Nelineárne optimalizačné modely a metódy

Téma prednášky č. 3

Konvexná analýza

(Časť 1)

Prof. Ing. Michal Fendek, CSc.

Katedra operačného výskumu a ekonometrie
Ekonomická univerzita
Dolnozemská 1
852 35 Bratislava

Continue F 10, 27

Definícia 3.1

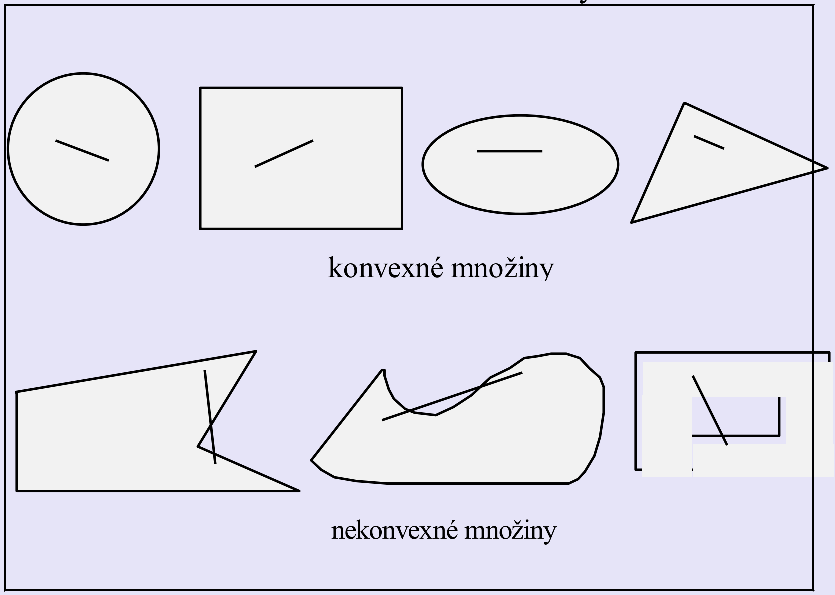
Hovoríme, že množina K je konvexná, ak pre ľubovoľné dva body tejto množiny $\mathbf{x}^1$, $\mathbf{x}^2 \in K$ a ľubovoľné $\lambda \in \langle 0,1 \rangle$ platí, že ich každá konvexná lineárna kombinácia

$$\mathbf{x} = \lambda \mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2$$

je tiež prvkom tejto množiny. Inými slovami, úsečka spájajúca ľubovoľné dva body konvexnej množiny leží v tejto množine.

Na obr. č. 3.1 sú uvedené príklady konvexných a nekonvexných množín.

Obr.č.3.1: Konvexné a nekonvexné množiny



Príklad

Vlastnosti riešenia úlohy nelineárneho programovania

Definícia 3.2

Nech $f(\mathbf{x}) : D \to R$, kde $D \subset R^n$ je neprázdna a konvexná množina. Hovoríme, že funkcia $f(\mathbf{x})$ je konvexná na množine D, ak pre ľubovoľné dva body tejto množiny $\mathbf{x}^1, \mathbf{x}^2 \in D$ a ľubovoľné $\lambda \in \langle 0,1 \rangle$ platí,

$$f[\lambda \mathrm{x}^1 + (1 - \lambda)\mathrm{x}^2] \leq \lambda f(\mathrm{x}^1) + (1 - \lambda) f(\mathrm{x}^2) \qquad (3.1)$$

Inými slovami, úsečka spájajúca ľubovoľné dva body krivky grafu funkcie leží na krivke, alebo nad krivkou grafu funkcie.

Poznámka 3.1

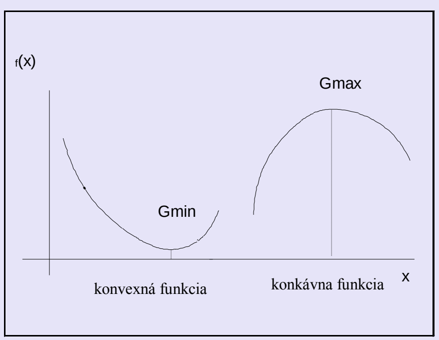
Hovoríme, že funkcia $f(\mathbf{x}) : D \to R$ je rýdzokonvexná na množine D, ak pre ľubovoľné dva body tejto množiny $\mathbf{x}^1, \mathbf{x}^2 \in D$ a ľubovoľné $\lambda \in (0,1)$ platí vo vzťahu (3.1) ostrá nerovnosť.

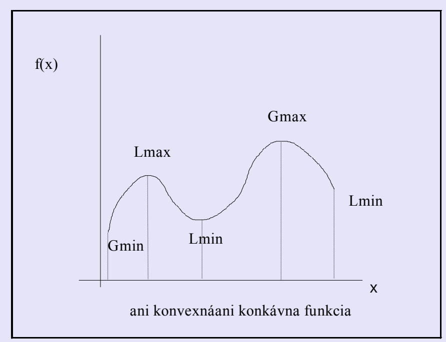
Poznámka 3.2

Hovoríme, že funkcia $f(\mathbf{x}) : D \to R$ je konkávna, resp. rýdzokonkávna na množine D, ak funkcia $-f(\mathbf{x})$ je konvexná, resp. rýdzokonvexná.

□□□

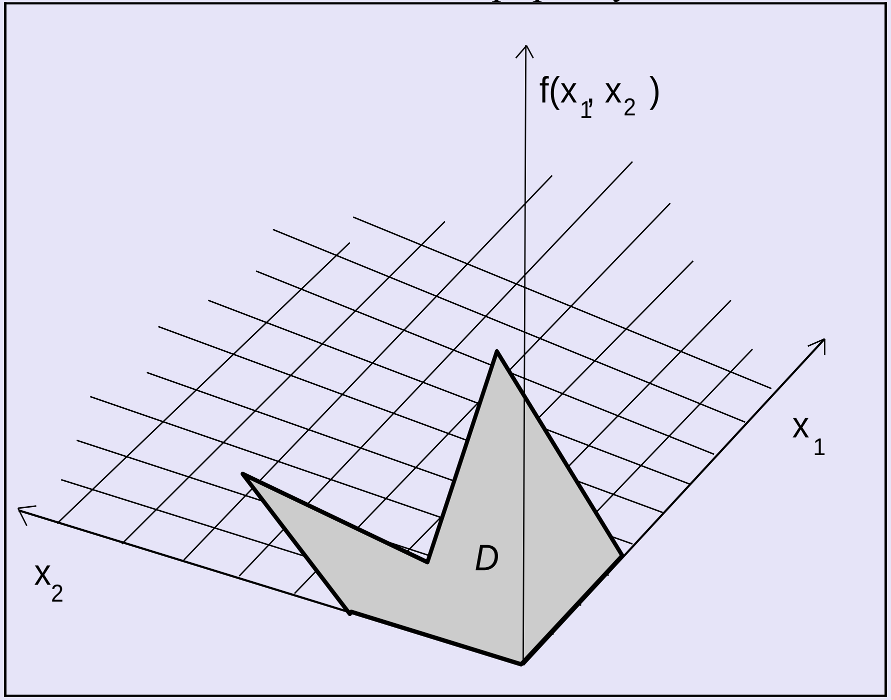
Obr.č.3.2a,b: Konvexné, konkávne a nekonvexné funkcie





