Core problem v úlohách nejmenších čtverců

doktorand:
ING. MARTIN PLEŠINGER
Fakulta mechatroniky a mezioborových inženýrských studií TUL
Hálkova 6
461 17 Liberec 1
martin.plesinger@vslib.cz

školitel:
PROF. ING. ZDENĚK STRAKOŠ, DRSC.
Ústav informatiky AV ČR
Pod Vodárenskou věží 2
182 07 Praha 8
strakos@cs.cas.cz

obor studia:
Vědecko-technické výpočty
číselné označení: M6

This work was supported by the National Program of Research "Information Society" under project 1ET400300415.

Abstrakt

Příspěvek stručně pojednává o core problému v úlohách lineárních nejmenších čtverců. Velmi zběžně se podíváme na klasickou úlohu nejmenších čtverců a její řešení. Dále se zaměříme na řešení úplného problému nejmenších čtverců a na komplikace, které mohou při řešení nastat.

Na úplný problém nejmenších čtverců se podíváme z poněkud odlišného úhlu, což povede k přirozené formulaci core problému. Ukážeme souvislost mezi řešením core problému a obvyklým řešením úplného problému nejmenších čtverců.

1. Úvod

V mnoha problémech matematických, fyzikálních i technických potřebujeme řešit soustavy lineárních algebraických rovnic (k popisu soustav budeme používat obvyklý maticový zápis). Je-li taková soustava čtvercová a regulární, máme k dispozici řadu metod k jejímu řešení. Často, například ve statistických aplikacích, se ale setkáváme s obecnějšími soustavami. Matice soustavy může být obecně obdélníková, hovoříme o nedourčených a přeurčených soustavách. Matice, ať už čtvercová nebo obdélníková, nemusí mít plný soupcový a/nebo řádkový rank. Vektor pravé strany obecně může ale nemusí ležet v oboru hodnot matice, může tedy ležet vně podprostoru generovaného sloupci matice.

Uvažujme tedy obecnou soustavu

$$Ax \approx b, \quad A \in \mathcal{R}^{n\times m}, \quad b \in \mathcal{R}^n, \quad \text{přičemž} \quad r \equiv \operatorname{rank}(A) \leq \min\{m,n\}, \quad m \lesseqgtr n \tag{1}$$

a platí buď $b \notin \mathcal{R}(A)$ (*nekompatibilní systém*) nebo $b \in \mathcal{R}(A)$ (*kompatibilní systém*), obvykle předpokládáme $b \neq \emptyset$, v opačném případě je řešení triviální.

2. Klasické řešení, úplný problém nejmenších čtverců

Za klasická bychom např. mohli označit řešení problému (1) pomocí *soustavy normálních rovnic* a pomocí *rozšířené soustavy rovnic*

$$A^TAx = A^Tb, \quad \text{respektive} \quad \begin{bmatrix} I & A \\ A^T & \emptyset \end{bmatrix} \begin{bmatrix} -g \\ x \end{bmatrix} = \begin{bmatrix} b \\ \emptyset \end{bmatrix},$$

kde $g = Ax - b$ je reziduum. Obecnou soustavu (1) jsme převedli na soustavu se čtvercovou, za jistých předpokladů regulární maticí, navíc s vektorem pravé strany ležícím v oboru hodnot matice soustavy. Předpokládáme, že takovou soustavu umíme řešit, viz například [1]. Řešení takto získané je řešením minimalizační úlohy

$$Ax = b + g, \quad \min_{g,x} \|g\|. \tag{2}$$

Minimalizujeme residuum g, vektor $x \perp \mathcal{N}(A)$, který nazýváme *řešení ve smyslu nejmenších čtverců minimální v normě*, je jednoznačně určen. Minimalizační úloha (2) se nazývá *problém nejmenších čtverců – LS*.

Při hledání řešení (2) v podstatě hledáme nejmenší možnou modifikaci vektoru b (pravé strany, vektoru pozorování, odezvy systému, ...). Připouštíme tak, že vektor b obsahuje chyby, ale zcela ignorujeme možnost, že chyby obsahuje i matice A (model). To je jistý nedostatek řešení (1) pomocí LS.

