

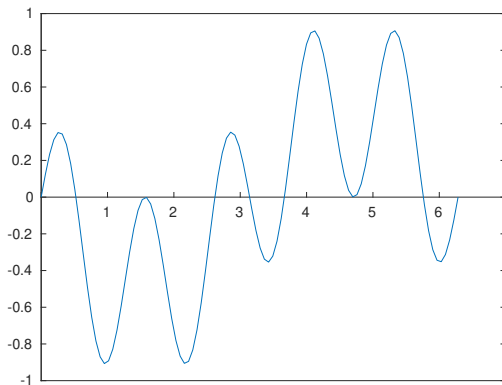
Alkalmazott matematika

Baran Ágnes

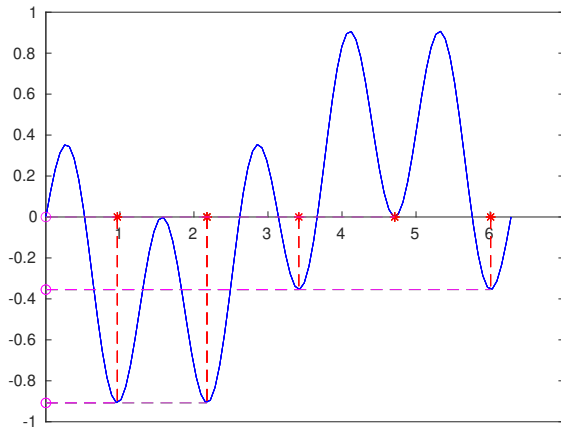
Optimalizáció

Egyváltozós függvény szélsőértéke (emlékeztető)

Az $f : \mathbb{R} \rightarrow \mathbb{R}$ függvény szélsőérték helyeit keressük.



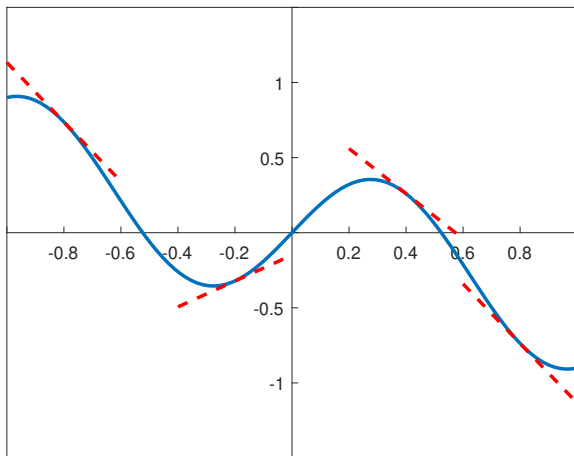
Egy függvénynek több **lokális szélsőérték**e is lehet.



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

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

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

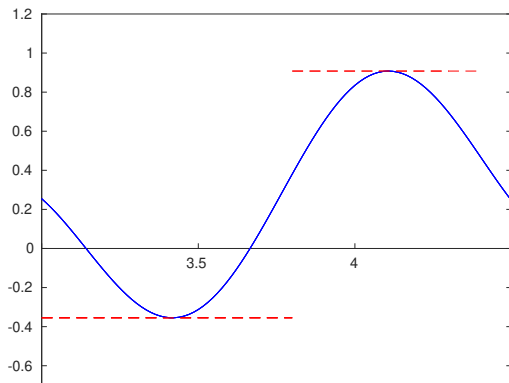


Emlékeztető: az f függvény x_0 -beli érintője:

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

Szükséges feltétel

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

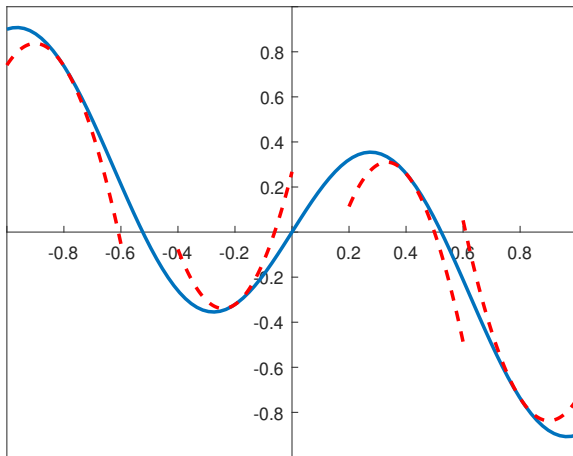


Fordítva nem igaz! Abból, hogy $f'(x) = 0$ NEM következik, hogy a függvénynek ott lokális szélsőértéke van.