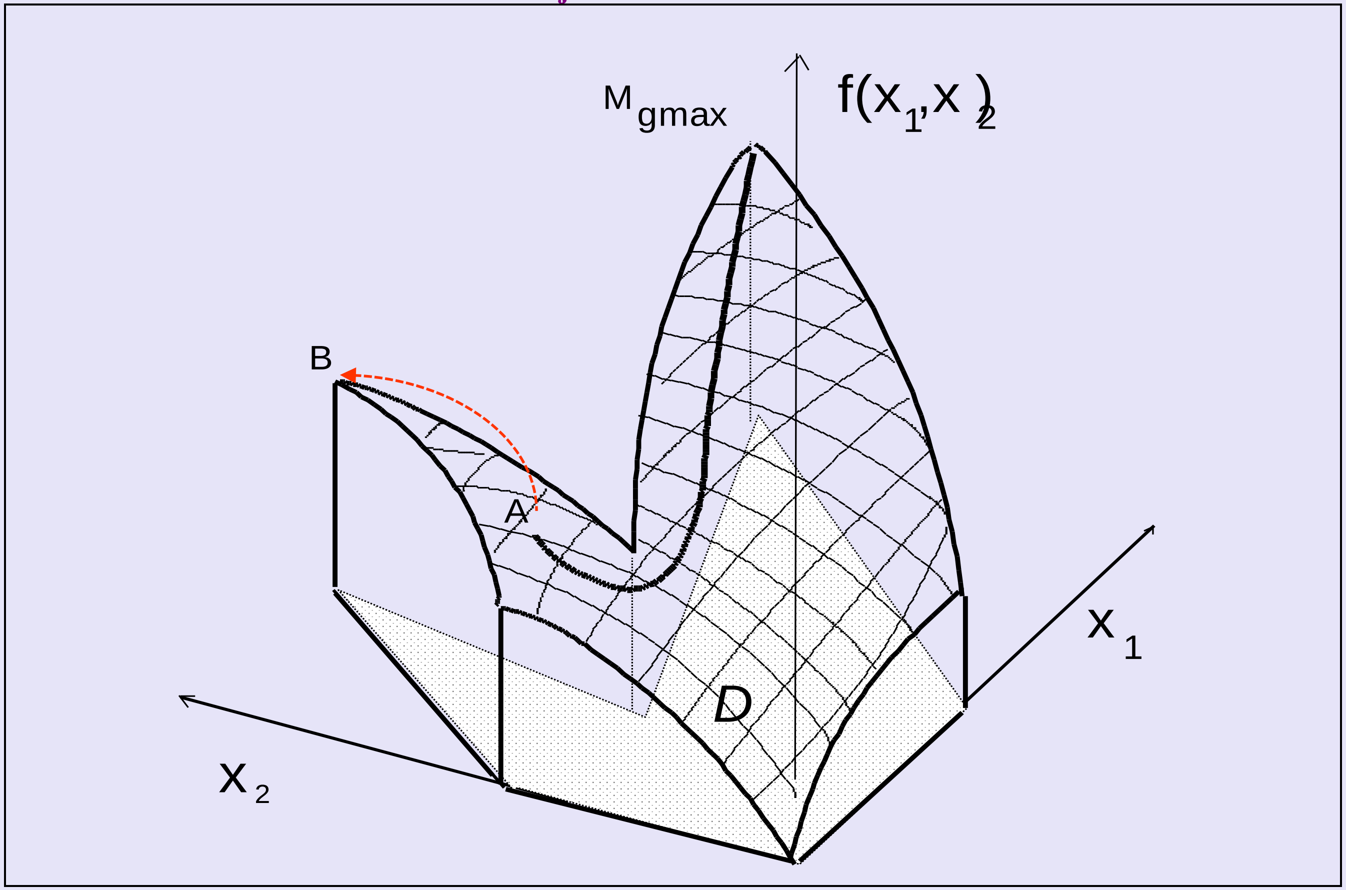
Ťažkosti vznikajúce pri riešení úloh, ktorých účelová funkcia, resp. funkcie sústavy ohraničení nie sú konvexné

Obr.č.3.3: Nekonvexná množina prípustných riešení



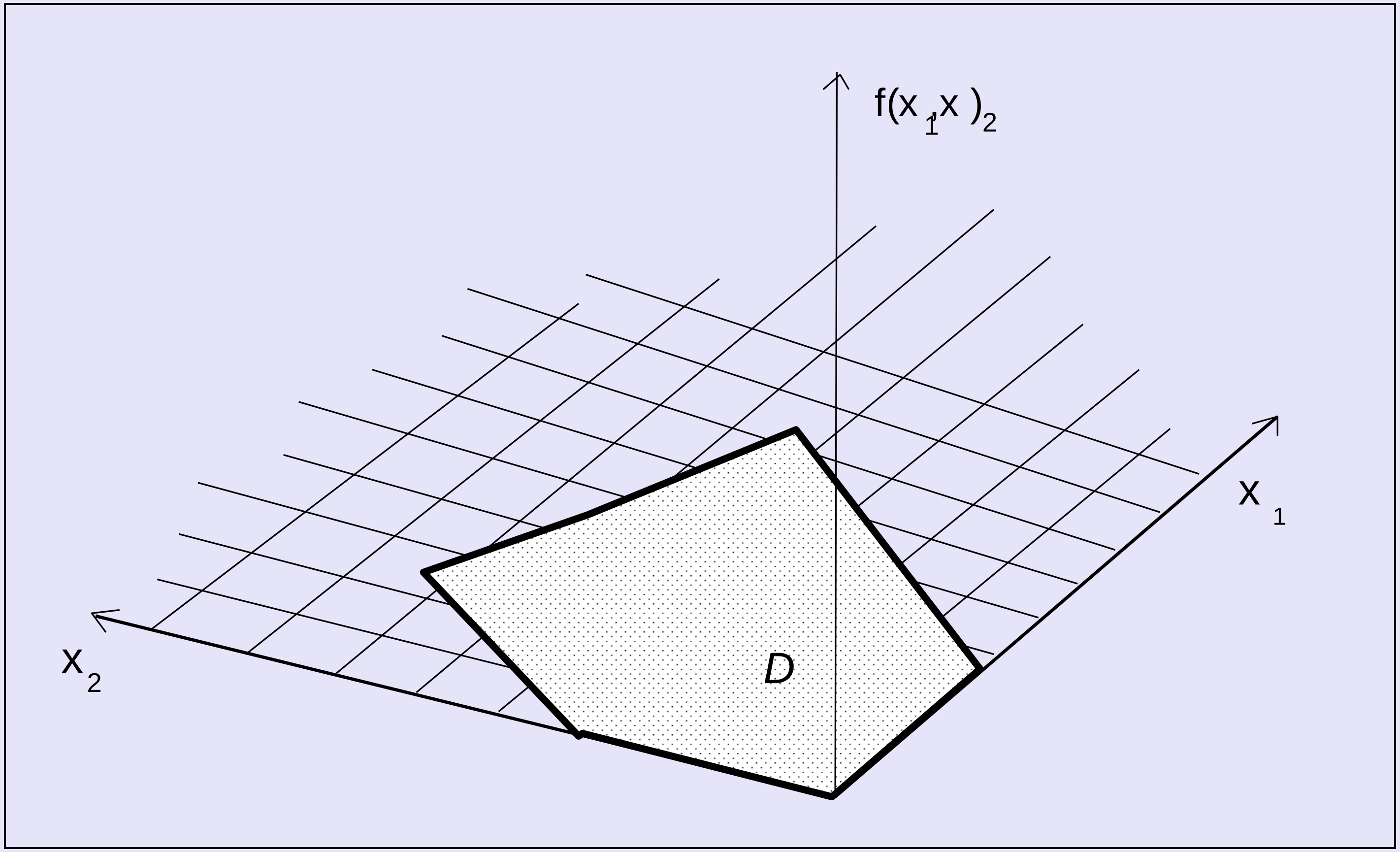
Vlastnosti riešenia úlohy nelineárneho programovania

