M5170: Matematické programování

Petr Zemánek (Masarykova Univerzita, Brno)

Kapitola 1: Úvod

(verze: 13. září 2021)

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 quaepiam 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)

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

$$\mathbb{F}(\mathbb{x}) \to \mathsf{opt}, \quad \mathbb{x} \in \mathbb{X}.$$

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

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

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

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

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 proměnné definovaných na jisté množině $\Omega \subseteq \mathbb{R}^n$, s jistými vlastnostmi (spojitost, diferencovatelnost 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 množiny prvek z $\mathbb{R}$)

V takovém případě tedy hledáme funkci z přípustné množiny, pro kterou daný funkcionál 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 spojnicí bodů $A = [x_1, y_1]$ a $B = [x_2, y_2]$, tj.

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

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

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

JOH. BERNOULLI SUPPLEMENTUM DEFECTUS GEOMEtriæ Carteſianæ circa Inventionem Locorum. Annotata quædam in Schediasmata Leibnitianum & Tſchirnhauſianum in ultimo Actorum Novemb. edita. De Complanatione ſuperficierum Conoidearum & Sphæroidearum. Problema novum Mathematicis propoſitum.

Quantum Geometriæ incrementum acceſſerit, ex quo variæ infinitorum Methodi, inter quas Differentialis Calculus nemine invidente primas tenebit, inclaruerunt, mille inventorum exempla comprobant, quæ, ſi verum dicere fas eſt, noſtro hoc ævo, ut ſuperiora taceam, etiamnum pro deſperatis habebantur; quemadmodum ex plurimis modernorum Mathematicorum ſcriptis conjicere licet. Non quidem deſunt, qui meliora non edocti audacter nimis aſſerunt, nihil eſſe in tota matheſi pro quo communis Geometria a Carteſio aliisque tradita non ſufficiat. Alii paulo æquioris judicii ab ea *infiniti* tantum conſiderationem excludunt, & hac in parte ejus imbecillitatem lubenter agnoſcunt, admittentes interim indiſcriminatim omnia quæ circa finitas & ordinarias quantitates verſantur. Sed quid ſi oſtendam, etiam hic multa deſiderari, quæ vulgarem Geometriam mirum quantum imperfectam relinquunt?

Notum eſt, potiſſimam Carteſianæ partem conſcriptam eſſe pro inveniendis locis & natura curvarum, tanquam materia in Geometricis ſummi momenti: Notum vero etiam eſt, quod methodus Auctoris ſemper ſupponat, certam quandam dari relationem punctorum in curva ad puncta in recta poſitione data quam axem vocat, pro qua relatione invenit æquationem algebraicam, unde curvæ naturam & conſtructionem determinat. Verum ſi ex ſola mutua relatione ipſorummet curvæ punctorum (nullis aliis conſideratis vel datis) natura vel conſtructio curvæ eruenda ſit, non video qua ratione id obtinere liceat per regulas Carteſii. Nec etiam methodus infiniteſimalis ibi quicquam præſtare poterit. Non itaque ingratum Geometris fore puto, ſi novam pro his & ſimilibus mihi repertam eſſe, per ſolutionem unius vel alterius exempli oſtendero.