Chyby může obsahovat vektor b a/nebo matice A, popřípadě mohou být chyby obsažené v b a v A v nějakém vztahu. Podrobnější analýza, viz [5], ukazuje, že významná úloha, jejímuž řešení se pokusíme porozumět, je *úplný problém nejmenších čtverců – TLS,*

$$(A + E)x = b + g, \quad \min_{g,E,x} \|[g|E]\|_F. \tag{3}$$

V dalším textu budeme tedy předpokládat, že chyby obsahuje celá soustava, celá matice $[b|A]$.

3. Analýza Goluba a van Loana

Pro jednoduchost budeme v této sekci uvažovat (1) s maticí A, která má plný sloupcový rank, navíc budeme předpokládat $b \notin \mathcal{R}(A)$.

Z analýzy Goluba a van Loana [2] vyplývá, že problém (3) lze za výše uvedených předpokladů a za jisté níže uvedené podmínky (6) přeformulovat a maticově zapsat

$$\left[\begin{array}{c|c} b + g & A + E \end{array}\right] \begin{bmatrix} -1 \\ x \end{bmatrix} = \emptyset. \tag{4}$$

Vidíme, že matice $[b + g|A + E]$ má netriviální nulový prostor (jádro), ve kterém leží vektor s nenulovou první komponentou.

Řešme tedy úlohu (3). Nechť ekonomický singulární rozklad matice $[b|A]$ je

$$[b|A] = U_q \Sigma_q V_q^T = \sum_{i=1}^{q} u_i \sigma_i v_i^T, \quad \text{přičemž} \quad q \equiv \operatorname{rank}([b|A]) \leq \min\{m + 1, n\}.$$

Obecně platí $r \leq q \leq r + 1$, kde r je hodnost matice A, z předpokladů na začátku sekce ovšem vyplývá $r = m < n$, $q = m + 1 \leq n$, tedy $r < q = r + 1$. Minimální perturbace $[g|E]$ matice $[b|A]$ taková, abychom dostali matici s netriviálním jádrem, je

$$[g|E] \equiv -u_q \sigma_q v_q^T, \tag{5}$$

vektor ležící v jádru matice je $v_q = (\nu_1, \ldots, \nu_{m+1})^T = (\nu_1, w^T)^T$. Za předpokladu (6) je $\nu_1 \neq 0$, viz [2], zřejmě platí

$$\begin{bmatrix} -1 \\ x \end{bmatrix} \equiv -\frac{1}{\nu_1} v_q, \quad \text{vektor} \quad x \equiv -\frac{1}{\nu_1} w$$

je řešením minimalizační úlohy (3), nazýváme ho *řešení ve smyslu TLS*, výše popsaný postup je *algoritmem Goluba a van Loana – AGVL*. Toto řešení můžeme také nazývat, zejména v kontextu další sekce, *generick é řešení*.

Obtíže nastanou pokud $\nu_1 = 0$, vektorově $e_1^T v_q = 0$. Minimální petrurbace (5) snižující hodnost matice $[b|A]$ sice lze sestrojit, ale řešení úlohy (3) neexistuje, místo minima existuje pouze infimum. Lze ukázat, že pro některá A, b ztrácí formulace (3) problému TLS zcela smysl, viz příklad v [2, str. 884].

Ona "jistá podmínka", za které lze problém (3) přeformulovat v (4), která zajistí $\nu_1 \neq 0$, je

$$\sigma_r \equiv \sigma_{\min}(A) > \sigma_{\min}([b|A]) \equiv \sigma_q. \tag{6}$$

Je-li (6) splněna, formulace úlohy (3) má vždy smysl a problém lze řešit pomocí AGVL. Bohužel tato podmínka je pouze podmínkou postačující, nikoliv podmínkou nutnou!

4. Negenerické řešení

Sabine Van Huffel a Joos Vandewalle [3] navazují na předchozí analýzu, ale zavádějí kvalitativně odlišné řešení aproximačního problému (1). Toto řešení je v jistém smyslu zobecněním řešení TLS problému (3).

