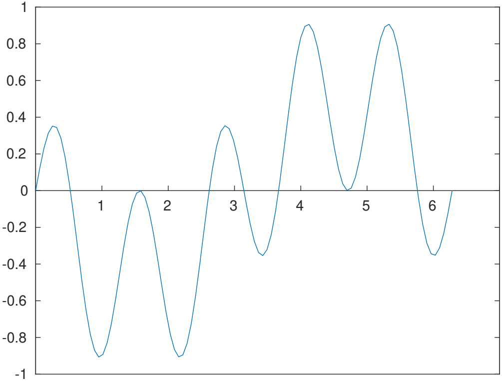
Numerikus matematika

Baran Ágnes

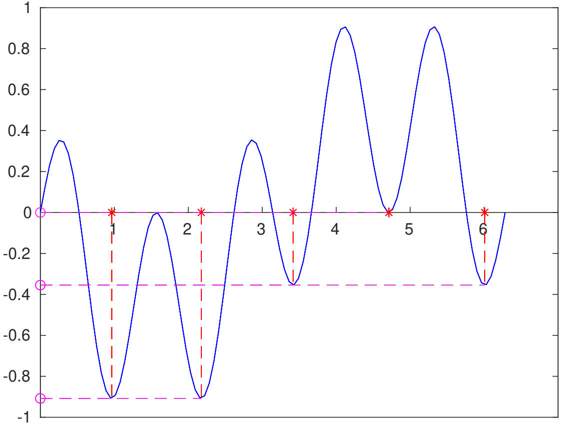
Optimalizálás

Egyváltozós függvény szélsőértéke (emlékeztető)

Az $f : \mathbb{R} \to \mathbb{R}$ függvény szélsőértékhelyeit keressük.



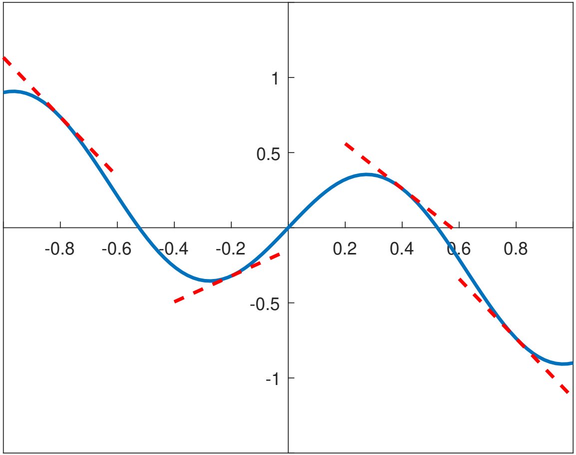
Egy függvénynek több lokális szélsőértéke is lehet.



Az $f(x) = \sin(2x)\cos(3x)$ függvény $[0, 2\pi]$-beli

- lokális minimumhelyei (*) és
- lokális minimumai (o)

Legyen $f : \mathbb{R} \to \mathbb{R}$ egy differenciálható függvény, mit jelent $f'(x_0)$?

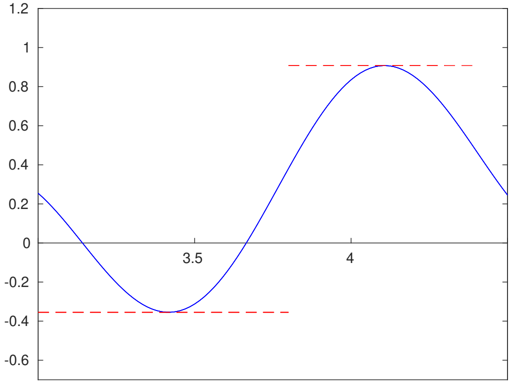


Emlékeztető: az f függvény x_0-beli érintője:

$$y(x) = f(x_0) + f'(x_0)(x - x_0)$$

Szélsőérték, szükséges feltétel

Az $f : \mathbb{R} \to \mathbb{R}$ differenciálható függvénynek csak ott lehet lokális szélsőértéke, ahol $f'(x) = 0$.

