ФУНКЦИЈЕ ВИШЕ ПРОМЕНЉИВИХ

Конвексност и конкавност функција више променљивих

Драган Ђорић

За реалне функције једне реалне променљиве конвексност и конкавност функције $f : (a, b) \to \mathbb{R}$ дефинисали смо односом графика функције f и тангенти на график у тачкама $x \in (a, b)$ [2]. Функција f је *конвексна* на (a, b) ако график функције нигде није испод било које тангенте на график у интервалу (a, b).

Овом услову еквивалентан је одговарајући однос графика функције и сечице на сваком одсечку $[x, y]$ за $a < x < y < b$. Функција је конвексна на (a, b) ако и само ако график те функције нигде није изнад одсечка $[x, y]$.

Ови услови могу да се уведу и за функције више променљивих.

1 Дефиниција конвексне и конкавне функције

Нека је функција $f : D \to \mathbb{R}$ дефинисана на конвексном скупу $D \subset \mathbb{R}^n$.

Дефиниција 1. *Функција f је конвексна на скупу D ако за свако $x, y \in D$ и свако $\lambda \in [0, 1]$ важи*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y). \tag{1}$$

Ако за $x \neq y$ и $\lambda \in (0, 1)$ у (1) важи строга неједнакост, функција f је строго конвексна *на скупу D. Функција f је конкавна на (a, b) ако је функција $-f$ конвексна на том интервалу.*

Пример 1. *Функција $f : \mathbb{R}^n \to \mathbb{R}$ дефинисана са $f(x) = |x - a|$, где је a нека тачка из R^n и где је $|x|$ норма вектора $x \in \mathbb{R}^n$, је конвесна на $\mathbb{R}^n$ јер је*

$$|\lambda x + (1 - \lambda)y - a| = |\lambda(x - a) + (1 - \lambda)(y - a)| \leq \lambda|x - a| + (1 - \lambda)|y - a|$$

за свако $x, y \in \mathbb{R}^n$ и свако $\lambda \in [0, 1]$.

2 Конвексност диференцијабилне функције

Диференцијабилне функције имају и друге карактеризације ковексности и конкавности. У даљем тексту за векторе $a = (a_1, a_2, \ldots, a_n)$ и $b = (b_1, b_2, \ldots, b_n)$ из $\mathbb{R}^n$ ознака (a, b) се користи за *скаларни производ* дефинисан са

$$(a, b) = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

Следећа теорема даје потребан и довољан услов за конвексност диференцијабилне функције на $\mathbb{R}^n$.

Теорема 1. *Функција $f : \mathbb{R}^n \to \mathbb{R}$ је конвексна на $\mathbb{R}^n$ ако и само ако је*

$$f(y) \geq f(x) + (\nabla f(x), y - x) \tag{2}$$

за све $x, y \in \mathbb{R}^n$.

Доказ. [`Потребност`] Претпоставимо да је функција f конвексна. Из неједнакости

$$f((1-\lambda)x+\lambda y) \leq (1-\lambda)f(x)+\lambda f(y)$$

за $\lambda \in (0,1]$ следи да је

$$\frac{f(x+\lambda(y-x))-f(x)}{\lambda} \leq f(y)-f(x). \tag{3}$$

За $\lambda \to 0$ на левој страни претходне неједнакости имамо извод функције f у тачки x у правцу вектора $y-x$. Како за диференцијабилне функције извод у правцу датог вектора v може да се изрази као скаларни производ градијента функција у посматраној тачки и датог вектора v (видети [1]), то из (3) закључујемо да је

$$(\nabla f(x), y-x) \leq f(y)-f(x).$$

Из ове неједнакости следи неједнакост (2) коју је и требало доказати.