Primum

ralem complanandi ſuperficies omnes conoideas & ſphæroideas habenets axem ad baſin rectum, & ex illis abſcindendi portionem dato ſpatio ſive rectilineo ſive curvilineo æqualem, utrumque ope ſolius quadratricis vulgaris. Ubi monendum tamen eſt, quod in aliquibus, ut in ſphærica, conica, conoidea, parabolica aliisque abſolute & ſine quadratrice id præſtari poſſit. Hinc ænigma Florentinum non ſolum in ſuperficie ſphærica, ſed in quavis alia conoidea vel ſphæroidea nullo labore & diverſimode ſolvere licet. Hinc etiam emergit inſignis coni recti proprietas, quam quia a nemine hactenus animadverſam ſcio, huc refero: *Si ſuper baſi coni recti elevetur prisma rectum, habens pro baſi figuram quamcunque ſive rectilineam ſive curvilineam, abſcindet hoc prisma ex ſuperficie conica portionem, quæ erit ad baſin prismatis, ut latus coni ad radium baſis coni.* Ex quo ultro patet, cuilibet ſpatio plano ſive quadrabili ſive non quadrabili poſſe ſumi abſolute ſpatium æquale ex ſuperficie coni recti, & viciſſim. Item omnis portio ſuperficiei conicæ rectæ terminata a tribus pluribusve hyperbolis in cono factis, quorum axes ſunt paralleli axi coni, eſt quadrabilis, utpote æqualis figuræ rectilineæ.

Problema novum ad cujus ſolutionem Mathematici invitantur.

Datis in plano verticali duobus punctis A & B (vid. Fig. 5) TAB. V.
aſſignare Mobili M, viam AMB, per quam gravitate ſua deſcendens & Fig. 5.
moveri incipiens a puncto A, breviſſimo tempore perveniat ad alterum punctum B.

Ut harum rerum amatores inſtigentur & propenſiori animo ferantur ad tentamen hujus problematis, ſciant non conſiſtere in nuda ſpeculatione, ut quidem videtur, ac ſi nullum haberet uſum; habet enim maximum etiam in aliis ſcientiis quam in mechanicis, quod nemo facile crediderit. Interim (ut forte quorundam præcipiti judicio obviam eam) quanquam recta AB ſit breviſſima inter terminos A & B, non tamen illa breviſſimo tempore percurritur; ſed eſt curva AMB Geometris notiſſima, quam ego nominabo, ſi elapſo hoc anno nemo alius eam nominaverit.

Ll 3 JOHAN-

Obrázek převzat z goo.gl/UqiNUc

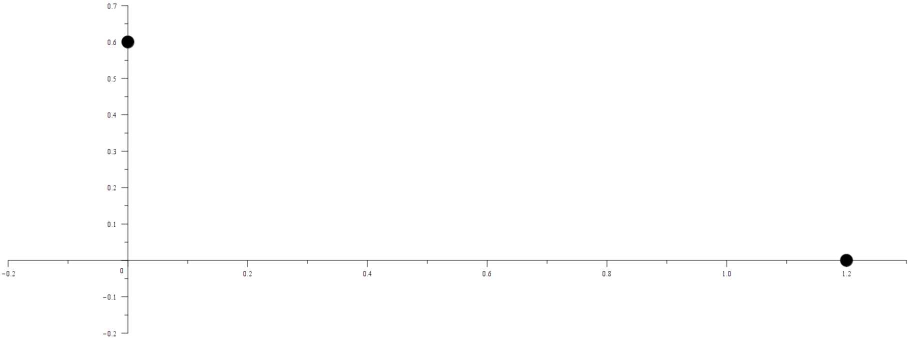
Já, Johann Bernoulli, oslovuji nejbystřejší matematiky na světě. Není nic přitažlivějšího pro inteligentní lidi něž čestný a náročný problém, jehož možné řešení přinese slávu a zůstane trvalým památníkem. Následujíce řady příkladů Pascala, Fermata atd. doufám, že získám vděčnost celé vědecké komunity 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.

Problema Novum, ad cujus 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*.*

