

Metódy voľnej optimalizácie - letný semester 2018

M. Trnovská, KAMŠ, FMFI UK

Informácie o predmete

METÓDY VOĽNEJ OPTIMALIZÁCIE

- **Prednášajúca:** M. Trnovská (M 267)
Cvičiaci: J. Komadel
- Informácie týkajúce sa predmetu na stránke

<http://www.iam.fmph.uniba.sk/institute/trnovska/mvo.html>

- Konzultácie sa dajú dohodnúť mailom.
trnovska@pc2.iam.fmph.uniba.sk

Podmienky hodnotenia

- **Domáce úlohy na cvičeniach**
 - 40 bodov
 - Výstup: matlabovský kód, spracovanie úlohy (výstupu, komentár, grafy, tabuľky) do pdf súboru (latex)
- **Skúška** - 60 bodov, na absolvovanie predmetu treba získať aspoň 15 bodov!
 - Minimálna kostra - základné znalosti
 - Písomka
 - Ústna časť (nepovinná)
- **Bonusové úlohy (?)**

Podmienky hodnotenia

body	známka
91-100	A
81-90	B
71-80	C
61-70	D
51-60	E

Literatúra

- Základná literatúra:
M. Hamala, M. Trnovská: Nelineárne programovanie (1. časť)
- Doplnková literatúra:
S. Boyd, L. Vandenberghe: Convex optimization
A. Antoniou, L. Wu-Sheng: Practical Optimization

Harmonogram prednášok

Úvod do predmetu

Metódy minimalizácie funkcie jednej premennej

- Metódy intervalovej aproximácie
- Metódy bodovej aproximácie

Metódy minimalizácie funkcie n premenných

- Klasické metódy (Gradientné metódy, Metóda CSR, Newtonova metóda)
- Moderné metódy (Združené gradienty, Kvázinewtonovské metódy)
- Metódy pre veľkorozmerné úlohy

Základná myšlienka metód pre minimalizáciu úloh s ohraničeniami

Úloha matematického programovania:

$$\min_{x \in K \subseteq \mathbb{R}^n} f_0(x)$$

- **Úlohy na voľný extrém:** K je otvorená množina alebo $K = \mathbb{R}^n$
- **Úlohy s ohraničeniami:** K je uzavretá množina, spravidla definovaná pomocou rovností a nerovností:
$$K = \{x \in \mathbb{R}^n \mid f_i(x) \leq 0, i \in I, h_j(x) = 0, j \in J\},$$
kde $I \neq \emptyset$ alebo $J \neq \emptyset$.
 - nekonvexné
 - konvexné (lineárne, nelineárne)

Metódy voľnej optimalizácie = metódy riešenia úloh na voľný extrém

- **Úlohy bez ohraničení** (rekonštrukcia signálu, klasifikácia dát - logistická regresia,...)
- **Úlohy s ohraničeniami** - riešia sa algoritmami, ktoré sú založené na postupnom riešení (parametrizovaných) úloh na voľný extrém, pričom postupnosť voľných miním takýchto úloh konverguje k minimu pôvodnej úlohy s ohraničeniami.

Aplikácia: Rekonštrukcia signálu

Signál je snímaný v pravidelných časových intervaloch a reprezentovaný nejakou funkciou závislou od času $x_i = x(t_i)$, $i = 0, 1, \dots, T$

Predpokladá sa, že signál sa príliš nemení - $x_i \approx x_{i+1}$

Audio signál - jednorozmerný signál

Obrázky/video - viacrozmerné signály

Prijatý signál je obyčajne znečistený šumom v :

$$x_{sum} = x + v$$

Úloha: odhadnúť pôvodný signál x zo zašumeného signálu x_{sum}
(signal de-noising)

Formulácia úlohy:

$$\text{Min} \left\{ \|x - x_{sum}\|_2 + \gamma\phi(x) \mid x \in \mathbb{R}^{T+1} \right\}$$

