Operációkutatás I.

Szegedi Tudományegyetem
Informatikai Intézet
Számítógépes Optimalizálás Tanszék

11. Előadás

Az optimalizálás alapfeladata

Keressük f függvény maximumát

$$\max_{\mathbf{x}} f(\mathbf{x})$$

ahol $f : \mathbb{R}^n \to \mathbb{R}$ és $\mathbf{x} \in \mathbb{R}^n$.

Lehetnek adottak korlátozó feltételek: $g_i(\mathbf{x}) \leq 0, i = 1, \ldots, m$ és/vagy $g_j(\mathbf{x}) = 0, j = m+1, \ldots, m+k$. Ha adottak, feltételes, különben feltétel nélküli optimalizálási feladatról beszélünk.

- f és g_i függvények lehetnek lineárisak (LP), vagy nemlineárisak (NLP),
- de feltesszük, hogy folytonosak
- $\mathbf{x}$-re nincs g_i-ken kívül megszorítás (pl. nem köthetjük meg, hogy egészek, mert akkor MINLP-vé válik a feladat)
- jelöljük $\mathcal{S}$-sel a lehetséges megoldások halmazát

Az optimalizálás alapfeladata

Mit jelent az, hogy $\mathbf{x}^*$ egy megoldás?

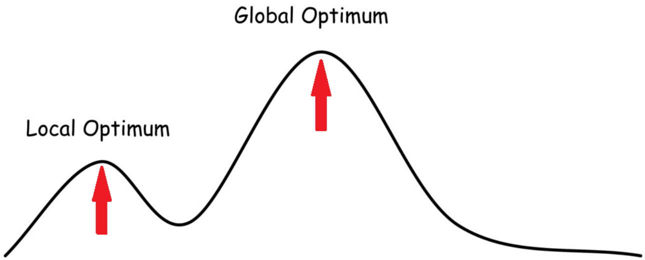
Lokális optimum

Egy $\mathbf{x}^*$ lokális maximum (optimum), ha $\exists \delta > 0$, úgy, hogy $f(\mathbf{x}^*) \geq f(\mathbf{x})$ $\forall \mathbf{x} \in \mathcal{S}$ és $||\mathbf{x} - \mathbf{x}^*|| < \delta$ esetén.

$||.||$ norma távolságot jelöl, pl. az Eulédeszi-távolságot, ami $||x - y||_2 = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}$

Globális optimum

Egy $\mathbf{x}^*$ globális maximum (optimum), ha $f(\mathbf{x}^*) \geq f(\mathbf{x})$ $\forall \mathbf{x} \in \mathcal{S}$ esetén.

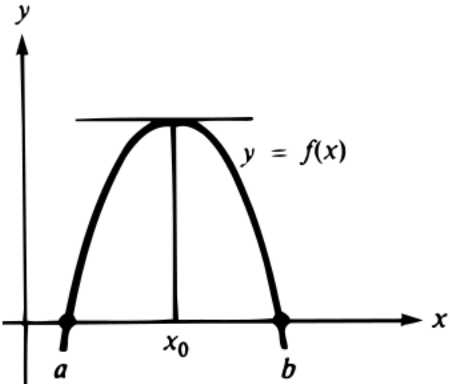


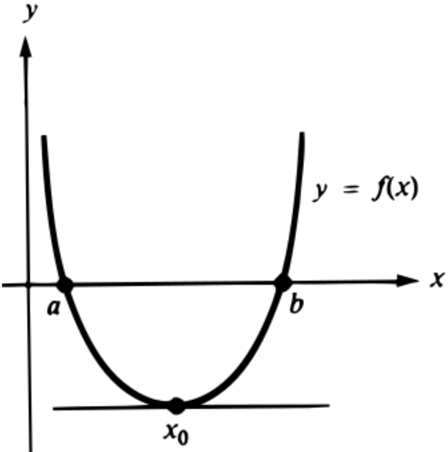
Lokális vs. globális optimum feltétel nélküli feladatnál

Kérdés

Mikor lesz minden lokális maximum (minimum) globális is?

Ha f konkáv (konvex) függvény, akkor biztosan.





Feltétel nélküli optimalizálás: Szükséges és elegendő feltétel optimalitásra

Tétel (szükséges feltétel). *Ha x^* az $f(x)$ függvény maximuma (minimuma), akkor $f'(x^*) = 0$.*

Ez nem elegendő, ld. pl. $f(x) = x^3$ függvény $x = 0$-ban.

