

M5170: MATEMATICKÉ PROGRAMOVÁNÍ

PETR ZEMÁNEK (MASARYKOVA UNIVERZITA, BRNO)

Kapitola 1: Úvod

(verze: 20. prosince 2024)



Úvod



Obrázek převzat z goo.gl/TBdxET

Cum enim Mundi universi fabrica sit perfectissima atque a Creatore sapientissimo absoluta, nihil omnino in mundo contingit, in quo non maximi minimive ratio quaequam eluceat.

/Jelikož uspořádání vesmíru je na nejvýše dokonalé a ve skutečnosti dílem nejmoudřejšího Stvořitele, neexistuje ve vesmíru zhola nic, z čeho by nevyzařoval nějaký princip maxima nebo minima./

Leonhard Euler (1707–1783)

Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici lattissimo sensu accepti (1744; viz Additamentum 1 /str. 245/ dostupné na adrese goo.gl/9u4An2)

1.1 ZÁKLADNÍ VYMEZENÍ

1.2 „OPAKOVÁNÍ“

1.3 NEOPTIMALIZOVA(TEL)NÝ HISTORICKÝ EXKURZ

Obecnou optimalizační úlohou můžeme rozumět problém nalezení takového bodu v dané (přípustné) množině, že v něm dané „úcelové zobrazení“ přiřazující prvkům přípustné množiny hodnoty z $\mathbb{R}$ nabývá v jistém smyslu optima (max./min.), tj. úlohu

$$F(x) \rightarrow \text{opt}, \quad x \in X.$$

Řešením této úlohy rozumíme prvek $x^* \in X$ takový, že

$$x^* \in \arg \text{opt}_{x \in X} F(x) := \{x \in X \mid F(x) \bowtie F(y) \text{ pro všechna } y \in X\}.$$

V některých případech ovšem není nutné (nebo možné) takový bod x^* najít a místo toho se spokojíme pouze s nalezením optimální hodnoty úcelového zobrazení, tj. číslo

$$F^* := \widetilde{\text{opt}}_{x \in X} F(x).$$

Takto formulované optimalizační úlohy můžeme podle charakteru neznámých veličin rozdělit do dvou skupin:

- funkční optimalizace ($\rightsquigarrow$ variační počet)
- parametrické optimalizace ($\rightsquigarrow$ matematické programování)

FUNKČNÍ OPTIMALIZACE

Přípustná množina = prostor funkcí (skalárních nebo vektorových) n -rozměrné pro-měnné definovaných na jisté množině $\Omega \subseteq \mathbb{R}^n$, s jistými vlastnostmi (spojitost, di-ferencovatelnost apod.) a splňujícími další (tzv. vedlejší) podmínky dané soustavou nějakých algebraických, transcendentních, diferenciálních nebo integrálních rovnic či nerovnic

Účelové zobrazení = funkcionál (tj. zobrazení přiřazující každé funkci z přípustné mno-žiny prvek z $\mathbb{R}$)

V takovém případě tedy hledáme funkci z přípustné množiny, pro kterou daný funk-cionál nabývá největší/nejmenší hodnotu mezi všemi funkcemi z přípustné množiny.

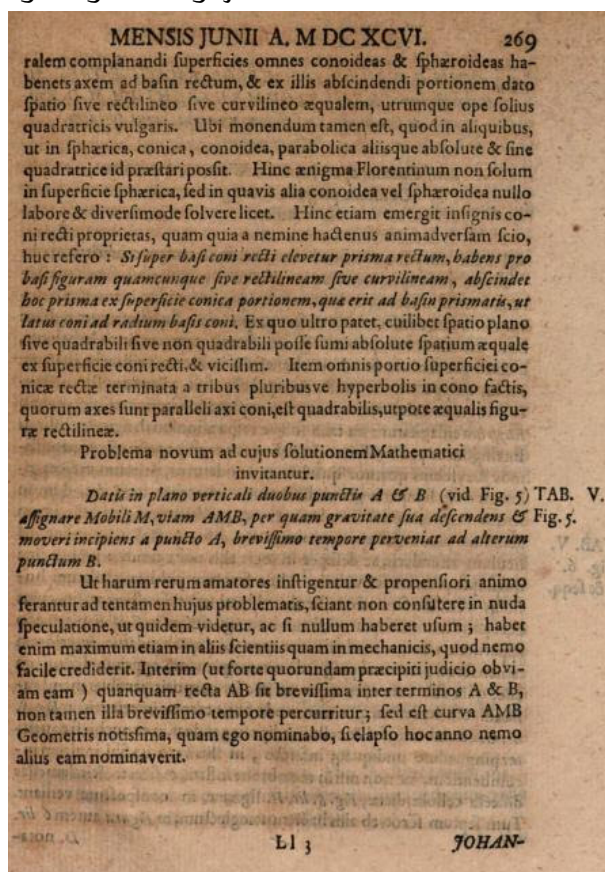
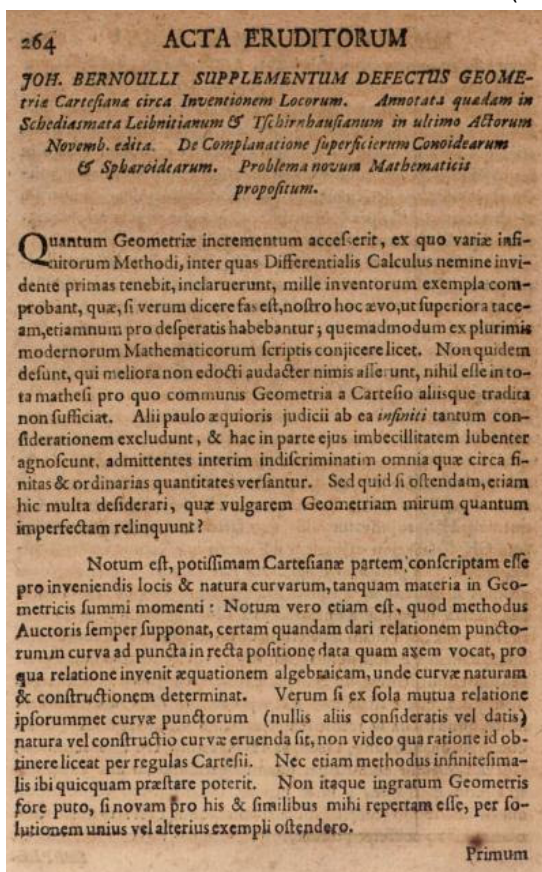