Obr.č.3.4: Maximalizácia na nekonvexnej množine

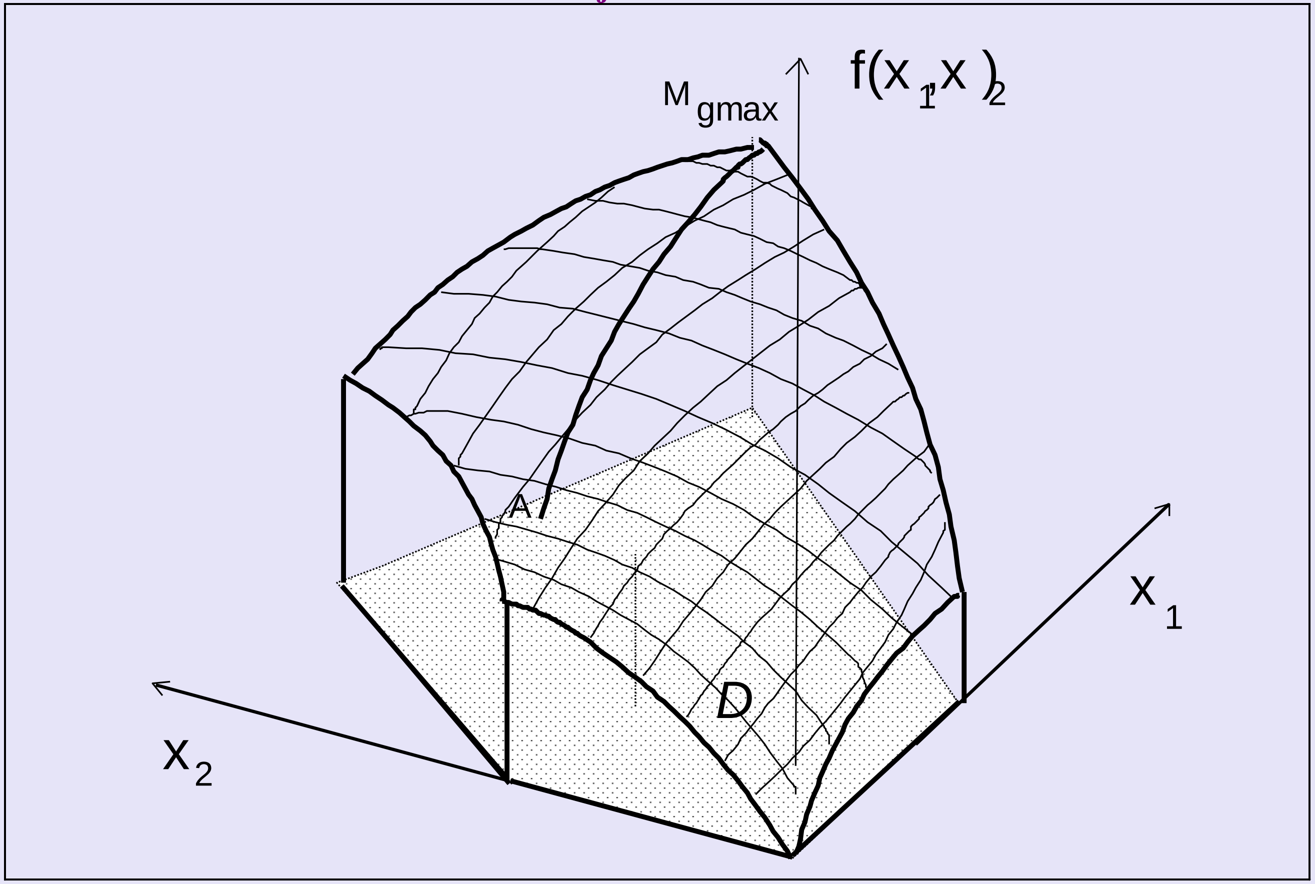


Vlastnosti riešenia úlohy nelineárneho programovania

Obr.č.3.5: Konvexná množina prípustných riešení

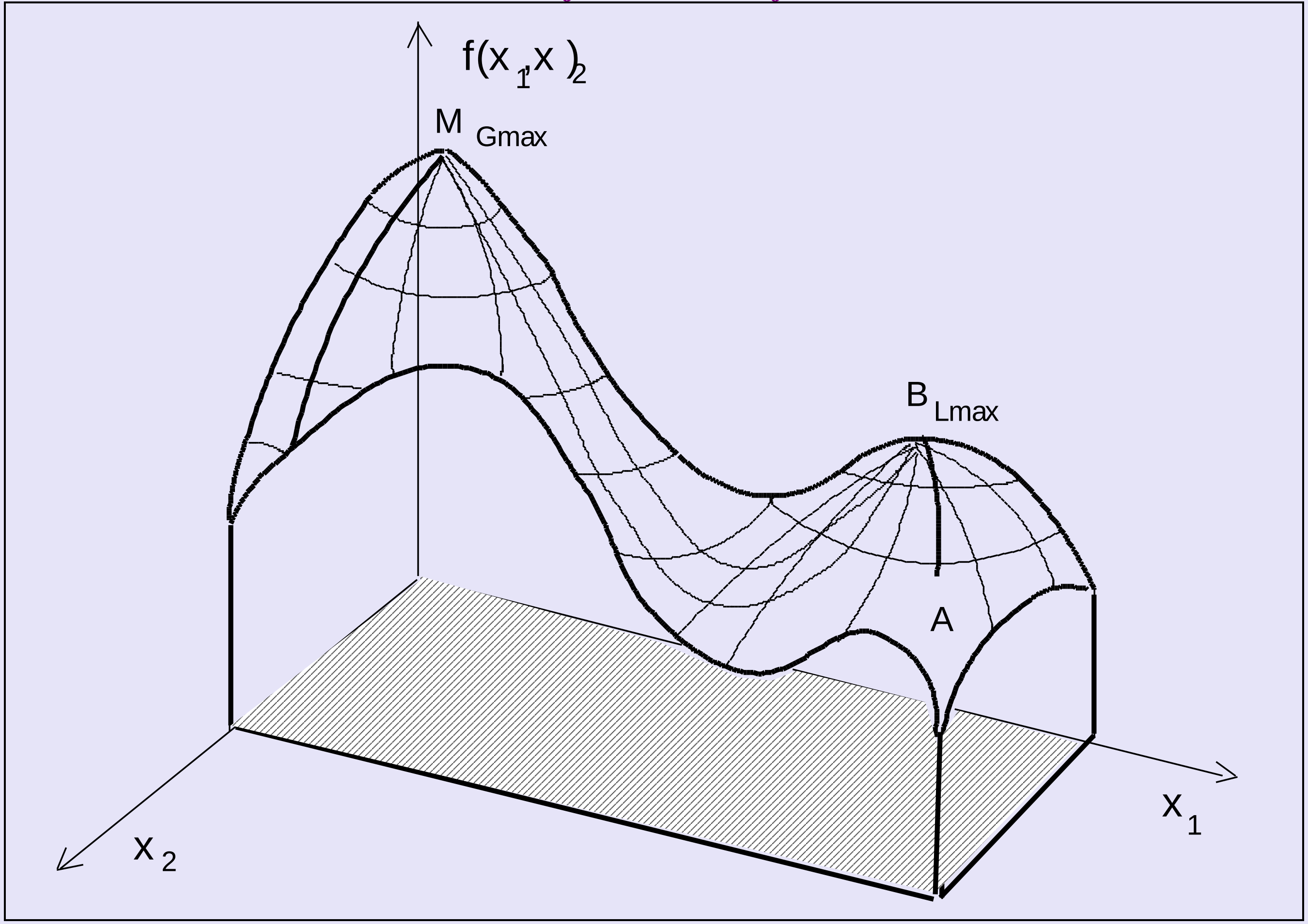


Obr.č.3.6: Maximalizácia na konvexnej množine D

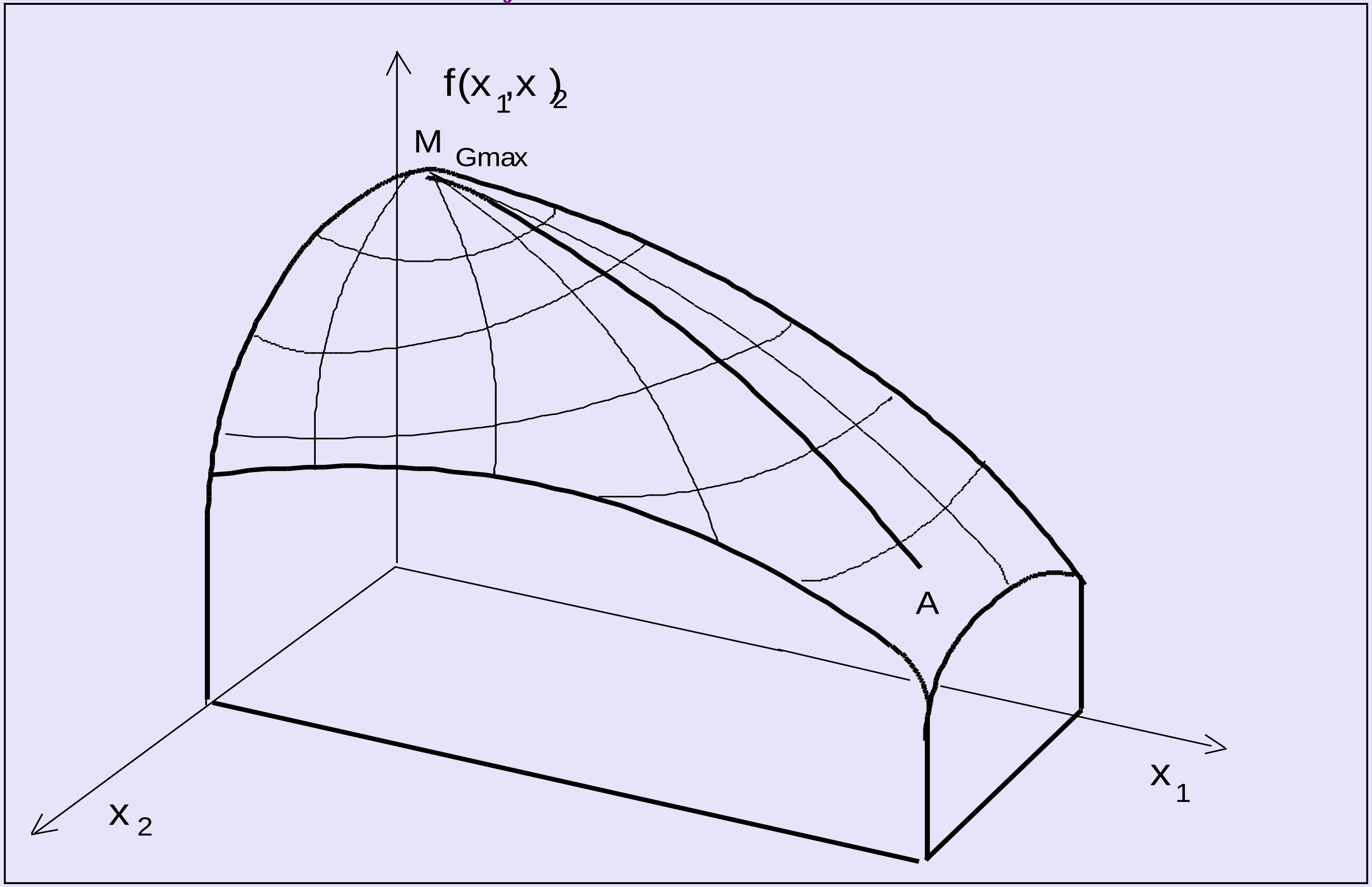


Vlastnosti riešenia úlohy nelineárneho programovania

Obr. č.3.7: Maximalizácia ani konvexnej ani konkávnej funkcie



Obr. č.3.8: Maximalizácia konkávnej fukcie



Vlastnosti riešenia úlohy nelineárneho programovania

Veta 2.2 (Nutná podmienka existencie extrému)