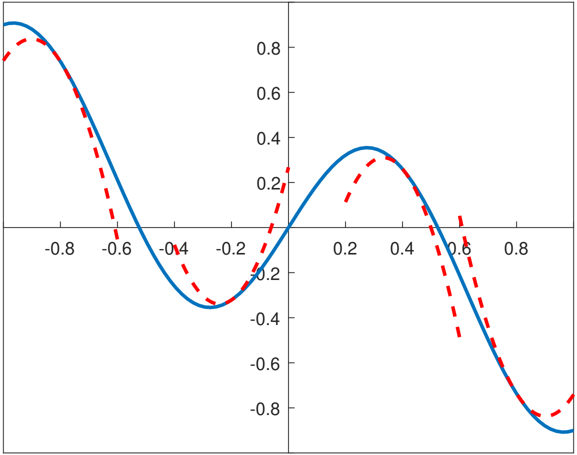


Fordítva nem igaz! Abból, hogy $f'(x) = 0$ NEM következik, hogy a függvénynek ott lokális szélsőértéke van.

Emlékeztető: ha $f : \mathbb{R} \to \mathbb{R}$ egy kétszer differenciálható függvény, akkor x_0 egy kis környezetében közelíthetjük az

$$y(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2$$

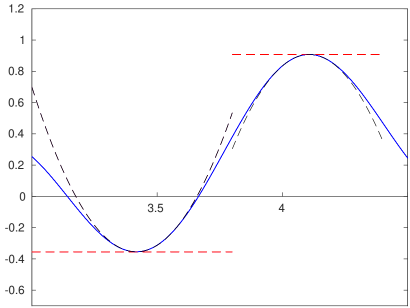
másodfokú polinommal.



Szélsőérték, elégséges feltételek

Ha az $f : \mathbb{R} \to \mathbb{R}$ függvény kétszer differenciálható és

- $f'(x^\star) = 0$, $f''(x^\star) > 0$, akkor f-nek $x^\star$-ban lokális minimuma van.
- $f'(x^\star) = 0$, $f''(x^\star) < 0$, akkor f-nek $x^\star$-ban lokális maximuma van.



piros szaggatott vonal: $f(x_0) + f'(x_0)(x - x_0)$

fekete szaggatott vonal: $f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2$

Egyváltozós függvény szélsőértéke

Az $f : \mathbb{R} \to \mathbb{R}$ függvény szélsőértékhelyeit keressük.

Szükséges feltétel

Az $f : \mathbb{R} \to \mathbb{R}$ differenciálható függvénynek csak ott lehet lokális szélsőértéke, ahol $f'(x) = 0$.

Elégséges feltételek

Ha az $f : \mathbb{R} \to \mathbb{R}$ függvény kétszer differenciálható és

- $f'(x^\star) = 0$, $f''(x^\star) > 0$, akkor f-nek $x^\star$-ban lokális minimuma van.
- $f'(x^\star) = 0$, $f''(x^\star) < 0$, akkor f-nek $x^\star$-ban lokális maximuma van.

Példa

Egy légitársaság A és B város közötti repülőjáratára 500 Euró egy jegy. A két város között egy 300 férőhelyes gép közlekedik, de átlagosan csak 180 utassal. Piackutatások szerint minden 5 Eurós engedmény a jegyárból átlagosan 3 plusz utast jelentene. Milyen jegyár mellett lenne maximális a légitársaság bevétele?

Tegyük fel, hogy a légitársaság $5n$ Eurót enged a jegyárból. Ekkor a várható bevétele:

$$f(n) = (180 + 3n)(500 - 5n) = -15n^2 + 600n + 90000$$

Az f maximumhelyét keressük.

$$f'(n) = -30n + 600$$

$$f'(n) = 0 \iff n = 20$$

Mivel

$$f''(n) = -30,$$

így $f''(20) < 0$, azaz $n = 20$ az f függvény maximumhelye.

1. feladat

Keresse meg az $f(x) = x^3 - 6x^2 + 9x + 15$ függvény lokális szélsőértékhelyeit!

2. feladat

Egy 108 dm^3 térfogatú, négyzet alapú, felül nyitott dobozt akarunk készíteni. Hogyan válasszuk meg a doboz méretét, ha a készítéséhez felhasznált anyag mennyiségét minimalizálni szeretnénk?

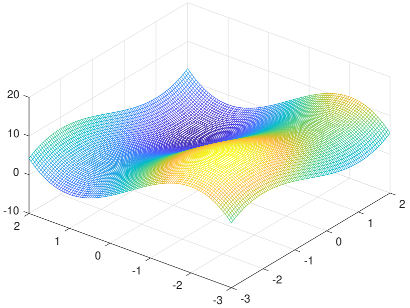
3. feladat

Egy folyó melletti telken szeretnénk egy 1800 m^2-es téglalap alakú részt elkeríteni úgy, hogy egyik oldalról a folyó határolja. Milyen méretű részt kerítsünk el, ha a felhasznált kerítés hosszát minimalizálni szeretnénk?

