NI-MPI přednáška 3

Analýza III: vázané extrémy

13. února 2025

FIT ČVUT

Věta 7.1 (Věta o implicitní funkci)

Nechť $f\colon \mathbb{R}^2 \to \mathbb{R}$ *a* $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ *jsou spojité v okolí bodu* (a,b). *Předpokládejme dále, že*

$$f(a,b) = 0 \qquad \text{a} \qquad \frac{\partial f}{\partial y}(a,b) \neq 0.$$

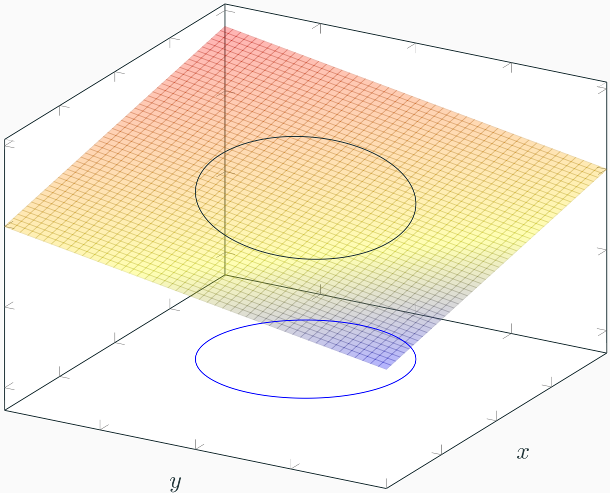
(a) *Pak existují* $\varepsilon > 0$, $\delta > 0$ *a obdélníkové okolí* $R := \{(x,y) : |x-a| < \varepsilon, |y-b| < \delta\}$ *bodu* (a,b) *takové, že pro každé* $x \in (a-\varepsilon, a+\varepsilon)$ *existuje právě jedno* $y \in (b-\delta, b+\delta)$, *které splňuje rovnici* $f(x,y) = 0$. *Tedy* y *je funkcí proměnné* x *a lze psát* $y = \varphi(x)$. *Definiční obor* φ *je interval* $(a-\varepsilon, a+\varepsilon)$.

(b) *Funkce* φ *určená v bodě (a) a její derivace* φ' *jsou spojité na intervalu* $(a-\varepsilon, a+\varepsilon)$ *a platí*

$$\varphi'(x) = -\frac{\frac{\partial f}{\partial x}(x,\varphi(x))}{\frac{\partial f}{\partial y}(x,\varphi(x))} \quad \text{pro } x \in (a-\varepsilon, a+\varepsilon).$$

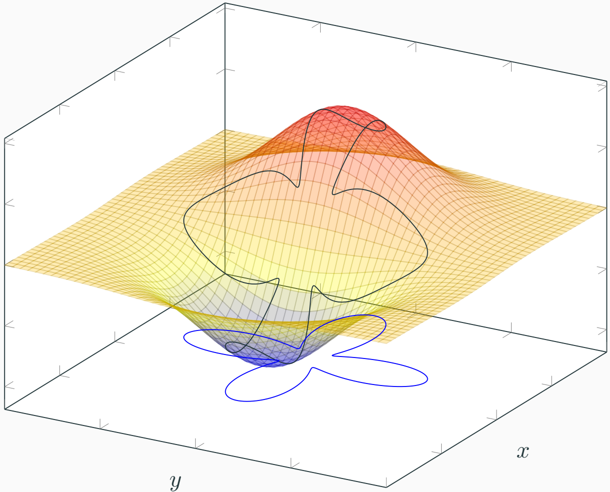
Ukázka - Obrázek 1

Nalezněte maximum a minimum, když se „pohybujeme“ na grafu pouze po černé křivce (= nad modrou křivkou):