[`Довољност`] Претпоставимо сада да важи неједнакост (2). Из те неједнакости имамо да је

$$f(z) \geq f(x)+(\nabla f(x), z-x), \quad f(w) \geq f(x)+(\nabla f(x), w-x). \tag{4}$$

За $x=(1-\lambda)z+\lambda w$ важи да је

$$z-x=\lambda(z-w), \quad w-x=(1-\lambda)(w-z),$$

па множењем прве неједнакости у (4) са $1-\lambda$ и друге са λ, добијамо да је

$$\begin{aligned}\lambda f(w)+(1-\lambda)f(z) \geq & f(x)+\lambda(\nabla f(x),(1-\lambda)(w-z))+(1-\lambda)(\nabla f(x), \lambda(z-w)) \\ =& f(x) \\ =& f(\lambda w+(1-\lambda)z).\end{aligned}$$

Како су z и w произвољне тачке из $\mathbb{R}^n$, функција f је конвексна у $\mathbb{R}^n$. ■

Из ове теореме следи да је функција f конкавна на $\mathbb{R}^n$ ако и само ако је

$$f(y) \leq f(x)+(\nabla f(x), y-x) \tag{5}$$

На исти начин се доказује да је функција f строго конвексна ако и само ако у (2) важи строга неједнакост за $x \neq y$, односно да је f конкавна ако у (5) важи строга неједнакост за $x \neq y$.

У случају конвексних и конкавних функција потребан услов првог реда за локални екстремум постаје и довољан услов.

Теорема 2. *Ако је функција $f:\mathbb{R}^n \to \mathbb{R}$ диференцијабилна и конвексна на R^n, тада она у тачки $a \in \mathbb{R}^n$ има глобални минимум ако и само ако је $df(A)=0$.*

Доказ. [`Потребност`] Претпоставимо да функција f има у тачки a глобални минимум. Како је функција диференцијабилна, према Фермаовој теореми имамо да је $df(a)=0$.

[`Довољност`] Претпоставимо да је $df(a)=0$. На основу Теореме 1 имамо да је

$$f(x) \geq f(a)+(\nabla f(a), x-a)=f(a)$$

за свако $x \in \mathbb{R}^n$. Према томе, функција f у тачки a има глобални минимум. ■

Ако је f конкавна функција на $\mathbb{R}^n$, тада применом претходне теореме на функцију $-f$ добијамо да важи и следеће тврђење.

Теорема 3. *Ако је функција $f:\mathbb{R}^n \to \mathbb{R}$ диференцијабилна и конкавна на R^n, тада она у тачки $a \in \mathbb{R}^n$ има глобални максимум ако и само ако је $df(A)=0$.*

3 Конвексност два пута диференцијабилне функције

За функције једне промењљиве конвексност и конкавност одређују изводи другог реда. Код функција више промењљивих улогу извода другог реда имају диференцијали другог реда.

Претпоставимо сада да је функција $f : \mathbb{R}^n \to \mathbb{R}$ два пута диференцијабилна на $\mathbb{R}^n$. Тада је конвексност и конкавност одређена диференцијалом другог реда, односно Хесеовом матрицом.

Теорема 4. *Два пута диференцијабилна функција $f : \mathbb{R}^n \to \mathbb{R}$ је конвексна ако и само ако је $d^2 f(x) \geq 0$ за свако $x \in \mathbb{R}^n$.*

Доказ. [`Потребност`] Претпоставимо да је функција f конвексна на $\mathbb{R}^n$ и да је $x \in \mathbb{R}^n$. Функција $g : \mathbb{R}^n \to \mathbb{R}$ дефинисана са

$$g(y) = f(y) - (\nabla f(x), y - x)$$

је такође конвексна јер је $y \mapsto (\nabla f(x), y - x)$ конвексна функција. Поред тога, за функцију g за свако $y \in \mathbb{R}^n$ важе и једнакости

$$\nabla g(y) = \nabla f(y) - \nabla f(x), \quad H_g(y) = H_f(y), \tag{6}$$

где су $H_g(y)$ и $H_f(y)$ Хесеове матрице функција g и f у тачки y.

Како из прве једнакости у (6) следи да је $\nabla g(x) = 0$, према Теореми 2 у тачки x је глобални минимум функције g. Из потребног услова другог реда (видети прилог [3]) имамо да је $d^2 g(x) \geq 0$. Сада из друге једнакости у (6) следи да је $d^2 f(x) = 0$.