Nech reálna funkcia $f(x_1,...,x_n)$ je diferencovateľná na množine $D \in R^n$. Ak v niektorom vnútornom bode $\mathbf{x^o} \in int\ D$ tejto množiny má funkcia lokálny extrém, t.j. maximum, alebo minimum, potom pre prvé parciálne derivácie funkcie platí

$$\frac{\partial f(x_1^o,\ x_2^o,\ \dots,\ x_n^o)}{\partial x_j} = 0 \qquad pre\ \forall\ j = 1,\dots,n \tag{2.3}$$

resp. *gradient funkcie* v tomto bode je rovný nulovému vektoru

$$\Delta f(x^o) = \left(\frac{\partial f(x^o)}{\partial x_1}, \frac{\partial f(x^o)}{\partial x_2}, \dots, \frac{\partial f(x^o)}{\partial x_n} \right)^T = 0$$

Dôkaz:

Dôkaz tejto vety uvedieme, nakoľko má aj vysvetľujúci charakter. Na základe *Taylorovej vety* pre každé $\theta \in (0,1)$ platí

$$f(\mathbf{x^0} + \mathbf{h}) - f(\mathbf{x^0}) = \nabla f(\mathbf{x^0})\mathbf{h} + 1/2\, \mathbf{h}^T \mathbf{H}(\mathbf{x}^0 + \theta \mathbf{h})\mathbf{h} \tag{2.4}$$

kde

h - vektor $\mathbf{h} \in \mathbf{R^n}$ taký, že jeho zložky platí $|h_j| \to 0$ prte $\forall$j;

H - Hessova matica druhých parciálnych derivácií funkcie $f(\mathbf{x})$ v bode $\mathbf{x^o}+\theta\mathbf{h}$.

Poznámka: Taylorov rozvoj:

$$f(x+h) - f(x) = \frac{1}{1!} f'(x)h + \frac{1}{2!} f''(x)h^2 + \dots + \frac{1}{n!} f^{(n)}(x)h^n + R_n$$

□

Ak , $|h_j| \to 0$ pre $\forall$j, tak člen Taylorovho rozvoja $1/2\,\mathbf{h}^{\mathrm{T}}\mathbf{H}(\mathbf{x}^0 + \theta\mathbf{h})\mathbf{h}$ predstavuje nekonečne malú veličinu, čo umožňuje vzťah (2.4) preformulovať v tvare nasledovnej približnej rovnosti

$$f(\mathbf{x^0} + \mathbf{h}) - f(\mathbf{x^0}) = \nabla f(\mathbf{x^0})\mathbf{h} \tag{2.5}$$

Dôkaz vety vykonáme sporom. Predpokladajme, že bod $\mathbf{x^0}$ je bodom minima funkcie *f(x)* a súčasne gradient funkcie v tomto bode sa nerovná nule, t.j. platí $\nabla f(\mathbf{x^0}) \neq 0$. Inými slovami, $\exists$j$\in\langle 1,n\rangle$, pre ktoré $\partial f(\mathbf{x^0})/\partial x_j < 0$, alebo $\partial f(\mathbf{x^0})/\partial x_j > 0$. Vždy môžeme, samozrejme, vybrať také znamienko h_j, aby platilo

$$h_j \frac{\partial f(\mathbf{x^0})}{\partial x_j} < 0$$

Ak položíme ostatné hodnoty h_k=0, k=1,...,n, k≠j, tak z Taylorovho rozvoja vyplýva platnosť nasledovnej nerovnosti

$$f(\mathbf{x^0} + \mathbf{h}) < f(\mathbf{x^0})$$

Tento záver je však v rozpore s predpokladom, že bod $\mathbf{x^0}$ je bodom minima funkcie *f(x)*, takže gradient funkcie musí byť rovný nulovému vektoru, t.j. $\nabla f(\mathbf{x^0}) = \mathbf{0}$

Vlastnosti riešenia úlohy nelineárneho programovania

- Nutná podmienka existencie extrému, ktorá bola formulovaná vo vete 2.2 totiž umožňuje identifikovať iba stacionárne body funkcie.

- Či je stacionárny bod bodom lokálneho minima funkcie, bodom jej lokálneho maxima, alebo nie je extremálnym bodom funkcie, na to dáva odpoveľ veta o postačujúcich podmienkach existencie extrému.

Ako veľmi jednoduchý a názorný príklad možno uviesť napr. funkciu

$$f(x)=x^3$$

Táto funkcia má v bode $x^o = 0$ prvú deriváciu rovnú nule, takže bod x^o je stacionárny bod funkcie.

Nie je však extremálnym bodom funkcie, nakoľko pre body z jeho ε-okolia platí:

$$f(x)<f(x^o) \quad \text{pre } x \in N_e(x^o),\, x<0$$

$$f(x)>f(x^o) \quad \text{pre } x \in N_e(x^o),\, x>0$$

Veta 3.1 (Postačujúce podmienky existencie extrému)

Nech reálna funkcia *f(x)* je dvakrát diferencovateľná na množine $D \subset R^n$ a bod x^o je stacionárnym bodom funkcie.

Potom, ak Hessova matica H(x) je v stacionárnom bode x^o kladne definitná, tak bod x^o je bodom minima funkcie,

Dôkaz:

Na základe *Talorovej vety* pre každé $\theta \in (0,1)$ platí

$$f(x^o + h) - f(x^o) = \Delta f^T(x^o)h + \frac{1}{2}h^T H(x^o + \theta h)h$$

Keďže bod x^o je stacionárnym bodom funkcie, tak podľa vety 2.2 platí $\nabla f(x^o) = 0$A. Takže platí vzťah:

$$f(x^o + h) - f(x^o) = \frac{1}{2}h^T H(x^o + \theta h)h$$

Nech stacionány bod x^o je bodom minima funkcie. Potom na základe definície minima funkcie ak pre všetky $h \neq 0$ platí:

$$\frac{1}{2}h^T H(x^o + \theta h)h > 0$$

potom

$$f(x^o + h) > f(x^o)$$

q.e.d.

Príklad č.3.3

Je daná funkcia

$$f(x_1, x_2, x_3) = x_1 + 2x_3 + x_2 x_3 - x_1^2 - x_2^2 - x_3^2$$

Vypočítajte extrémy funkcie (ak existujú)!

Riešenie

a) Z nutnej podmienky existencie extrému $\nabla f(x)=0$ vyplýva, že:

$\partial f/\partial x_1 = 1 - 2x_1 = 0$

$\partial f/\partial x_2 = x_3 - 2x_2 = 0$

$\partial f/\partial x_3 = 2 + x_2 - 2x_3 = 0$

Riešením tejto sústavy rovníc získame stacionárny bod funkcie

$$x^o = (1/2,\ 2/3,\ 4/3)$$

Vlastnosti riešenia úlohy nelineárneho programovania

b) Je zrejmé, že typ extrému v stacionárnom bode funkcie je implikovaný znamienkom druhého člena Taylorovho rozvoja v tomto bode. Preskúmajme hodnotu tohto člena

$$\mathbf{h}^{\mathrm{T}}\mathbf{H}(\mathbf{x}^{\mathbf{0}})\mathbf{h} = \mathbf{h}^{\mathrm{T}}\begin{pmatrix} \partial^2 f(\mathbf{x})/\partial x_1^2 & \partial^2 f(\mathbf{x})/\partial x_1 x_2 & \partial^2 f(\mathbf{x})/\partial x_1 x_3 \\ \partial^2 f(\mathbf{x})/\partial x_2 x_1 & \partial^2 f(\mathbf{x})/\partial x_2^2 & \partial^2 f(\mathbf{x})/\partial x_2 x_3 \\ \partial^2 f(\mathbf{x})/\partial x_3 x_1 & \partial^2 f(\mathbf{x})/\partial x_3 x_2 & \partial^2 f(\mathbf{x})/\partial x_3^2 \end{pmatrix}\mathbf{h}$$

$$\mathbf{h}^{\mathrm{T}}\mathbf{H}(\mathbf{x}^{\mathbf{0}})\mathbf{h} = (h_1, h_2, h_3)\begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} =$$

$$= -2h_1^2 - h_2^2 - h_3^2 - (h_2^2 - 2h_2h_3 + h_3^2) =$$

$$= -2h_1^2 - h_2^2 - h_3^2 - (h_2 - h_3)^2 < 0 \qquad \text{ak } \underset{\mathrm{i}}{\exists} h_i \neq 0 \quad i = 1,2,3$$

Vidíme, že pre $\mathbf{h} \neq \mathbf{0}$ má druhý člen Taylorovho rozvoja evidentne záporné znamienko, to znamená, že v okolí stacionárneho bodu $\mathbf{x}^{\mathbf{0}}$ hodnoty funkcie klesajú, a teda stacionárny bod predstavuje bod maxima funkcie.