Kétváltozós függvények

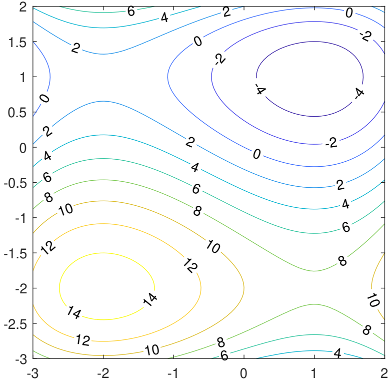
Példa

Az
$$f(x_1, x_2) = \frac{1}{4}(2x_1^3 + 3x_1^2 - 12x_1) + \frac{1}{2}(2x_2^3 + 3x_2^2 - 12x_2)$$
függvény a $[-2, 2] \times [-2, 3]$ tartomány felett.



Példa

Rajzoltassuk ki az előző függvény szintvonalait is. (Mikroökonómia: szintvonal = közömbösségi görbe.)



Kétváltozós függvények minimalizálása

Az $f : \mathbb{R}^2 \to \mathbb{R}$ függvény lokális minimumhelyeit keressük.

Gradiens

Az $f : \mathbb{R}^2 \to \mathbb{R}$ függvény x-beli gradiense

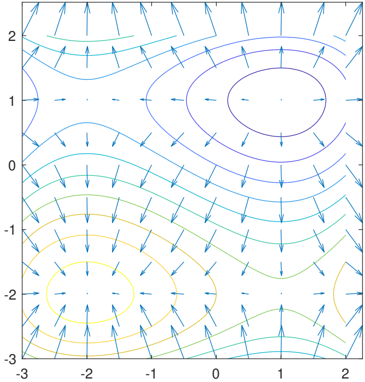
$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \end{bmatrix}$$

Példa

$$f(x_1, x_2) = \frac{1}{4}(2x_1^3 + 3x_1^2 - 12x_1) + \frac{1}{2}(2x_2^3 + 3x_2^2 - 12x_2)$$

$$\nabla f(x) = \begin{bmatrix} \frac{3}{2}x_1^2 + \frac{3}{2}x_1 - 3 \\ 3x_2^2 + 3x_2 - 6 \end{bmatrix}$$

Látjuk, hogy a gradiensvektor értéke pontonként más-más lehet. Rácsozzuk be a $[-3, 2]^2$ tartományt (mindkét tengely mentén 11-11 részre osztva) és számítsuk ki az előző függvény gradiensét ezekben a pontokban, majd rajzoltassuk rá ezeket a vektorokat a szintvonalakra!

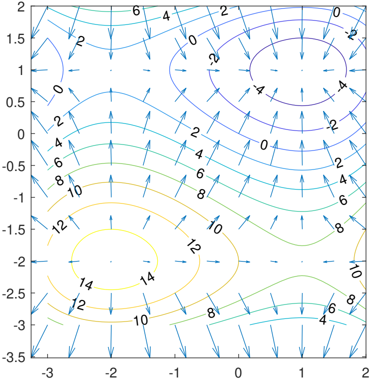


Az előző ábrán megfigyelhetjük, hogy

- a gradiensvektor merőleges az adott pontbeli szintvonalra
- a vektorok hossza a gradiens nagyságát, az iránya a gradiens irányát mutatja
- bizonyos pontokban a gradiensvektor hossza 0, vagy 0 közeli

A gradiensvektor az adott pontban a legmeredekebb emelkedés irányába mutat, a (-1)-szerese (a negatív gradiens) pedig a legmeredekebb csökkenés irányába.

Ha a gradiensmező helyett a negatív gradiensmezőt rajzoltatjuk ki, akkor a nyilak a csökkenés irányába mutatnak.



Az
$$f(x_1, x_2) = \frac{1}{4}(2x_1^3 + 3x_1^2 - 12x_1) + \frac{1}{2}(2x_2^3 + 3x_2^2 - 12x_2)$$
függvény szintvonalai és a negatív gradiens mező.

A lokális szélsőérték feltételei

Elsőrendű szükséges feltétel

Ha x^* az $f : \mathbb{R}^2 \to \mathbb{R}$ lokális minimumhelye, és f folytonosan differenciálható az x^* egy nyílt környezetében, akkor $\nabla f(x^*) = 0$.

Definíció (Stacionárius pont)

Legyen $f : \mathbb{R}^2 \to \mathbb{R}$. Az x^* pontot stacionárius pontnak hívjuk, ha $\nabla f(x^*) = 0$.

