Verified computations for all generalized singular values

Shinya Miyajima

Faculty of Engineering, Gifu University

15th GAMM - IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics

September 27, 2012

miyajima@gifu-u.ac.jp

Singular value decomposition (SVD)

Theorem 1 [e.g. Golub & Van Loan (1996)] Let $A \in \mathbb{R}^{m\times n}$. There exist orthogonal $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ s.t.

$$U^T AV = \Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_q), \quad q := \min(m, n),$$
$$\sigma_1 \geq \cdots \geq \sigma_{r^*} > \sigma_{r^*+1} = \cdots = \sigma_q = 0, \quad r^* = \mathrm{rank}(A).$$

$A = U\Sigma V^T$: SVD of A, $\sigma_1, \ldots, \sigma_q$: Singular value (SV) of A

Applications: Pseudoinverse, Condition number, Spectral norm, etc.

Generalized SVD (GSVD)

Theorem 2 [Van Loan (1976)] Let $A \in \mathbb{R}^{m\times n}$ with $m \geq n$ and $B \in \mathbb{R}^{p\times n}$. There exist orthogonal $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{p\times p}$ and nonsingular $X \in \mathbb{R}^{n\times n}$ s.t.

$$U^T AX = \Sigma_A = \mathrm{diag}(c_1, \ldots, c_n),\ V^T BX = \Sigma_B = \mathrm{diag}(s_1, \ldots, s_q),$$

$$q = \min(p, n),\ \ 1 \geq s_1 \geq \cdots \geq s_{r^*} > s_{r^*+1} = \cdots = s_q = 0,$$

$$r^* = \mathrm{rank}(B),\ 0 \leq c_1 \leq \cdots \leq c_n \leq 1,\ \ c_i^2 + s_i^2 = 1, \quad i = 1, \ldots, q.$$

$\mu_j = c_j/s_j$, $j = 1, \ldots, r^*$: generalized SV (GSV) of A and B

Applications: damped least squares, weighted least squares, etc.

Purpose

Numerically computing verified bounds of all SV or GSV

$\Rightarrow$ Numerically computing $\underline{\varepsilon}_i$ and $\overline{\varepsilon}_i$ satisfying $\underline{\varepsilon}_i \leq \sigma_i \leq \overline{\varepsilon}_i$ or $\underline{\varepsilon}_i \leq \mu_i \leq \overline{\varepsilon}_i$ for all i

Preferable

- Radii $(\overline{\varepsilon}_i - \underline{\varepsilon_i})/2$ is small
- Fast algorithm

Previous work for SV

- Oishi (2001)
  - full SVD of A is numerically computed
- Rump (2011) -1st algorithm
  - full SVD of A is numerically computed
- Rump (2011) -2nd algorithm
  - eigen-decomposition (ED) of A^TA is numerically computed
  - faster than the above when $m \geq n$

Our contribution (1/3)

4 algorithms for computing verified bounds of all SV

- 1st algorithm
  - economy SVD of A is numerically computed
  - smaller radii than Oishi (2001) (shown theoretically)
  - faster than Oishi (2001)
- 2nd algorithm
  - economy SVD of A is numerically computed
  - smaller radii than Rump (2011) -1st (shown theoretically)
  - faster than Rump (2011) -1st

Our contribution (2/3)

- 3rd algorithm
  - economy SVD of A is numerically computed
  - faster than the proposed 1st and 2nd
- 4th algorithm
  - ED of A^TA or AA^T is numerically computed
  - smaller radii than Rump (2011) -2nd when $m \geq n$ (shown theoretically)
  - fastest of all

Our contribution (3/3)

8 algorithms for computing for verified bounds of all GSV

- 1st - 4th algorithms
  - Extension of the 4 algorithms for SV
  - Applicable when $B^T B$ is nonsingular
- 5th – 8th algorithms
  - Extension of the 4 algorithms for SV
  - Applicable when $A^T A$ is nonsingular

Oishi's bound

$A \approx U\Sigma V^T$: full SVD ($V^TV \approx I_n$, $U^TU \approx I_m$)

Theorem 3 [Oishi (2001)] Let $q := \min(m, n)$,
$E := U\Sigma V^T - A$, $F := U^TU - I_m$, $G := V^TV - I_n$.
If $\|F\|_2 < 1$ and $\|G\|_2 < 1$, $\Sigma_{ii} - \delta_i \le \sigma_i(A) \le \Sigma_{ii} + \delta_i$,
$i = 1, \ldots, q$, where $\delta_i := \Sigma_{ii} \max(\|F\|_2, \|G\|_2) + \|E\|_2$.