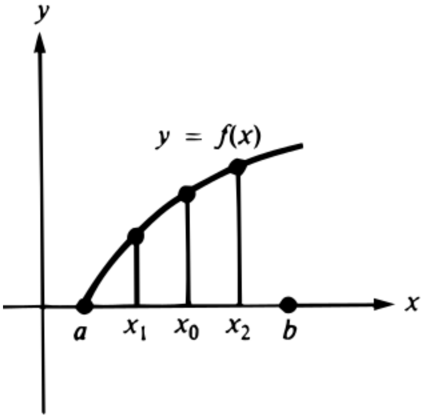
Tétel (elegendő feltétel). *Ha $f'(x^*) = 0$ és $f''(x^*)$ negatív (pozitív), akkor x^* lokális maximuma (minimuma) f-nek.*

Azokat a pontokat, ahol $f'(x) = 0$ **stacionárius pont**oknak nevezzük. A

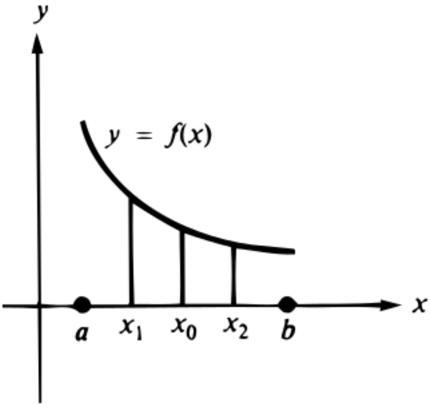
$$\max_{x \in [a,b]} f(x)$$

feladat megoldását a stacionárius pontjai illetve a és b közt keressük. A következő fóliákon erre a feladatra nézünk példákat különböző tulajdonságú függvények mellett.

Példák monoton függvényre

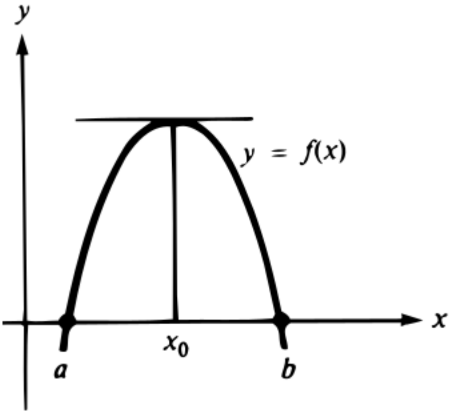


(a) $f'(x_0) > 0$
$f(x_1) < f(x_0)$
$f(x_2) > f(x_0)$
x_0 nem lokális szélsőértékhely

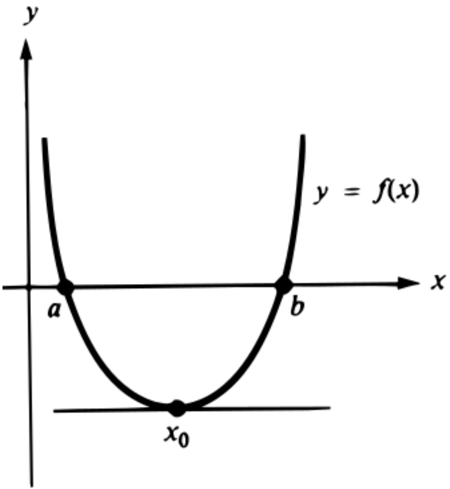


(b) $f'(x_0) < 0$
$f(x_1) > f(x_0)$
$f(x_2) < f(x_0)$
x_0 nem lokális szélsőértékhely

Példák konkáv és konvex függvényekre

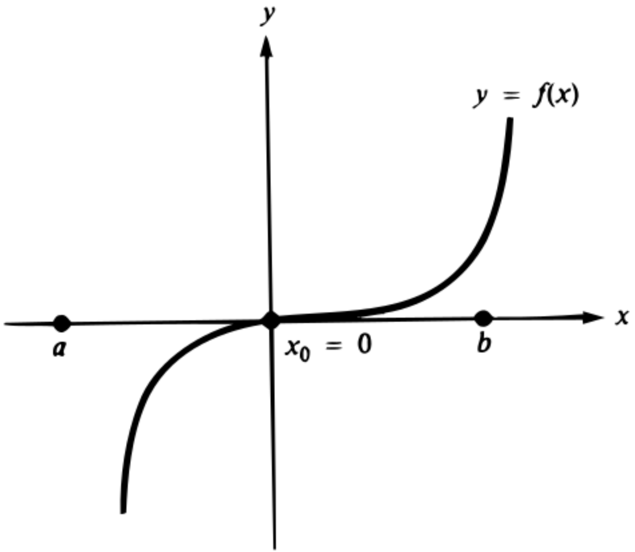


(c) $f'(x_0) = 0$
Ha $x < x_0$, akkor $f'(x) > 0$
Ha $x > x_0$, akkor $f'(x) < 0$
x_0 lokális maximum