Např. mezi všemi funkcemi třídy $C^1[a, b]$ najít tu, jejíž graf je nejkratší spojnici bodů $A = [x_1, y_1]$ a $B = [x_2, y_2]$, tj.

$$\mathcal{J}[f] = \int_{x_1}^{x_2} \sqrt{1 + [f'(x)]^2} dx \rightarrow \min.$$

VARIAČNÍ POČET (M6800; JARO 2018 + 2k, $k \in \mathbb{N} \cup \{0\}$)

Johan Bernoulli, Acta Eruditorum (1696), viz goo.gl/26Fgvj





Já, Johann Bernoulli, oslovuji nejbystřejší matematiky na světě. Není nic přitažlivějšího pro inteligentní lidi než čestný a náročný problém, jehož možné řešení přinese slávu a zůstane trvalým památníkem. Následující řady příkladů Pascala, Fermata atd. doufám, že získám vděčnost celé vědecké komunity tím, že před nejlepší matematiky naší doby postavím problém, který otestuje jejich metody a sílu jejich intelektu. Jestliže mi kdokoli sdělí řešení navrhovaného problému, veřejně prohlásím, že je hoden uznání.

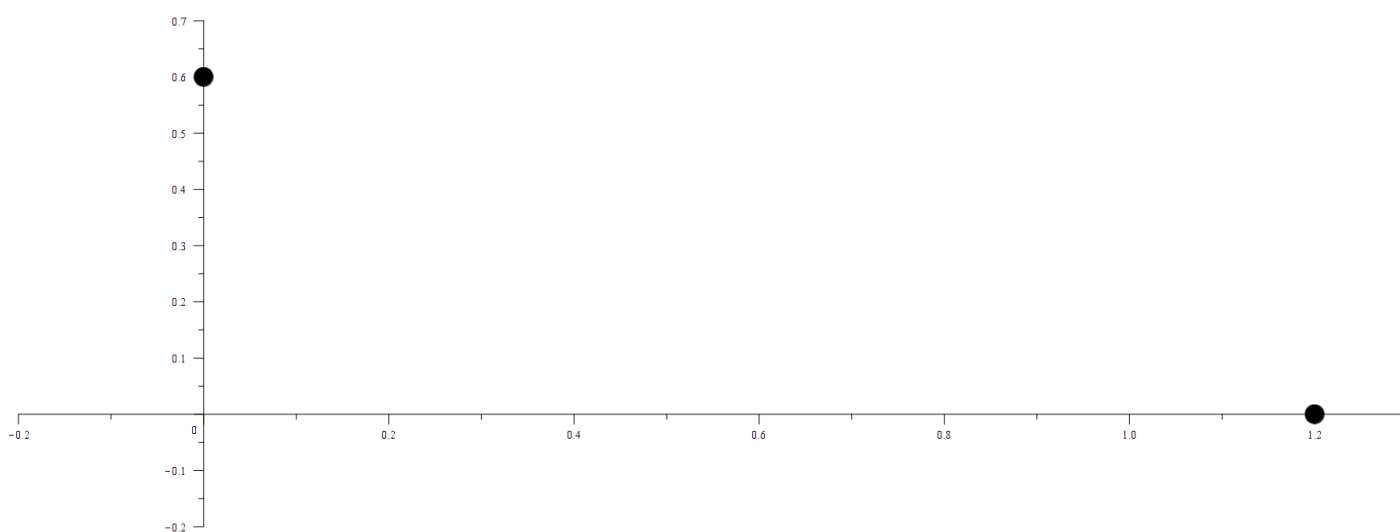
Obrázek převzat z goo.gl/UqiNUc

Problema Novum, ad cuius Solutionem Mathematici invitantur.

Datis in plano verticali duobus punctis A et B assignare mobili M viam AMB, per quam gravitate sua descendens et moveri incipiens a puncto A, brevissimo tempore perveniat ad alterum punctum B.

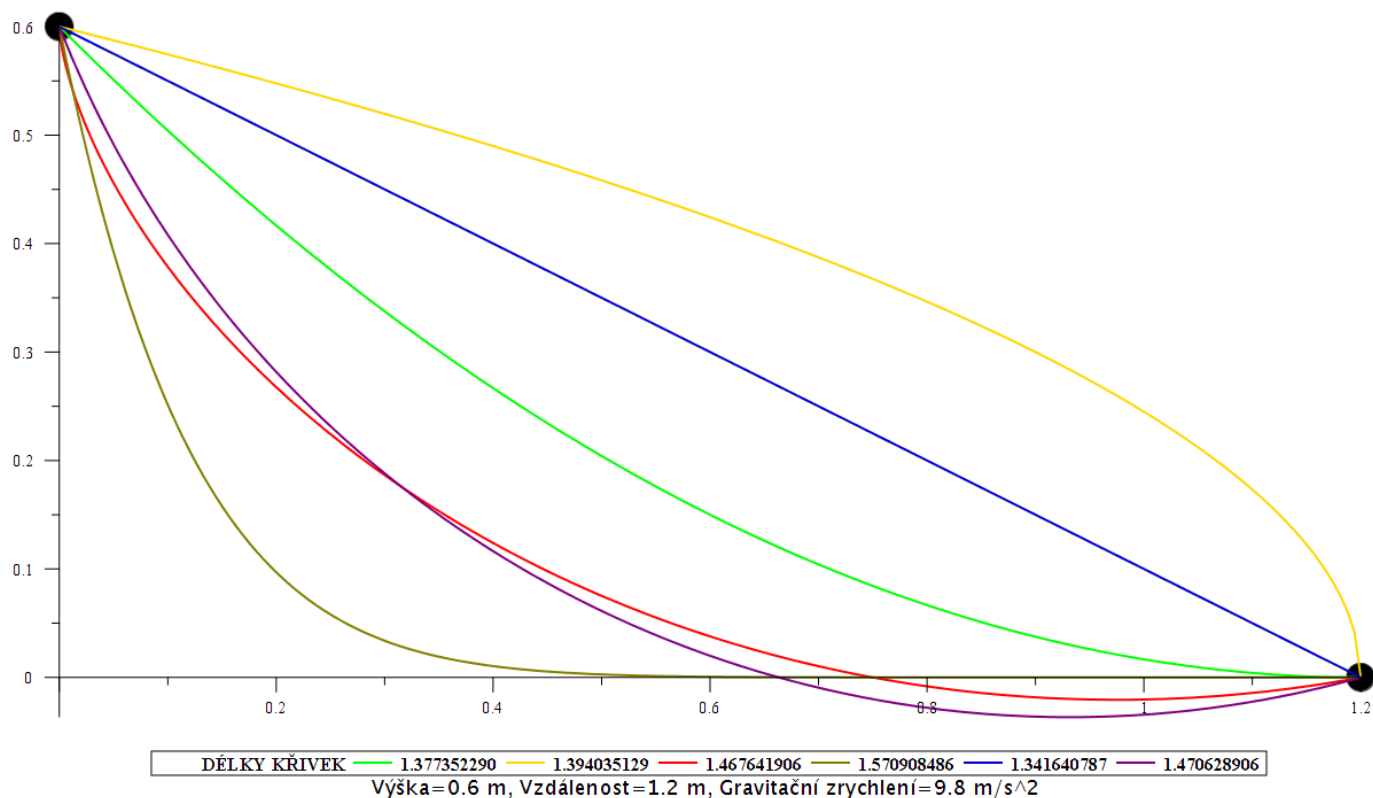
/Nová úloha, k jejímuž řešení se vyzývají matematikové: Nechť jsou dány dva body A a B ve svislé rovině. Určete křivku vyznačenou hmotným bodem, který startuje v bodě A a pouhým působením gravitace dosáhne bodu B v nejkratším možném čase (=brachistochrona)./