3.2 Kvadratické formy a ich vlastnosti

Definícia 3.3

Nech C je štvorcová matica *n*-tého stupňa s prvkami c_{ij} pre $i,j=1, \dots, n$. Kvadratickou formou *n* premenných $x_1, x_2, \dots, x_n$ nazývame funkciu

$$Q(\mathbf{x}) = Q(x_1, \dots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_i x_j = \mathbf{x}^{\mathrm{T}}\mathbf{C}\mathbf{x}$$

Maticu C potom nazývame maticou kvadratickej formy.

Príklad č.3.4

Zapíšte kvadratickú formu $Q(x_1,x_2,x_3)$ nasledovnej štvorcovej matice

$$\mathbf{C} = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix}$$

Riešenie:

$$\mathbf{Q}(\mathbf{x}) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 2 & 1 & 3 \\ 4 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$$

$$= 2x_1^2 + x_2^2 + 2x_3^2 + 5x_1x_2 + 5x_1x_3 + 2x_2x_3$$

Vzhľadom na to, že platí pre všetky *i,j* = 1,2,...,*n*

$$x_i x_j = x_j x_i$$

môžeme kvadratickú formu preformulovať do nasledovného "trojuholníkového" tvaru

$$\begin{aligned}\mathbf{x}^{\mathrm{T}}\mathbf{C}\mathbf{x} = c_{11}x_1^2 + (c_{12}+c_{21})x_1x_2 + (c_{13}+c_{31})x_1x_3 + \ldots + (c_{1n}+c_{n1})x_1x_n &+ \\ c_{22}x_1^2 + (c_{23}+c_{32})x_2x_3 + \ldots + (c_{1n}+c_{n1})x_1x_n &+ \\ c_{33}x_3^2 + \ldots + (c_{1n}+c_{n1})x_1x_n &+ \\ \ldots& \\ c_{nn}x_{nn}^2&\end{aligned}$$

Z uvedeného tvaru bezprostredne vyplýva, že matica C kvadratickej formy nie je jednoznačne určená príslušnou kvadratickou formou, to znamená, že rôzne štvorcové matice môžu implikovať tú istú kvadratickú formu.

<u>Príklad č. 3.5</u>

Zapíšte kvadratickú formu, ak je daná nasledovná matica kvadratickej formy

$$\mathbf{C} = \begin{pmatrix} 3 & 6-4y \\ 2+4y & 1 \end{pmatrix}$$

Riešenie:

$$Q(\mathbf{x}) = \begin{pmatrix} x_1 & x_2 \end{pmatrix}\begin{pmatrix} 3 & 6-4y \\ 2+4y & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} =$$

$$= 3x_1^2 + x_2^2 + 8x_1x_2$$

Vidíme, že pre ľubovoľnú hodnotu $y \in R$ generuje matica C rovnakú kvadratickú formu *Q(x)*. Táto nejednoznačnosť v priradení kvadratickej formy k matici kvadratickej formy sa odstráni prijatím predpokladu o nejakej dodatočnej vlastnosti matice kvadratickej formy C, napr. o jej *symetričnosti,* t.j.

$$c_{ij} = c_{ji} \quad \text{pre } i, j = 1,2,\ldots,n$$

Príklad č. 3.6

K nasledovnej kvadratickej forme určte symetrickú maticu kvadratickej formy

$$Q(x_1, x_2, x_3) = 2x_1^2 + 5x_2^2 + x_3^2 + 3x_1x_2 + x_1x_3 - 3x_2x_3 + x_2x_1 + \\ + 7x_3x_1 - 5x_2x_3$$

Riešenie:

Nakoľko $x_i\, x_j = x_j\, x_i$, tak môžeme zlúčiť tieto dvojice členov, pričom súčiny $x_i\, x_j$ zapíšeme tak, aby $i<j$ a lexikograficky ich usporiadame. Dostávame

$$Q(\mathbf{x}) = 2x_1^2 + 4x_1x_2 + 8x_1x_3 + 5x_2^2 - 8x_3x_2 + x_3^2$$

Odtiaľ bezprostredne dostávame "trojuholníkový" tvar kvadratickej formy

$$\begin{aligned} Q(x_1, x_2, x_3) = 2x_1^2 + 4x_1x_2 + 8x_1x_3 &+ \\ 5x_2^2 - 8x_2x_3 &+ \\ + x_3^2 & \end{aligned}$$

a napokon symetrická matica kvadratickej formy má tvar

$$\mathbf{C} = \begin{pmatrix} 2 & 2 & 4 \\ 2 & 5 & -4 \\ 4 & -4 & 1 \end{pmatrix}$$

Definícia 3.4

Kvadratická forma $Q(x)=\mathbf{x}^T\mathbf{C}\mathbf{x}$ sa nazýva:

a) *kladne semidefinitná*, resp. *záporne semidefinitná*, ak pre všetky $\mathbf{x} \in R^n$ *platí*

$$\mathbf{x}^T\mathbf{C}\mathbf{x} \geq 0, \text{ resp. } \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} \leq 0,$$

b) *kladne definitná*, resp. *záporne definitná*, ak pre všetky nenulové $\mathbf{x} \in R^n$ *platí*

$$\mathbf{x}^T\mathbf{C}\mathbf{x} > 0, \text{ resp. } \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} < 0,$$

c) *indefinitná*, ak nie je definitná, t.j. ak $\exists\ \mathbf{x},\mathbf{y} \in \mathrm{R}^n$ také, že platí

$$\mathbf{x}^T\mathbf{C}\mathbf{x} > 0, \ \boldsymbol{y}^T\boldsymbol{C}\boldsymbol{y} < 0.$$

Poznámka 3.3

Zároveň s kvadratickou formou $\mathbf{x}^T\mathbf{C}\mathbf{x}$ nazývame aj príslušnú štvorcovú maticu *kladne*, resp . *záporne (semi)definitnou*.

Vlastnosti riešenia úlohy nelineárneho programovania

Praktické skúmanie definitnosti kvadratickej formy predstavuje samostatný výpočtový problém, s ktorým sa tu nebudeme podrobnejšie zaoberať. Napokon pre matice menších rozmerov možno definitnosť určiť aj priamo na základe analytického tvaru kvadratickej formy, ako to bolo ukázané v príklade č.3.3. Pre určenie typu definitnosti matíc väčších rozmerov možno použiť nasledovné algoritmy

a) skúmanie znamienok hlavných subdeterminantov matice C

b) *Lagrangeova* metóda transformácie kvadrarickej formy k hlavným osiam

c) skúmanie vlastných čísel matice C

Vysvetlíme si krátko prvý z týchto prístupov. Preveríme, ži matica **C** je kladne definitná, alebo kladne semidefinitná. Skúmajme štvorcovú symetrickú maticu n-tého stupňa

$$C = [c_{ij}] = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ . & . & \cdots & . \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

Definícia 3.5

Matica **C** je kladne definitná, ak sú všetky hlavné subdeterminanty matice kladné.

Poznámka 3.5

Hlavným subdeterminantom matice **C** je determinant matice tvaru

$$\det C = \begin{vmatrix} c_{i_1,i_1} & c_{i_1,i_2} & \cdots & c_{i_1,i_k} \\ c_{i_2,i_1} & c_{i_2 i_2} & \cdots & c_{i_2,i_k} \\ . & . & \cdots & . \\ c_{i_k,i_1} & c_{i_k,i_2} & \cdots & c_{i_k,i_k} \end{vmatrix}$$

kde $1 \leq i_1 < i_2 < \ldots < i_k \leq n$.

Tento subdeterminant vypočítame tak, že v matici **C** vyškrtneme všetky riadky a stĺpce okrem riadkov a stĺpcov $i_1, i_2, \ldots, i_k$ a z týchto nevyškrtnutých riadkov a stĺpcov zostavíme determinant, pričom nemeníme vzájomné postavenie prvkov matice **C**.

Definícia 3.6

Matica C je kladne semidefinitná, ak sú všetky hlavné subdeterminanty matice nezáporné.

Dôsledok 3.1

Elementy na hlavnej diagonále kladne definitnej, resp. kladne semidefinitnej matice sú kladné, resp. nezáporné.

Dôsledok 3.2

Elementy na hlavnej diagonále záporne definitnej, resp. záporne semidefinitnej matice sú záporné, resp. nekladné.