Tentokrát ovšem budeme vycházet ze zápisu (4) (uvažujeme zcela obecný problém (1)). Chceme získat nejmenší petrubaci $[g|E]$ takovou, aby byla splněna rovnost (4), hledáme matici, v jejímž jádru leží vektor s nenulovou první komponentou. Nejmenší taková petrubace $[g|E]$ je zřejmě

$$[g|E] \equiv -u_s \sigma_s v_s^T, \quad v_s = \begin{bmatrix} \mu_1 \\ z \end{bmatrix}, \tag{7}$$

kde σ_s je nejmenší singulární číslo takové, že $\mu_1 \neq 0$. Řešení

$$x \equiv -\frac{1}{\mu_1} z$$

nazýváme *negenerické řešení TLS*, postup nazýváme *rozšířením Van Huffel a Vandewalle – EVHV*.

Negenerické řešní TLS odpovídá řešení původní minimalizační úlohy (3) s rozšiřující podmínkou

$$(A + E)x = b + g, \quad \min_{g,E,x} \|[g|E]\|_F \quad \wedge \quad [g|E][v_{s+1}, \ldots, v_q] = \emptyset. \tag{8}$$

Negenerické řešení vždy existuje (pro libovolná A, b z (1)) a v případě, že je splněna podmínka (6), je identické s řešením nalezeným pomocí AGVL, podrobněji viz [3]. Toto řešení je tedy zobecněním (rozšířením) generického řešení na data nesplňující podmínku (6). Ve skutečnosti však hledáme negenerické řešení minimální v normě, viz [3], zde to pro jednoduchost vynecháváme.

5. Odlišný úhel pohledu

Uvažujme aproximační problém (1) a ortogonální matice P, Q odpovídajících rozměrů. Transformovaný problém

$$\tilde{A}\tilde{x} \equiv \left[P^T A Q\right] \left[Q^T x\right] \approx \left[P^T b\right] \equiv \tilde{b}, \qquad P^{-1} = P^T, \; Q^{-1} = Q^T \tag{9}$$

má, až na transformaci $x = Q\tilde{x}$, stejné generické, resp. negenerické řešení jako problém původní (neboť Frobeniova norma i singulární rozklad jsou ortogonálně invariantní). Předpokládejme, že transformovaný problém má následující strukturu

$$P^T[b|AQ] = [\tilde{b}|\tilde{A}] = \left[\begin{array}{c||c|c} b_1 & A_{11} & \emptyset \\ \hline \emptyset & \emptyset & A_{22} \end{array}\right]. \tag{10}$$

Původní problém $[b|A]$ se rozpadl na dva podproblémy $[b_1|A_{11}]$ a $[\emptyset|A_{22}]$, z nichž druhý má zřejmě jediné smysluplné řešení $x_2 = \emptyset$. Intuice napoví, že řešení původního problému je bezpodmínečně

$$x = Q \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = Q \begin{bmatrix} x_1 \\ \emptyset \end{bmatrix}, \tag{11}$$

kde x_1 je řešení prvního podproblému. Bez újmy na obecnosti, jak uvidíme později, můžeme předpokládat, že první podproblém $A_{11}x_1 \approx b_1$ vyhovuje (6) (je řešitelný pomocí AGVL a x_1 je jeho generickým řešením).

Abychom nahlédli, kde se skrývá obtíž při řešení problému pomocí AGVL, budeme předpokládat

$$\sigma_q = \sigma_{\min}(A_{22}) < \sigma_{\min}([b_1|A_{11}]),$$

intuitivní řešení (11) se tím nijak nezmění. Pokusme se celý problém vyřešit pomocí AGVL (hned si všimněme, že není splněna (ovšem postačující, ne nutná) podmínka (6)). Platí následující rovnost

$$\left[\begin{array}{c||c|c} b_1 & A_{11} & -r_1\theta^{-1}v_q^T \\ \hline \emptyset & \emptyset & A_{22}-u_q\sigma_q v_q^T \end{array}\right]\left[\begin{array}{c} -1 \\ \tilde{x} \end{array}\right] = \emptyset, \quad \text{kde} \quad \tilde{x} = \left[\begin{array}{c} z \\ v_q\theta \end{array}\right] \quad \text{a} \quad x = Q\left[\begin{array}{c} z \\ v_q\theta \end{array}\right], \tag{12}$$

