

# Úplný problém nejmenších čtverců v úlohách s násobnou pravou stranou

doktorand:

ING. MARTIN PLEŠINGER

Fakulta mechatroniky  
Technická univerzita Liberec  
Hálkova 6  
461 17 Liberec 1

martin.plesinger@tul.cz

školitel:

PROF. ING. ZDENĚK STRAKOŠ, DRSC.

Ústav informatiky AV ČR v.v.i.  
Pod Vodárenskou věží 2  
182 07 Praha 8

strakos@cs.cas.cz

obor studia:  
Přírodovědné inženýrství

Tato práce byla podpořena grantem národního programu výzkumu "Informační společnost" č. 1ET400300415.

## Abstrakt

V tomto příspěvku se budeme zabývat klasifikací lineárních aproximačních úloh s ohledem na jejich řešitelnost ve smyslu formulace tzv. úplného problému nejmenších čtverců.

## 1. Úvod

Uvažujme lineární aproximační úlohu

$$AX \approx B, \quad (1)$$

kde  $A \in \mathbb{R}^{m \times n}$  a  $B \in \mathbb{R}^{m \times d}$  jsou matice systému a  $d$ -násobná pravá strana,  $X \in \mathbb{R}^{n \times d}$  je matice neznámých. (Pokud  $d = 1$  značíme pravou stranu  $b$  a vektor neznámých  $x$ .) Úloha (1) lze ekvivalentně přeformulovat

$$[B|A] \begin{bmatrix} -I_d \\ X \end{bmatrix} \approx 0.$$

Způsob aproximace upřesní následující definice.

### Definice 1 Minimalizační úlohu

$$\min_{G,E,X} \|[G|E]\|_F, \quad (A+E)X = (B+G), \quad (2)$$

nazveme *úplným problémem nejmenších čtverců (total least squares problem, TLS)*. Řešením úplného problému nejmenších čtverců nazveme libovolné  $X$  splňující (2).

Bez újmy na obecnosti budeme předpokládat  $m \geq n + d$  (v opačném případě matici  $[B|A]$  doplníme nulovými řádky). Dále předpokládejme  $A^T B \neq 0$  (v opačném případě, tedy jsou-li obory hodnot matic  $A$  a  $B$  vzájemně ortogonální, snaha aproximovat  $B$  pomocí sloupců matice  $A$  postrádá smysl a lze ukázat, že problém (2) má triviální řešení).

## 2. Úlohy s jednou pravou stranou

Úplný problém nejmenších čtverců pro  $d = 1$  (úlohu s jednou pravou stranou) prvně analyzovali Gene Golub a Charles Van Loan [1], 1980. Ukázali, že úloha (2) nemá obecně řešení. Za předpokladu, že řešení existuje, nemusí být jednoznačné; zavádí se řešení minimální v normě.

Sabine Van Huffel a Joos Vandewalle [5], 1991, nazývají problém (2) s  $d = 1$  jež nemá řešení *negenerickým*. Zavádějí pro něj tzv. *negenerické řešení*, které vždy existuje a lze definovat tak, že je jednoznačné. Význam negenerického řešení ovšem není příliš zřejmý.

Analýzu problému s jednou pravou stranou uzavírá článek Christophera Paige a Zdeňka Strakoše [6], 2006. Za předpokladu ortogonální invariance úlohy (1) lze ukázat, že existují matice  $P \in \mathbb{R}^{m \times m}$ ,  $P^{-1} = P^T$  a  $Q \in \mathbb{R}^{n \times n}$ ,  $Q^{-1} = Q^T$  takové, že

$$P^T [b|AQ] = \left[ \begin{array}{c|c|c} b_1 & A_{11} & 0 \\ \hline 0 & 0 & A_{22} \end{array} \right] \quad (3)$$

přičemž matice  $[b_1|A_{11}]$  má minimální dimenzi přes všechny ortogonální transformace vedoucí na blokově diagonální strukturu (3). Původní úloha (1) se tak rozpadne na dva nezávislé podproblémy