Dôsledok 3.3

Ak na hlavnej diagonále matice sa vyskytujú kladné aj záporné čísla, matica je indefinitná.

(Dôkazy uvedených dôsledkov vyplývajú priamo z poznatku, že prvky na hlavnej diagonále matice sú aj jej subdeterminanty *prvého* stupňa.)

Preskúmanie všetkých hlavných subdeterminantov matice C je však z výpočtového hľadiska neúnosné.

Vieme však, že transformovaním matice na maticu, ktorej riadky sú lineárne kombinácie riadkov pôvodnej matice, sa hlavné subdeterminanty nemenia.

Maticu C preto upravíme na *hornotrojuholníkovú maticu*, pre ktorú ľahko preskúmame znamienka hlavných subdeterminantov.

Ak sú znamienka všetkých prvkov na hlavnej diagonále nezáporné, matica je kladne semidefinitná. Podobne, ak sú znamienka všetkých prvkov na hlavnej diagonále kladné, matica je kladne definitná.

- **Charakter stacionárneho bodu je implikovaný typom definitnosti *Hessovej* matice, a teda znamienkom druhého člena *Taylorovho* rozvoja funkcie.**

Nech x^o je stacionárnym bodom funkcie *f(x)* a nech sú ďalej všetky druhé parciálne derivácie funkcie v okolí bodu x^o spojité. Potom môžeme formulovať nasledovné kritérium pre určenie charakteru stacionárneho bodu:

a) ak matica $H(x^o)$ je kladne definitná, potom x^o je bodom *lokálneho minima*;

b) ak matica $H(x^o)$ je záporne definitná, potom x^o je bodom *lokálneho maxima*;

c) ak matica $H(x^o)$ je indefinitná, potom x^o nie je bodom *lokálneho extrému*;

d) ak matica $H(x^o)$ je semidefinitná, potom toto kritérium *neumožňuje* identifikovať charakter stacionárneho bodu x^o.

Príklad č.3.8

Vyšetrite extrémy nasledovných funkcií

a) $f(x_1, x_2) = 1/2x_1^2 + 5x_2^2 + 2x_1x_2 - 6x_2$

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

c) $f(x_1, x_2) = x_1^2 - 1/2x_2^2 + x_1x_2 - 3x_1$

d) $f(x_1, x_2) = x_1^2 + x_2^4$

e) $f(x_1, x_2) = x_1^2 - x_2^4$

Zovšeobecnenie pojmu konvexnej funkcie

Konvexná funkcia

Nech $f(x) : D \rightarrow R$, kde D Ě R^n je neprázdna a konvexná množina. Hovoríme, že funkcia $f(x)$ je konvexná na množine D, ak pre ľubovoľné dva body tejto množiny $x^1, x^2 \in D$ a ľubovoľné $\lambda \in <0,1>$ platí,

$$\boxed{f[\lambda x^1 + (1 - \lambda)x^2] \leq \lambda f(x^1) + (1 - \lambda)f(x^2)}$$

Inými slovami, úsečka spájajúca ľubovoľné dva body krivky grafu funkcie leží na krivke, alebo nad krivkou grafu funkcie.

Rýdzokonvexná funkcia

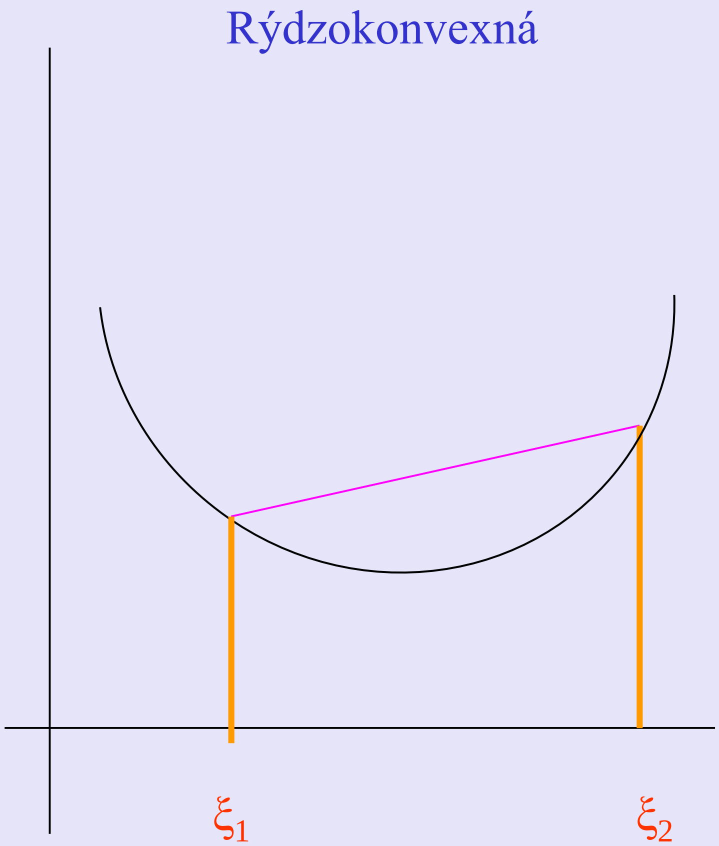
Hovoríme, že funkcia $f(x) : D \rightarrow R$ je rýdzokonvexná na množine D, ak pre ľubovoľné dva body tejto množiny $x^1, x^2 \in D$ a ľubovoľné $\lambda \in (0,1)$ platí

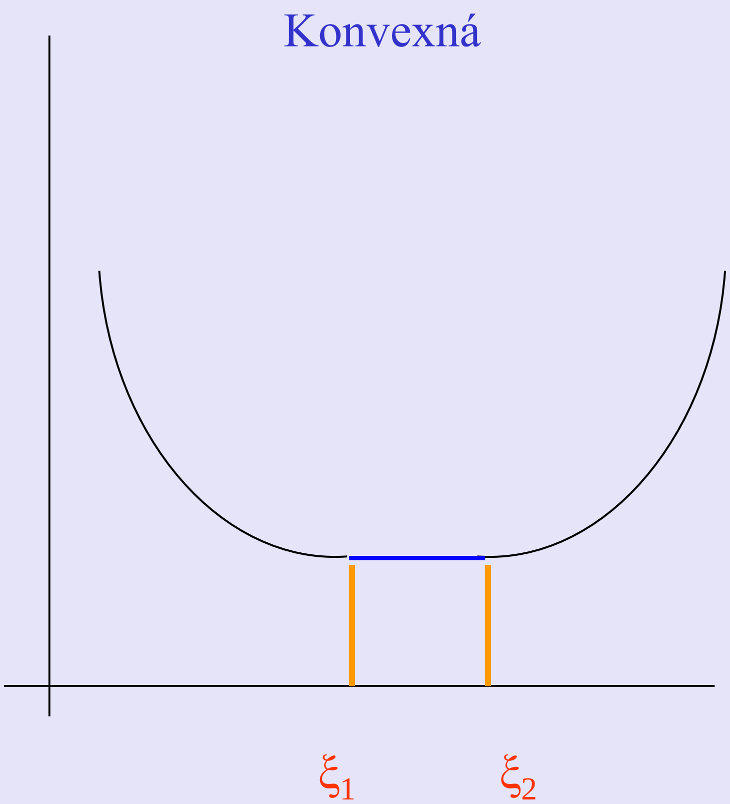
$$\boxed{f[\lambda x^1 + (1-\lambda)x^2] < \lambda f(x^1) + (1-\lambda)f(x^2)}$$

Poznámka

Hovoríme, že funkcia $f(x) : D \rightarrow R$ je konkávna, resp. rýdzokonkávna na množine D, ak funkcia -f(x) je konvexná, resp. rýdzokonvexná.

