KONVEX FÜGGVÉNY
KVÁZIKONVEX FÜGGVÉNY

DR. NAGY TAMÁS
egyetemi docens

Miskolci Egyetem
Alkalmazott Matematikai Tanszék

„A bemutatott kutató munka a TÁMOP-4.2.1.B-10/2/KONV-2010-0001 jelű projekt részeként az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg”
„This research was carried out as part of the TAMOP-4.2.1.B-10/2/KONV-2010-0001 project with support by the European Union, co-financed by the European Social Fund.”

Miskolc, 2012

Tartalomjegyzék

1. Bevezetés

A konvex függvények, hasonlóan a konvex halmazokhoz, nagy jelentőséggel bírnak az optimalizálás területén.

Mielőtt a konvex függvények vizsgálatára rátérnénk, ismertetünk két fogalmat.

A tananyagban sok tételben feltesszük a függvények differenciálhatóságát, tehát használjuk a gradiens vektor és a Hesse mátrix fogalmakat. A többváltozós függvények differenciálhatóságának kérdésével, így ezekkel a fogalmakkal is egy másik tananyag foglalkozik.

1.1. Gráf, epigráf, hipergráf fogalma

Legyen $S \subseteq \mathbb{R}^n$ egy nemüres halmaz és legyen $f : S \to \mathbb{R}$ függvény.

Az f függvény gráfjának (jelölése: $gr\ f$) az alábbi módon definiált $\mathbb{R}^{n+1}$-beli halmazt nevezzük:

$$gr\ f = \left\{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : \mathbf{x} \in S, y \in \mathbb{R}, y = f(\mathbf{x})\right\}.$$

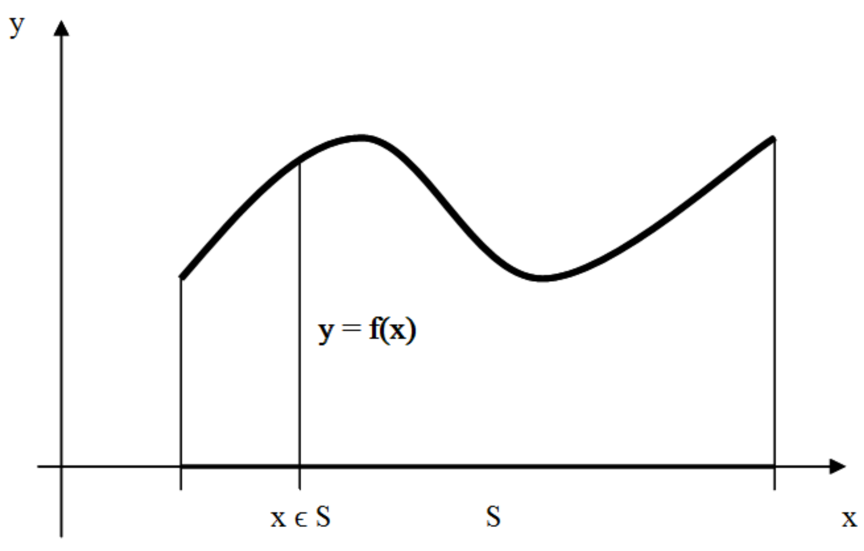
Az f függvény epigráfjának (jelölése: $epi\ f$) az alábbi módon definiált $\mathbb{R}^{n+1}$-beli halmazt nevezzük:

$$epi\ f = \left\{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : \mathbf{x} \in S, y \in \mathbb{R}, y \geqq f(\mathbf{x})\right\}.$$

Az f függvény hypográfjának (jelölése: $hip\ f$) az alábbi módon definiált $\mathbb{R}^{n+1}$-beli halmazt nevezzük:

$$hyp\ f = \left\{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : \mathbf{x} \in S, y \in \mathbb{R}, y \leqq f(\mathbf{x})\right\}.$$

Az egyváltozós esetben ezek a halmazok jól ismertek. Az alábbi ábrán a vastag görbe a függvény képét mutatja. A görbe pontjai és a görbe feletti pontok halmaza az epigráf. A görbe pontjai és a görbe alatti pontok halmaza a hypográf.

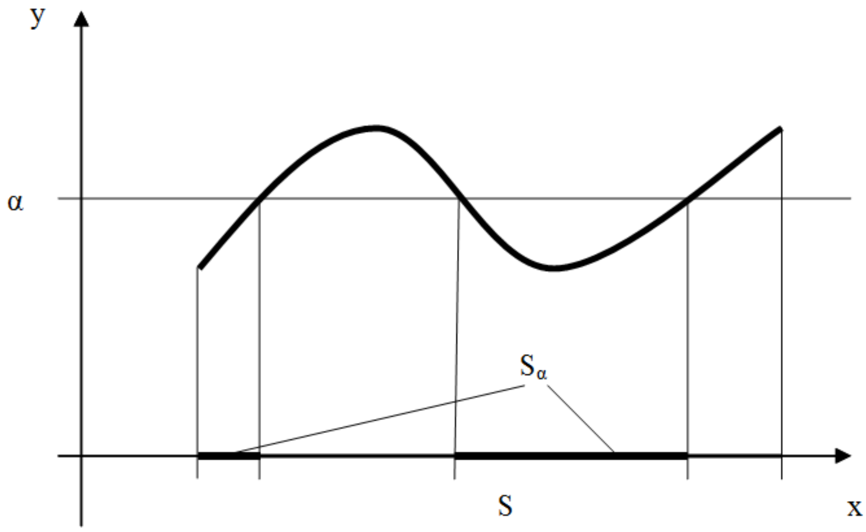


1.2. Alsó nívóhalmaz fogalma

Legyen $S \subseteq \mathbb{R}^n$ egy nemüres halmaz, $f : S \to \mathbb{R}$ függvény és $\alpha \in \mathbb{R}$ valós szám. Ekkor az $S_\alpha = \{\mathbf{x} \in S : f(\mathbf{x}) \leqq \alpha\}$ halmazt az f függvény alsó nívóhalmazának nevezzük.

Az alsó nívóhalmaz tehát mindazon S-beli pontok összessége, amelyekben a függvény értéke egy megadott α szám vagy ennél kisebb. Hasonlóan definiálhatjuk felső nívóhalmazt is.

Az alábbi ábra az S_α alsó nívóhalmazt szemlélteti.



2. Konvex függvény

2.1. Konvex függvény definiciója

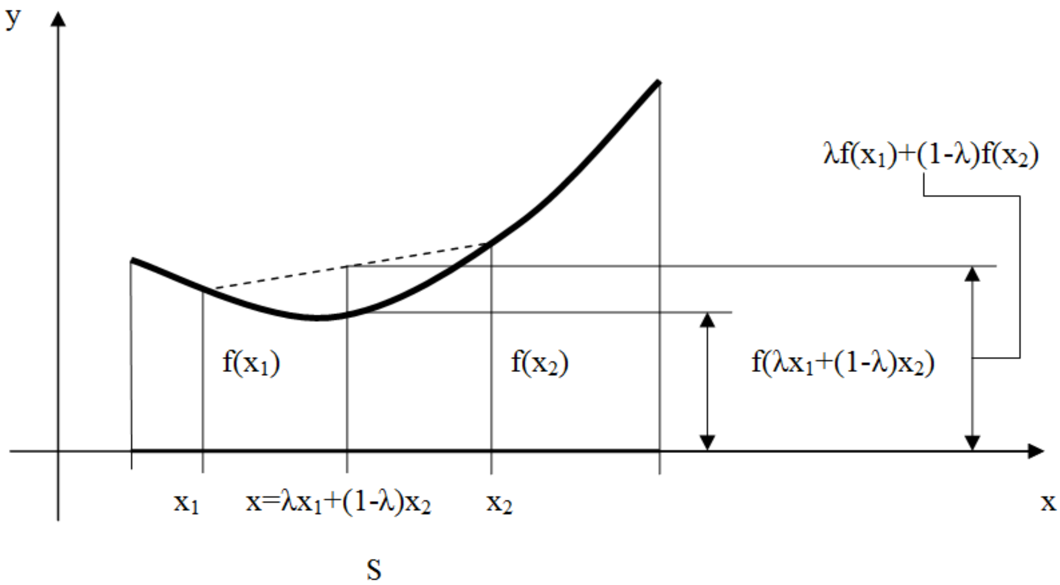
Legyen $f : S \to \mathbb{R}$, ahol $S \subseteq \mathbb{R}^n$ egy nemüres konvex halmaz. Az $f(\mathbf{x})$ függvényt akkor mondjuk konvexnek az S halmazon, ha

$$f(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \leqq \lambda f(\mathbf{x}_1) + (1-\lambda) f(\mathbf{x}_2)$$

bármely $\mathbf{x}_1$, $\mathbf{x}_2 \in S$ vektorra és minden $\lambda \in (0,1)$ számra.