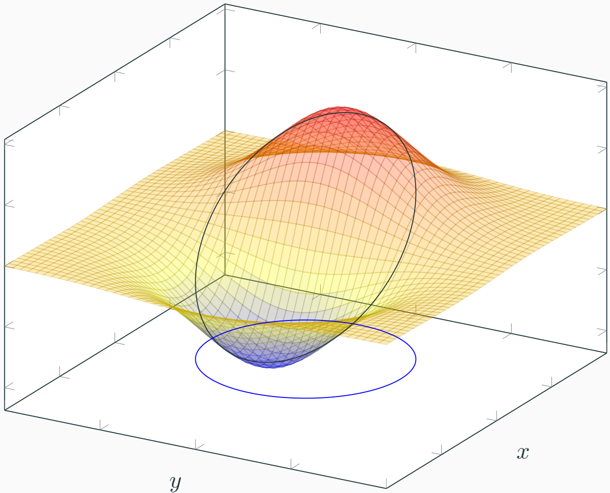
Ukázka - Obrázek 2

Nalezněte maximum a minimum, když se „pohybujeme“ na grafu pouze po černé křivce (= nad modrou křivkou):



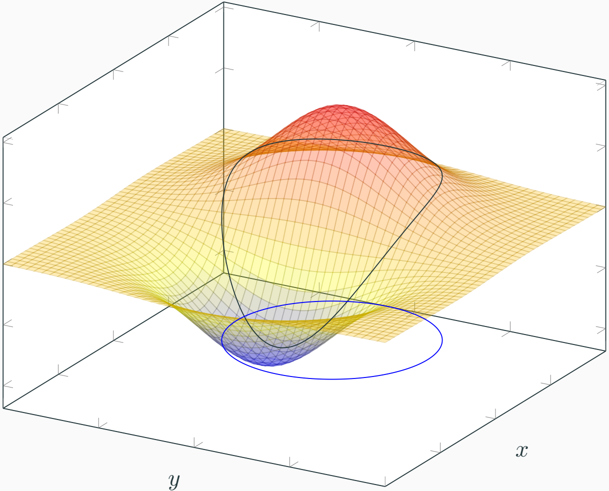
Ukázka - Obrázek 3

Nalezněte maximum a minimum, když se „pohybujeme“ na grafu pouze po černé křivce (= nad modrou křivkou):



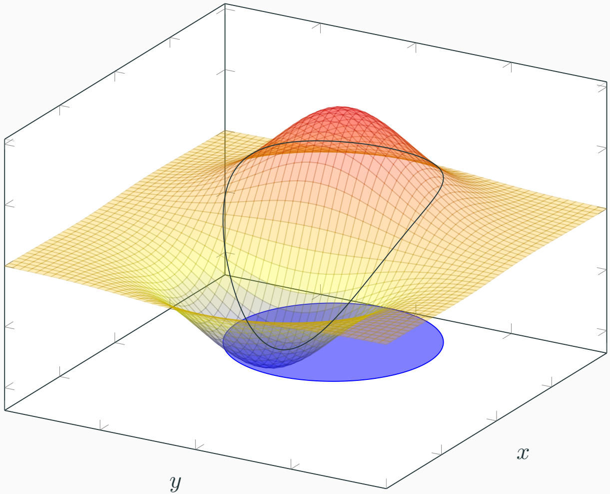
Ukázka - Obrázek 4

Nalezněte maximum a minimum, když se „pohybujeme“ na grafu pouze po černé křivce (= nad modrou křivkou):



Ukázka - Obrázek 5

Nalezněte maximum a minimum, když se „pohybujeme“ na grafu pouze **nad modrou množinou**:



Formulace problému

Označme $\hat{m} = \{1, \dots, m\}$ a $\hat{p} = \{1, \dots, p\}$ $(m, p \in \mathbb{N})$.

Úloha vázaného extrému (minima) (*constrained optimization/minimization*) je obecně následující úloha:

$$\begin{cases} \text{minimalizuj} & f(\mathbf{x}), \\ \text{za podmínek} & g_j(\mathbf{x}) = 0, \ j \in \hat{m}, \\ & h_k(\mathbf{x}) \leq 0, \ k \in \hat{p}, \end{cases} \tag{1}$$

kde f, g_j, h_k jsou funkce $D \to \mathbb{R}$, kde $D \subset \mathbb{R}^n$.