Proposed bound for SV -1st

$A \approx \hat{U}\hat{\Sigma}\hat{V}^T$: economy SVD ($\hat{U}$, $\hat{\Sigma}$, $\hat{V}$: submatrix)

Theorem 4 Let $\hat{E} := \hat{U}\hat{\Sigma}\hat{V}^T - A$, $\hat{F} := \hat{U}^T\hat{U} - I_q$, $\hat{G} := \hat{V}^T\hat{V} - I_q$. If $\|\hat{F}\|_2 < 1$ and $\|\hat{G}\|_2 < 1$, $\underline{\delta}_i^M \leq \sigma_i(A) \leq \overline{\delta}_i^M$,

$$\underline{\delta}_i^M := \sqrt{(1 - \|\hat{F}\|_2)(1 - \|\hat{G}\|_2)}\hat{\Sigma}_{ii} - \|\hat{E}\|_2,$$

$$\overline{\delta}_i^M := \sqrt{(1 + \|\hat{F}\|_2)(1 + \|\hat{G}\|_2)}\hat{\Sigma}_{ii} + \|\hat{E}\|_2.$$

Theorem 5 $\Sigma_{ii} - \delta_i \leq \underline{\delta}_i^M$ and $\overline{\delta}_i^M \leq \Sigma_{ii} + \delta_i$ hold.

Rump's bound -1st

$A \approx U\Sigma V^T$: full SVD

Theorem 6 [Rump (2011)] Let $q := \min(m, n)$, $F := U^T U - I_m$, $G := V^T V - I_n$, $U^T AV =: D + E$, D be diagonal. If $\|F\|_2 < 1$ and $\|G\|_2 < 1$, there is a numbering $\nu : \{1, \dots, q\} \to \{1, \dots, q\}$ with $\underline{\varepsilon}_i \le \sigma_{\nu(i)}(A) \le \overline{\varepsilon}_i$, where

$$\underline{\varepsilon}_i := \frac{|D_{ii}| - \|E\|_2}{\sqrt{(1 + \|F\|_2)(1 + \|G\|_2)}}, \quad \overline{\varepsilon}_i := \frac{|D_{ii}| + \|E\|_2}{\sqrt{(1 - \|F\|_2)(1 - \|G\|_2)}}.$$

Proposed bound for SV -2nd

$A \approx \hat{U}\hat{\Sigma}\hat{V}^T$: economy SVD

Theorem 7 Let $\hat{F} := \hat{U}^T\hat{U} - I_q$, $\hat{G} := \hat{V}^T\hat{V} - I_q$, $\hat{U}^T A \hat{V} =: \hat{D} + \hat{E}$, $\hat{D}$ be diagonal. If $\|\hat{F}\|_2 < 1$ and $\|\hat{G}\|_2 < 1$, $\underline{\varepsilon}_i^M \le \sigma_{\nu(i)}(A) \le \overline{\varepsilon}_i^M$, where

$$\underline{\varepsilon}_i^M := \frac{|\hat{D}_{ii}| - \|\hat{E}\|_2}{\sqrt{(1 + \|\hat{F}\|_2)(1 + \|\hat{G}\|_2)}}, \quad \overline{\varepsilon}_i^M := \frac{|\hat{D}_{ii}| + \|\hat{E}\|_2}{\sqrt{(1 - \|\hat{F}\|_2)(1 - \|\hat{G}\|_2)}}.$$

Theorem 8 $\underline{\varepsilon}_i \le \underline{\varepsilon}_i^M$ and $\overline{\varepsilon}_i^M \le \overline{\varepsilon}_i$ hold.