Megjegyzés

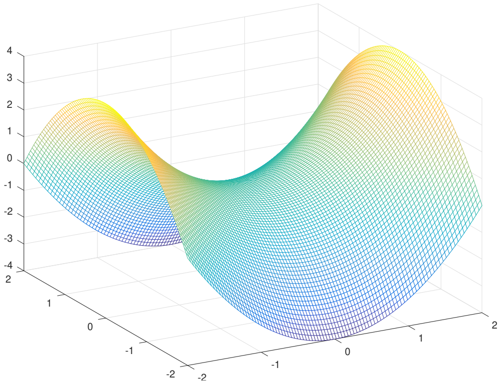
Ha x^* stacionárius pontja f-nek, akkor stacionárius pontja $-f$-nek is, azaz a stacionárius pont lokális maximum is lehet.

Definíció (Nyeregpont)

Ha x^* olyan stacionárius pontja f-nek, amely se nem lokális minimum, se nem lokális maximum, akkor nyeregpontnak hívjuk.

Példa

Legyen $f(x) = x_1^2 - x_2^2$. Ekkor $\nabla f(x) = (2x_1, -2x_2)^T$, így $x = (0,0)$ az egyetlen stacionárius pont, amely nyeregpont.



Példa

Határozzuk meg az

$$f(x_1, x_2) = \frac{1}{4}(2x_1^3 + 3x_1^2 - 12x_1) + \frac{1}{2}(2x_2^3 + 3x_2^2 - 12x_2)$$

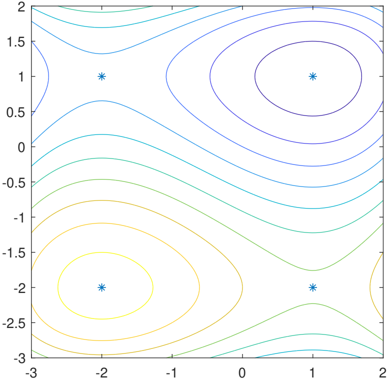
függvény stacionárius pontjait!

$$\nabla f(x) = \begin{bmatrix} \frac{3}{2}x_1^2 + \frac{3}{2}x_1 - 3 \\ 3x_2^2 + 3x_2 - 6 \end{bmatrix}$$

$$\nabla f(x) = 0 \iff x_1^2 + x_1 - 2 = 0 \text{ és } x_2^2 + x_2 - 2 = 0$$

A stacionárius pontok:

$$(1, 1), \quad (1, -2), \quad (-2, 1), \quad (-2, -2)$$



Az

$$f(x_1, x_2) = \frac{1}{4}(2x_1^3 + 3x_1^2 - 12x_1) + \frac{1}{2}(2x_2^3 + 3x_2^2 - 12x_2)$$

függvény szintvonalai és stacionárius pontjai.

A stacionárius pont típusai

Hesse-mátrix

Vezessük be a következő jelölést:

$$f_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}, \quad i, j = 1, 2.$$

Az $f : \mathbb{R}^2 \to \mathbb{R}$ függvény Hesse-mátrixa:

$$H(x) = \begin{bmatrix} f_{11}(x) & f_{12}(x) \\ f_{21}(x) & f_{22}(x) \end{bmatrix}$$

Legyen $\Delta_1 := f_{11}(x)$ és $\Delta_2 := \det(H(x))$.

A stacionárius pont típusai

Tétel

Ha az $x \in \mathbb{R}^2$ stacionárius pontban

- $\Delta_2 > 0$ és $\Delta_1 > 0$, akkor x lokális minimumhely.
- $\Delta_2 > 0$ és $\Delta_1 < 0$, akkor x lokális maximumhely.
- $\Delta_2 < 0$, akkor x nyeregpont.
- $\Delta_2 = 0$, akkor további vizsgálat szükséges.

Példa

Határozzuk meg az

$$f(x_1, x_2) = \frac{1}{4}(2x_1^3 + 3x_1^2 - 12x_1) + \frac{1}{2}(2x_2^3 + 3x_2^2 - 12x_2)$$

függvény stacionárius pontjainak típusát.

$$f(x_1, x_2) = \frac{1}{4}(2x_1^3 + 3x_1^2 - 12x_1) + \frac{1}{2}(2x_2^3 + 3x_2^2 - 12x_2)$$

A gradiens:

$$\nabla f(x) = \begin{bmatrix} \frac{3}{2}x_1^2 + \frac{3}{2}x_1 - 3 \\ 3x_2^2 + 3x_2 - 6 \end{bmatrix}$$