přičemž u_q, v_q jsou levý, resp. pravý singulární vektor odpovídající singulárnímu číslu σ_q ze singulárního rozkladu bloku A_{22}, z je libovolně zvolený vektor a $r_1 \equiv A_{11}z - b_1$, $\theta > 0$ je kladné číslo. Perturbace $[g|E]$ původního systému $[b|A]$ tak je

$$[g|E] = P\left[\begin{array}{c||c|c} \emptyset & \emptyset & -r_1\theta^{-1}v_q^T \\ \hline \emptyset & \emptyset & -u_q\sigma_q v_q^T \end{array}\right]\left[\begin{array}{c|c} 1 & \emptyset \\ \hline \emptyset & Q^T \end{array}\right].$$

Pokud se pokusíme nalézt minimální perturbaci, narazíme na problém. Zřejmě pro $\theta \to +\infty$

$$\|[g|E]\|_F = \sqrt{\|r_1\|^2\theta^{-2} + \sigma_q^T} \quad \longrightarrow \quad \sigma_q = \sigma_{\min}(A_{22})$$

a zároveň

$$\|x\| = \|\tilde{x}\| = \left\|\left[\begin{array}{c} z \\ v_q\theta \end{array}\right]\right\| \quad \longrightarrow \quad +\infty.$$

Minimum tedy neexistuje, existje pouze infimum. Narazili jsme právě na případ, kdy formulace (3) TLS problému zcela postrádá smyslu (ostatně snadno nahlédneme, že pravý singulární vektor odpovídající σ_q v rozkladu matice $[\tilde{b}|\tilde{A}]$ musí mít první složku nulovou). Řešení (3) v tomto případě neexistuje, a tak ani x řešící (12) nemůže být ani vzdálenou aproximací řešení (3). Nedosti na tom, x řešící (12) je, krom libovolně zvolených komponent z a θ, tvořeno pravým singulárním vektorem bloku A_{22}, který jsme v intuitivním přístupu celý zanedbali.

Důležité je následující pozorování, jež nám bude motivací při formulování core problému:

- Pokud se nám původní problém podaří transformovat na (10) a řešení budeme hledat ve tvaru (11), zcela potlačíme vliv dat obsažených v bloku A_{22}. Jinými slovy: z problému odfiltrujeme komponenty související se singulárními čísly $\sigma_i(A_{22})$ bloku A_{22}.

  Tento blok není při intuitivním přístupu vůbec nutný k řešení problému, informace v něm obsažené jsou zcela irelevantní a s řešením problému, tedy i s problémem samotným, vůbec nesouvisí.

- Řešíme-li původní problém pomocí EVHV dospějeme k řešení shodnému s intuitivním řešením (11) (všechny pravé singulární vektory z rozkladu matice $[\tilde{b}|\tilde{A}]$ odpovídající singulárním číslům bloku A_{22} mají zřejmě nulovou první složku). Pomocí EVHV odfiltrujeme část informace obsažené v bloku A_{22}, ovšem odfiltrujeme pouze ta data, která souvisejí se singulárnímy čísly $\sigma_i(A_{22}) < \sigma_s$ (tedy pro $i > s$), viz (8).

6. Core problem

Vhodné by bylo pro řešení (1) využít pouze informace nutné a postačující k řešení. Naší snahou bude odfiltrovat veškerou informaci, která s problémem nesouvisí a na jeho řešení nemá vliv. V dalším textu budeme předpokládat $A^Tb \neq 0$, tj. vektor pravé strany není ortogonální na podprostor generovaný sloupci matice A (v opačném případě existuje pouze negenerické řešení a je triviální $x = \emptyset$, přičemž $[g|E] = [-b|\emptyset]$; předpoklad přirozeně zahrnuje i případ, kdy b je identicky nulový vektor).

