Verified solutions of saddle point linear systems

Shinya Miyajima

Faculty of Engineering, Gifu University

16th GAMM - IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics

September 22, 2014, miyajima@gifu-u.ac.jp

Saddle point linear systems and purpose

$$\mathcal{H}u = b, \quad \mathcal{H} := \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix}, u := \begin{pmatrix} x \\ y \end{pmatrix}, b := \begin{pmatrix} f \\ g \end{pmatrix},$$

$A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{m\times m}$, $f \in \mathbb{R}^n$, $g \in \mathbb{R}^m$: given,

$u \in \mathbb{R}^{n+m}$: unknown, $n \geq m$, A: SPD, B: full rank, C: SPSD

Purpose Computing upper bounds for $\|\tilde{u} - u^*\|_\infty$ without using $\mathrm{fl}(\mathcal{H}^{-1})$, where $u^* = ({x^*}^T, {y^*}^T)^T$ and $\tilde{u} = (\tilde{x}^T, \tilde{y}^T)^T$ denote the exact and numerical solutions, respectively.

Preferable: smaller bound, fast algorithm

Notations

$$r_u = \begin{pmatrix} r_f \\ r_g \end{pmatrix} := \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} - \begin{pmatrix} f \\ g \end{pmatrix}$$

$A = L_A L_A^T$, $BB^T = L_B L_B^T$: Cholesky decomposition

$L_A^{-1} B^T = QR$: thin QR factorization

$$\zeta := \frac{\|A\|_2 \|(BB^T)^{-1}\|_2}{1 + \|A\|_2 \|(BB^T)^{-1}\|_2 \sigma_{\min}(C)},\ \eta := \frac{\|R^{-1}\|_2^2}{1 + \|R^{-1}\|_2^2 \sigma_{\min}(C)}$$

(we can prove $\zeta \geq \eta$)

$s^{(n)} := (1, \dots, 1)^T \in \mathbb{R}^n$

Previous and our error estimations

Chen-Hashimoto (2003)

$$\|\tilde{x}-x^*\|_2 \le \|A^{-1}\|_2((1+\zeta\|B^T\|_2\|BA^{-1}\|_2)\|r_f\|_2+\zeta\|B^T\|_2\|r_g\|_2),$$

$$\|\tilde{y}-y^*\|_2 \le \zeta(\|BA^{-1}\|_2\|r_f\|_2+\|r_g\|_2).$$

Our

$$\|\tilde{x}-x^*\|_2 \le \|L_A^{-1}\|_2\|L_A^{-1}r_f\|_2+\eta\|A^{-1}B^T\|_2\|r_g\|_2,$$

$$\|\tilde{y}-y^*\|_2 \le \eta(\|BA^{-1}r_f\|_2+\|r_g\|_2).$$

We can prove our $\le$ Chen-Hashimoto.

Previous and our error estimations ($C = 0$) (1/2)

Chen-Hashimoto (2003)

$$\|\tilde{x} - x^*\|_2 \leq \|A^{-1}\|_2((1 + \|A\|_2\|B^T\|_2\|(BB^T)^{-1}\|_2\|BA^{-1}\|_2)\|r_f\|_2 + \|A\|_2\|B^T\|_2\|(BB^T)^{-1}\|_2\|r_g\|_2).$$

Hashimoto (2007)

$$\|\tilde{x} - x^*\|_2 \leq \|A^{-1}\|_2\|r_f\|_2 + \|A\|_2\|A^{-1}\|_2\|L_B^{-1} r_g\|_2.$$

Our $\|\tilde{x} - x^*\|_2 \leq \|L_A^{-1}\|_2(\|L_A^{-1} r_f\|_2 + \|R^{-T} r_g\|_2)$.

We can prove our $\leq$ Hashimoto $\leq$ Chen-Hashimoto.

Previous and our error estimations ($C = 0$) (2/2)

Chen-Hashimoto (2003)

$$\|\tilde{y} - y^*\|_2 \leq \|A\|_2 \|(BB^T)^{-1}\|_2 (\|BA^{-1}\|_2 \|r_f\|_2 + \|r_g\|_2).$$

Hashimoto (2007)

$$\|\tilde{y} - y^*\|_2 \leq \|L_B^{-1}\|_2 (\|A\|_2 \|A^{-1}\|_2 \|r_f\|_2 + \|A\|_2 \|L_B^{-1} r_g\|_2).$$

Kimura-Chen (2009)

$$\|\tilde{u} - u^*\|_2 \leq 2 \max(\|A^{-1}\|_2, \|A\|_\infty \|(BB^T)^{-1}\|_2) \|r_u\|_2 / (\sqrt{5} - 1).$$

Our $\|\tilde{y} - y^*\|_2 \leq \|R^{-T}\|_2 \|L_A^{-1} r_f\|_2 + \|R^{-1} R^{-T} r_g\|_2.$

We can prove our $\leq$ Hashimoto $\leq$ Chen-Hashimoto.

