Optimalizace

Nelineární metoda nejmenších čtverců

Tomáš Kroupa, Mirko Navara

24. 4. 2024

FEL ČVUT

Obsah

Nelineární metoda nejmenších čtverců pro řešení soustavy rovnic

Dosud jsme hledali ideální přesné řešení; teď se vrátíme do reality.

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

Úloha: Minimalizujte funkci

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

Gaussova-Newtonova metoda řešení soustavy rovnic

Speciální případ

Přeurčená nehomogenní soustava lineárních rovnic
$\mathbf{T}_1(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b} = \mathbf{0}$ má optimální řešení $\mathbf{x} = \underbrace{(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top}_{\mathbf{A}^+} \mathbf{b}\,.$

Zobrazení $\mathbf{g}$ aproximujeme v okolí bodu $\mathbf{x}_k$ afinním zobrazením $\mathbf{T}_{1,k}$,

$$\mathbf{g}(\mathbf{y}) \approx \mathbf{T}_{1,k}(\mathbf{y}) = \underbrace{\mathbf{g}(\mathbf{x}_k)}_{-\mathbf{b}} + \underbrace{\mathbf{g}'(\mathbf{x}_k)}_{\mathbf{A}}\,\underbrace{(\mathbf{y} - \mathbf{x}_k)}_{\mathbf{x}}$$

a místo hledání minima $\|\mathbf{g}(\mathbf{y})\|^2$ minimalizujeme $\|\mathbf{T}_{1,k}(\mathbf{y})\|^2$,

$$\begin{aligned}\mathbf{x}_{k+1} - \mathbf{x}_k &= \underset{\mathbf{y}}{\operatorname{argmin}} \|\mathbf{T}_{1,k}(\mathbf{y})\|^2 = \mathbf{A}^+ \mathbf{b} = \\ &= -\mathbf{g}'(\mathbf{x}_k)^+ \mathbf{g}(\mathbf{x}_k) = -(\mathbf{g}'(\mathbf{x}_k)^\top \mathbf{g}'(\mathbf{x}_k))^{-1} \mathbf{g}'(\mathbf{x}_k)^\top \mathbf{g}(\mathbf{x}_k)\,.\end{aligned}$$

Gaussova-Newtonova metoda

Iterace Gaussovy-Newtonovy metody (přibližného řešení soustavy m rovnic o n neznámých)

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

Speciální případ:

Iterace Newtonovy metody (řešení soustavy n rovnic o n neznámých)

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{g}'(\mathbf{x}_k)^{-1}\,\mathbf{g}(\mathbf{x}_k)\,.$$

Možné zobecnění Gaussovy-Newtonovy metody

Iterace Gaussovy-Newtonovy metody (obecněji)

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{x}_k - \alpha_k\, \mathbf{g}'(\mathbf{x}_k)^+\, \mathbf{g}(\mathbf{x}_k) = \\ &= \mathbf{x}_k - \alpha_k\, (\mathbf{g}'(\mathbf{x}_k)^\mathsf{T}\, \mathbf{g}'(\mathbf{x}_k))^{-1}\, \mathbf{g}'(\mathbf{x}_k)^\mathsf{T}\, \mathbf{g}(\mathbf{x}_k)\,.\end{aligned}$$

Jacobiho matice $\mathbf{g}'(\mathbf{x}_k)$ musí mít LN sloupce, pak je $\mathbf{g}'(\mathbf{x}_k)^\mathsf{T}\, \mathbf{g}'(\mathbf{x}_k)$ pozitivně definitní.

Nelineární metoda nejmenších čtverců jako optimalizační úloha

Minimalizujeme $f(\mathbf{x}) = \|\mathbf{g}(\mathbf{x})\|^2 = \mathbf{g}(\mathbf{x})^\mathsf{T}\,\mathbf{g}(\mathbf{x})$, kde

$$
\begin{aligned}
f'(\mathbf{x}) &= 2\,\mathbf{g}(\mathbf{x})^\mathsf{T}\,\mathbf{g}'(\mathbf{x})\,, \\
f''(\mathbf{x}) &= 2\,\mathbf{g}'(\mathbf{x})^\mathsf{T}\,\mathbf{g}'(\mathbf{x}) + 2\sum_i g_i(\mathbf{x})\,g_i''(\mathbf{x})\,.
\end{aligned}
$$

Iterace Newtonovy metody (optimalizační)

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

Pro srovnání:

Iterace Gaussovy-Newtonovy metody

$$
\begin{aligned}
\mathbf{x}_{k+1} &= \mathbf{x}_k - \left(\mathbf{g}'(\mathbf{x}_k)^\mathsf{T}\,\mathbf{g}'(\mathbf{x}_k)\right)^{-1}\overbrace{\mathbf{g}'(\mathbf{x}_k)^\mathsf{T}\,\mathbf{g}(\mathbf{x}_k)}^{\frac{1}{2}\,f'(\mathbf{x}_k)^\mathsf{T}} \\
&= \mathbf{x}_k - \left(2\,\mathbf{g}'(\mathbf{x}_k)^\mathsf{T}\,\mathbf{g}'(\mathbf{x}_k)\right)^{-1} f'(\mathbf{x}_k)^\mathsf{T}\,.
\end{aligned}
$$

Nelineární metoda nejmenších čtverců jako optimalizační úloha

Iterace Gaussovy-Newtonovy metody

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \left(2\,\mathbf{g}'(\mathbf{x}_k)^\mathsf{T}\,\mathbf{g}'(\mathbf{x}_k)\right)^{-1} f'(\mathbf{x}_k)^\mathsf{T}\,.$$

Pro srovnání:

Iterace gradientní metody

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k\,\mathbf{I}\,f'(\mathbf{x}_k)^\mathsf{T}\,.$$

Levenbergova-Marquardtova metoda řešení soustavy rovnic

Iterace Levenbergovy-Marquardtovy metody

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

Iterace Gaussovy-Newtonovy metody

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

Iterace gradientní metody

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k\,\mathbf{I}\,f'(\mathbf{x}_k)^\mathsf{T} = \mathbf{x}_k - \left(\tfrac{1}{\alpha_k}\,\mathbf{I}\right)^{-1} f'(\mathbf{x}_k)^\mathsf{T}\,.$$

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

- Gaussovou-Newtonovou metodou (pro $\mu_k \to 0$),
- gradientní metodou (pro $\mu_k \to \infty$ je $\alpha_k \to \frac{1}{\mu_k}$).

Levenbergova-Marquardtova metoda (pokračování)

Zvolíme konstantu $q > 1$, např. 2, začneme velkou hodnotou μ_0,

- když se účelová funkce sníží, krok přijmeme a $\mu_{k+1} := \mu_k/q$,
- když se účelová funkce zvýší, krok odmítneme a $\mu_{k+1} := \mu_k\, q$.

Reference

- T. Werner. *Optimalizace* (kapitola 8). Elektronická skripta. FEL ČVUT, 17. 10. 2022. https://cw.fel.cvut.cz/wiki/_media/courses/b0b33opt/opt.pdf
- J. Hamhalter a J. Tišer. *Diferenciální počet funkcí více proměnných*. FEL ČVUT, 2005. https://math.fel.cvut.cz/en/people/tiser/dipocet.pdf
- G. Goh. *Why Momentum Really Works.* https://distill.pub/2017/momentum/
- *Matlab Activities for Multivariable Calculus.* https://mse.redwoods.edu/darnold/math50c/matlab/
- M. Navara. *Matematika spojitého světa*. Elektronická skripta. FEL ČVUT, 2011. https://cmp.felk.cvut.cz/~navara/MSS/MSS_print.pdf