$$A_{11} x_1 \approx b_1, \quad A_{22} x_2 \approx 0.$$

Lze ukázat, [6, 1], že první z obou podproblémů je *vždy řešitelný* ve smyslu Definice 1 a navíc jeho řešení  $x_1$  je *jednoznačné*. Podproblém  $A_{11} x_1 \approx b_1$  nazýváme *core problémem*. Dále lze ukázat, že vektor

$$x \equiv Q \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$$

je identický s řešením minimálním v normě, dle Goluba a Van Loana [1], existuje-li, respektive s řešením negenerickým [5] v případě, že (2) řešení nemá. Teorie core problému tak dává konceptu negenerického řešení dobře interpretovatelný význam. Pro podrobnější výklad viz [6, 2, 3] případně [7].

### 2.1. Klasická analýza úloh s jednou pravou stranou

O tom, zda problém (2) s jednou pravou stranou má nebo nemá řešení, případně o tom, zda řešení, existuje-li, je jednoznačné či nikoliv, lze rozhodnout na základě pravých singulárních vektorů a distribuce singulárních čísel rozšířené matice  $[b|A]$ . Uvažujme tedy singulární rozklad

$$[b|A] = U \Sigma V^T \quad (4)$$

a

$$\sigma_1 \geq \dots \geq \sigma_p > \sigma_{p+1} = \dots = \sigma_{n+1} \geq 0 \quad (5)$$

singulární čísla matice  $[b|A]$ . Dále uvažujme následující dělení matice pravých singulárních vektorů

$$V = \left[ \begin{array}{c|c} V_{11} & V_{12} \\ \hline V_{21} & V_{22} \end{array} \right], \quad (6)$$

kde  $V_{11} \in \mathbb{R}^{1 \times p}$ ,  $V_{12} \in \mathbb{R}^{1 \times (n-p+1)}$ ,  $V_{21} \in \mathbb{R}^{n \times p}$ ,  $V_{22} \in \mathbb{R}^{n \times (n-p+1)}$ . (Pokud  $\sigma_1 = \sigma_{n+1}$ , pak  $p = 0$  a  $\sigma_p$ ,  $V_{11}$  a  $V_{21}$  neexistují. V [5] mají bloky matice (6) jiné pořadí.) Platí následující věta.

**Věta 1** *Nechť je dána lineární aproximační úloha (1) a  $d = 1$ . Uvažujme singulární rozklad (4), se značením zavedeným v (5)–(6).*

*Úplný problém nejmenších čtverců (2) má řešení tehdy a jen tehdy, když  $V_{12} \neq 0$ .*

*Navíc pokud  $p = n$ , pak je toto řešení jednoznačné, pokud  $p > n$  pak lze zkonstruovat řešení minimální v normě.*

Důkaz viz [1, 5]. (V opačném případě  $V_{12} = 0$  lze vždy, jak již bylo řečeno, zkonstruovat jednoznačné negenerické řešení, viz [4, 5], které však není řešením úlohy (2) ve smyslu Definice 1.)

### 3. Úlohy s násobnou pravou stranou

V článku [6] je pro úlohy s jednou pravou stranou využito klasické analýzy, zde shrnuté ve Větě 1, k důkazu jednoznačné řešitelnosti core problému ve smyslu Definice 1. Naší snahou je rozšířit tuto teorii,

zejména ideu redukce úlohy na core problém, na úlohy s násobnou pravou stranou, tedy pro  $d > 1$ . Je tedy nutné vědět, kdy je daný úplný problém nejmenších čtverců (2) řešitelný.

