Optimalizace

Nelineární metody nejmenších čtverců

Tomáš Kroupa, Lukáš Adam

FEL ČVUT

Obsah

Nelineární metoda nejmenších čtverců

Nechť $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$ je dif. zobrazení. Hledáme přibližné řešení přeurčené soustavy $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ ve smyslu nejmenších čtverců.

Minimalizuj funkci

$$f(\mathbf{x}) := \|\mathbf{g}(\mathbf{x})\|^2 = \sum_{i=1}^{m} g_i(\mathbf{x})^2, \qquad \mathbf{x} \in \mathbb{R}^n$$

Řešení lineární nehomogenní soustavy je speciálním případem:

$$\mathbf{g}(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$$

Gauss-Newtonova metoda

Zobrazení $\mathbf{g}$ v okolí bodu $\mathbf{x}_k$ aproximujeme afinním zobrazením $\mathbf{T}_1$ a místo funkce $\|\mathbf{g}(\mathbf{x})\|^2$ tak minimalizujeme $\|\mathbf{T}_1(\mathbf{x})\|^2$.

Iterace

$$\mathbf{x}_{k+1} = \mathbf{x}_k - (\mathbf{g}'(\mathbf{x}_k)^\mathsf{T}\mathbf{g}'(\mathbf{x}_k))^{-1}\mathbf{g}'(\mathbf{x}_k)^\mathsf{T}\mathbf{g}(\mathbf{x}_k)$$

- Jacobiho matice $\mathbf{g}'(\mathbf{x}_k)$ musí mít LN sloupce
- Platí $f'(\mathbf{x}_k) = 2\mathbf{g}(\mathbf{x}_k)^\mathsf{T}\mathbf{g}'(\mathbf{x}_k)$

Levenberg-Marquardtova metoda

Iterace G-N metody

$$\mathbf{x}_{k+1} = \mathbf{x}_k - (\mathbf{g}'(\mathbf{x}_k)^\mathsf{T}\mathbf{g}'(\mathbf{x}_k))^{-1}\mathbf{g}'(\mathbf{x}_k)^\mathsf{T}\mathbf{g}(\mathbf{x}_k)$$

Iterace L-M metody

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

Regularizační parametr $\mu_k > 0$ umožňuje plynule kombinovat mezi

- G-N metodu (μ_k je malé)
- gradientní metodou (μ_k je velké)

Stochastic gradient descent

Alternativní zápis (lepší pro numerické metody)

$$\text{minimalizuj} \qquad \frac{1}{2m}\sum_{i=1}^{m} g_i(\mathbf{x})^2.$$

Iterace gradient descent metody

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \frac{1}{m}\sum_{i=1}^{m} g_i(\mathbf{x}) g_i'(\mathbf{x})^\top$$

Iterace stochastic gradient descent metody

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \frac{1}{|I|}\sum_{i\in I} g_i(\mathbf{x}) g_i'(\mathbf{x})^\top$$

Reference

- T. Werner. *Optimalizace* (kapitola 9). Elektronická skripta. FEL ČVUT, 2020.
- G. Goh. *Why Momentum Really Works.* `https://distill.pub/2017/momentum/`
- *Matlab Activities for Multivariable Calculus.* `https://mse.redwoods.edu/darnold/math50c/matlab/`