Optimalizace a teorie her

Konvexní funkce

Martin Bohata

Katedra matematiky
FEL ČVUT v Praze
bohata@math.feld.cvut.cz

Definice konvexní funkce

Definice

Nechť $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ a $C \subseteq D$ je neprázdná konvexní množina. Řekneme, že f je

1. konvexní na C, jestliže pro každé $x, y \in C$ a každé $\lambda \in [0, 1]$ je
$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$
2. ryze konvexní na C, jestliže pro každé dva různé body $x, y \in C$ a každé $\lambda \in (0, 1)$ je
$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y).$$
3. konkávní (resp. ryze konkávní) na C, jestliže $-f$ je konvexní (resp. ryze konvexní) na C.

Úmluva: U pojmů z předchozí definice budeme vynechávat „na C“, jestliže $C = D$.

Základní příklady

Příklad

Z úvodního kurzu matematické analýzy víme, že

1. $f(x) = x^2$ je ryze konvexní;
2. $f(x) = x^3$ není konvexní ani konkávní. Avšak je ryze konvexní na $[0, \infty)$ a ryze konkávní na $(-\infty, 0]$;
3. $f(x) = e^x$ je ryze konvexní;
4. $f(x) = \ln x$ je ryze konkávní;
5. $f(x) = \sqrt{x}$ je ryze konkávní.

Příklad

1. afinní zobrazení $f : \mathbb{R}^n \to \mathbb{R}$ (tj. $f(x) = \langle x, a \rangle + b$, kde $a \in \mathbb{R}^n$ a $b \in \mathbb{R}$) je konvexní a také konkávní.
2. funkce $f(x) = \|x\|$ je konvexní.

Dolní úrovňová množina

Definice

Dolní úrovňová množina funkce $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ hladiny $\alpha \in \mathbb{R}$ je množina

$$\operatorname{lev}_{\leq}(f; \alpha) := \{x \in D \,|\, f(x) \leq \alpha\}\,.$$

Tvrzení

Je-li f konvexní na $C \subseteq \mathbb{R}^n$, potom $\operatorname{lev}_{\leq}(f|_C; \alpha)$ je konvexní pro každé $\alpha \in \mathbb{R}$.

Důkaz: Viz přednáška. ■

- Opačné tvrzení neplatí (viz např. $f(x) = x^3$).

Příklad

Množina $M = \left\{ x \in \mathbb{R}^2 \,\middle|\, \|x\| \leq 1, \left\langle x, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\rangle \leq 1 \right\}$. je konvexní.

Operace zachovávající konvexitu

Tvrzení

Jestliže f, g jsou konvexní funkce na $C \subseteq \mathbb{R}^n$ a $\alpha \geq 0$, potom $f + g$ a αf jsou konvexní na C.

Důkaz: Viz přednáška. ■

Příklad

Funkce

$$f(x) = e^x - 3\ln x + 2x$$

je konvexní.

- Součin konvexních funkcí již není obecně konvexní funkce ($f(x) = x$ a $g(x) = -x$ jsou konvexní, ale $(fg)(x) = -x^2$ není konvexní).

Operace zachovávající konvexitu

Tvrzení

Nechť f je konvexní na $K \subseteq \mathbb{R}^m$, $C \subseteq \mathbb{R}^n$ je neprázdná konvexní a $g : \mathbb{R}^n \to \mathbb{R}^m$ je afinní. Jestliže $g(C) \subseteq K$, pak $f \circ g$ je konvexní na C.

Důkaz: Viz přednáška. ∎

Příklad

1. Funkce $f(x) = \|Ax + b\|$, kde $A \in \mathbb{M}_{m,n}(\mathbb{R})$ a $b \in \mathbb{R}^m$, je konvexní.
2. Funkce $f(x) = e^{x_1 - 2x_2 + 3x_3} + e^{2x_1} - x_3$ je konvexní.

- Skládání konvexních funkcí není obecně konvexní funkce. Funkce $f(x) = x^2$ a $g(x) = x^2 - 1$ jsou konvexní, ale

$$(f \circ g)(x) = (f(g(x)) = (x^2 - 1)^2$$

není konvexní.

Operace zachovávající konvexitu

Tvrzení

Jestliže f je konvexní a neklesající na intervalu I, g je konvexní na $C \subseteq \mathbb{R}^n$ a $g(C) \subseteq I$, pak $f \circ g$ je konvexní na C.

Důkaz: Viz přednáška. ■

Příklad

1. Funkce $f(x) = \|x\|^2$ je konvexní.
2. Funkce $f(x) = e^{\|x\|}$ je konvexní.

Body minima konvexních funkcí

Věta (O extrémech a konvexitě)

Nechť f je konvexní na $C \subseteq \mathbb{R}^n$. Potom platí:

1. *Každý bod lokálního minima f na C je bodem minima f na C.*
2. *Množina $\operatorname{argmin}_{x \in C} f(x)$ je konvexní. Je-li navíc f ryze konvexní na C, pak existuje nejvýše jeden bod minima funkce f na C.*

Důkaz: Viz přednáška. ∎

Příklad

Nechť $f(x) = \max\{1, \|x\|\}$. V souladu s předchozí větou je

$$\operatorname*{argmin}_{x \in \mathbb{R}^n} f(x) = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$$

konvexní množina. (Navíc odtud vidíme, že f nemůže být ryze konvexní.)

Diferencovatelné funkce a konvexita

Věta (O konvexitě a první derivaci)

Nechť $\Omega \subseteq \mathbb{R}^n$ je otevřená, $C \subseteq \Omega$ neprázdná konvexní a $f \in C^1(\Omega)$. Potom platí:

1. *f je konvexní na C právě tehdy, když pro každé $x, y \in C$ je*
$$f(x) + \langle \nabla f(x), y - x \rangle \leq f(y).$$
2. *f je ryze konvexní na C právě tehdy, když pro každé dva různé body $x, y \in C$ je*
$$f(x) + \langle \nabla f(x), y - x \rangle < f(y).$$

Důkaz: Důkaz 1 viz přednáška. Důkaz 2 vynecháváme. ∎

- Geometrická interpretace: graf leží nad tečnou nadrovinou.

Diferencovatelné funkce a konvexita

Věta (O konvexitě a druhé derivaci)

Nechť $\Omega \subseteq \mathbb{R}^n$ *je otevřená,* $C \subseteq \Omega$ *neprázdná konvexní a* $f \in C^2(\Omega)$*. Potom platí:*

1. *Jestliže pro každé* $x \in C$ *je* $\nabla^2 f(x)$ *pozitivně semidefinitní matice, pak* f *je konvexní na* C*.*
2. *Jestliže* f *je konvexní na* C *a* C *je otevřená, potom* $\nabla^2 f(x)$ *je pozitivně semidefinitní matice pro každé* $x \in C$*.*
3. *Jestliže pro každé* $x \in C$ *je* $\nabla^2 f(x)$ *pozitivně definitní matice, pak* f *je ryze konvexní na* C*.*

Důkaz: Viz přednáška. ■

Příklad

Funkce $f(x_1, x_2) = x_1^2 + x_1 x_2 + x_2^2$ je ryze konvexní.