Proposed algorithm

The algorithm for computing rigorous upper bounds of our error estimations when $C = 0$.

Computational cost

$$\begin{aligned}
\text{Chen-Hashimoto:}\quad & 44m^3/3 + 4m^2n + 4mn^2 + 5n^3 + \mathcal{O}(n^2) \\
\text{Kimura-Chen:}\quad & 44m^3/3 + 4m^2n + 5n^3 + \mathcal{O}(n^2) \\
\text{Our:}\quad & -m^3/3 + 6m^2n + 2mn^2 + 2n^3/3 + \mathcal{O}(n^2)
\end{aligned}$$

We can show our $<$ Kimura-Chen $<$ Chen-Hashimoto.

Preliminaries

$$\mathcal{H}^{-1} = \begin{pmatrix} A^{-1} - A^{-1}B^TS^{-1}BA^{-1} & A^{-1}B^TS^{-1} \\ S^{-1}BA^{-1} & -S^{-1} \end{pmatrix},$$

where $S := BA^{-1}B^T + C = BL_A^{-T}L_A^{-1}B^T + C = R^TR + C$. Hence

$$\begin{pmatrix} \tilde{x} - x^* \\ \tilde{y} - y^* \end{pmatrix} = \begin{pmatrix} (A^{-1} - A^{-1}B^TS^{-1}BA^{-1})r_f + A^{-1}B^TS^{-1}r_g \\ S^{-1}BA^{-1}r_f - S^{-1}r_g \end{pmatrix},$$

so we estimate $\|(A^{-1} - A^{-1}B^TS^{-1}BA^{-1})r_f\|_2$, $\|A^{-1}B^TS^{-1}r_g\|_2$, $\|S^{-1}BA^{-1}r_f\|_2$ and $\|S^{-1}r_g\|_2$.

$S^{-1} = R^{-1}(I_m + R^{-T}CR^{-1})^{-1}R^{-T}$. $S^{-1} = R^{-1}R^{-T}$ if $C = 0$.

Derivation ($C = 0, S^{-1} = R^{-1}R^{-T}$) (1/2)

$$\|(A^{-1} - A^{-1}B^TS^{-1}BA^{-1})r_f\|_2$$

$$= \|(L_A^{-T}L_A^{-1} - L_A^{-T}L_A^{-1}B^TR^{-1}R^{-T}BL_A^{-T}L_A^{-1})r_f\|_2$$

$$= \|(L_A^{-T}L_A^{-1} - L_A^{-T}QRR^{-1}R^{-T}R^TQ^TL_A^{-1})r_f\|_2$$

$$= \|L_A^{-T}(I_n - QQ^T)L_A^{-1}r_f\|_2 \leq \|L_A^{-T}\|_2\|I_n - QQ^T\|_2\|L_A^{-1}r_f\|_2$$

$$= \|L_A^{-T}\|_2\|L_A^{-1}r_f\|_2.$$

$$\|S^{-1}r_g\|_2 = \|R^{-1}R^{-T}r_g\|_2.$$

Derivation ($C = 0, S^{-1} = R^{-1}R^{-T}$) (2/2)

$$\|A^{-1}B^TS^{-1}r_g\|_2 = \|L_A^{-T}L_A^{-1}B^TR^{-1}R^{-T}r_g\|_2$$

$$= \|L_A^{-T}QRR^{-1}R^{-T}r_g\|_2 = \|L_A^{-T}QR^{-T}r_g\|_2$$

$$\leq \|L_A^{-T}\|_2\|Q\|_2\|R^{-T}r_g\|_2 = \|L_A^{-T}\|_2\|R^{-T}r_g\|_2.$$

$$\|S^{-1}BA^{-1}r_f\|_2 = \|R^{-1}R^{-T}BL_A^{-T}L_A^{-1}r_f\|_2$$

$$= \|R^{-1}R^{-T}R^TQ^TL_A^{-1}r_f\|_2 = \|R^{-1}Q^TL_A^{-1}r_f\|_2$$

$$\leq \|R^{-1}\|_2\|Q^T\|_2\|L_A^{-1}r_f\|_2 = \|R^{-1}\|_2\|L_A^{-1}r_f\|_2.$$

Derivation (general C) (1/5)

$$(A^{-1} - A^{-1}B^TS^{-1}BA^{-1})r_f$$

$$= (L_A^{-T}L_A^{-1} - L_A^{-T}QRR^{-1}(I_m + R^{-T}CR^{-1})^{-1}R^{-T}R^TQ^TL_A^{-1})r_f$$

$$= L_A^{-T}(I_n - Q(I_m + R^{-T}CR^{-1})^{-1}Q^T)L_A^{-1}r_f$$

($R^{-T}CR^{-1} = V\Sigma V^T$: SVD ($\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_m)$))