A stacionárius pontok:

$$(1,1), \quad (1,-2), \quad (-2,1), \quad (-2,-2)$$

A Hesse-mátrix:

$$H(x) = \begin{bmatrix} 3x_1 + \frac{3}{2} & 0 \\ 0 & 6x_2 + 3 \end{bmatrix}$$

$$H(1,1) = \begin{bmatrix} \frac{9}{2} & 0 \\ 0 & 9 \end{bmatrix}$$

$\Delta_1 > 0$ és $\Delta_2 > 0$, így az $(1,1)$ lokális minimumhely

$$H(1,-2) = \begin{bmatrix} \frac{9}{2} & 0 \\ 0 & -9 \end{bmatrix}$$

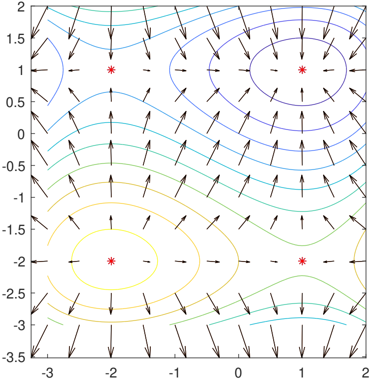
$\Delta_2 < 0$, így az $(1,-2)$ nyeregpont.

$$H(-2,1) = \begin{bmatrix} -\frac{9}{2} & 0 \\ 0 & 9 \end{bmatrix}$$

$\Delta_2 < 0$, így az $(-2,1)$ nyeregpont.

$$H(-2,-2) = \begin{bmatrix} -\frac{9}{2} & 0 \\ 0 & -9 \end{bmatrix}$$

$\Delta_1 < 0$ és $\Delta_2 > 0$, így az $(-2,-2)$ lokális maximumhely

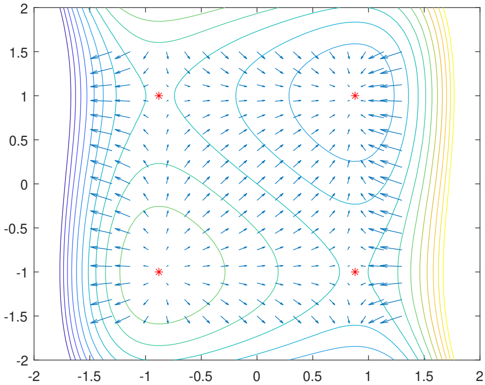


Az

$$f(x_1, x_2) = \frac{1}{4}(2x_1^3 + 3x_1^2 - 12x_1) + \frac{1}{2}(2x_2^3 + 3x_2^2 - 12x_2)$$

függvény szintvonalai, stacionárius pontjai és a negatív gradiens mező.

Az ábrán az $f(x) = x_1^5 + x_2^3 - 3x_1 - 3x_2$ függvény szintvonalai láthatók a $[-2, 2]^2$ tartományon, a negatív gradiensmezővel együtt. A függvénynek ebben a tartományban 4 stacionárius pontja van (*).
Adja meg a stacionárius pontok típusát, ha a negatív gradiensmezőt elég sűrű rácson ábrázoltuk ahhoz, hogy jól jellemezze a függvényt.



4. feladat

Határozza meg az

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

függvény stacionárius pontjait, és azok típusát.

5. feladat

Határozza meg az

$$f(x) = x_1^3 - x_2^3 + 6x_1x_2$$

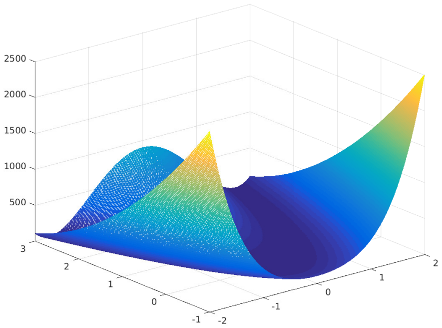
függvény stacionárius pontjait, és azok típusát.

6. feladat (Rosenbrock függvény)

Határozza meg az

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

függvény stacionárius pontjait, és azok típusát.



7. feladat

Egy autókereskedés egy adott autómárkából kombi és szedán típust is értékesít. Egy piackutatás eredménye azt mutatja, hogy ha a kombik ára x_1, a szedánoké x_2, akkor a kereslet a két autótípus iránt rendre

$$k = 10000 - 2x_1 + 2.5x_2$$
$$s = 16000 + 1.5x_1 - 3x_2$$

(ha az egyik típus ára emelkedik, akkor az ezirányú kereslet csökken, viszont a másik típusé nő). Hogyan érdemes megválasztani az egyes típusok árait, ha a bevételt maximalizálni szeretnénk?

