf{x}_{02} = (-0.5, 0.3, 0.3)^T$.



Jako výsledek dostaneme dva stacionární body $\mathbf{x}_1 = (-0.153134, 0, 0.548761)$ a $\mathbf{x}_2 = (0, 0.209667, 0.571324)$.

Jedná se o minima, jelikož k nim dokonvergovala Gaussova-Newtonova metoda, která se vždy pohybuje ve směru největšího spádu funkce, tj. směrem k lokálnímu minimu.

Funkční hodnoty (hodnoty kritéria) po 20 iteracích GN metodou jsou

$$f(\mathbf{x}_1) = 0.25155, f(\mathbf{x}_2) = 0.22774$$

Gradient (vektor) dostaneme jako transpozici totální derivace (lin. funkce), tedy

$$\nabla f(\mathbf{x}) = f'(\mathbf{x})^T = 2\mathbf{g}'(\mathbf{x})^T \mathbf{g}(\mathbf{x})$$

Hodnoty gradientu ∇f v těchto bodech po 20 iteracích jsou

$$\nabla f(\mathbf{x}_1) = \begin{bmatrix} 0\\ 1.4438450435 \cdot 10^{-13}\\ 1.1102230246 \cdot 10^{-16} \end{bmatrix}$$
$$\nabla f(\mathbf{x}_2) = \begin{bmatrix} 2.7006113366 \cdot 10^{-6}\\ -1.7883805548 \cdot 10^{-11}\\ -6.1739502399 \cdot 10^{-12} \end{bmatrix}$$