Az f függvény szigorúan konvex, ha a definíciós formulában szigorú egyenlőtlenség áll fenn.

Az alábbi ábra egy egyváltozós konvex függvényt szemléltet.



A definíció szemléletesen azt fejezi ki, hogy konvex függvénynél bármely két pont esetén a függvénypontokat összekötő húr (szaggatott vonal) alatt halad a függvény görbéje, legalábbis nem megy fölé. Többváltozós esetben is ugyanez mondható el, csak ott $n = 2$ esetén tudunk szemléletet adni.

A konvex függvény mellett a gyakorlatban szintén nagy jelentősége van a konkáv függvénynek, amelyet az alábbiakban definiálunk.

Legyen $f : S \to \mathbb{R}$, ahol $S \subseteq \mathbb{R}^n$ egy nemüres konvex halmaz. Az $f(\mathbf{x})$ függvényt akkor mondjuk konkávnak az S halmazon, ha

$$f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \geqq \lambda f(\mathbf{x}_1) + (1-\lambda)f(\mathbf{x}_2)$$

bármely $\mathbf{x}_1$, $\mathbf{x}_2 \in S$ vektorra és minden $\lambda \in (0, 1)$ számra.

Az f függvény szigorúan konkáv, ha a definíciós formulában szigorú egyenlőtlenség áll fenn.

Sokszor megelégszünk a konkávitásnak a fentiekkel azonos eredményt adó definíciójával is, amely szerint az f függvényt (szigorúan) konkávnak nevezzük, ha a $-f$ függvény (szigorúan) konvex.

2.2. Alapvető tételek konvex függvényekre

TÉTEL (konvex függvény belső pontban folytonos)

Legyen $S \subseteq \mathbb{R}^n$ nemüres konvex halmaz és legyen $f : S \to \mathbb{R}$ konvex függvény. Ekkor az f függvény folytonos az S halmaz minden belső pontjában.

TÉTEL (konvex függvények nemnegatív kombinációja és maximuma konvex függvény)

Legyen $S \subseteq \mathbb{R}^n$ nemüres konvex halmaz és legyenek $f_1, f_2, \ldots, f_k : S \to \mathbb{R}$ konvex függvények, $\lambda_1, \lambda_2, \ldots, \lambda_k \geqq 0$ valós számok. Ekkor az

$$\begin{aligned} f(\mathbf{x}) &= \sum_{i=1}^{k} \lambda_i f_i(\mathbf{x}), \\ f(\mathbf{x}) &= \max\{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x})\} \end{aligned}$$

függvények szintén konvex függvények.

TÉTEL (Jensen egyenlőtlenség)
Legyen $S \subseteq \mathbb{R}^n$ nemüres konvex halmaz, $f : S \to \mathbb{R}$ konvex függvény, továbbá legyenek $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in S$, $\lambda_1, \lambda_2, \dots, \lambda_k \geqq 0$, $\sum_{i=1}^{k} \lambda_i = 1$. Ekkor igaz az alábbi egyenlőtlenség:

$$f\left(\sum_{i=1}^{k} \lambda_i \mathbf{x}_i\right) \leqq \sum_{i=1}^{k} \lambda_i f(\mathbf{x}_i).$$

TÉTEL (alsó nívóhalmaz konvex)
Legyen $S \subseteq \mathbb{R}^n$ nemüres konvex halmaz, $f : S \to \mathbb{R}$ konvex függvény és $\alpha \in \mathbb{R}$ valós szám. Ekkor az $S_\alpha = \{\mathbf{x} \in S : f(\mathbf{x}) \leqq \alpha\}$ ún. alsó nívóhalmaz konvex halmaz.

TÉTEL (alsó nívóhalmazok metszete konvex)
Legyen $S \subseteq \mathbb{R}^n$ nemüres konvex halmaz, legyenek $f_1, f_2, \dots, f_k : S \to \mathbb{R}$ konvex függvények, $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}$ valós számok. Ekkor az

$$S_\alpha = \{\mathbf{x} : \mathbf{x} \in S,\ f_i(\mathbf{x}) \leqq \alpha_i, i = 1, 2, \dots, k\}$$

halmaz konvex.

TÉTEL (iránymenti derivált létezése)
Legyen $f : \mathbb{R}^n \to \mathbb{R}$ konvex függvény. Tekintsünk egy tetszőleges $\bar{\mathbf{x}} \in \mathbb{R}^n$ pontot és egy tetszőleges $\mathbf{d} \in \mathbb{R}^n, \mathbf{d} \neq \mathbf{0}$ irányvektort. Ekkor az f függvénynek az $\bar{\mathbf{x}}$ pontban a $\mathbf{d}$ irány mentén vett $f'(\bar{\mathbf{x}}, \mathbf{d})$ iránymenti deriváltja létezik.

2.3. Karakterizációs tételek konvex függvényekre

Most néhány karakterizációs tétel következik. Ezeknek a tételeknek a segítségével fogjuk eldönteni egy fügyvény konvex vagy konkáv voltát.

TÉTEL (karakterizációs tétel, egyváltozós függvény konvex)
Legyen $S \subseteq \mathbb{R}^n$ nemüres konvex halmaz és legyen $f : S \to \mathbb{R}$ függvény. Legyen $\mathbf{x} \in S$ és legyen $\mathbf{d}$ olyan, hogy $\mathbf{x} + \mathbf{d} \in S$. Legyen $\varphi(\lambda) = f(\mathbf{x} + \lambda \mathbf{d})$, ahol $0 \leqq \lambda \leqq 1$. Az $f(\mathbf{x})$ függvény akkor és csak akkor konvex, ha a $\varphi(\lambda)$ egyváltozós függvény konvex.

TÉTEL (karakterizációs tétel konvex függvényre, epigráf konvex)
Legyen $S \subseteq \mathbb{R}^n$ nemüres konvex halmaz és legyen $f : S \to \mathbb{R}$ függvény. Ekkor az f függvény akkor és csak akkor konvex, ha az epigráfja konvex halmaz.

TÉTEL (karakterizációs tétel konkáv függvényre, hypográf konvex)
Legyen $S \subseteq \mathbb{R}^n$ nemüres konvex halmaz és legyen $f : S \to \mathbb{R}$ függvény. Ekkor az f függvény akkor és csak akkor konkáv, ha a hypográfja konvex halmaz.

TÉTEL (1. karakterizációs tétel differenciálható konvex függvényre)

Legyen $S \subseteq \mathbb{R}^n$ nemüres nyílt konvex halmaz és legyen $f : S \to \mathbb{R}$ differenciálható függvény az S halmazon. Ekkor az f függvény akkor és csak akkor konvex, ha bármely $\bar{\mathbf{x}} \in S$ pontra

$$f(\mathbf{x}) \geqq f(\bar{\mathbf{x}}) + \boldsymbol{\nabla} f(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}) \quad \text{minden } \mathbf{x} \in S \text{ esetén.}$$

Ez a fontos tétel azt fejezi ki, hogy a konvex függvény mindig az érintője felett halad.