$$= L_A^{-T}(I_n - Q(I_m + V\Sigma V^T)^{-1}Q^T)L_A^{-1}r_f$$

$$= L_A^{-T}(I_n - QV(I_m + \Sigma)^{-1}V^TQ^T)L_A^{-1}r_f$$

Derivation (general C) (2/5)

$$= L_A^{-T}\left(I_n - QV\mathrm{diag}\left(\frac{1}{1+\sigma_1},\dots,\frac{1}{1+\sigma_m}\right)V^TQ^T\right)L_A^{-1}r_f$$

($\exists U = (QV, U_\perp)$: orthogonal, $U_\perp$: orthogonal complement of QV)

$$\left(QV(I_m+\Sigma)^{-1}V^TQ^T = U\mathrm{diag}\left(\frac{1}{1+\sigma_1},\dots,\frac{1}{1+\sigma_m},\underbrace{0,\dots,0}_{n-m}\right)U^T\right)$$

$$= L_A^{-T}\left(I_n - U\mathrm{diag}\left(\frac{1}{1+\sigma_1},\dots,\frac{1}{1+\sigma_m},\underbrace{0,\dots,0}_{n-m}\right)U^T\right)L_A^{-1}r_f$$

Derivation (general C) (3/5)

$$= L_A^{-T} U \operatorname{diag}\left(\frac{\sigma_1}{1+\sigma_1}, \ldots, \frac{\sigma_m}{1+\sigma_m}, \underbrace{1, \ldots, 1}_{n-m}\right) U^T L_A^{-1} r_f,$$

so $\|(A^{-1} - A^{-1}B^T S^{-1} B A^{-1}) r_f\|_2$

$$\leq \|L_A^{-T}\|_2 \|U\|_2 \left\|\operatorname{diag}\left(\frac{\sigma_1}{1+\sigma_1}, \ldots, \frac{\sigma_m}{1+\sigma_m}, \underbrace{1, \ldots, 1}_{n-m}\right)\right\|_2 \|U^T\|_2 \|L_A^{-1} r_f\|_2$$

$$= \|L_A^{-T}\|_2 \|L_A^{-1} r_f\|_2.$$

Derivation (general C) (4/5)

$$\sigma_{\min}(S) = \min_{v\in\mathbb{R}^m, v\neq 0} \frac{v^T(BA^{-1}B^T + C)v}{v^Tv}$$

$$\geq \min_{v\in\mathbb{R}^m, v\neq 0} \frac{v^TBA^{-1}B^Tv}{v^Tv} + \min_{v\in\mathbb{R}^m, v\neq 0} \frac{v^TCv}{v^Tv}$$

$$= \sigma_{\min}(BA^{-1}B^T) + \sigma_{\min}(C) = \sigma_{\min}(R^TR) + \sigma_{\min}(C)$$

$$= \sigma_{\min}(R)^2 + \sigma_{\min}(C) = \frac{1}{\|R^{-1}\|_2^2} + \sigma_{\min}(C) = \frac{1}{\eta},$$

so that $\|S^{-1}\|_2 = 1/\sigma_{\min}(S) \leq \eta$.

Derivation (general C) (5/5)

$$\|A^{-1}B^TS^{-1}r_g\|_2 \le \|A^{-1}B^T\|_2\|S^{-1}\|_2\|r_g\|_2 \le \eta\|A^{-1}B^T\|_2\|r_g\|_2,$$

$$\|S^{-1}BA^{-1}r_f\|_2 \le \|S^{-1}\|_2\|BA^{-1}r_f\|_2 \le \eta\|BA^{-1}r_f\|_2,$$

$$\|S^{-1}r_g\|_2 \le \|S^{-1}\|_2\|r_g\|_2 \le \eta\|r_g\|_2.$$

Numerical results ($C = 0$)

Intel Xeon 2.66GHz Dual CPU, 4.00GB RAM, MATLAB 7.5 with Intel MKL, and IEEE 754 double precision

- `CH`: the Chen-Hashimoto algorithm
- `I`: the INTLAB function `verifylss` ($\mathrm{fl}(\mathcal{H}^{-1})$ is used)
- `KC`: the Kimura-Chen algorithm
- `M`: the proposed algorithm
- `Mi`: `M` with the iterative refinement ($\tilde{x}$ and $\tilde{y}$ are stored by 2 terms)

We obtained b, i.e. f and g, such that $b = \mathrm{fl}(\mathcal{H}s^{(m+n)})$.

We show the upper bound for $\|\tilde{u} - u^*\|_\infty$ to assess the qualities.

