B0B33OPT
Karina Balagazova

NLLS: Kružnice

1 Proložení bodů kružnicí

1.1 dist(x, a)

Máme $a_1, ..., a_m \in \mathbb{R}^2$. Označme jako $x = (c, r) \in \mathbb{R}^3$ vektor parametrů kružnice. Funkce $dist(x, a)$ je orientovaná vzdálenost bodu a od kružnice s parametry x. Spočitáme ji jako

$$dist(x, a) = \|c - a_i\| - r$$

přičemž pro a vně kružnice je $dist(x, a) > 0$ a pro a uvnitř kružnice je $dist(x, a) < 0$.

Jacobiho matice této funkce bude vypadat takto:

$$dist'(x, a) = \begin{bmatrix} \frac{c_1 - a_{i_1}}{\|c - a_i\|} & \frac{c_2 - a_{i_2}}{\|c - a_i\|} & -1 \end{bmatrix}$$

1.2 Gauss-Newtonova metoda

Iterační krok:

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

1.3 Levenberg-Marquardtova metoda

Iterační krok:

$$\mathbf{x_{new}} = \mathbf{x} - (\mathbf{g}'(\mathbf{x})^{\mathbf{T}}\mathbf{g}'(\mathbf{x}) + \mu_{\mathbf{k}}\mathbf{I})^{-\mathbf{1}}\mathbf{g}'(\mathbf{x})^{\mathbf{T}}\mathbf{g}(\mathbf{x})$$

Úkoly

1 Diferencovatelnost funkce f

$$f(x) = \sum_{i=1}^{m} dist(x, a_i)^2 = \|dist(x, a)\|^2$$

Postačující podmínkou pro diferencovatelnost funkce v bodě x je existence a spojitost všech parciálních derivací funkce v tomto bodě. Zkusíme tedy funkci f zderivovat

$$f'(x) = 2dist(x, a)^T dist'(x, a)$$

Takže musí platit

$$\|c - a_i\| \neq 0$$

tj. funkce f je diferencovatelná tehdy, když $c \neq a_i$. Jinými slovy střed kružnice by neměl ležet v žádném z bodů $a_1, ..., a_m$.

2 Omezení $r \geq 0$

Ignorování této podmínky by nemělo nic změnit. Při počátečním zvolení $r < 0$ vždycky bude existovat lepší kladné r, ke kterému pak bude konvergovat.

3 Lokální minima s různou funkční hodnotou

Zkusíme spustit Gauss-Newtonovou metodu pro body $a_1, ..., a_5$ a počáteční odhady x_{0_1} a x_{0_2}.

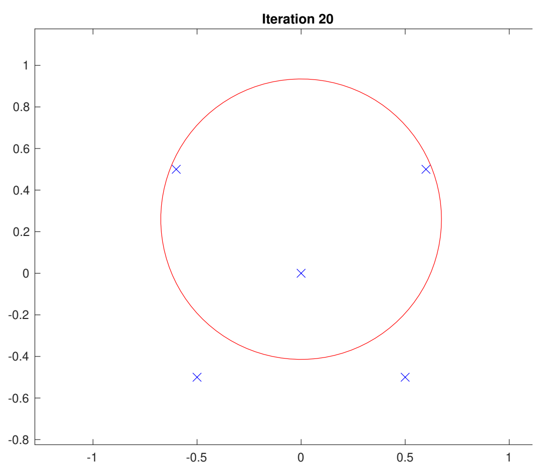
$$a = \begin{bmatrix} 0.6 & 0.5 & -0.6 & -0.5 & 0 \\ 0.5 & -0.5 & 0.5 & -0.5 & 0 \end{bmatrix}, x_{0_1} = \begin{bmatrix} 0 \\ 0.5 \\ 0.7 \end{bmatrix}, x_{0_2} = \begin{bmatrix} 0 \\ -0.5 \\ 0.7 \end{bmatrix}$$

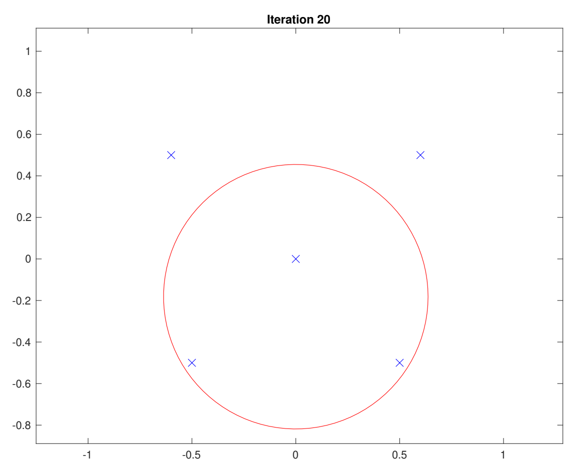
Metoda zkonverguje do dvou různých stacionárních bodů, a vzhledem k tomu, že jsme použili Gauss-Newtonovou metodu, tak téměř jistě víme, že jsou to lokální minima. Po 20 iterací dostaneme parametry:

$$x_1 = \begin{bmatrix} 0.0000 \\ -0.1818 \\ 0.6367 \end{bmatrix}, x_2 = \begin{bmatrix} 0.0000 \\ 0.2601 \\ 0.6744 \end{bmatrix}$$

$$f(x_1) = 0.2841, f(x_2) = 0.3583$$

$$\nabla f(x_2) = \begin{bmatrix} 0.0134 \\ 0.2220 \\ -0.7772 \end{bmatrix}, \nabla f(x_2) = \begin{bmatrix} -0.0073 \\ 0.4441 \\ 0.2220 \end{bmatrix}$$





4 Minimalizace funkce přes dvě proměnné

Abychom mohli najít minimum funkce, její derivace v tomto bodě musí být 0:

$$\frac{\partial f}{\partial r} = 2\sum_{i=1}^{m} dist(x, a_i) * (-1) = -2\sum_{i=1}^{m}(\|c - a_i\| - r) = 0$$

$$\sum_{i=1}^{m} \|c - a_i\| - mr = 0$$

A odvodíme tak vzoreček pro r:

$$r = \frac{1}{m} \sum_{i=1}^{m} \|c - a_i\|$$

Dosadíme vysledek do funkci $dist$:

$$dist(c, a) = \|c - a\| - \frac{1}{m} \sum_{i=1}^{m} \|c - a_i\|$$

Funkce f, vyjadřená přes dvě proměnné bude vypadat takto:

$$g(c) = \sum_{i=1}^{m} dist(c, a_i)^2 = \|dist(c, a)\|^2 = \|\|c - a\| - \frac{1}{m} \sum_{i=1}^{m} \|c - a_i\|\|^2 =$$

$$= \sum_{i=1}^{m} \|c - a_i\|^2 - \frac{2}{m} (\sum_{i=1}^{m} \|c - a_i\|)^2 + m \frac{1}{m^2} (\sum_{i=1}^{m} \|c - a_i\|)^2$$

tj. minimalizujeme funkce

$$g(c) = \sum_{i=1}^{m} \|c - a_i\|^2 - \frac{1}{m} (\sum_{i=1}^{m} \|c - a_i\|)^2$$

Velkých rozdílu mezi těmito dvěma formulacemi úlohy jsem si nevšimla. Při porovnání na konkretních příkladech metoda vždy konvergovala ke stejným řešením a přibližně ve stejném počtu iterací.