Bizonyítás:

i) Először induljunk ki abból, hogy az $f(\mathbf{x})$ függvény konvex, ekkor a definíció szerint

$$f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \leqq \lambda f(\mathbf{x}_1) + (1-\lambda) f(\mathbf{x}_2)$$

Rendezzük az egyenlőtlenséget az alábbiak szerint

$$\frac{f(\mathbf{x}_2 + \lambda(\mathbf{x}_1 - \mathbf{x}_2)) - f(\mathbf{x}_2)}{\lambda} \leqq f(\mathbf{x}_1) - f(\mathbf{x}_2)$$

A $\lambda \to 0$ határátmenetet képezve a bal oldal az $\mathbf{x}_2$ pontban az $\mathbf{x}_1 - \mathbf{x}_2$ irányban vett iránymenti deriválت, amelyről tudjuk, hogy differenciálhatóság esetén $\boldsymbol{\nabla} f(\mathbf{x}_2)(\mathbf{x}_1 - \mathbf{x}_2)$. Ezt figyelembe véve kapjuk, hogy

$$\boldsymbol{\nabla} f(\mathbf{x}_2)(\mathbf{x}_1 - \mathbf{x}_2) \leqq f(\mathbf{x}_1) - f(\mathbf{x}_2),$$

amelyből rendezés és a változók átnevezése után ($\mathbf{x}_2 = \bar{\mathbf{x}}$, $\mathbf{x}_1 = \mathbf{x}$) a bizonyítandó összefüggés adódik.

ii) Másodszor induljunk ki az $f(\mathbf{x})$ függvényre adott $f(\mathbf{x}) \geqq f(\bar{\mathbf{x}}) + \boldsymbol{\nabla} f(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})$ egyenlőtlenségből. Írjunk az $\bar{\mathbf{x}}$ helyébe $\mathbf{x}$-et, $\mathbf{x}$ helyébe pedig $\mathbf{x}_1$, ill. $\mathbf{x}_2$-t, ekkor kapjuk, hogy

$$\begin{aligned} f(\mathbf{x}_1) &\geqq f(\mathbf{x}) + \boldsymbol{\nabla} f(\mathbf{x})(\mathbf{x}_1 - \mathbf{x}) \\ f(\mathbf{x}_2) &\geqq f(\mathbf{x}) + \boldsymbol{\nabla} f(\mathbf{x})(\mathbf{x}_2 - \mathbf{x}) \end{aligned}$$

Szorozzuk az első összefüggést λ-val, a másodikat $(1-\lambda)$-val, majd a két egyenlőtlenséget adjuk össze, majd $\mathbf{x} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ összefüggésbőlaz kapjuk, hogy

$$\begin{aligned} \lambda f(\mathbf{x}_1) + (1-\lambda) f(\mathbf{x}_2) &\geqq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{x}) + \boldsymbol{\nabla} f(\mathbf{x}) \left(\lambda(\mathbf{x}_1 - \mathbf{x}) + (1-\lambda)(\mathbf{x}_2 - \mathbf{x})\right) = \\ &= f(\mathbf{x}) + \boldsymbol{\nabla} f(\mathbf{x})(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 - \mathbf{x}). \end{aligned}$$

Ez pedig az $\mathbf{x} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ választással a konvex függvény definíciójával azonos.

Q.e.d.

Példa:

Tekintsük az alábbi n változós függvényt. Állapítsuk meg, hogy a függvény konvex az $S = \mathbb{R}^n$ halmazon!

$$f(x_1, x_2, \ldots, x_n) = x_1^2 + x_2^2 + \ldots + x_n^2 = \sum_{i=1}^{n} x_i^2$$

Megoldás:

A függvény differenciálható, a gradiense az alábbi

$$\boldsymbol{\nabla} f(\mathbf{x}) = \begin{bmatrix} 2x_1 \\ 2x_2 \\ \vdots \\ 2x_n \end{bmatrix} = 2\mathbf{x}$$

A fenti karakterizációs tétel szerint, ha minden $\mathbf{x}, \bar{\mathbf{x}} \in \mathbb{R}^n$ esetén fennáll az alábbi összefüggés, akkor a függvény konvex. Esetünkben a megvizsgálandó összefüggés

$$\sum_{i=1}^{n} x_i^2 \geqq \sum_{i=1}^{n} \bar{x}_i^2 + 2\bar{\mathbf{x}}(\mathbf{x} - \bar{\mathbf{x}}),$$

elvégezve a kijelölt skaláris szorzást, adódik, hogy

$$\sum_{i=1}^{n} x_i^2 \geqq \sum_{i=1}^{n} \bar{x}_i^2 + 2\sum_{i=1}^{n} \bar{x}_i(x_i - \bar{x}_i) = \sum_{i=1}^{n} \bar{x}_i^2 + 2\sum_{i=1}^{n} \bar{x}_i x_i - 2\sum_{i=1}^{n} \bar{x}_i^2 = 2\sum_{i=1}^{n} \bar{x}_i x_i - \sum_{i=1}^{n} \bar{x}_i^2,$$

egy oldalra rendezve pedig kapjuk, hogy

$$\sum_{i=1}^{n} \left(x_i^2 - 2\bar{x}_i x_i + \bar{x}_i^2\right) = \sum_{i=1}^{n} (x_i - \bar{x}_i)^2 = \|\mathbf{x} - \bar{\mathbf{x}}\|^2 \geqq 0,$$

amely valóban minden vektorpárra igaz, így a karakterizációs tétel értelmében a függvény konvex.

TÉTEL (2. karakterizációs tétel differenciálható konvex függvényre)
Legyen $S \subseteq \mathbb{R}^n$ nemüres nyílt konvex halmaz és legyen $f : S \to \mathbb{R}$ differenciálható függvény az S halmazon. Ekkor az f függvény akkor és csak akkor konvex, ha bármely $\mathbf{x}_1, \mathbf{x}_2 \in S$ pontra

$$(\boldsymbol{\nabla} f(\mathbf{x}_2) - \boldsymbol{\nabla} f(\mathbf{x}_1))(\mathbf{x}_2 - \mathbf{x}_1) \geqq 0.$$

TÉTEL (karakterizációs tétel kétszer differenciálható konvex függvényre)
Legyen $S \subseteq \mathbb{R}^n$ nemüres nyílt konvex halmaz és legyen $f : S \to \mathbb{R}$ kétszer differenciálható függvény az S halmazon.

Az f függvény akkor és csak akkor konvex, ha az f függvény $\mathbf{H}(\mathbf{x})$ Hesse mátrixa pozitív szemidefinit minden $\mathbf{x} \in S$ pontban.

Az f függvény akkor és csak akkor konkáv, ha az f függvény $\mathbf{H}(\mathbf{x})$ Hesse mátrixa negatív szemidefinit minden $\mathbf{x} \in S$ pontban.

Példa:
Tekintsük az alábbi kétváltozós valós függvényt az $S = \mathbb{R}^2$ halmazon. Igazoljuk, hogy ez a függvény konvex.

$$f(x_1, x_2) = e^{2x_1 + x_2}.$$

Megoldás:
A függvény kétszer differenciálható, a függvény gradiensvektora és a Hesse mátrixa az alábbiak szerint írható:

$$\boldsymbol{\nabla} f(\mathbf{x}) = \begin{bmatrix} 2e^{2x_1+,x_2} \\ e^{2x_1+,x_2} \end{bmatrix}, \quad \mathbf{H}(\mathbf{x}) = \begin{bmatrix} 4e^{2x_1+,x_2} & 2e^{2x_1+,x_2} \\ 2e^{2x_1+,x_2} & e^{2x_1+,x_2} \end{bmatrix}$$

El kell dönteni a Hesse mátrix definitségét, ezt inercia teszt segítségével végezzük. A számítások eredménye:

$$\mathbf{H}(\mathbf{x}) = \begin{bmatrix} \boxed{4e^{2x_1+,x_2}} & 2e^{2x_1+,x_2} \\ 2e^{2x_1+,x_2} & e^{2x_1+,x_2} \end{bmatrix} \to \begin{bmatrix} 0 \end{bmatrix}$$

Mivel a pivotelem minden $\mathbf{x}$-re pozitív, így $Iner(\mathbf{H}(\mathbf{x})) = (1, 1, 0)$, az inerciateszt alapján a $\mathbf{H}(\mathbf{x})$ mátrix pozitív szemidefinit, a tételből pedig az következik, hogy az $f(\mathbf{x})$ függvény konvex az $\mathbb{R}^2$ halmazon.