Example 1 ($n = 2p^2$, $m = p^2$)

$$A = \begin{pmatrix} I_p \otimes T + T \otimes I_p & 0 \\ 0 & I_p \otimes T + T \otimes I_p \end{pmatrix} \in \mathbb{R}^{2p^2 \times 2p^2},$$

$$B = \begin{pmatrix} I_p \otimes F \\ F \otimes I_p \end{pmatrix}^T \in \mathbb{R}^{p^2 \times 2p^2},$$

$$T := \mathrm{tridiag}(-1, 2, -1)/h^2 \in \mathbb{R}^{p \times p},$$

$$F := \mathrm{tridiag}(-1, 1, 0)/h \in \mathbb{R}^{p \times p}, \;\; h := 1/(p+1).$$

The upper bounds and computing times (sec)

p	m	n	CH	I	KC	M	Mi
16	256	512	5.2e−8	2.3e−16	9.5e−9	8.6e−12	6.7e−28
24	576	1152	7.7e−7	2.3e−16	6.5e−8	2.9e−11	2.1e−27
32	1024	2048	6.4e−6	2.3e−16	3.1e−7	8.9e−11	6.9e−27
40	1600	3200	2.4e−5	MO	7.3e−7	1.3e−10	MO
16	256	512	0.37	0.83	0.33	0.26	0.79
24	576	1152	3.41	6.90	3.02	2.00	5.25
32	1024	2048	18.7	34.7	16.3	10.3	29.3
40	1600	3200	63.8	MO	56.1	33.4	MO

Example 2

$$A_{ij} = \begin{cases} i+1, & i=j, \\ 1, & |i-j|=1, \quad i,j=1,\ldots,n, \\ 0, & \text{others}, \end{cases}$$

$$B_{ij} = \begin{cases} i, & j=i+n-m, \\ 0, & \text{others}, \end{cases} \quad i=1,\ldots,m,\ j=1,\ldots,n.$$

The upper bounds and computing times (sec)

m	n	CH	I	KC	M	Mi
10	1000	4.1e–7	2.3e–16	8.0e–9	6.4e–12	5.7e–28
10	3000	7.1e–6	2.3e–16	1.4e–7	3.6e–11	3.1e–27
100	1000	4.5e–5	2.3e–16	8.8e–9	6.5e–12	5.4e–28
100	3000	7.4e–4	2.3e–16	1.5e–7	3.7e–11	3.1e–27
10	1000	2.58	2.83	2.49	0.75	1.81
10	3000	41.4	41.4	40.5	11.9	22.3
100	1000	2.67	3.57	2.54	0.88	2.04
100	3000	41.7	45.0	40.9	12.6	23.4

Example 3 ($n = 2p^2$, $m = p^2$)

$$A = \begin{pmatrix} A_{1,1} & 0 \\ 0 & A_{2,2} \end{pmatrix} \in \mathbb{R}^{2p^2 \times 2p^2},\ B = \begin{pmatrix} A_{1,3} \\ A_{2,3} \end{pmatrix}^T \in \mathbb{R}^{p^2 \times 2p^2},$$

$$A_{1,1} := \mathrm{Tridiag}(-\beta I_p, \beta T_p, -\beta I_p) \in \mathbb{R}^{p^2 \times p^2},$$

$$A_{2,2} := \mathrm{Tridiag}(-\beta I_p, \beta T_p, -\beta I_p) \in \mathbb{R}^{p^2 \times p^2},$$

$$A_{1,3} := \mathrm{Diag}(S_p) \in \mathbb{R}^{p^2 \times p^2},\ A_{2,3} := \frac{h}{2}\mathrm{Tridiag}(I_p, 0, -I_p) \in \mathbb{R}^{p^2 \times p^2},$$

$$T_p := \mathrm{tridiag}(-\beta, 4\beta, -\beta) \in \mathbb{R}^{p \times p},$$

$$S_p := \frac{h}{2}\mathrm{tridiag}(1, 0, -1) \in \mathbb{R}^{p \times p},\ \ h := 1/(p+1).$$

The upper bounds and computing times (sec)

p	m	n	CH	I	KC	M	Mi
16	256	512	7.8e–9	2.3e–16	5.8e–9	2.9e–11	6.9e–27
24	576	1152	1.2e–7	2.3e–16	4.1e–8	1.8e–10	4.3e–26
32	1024	2048	8.6e–7	2.3e–16	1.7e–7	4.2e–10	1.1e–25
40	1600	3200	4.2e–6	MO	5.2e–7	1.2e–9	MO
16	256	512	0.40	0.78	0.33	0.26	0.79
24	576	1152	3.44	6.63	2.98	2.03	5.26
32	1024	2048	18.6	33.9	16.1	10.3	28.9
40	1600	3200	63.5	MO	56.2	33.4	MO