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

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

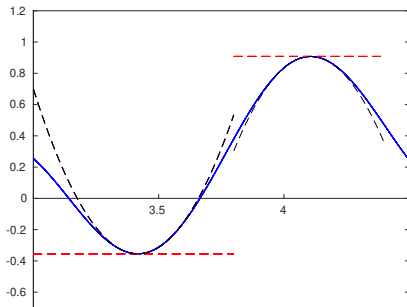
másodfokú polinommal.



Szélsőérték, elégséges feltételek

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

- $f'(x^*) = 0$, $f''(x^*) > 0$, akkor f -nek x^* -ban lokális minimuma van.
- $f'(x^*) = 0$, $f''(x^*) < 0$, akkor f -nek x^* -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 $f : \mathbb{R} \rightarrow \mathbb{R}$ függvény szélsőértékhelyeit keressük.

Szükséges feltétel

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

Elégséges feltételek

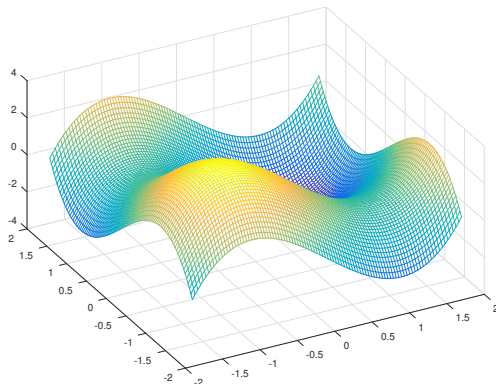
Ha az $f : \mathbb{R} \rightarrow \mathbb{R}$ függvény kétszer differenciálható és

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

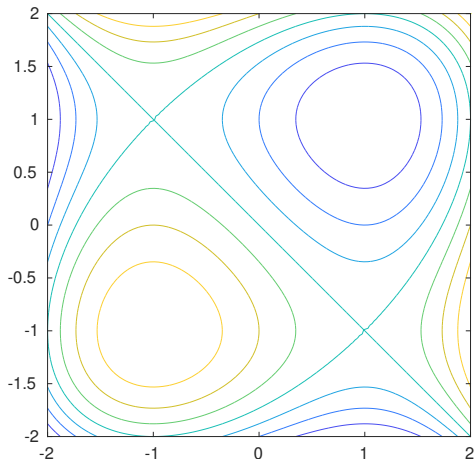
Többváltozós függvények minimalizálása

Egy $f : \mathbb{R}^n \rightarrow \mathbb{R}$ függvény (lokális) minimumhelyét keressük.

Például: $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = x_1^3 + x_2^3 - 3x_1 - 3x_2$.



Rajzoljuk ki a függvény **szintvonalait** is!



Gradiens

Az $f : \mathbb{R}^n \rightarrow \mathbb{R}$ függvény x -beli **gradiense**

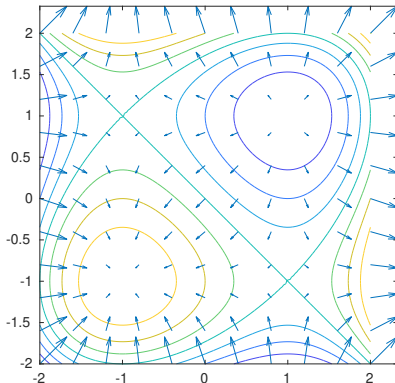
$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{bmatrix}$$

Példa

Az $f(x) = x_1^3 + x_2^3 - 3x_1 - 3x_2$ függvény gradiense:

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

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



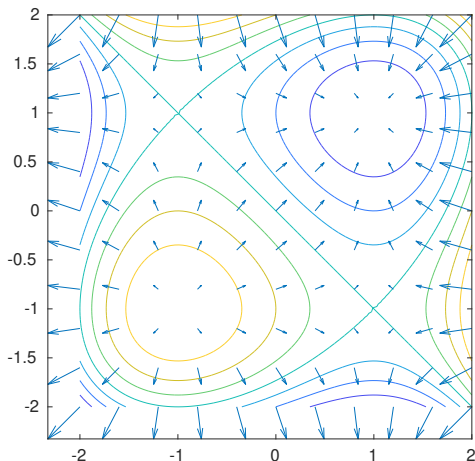
Az $f(x) = x_1^3 + x_2^3 - 3x_1 - 3x_2$ függvény szintvonalai és a gradiensmező.

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 $f(x) = x_1^3 + x_2^3 - 3x_1 - 3x_2$ függvény szintvonalai és a **negatív** gradiensmező.

A lokális szélsőérték feltételei

Elsőrendű szükséges feltétel

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

Definíció (Stacionárius pont)

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

Megjegyzés

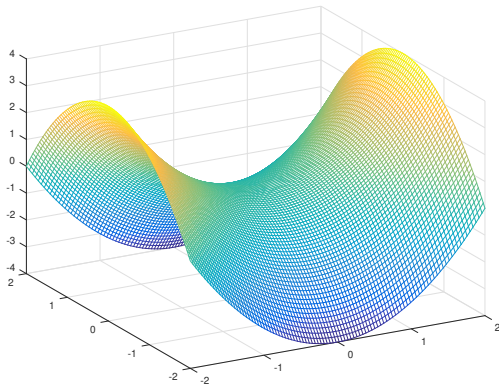
Ha x^* stacionárius pontja f -nek, akkor stacionárius pontja $-f$ -nek is, azaz a stacionárius pont lokális maximum is lehet.

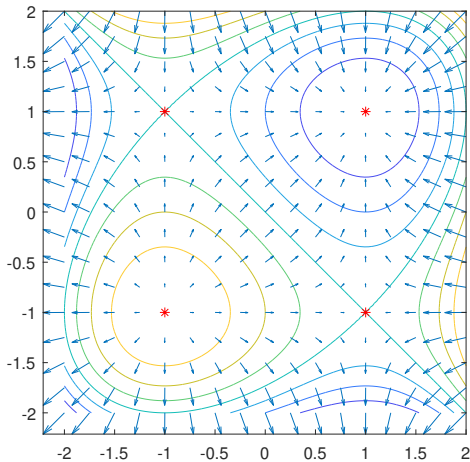
Definíció (Nyeregpont)

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





Az $f(x) = x_1^3 - 3x_1 + x_2^3 - 3x_2$ függvény szintvonalai, a **negatív** gradiensmező és a stacionárius pontok.

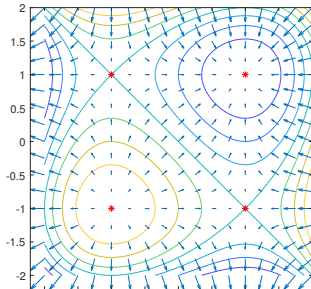
Többváltozós függvény minimuma Matlab-bal

Az `fminunc` és `fminsearch` függvényeket használhatjuk.

Mindkettő egy sorozattal közelíti a minimumhelyet, ehhez egy kezdővektorra van szükségük.

A kezdővektor megválasztása erősen befolyásolhatja az algoritmus sikeres lefutását.

Vizsgáljuk meg milyen eredményt kapunk a fenti függvényeket az $f(x) = x_1^3 - 3x_1 + x_2^3 - 3x_2$ függvénnyel és az alábbi kezdővektorokkal meghívva:



$$x_0 = [-0.5, -0.5],$$

$$x_0 = [-0.5, 0],$$

$$x_0 = [-1, -0.5],$$

$$x_0 = [-1.5, -1.5]$$

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.