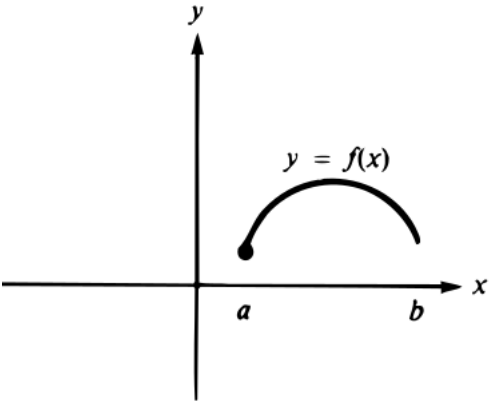
(d) $f'(x_0) = 0$
Ha $x < x_0$, akkor $f'(x) < 0$
Ha $x > x_0$, akkor $f'(x) > 0$
x_0 lokális minimum

Példa se nem konkáv se nem konvex függvényre

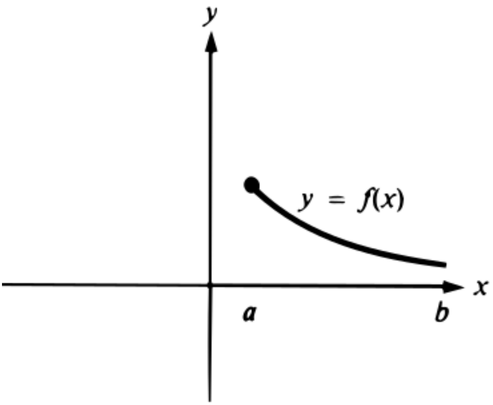


(e) $x_0 = 0$ nem lokális maximum vagy lokális minimum, de $f'(x_0) = 0$

Példa optimumra nem 0 deriválttal



(a) $f'(a) > 0$
a lokális minimum



(b) $f'(a) < 0$
a lokális maximum

Legmeredekebb lejtő módszere

Több dimenziós feladatoknál az $f'(x) = 0$ feltételt felváltja a

$$\nabla f(\mathbf{x}) = (f'_{x_1}(\mathbf{x}), f'_{x_2}(\mathbf{x}), \ldots, f'_{x_n}(\mathbf{x})) = \mathbf{0}$$

szükséges feltétel.
Ez nem feltétlenül vezet könnyen megoldható egyenletrendszerre, ezért iteratív megoldáshoz folyamodunk.

Lokális kereső eljárás

Legyen $\mathbf{x}_0$ a kezdő megoldásunk, az iterációs lépés pedig

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \mu \mathbf{d}_i, \quad i = 0{,}1{,}2, \ldots$$

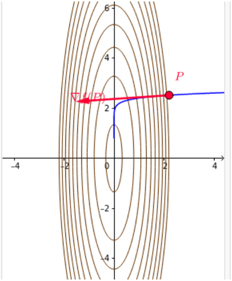
ahol $\mathbf{d}_i$ egy növekvő irány maximalizálásnál, μ pedig a lépéshossz.

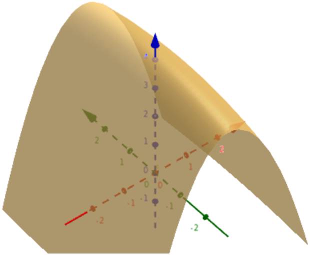
Ha $||\mathbf{x}_{i+1} - \mathbf{x}_i|| \le \epsilon$, azaz az új lépes nem elég nagy, STOP.

Itt $\mathbf{x}_i$ az i. lépésben kapott n-dimenziós vektor.

Legmeredekebb lejtő módszere

Ahogy $f'(x)$ a növekvő irányba mutat 1D-ben, úgy $\nabla f(\mathbf{x})$ a legmeredekebb növekedés irányába mutat nD-ben. Így $\mathbf{d}_i = \nabla f(\mathbf{x}_i)$ jó választás.