ÚLOHA O BRACHISTOCHRONĚ – VOLBA BODŮ

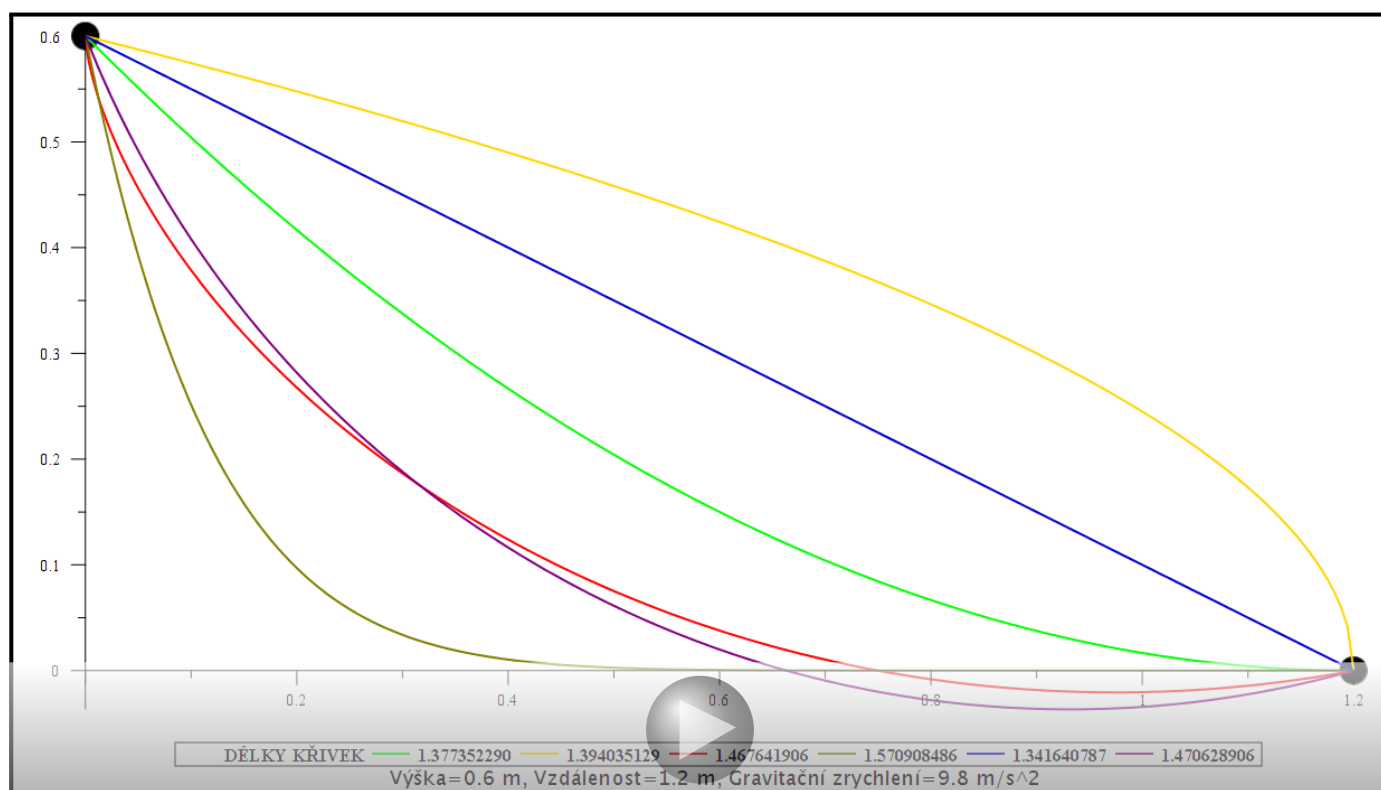


Výška=0.6 m, Vzdálenost=1.2 m, Gravitační zrychlení=9.8 m/s²

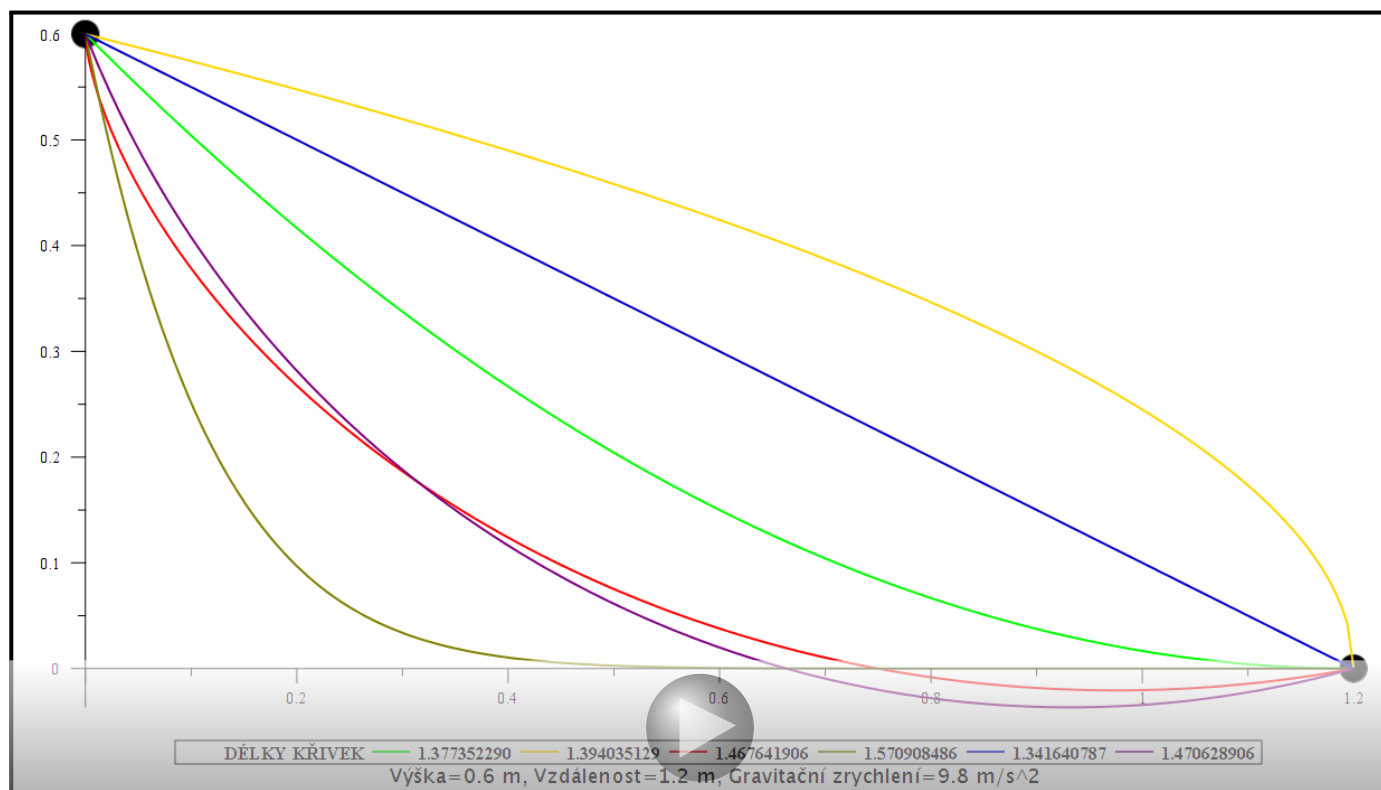
MOŽNÁ ŘEŠENÍ



ANIMACE



ANIMACE



KŘÍVKA ["PŘÍMKA", 0.7824607963]

KŘÍVKA ["POLYNOM", 0.6406276898]

