Optimalizace

Metody hledání volných lokálních extrémů

Tomáš Kroupa

FEL ČVUT

Obsah

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 délku kroku $\alpha_k > 0$.

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 $f'(\mathbf{x}_k)\mathbf{v}_k < 0$
- Optimální délku kroku α_k lze nalézt minimalizací funkce

$$\varphi(\alpha_k) := f(\mathbf{x}_k + \alpha_k \mathbf{v}_k)$$

Gradientní metoda

Iterace

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \nabla f(\mathbf{x}_k)$$

- Směr $-\nabla f(\mathbf{x}_k)$ je sestupný
- Robustní
- Může konvergovat velmi pomalu

Newtonova metoda

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

Iterace

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

- Jacobiho matice musí být regulární
- Rychlá konvergence
- Nutno začít s dobrou aproximací $\mathbf{x}_0$ řešení

Newtonova metoda na minimalizaci funkce

Hledáme lokální minimum dvakrát diferencovatelné funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ pomocí řešení rovnice $f'(\mathbf{x})^\mathsf{T} = \mathbf{0}$.

Iterace

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

- 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
- Čistá Newtonova metoda ($\alpha_k := 1$)
- Pro velká n je drahé počítat Hessián a jeho inverzi

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)

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/`