Terminologie

f: objektivní/účelová/minimalizovaná/optimalizovaná funkce

g_j: rovnostní podmínka/vazba | podmínky/vazby

h_k: nerovnostní podmínka/vazba | podmínky/vazby

Jsou-li všechny funkce lineární (a je-li alespoň jedna nerovnostní podmínka): **úloha lineárního programování**

Jsou-li všechny vazby lineární (a je-li alespoň jedna nerovnostní podmínka) a f kvadratická: **úloha kvadratického programování**

Jinak obecně: **úloha nelineárního programování**

Přípustná řešení a jiná formulace

Označme **množinu přípustných řešení:**

$$\mathcal{M} = \{\mathbf{x} \in D \colon (\forall j \in \hat{m})(g_j(\mathbf{x}) = 0) \ \wedge \ (\forall k \in \hat{p})(h_k(\mathbf{x}) \leq 0)\}$$

Definice 7.2

Mějme množiny $D \subset \mathbb{R}^n$ a $\mathcal{M} \subset D$. Řekneme, že funkce $f : D \to \mathbb{R}$ má v bodě $\mathbf{x}^* \in \mathcal{M}$ lokální minimum **vzhledem k množině** $\mathcal{M}$, pokud existuje okolí $H(\mathbf{x}^*)$ takové, že platí

$$\forall \mathbf{x} \in (H(\mathbf{x}^*) \cap \mathcal{M}) \quad f(\mathbf{x}^*) \leq f(\mathbf{x}).$$

Bod $\mathbf{x}^*$ se zove bodem lokálního minima funkce f vzhledem k množině $\mathcal{M}$.

Analogicky se definuje maximum a ostré extrémy vzhledem k množině, a body těchto extrémů.

Obecně

Obecná metoda závisí na funkcích f, g_j a h_k:

- substituce,
- lineární programování,
- kvadratické programování,
- konvexní optimalizace,
- ...

Máme-li k dispozici (druhé) parciální derivace všech funkcí, lze je zkusit (pro nelineární problémy) použít. Tyto analytické metody se opět opírají o geometrický význam první a druhé parciální derivace.

Metoda řešení při rovnostních vazbách

Lagrangeova funkce

Mějme úlohu nalézt lokální extrémy funkce $f : D \to \mathbb{R}$, kde $D \subset \mathbb{R}^n$, vzhledem k množině

$$\mathcal{M} = \{\mathbf{x} \in D\colon g_i(\mathbf{x}) = 0, i = 1, 2, \ldots, m\}, \text{kde } g_i : \mathbb{R}^n \to \mathbb{R}.$$

Definice 7.3

Funkci $L : \mathcal{M} \times \mathbb{R}^m \to \mathbb{R}$ definovanou jako

$$L(\mathbf{x}; \lambda) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j g_j(\mathbf{x})$$

nazýváme **Lagrangeovou funkcí** pro danou úlohu.

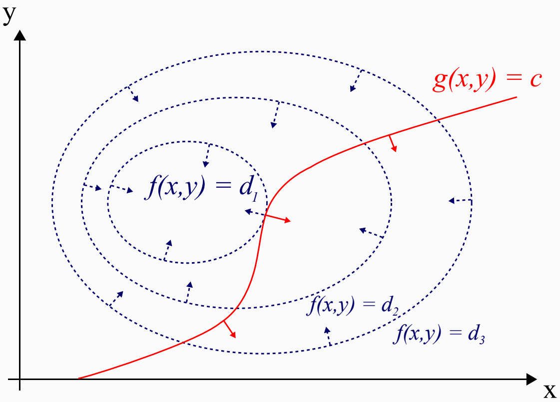
Koeficienty

$$\lambda = (\lambda_1, \ldots, \lambda_m)$$

nazýváme **Lagrangeovy multiplikátory** (*Lagrange multipliers*).

K čemu Lagrangeova funkce?

Uvažujme jednu rovnostní vazbu (tedy $m = 1$). Pokud v bodě x^* platí, že gradient vazby i naší funkce mají stejný směr, tedy $\nabla f(x^*) = -\lambda_1^* \nabla g_1(x^*)$, pak „vazba v daném bodě neprotíná vrstevnici".