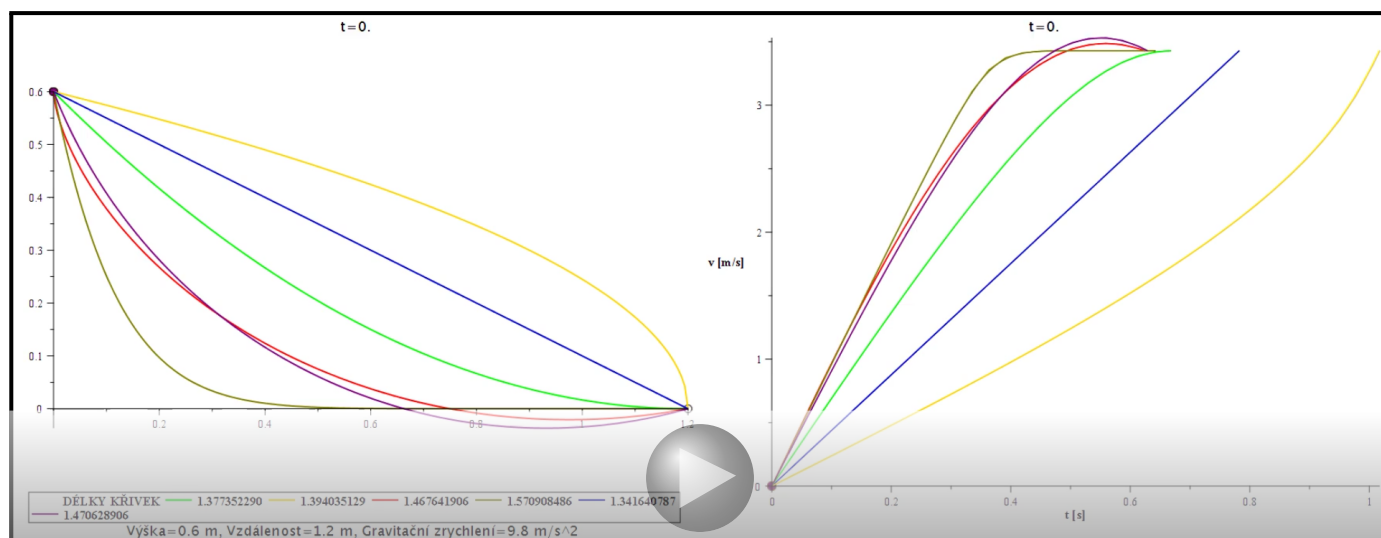
KŘÍVKA ["PARABOLA1", 0.6683953784]

KŘÍVKA ["KRUŽNICE", 0.6286226767]

KŘÍVKA ["PARABOLA2", 1.017818985]

KŘÍVKA ["BRACHISTOCHRONA", 0.6243053388]

ANIMACE #2



KŘÍVKA ["PŘÍMKA", 0.7824607963]

KŘÍVKA ["POLYNOM", 0.6406276898]