Budeme hledat ortogonální matice P, $P^{-1} = P^T$, Q, $Q^{-1} = Q^T$ transformující původní problém na problém se strukturou (10), navíc budeme požadovat aby blok A_{22} měl maximální možnou dimenzi (a blok $[b_1|A_{11}]$ minimální).

Definice 1 *Podproblém $A_{11}x_1 \approx b_1$ z rozkladu (10) nazveme* **core problem** *v aproximačním problému $[b|A]$ v případě, že $[b_1|A_{11}]$ má minimální dimenzi (a A_{22} maximální dimenzi).*

Pokusíme se core problém nalézt. Nechť

$$A = U\Sigma V^T = U \left[\begin{array}{c|c} \Sigma_r & \emptyset \\ \hline \emptyset & \emptyset \end{array}\right] V^T = \sum_{i=1}^{r} u_i \sigma_i v_i^T, \quad \Sigma_r = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$$

je singulární rozklad matice A. Zaveďme pomocné značení $U = [U_1|U_2]$, kde $U_1 = (u_1, \ldots, u_r)$, $U_2 = (u_{r+1}, \ldots, u_n)$, a analogicky $V = [V_1|V_2]$, kde $V_1 = (v_1, \ldots, v_r)$, $V_2 = (v_{r+1}, \ldots, v_m)$. Platí

$$U^T[b|AV] = \left[\begin{array}{c||c|c} c & \Sigma_r & \emptyset \\ \hline d & \emptyset & \emptyset \end{array}\right]. \tag{13}$$

Vektory c, d mohou obsahovat nulové prvky, pomocí ortogonálních transformací se budeme snažit problém upravit tak, aby tyto vektory obsahovaly maximum nulových prvků (tím nalezneme core problém). Nejprve ortogonální maticí H_{22} (Householderova reflexe) modifikujeme vektor d tak, že $H_{22}^T d = e_1\delta$, $\delta = \|d\|$. Podmatici U_2 v rozkladu (13) nahradíme maticí U_2H_{22}.

Nyní budeme upravovat vektor $c = (\gamma_1, \ldots, \gamma_r)$. Uvažujme $\sigma_i = \ldots = \sigma_j$ singulární čísla matice A, jinak řečeno, σ_i je singulární číslo s násobností $j - i + 1$, přičemž $j \geq i$. Pomocí ortogonální matice H_{ij} (Householderova reflexe) transformujeme jim odpovídající podvektor $c_{ij} = (\gamma_i, \ldots, \gamma_j)$ tak, že $H_{ij}^T c_{ij} = e_1\gamma_{ij}$, $\gamma_{ij} = \|c_{ij}\|$. Označme H_{11} ortogonální, blokově diagonální matici, která má na diagonále matice H_{ij} (signulárním číslům s násobností jedna tedy odpovídají jednotky na diagonále). Touto transformací získáme na místě vektoru c vektor s maximálním možným počtem nulových prvků. Nalezneme permutaci P_{11}^T řádků matice $H_{11}^T[c|\Sigma_r H_{11}]$ tak, že (lidově řečeno) řádky s nenulovou první složkou seřadíme pod sebe a všechny řádky začínající nulou odsuneme dolů. V rozkladu (13) nahradíme podmatici U_1 maticí $U_1H_{11}P_{11}$ a podmatici V_1 maticí $V_1H_{11}P_{11}$.

V případě, že $\delta \neq 0$, provedeme ještě permutaci řádků P_0^T tak, že tento řádek zařadíme přímo pod ostatními řádky s nenulovou první složkou, zřejmě $\delta \neq 0$ tehdy a jen tehdy, je-li původní problém nekompatibilní, tedy když $b \notin \mathcal{R}(A)$. Získáme tak rozklad

$$P^T[b|AQ] \equiv \left(P_0^T [U_1H_{11}P_{11}|U_2H_{22}]^T \right) \left[b|A \left([V_1H_{11}P_{11}|V_2] \right) \right] = \left[\begin{array}{c||c|c} \tilde{c} & \Sigma_1 & \emptyset \\ \delta & \emptyset & \emptyset \\ \hline \emptyset & \emptyset & \Sigma_2 \end{array}\right], \tag{14}$$

kde Σ_1 obsahuje pouze vzájemně různá signulární čísla matice A (ovšem obecně ne všechna), Σ_2 obsahuje všechna ostatní singulární čísla matice A, všechna opakování singulárních čísel a navíc (pro přehlednost zápisu) obsahuje nulová singulární čísla. Tím jsme problém transformovali na blokovou strukturu (10).