Zdroj: https://en.wikipedia.org/wiki/File:LagrangeMultipliers2D.svg

Věta 7.4 (Postačující podmínka existence ostrého lokálního minima pro rovnostní vazby)

Nechť $f, g_j, j \in \{1, \dots, m\}$ *mají spojité všechny druhé parciální derivace na nějaké otevřené nadmnožině* $\tilde{\mathcal{M}} \supset \mathcal{M}$. *Pokud dvojice* $(x^*; \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$ *splňuje podmínky:*

(0) *(0. derivace)* $x^* \in \mathcal{M}$;

(1) *(1. derivace)* $\forall i, \frac{\partial L}{\partial x_i}(x^*; \lambda^*) = 0$;

(2) *(2. derivace) pro každý (sloupcový) vektor* $0 \neq v \in \mathbb{R}^n$ *splňující*

$$\nabla g_j(x^*) \cdot v = 0, \quad \text{pro } \forall j \in \{1, \dots, m\},$$

platí

$$v^T \cdot \nabla_x^2 L(x^*; \lambda^*) \cdot v > 0;$$

kde $\nabla_x^2 L$ *je Hessova matice funkce* L *vzhledem k proměnným* $x = (x_1, x_2, \dots, x_n)$,

potom je x^* *bodem ostrého lokálního minima.*

Všimněme si, že body (0) a (1) jsou ekvivalentní rovnosti $\nabla L(x^*; \lambda^*) = 0$.

Metoda řešení při rovnostních vazbách i nerovnostních vazbách

Lagrangeova funkce

Mějme úlohu nalézt lokální extrémy funkce $f : D \to \mathbb{R}$, kde $D \subset \mathbb{R}^n$, vzhledem k množině

$$\mathcal{M} = \{\mathbf{x} \in D \colon g_i(\mathbf{x}) = 0, h_j(\mathbf{x}) \leq 0, i = 1, 2, \ldots, m, j = 1, \ldots, p\}, \text{kde } g_i, h_j : \mathbb{R}^n \to \mathbb{R}.$$

Definice 7.5 (Lagrangeova funkce)

Funkci $L : \mathcal{M} \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ definovanou

$$L(x; \lambda; \mu) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) + \sum_{k=1}^{p} \mu_k h_k(x)$$

nazýváme **Lagrangeovou funkcí** pro danou úlohu (1). Koeficienty

$$\lambda = (\lambda_1, \ldots, \lambda_m) \quad \text{a} \quad \mu = (\mu_1, \ldots, \mu_p)$$

nazýváme **Lagrangeovy multiplikátory** (*Lagrange multipliers*).

Technické předpoklady diferencovatelnosti

Nechť $\widetilde{\mathcal{M}}$ je otevřená **nad**množina $\mathcal{M}$, na níž pro všechny funkce f, g_j pro $j \in \hat{m}$, h_k pro $k \in \hat{p}$ existují druhé parciální derivace.

Funkci L pak chápeme jako funkci $L : \widetilde{\mathcal{M}} \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$.

Pro jednoduchost lze uvažovat $\mathcal{M} \subset \widetilde{\mathcal{M}} = \mathbb{R}^n$.

(Množina $\mathcal{M}$ je uzavřená (díky spojitosti g_j a h_k).)

Aktivní omezení

Pokud máme nerovnostní vazby h_k, je třeba vědět, kdy je splněna rovnost:

Definice 7.6 (Množina aktivních omezení)

Pro bod $x \in \mathcal{M}$ definujeme **množinu aktivních omezení**

$$\mathcal{B}(x) = \{k \in \hat{p} \colon h_k(x) = 0\}\,.$$

Tedy $\mathcal{B}(x)$ jsou indexy nerovnostních vazeb takových, že x je na hranici množiny $\{x \colon h_k(x) \leq 0\}$.

Věta 7.7 (Postačující podmínka existence ostrého lokálního minima)

Nechť f, g_j, h_k *pro* $j \in \hat{m}, k \in \hat{p}$ *mají spojité všechny druhé parciální derivace na nějaké otevřené nadmnožině* $\tilde{\mathcal{M}} \supset \mathcal{M}$. *Pokud trojice* $(x^*; \lambda^*; \mu^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ *splňuje podmínky:*

1. *(0. derivace)* $x^* \in \mathcal{M}$;
2. *(1. derivace)* $\forall i, \frac{\partial L}{\partial x_i}(x^*; \lambda^*; \mu^*) = 0$;
3. *(aktivní a neaktivní vazby) Pro každé* $k \in \hat{p}$, $\mu_k^* = 0$ *nebo* $h_k(x^*) = 0$;
4. *(2. derivace) pro každý (sloupcový) vektor* $0 \neq v \in \mathbb{R}^n$ *splňující*