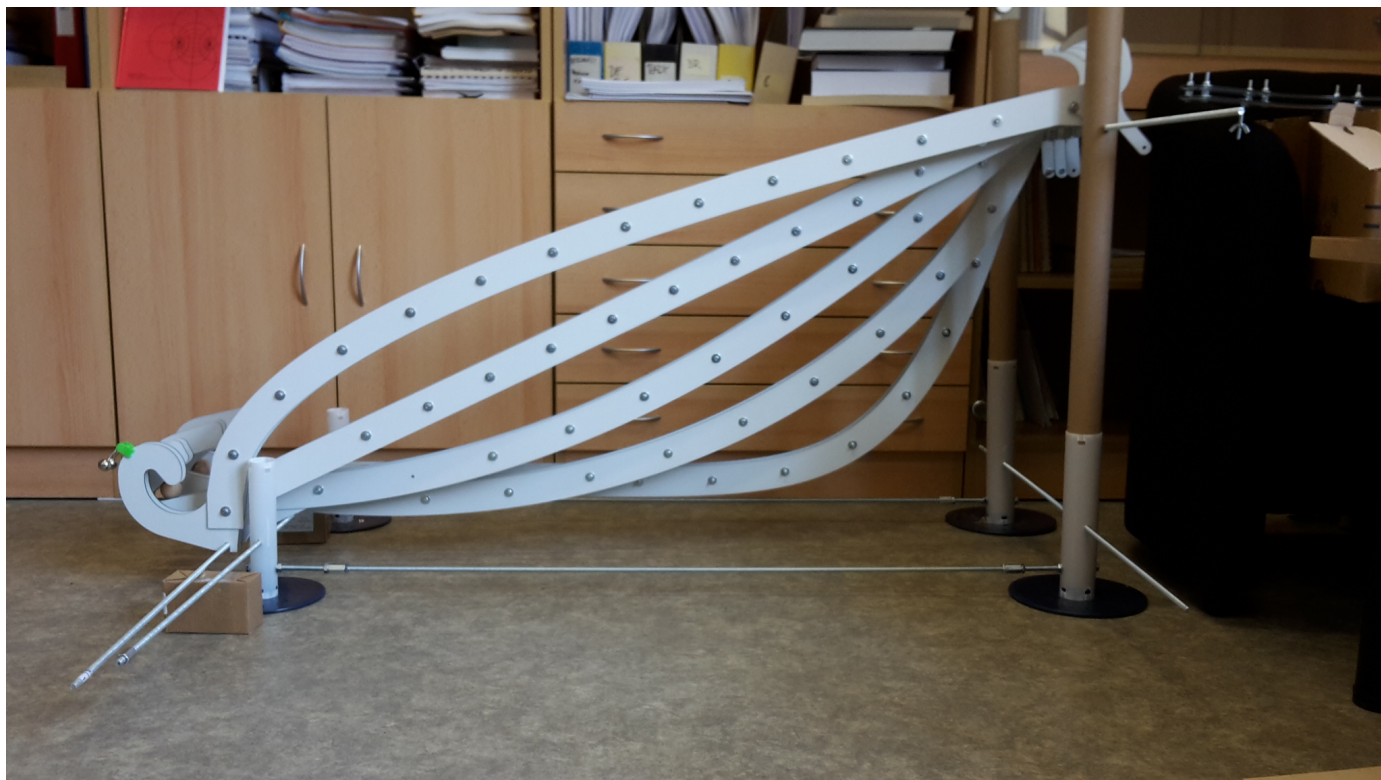
KŘÍVKA ["PARABOLA1", 0.6683953784]

KŘÍVKA ["KRUŽNICE", 0.6286226767]

KŘÍVKA ["PARABOLA2", 1.017818985]

KŘÍVKA ["BRACHISTOCHRONA", 0.6243053388]

„NA ŽIVO:“ 02021A, ÚMS, PŘF MU



↑ Obrázek převzat z goo.gl/rV9esA

BVV – Pavilon A →
Obrázek převzat z goo.gl/hCHQc6





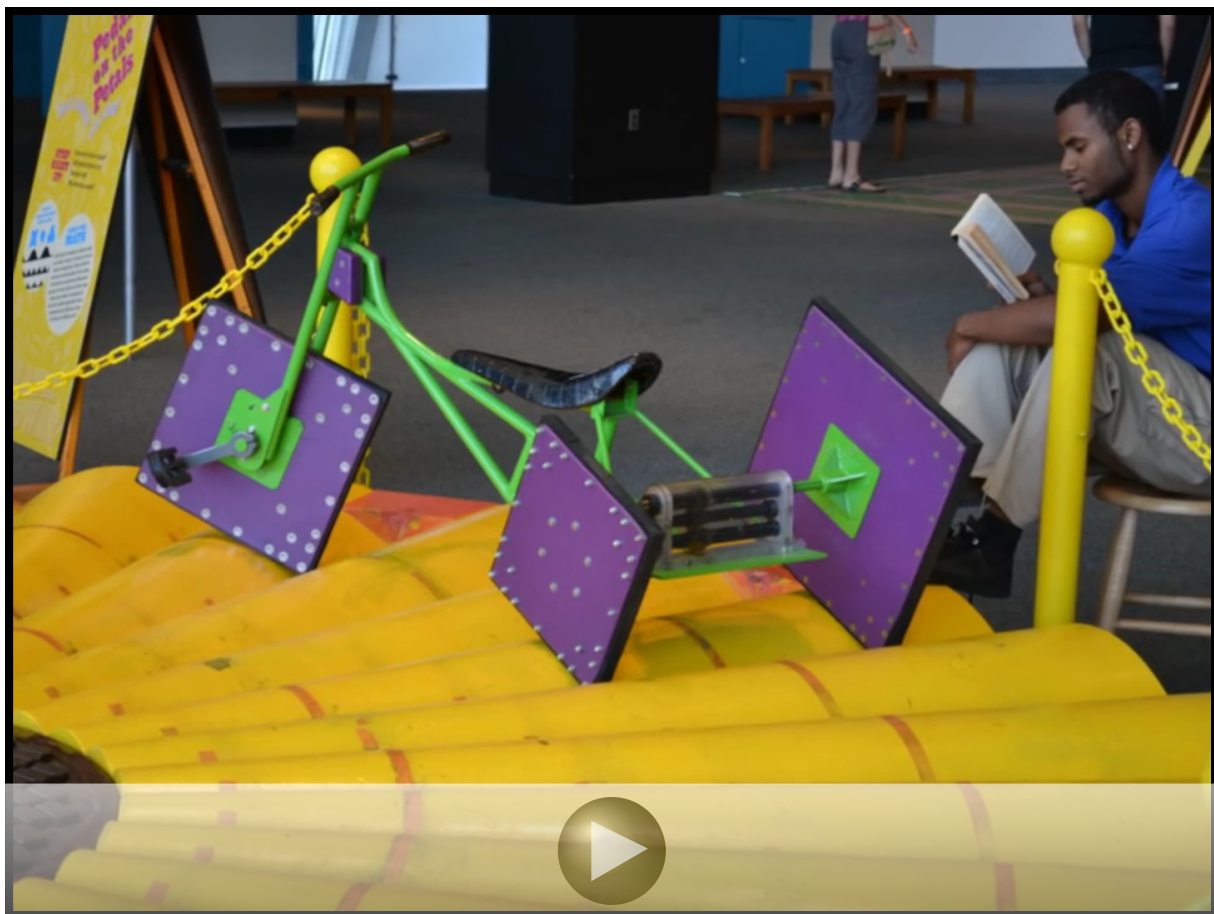
↑ BVV – Pavilon A
Obrázek převzat z goo.gl/EmGpzh