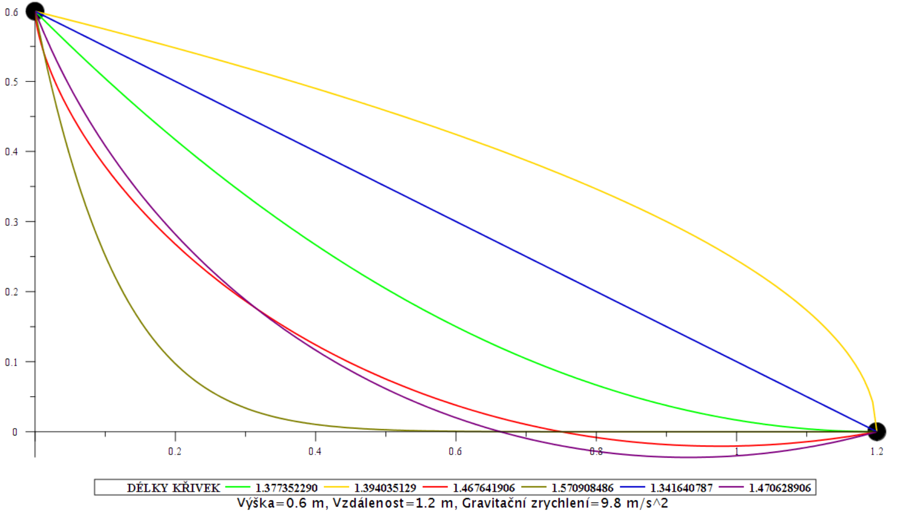
/Nová úloha, k jejímuž řešení se vyzývají matematikové: Nechť jsou dány dva body A a B ve svislé rovině. Určete křivku vyznačenou hmotným bodem, který startuje v bodě A a pouhým působením gravitace dosáhne bodu B v nejkratším možném čase (=brachistochrona)./