Analýzou existence a jednoznačnosti řešení pro úlohy s násobnou pravou stranou se zabývali Sabine Van Huffel a Joos Vandewalle, [5]. Analyzovali však jen některé *speciální případy*, obecná analýza chybí, viz [5, poznámka na str. 66]. Navzdory tomu algoritmus pro řešení úplného problému nejmenších čtverců, tzv. *TLS algoritmus*, viz [4], [5, Algoritmus 3.1, str. 87–88], vrátí „řešení“ pro libovolnou úlohu (1). Jedním z kroků vedoucích k rozšíření teorie core problému na úlohy s násobnou pravou stranou tak je zúplnění analýzy řešitelnosti problému (2).

### 3.1. Klasická analýza úloh s více pravými stranami

Speciální případy analyzované v [5] budeme opět identifikovat pomocí pravých singulárních vektorů a distribuce singulárních čísel rozšířené matice  $[B|A]$ . Uvažujme tedy singulární rozklad

$$[B|A] = U \Sigma V^T \quad (7)$$

a

$$\begin{aligned} \sigma_1 \geq \dots \geq \sigma_p > \sigma_{p+1} = \dots = \sigma_{n+1} = \dots \\ = \sigma_{n+e} > \sigma_{n+e+1} \geq \dots \geq \sigma_{n+d} \geq 0 \end{aligned} \quad (8)$$

singulární čísla matice  $[B|A]$ . Dále uvažujme následující dělení matice pravých singulárních vektorů

$$V = \left[ \begin{array}{c|c|c} V_{11} & V_{12} & V_{13} \\ \hline V_{21} & V_{22} & V_{23} \end{array} \right], \quad (9)$$

kde  $V_{11} \in \mathbb{R}^{d \times p}$ ,  $V_{12} \in \mathbb{R}^{d \times (n-p+e)}$ ,  $V_{13} \in \mathbb{R}^{d \times (d-e)}$ ,  $V_{21} \in \mathbb{R}^{n \times p}$ ,  $V_{22} \in \mathbb{R}^{n \times (n-p+e)}$ ,  $V_{23} \in \mathbb{R}^{n \times (d-e)}$ . (Pokud  $\sigma_1 = \sigma_{n+1}$ , pak  $p = 0$  a  $\sigma_p$ ,  $V_{11}$  a  $V_{21}$  neexistují, obdobně pokud  $\sigma_{n+1} = \sigma_{n+d}$ , pak  $e = d$  a  $\sigma_{n+e+1}$ ,  $V_{13}$  a  $V_{23}$  neexistují. V [5] je dělení matice (9) zavedeno jinak, bloky mají navíc jiné pořadí.) Platí následující věta.

**Věta 2** *Nechť je dána lineární aproximační úloha (1). Uvažujme singulární rozklad (7), se značením zavedeným v (8)–(9).*

*Nechť  $\text{rank}([V_{12}|V_{13}]) = d$  a zároveň  $p = n$  (tedy  $[V_{12}|V_{13}]$  je čtvercová nesingulární matice). Pak úplný problém nejmenších čtverců (2) má řešení a toto řešení je jednoznačné.*

*Nechť  $\text{rank}([V_{12}|V_{13}]) = d$  a zároveň  $e = d$  (tedy všechna singulární čísla počínaje  $\sigma_{p+1}$  jsou*

si rovna). Pak úplný problém nejmenších čtverců (2) má nekonečně mnoho řešení, lze zkonstruovat řešení minimální v normě (spektrální i Frobeniově).

Nechť  $\text{rank}([V_{12} | V_{13}]) < d$ . Pak úplný problém nejmenších čtverců (2) nemá řešení.

Důkaz viz [5]. V případě, že  $\text{rank}([V_{12} | V_{13}]) < d$ , lze vždy zkonstruovat jednoznačné negenerické řešení, viz [4, 5].

Všimněme si, že zatímco Věta 1 říká, že existence řešení je ekvivalentní s nenulovostí jistého bloku matice  $V$ , Věta 2 pouze implikuje existenci řešení ve dvou speciálních případech.