Gradiens-módszer

Az $f : \mathbb{R}^2 \to \mathbb{R}$ függvény lokális minimumhelyeit keressük.

Egy $x^{(k)}$, $k = 0, 1, \ldots$ sorozatot definiálunk, mely optimális esetben közelíti a függvény egy lokális minimumhelyét.

A módszer adott, ha

- az $x^{(0)}$ kezdővektor adott,
- ismert az $x^{(k)} \mapsto x^{(k+1)}$ stratégia,
- adott a leállási feltétel.

A gradiens-módszer esetén az $x^{(k)}$ pontból a legmeredekebb leereszkedés irányában lépünk tovább.

Az $x^{(k)}$-beli legmeredekebb leereszkedés iránya: $-\nabla f(x^{(k)})$.

Gradiens módszer

- $x^{(0)}$ adott,
- ha $x^{(k)}$ adott, akkor

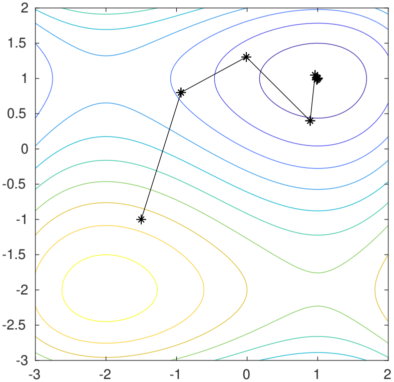
$$x^{(k+1)} = x^{(k)} - \alpha_k \nabla f(x^{(k)}), \quad k = 0, 1, \dots,$$

  ahol $\alpha_k > 0$ a lépéshossz. α_k értékét úgy határozzuk meg, hogy $x^{(k)}$-ból indulva $p_k = -\nabla f(x^{(k)})$ irányban meghatározzuk az f minimumhelyét, vagy annak egy elég jó közelítését.
- Leállási feltétel: ha $\|\nabla f(x^{(k)})\| < \varepsilon$, ahol $\varepsilon > 0$ adott paraméter.

Megjegyzés: Az α_k lépéshossz meghatározására különféle algoritmusok léteznek.

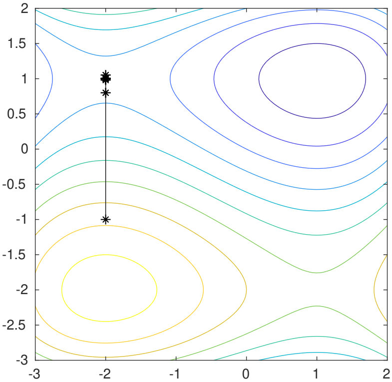
Példa

Vizsgáljuk meg a gradiens-módszer viselkedését az $f(x_1, x_2) = \frac{1}{4}(2x_1^3 + 3x_1^2 - 12x_1) + \frac{1}{2}(2x_2^3 + 3x_2^2 - 12x_2)$ függvény esetén.



$x^{(0)} = [-1.5, -1]^T$, $\varepsilon = 0.001$, az elvégzett lépések száma 10, $x^{(10)} = [1.0000, 0.9999]^T$.

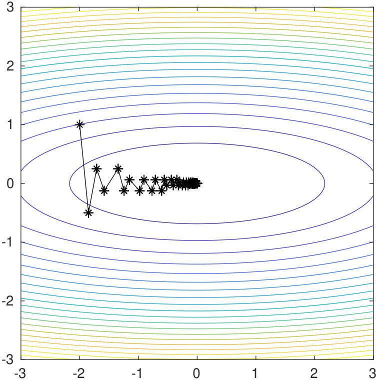
Előfordulhat, hogy a gradiens-módszer a függvény egy nyeregpontjában áll meg.



$x^{(0)} = [-2, -1]^T$, $\varepsilon = 0.001$, az elvégzett lépések száma 7, $x^{(7)} = [-2.0000, 1.0000]^T$.