Úloha o brachistochroně – volba bodů

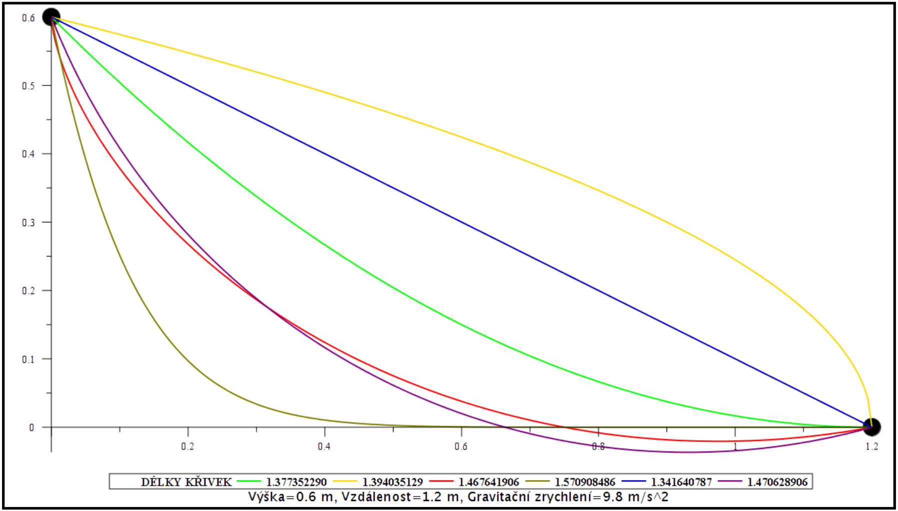


Výška=0.6 m, Vzdálenost=1.2 m, Gravitační zrychlení=9.8 m/s^2

Možná řešení

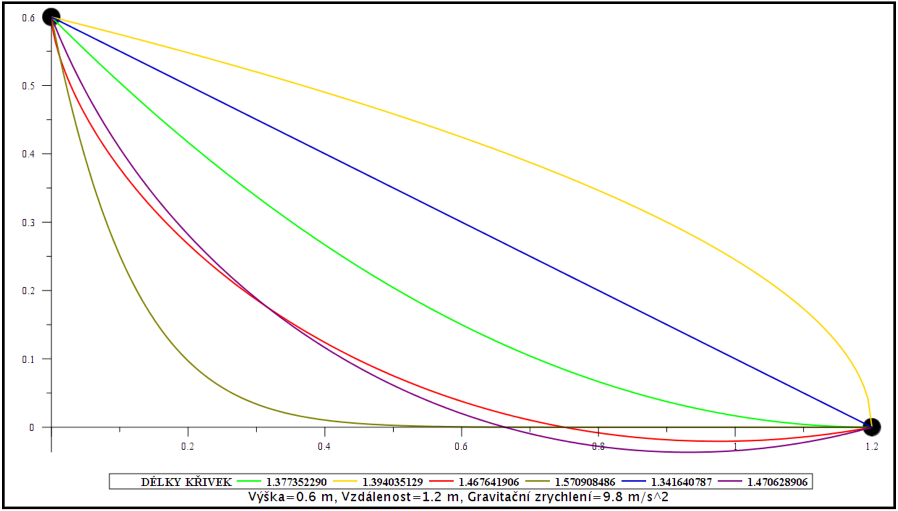


Animace



Play/Pause

Animace



KŘIVKA ["PŘÍMKA", 0.7824607963]

KŘIVKA ["PARABOLA1", 0.6683953784]

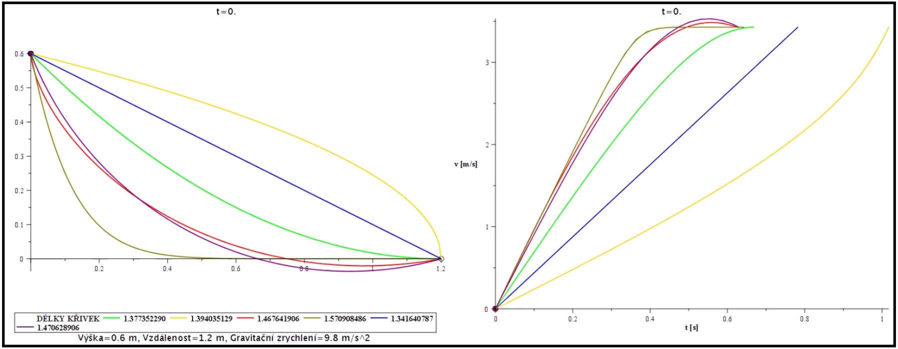
KŘIVKA ["PARABOLA2", 1.017818985]

KŘIVKA ["POLYNOM", 0.6406276898]

KŘIVKA ["KRUŽNICE", 0.6286226767]

KŘIVKA ["BRACHISTOCHRONA", 0.6243053388]

Animace #2



KŘIVKA ["PŘÍMKA", 0.7824607963]

KŘIVKA ["PARABOLA1", 0.6683953784]

KŘIVKA ["PARABOLA2", 1.017818985]

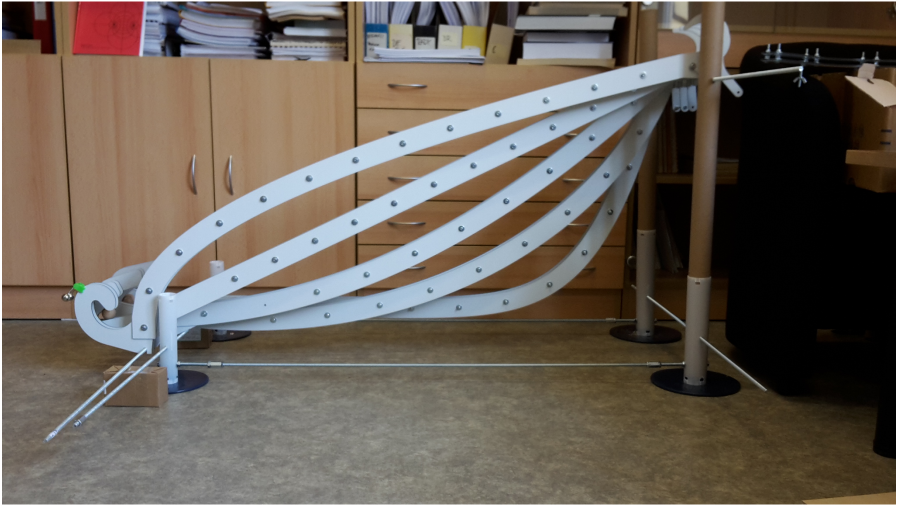
KŘIVKA ["POLYNOM", 0.6406276898]