$\phi(x)$ je nejaká regularizačná funkcia

Kvadratické vyhladenie

$$\phi_{quad}(x) = \sum_{i=0}^T (x_{i+1} - x_i)^2 = \|Dx\|_2^2$$

Totálna varácia

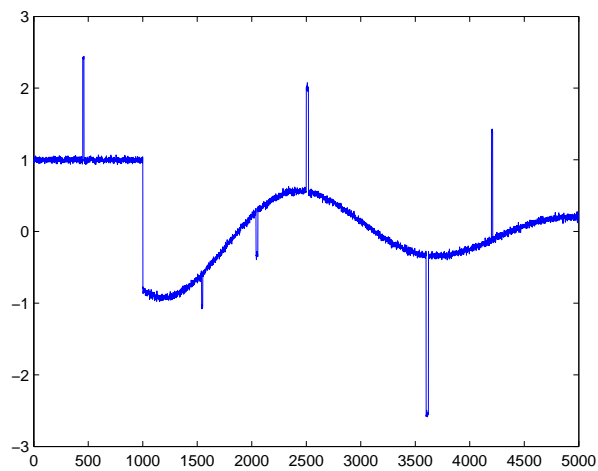
$$\phi_{tv}(x) = \sum_{i=0}^T |x_{i+1} - x_i| = \|Dx\|_1$$

$$D = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{bmatrix} \in \mathbb{R}^{T \times (T+1)}$$

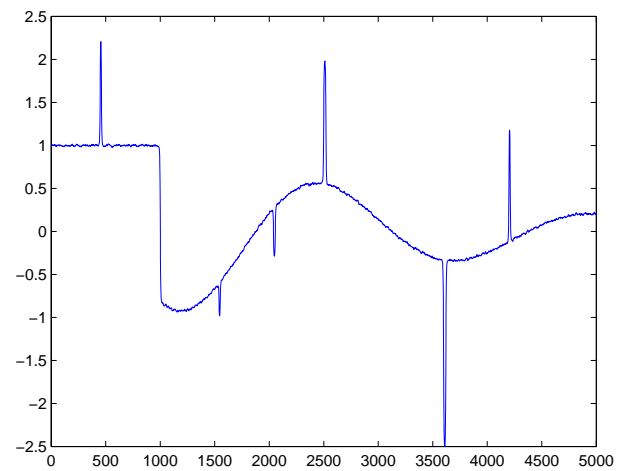
Ak je pôvodný signál hladký a šum sa prudko mení - $\phi_{quad}(x)$

Ak sa skokovito menia hodnoty pôvodného signálu - $\phi_{tv}(x)$

Kvadratické vyhladenie

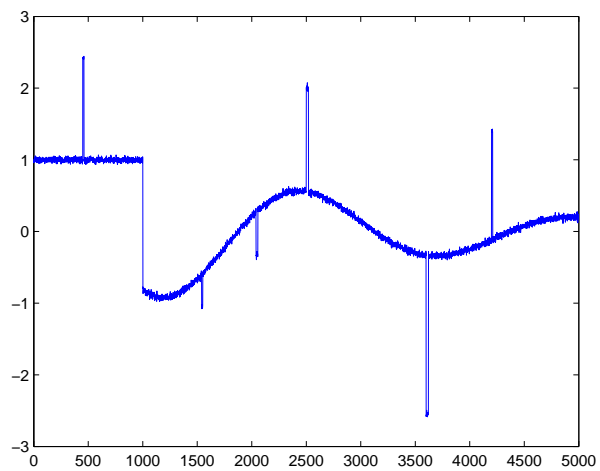


zašumený signál

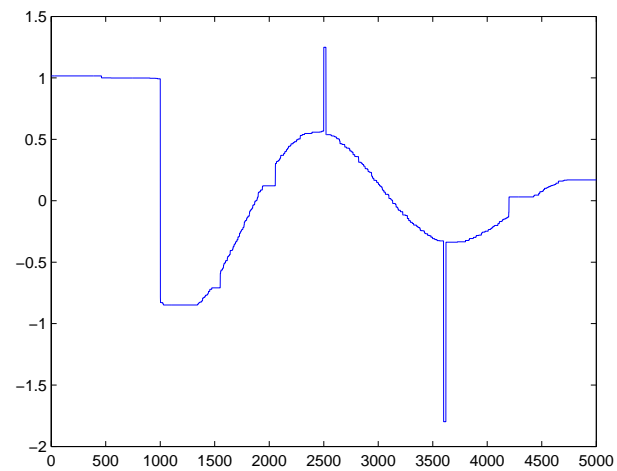


KV rekonštruovaný signál $\gamma = 0.9$

Totálna variácia



zašumený signál



TV rekonštruovaný signál $\gamma = 0.9$

Rekonštrukcia voacrozmerného signálu

$$\text{Min} \left\{ \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|Wx\|_1 \mid x \in \mathbb{R}^n \right\}$$

b - rozmazaný $m \times n$ obrázok reprezentovaný ako mn rozmerný vektor

A lineárna transformácia reprezentujúca "rozmazanie"

W ortogonálna matica reprezentujúca vlnkovú transformáciu (wavelet)

Observed Image



Reconstructed Image (lambda=0.001)