[`Довољност` Претпоставимо да је $d^2 f(x) \geq 0$ за свако $x \in \mathbb{R}^n$. За $x, y \in \mathbb{R}^n$ из Тејлорове формуле имамо да је

$$f(y) = f(x) + df(x) + \frac{1}{2} d^2 f(x + t(y - x)) \geq f(x) + df(x) = f(x) + (\nabla f(x), y - x)$$

за неко $t \in [0, 1]$. Према Теореми 1 то значи да је функција f конвексна ■.

Из ове теореме и Силвестеровог критеријума следи да је два пута диференцијабилна функција $f : \mathbb{R}^n \to \mathbb{R}$ конвексна ако и само ако је Хесијан $H_f(x)$ за свако $x \in \mathbb{R}^n$ ненегативно дефинитна матрица.

Што се тиче строге конвексности, она следи из претпоставке да је $d^2 f$ позитивно дефинитна форма у свакој тачки, односно да је Хесијан H_f позитивно дефинитна матрица у свакој тачки. Међутим, као и код функција једне промењљиве, обратно не важи. На пример, функција $f : (x, y) \mapsto x^4 + y^4$ је строго конвексна на $\mathbb{R}^2$, а $d^2 f(x, y)$ није позитивно дефинитна форма јер је $d^2 f(0, 0) = 0$.

За конкавне функције више промењљивих важи аналогно тврђење и одговарајући коментари.

Теорема 5. *Два пута диференцијабилна функција $f : \mathbb{R}^n \to \mathbb{R}$ је конкавна ако и само ако је $d^2 f(x) \leq 0$ за свако $x \in \mathbb{R}^n$.*

Наведена тврђења су формулсана за функције које су дефинисане и испуњавају одређене услове на $\mathbb{R}^n$, а слично важи и за функције које су дефинисане на неком конвексном подскупу од $\mathbb{R}^n$.

Задаци

1. Доказати да је збир две конвексне функције $f, g : \mathbb{R}^n \to \mathbb{R}$ такође конвексна функција.

2. Доказати да је прозвод конвексне функције $f : \mathbb{R}^n \to \mathbb{R}$ и реалнпг броја $c \geq 0$ такође конвексна функција.

3. Доказати да је за ковексне функције $f, g : \mathbb{R}^n \to \mathbb{R}$ функција $h : (x, y) \mapsto \max\{f(x), g(x)\}$ такође конвексна.

4. Нека је функција f конвексна и ненегативна на скупу $X \subset \mathbb{R}^2$. Доказати да је и функција f^2 такође конвексна на скупу X.

5. Нека је функција f конвексна на $\mathbb{R}^n$ и нека је $a, b \in \mathbb{R}^n$, Доказати да је функција $G : \mathbb{R} \to \mathbb{R}$ дефинисана са $g(t) = f(a + tb)$ конвексна на $\mathbb{R}$.

6. Нека је f конвексна функција на скупу $X \subset \mathbb{R}^n$. Доказати да за функцију f важи *Јенсенова неједнакост*

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i),$$

где је $x_1, x_2, \ldots, x_n \in X$ и где су $\lambda_1, \lambda_2, \ldots, \lambda_n$ ненегативни реални бројеви, такви да је $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$.

У следећим задацима испитати конвексност дате функције f на скупу $\mathbb{R}^2$.

7. $f(x, y) = (x^2 - y)^2$.

8. $f(x, y) = e^{2x+y}$.

9. $f(x, y) = xe^{-x-y}$.

10. $f(x, y) = \sqrt{1 + x^2 + y^2}$.

У следећим задацима испитати конвексност дате функције f на скупу $\mathbb{R}^3$.

11. $f(x, y, z) = x^2 + y^2 + \dfrac{1}{2}z^2 + xy - z + 10$.

12. $f(x, y, z) = -x^2 - y^2 - 2z^2 + xy + xz + yz + 5y$.

13. $f(x, y, z) = 5x^4 + y^6 + z^3 - 13x + 7y - 8$.

Литература

[1] Стојановић, М., Михић, О., *Математика 2*, ФОН, Београд, 2013.

[2] Ђорић, Д., Лазовић, Р., *Математика 1*, ФОН, Београд, 2014

[3] Ђорић, *Потребни услови другог реда за локални екстремум функције више промењљивих*, http://math.fon.rs/matematika-dva