Feladat:
A definitséget döntse el főminor teszt alkalmazásával is!

Példa:
Tekintsük a közgazdaságban jól ismert Cobb-Douglas féle termelési függvényt:

$$f(x_1, x_2) = Ax_1^{a_1} x_2^{a_2},$$

ahol $A > 0, a_1, a_2 > 0$, $a_1 + a_2 = 1$ állandók, $x_1, x_2 > 0$ változók.
Mutassuk meg, hogy ez a függvény konkáv

Megoldás:
A függvény kétszer differenciálható, képezzük függvény gradiensét és Hesse mátrixát:

$$\begin{aligned}
\boldsymbol{\nabla} f(x_1, x_2) &= \begin{bmatrix} Aa_1 x_1^{a_1-1} x_2^{a_2} \\ Aa_2 x_1^{a_1} x_2^{a_2-1} \end{bmatrix} = f(x_1, x_2) \begin{bmatrix} \frac{a_1}{x_1} \\ \frac{a_2}{x_2} \end{bmatrix}, \\
\mathbf{H}(x_1, x_2) &= \begin{bmatrix} Aa_1 (a_1 - 1) x_1^{a_1-2} x_2^{a_2} & Aa_1 a_2 x_1^{a_1-1} x_2^{a_2-1} \\ Aa_1 a_2 x_1^{a_1-1} x_2^{a_2-1} & Aa_2 (a_2 - 1) x_1^{a_1} x_2^{a_2-2} \end{bmatrix} = \\
&= f(x_1, x_2) \cdot \begin{bmatrix} \frac{a_1(a_1-1)}{x_1^2} & \frac{a_1 a_2}{x_1 x_2} \\ \frac{a_1 a_2}{x_1 x_2} & \frac{a_2(a_2-1)}{x_2^2} \end{bmatrix}.
\end{aligned}$$

Mivel $n = 2$, így a legalkalmasabb teszt a főminor teszt. A feltételek miatt $f(x_1, x_2) > 0$ minden $x_1, x_2 > 0$ esetén, ezt felhasználva a főminorokra a következőket mondhatjuk.
A Hesse mátrix első főminora negatív, mivel $0 < a_1 < 1$.
A Hesse mátrix második főminora, azaz a Hesse mátrix determinánsa

$$\det(\mathbf{H}(x_1, x_2)) = [f(x_1, x_2)]^2 \frac{a_1 a_2 (1 - a_1 - a_2)}{x_1^2 x_2^2},$$

ez pedig $a_1 + a_2 = 1$ miatt zérus. A Hesse mátrix tehát negatív szemidefinit. A tételeink következtében a fenti Cobb-Douglas féle termelési függvény konkáv függvény.

Megjegyezések:
Ha $a_1 + a_2 < 1$, akkor a második főminor pozitív, így ebben az esetben a Cobb-Douglas féle termelési függvény szigorúan konkáv függvény.
Ha $a_1 + a_2 > 1$, akkor a második főminor negatív, így függetlenül az első főminortól ($a_1 > 1$ vagy $a_1 < 1$) a Hesse mátrix indefinit, így ebben az esetben a Cobb-Douglas féle termelési függvény sem nem konkáv sem nem konvex függvény.

Példa:
Tekintsük az alábbi kétváltozós valós függvényt az $\mathbb{R}^2$ halmaz azon részén, amelyen $2x_1 + x_2 > 0$. Igazoljuk, hogy ez a függvény konvex.

$$f(x_1, x_2) = 2^{x_1 + x_2^4} - \ln(2x_1 + x_2) + e^{x_1^2 + x_2 + 1}.$$

Megoldás:

A függvény három tagból áll. A második tagot a $-\ln(2x_1+x_2)$ függvényt írjuk át $+1(-\ln(2x_1+x_2))$ alakra, így a függvényünk három függvény pozitív lineáris kombinációja. Ha mindegyik tag konvex függvény, akkor az eredeti függvény is konvex, hisz most már az eredeti függvény függvények pozitív lineáris kombinációja. Ismert tétel pedig azt mondja, hogy konvex függvények nemnegatív kombinációja is konvex. Ezzel egyszerűbbé tehetjük a konvexitás vizsgálatát. Legyen a három függvény f_1, f_2, f_3, ekkor a függvények gradiense és a Hesse mátrixa az alábbi

$$\begin{array}{ll}
\boldsymbol{\nabla} f_1 = \begin{bmatrix} 2^{x_1+x_2^4}\ln 2 \\ 2^{x+x_2^4}4x_2^3\ln 2 \end{bmatrix} & \mathbf{H}_1 = 2^{x_1+x_2^4}\ln 2 \begin{bmatrix} \ln 2 & 4x_2^3\ln 2 \\ 4x_2^3\ln 2 & 16x_2^6\ln 2 + 12x_2^2 \end{bmatrix} \\
\boldsymbol{\nabla} f_2 = \begin{bmatrix} -\frac{2}{2x_1+x_2} \\ -\frac{1}{2x_1+x_2} \end{bmatrix} & \mathbf{H}_2 = \frac{1}{(2x_1+x_2)^2}\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \\
\boldsymbol{\nabla} f_3 = \begin{bmatrix} 2x_1 e^{x_1^2+x_2+1} \\ e^{x_1^2+x_2+1} \end{bmatrix} & \mathbf{H}_3 = e^{x_1^2+x_2+1}\begin{bmatrix} 2+4x_1^2 & 2x_1 \\ 2x_1 & 1 \end{bmatrix}
\end{array}.$$

A főminor tesztet alkalmazzuk. Mindegyik Hesse mátrix első főminora, azaz a bal felső eleme pozitív. A második főminorokra, azaz a Hesse mátrix determinánsaira az alábbiak adódnak

$$\det(\mathbf{H}_1) = \left(2^{x_1+x_2^4}\right)^2 12x_2^2\ln^3 2, \qquad \det(\mathbf{H}_2) = 0, \qquad \det(\mathbf{H}_3) = 2\left(e^{x_1^2+x_2+1}\right)^2.$$

A harmadik determináns minden (x_1, x_2) vektorra pozitív, tehát a harmadik függvény Hesse mátrixa pozitív definit, így ez a függvény konvex (sőt a következő tétel szerint szigorúan konvex). Az első determináns pozitív vagy lehet zérus is, de negatív nem lehet. Az első és második függvény Hesse mátrixa pozitív szemidefinit, így ezek a függvények konvex függvények. A tétel szerint az eredeti függvény konvex, mivel a pozitív kombinációban szereplő függvények konvexek.

TÉTEL (féloldali karakterizációs tétel kétszer differenciálható konvex függvényre)

Legyen $S \subseteq \mathbb{R}^n$ nemüres nyílt konvex halmaz és legyen $f : S \to \mathbb{R}$ kétszer differenciálható függvény az S halmazon.

Ha az f függvény $\mathbf{H}(\mathbf{x})$ Hesse mátrixa pozitív definit minden $\mathbf{x} \in S$ pontban, akkor az f függvény szigorúan konvex. Fordítva, ha az f függvény szigorúan konvex, akkor a Hesse mátrix pozitív szemidefinit minden $\mathbf{x} \in S$ pontban. Azonban, ha az f függvény szigorúan konvex és kvadratikus, akkor a Hesse mátrix pozitív definit.

Ha az f függvény $\mathbf{H}(\mathbf{x})$ Hesse mátrixa negatív definit minden $\mathbf{x} \in S$ pontban, akkor az f függvény szigorúan konkáv. Fordítva, ha az f függvény szigorúan konkáv, akkor a Hesse mátrix negatív szemidefinit minden $\mathbf{x} \in S$ pontban. Azonban, ha az f függvény szigorúan konkáv és kvadratikus, akkor a Hesse mátrix negatív definit.