### 3.2. Revize analýzy úloh s více pravými stranami

Případ

$$\text{rank}([V_{12} | V_{13}]) = d, \quad p < n, \quad e < d \quad (10)$$

není z hlediska řešitelnosti ve smyslu Definice 1 v nám známé literatuře vůbec analyzován. Obvykle je tento případ interpretován následujícím způsobem: Protože aproximační úloha (1) je chápána jako *perturbace původně kompatibilního systému*, jsou všechna singulární čísla, tedy i  $\sigma_{n+e+1}, \dots, \sigma_{n+d}$ , zatížena chybou. Nahradíme-li tato singulární čísla čísla  $\tilde{\sigma}_{n+e+1}, \dots, \tilde{\sigma}_{n+d}$  rovnými  $\sigma_{n+1}$ , zredukujeme obecný případ (10) vždy na druhý speciální případ popsaný Větou 2, tzv. *truncated TLS* koncept, viz např. [5, poslední odstavec na str. 77]. Tím může být diskuze o existenci řešení uzavřena. S touto myšlenkou pak přirozeně koresponduje fakt, že klasický přístup, prezentovaný např. v [5], pracuje téměř výhradně s celým blokem  $[V_{12} | V_{13}]$ . Chceme-li ovšem diskutovat řešitelnost problému (2) obecně, bez uvažování dalších perturbací úlohy, musíme provést jemnější analýzu úlohy s více pravými stranami. Pro potřeby této analýzy navrhuje pracovat s bloky  $V_{12}$  a  $V_{13}$  odděleně.

Analýzu provedeme pro nejobecnější případ úlohy (1) splňující podmínku  $\text{rank}([V_{12} | V_{13}]) = d$ . Speciální případy úloh popsaných Větou 1 (úlohy s jednou pravou stranou) a Větou 2 přirozeně vyplynou jako případy se speciálními hodnotami  $d$ ,  $p$  a nebo  $e$ . Snadno nahlédneme, že matice  $V_{13} \in \mathbb{R}^{d \times (d-e)}$  nemůže mít za předpokladu  $\text{rank}([V_{12} | V_{13}]) = d$  rank větší než  $d - e$ , z čehož vyplývá, že matice  $V_{12} \in \mathbb{R}^{d \times (n-p+e)}$  nemůže mít rank menší než  $e$ . Plný (řádkový) rank matice  $[V_{12} | V_{13}]$  je tak možno mezi bloky  $V_{12}$  a  $V_{13}$  „rozdělit“ třemi různými způsoby. Když  $\text{rank}(V_{12}) = e$ , pak nutně  $V_{13}$  musí mít plný (sloupcový) rank, tedy

$\text{rank}(V_{13}) = d - e$ , opačná implikace ovšem neplatí. Může se tedy stát, že  $\text{rank}(V_{13}) = d - e$  a zároveň  $\text{rank}(V_{12}) > e$ . Třetí a poslední možností je, že matice  $V_{13}$  nebude mít plný (sloupcový) rank, tedy  $\text{rank}(V_{13}) < d - e$ , pak ale nutně  $\text{rank}(V_{12}) > e$ . Klasifikaci úloh zpřehlední následující definice.

**Definice 2** Uvažujme značení zavedené v (7)–(9). Množinu všech úloh (1), které splňují podmínku  $\text{rank}([V_{12} | V_{13}]) = d$ , označíme  $\mathcal{F}$ . Množinu všech úloh (1), které podmínku  $\text{rank}([V_{12} | V_{13}]) = d$  nesplňují, označíme  $\mathcal{S}$ . Dále

- množinu všech úloh z  $\mathcal{F}$ , pro něž  $\text{rank}(V_{12}) = e$  (a tedy  $\text{rank}(V_{13}) = d - e$ ), označíme  $\mathcal{F}_1$ .
- Množinu všech úloh z  $\mathcal{F}$ , pro něž  $\text{rank}(V_{13}) = d - e$  a zároveň  $\text{rank}(V_{12}) > e$ , označíme  $\mathcal{F}_2$ .
- Množinu všech úloh z  $\mathcal{F}$ , pro něž  $\text{rank}(V_{13}) < d - e$  (a tedy  $\text{rank}(V_{12}) > e$ ), označíme  $\mathcal{F}_3$ .