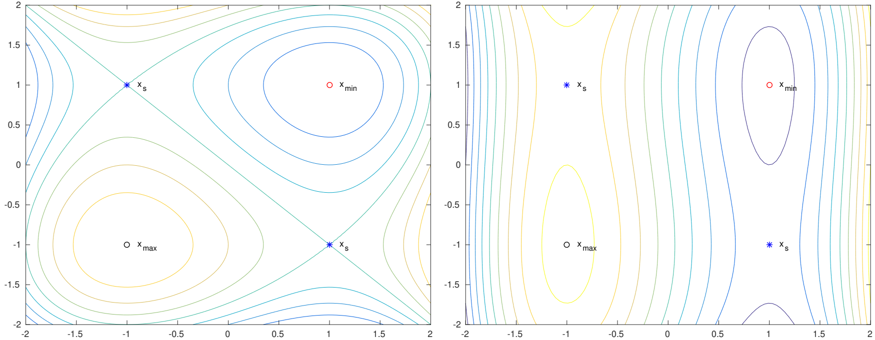
Megjegyzés

Ha a felület elnyújtott völgyeket tartalmaz, akkor a gradiens-módszer konvergenciája lassú lehet.



$f(x) = x_1^2 + 10x_2^2$,
$x^{(0)} = [-2, 1]^T$, $\varepsilon = 0.001$, az elvégzett lépések száma 78,
$x^{(78)} = [0.0000, 0.0000]^T$.

Gradiens módszer



Az

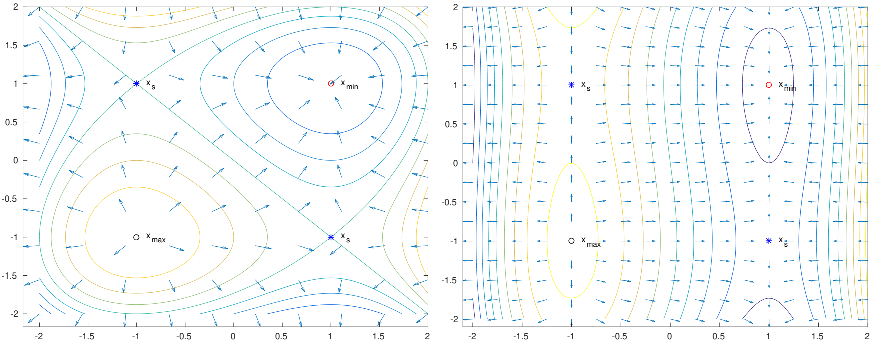
$$f(x_1, x_2) = x_1^3 + x_2^3 - 3x_1 - 3x_2$$

és a

$$f(x_1, x_2) = 10x_1^3 + x_2^3 - 30x_1 - 3x_2$$

függvény szintvonalai.

Gradiens módszer



Az

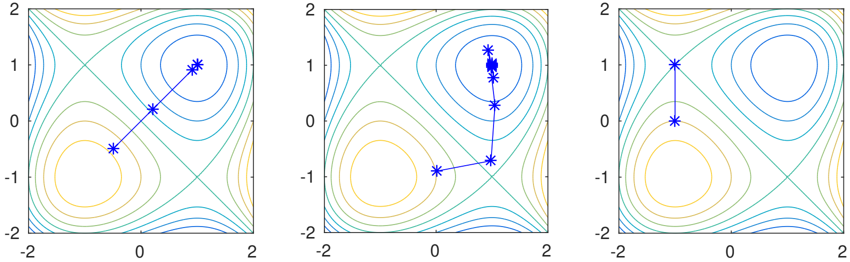
$$f(x_1, x_2) = x_1^3 + x_2^3 - 3x_1 - 3x_2$$

és a

$$f(x_1, x_2) = 10x_1^3 + x_2^3 - 30x_1 - 3x_2$$

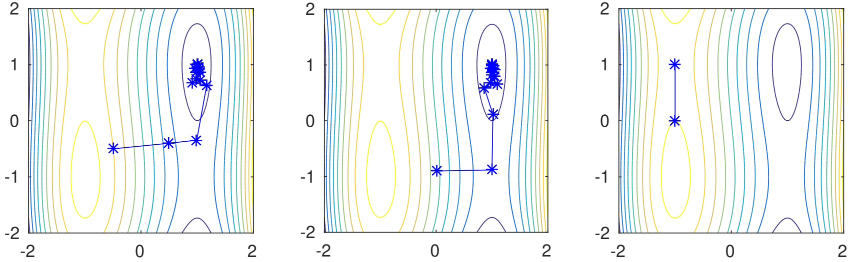
függvény szintvonalai és a negatív gradiensmezők.

Gradiens módszer



A gradiens módszer az $f(x_1, x_2) = x_1^3 + x_2^3 - 3x_1 - 3x_2$ függvény esetén.