Példa:

Tekintsük az alábbi kétváltozós valós függyvényt az $S = \mathbb{R}^2$ halmazon. Igazoljuk, hogy ez a függvény szigorúan konvex.

$$f(x_1, x_2) = x_1^2 - 2x_1x_2 + 2x_2^2 - x_1 + 4x_2 + 5$$

Megoldás:

A függvény kétszer differenciálható, a függyvény gadiens vektora és a Hesse mátrixa az alábbiak szerint írható:

$$\boldsymbol{\nabla} f(\mathbf{x}) = \begin{bmatrix} 2x_1 - 2x_2 - 1 \\ -2x_1 + 4x_2 + 4 \end{bmatrix}, \quad \mathbf{H}(\mathbf{x}) = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}$$

El kell dönteni a Hesse mátrix definitségét, ezt inercia teszt segítségével végezzük. A számítások a következők:

$$\mathbf{H}(\mathbf{x}) = \begin{bmatrix} \boxed{2} & -2 \\ -2 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 2 \end{bmatrix}$$

Mivel a Hesse mátrix minden $\mathbf{x}$-re azonos, így $Iner(\mathbf{H}(\mathbf{x})) = (2, 0, 0)$, az inerciateszt alapján a $\mathbf{H}(\mathbf{x})$ mátrix pozitív definit, a tételből pedig az következik, hogy az $f(\mathbf{x})$ függvény szigorúan konvex az $\mathbb{R}^2$ halmazon.

Példa:
Tekintsük az alábbi n változós függvényt. Állapítsuk meg, hogy a függvény konkáv az $S = \mathbb{R}^n$ halmazon!

$$f(x_1, x_2, \ldots x_n) = -x_1^2 - x_2^2 - \ldots - x_n^2 = -\sum_{i=1}^{n} x_i^2$$

Megoldás:
Az előzőekben ennek a függvénynek a (-1)-szereséről eldöntöttük, hogy konvex, a konkáv függvény egyik definíciója értelmében az eredeti függvény tehát konkáv. Az előző vizsgálatunkban a függvény differenciálhatóságát használtuk fel, most a kétszer differenciálhatósági tulajdonságot felhasználva fogjuk kimutatni a már ismert eredményt. A függvény kétszer differenciálható, a gradiense és a Hesse mátrixa az alábbi

$$\boldsymbol{\nabla} f(\mathbf{x}) = \begin{bmatrix} -2x_1 \\ -2x_2 \\ \vdots \\ -2x_n \end{bmatrix} = -2\mathbf{x}, \quad \mathbf{H}(\mathbf{x}) = \begin{bmatrix} -2 & 0 & \cdots & 0 \\ 0 & -2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -2 \end{bmatrix}$$

A Hesse mátrix diagonális mátrix negatív diagonális elemekkel, így minden sajátértéke negatív, a sajátérték teszt értelmében a Hesse mátrix negatív definit, így a karakterizációs tétel értelmében a függvény szigorúan konkáv az $\mathbb{R}^2$ halmazon.

2.4. Konvex függvény optimuma

2.4.1. Függvény optimumának definíciója

Mielőtt a konvex függvények optimumának vizsgálatára rátérnénk, definiáljuk egy tetszőleges függvény minimumát. Legyen $S \subseteq \mathbb{R}^n$ tetszőleges halmaz és $f : S \rightarrow \mathbb{R}$ tetszőleges függvény. Tekintsük az alábbi matematikai programozási feladatot, amelyben keressük az $f(\mathbf{x})$ függvény minimumát az S halmazon, jelben:

$$\min \{f(\mathbf{x}) : \mathbf{x} \in S\}.$$

Az $\mathbf{x} \in S$ pontot, azaz a feltételt kielégítő pontot az optimalizálási probléma lehetséges (megengedett) megoldásának nevezzük.

Egy $\bar{\mathbf{x}} \in S$ lehetséges megoldást optimális megoldásnak (minimumpontnak) nevezünk, ha $f(\bar{\mathbf{x}}) \leqq f(\mathbf{x})$ minden $\mathbf{x} \in S$ esetén. Szokás ezt a pontot globális optimális megoldásnak vagy egyszerűen megoldásnak is nevezni.

Egy $\bar{\mathbf{x}} \in S$ lehetséges megoldást lokális optimális megoldásnak nevezünk, ha az $\bar{\mathbf{x}}$ pontnak van olyan $K_\varepsilon(\bar{\mathbf{x}})$ környezete, amelynél $f(\bar{\mathbf{x}}) \leqq f(\mathbf{x})$ minden $\mathbf{x} \in S \cap K_\varepsilon(\bar{\mathbf{x}})$ esetén.

Amennyiben az optimális megoldások mindkét definíciójában $f(\bar{\mathbf{x}}) < f(\mathbf{x})$ $(\mathbf{x} \neq \bar{\mathbf{x}})$ szerepel, szigorú optimális megoldásról beszélünk.

Szokásos az optimális megoldást az alábbiak szerint definiálni:

Egy $\bar{\mathbf{x}} \in S$ lehetséges megoldást optimális megoldásnak (minimumpontnak) nevezünk, ha **nincs** olyan $\mathbf{x} \in S$ pont, amelyre $f(\bar{\mathbf{x}}) > f(\mathbf{x})$.

A függvény maximumát a fentiek alapján értelemszerűen definiálhatjuk. Szokás úgy is definiálni a maximumpontot, hogy az a $-f(\mathbf{x})$ függvény minimumpontja.

2.4.2. Konvex függvény optimumára vonatkozó tételek

TÉTEL (konvex függvény optimuma globális)

Legyen $S \subseteq \mathbb{R}^n$ nemüres konvex halmaz és $f : S \to \mathbb{R}$ függvény. Tekintsük az alábbi optimalizálási feladatot:

$$\min \{f(\mathbf{x}) : \mathbf{x} \in S\}$$

Tegyük fel, hogy az $\bar{\mathbf{x}} \in S$ lokális optimális megoldás. Ekkor

a) Ha f konvex, akkor az $\bar{\mathbf{x}}$ lokális optimális megoldás egyben globális optimális megoldás is.

b) Ha f szigorúan konvex, akkor $\bar{\mathbf{x}}$ az egyetlen globális optimális megoldás.

TÉTEL (karakterizációs tétel, konvex függvény optimuma és a gradiens)

Legyen $S \subseteq \mathbb{R}^n$ nemüres konvex halmaz és $f : S \to \mathbb{R}$ differenciálható konvex függvény. Tekintsük a $\min \{f(\mathbf{x}) : \mathbf{x} \in S\}$ optimalizálási feladatot.

Az $\bar{\mathbf{x}} \in S$ akkor és csak akkor optimális megoldása a feladatnak, ha $\boldsymbol{\nabla} f(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}) \geqq 0$ minden $\mathbf{x} \in S$ esetén. Amennyiben S nyílt halmaz, úgy az $\bar{\mathbf{x}} \in S$ akkor és csak akkor optimális megoldása a feladatnak, ha $\boldsymbol{\nabla} f(\bar{\mathbf{x}}) = \mathbf{0}$.

TÉTEL (konvex függvény maximuma kompakt konvex poliéderen)

Legyen $S \subseteq \mathbb{R}^n$ nemüres kompakt konvex poliéder és $f : \mathbb{R}^n \to \mathbb{R}$ konvex függvény. Tekintsük a $\max \{f(\mathbf{x}) : \mathbf{x} \in S\}$ optimalizálási feladatot.

Ekkor létezik $\bar{\mathbf{x}} \in S$ optimális megoldás és az $\bar{\mathbf{x}}$ optimális megoldás a konvex poliéder valamely extremális pontja.

A fenti tételek közül csak egy tételt bizonyítottunk be, ezek a tételek általában könnyen bizonyíthatók. Gyakorlásképpen a legutolsó tételt is bebizonyítjuk.