Množiny  $\mathcal{F}_1$ ,  $\mathcal{F}_2$ ,  $\mathcal{F}_3$  a  $\mathcal{S}$  jsou zřejmě disjunktní a zřejmě platí  $\bigcup_{j=1}^3 \mathcal{F}_j = \mathcal{F}$ . Úlohy mající řešení ve smyslu Definice 1 popsané Větou 1 (případ  $d = 1$ ) a Větou 2 (případy  $p = n$  nebo  $e = d$ ) vždy splňují podmínku  $\text{rank}(V_{12}) = e$ , patří tedy do množiny  $\mathcal{F}_1$ . V [8] bylo ukázáno, že platí následující věta zobecňující toto pozorování.

**Věta 3** Nechť je dána lineární aproximační úloha (1). Uvažujme singulární rozklad (7), se značením zavedeným v (8)–(9). Předpokládejme  $\text{rank}([V_{12} | V_{13}]) = d$ .

Je-li daná úloha z množiny  $\mathcal{F}_1$  (tedy pokud  $\text{rank}(V_{12}) = e$ ), pak úplný problém nejmenších čtverců (2) má řešení. Má-li úloha více než jedno řešení, pak lze zkonstruovat řešení minimální v normě (spektrální i Frobeniově).

Je-li daná úloha z množiny  $\mathcal{F}_2$  (tedy pokud  $\text{rank}(V_{13}) = d - e$  a zároveň  $\text{rank}(V_{12}) > e$ ), pak úplný problém nejmenších čtverců (2) má řešení.

Je-li daná úloha z množiny  $\mathcal{F}_3$  (tedy pokud  $\text{rank}(V_{13}) < d - e$ ), pak úplný problém nejmenších čtverců (2) nemá řešení.

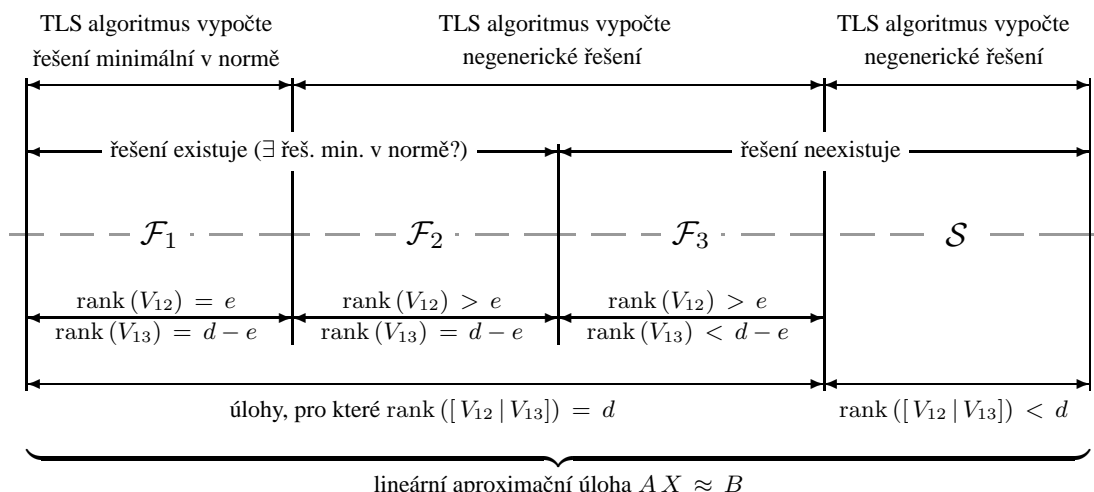
Naznačíme pouze základní myšlenku důkazu, úplný důkaz viz [8]. Lze ukázat, že existence řešení úplného problému nejmenších čtverců je ekvivalentní s existencí ortogonální matice  $W \in \mathbb{R}^{(n-p+e) \times (n-p+e)}$  takové, že  $V_{12}W = [\Delta | \Gamma_1]$ ,  $\Gamma_1 \in \mathbb{R}^{d \times e}$ , kde čtvercová matice  $\Gamma \equiv [\Gamma_1 | V_{13}]$  je *nesingulární*. Pro úlohy z množiny