Věta 1 *Existuje taková ortogonální transformace (9), tedy takové ortogonální matice P, Q, že blok A_{11}, resp. A_{22} z rozkladu (10) má minimální, resp. maximální možnou dimenzi přes všechny ortogonální transformace vedoucí na danou strukturu. Blok A_{11} nemá žádné opakující se a ani nulové singulární číslo a blok A_{22} obsahuje všechna opakování singulárních čísel (redundance), všechna nerelevantní data (singulární čísla jímž odpovídající levé singulární podprostory jsou ortogonální na vektor b) a všechna nulová singulární čísla.*

Podproblém $A_{11}x_1 \approx b_1$ je core problém a je vždy řešitelný pomocí AGVL a podproblém $A_{22}x_2 \approx \emptyset$ má jediné smysluplné řešení $x_2 = \emptyset$. Řešení původního problému je

$$x = Q \begin{bmatrix} x_1 \\ \emptyset \end{bmatrix}. \tag{15}$$

Důkaz věty 1 byl z části proveden v přechozím textu, podrobněji viz [4].

Věta 2 *Matice A_{11} z věty 1 je čtvercová matice $\mathcal{R}^{p\times p}$, resp. obdélníková matice $\mathcal{R}^{p+1\times p}$ tehdy a jen tehdy, když vektor b má právě p nenulových projekcí do levých singulárních podprostorů odpovídajících různým singulárním číslům matice A a zároveň $b \in \mathcal{R}(A)$, resp. $b \notin \mathcal{R}(A)$.*

Věta 3 *Podproblém $A_{11}x_1 \approx b_1$ z rozkladu (10) je kompatibilní tehdy a jen tehdy, když je kompatibilní celý problém $Ax \approx b$, tedy*

$$b \in \mathcal{R}(A) \quad \Longleftrightarrow \quad b_1 \in \mathcal{R}(A_{11}).$$

Jinými slovy: v kompatibilním případě je matice $A_{11} \in \mathcal{R}^{p\times p}$ z věty 1 čtvercová a regulární.

Důkazy obou vět 2, 3 vyplývají z textu, z konstrukce rozkladu (14), podrobněji opět viz [4].

7. Nalezení core problému

Ortogonální matice P, Q z věty 1 dokážeme nalézt přímo, transformací matice $[b|A]$ na horní bidiagonální tvar (Householderovy reflexe, nebo Lanczos-Golub-Kahanova bidiagonalizace). Důkaz využívající singulární rozklady jednotlivých bloků rozkladu (10) nalezneme v článku [4].

S výhodou využijeme faktu, že bidiagonalizaci provádíme postupně. Díky tomu můžeme proces zastavit v momentě, když dojde k oddělení obou podproblémů, jinak řečeno: blok A_{22} nemusí být bidagonalizován. K oddělení obou podproblémů dojde:

- V kompatibilním případě $b \in \mathcal{R}(A)$

$$[b_1|A_{11}] = \left[\begin{array}{c|cccc} \beta_1 & \alpha_1 & & & \\ & \beta_2 & \alpha_2 & & \\ & & \ddots & \ddots & \\ & & & \beta_p & \alpha_p \end{array}\right] \in \mathcal{R}^{p\times p+1}, \quad \alpha_i\beta_i \neq 0, \quad i = 1 \ldots p.$$

  pokud $\beta_{p+1} = 0$ nebo $p = n$. Matice A_{11} je čtvercová dimenze p, nesingulární. Core problém $A_{11}x_1 = b_1$ je kompatibilní.