↑ Casa Milà (Barcelona © Antonio Gaudí)
Obrázek převzat z goo.gl/CAfYRe



Gateway Arch (St. Louis)



Video převzato z goo.gl/RSeshr

PARAMETRICKÁ OPTIMALIZACE

Přípustná množina = množina $X \subseteq \mathbb{R}^n$ vymezená jistými podmínkami danými soustavou nějakých algebraických nebo transcendentních rovnic či nerovnic

účelové zobrazení = reálná funkce f definovaná (alespoň) na přípustné množině. V tomto kontextu bývá (v závislosti na oblasti) funkce f označována jako účelová, cílová, ztrátová, nákladová, nepřímá užitková, užitková, energetická atd.

V takovém případě tedy hledáme bod $x^* \in X$, ve kterém daná reálná funkce f nabývá své nejmenší hodnoty mezi všemi body $x \in X$, tj. máme řešit úlohu

$$f(x) \rightarrow \min, \quad x \in X, \quad (1.1)$$

přičemž množinu X můžeme vyjádřit jako

$$X := \{x \in \mathbb{R}^n \mid g_1(x) \leq 0, \dots, g_L(x) \leq 0 \text{ \& } h_1(x) = 0, \dots, h_K(x) = 0\}$$

pro $K, L \in \mathbb{N} \cup \{0\}$ a dané funkce $g_1, \dots, g_L, h_1, \dots, h_K$, tj. hledáme

$$x^* \in \arg \min_{x \in X} f(x) = \{x \in X \mid f(x) \leq f(y) \text{ pro všechna } y \in X\}.$$

Úlohu (1.1) pak čteme jako: Najdi bod x^* , v němž reálná funkce f dosahuje nejmenší hodnoty (tj. minima) mezi všemi body $x \in D(f) \subseteq \mathbb{R}^n$, které vyhovují soustavě L nerovností

$$g_1(x) \leq 0 \quad \& \quad g_2(x) \leq 0 \quad \& \quad \dots \quad \& \quad g_L(x) \leq 0$$

a soustavě K rovností

$$h_1(x) = 0 \quad \& \quad h_2(x) = 0 \quad \& \quad \dots \quad \& \quad h_k(x) = 0.$$

Případně můžeme místo (1.1) řešit „pouze“ úlohu

$$f^* = \inf_{x \in X} f(x) \tag{1.2}$$

(1.1) vs. (1.2)

Řešení úlohy (1.1) pak není nic jiného než nalezení vrstevnice funkce f na úrovni f^* .

Další
klasifikace

Úlohu (1.1) navíc můžeme rozdělit do dvou subkategorií

- úlohy bez omezujících podmínek (tzv. volné extrém), tj. $X = \mathbb{R}^n$, např. $x^2 + 1 \rightarrow \min$
- úlohy s omezujícími podmínkami (tzv. vázané extrém), tj. $X \subsetneq \mathbb{R}^n$, např. $x \cos y \rightarrow \min$, $[x, y] \in [-5, 5] \times \mathbb{R}$.

Co již známe?

1.1 ZÁKLADNÍ VYMEZENÍ

1.2 „OPAKOVÁNÍ“

1.3 NEOPTIMALIZOVA(TEL)NÝ HISTORICKÝ EXKURZ

Definice 1.1

Bod $x^* \in X$ nazveme bodem globálního minima funkce f na X , přičemž $X \subseteq D(f)$, tj. řešením úlohy (1.1), pokud $f(x^*) \leq f(x)$ pro všechny $x \in X$. Bod $x^* \in X$ nazveme bodem lokálního minima funkce f na X , neboli lokálním řešením úlohy (1.1), jestliže existuje takové ε -okolí bodu x^* , tj.

$$\mathcal{O}_\varepsilon(x^*) := \{x \in \mathbb{R}^n \mid \|x - x^*\| < \varepsilon\}, \quad \varepsilon > 0,$$

že $f(x^*) \leq f(x)$ pro každé $x \in X \cap \mathcal{O}_\varepsilon(x^*)$.

V případě ostrých nerovností hovoříme o ostrých globálních/lokálních minimech. Analogicky definujeme globální/lokální (ostrá) maxima.

Jak je najít?

Věta 1.2
(Fermat)

Nechť pro funkci $f : \mathbb{R}^n \rightarrow \mathbb{R}$ existují všechny parciální derivace 1. řádu v bodě x^* , ve kterém má funkce f lokální extrém na $\mathbb{R}^n$. Potom x^* je stacionárním bodem funkce f , tj.

$$\text{grad } f(x^*) = \begin{pmatrix} f_{x_1}(x^*) \\ f_{x_2}(x^*) \\ \vdots \\ f_{x_n}(x^*) \end{pmatrix} = 0.$$

Věta 1.2a

Má-li funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$ v bodě $x^* \in \mathbb{R}^n$ a nějakém jeho okolí spojitě parciální derivace 2. řádu, $\text{grad } f(x^*) = 0$ a symetrická $n \times n$ (Hessova) matice

$$\nabla^2 f(x^*) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x^*) \right)_{i,j=1,\dots,n}$$

je pozitivně definitní, pak bod x^* je bodem ostrého lokálního minima funkce f na $\mathbb{R}^n$.

Jak se pozná pozitivní definitnost?

Platí také opačná implikace? Jak se pozná pozitivní semidefinitnost?

Existence globálních extrémů?