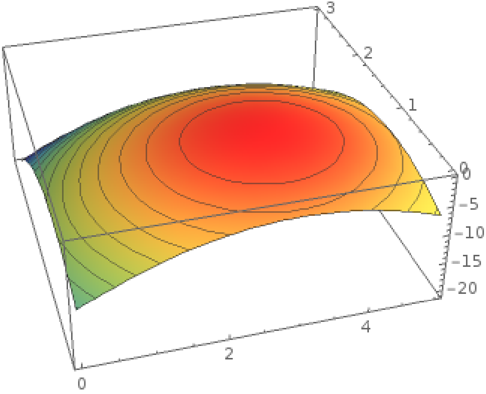
Geogebra

Lépéshossz meghatározása

μ-t vagy az egydimenziós $\max_\mu f(\mathbf{x}_i + \mu\mathbf{d}_i)$ feladat megoldásával, vagy közelítéssel határozzuk meg.

Legmeredekebb lejtő: példa

Példa. Maximalizáljuk az $f(x,y) = -(x-3)^2 - 3(y-1)^2$ függvényt!



$$\nabla f = \begin{pmatrix} f'_x \\ f'_y \end{pmatrix} =$$

$$= \begin{pmatrix} -2(x-3) \\ -6(y-1) \end{pmatrix} = \begin{pmatrix} 6-2x \\ 6-6y \end{pmatrix}$$

$\mathbf{x}_0 = (0{,}0)$

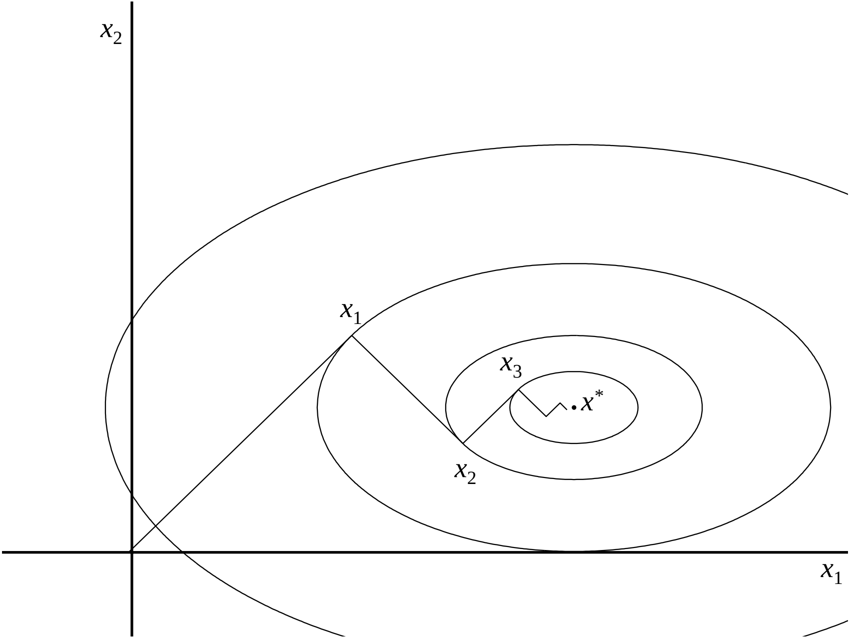
$\mathbf{x}_1 = (0{,}0) + \mu(6{,}6) = (6\mu{,}6\mu)$,

vagyis

$\max_\mu f(6\mu{,}6\mu)$-t keressük!

$$\max_\mu f(6\mu{,}6\mu) = \max -(6\mu-3)^2 - 3(6\mu-1)^2 = \max -144(\mu-1/4)^2 - 3$$

$$\mu = 1/4 \quad \rightarrow \quad \mathbf{x}_1 = (6/4{,}6/4)\,!!!$$


x_2
x_1
x_3
x^*
x_2
x_1

Feltételes optimalizálás: Lagrange függvény

Adott a következő feltételes probléma

$$\begin{aligned} \max_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) \\ & g_j(\mathbf{x}) = 0, \quad j = 1, \dots, m \end{aligned} \tag{1}$$

Az

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \sum_{j=1}^{m} \lambda_j g_j(\mathbf{x})$$

függvényt a feladat **Lagrange függvény**ének nevezzük.

Tétel (szükséges feltétel). *Ha $\mathbf{x}^*$ az (1) feladat maximuma, akkor létezik λ^* amire $\mathbf{x}^*$ a Lagrange függvény stacionárius pontja, azaz $\nabla\mathcal{L}(\mathbf{x}^*, \lambda^*) = 0$.*

A λ vektor elemeit **Lagrange multiplikátor**oknak nevezzük.