- V nekompatibilním případě $b \notin \mathcal{R}(A)$

$$[b_1|A_{11}] = \left[\begin{array}{c|cccc} \beta_1 & \alpha_1 & & & \\ & \beta_2 & \alpha_2 & & \\ & & \ddots & \ddots & \\ & & & \beta_p & \alpha_p \\ & & & & \beta_{p+1} \end{array}\right] \in \mathcal{R}^{p+1\times p+1}, \quad \begin{array}{l} \alpha_i\beta_i \neq 0, \quad i = 1 \ldots p, \\ \beta_{p+1} \neq 0. \end{array}$$

  pokud $\alpha_{p+1} = 0$ nebo $p = m$. Matice $[b_1|A_{11}]$ je čtvercová dimenze $p + 1$, nesingulární. Core problém $A_{11}x_1 \approx b_1$ je nekompatibilní.

V obou případech má matice A_{11} plný sloupcový rank a matice $[b_1|A_{11}]$ má plný řádkový rank.

8. Shrnutí a závěr

Řešení původního problému získané technikou core problému je obecně shodné s negenerickým řešením (s řešením pomocí EVHV), za předpokladu (6) je shodné s řešením generickým (pomocí AGVL). Výhodou přístupu pomocí core problému je využití pouze nutných a postačujících informací k řešení. Oprávněnost řešení (15) core problému (10), narozdíl od Golubova, van Loanova řešení podmíněného (6) a rozšířeného na negenerické řešení (8), je snadno nahlédnutelná a význam řešení (15) zcela odpovídá naší intuici. Celkový vhled do problematiky prostřednictvím core problému je jasný a zřejmý, čehož není možné dosháhnout prostřednictvím AGVL a EVHV.

Celý postup přitom není složitější než přístup pomocí AGVL a EVHV, ba právě naopak. Abychom zjistili, zda je splněna podmínka (6), potřebujeme spočítat obě singulární čísla $\sigma_{\min}(A)$ a $\sigma_{\min}([b|A])$. Teprve potom můžeme rozhodnout zda hledat generické nebo negenerické řešení. K nalezení core problému potřebujeme provést bidiagonalizaci matice $[b|A]$, ovšem neúplnou, nebot' blok A_{22} nemusí být bidiagonalizován. Core problém pak lze bezpodmínečně řešit pomocí AGVL. K porovnání složitostí obou přístupů si stačí uvědomit, že bidiagonalizace je první (finitní) částí algoritmu pro výpočet singulárního rozkladu.

Kolem problematiky lineárních nejmenších čtverců řešených prostřednictvím core problému je řada důležitých, otevřených otázek:

- Jak vytvořit analogickou teorii pro úlohy s vícenásobnou pravou stranou?
- Při řešení ill-posed problémů provádíme regularizaci, hledáme minimum

  $$\min\left(\|Ax-b\|+\|Lx\|\right),$$

  kde L je matice určená pro daný problém. V případě, že matice L kombinuje řešení obou podproblémů $A_{11}x_1 \approx b_1$ a $A_{22}x_2 \approx \emptyset$, nemůžeme položit $x_2 = \emptyset$.
- Celá teorie byla odvozena v přesné aritmetice. Jak se ovšem bude core problém chovat v aritmetice s konečnou přesností? Jaká je citlivost a jaké je numerické chování core problému?
- Neposledním důležitým úkolem je numericky stabilní implementace software pro řešení úloh (1) pomocí core problému.

Literatura

[1] G. Dahlquist, Å Björck, “Numerical Mathematics and Scientific Computation”, Vol. 2, Chap. 8, Linear Least Squares Problem, Web Draft, 2005.

[2] G. H. Golub, C. F. van Loan, “An Analysis of The Total Least Squares Problem”, *Numer. Anal.*, vol 17, pp. 883–893, 1980.

[3] S. Van Huffel, J. Vandewalle, “The Total Least Squares Problem: Computational Aspects and Analysis”, SIAM Publications, Philadelphia PA, 1991.

[4] C. C. Paige, Z. Strakoš, “Core Problems in Linear Algebraic Systems”, Accepted for publications in *Matrix Anal.*, 2005.

[5] C. C. Paige, Z. Strakoš, “Scaled Total Least Squares Fundamentals”, *Numerische Mathematik*, vol. 91, pp. 117–146, 2002.