Věta 1.3
(Weierstrass)

Nechť $X \subseteq \mathbb{R}^n$ je kompaktní množina a $f : X \rightarrow \mathbb{R}$ je spojitá funkce na X . Pak funkce f nabývá své největší a nejmenší hodnoty na X (tj. globálních extrémů).

Dokážeme extrémy „lokalizovat“?

Věta 1.4

Nechť $X \subseteq \mathbb{R}^n$ je kompaktní množina a funkce $f : X \rightarrow \mathbb{R}$ je diferencovatelná na X . Pak f nabývá svého maxima a minima na X buď ve stacionárním bodě ležícím uvnitř X nebo v některém hraničním bodě množiny X .

Praktické?

Věta 1.5

Nechť $f, g_1, \dots, g_m$ jsou funkce třídy C^1 na otevřené množině $\mathcal{U} \subseteq \mathbb{R}^n$, kde $m \in \{1, \dots, n\}$. Uvažme množinu $\mathcal{M}$ danou jako

$$\mathcal{M} := \bigcap_{i=1}^m \{x \in \mathbb{R}^n \mid g_i(x) = 0\} \subseteq \mathcal{U},$$

přičemž vektory $\text{grad } g_1(x), \dots, \text{grad } g_m(x)$ jsou lineárně nezávislé (a tedy nutně $m \leq n$) pro všechna $x \in \mathcal{M}$, tj. Jacobiho matice

$$DG(x) = D(g_1(x), \dots, g_m(x))^T = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x) & \dots & \frac{\partial g_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1}(x) & \dots & \frac{\partial g_m}{\partial x_n}(x) \end{pmatrix} \in \mathbb{R}^{m \times n}$$

má plnou hodnost, tj. $\text{rank } DG(x) = m$. Je-li bod $x^* \in \mathcal{M}$ lokálním extrémem funkce f na množině $\mathcal{M}$, pak existují taková čísla $\lambda_1^*, \dots, \lambda_m^* \in \mathbb{R}$ (tzv. Lagrangeovy multiplikátory), že x^* je stacionárním bodem Lagrangeovy funkce

$$L(x, \lambda^*) = f(x) + \sum_{i=1}^m \lambda_i^* g_i(x),$$

tj. $\text{grad}_x L(x^*, \lambda^*) = 0$ neboli

$$\frac{\partial f}{\partial x_j}(x^*) + \sum_{i=1}^m \lambda_i^* \frac{\partial g_i}{\partial x_j}(x^*) = 0 \quad \text{pro všechna } j = 1, \dots, n.$$

Co je množina $\mathcal{M}$? Množina

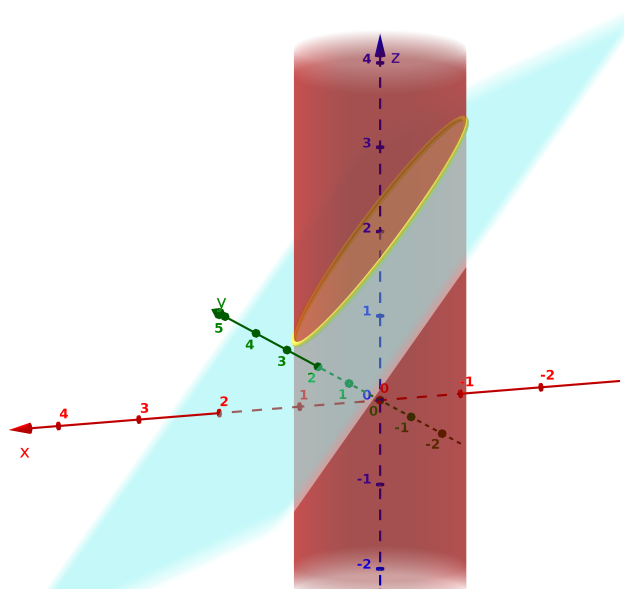
$$\{x \in \mathbb{R}^n \mid g_1(x) = 0\}$$

obvykle představuje nadplochu, tj. „útvár“ dimenze $n - 1$. Přidáním další nadplochy dostaneme

$$\{x \in \mathbb{R}^n \mid g_1(x) = 0\} \cap \{x \in \mathbb{R}^n \mid g_2(x) = 0\},$$

což bude mít dimenzi $n - 2$ atd. Tedy $\mathcal{M}$ bude $(n - m)$ -dimenzionální plocha v $\mathbb{R}^n$. Např.

$$g_1(x, y, z) = x + y + z - 2, \quad g_2(x, y, z) = x^2 + y^2 - 1$$



Věta 1.6

Co znamená, že x^* je stacionárním bodem funkce $L(x, \lambda^*)$?

Nechť $f, g_1, \dots, g_m$ jsou funkce třídy C^2 na otevřené množině $\mathcal{U} \subseteq \mathbb{R}^n$, kde $m \in \{1, \dots, n\}$. Uvažme množinu $\mathcal{M}$ danou jako

$$\mathcal{M} := \bigcap_{i=1}^m \{x \in \mathbb{R}^n \mid g_i(x) = 0\} \subseteq \mathcal{U},$$

přičemž vektory $\text{grad } g_1(x), \dots, \text{grad } g_m(x)$ jsou lineárně nezávislé pro všechna $x \in \mathcal{M}$. Jestliže pro bod $x^* \in \mathcal{M}$ existují multiplikátory $\lambda_1^*, \dots, \lambda_m^* \in \mathbb{R}$ takové, že platí následující podmínky

- (i) Lagrangeova funkce $L(x, \lambda^*)$ má v bodě x^* stacionární bod, tj. $\text{grad}_x L(x^*, \lambda^*) = 0$,
- (ii) Hessova matice $\nabla^2 L(x^*, \lambda^*)$ je pozitivně definitní na

$$\begin{aligned} \text{Ker}(DG(x^*)) &= \text{span}\{\text{grad } g_1(x^*), \dots, \text{grad } g_m(x^*)\}^\perp \\ &= \bigcap_{i=1}^m \{x \in \mathbb{R}^n \mid x \perp \text{grad } g_i(x^*)\}, \end{aligned}$$

pak funkce f má v bodě x^* ostré vázané lokální minimum.

Poznámka

K tomu, aby bod x^* byl lokálním vázaným extrémem není nutné (ale postačující), aby byl současně lokálním extrémem funkce $L(x, \lambda^*)$. Avšak je-li x^* vázaným lokálním minimem, nutně $\nabla^2 L(x^*, \lambda) \geq 0$ na $\text{Ker}(DG(x^*))$.