Vlastnosti riešenia úlohy nelineárneho programovania





Zovšeobecnenie pojmu konvexnej funkcie

a) Kvázikonvexné funkcie

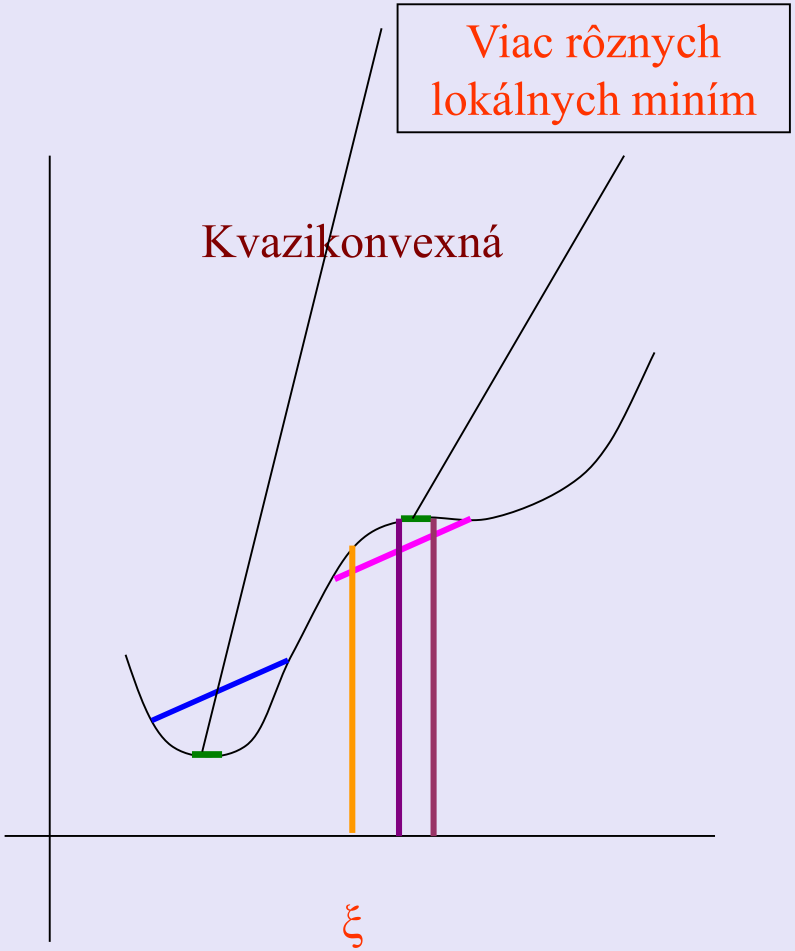
Definícia

Nech f($\mathbf{x}$) : $D \to R$, kde D je neprázdna a konvexná množina v R^n. Hovoríme, že funkcia $f(\mathbf{x})$ je kvázikonvexná na množine D, ak pre ľubovoľné dva body tejto množiny $\mathbf{x}^1, \mathbf{x}^2 \in D$ a ľubovoľné $\lambda \in (0, 1)$ platí,

$$f(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2) \leq \max \{f(\mathbf{x}^1), f(\mathbf{x}^2)\}$$

Poznámka:
Z uvedenej definície *kvázikonvexnej funkcie* je zrejmé, že každá konvexná funkcia je zároveň aj kvázikonvexnou.

Vlastnosti riešenia úlohy nelineárneho programovania



$$f(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2) \leq \max \{f(\mathbf{x}^1), f(\mathbf{x}^2)\}$$

b) Rýdzo kvázikonvexné funkcie

Rýdzo kvázikonvexné a rýdzo kvázikonkávne funkcie majú zvlášť v nelineárnom programovaní značný význam, nakoľko práve pre tieto funkcie má *lokálne minimum* a *lokálne maximum* na konvexnej množine zároveň charakter *globálneho maxima* a *globálneho minima*.

Definícia 3.12

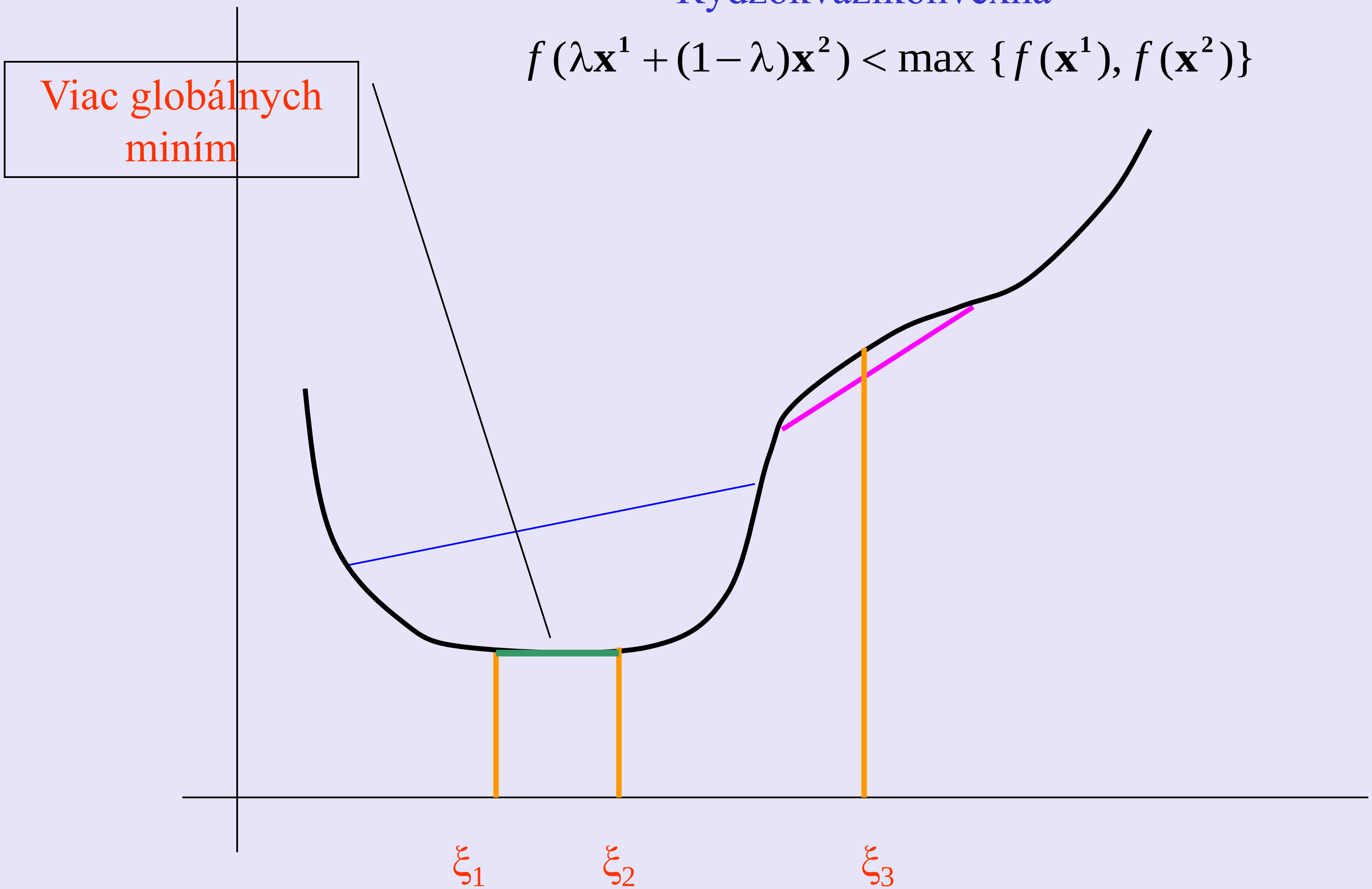
Nech $f(x) : D \to R$, kde D je neprázdna a konvexná množina v R^n. Hovoríme, že funkcia $f(x)$ je rýdzo kvázikonvexná na množine D, ak pre ľubovoľné dva body tejto množiny $x^1, x^2 \in D$ také, že $f(x^1) \neq f(x^2)$ pre všetky $\lambda \in (0, 1)$ platí,

$$f(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2) < \max \{f(\mathbf{x}^1), f(\mathbf{x}^2)\}$$

Funkcia $f(x)$ sa nazýva *rýdzo kvázikonkávnou*, ak funkcia $-f(x)$ je *rýdzo kvázikonvexnou*.