Bizonyítás:

Az f függvény konvexitása az $\mathbb{R}^n$ térben azt jelenti, hogy az $\mathbb{R}^n$ térben folytonos. Folytonos függvény pedig felveszi maximumát az $S \subseteq \mathbb{R}^n$ kompakt halmazon, legyen ez a maximumpont az $\bar{\mathbf{x}} \in S$. Legyenek a kompakt konvex poliéder extremális pontjai $\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_k$. A korlátosság miatt extremális irány nincs. A konvex poliéder karakterizációs tétele szerint a poliéder tetszőleges pontja felírható az extremális pontok konvex kombinációjaként, így az $\bar{\mathbf{x}} \in S$ pont is, amely szerint:

$$\bar{\mathbf{x}} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i, \quad \lambda_1, \lambda_2, ..., \lambda_k \geqq 0, \quad \lambda_1 + \lambda_2 + ... + \lambda_k = 1.$$

Mivel az f függvény konvex, így érvényes a Jensen egyenlőtlenség, amelyet az $\bar{\mathbf{x}} \in S$ pontra felírva

$$f(\bar{\mathbf{x}}) = f\left(\sum_{i=1}^{k} \lambda_i \mathbf{x}_i\right) \leqq \sum_{i=1}^{k} \lambda_i f(\mathbf{x}_i)$$

Indirekte tegyük fel, hogy extremális pontban nincs optimális megoldás, azaz

$$f(\bar{\mathbf{x}}) > f(\mathbf{x}_i), \quad i = 1, ..., k.$$

Szorozzuk meg a fenti egyenlőtlenségeket a λ_i nemnegatív számokkal és adjuk össze az egyenlőtlenség mindkét oldalát. Kihasználva, hogy a λ_i számok összege 1, kapjuk, hogy

$$f(\bar{\mathbf{x}}) > \sum_{i=1}^{k} \lambda_i f(\mathbf{x}_i),$$

ez pedig ellentmond a Jensen egyenlőtlenségnek, tehát az f konvex függvény extremális pontban is felveszi maximumát.

3. Kvázikonvex függvény

3.1. Kvázikonvex függvény definíciója

Az alsó nívóhalmazra vonatkozó megállapítás a konvex függvény egy fontos tulajdonságát mutatja. A probléma az, hogy fordítva nem igaz az állítás, azaz ha az alsó nívóhalmaz konvex, akkor nem biztos, hogy a függvény is konvex. Most a konvex függvénynek egy olyan általánosítását adjuk, amelynél mindkét irányban igaz lesz az állítás. Ezt a függvényt kvázikonvex függvénynek nevezzük.

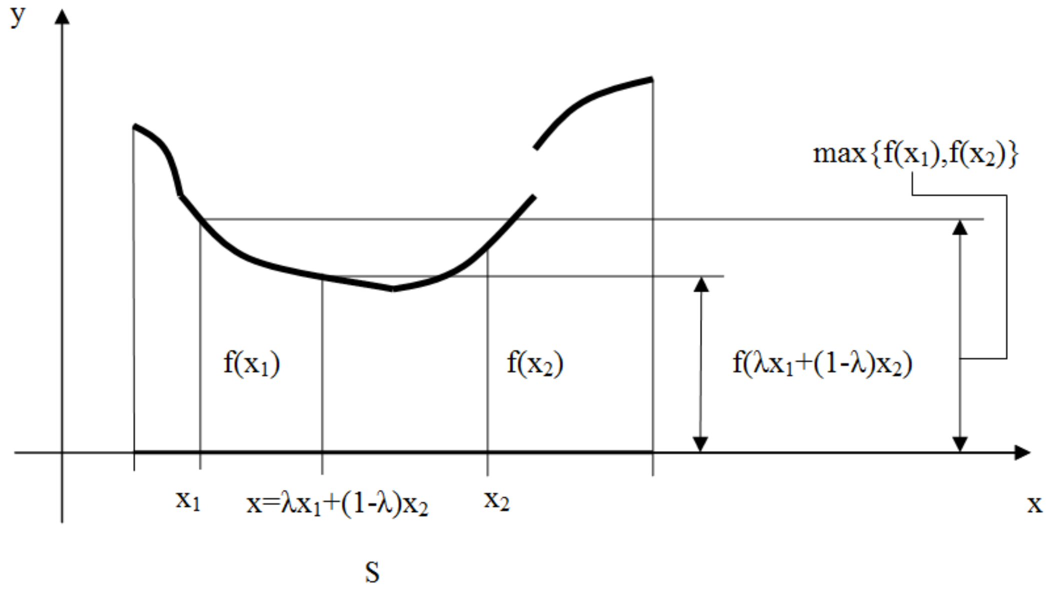
Legyen $f : S \to \mathbb{R}$, ahol $S \subseteq \mathbb{R}^n$ nemüres konvex halmaz. Az $f(\mathbf{x})$ függvényt akkor mondjuk kvázikonvexnek az S halmazon, ha

$$f(\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2) \leqq \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}$$

bármely $\mathbf{x}_1$, $\mathbf{x}_2 \in S$ vektorra és minden $\lambda \in (0, 1)$ számra.

Az f függvény szigorúan kvázikonvex, ha a formulában szigorú egyenlőtlenség áll fenn.

Az alábbi ábra egy egyváltozós kvázikonvex függvényt szemléltet.



A definíció szemléletesen azt fejezi ki, hogy kvázikonvex függvénynél bármely két pont esetén a két függvényérték közül a nagyobbik értéke alatt halad a függvény görbéje, legalábbis nem megy fölé. Többváltozós esetben is ugyanez mondható el, csak ott $n = 2$ esetén tudunk szemléletet adni.

A kvázikonvex függvény mellett a gyakorlatban nagy jelentősége van a kvázikonkáv függvénynek is, amelyet az alábbiakban definiálunk.

Legyen $f : S \to \mathbb{R}$, ahol $S \subseteq \mathbb{R}^n$ nemüres konvex halmaz. Az $f(\mathbf{x})$ függvényt akkor mondjuk kvázikonkávnak az S halmazon, ha

$$f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \geqq \min\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}$$

bármely $\mathbf{x}_1$, $\mathbf{x}_2 \in S$ vektorra és minden $\lambda \in (0, 1)$ számra.

Az f függvény szigorúan kvázikonkáv, ha a definíciós formulában szigorú egyenlőtlenség áll fenn.

Sokszor megelégszünk a fentiekkel azonos eredmény adó definícióval is, amely szerint az f függvényt (szigorúan) kvázikonkávnak nevezzük, ha a $-f$ függvény (szigorúan) kvázikonvex.

3.2. Karakterizációs tételek kvázikonvex függvényekre

TÉTEL (karakterizációs tétel, alsó nívó halmaz)

Legyen $S \subseteq \mathbb{R}^n$ nemüres konvex halmaz, $f : S \to \mathbb{R}$ függvény, $\alpha \in \mathbb{R}$ valós szám. Az f függvény akkor és csak akkor kvázikonvex az S halmazon, ha az $S_\alpha = \{\mathbf{x} \in S : f(\mathbf{x}) \leqq \alpha\}$ ún. alsó nívóhalmaz konvex halmaz minden $\alpha \in \mathbb{R}$ valós számra.