Lagrange függvény – példa

Példa. Maximalizáljuk $f(x,y) = x + y$ függvényt, ha $x^2 + y^2 = 1$.
$g(x,y) = 0$, ahol $g(x,y) = x^2 + y^2 - 1$.

Lagrange függvény.

$$\mathcal{L}(x,y,\lambda) = x + y - \lambda(x^2 + y^2 - 1) = x + y - \lambda x^2 - \lambda y^2 + \lambda$$

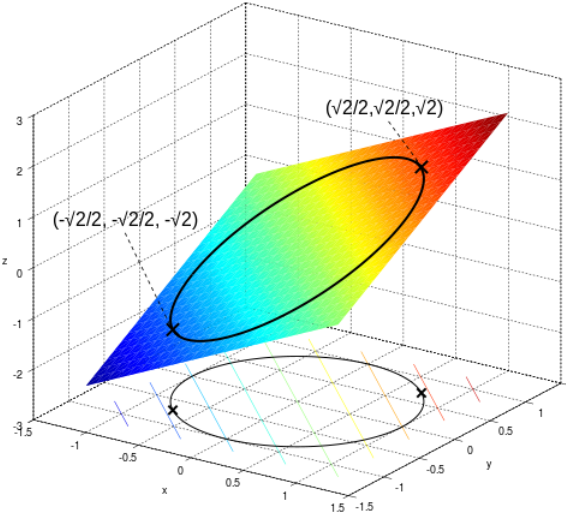
A szükséges feltétel

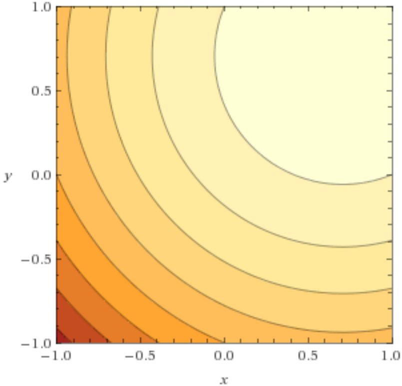
$$\nabla\mathcal{L}(x,y,\lambda) = \begin{pmatrix} \mathcal{L}'_x \\ \mathcal{L}'_y \\ \mathcal{L}'_\lambda \end{pmatrix} = \begin{pmatrix} 1 - 2\lambda x \\ 1 - 2\lambda y \\ -x^2 - y^2 + 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Megoldva: $x = y = 1/(2\lambda)$, $\lambda = \pm 1/\sqrt{2}$
így $(x,y) = (\sqrt{2}/2, \sqrt{2}/2)$, illetve $(x,y) = (-\sqrt{2}/2, -\sqrt{2}/2)$.

Ebből $f(\sqrt{2}/2, \sqrt{2}/2) = \sqrt{2}$ a maximum, $f(-\sqrt{2}/2, -\sqrt{2}/2) = -\sqrt{2}$ a minimum.

Lagrange függvény – példa



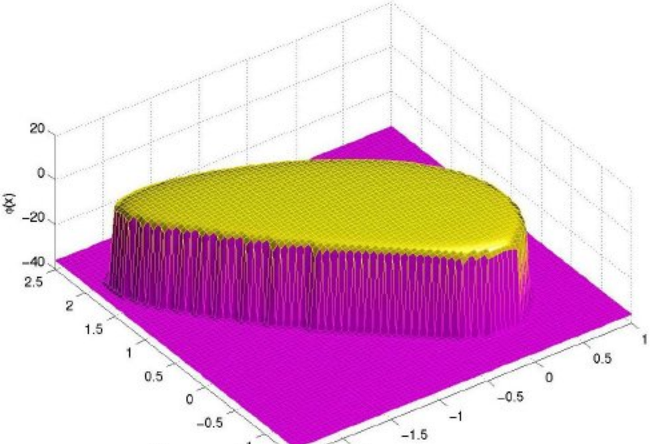


A Lagrange fv $\lambda = 1/\sqrt{2}$-vel

Korlátozófüggvény módszer

Motiváció: ahelyett, hogy megoldanánk egy feltételes optimalizálási feladatot, beépítjük a feltételeket a célfüggvénybe. (hasonlóan mint a Lagrange függvénynél)

Ötlet: Tegyük fel, hogy adott egy kezdő lehetséges megoldásunk ($\mathbf{x}_0 \in \mathcal{S}$). „Készítsünk szakadékot" az $\mathcal{S}$ határán.