Rýdzokvazikonvexná

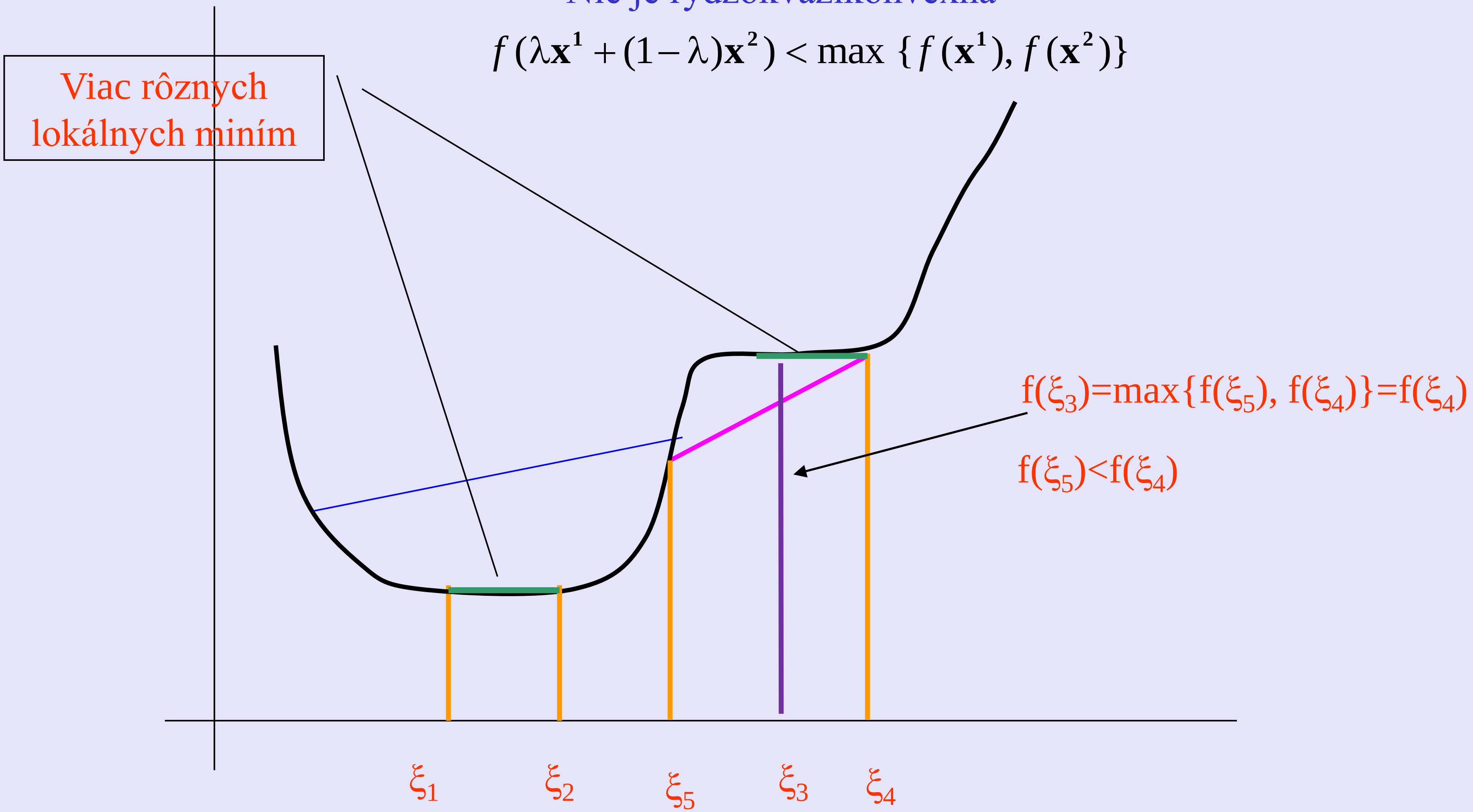
$$f(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2) < \max\{f(\mathbf{x}^1), f(\mathbf{x}^2)\}$$



Vlastnosti riešenia úlohy nelineárneho programovania

Nie je rýdzokvazikonvexná

$$f(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2) < \max\{f(\mathbf{x}^1), f(\mathbf{x}^2)\}$$



c) Silno kvázikonvexné funkcie

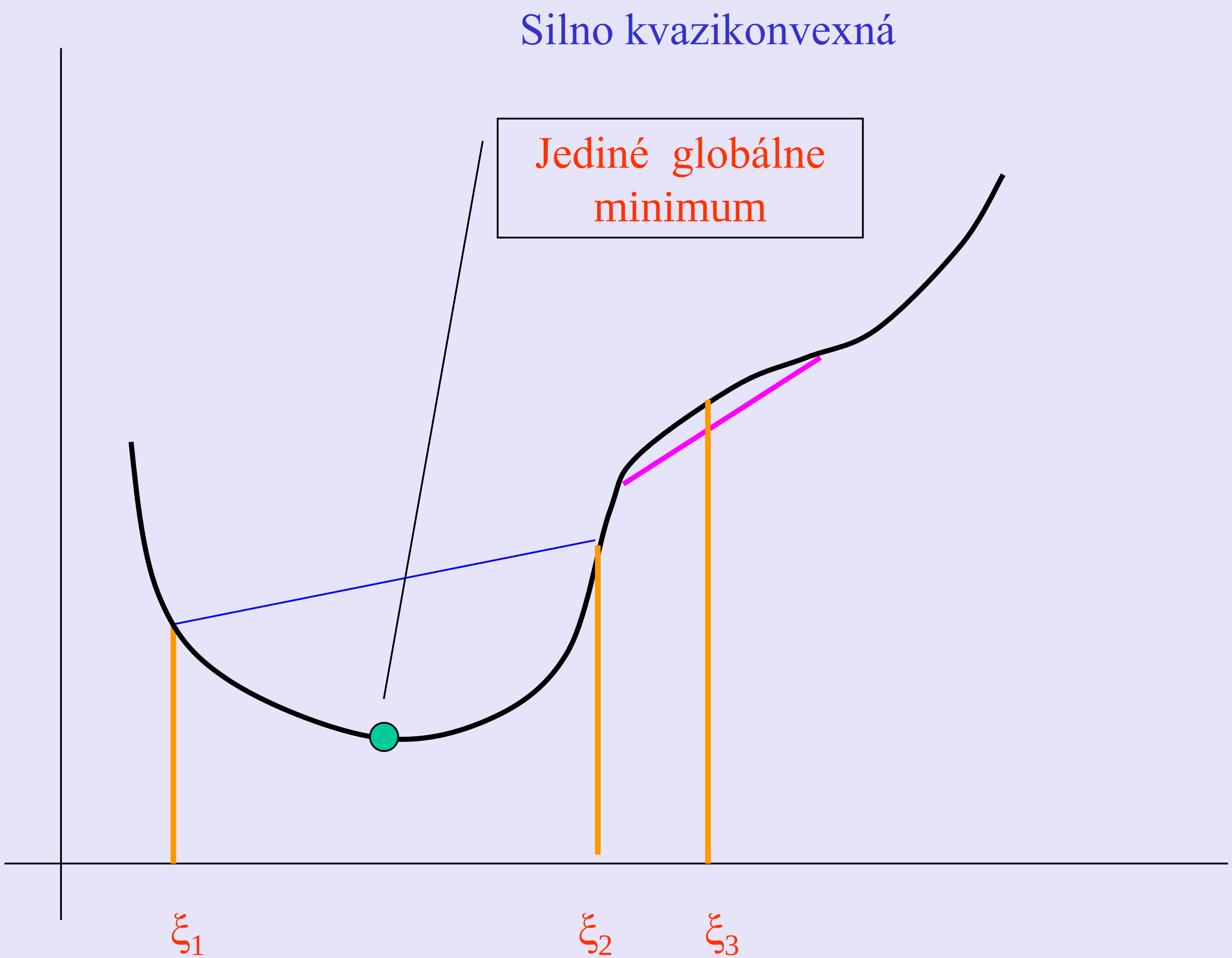
Ukázali sme, že bod lokálneho minima rýdzo kvázikonvexnej funkcie na konvexnej množine D je zároveň globálnym optimálnym riešením príslušnej úlohy. Avšak rýdza kvázikonvexnosť funkcie nezaručuje, aby globálne riešenie bolo jediným globálnym optimálnym riešením, ďalej budeme definovať iný variant kvázikonvexnosti funkcie, tzv. ***silnú kvázikonvexnosť***, ktorá bude garantovať existenciu jediného globálneho minima funkcie na konvexnej množine.

<u>Definícia 3.13</u>

Nech $f(\mathbf{x}) : D \to R$, kde D je neprázdna a konvexná množina v R^n. Hovoríme, že funkcia $f(\mathbf{x})$ je silno kvázikonvexná na množine D, ak pre ľubovoľné dva body tejto množiny $\mathbf{x}^1, \mathbf{x}^2 \in D$ také, že $\mathbf{x}^1 \neq \mathbf{x}^2$ pre všetky $\lambda \in (0,1)$ platí,

$$\boxed{f(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2) < \max\{f(\mathbf{x}^1), f(\mathbf{x}^2)\}}$$

Funkcia $f(\mathbf{x})$ sa nazýva *silno kvázikonkávnou*, ak funkcia $-f(\mathbf{x})$ je *silno kvázikonvexnou*.


Silno kvazikonvexná
Jediné globálne minimum
ξ_1
ξ_2
ξ_3

Pseudokonvexné funkcie

Definícia

Nech *f(x) : D* $\to$ *R* je diferencovateľná na *D*, kde *D* je neprázdna a otvorená množina v R^n. Hovoríme, že funkcia *f(x)* je pseudokonvexná, ak pre ľubovoľné dva body x^1, $x^2 \in D$, pre ktoré platí

$$\nabla f(\mathbf{x}^1)^T(\mathbf{x}^2 - \mathbf{x}^1) \geq 0$$

platí vzťah

$$f(x^2) \geq f(x^1),$$

alebo, čo je ekvivalentné, platí implikácia

$$f(\mathbf{x}^2) < f(\mathbf{x}^1) \quad \Rightarrow \quad \nabla f(\mathbf{x}^1)^T(\mathbf{x}^2 - \mathbf{x}^1) < 0$$

Hovoríme, že funkcia *f(x)* je *pseudokonkávnou*, ak funkcia *-f(x)* je *pseudokonvexnou*.

Vlastnosti riešenia úlohy nelineárneho programovania