Példa:

Tekintsük az alábbi kétváltozós függvényt. Állapítsuk meg, hogy függvény kvázikonvex az $S = \mathbb{R}^2$ halmazon!

$$f(x_1, x_2) = -e^{-(x_1^2+x_2^2)}$$

Megoldás:

A függvény alsó nívóhalmaza

$$S_\alpha = \left\{(x_1, x_2) \in \mathbb{R}^2 : -e^{-(x_1^2+x_2^2)} \leqq \alpha\right\}$$

Az $e^{-(x_1^2+x_2^2)} \geqq -\alpha$ egyenlőtlenséget vizsgálva,

ha $\alpha \geqq 0$, akkor minden $\mathbf{x}$ vektorra igaz az egyenlőtlenség, tehát $S_\alpha = \mathbb{R}^2$, ez pedig konvex halmaz,

ha $\alpha < -1$, akkor nincs olyan $\mathbf{x}$ vektor, amelyre igaz az egyenlőtlenség, tehát $S_\alpha = \emptyset$, a definíció szerint az üres halmazt is konvexnek tekintjük,

ha $-1 \leqq \alpha < 0$, akkor - felhasználva az exponenciális függvény szigorúan monoton növekedését - az egyenlőtlenség az alábbiak szerint írható

$$x_1^2 + x_2^2 \leqq -\ln(-\alpha).$$

A vizsgált tartományban $-\ln(-\alpha) \geqq 0$, tehát az alsó nívóhalmazok $\sqrt{-\ln(-\alpha)}$ sugarú körök és azok belseje. A körtartomány pedig köztudottan konvex halmaz. Az $\alpha = -1$ esetben kör helyett egyetlen pontot kapunk, de definíció szerint az egyetlen pontból álló halmazt is konvexnek tekintjük. Összességében tehát az alsó nívóhalmazok konvexek, így a karakterizációs tétel értelmében a függvény kvázikonvex.

TÉTEL (karakterizációs tétel, differenciálható kvázikonvex függvényre)

Legyen $S \subseteq \mathbb{R}^n$ nemüres nyílt konvex halmaz és legyen $f : S \to \mathbb{R}$ differenciálható függvény az S halmazon. Az f függvény akkor és csak akkor kvázikonvex, ha

$$\text{bármely } \mathbf{x}_1, \mathbf{x}_2 \in S \text{ és } f(\mathbf{x}_1) \leqq f(\mathbf{x}_2) \text{ esetén } \boldsymbol{\nabla} f(\mathbf{x}_2)(\mathbf{x}_1 - \mathbf{x}_2) \leqq 0.$$

Bizonyítás:

Csak a tétel egyik irányát bizonyítjuk, tehát a függvény kvázikonvex voltából indulunk ki. Az $f(\mathbf{x})$ függvény differenciálható, így írható, hogy

$$f(\mathbf{x}_2 + \lambda(\mathbf{x}_1 - \mathbf{x}_2)) = f(\mathbf{x}_2) + \lambda \boldsymbol{\nabla} f(\mathbf{x}_2)(\mathbf{x}_1 - \mathbf{x}_2) + |\lambda| \, \|\mathbf{x}_1 - \mathbf{x}_2\| \, \alpha(\mathbf{x}_2, \lambda(\mathbf{x}_1 - \mathbf{x}_2)).$$

A baloldali függvény átírható, vagyis $f(\mathbf{x}_2 + \lambda(\mathbf{x}_1 - \mathbf{x}_2)) = f(\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2)$.

Az $f(\mathbf{x})$ függvény kvázikonvex, ami azt jelenti, hogy $f(\lambda\mathbf{x}_1+(1-\lambda)\mathbf{x}_2) \leqq \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}$. A tételben szereplő $f(\mathbf{x}_1) \leqq f(\mathbf{x}_2)$ feltételből azt kapjuk, hogy $f(\lambda\mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2) \leqq f(\mathbf{x}_2)$, ekkor a differenciálhatósági összefüggés az alábbi szerint alakul

$$\lambda \boldsymbol{\nabla} f(\mathbf{x}_2)(\mathbf{x}_1 - \mathbf{x}_2) + |\lambda| \, \|\mathbf{x}_1 - \mathbf{x}_2\| \, \alpha(\mathbf{x}_1, \mathbf{x}_1 - \mathbf{x}_2) \leqq 0$$

Az egyenlőtlenséget λ-val osztva, majd a $\lambda \to 0$ határátmenetet képezve, a bizonyítandó egyenlőtlenséget kapjuk.

Q.e.d.

TÉTEL (karakterizációs tétel, kétszer differenciálható kvázikonvex függvényre)

Legyen $S \subseteq \mathbb{R}^n$ nemüres nyílt konvex halmaz és legyen $f : S \to \mathbb{R}$ kétszer differenciálható függvény az S halmazon.

Az f függvény akkor és csak akkor **kvázikonvex**, ha az f függvény $\mathbf{H}(\mathbf{x})$ Hesse mátrixa minden $\mathbf{x} \in S$ pontban f**eltételesen pozitív szemidefinit** minden $\mathbf{d} \neq \mathbf{0}$, $\nabla f(\mathbf{x})\mathbf{d} = 0$ vektorra nézve.

Az f függvény akkor és csak akkor **kvázikonkáv**, ha az f függvény $\mathbf{H}(\mathbf{x})$ Hesse mátrixa minden $\mathbf{x} \in S$ pontban **feltételesen negatív szemidefinit** minden $\mathbf{d} \neq \mathbf{0}$, $\nabla f(\mathbf{x})\mathbf{d} = 0$ vektorra nézve.

Példa:
Tekintsük az alábbi kétváltozós függvényt. Állapítsuk meg, hogy függvény kvázikonvex vagy kvázikonkáv!

$$f(x_1, x_2) = \frac{x_1^2 + x_2^2}{x_1^2 + x_2^2 + 1}$$

Megoldás:
A függvény kétszer differenciálható, a függvény gradiens vektora és a Hesse mátrixa az alábbiak szerint írható:

$$\boldsymbol{\nabla} f(x_1, x_2) = \begin{bmatrix} \frac{2x_1}{\left(x_1^2+x_2^2+1\right)^2} \\ \frac{2x_2}{\left(x_1^2+x_2^2+1\right)^2} \end{bmatrix}, \qquad \mathbf{H}(x_1, x_2) = \begin{bmatrix} \frac{2\left(x_2^2-3x_1^2+1\right)}{\left(x_1^2+x_2^2+1\right)^3} & \frac{-8x_1x_2}{\left(x_1^2+x_2^2+1\right)^3} \\ \frac{-8x_1x_2}{\left(x_1^2+x_2^2+1\right)^3} & \frac{2\left(x_1^2-3x_2^2+1\right)}{\left(x_1^2+x_2^2+1\right)^3} \end{bmatrix}$$

A Hesse mátrix indefinit, így sem nem konvex sem nem konkáv a függvény. A Hesse mátrix determinánsa

$$\frac{4 - 12x_1^2 - 12x_2^2}{(x_1^2 + x_2^2 + 1)^5}.$$

A főminor teszt szerint akkor lenne konvex vagy konkáv a függvény, ha minden $\mathbf{x}$ vektorra a második főminor, vagyis a Hesse mátrix determinánsa pozitív lenne, de ez nincs így. Most nézzük meg lehet-e a függvény kvázikonvex vagy kvázikonkáv. Ez esetben a tétel alapján azt kellene megvizsgálni, hogy a Hesse mátrix feltételesen pozitív vagy negatív definit. A feltételes definitség, mint a mátrix definitségével foglalkozó tananyagban látható, a szegélyezett mátrix segítségével dönthető el. Képezzük a szegélyezett Hesse mátrixot, amelynél a szegélyt a $\boldsymbol{\nabla} f(x_1, x_2)$ gradiens vektor adja.

$$\mathbf{C}(x_1, x_2) = \begin{bmatrix} \frac{2\left(x_2^2-3x_1^2+1\right)}{\left(x_1^2+x_2^2+1\right)^3} & \frac{-8x_1x_2}{\left(x_1^2+x_2^2+1\right)^3} & \frac{2x_1}{\left(x_1^2+x_2^2+1\right)^2} \\ \frac{-8x_1x_2}{\left(x_1^2+x_2^2+1\right)^3} & \frac{2\left(x_1^2-3x_2^2+1\right)}{\left(x_1^2+x_2^2+1\right)^3} & \frac{2x_2}{\left(x_1^2+x_2^2+1\right)^2} \\ \frac{2x_1}{\left(x_1^2+x_2^2+1\right)^2} & \frac{2x_2}{\left(x_1^2+x_2^2+1\right)^2} & 0 \end{bmatrix}$$