$\mathcal{F}_1 \cup \mathcal{F}_2$  vždy taková matice  $W$  existuje. Navíc pro úlohy z množiny  $\mathcal{F}_1$  vždy existuje  $W = W_0$  taková, že  $\Delta = 0$ ; jejím užitím konstruujeme řešení minimální v normě, které je jednoznačné (nezávislé na volbě  $W_0$ ). Pro úlohy z množiny  $\mathcal{F}_2$  je  $\Delta \neq 0$  pro libovolné  $W$ ; existence a případná jednoznačnost řešení minimálního v normě zatím *není jasná*. Pro úlohy z množiny  $\mathcal{F}_3$  zřejmě žádná matice  $W$  dávající nesingulární  $\Gamma$  neexistuje.

Oproti tomu řešení úlohy z množiny  $\mathcal{F}$  spočtené TLS algoritmem je jednoznačně určeno libovolnou ortogonální maticí  $\bar{W} \in \mathbb{R}^{(n-p+d) \times (n-p+d)}$  takovou, že  $[V_{12} | V_{13}] \bar{W} = [0 | \bar{\Gamma}]$ , kde čtvercová matice  $\bar{\Gamma}$  je nesingulární. Taková matice  $\bar{W}$  existuje pro libovolnou úlohu z množiny  $\mathcal{F}$ . Pro úlohy z množiny  $\mathcal{F}_1$  TLS al-

goritmus spočte právě řešení problému (2) minimální v normě (např.  $\bar{W} = \text{diag}(W_0, I_{d-e})$ ). Dále lze ukázat, že řešení libovolné úlohy z množiny  $\mathcal{F}_2 \cup \mathcal{F}_3$  vypočtené TLS algoritmem není řešením odpovídajícího problému (2); má charakter negenerického řešení, které TLS algoritmus vrací i pro úlohy z množiny  $\mathcal{S}$ . (Zde je třeba dát pozor na užitou terminologii, např. v [5] se termínu negenerické řešení používá výhradně v kontextu úloh z množiny  $\mathcal{S}$ , v [8] a také zde se o negenerickém řešení hovoří navíc i v kontextu úloh z množiny  $\mathcal{F}_2 \cup \mathcal{F}_3$ ). Negenerické řešení není řešením úlohy (2) ve smyslu Definice 1.

Vzájemný vztah mezi vlastnostmi dané úlohy a existencí řešení problému (2) a řešením spočteným TLS algoritmem shrnuje následující schema.



#### 4. Závěr

Zda je pro úlohy z množiny  $\mathcal{F}_2$  smysluplnější preferovat skutečné řešení problému (2), které vždy existuje, či řešení negenerické vypočtené pomocí TLS algoritmu, není zatím jasné. Obdobně význam samotného negenerického řešení pro  $d > 1$  není příliš zřejmý.

Užitím datové redukce, která zobecňuje pojem core problému pro úlohy s  $d \geq 1$ , viz [8], lze v jistých typických případech (analogických s úlohami s jednou pravou stranou) smysluplnost negenerického řešení interpretovat obdobně jako v případě  $d = 1$ . Zda mezi tyto případy mohou patřit i úlohy z množiny  $\mathcal{F}_2$  zatím není jasné.