Korlátozófüggvény módszer

Az eredeti feladat:

$$\begin{array}{ll} \max_{\mathbf{x}\in\mathbb{R}^n} & f(\mathbf{x}) \\ \text{feltéve} & g_i(\mathbf{x}) \leq 0, \quad i = 1,\ldots,m \end{array}$$

A korlátozófüggvény (barrier):

$$B_\mu(\mathbf{x}) = f(\mathbf{x}) + \mu\phi(\mathbf{x})$$

ahol $\phi(\mathbf{x})$ tartson $-\infty$-hez ha $\mathbf{x}$ tart a megengedett tartomány határához, de legyen kis negatív ha még megengedett megoldás!

A módszer

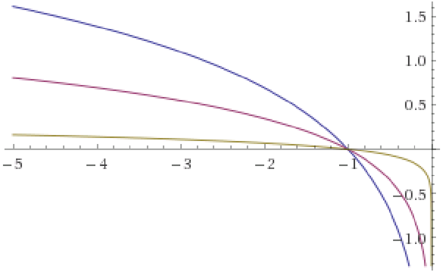
Induljunk ki egy $\mathbf{x}_0$ megengedett kezdő megoldásunkból, oldjuk meg a $\max B_\mu$ feladatot, majd a kapott megoldásból μ-t fokozatosan csökkentve újra meg újra oldjuk meg a $\max B_\mu$ feladatot!

Korlátozófüggvény módszer : A ϕ függvény lehetséges választásai

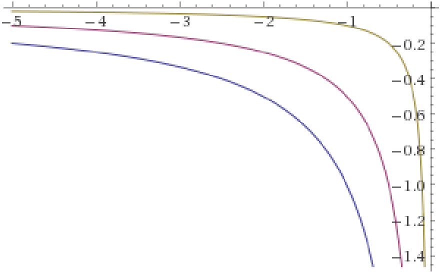
$$\phi(\mathbf{x}) = \sum_{i=1}^{m} \log(-g_i(\mathbf{x}))$$

$$\phi(\mathbf{x}) = \sum_{i=1}^{m} \frac{1}{g_i(\mathbf{x})}$$

Az $x \leq 0$ feltétel esetén
$\mu\phi(x) = \mu \log(-x)$



Az $x \leq 0$ feltétel esetén
$\mu\phi(x) = \mu\frac{1}{x}$



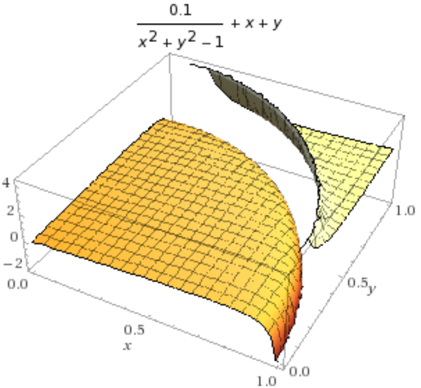
A μ paraméter megválasztása kontrollálja az akadály szigorúságát :

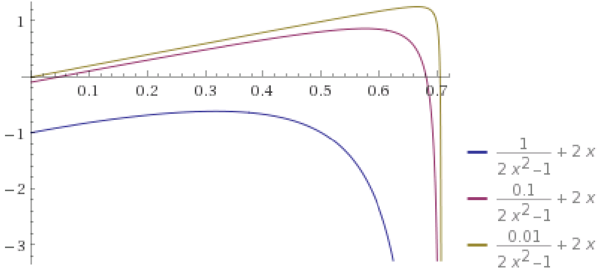
- μ nagy : fokozatos akadály (kék)
- μ kicsi : éles akadály (barna)

Korlátozófüggvény módszer: példa

Példa. $\max x + y$ ha $x^2 + y^2 \leq 1$

$B_\mu = x + y + \frac{\mu}{x^2+y^2-1}$





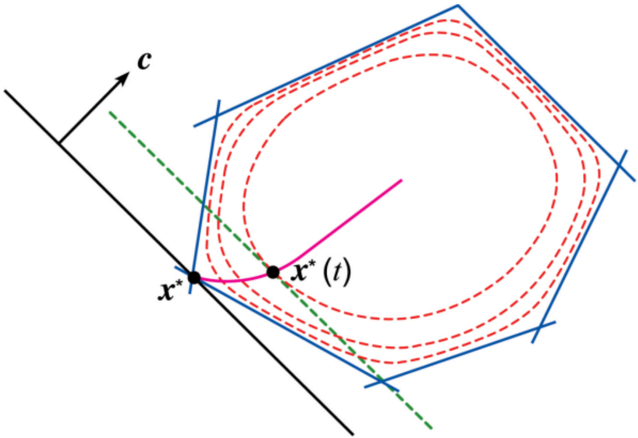
Induljunk ki egy $\mathbf{x}_0$ megengedett kezdő megoldásból, maximalizáljuk a kék B_1 függvényt, majd annak optimumából indulva μ-t csökkentve közelítsük $B_{0.1}$-et, majd annak a megoldásából indulva $B_{0.01}$-et, és így tovább!