A főminor teszt segítségével fogjuk eldönteni a definitséget. Mivel esetünkben $n = 2, m = 1$, így csak $k = 2$ esetében kell determinánst számolni, amely valójában a szegélyezett Hesse mátrix determinánsa. A számolást megkönnyíthetjük, ha átrendezzük a szóbanforgó mátrixot:

$$\mathbf{C}(x_1, x_2) = \frac{2}{(x_1^2 + x_2^2 + 1)^3} \begin{bmatrix} (x_2^2 - 3x_1^2 + 1) & -4x_1x_2 & x_1^3 + x_1x_2^2 + x_1 \\ -4x_1x_2 & (x_1^2 - 3x_2^2 + 1) & x_1^2x_2 + x_2^3 + x_2 \\ x_1^3 + x_1x_2^2 + x_1 & x_1^2x_2 + x_2^3 + x_2 & 0 \end{bmatrix}$$

Ennek determinánsa:

$$\det(\mathbf{C}) = (-8)\frac{x_1^2 + x_2^2}{(x_1^2 + x_2^2 + 1)^6}$$

A determináns minden $\mathbf{x} \neq \mathbf{0}$ vektorra negatív. Azt kaptuk, hogy $(-1)^m \det(\mathbf{C}_k) > 0,\ k = 2$-re, így a főminor teszt alapján a Hesse mátrix feltételesen pozitív definit. Ez pedig a fenti tétel értelmében azt jelenti, hogy a példabeli függvény kvázikonvex.

Példa:

Tekintsük az alábbi kétváltozós függvényt. Állapítsuk meg, hogy függvény kvázikonvex vagy kvázikonkáv!

$$f(x_1, x_2) = e^{-x_1^2 - x_2^2}$$

Megoldás:

Az előzőekben ennek a függvénynek a (-1)-szereséről eldöntöttük, hogy kvázikonvex, a kvázikonkávitás másik definíciója értelmében ez a függvény kvázikonkáv. Abban a vizsgálatban nem használtuk fel a függvény differenciálhatóságát, most a kétszer differenciálhatósági tulajdonságot felhasználva fogjuk kimutatni a már ismert eredményt. A függvény kétszer differenciálható, a függvény gradiens vektora és a Hesse mátrixa az alábbiak szerint írható:

$$\nabla f(x_1, x_2) = \begin{bmatrix} -2x_1 e^{-x_1^2 - x_2^2} \\ -2x_2 e^{-x_1^2 - x_2^2} \end{bmatrix}, \qquad \mathbf{H}(x_1, x_2) = \begin{bmatrix} 2\left(2x_1^2 - 1\right) e^{-x_1^2 - x_2^2} & 4x_1 x_2 e^{-x_1^2 - x_2^2} \\ 4x_1 x_2 e^{-x_1^2 - x_2^2} & 2\left(2x_2^2 - 1\right) e^{-x_1^2 - x_2^2} \end{bmatrix}$$

Egy kis számolás után meggyőződhet az olvasó, hogy a Hesse mátrix indefinit, így sem nem konvex sem nem konkáv a függvény.

A feltételes definitség eldöntéséhez most is képezzük a ∇f vektorral szegélyezett Hesse mátrixot, amely átrendezve az alábbi

$$\mathbf{C}(x_1, x_2) = 2e^{-x_1^2 - x_2^2} \begin{bmatrix} 2x_1^2 - 1 & 2x_1 x_2 & -x_1 \\ 2x_1 x_2 & 2x_2^2 - 1 & -x_2 \\ -x_1 & -x_2 & 0 \end{bmatrix}$$

Mivel e példában is $n = 2, m = 1$, így csak $k = 2$ esetében kell determinánst számolni, amely valójában a szegélyezett Hesse mátrix determinánsa, ez pedig az alábbi

$$\det(\mathbf{C}) = 8e^{3(-x_1^2 - x_2^2)} \left(x_1^2 + x_2^2\right).$$

A determináns minden $\mathbf{x} \neq \mathbf{0}$ vektorra pozitív. Azt kaptuk, hogy $(-1)^k \det(\mathbf{C}_k) > 0,\ k = 2$-re, így a főminor teszt alapján a Hesse mátrix feltételesen negatív definit. Ez pedig a fenti tétel értelmében azt jelenti, hogy a példabeli függvény kvázikonkáv.

Példa:

Tekintsük az alábbi kétváltozós függvényt. Állapítsuk meg, hogy függvény kvázikonvex vagy kvázikonkáv!

$$f(x_1, x_2) = \sqrt[4]{x_1^2 + x_2^2}$$

Megoldás:

A függvény kétszer differenciálható, a függvény gradiens vektora és a Hesse mátrixa az alábbiak szerint írható:

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{1}{2} \frac{x_1}{\left(x_1^2 + x_2^2\right)^{\frac{3}{4}}} \\ \frac{1}{2} \frac{x_2}{\left(x_1^2 + x_2^2\right)^{\frac{3}{4}}} \end{bmatrix}, \qquad H(x_1, x_2) = \begin{bmatrix} \frac{1}{4} \frac{2x_2^2 - x_1^2}{\left(x_1^2 + x_2^2\right)^{\frac{7}{4}}} & -\frac{1}{4} \frac{3x_1 x_2}{\left(x_1^2 + x_2^2\right)^{\frac{7}{4}}} \\ -\frac{1}{4} \frac{3x_1 x_2}{\left(x_1^2 + x_2^2\right)^{\frac{7}{4}}} & \frac{1}{4} \frac{2x_1^2 - x_2^2}{\left(x_1^2 + x_2^2\right)^{\frac{7}{4}}} \end{bmatrix}$$

Egy kis számolás után meggyőződhet az olvasó, hogy a Hesse mátrix indefinit, így sem nem konvex sem nem konkáv a függvény.

A feltételes definitség eldöntéséhez most is képezzük a ∇f vektorral szegélyezett Hesse mátrixot, amely átrendezve az alábbi

$$\mathbf{C}(x_1, x_2) = \frac{1}{4} \frac{1}{(x_1^2 + x_2^2)^{\frac{7}{4}}} \begin{bmatrix} 2x_2^2 - x_1^2 & -3x_1x_2 & 2\left(x_1^3 + x_1x_2^2\right) \\ -3x_1x_2 & 2x_1^2 - x_2^2 & 2\left(x_1^2x_2 + x_2^3\right) \\ 2\left(x_1^3 + x_1x_2^2\right) & 2\left(x_1^2x_2 + x_2^3\right) & 0 \end{bmatrix}$$

Mivel e példában is $n = 2, m = 1$, így csak $k = 2$ esetében kell determinánst számolni, amely valójában a szegélyezett Hesse mátrix determinánsa, ez pedig az alábbi

$$\det(\mathbf{C}) = (-\frac{1}{8}) \frac{1}{(x_1^2 + x_2^2)^{\frac{5}{4}}}$$

A determináns minden $\mathbf{x} \neq \mathbf{0}$ vektorra negatív. Azt kaptuk, hogy $(-1)^m \det(\mathbf{C}_k) > 0,\ k = 2$-re, így a főminor teszt alapján a Hesse mátrix feltételesen pozitív definit. Ez pedig a fenti tétel értelmében azt jelenti, hogy a példabeli függvény kvázikonvex.

3.3. Kvázikonvex függvény maximuma

TÉTEL (kvázikonvex függvény maximuma kompakt konvex poliéderen)

Legyen $S \subseteq \mathbb{R}^n$ nemüres kompakt konvex poliéder és $f : \mathbb{R}^n \to \mathbb{R}$ kvázikonvex és folytonos függvény az S halmazon. Tekintsük a $\max\{f(\mathbf{x}) : \mathbf{x} \in S\}$ optimalizálási feladatot.

Ekkor létezik $\bar{\mathbf{x}} \in S$ optimális megoldás és az $\bar{\mathbf{x}}$ optimális megoldás a konvex poliéder valamely extremális pontja.