1.1 ZÁKLADNÍ VYMEZENÍ

1.2 „OPAKOVÁNÍ“

1.3 NEOPTIMALIZOVA(TEL)NÝ HISTORICKÝ EXKURZ

Leonid Vitaliyevich KANTOROVICH (1912–1986)

George Bernard DANTZIG (1914–2005)

William KARUSH (1917–1997)

Harold William KUHN (1925–2014)

Albert William TUCKER (1905–1995)

Fritz JOHN (1910–1994)

Moritz Werner FENCHEL (1905–1988)

R. Tyrrell ROCKAFELLAR (★1935, viz goo.gl/kvDL8y)

*V této oblasti (na rozdíl od některých jiných) není klíčové rozlišovat mezi lineární a nelineární úlohou, ale mezi **konvexní** a **nekonvexní** úlohou.*

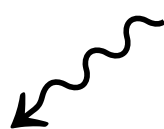


KONVEXNÍ ANALÝZA

V případě $X \subseteq \mathbb{R}^n$ představuje bod x^* v úloze (1.1) optimální řešení (strategii, též návrhový vektor aj.) tzv. matematického programu (IT?!)



Matematické programování (M5170)

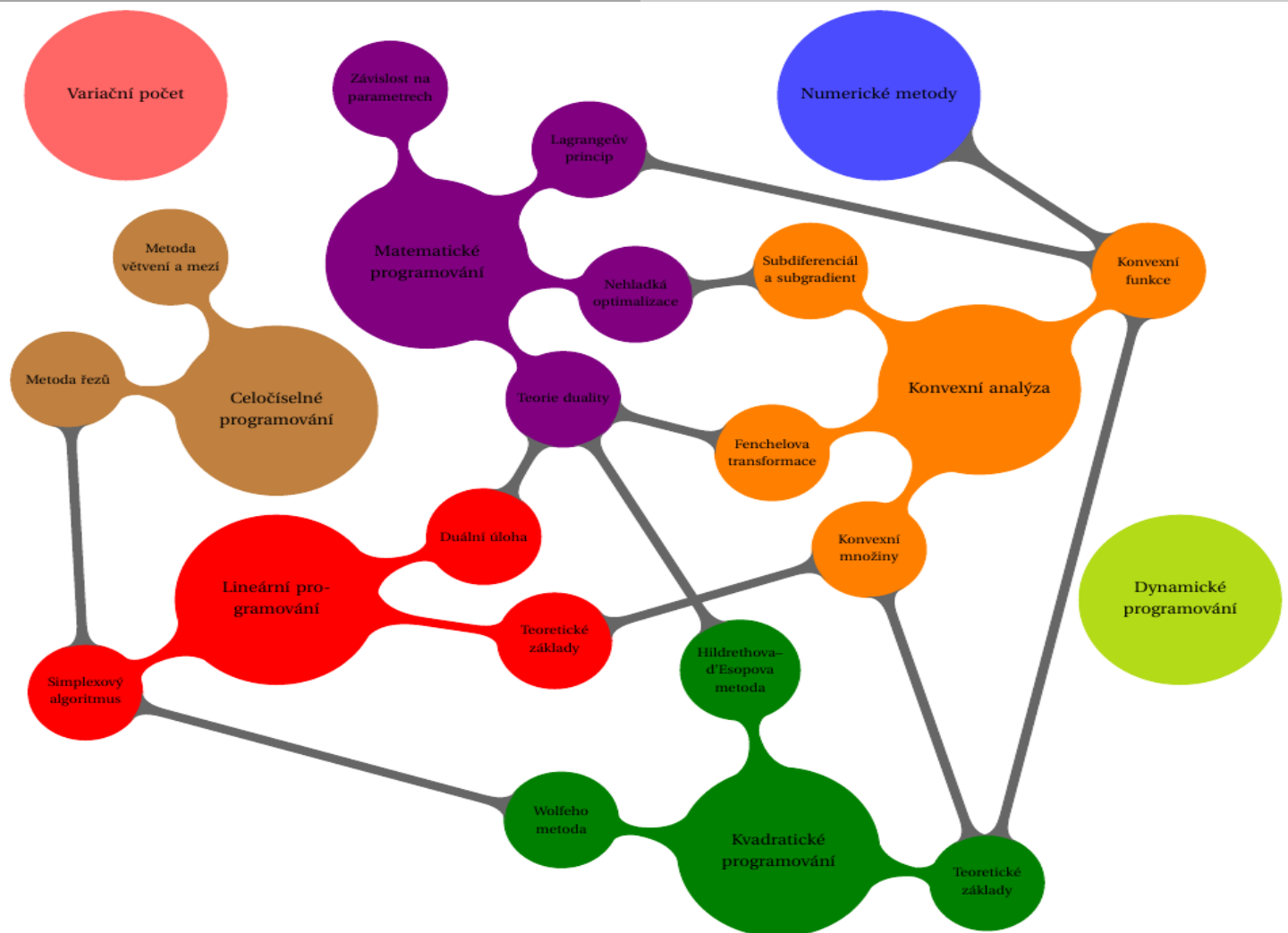


Lineární programování
(M411/0)

Kvadratické programování

Celočíselné programování
(M8150)

Optimalizace (M0160; každý jarní semestr)



LITERATURA

-  O. Došlý, *Základy konvexní analýzy a optimalizace v $\mathbb{R}^n$* , Masarykova univerzita, 2005.
-  M. Hamala, *Nelineárne programovanie*, druhé vydání, Alfa, Bratislava, 1976.
-  A. Sucharev, A. Timochov, V. Fedorov, *Kurs metodov optimizacii*, Nauka, Moskva, 1986.
-  R. Rockafellar, *Convex Analysis*, Princeton University Press, New Jersey, 1970.
-  S. M. Sinha, *Mathematical Programming – Theory and Methods*, Elsevier, Delhi, 2006.
-  W. Sun, Y.-X. Yuan, *Optimization Theory and Methods – Nonlinear Programming*, Springer, New York, 2006.
-  D. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, 1999.
-  D. Bertsekas, A. Nedic, A. Ozdeglar, *Convex Analysis and Optimization*, Athena Scientific, Belmont, 2003.
-  D. Bertsekas, *Convex Optimization Theory*, Athena Scientific, Belmont, 2009.

Konec.