Korlátozófüggvény módszer: Megjegyzések

- A B_μ feladatok megoldása tart az eredeti feladat megoldásához ha $\mu \to 0$, de nem éri el ha $g_i(\mathbf{x}^*) = 0$ ($B_\mu(\mathbf{x}^*) = -\infty\ \forall \mu > 0$).
- Csak egyenlőtlenség feltételek mellett alkalmazható, különben nincs lehetséges megoldása!
- Kis μ esetén a korlátozó (barrier) függvény rosszul kondícionált, azaz numerikusan nehezen megoldható (túl nagy a célfüggvényérték különbség kis lépésre is).
- A feltételeket nem hagyhatjuk figyelmen kívül: ha kilépünk a lehetséges tartományból a függvényünk vagy nem értelmezett (logaritmus), vagy rossz megoldást ad (reciprok).

Szimplex vs. belső pontos módszer

A korlátozófüggvény módszer lineáris programozási feladatra való alkalmazásából született az ú.n. **belső pontos módszer**!



Youtube

Büntetőfüggvény

Mi a helyzet akkor, ha **nincs meg a kezdő lehetséges megoldás** (azaz egy belső pont...)?

Ötlet: Legyen a **büntető függvény**

$$P_\rho(\mathbf{x}) = f(\mathbf{x}) + \rho\psi(\mathbf{x}),$$

ahol $\rho < 0$ szám a büntető paraméter és

$$\psi(\mathbf{x}) = \begin{cases} 0, & \text{ha } \mathbf{x} \text{ lehetséges megoldás} \\ > 0, & \text{különben.} \end{cases}$$

Cél:

$$\max_{\mathbf{x}} P_\rho(\mathbf{x})$$

problémák sorozatának megoldása, ahol $\rho \to -\infty$.
Vagyis: a nem lehetséges megoldások esetén a célfüggvény legyen egyre kisebb.

Büntetőfüggvény

Tekintsük a

$$\begin{array}{lll} \max_{\mathbf{x}\in\mathbb{R}^n} & f(\mathbf{x}) & \\ \text{feltéve} & g_j(\mathbf{x}) \le 0, & j = 1,\ldots,m \\ \text{feltéve} & g_j(\mathbf{x}) = 0, & j = m+1,\ldots,l \end{array}$$

feltételekkel adott feladatot.

A ψ függvény lehetséges választásai:

$$\psi(\mathbf{x}) = \sum_{i=1}^{m} \max\{0, g_i(\mathbf{x})\} + \sum_{i=m+1}^{l} |g_i(\mathbf{x})|$$

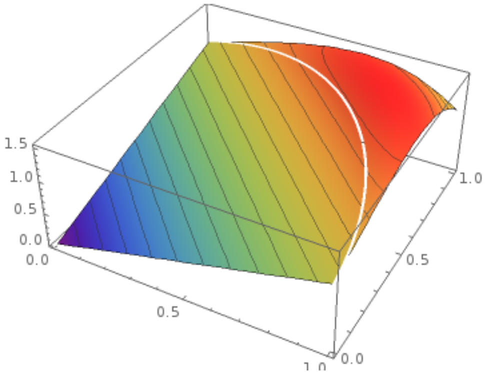
Vagy a négyzetes büntetőfüggvény

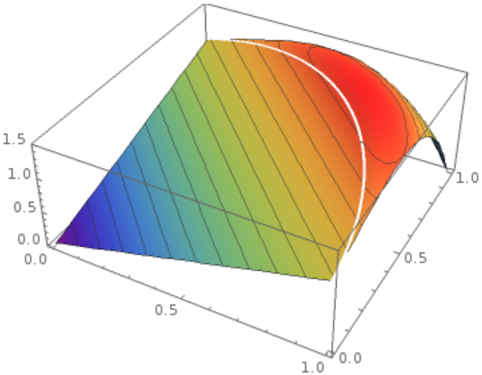
$$\psi(\mathbf{x}) = \sum_{i=1}^{m} \max\{0, g_i(\mathbf{x})\}^2 + \sum_{i=m+1}^{l} g_i(\mathbf{x})^2$$