KŘIVKA ["KRUŽNICE", 0.6286226767]

KŘIVKA ["BRACHISTOCHRONA", 0.6243053388]

Play/Pause

„Na živo:“ 02021a, ÚMS, PřF MU



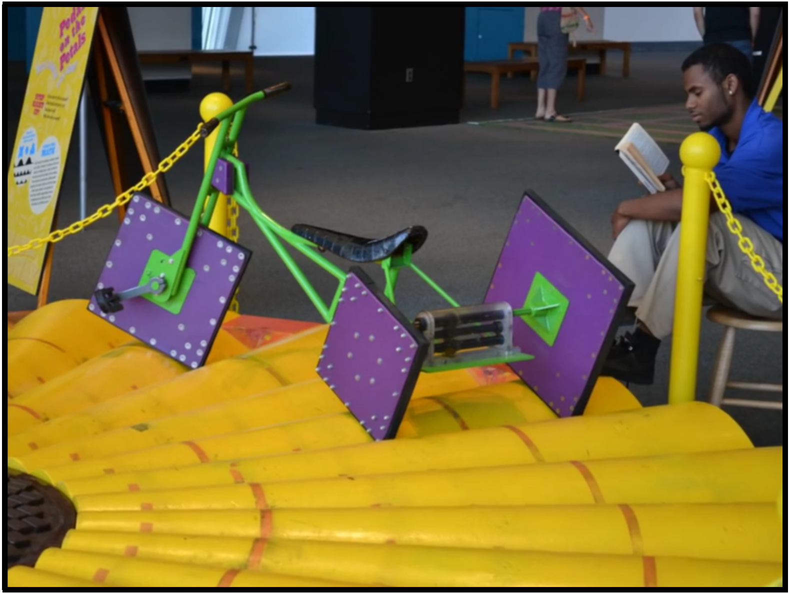
↑ Obrázek převzat z goo.gl/rV9esA

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)



Play/Pause

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

$$f(x) \to \min, \quad x \in X, \tag{1.1}$$

přičemž množinu X můžeme vyjádřit jako

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

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

$$x^* \in \operatorname*{arg\,min}_{x \in X} f(x) = \{x \in X \mid f(x) \leqslant f(y) \text{ pro všechna } y \in X\}.$$

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

$$g_1(x) \leqslant 0 \quad \& \quad g_2(x) \leqslant 0 \quad \& \quad \cdots \quad \& \quad g_L(x) \leqslant 0$$

a soustavě K rovností

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

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

Co již známe?

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 $f(x^*) \leqslant f(x)$ 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 $f(x^*) \leqslant f(x)$ 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 $f : \mathbb{R}^n \to \mathbb{R}$ 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.

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

Věta 1.2a

Má-li funkce $f : \mathbb{R}^n \to \mathbb{R}$ v bodě $x^* \in \mathbb{R}^n$ a nějakém jeho okolí spojité parciální derivace 2. řádu, $\operatorname{grad} f(x^*) = 0$ 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 (Weiestrass)

Nechť $X \subseteq \mathbb{R}^n$ je kompaktní množina a $f : X \to \mathbb{R}$ 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 $f : X \to \mathbb{R}$ 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

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

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

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

má plnou hodnost, tj. $\operatorname{rank} DG(x) = m$. 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. $\operatorname{grad}_x L(x^*, \lambda^*) = 0$ 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

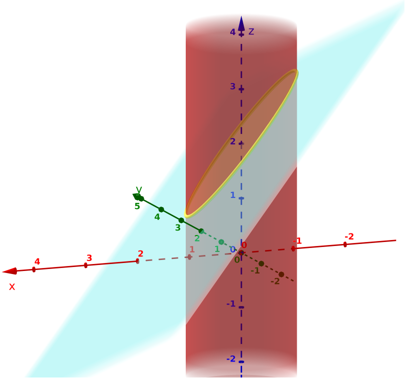
$$\{x \in \mathbb{R}^n \mid g_1(x) = 0\}$$

obvykle představuje nadplochu, tj. „útvar“ 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.