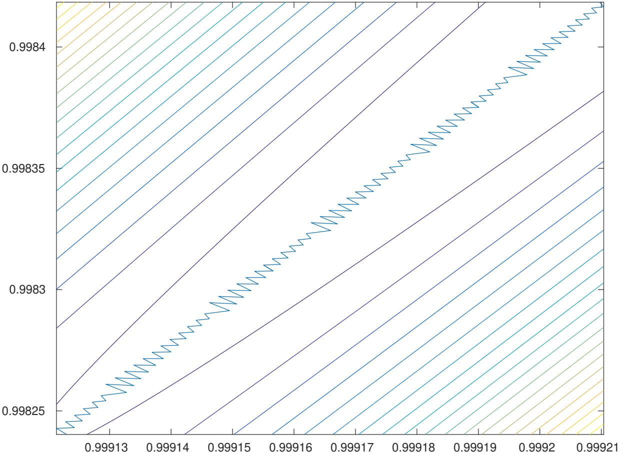
x_0	$(-0.5, -0.5)$	$(0, -0.9)$	$(-1, 0)$
lépés	6	11	2
x^*	$(1.0000, 1.0000)$	$(1.0000, 0.9999)$	$(-1, 1)$



A gradiens módszer az $f(x_1, x_2) = 10x_1^3 + x_2^3 - 30x_1 - 3x_2$ függvény esetén.

x_0	$(-0.5, -0.5)$	$(0, -0.9)$	$(-1, 0)$
lépésszám	36	33	2
x^*	$(1.0000, 0.9999)$	$(1.0000, 1.0001)$	$(-1, 1)$

Gradiens módszer



A gradiens módszer a Rosenbrock-függvényre, az utolsó 130 iterált. $x^{(0)} = (-1.2, 1)$, $\varepsilon = 10^{-3}$. Az elvégzett lépések száma 5231.

Newton-módszer optimalizálásra

Newton-módszer nemlineáris egyenletek gyökeinek közelítésére, emlékeztető:

Az $f(x) = 0$ (ahol $f : \mathbb{R} \to \mathbb{R}$) **nemlineáris egyenlet** gyökének közelítésére szolgáló Newton-iteráció:

$$x_0 \text{ adott}, \qquad x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \ldots$$

Az $F(x) = 0$ (ahol $F : \mathbb{R}^n \to \mathbb{R}^n$) **nemlineáris egyenletrendszer** gyökének közelítésére szolgáló Newton-iteráció:

$$x^{(0)} \text{ adott}, \qquad F'(x^{(k)})(x^{(k+1)} - x^{(k)}) = -F(x^{(k)}), \quad k = 0, 1, 2, \ldots$$

Newton-módszer optimalizálásra

Az f függvény minimumhelye megoldása a $\nabla f(x) = 0$ egyenletnek. Mivel $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$, ezért ez egy nemlineáris egyenletrendszer.

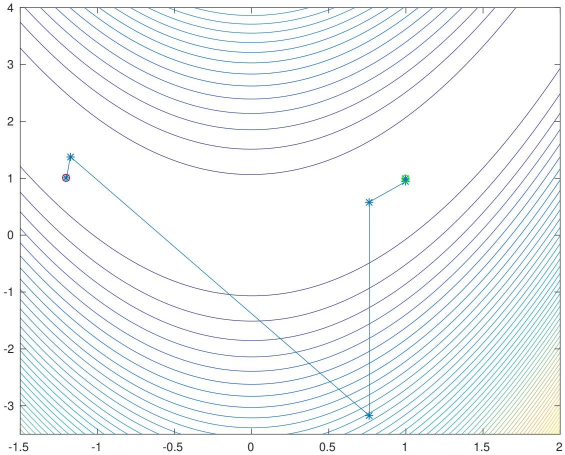
Ha f kétszer folytonosan differenciálható, akkor a Newton-módszer a $\nabla f(x) = 0$ egyenletre:

$$x^{(0)} \text{ adott}, \quad H(x^{(k)})(x^{(k+1)} - x^{(k)}) = -\nabla f(x^{(k)}), \quad k = 0, 1, 2, \dots$$

ahol H az f függvény Hesse-mátrixa.

- $x^{(0)}$ adott,
- $H(x^{(k)})p_k = -\nabla f(x^{(k)})$, (azaz $p_k = -(H_k)^{-1}\nabla f_k$)
- $x^{(k+1)} = x^{(k)} + p_k$
- ha $\|\nabla f_k\| < \varepsilon$, akkor leállás

Newton-módszer, példa



A Newton-módszer a Rosenbrock-függvényre. $x^{(0)} = (-1.2, 1)$,
$x_{opt} = (0.999996, 0.999991)$, $f(x_{opt}) = 1.8 \cdot 10^{-11}$, $k = 5$.