Optimalizace

Metody hledání volných lokálních extrémů

Tomáš Kroupa, (Mirko Navara)

28. 3. 2023

FEL ČVUT

Obsah

Podmínka prvního řádu

Stacionární bod $\mathbf{x} \in \mathbb{R}^n$ funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ splňuje $f'(\mathbf{x}) = \mathbf{0}$.

Věta

Nechť $X \subseteq \mathbb{R}^n$, $\mathbf{x} \in X$ je vnitřní bod množiny X a $f\colon \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná v $\mathbf{x}$. Je-li $\mathbf{x}$ lokální extrém funkce f na X, platí

$$f'(\mathbf{x}) = \mathbf{0}\,,$$

tj. $\mathbf{x}$ je stacionární bod funkce f.

Opakování: Taylorův polynom pro funkci $f\colon \mathbb{R}^n \to \mathbb{R}$

Polynom k-tého stupně $T_k\colon \mathbb{R}^n \to \mathbb{R}$, který má v bodě $\mathbf{x}$ všechny parciální derivace až do řádu k stejné jako funkce f.

Taylorovy polynomy v bodě $\mathbf{x}$ do stupně 2

$$T_0(\mathbf{y}) = f(\mathbf{x})$$

$$T_1(\mathbf{y}) = f(\mathbf{x}) + f'(\mathbf{x})\,(\mathbf{y} - \mathbf{x})$$

$$T_2(\mathbf{y}) = f(\mathbf{x}) + f'(\mathbf{x})\,(\mathbf{y} - \mathbf{x}) + \frac{1}{2}\,(\mathbf{y} - \mathbf{x})^\mathsf{T}\, f''(\mathbf{x})\,(\mathbf{y} - \mathbf{x})$$

Podmínky druhého řádu

Věta

Nechť $X \subseteq \mathbb{R}^n$, $\mathbf{x} \in X$ je vnitřní bod množiny X a $f\colon \mathbb{R}^n \to \mathbb{R}$ je dvakrát diferencovatelná v $\mathbf{x}$. Je-li $\mathbf{x}$ stacionární bod, platí:

- Pokud je $\mathbf{x}$ lokální minimum, pak je Hessova matice $f''(\mathbf{x})$ pozitivně semidefinitní.
- Pokud je $f''(\mathbf{x})$ pozitivně definitní, pak je $\mathbf{x}$ ostré lokální minimum.
- Pokud je $f''(\mathbf{x})$ indefinitní, pak $\mathbf{x}$ není lokální extrém, ale sedlový bod.

Iterační metody

Hledáme lokální minimum funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ pomocí konstrukce posloupnosti bodů $\mathbf{x}_0, \mathbf{x}_1, \ldots \in \mathbb{R}^n$. Volíme počáteční bod $\mathbf{x}_0 \in \mathbb{R}^n$, směr hledání $\mathbf{v}_k \in \mathbb{R}^n$ a parametr $\alpha_k > 0$, určující délku kroku $\alpha_k \|\mathbf{v}\|$.

Iterace

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{v}_k\,.$$

- Sestupná metoda splňuje $f(\mathbf{x}_{k+1}) < f(\mathbf{x}_k)$.
- Sestupný směr $\mathbf{v}_k$ splňuje $f'(\mathbf{x}_k)\,\mathbf{v}_k < 0$ (tj. $\nabla f(\mathbf{x}_k)^{\mathrm{T}}\,\mathbf{v}_k < 0$).
- Optimální parametr α_k lze nalézt minimalizací funkce

$$\varphi_k(\alpha) := f(\mathbf{x}_k + \alpha\,\mathbf{v}_k)\,.$$

Gradientní metoda

Iterace gradientní metody

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

- Směr $-\nabla f(\mathbf{x}_k)$ je sestupný,
- robustní metoda,
- může konvergovat velmi pomalu.

Newtonova metoda řešení soustavy rovnic

Hledáme řešení soustavy rovnic $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, kde $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^n$ je diferencovatelné zobrazení.

$$T_1(\mathbf{x}_{k+1}) = \mathbf{g}(\mathbf{x}_k) + \mathbf{g}'(\mathbf{x}_k)\,(\mathbf{x}_{k+1} - \mathbf{x}_k) = 0\,.$$

Iterace

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

- Jacobiho matice musí být regulární (raději dobře podmíněná),
- rychlá konvergence, ale nezaručená,
- nutno začít s dobrou počáteční aproximací řešení, $\mathbf{x}_0$.

Newtonova metoda minimalizace funkce

Hledáme možné lokální minimum dvakrát diferencovatelné funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ jako řešení rovnice $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, kde $\mathbf{g}(\mathbf{x}) := \nabla f(\mathbf{x}) = f'(\mathbf{x})^{\mathsf{T}}$.

Iterace Newtonovy metody

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

Iterace gradientní metody

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

- Newtonův směr $\mathbf{v}_k := -f''(\mathbf{x}_k)^{-1}\,f'(\mathbf{x}_k)^{\mathsf{T}}$ je sestupný, pokud $\mathbf{x}_k$ není stacionární bod a matice $f''(\mathbf{x}_k)$ je pozitivně definitní.
- Pro velká n je pracné počítat inverzi Hessovy matice $f''(\mathbf{x}_k)$.

Nelineární metoda nejmenších čtverců pro řešení soustavy rovnic

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.

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 bylo speciálním případem:

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

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

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

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

Levenbergova-Marquardtova metoda

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 Levenbergovy-Marquardtovy metody

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

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

- Gaussovou-Newtonovou metodou (μ_k je malé),
- gradientní metodou (μ_k je velké).

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