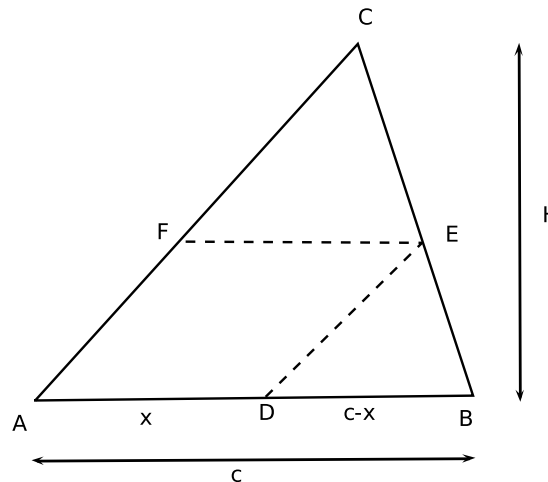


História optimalizačných úloh

Euklidova úloha:

Do daného trojuholníka ABC vpíšte rovnobežník $ADEF$ tak, že $AF \parallel DE$, $AD \parallel EF$ a jeho obsah je maximálny.

Ako naformulujeme Euklidov problém ako optimalizačnú úlohu?



Euklides našiel riešenie geometrickou úvahou

Dnes vieme, že stačí riešiť $f'(x) = 0$.

Fermat (1601-1665)

Newton (1643-1727)

Leibnitz (1646-1716)

Lagrange (1736-1813)

Cauchy (1789-1857)

Dantzig - 1947 - simplexová metóda

Kuhn, Tucker - 1950 - nelineárne programovanie

Klee, Minty - 1972

Karmarkar - 1984

Nesterov, Nemirovski - 1994

Základné princípy

Majme úlohu matematického programovania

$$\min_{x \in K \subseteq \mathbb{R}^n} f(x) \quad (MP)$$

Bod $\hat{x} \in K$ sa nazýva

globálne minimum úlohy (MP) ak $f(\hat{x}) \leq f(x) \forall x \in K$

ostré globálne minimum úlohy (MP) ak $f(\hat{x}) < f(x) \forall x \in K$

lokálne minimum úlohy (MP) ak

$$\exists \varepsilon > 0 : f(\hat{x}) \leq f(x) \quad \forall x \in K \cap \mathcal{O}_\varepsilon(\hat{x})$$

ostré lokálne minimum úlohy (MP) ak

$$\exists \varepsilon > 0 : f(\hat{x}) < f(x) \quad \forall x \in K \cap \mathcal{O}_\varepsilon(\hat{x})$$

Budeme používať značenie

$$g(x) = \nabla f(x), \quad H(x) = \nabla^2 f(x)$$

Nutné a postačujúce podmienky pre lokálne minimum úlohy na voľný extrém ($K = \mathbb{R}^n$):

Nutná podmienka 1. rádu: $g(\hat{x}) = 0$

Nutné podmienky 2. rádu: $g(\hat{x}) = 0, \quad H(\hat{x}) \succeq 0$

Postačujúce podmienky 2. rádu: $g(\hat{x}) = 0, \quad H(\hat{x}) \succ 0$

Ak je funkcia f konvexná: lokálne minimum=globálne minimum

Kritériá konvexnosti funkcie f

(Ak treba diferencovateľná) funkcia $f : \mathcal{D} \rightarrow \mathbb{R}$ je konvexná na \mathcal{D} ak množina \mathcal{D} je konvexná a

$$\forall x, y \in \mathcal{D}, \forall \lambda \in [0, 1] : f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

Kritérium 1. rádu

$$\forall x, y \in \mathcal{D} : f(x) \geq f(y) + \nabla f(y)^T (x - y)$$

Kritérium 2. rádu

$$\forall x \in \mathcal{D} : \nabla^2 f(x) \succeq 0$$

Úloha

$$\min_{x \in \mathbb{R}^n} f(x)$$

sa spravidla rieši **iteračným algoritmom**, ktorý generuje postupnosť bodov:

$$x^0, x^1, x^2, \dots \in \mathbb{R}^n \quad f(x^k) \rightarrow f(\hat{x}).$$

Algoritmus končí v tzv. **ε -presnom riešení**, t.j. platí

$$f(x^k) - f(\hat{x}) < \varepsilon \text{ alebo } \|g(\hat{x})\| < \varepsilon.$$

Všeobecná schéma algoritmov:

$$x^{k+1} = x^k + \lambda_k s^k,$$

kde s^k je smer (nenulový vektor) v k -tej iterácii, λ_k je dĺžka smeru s^k