Büntetőfüggvény: példa

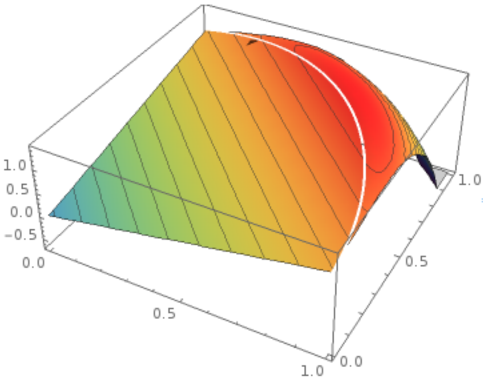
Példa. $\max x + y$ ha $x^2 + y^2 \leq 1$

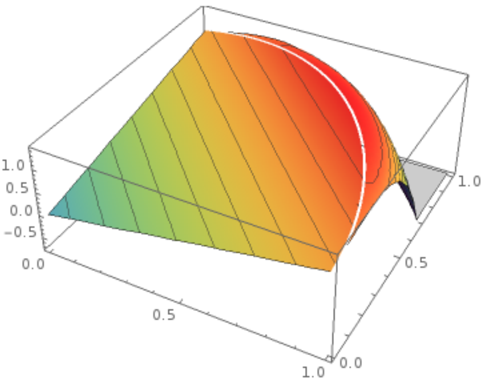
$$P_\rho = x + y + \rho \max\{0, x^2 + y^2 - 1\}^2$$

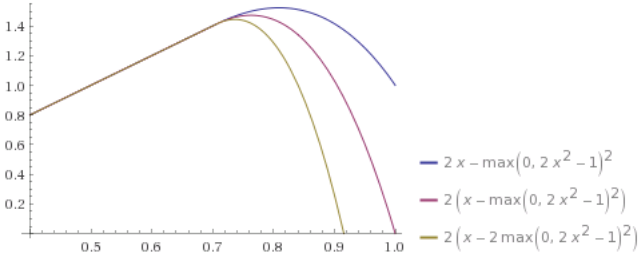




Büntetőfüggvény: példa





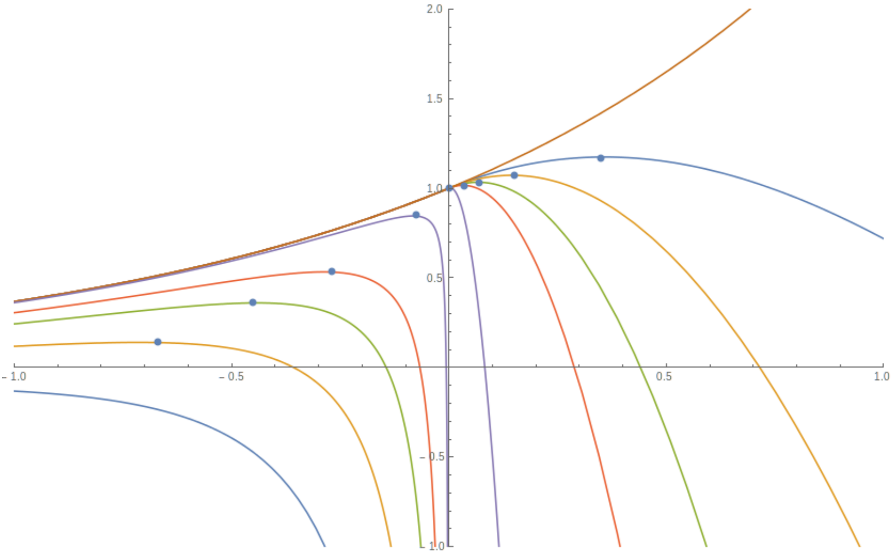


Büntetőfüggvény

- Nem tudjuk előre, mekkora ρ fog kelleni
- $\rho < 0$, ha $|\rho|$ túl nagy, a feladat rosszul kondicionált
- Csökkentsük fokozatosan ρ értékét
- A keresést mindig az előző megoldásból indítsuk
- Ha konvergál a függvényérték, megállunk
- Van olyan feladat amire csak nem megengedett megoldást ad (ld. következő ábra)

Büntetőfüggvény – ábra

Példa. $\max e^x$ f.h. $x \leq 0$



Korlátozófüggvény vs büntetőfüggvény – ábra

Korlátozó függvények Büntető függvények