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



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

Věta 1.6

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

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

přičemž vektory $\operatorname{grad} g_1(x), \dots, \operatorname{grad} g_m(x)$ 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. $\operatorname{grad}_x L(x^*, \lambda^*) = 0$,

(ii) Hessova matice $\nabla^2 L(x^*, \lambda^*)$ je pozitivně definitní na

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

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 $\nabla^2 L(x^*, \lambda) \geqslant 0$ na $\operatorname{Ker}(DG(x^*))$.

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

$$\Downarrow$$

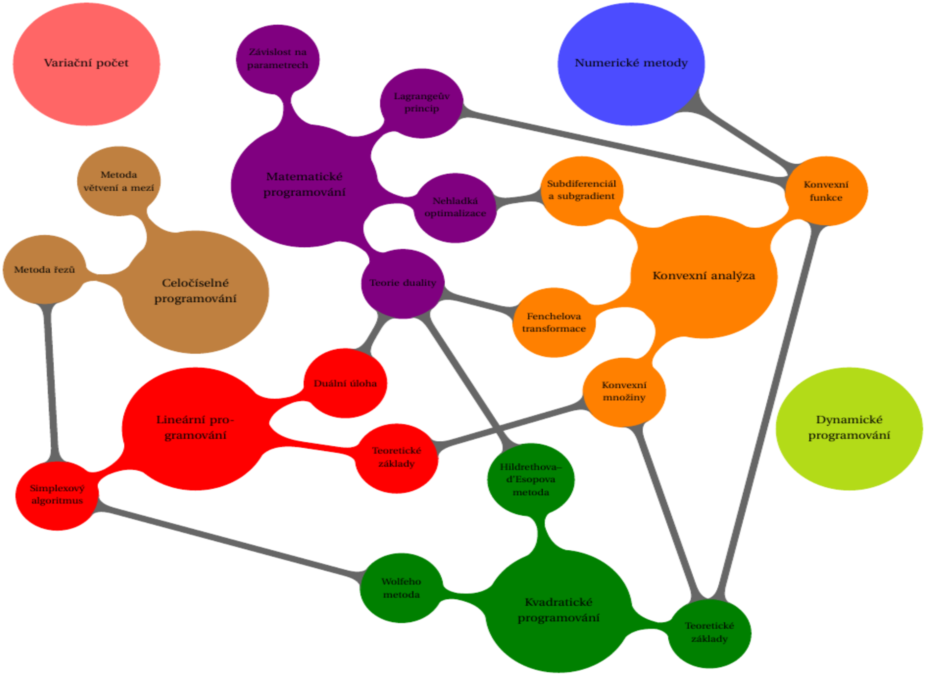
$\boxed{\text{Matematické programování (M5170)}}$

Lineární programování (~~M4110~~)

Kvadratické programování

Celočíselné programování (~~M8150~~)

Teorie optimalizace (M0160; každý jarní semestr)


Variační počet
Závislost na parametrech
Lagrangeův princip
Numerické metody
Matematické programování
Metoda větvení a mezí
Nehladká optimalizace
Subdiferenciál a subgradient
Konvexní funkce
Metoda řezů
Celočíselné programování
Konvexní analýza
Teorie duality
Fenchelova transformace
Duální úloha
Konvexní množiny
Dynamické programování
Lineární programování
Teoretické základy
Hildrethova–d'Esopova metoda
Simplexový algoritmus
Wolfeho metoda
Kvadratické programování
Teoretické základy

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.