Na druhou stranu je v [8] ukázáno, že existuje celá třída problémů, pro něž vykazuje koncept negenerického řešení ne zcela uspokojivé chování. Ilustrujeme tento jev na příkladu.

**Příklad 1** Uvažujme dvě (nezávislé) úlohy (1) s jednou pravou stranou, které jsou dány ve formě core problémů

$$A_{11}^I x_1^I \approx b_1^I, \quad A_{11}^{II} x_1^{II} \approx b_1^{II}$$

a mají tedy dle [6] jednoznačné řešení. Předpokládejme, že

$$\sigma_{\min}([b_1^I | A_{11}^I]) > \sigma_{\max}([b_1^{II} | A_{11}^{II}]).$$

Jinak řečeno, všechna singulární čísla prvního core problému jsou ostře větší než největší singulární číslo druhého core problému. V [8] je ukázáno, že rozšířená matice aproximačního problému

$$\left[ \begin{array}{c|c} A_{11}^I & 0 \\ \hline 0 & A_{11}^{II} \end{array} \right] X \approx \left[ \begin{array}{c|c} b_1^I & 0 \\ \hline 0 & b_1^{II} \end{array} \right], \quad (11)$$

má minimální dimenzi a představuje tak analogii core problému v dané úloze se dvěma pravými stranami. Dále je v [8] ukázáno, že TLS algoritmus aplikován na (11)

vrátí negenerické řešení

$$X = \left[ \begin{array}{c|c} x_1^I & 0 \\ \hline 0 & 0 \end{array} \right],$$

namísto intuitivně očekávaného

$$X = \left[ \begin{array}{c|c} x_1^I & 0 \\ \hline 0 & x_1^{II} \end{array} \right].$$

Zde prezentovaná analýza a dosavadní výsledky naznačují, že zdánlivě elementární formulace úplného problému nejmenších čtverců (2) pro úlohy s více pravými stranami je mnohem komplikovanější a komplexnější než je tomu u úloh s jednou pravou stranou. Vzhledem k tomu, že existuje reálná potřeba řešit aproximační úlohy s více pravými stranami (např. multi-input multi-output dynamické systémy v teorii řízení), bude třeba dalšího studia dané problematiky.

#### Literatura

- [1] G. H. Golub, C. F. Van Loan, “An analysis of the total least squares problem”, *Numer. Anal.* vol. 17, pp. 883–893, 1980.
- [2] I. Hnětynková, Z. Strakoš, “Lanczos tridiagonalization and core problems”, *Linear Algebra Appl.*, vol. 421, pp. 243–251, 2007.
- [3] I. Hnětynková, M. Plešinger, Z. Strakoš, “Lanczos tridiagonalization, Golub-Kahan bidiagonalization and core problem”, *PAMM, Proc. Appl. Math. Mech.* 6, pp. 717–718, 2006.
- [4] S. Van Huffel, “Documented Fortran 77 programs of the extended classical total least squares algorithm, the partial singular value decomposition algorithm and the partial total least squares algorithm”, *Internal. Report ESAT-KUL 88/1*, ESAT Lab., Dept. of Electrical Engrg., Katholieke Universiteit Leuven, 1988.
- [5] S. Van Huffel, J. Vandewalle, “The Total Least Squares Problem, Computational Aspects and Analysis”, *SIAM Publications*, 1991.
- [6] C. C. Paige, Z. Strakoš, “Core problem in linear algebraic systems”, *SIAM J. Matrix Anal. Appl.* vol. 27, pp. 861–875, 2006.
- [7] M. Plešinger, Z. Strakoš, “Core problém v úlohách nejmenších čtverců”, *Sborník konference: Doktorandský den '05, ÚI AV ČR*, pp. 102–108, 2005.
- [8] M. Plešinger, I. Hnětynková, D. M. Sima, Z. Strakoš, S. Van Huffel, “The total least squares problem and reduction of data in  $AX \approx B$ ”, *work in progress*.