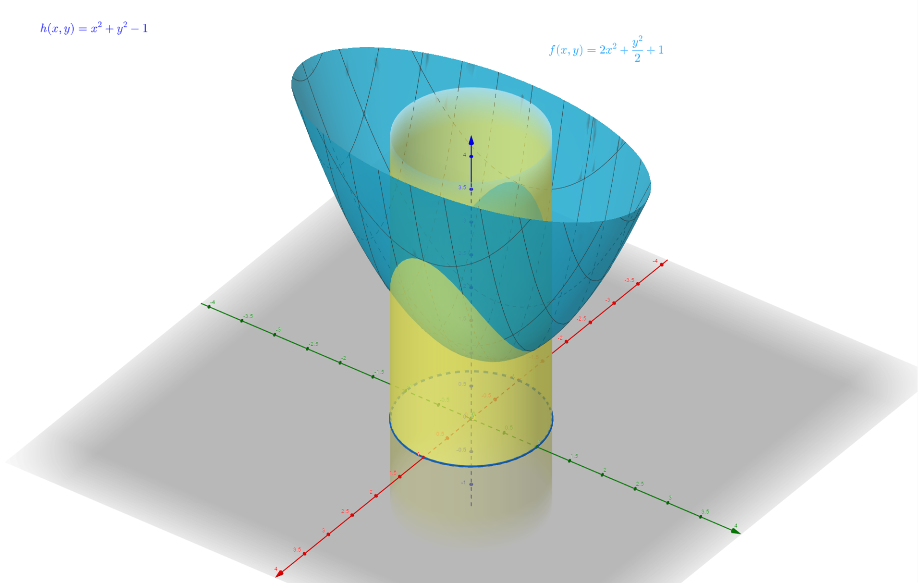
   $$\begin{aligned} \nabla g_j(x^*) \cdot v &= 0, \quad \text{pro všechna } j \in \hat{m}, \\ \nabla h_k(x^*) \cdot v &= 0, \quad \text{pro všechna } k \in \hat{p}, \mu_k^* \neq 0, \text{ platí} \end{aligned}$$

   $$v^T \cdot \nabla_x^2 L(x^*; \lambda^*; \mu^*) \cdot v > 0,$$

   kde $\nabla_x^2 L$ *je Hessova matice funkce* L *vzhledem k proměnným* $x = (x_1, x_2, \ldots, x_n)$;
5. *(správný „směr“ od hranice* $\mathcal{M}$*)* $\mu_k^* \geq 0$, *pro každé* $k \in \hat{p}$.

Potom je x^* *bodem ostrého lokálního minima úlohy* (1).

Znaménko multiplikátoru



Věta 7.8 (Postačující podmínka existence ostrého lokálního maxima)

Nechť f, g_j, h_k pro $j \in \hat{m}, k \in \hat{p}$ mají spojité všechny druhé parciální derivace na nějaké otevřené nadmnožině $\tilde{\mathcal{M}} \supset \mathcal{M}$. Pokud trojice $(x^*; \lambda^*; \mu^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ splňuje podmínky:

1. *(0. derivace) $x^* \in \mathcal{M}$;*
2. *(1. derivace) $\forall i, \frac{\partial L}{\partial x_i}(x^*; \lambda^*; \mu^*) = 0$;*
3. *(aktivní a neaktivní vazby) Pro každé $k \in \hat{p}$, $\mu_k^* = 0$ nebo $h_k(x^*) = 0$;*
4. *(2. derivace) pro každý (sloupcový) vektor $0 \neq v \in \mathbb{R}^n$ splňující*

$$\begin{aligned} \nabla g_j(x^*) \cdot v &= 0, \quad \text{pro všechna } j \in \hat{m}, \\ \nabla h_k(x^*) \cdot v &= 0, \quad \text{pro všechna } k \in \hat{p}, \mu_k^* \neq 0, \ \text{platí} \end{aligned}$$

$$v^T \cdot \nabla_x^2 L(x^*; \lambda^*; \mu^*) \cdot v < 0,$$

   kde $\nabla_x^2 L$ je Hessova matice funkce L vzhledem k proměnným $x = (x_1, x_2, \ldots, x_n)$;

5. *(správný „směr“ od hranice $\mathcal{M}$) $\mu_k^* \leq 0$, pro každé $k \in \hat{p}$.*

Potom je x^* bodem ostrého lokálního maxima úlohy (1).