Proposed bound for SV -3rd (1/2)

$A \approx \hat{U}\hat{\Sigma}\hat{V}^T$: economy SVD

Theorem 9 Let $\hat{F} := \hat{U}^T\hat{U} - I_q$, $\hat{G} := \hat{V}^T\hat{V} - I_q$,

$$\underline{\Sigma} := \begin{cases} (\sqrt{1-\|\hat{F}\|_2}/\sqrt{1+\|\hat{G}\|_2})\hat{\Sigma} & (\text{if } m \geq n) \\ (\sqrt{1-\|\hat{G}\|_2}/\sqrt{1+\|\hat{F}\|_2})\hat{\Sigma} & (\text{if } m < n) \end{cases}$$

$$\overline{\Sigma} := \begin{cases} (\sqrt{1+\|\hat{F}\|_2}/\sqrt{1-\|\hat{G}\|_2})\hat{\Sigma} & (\text{if } m \geq n) \\ (\sqrt{1+\|\hat{G}\|_2}/\sqrt{1-\|\hat{F}\|_2})\hat{\Sigma} & (\text{if } m < n) \end{cases}$$

Proposed bound for SV -3rd (2/2)

$$\rho := \begin{cases} \|A\hat{V} - \hat{U}\hat{\Sigma}\|_2 / \sqrt{1 - \|\hat{G}\|_2} & (\text{if } m \geq n) \\ \|\hat{U}^T A - \hat{\Sigma}\hat{V}^T\|_2 / \sqrt{1 - \|\hat{F}\|_2} & (\text{if } m < n) \end{cases}.$$

If $\|\hat{F}\|_2 < 1$ and $\|\hat{G}\|_2 < 1$, $\underline{\Sigma}_{ii} - \rho \leq \sigma_i(A) \leq \overline{\Sigma}_{ii} + \rho$.

Rump's bound -2nd

$A^TA \approx V\Lambda V^T$: numerical ED ($V^TV \approx I_n$, $(AV)^TAV \approx \Lambda$)

Theorem 10 [Rump (2011)] Let $q := \min(m,n)$, $G := I_n - V^TV$, $(AV)^TAV =: D + E$, D be diagonal. If $\|G\|_2 < 1$, there is a numbering $\nu : \{1,\ldots,q\} \to \{1,\ldots,q\}$ with $\underline{\zeta}_i \le \sigma_{\nu(i)}(A) \le \overline{\zeta}_i$, where

$$\underline{\zeta}_i := \sqrt{\frac{D_{ii} - \|E\|_2}{1 + \|G\|_2}}, \quad \overline{\zeta}_i := \sqrt{\frac{D_{ii} + \|E\|_2}{1 - \|G\|_2}}.$$

Proposed bound for SV -4th (1/2)

requires one of $A^TA \approx V_1\Lambda_1V_1^T$, $AA^T \approx V_2\Lambda_2V_2^T$

Theorem 11 Let $D+E := \begin{cases} (AV)^TAV & (\text{if } m \geq n) \\ V^TAA^TV & (\text{if } m < n) \end{cases}$,
D be diagonal, $q := \min(m,n)$, $G := I_n - V^TV$, $\hat{f} := |E|s$,
$f_i := \begin{cases} \hat{f}_i & (\text{if } < D_{ii}, \hat{f}_i > \text{ is isolated from the other}) \\ \|E\|_\infty & (\text{otherwise}) \end{cases}$.

Proposed bound for SV -4th (2/2)

If $\|G\|_2 < 1$, $\underline{\zeta}_i^M \le \sigma_{\nu(i)}(A) \le \overline{\zeta}_i^M$, where

$$\underline{\zeta}_i^M := \begin{cases} \sqrt{\dfrac{D_{ii} - f_i}{1 + \|G\|_2}} & (\text{if } D_{ii} > f_i) \\ 0 & (\text{otherwise}) \end{cases}, \quad \overline{\zeta}_i^M := \sqrt{\dfrac{D_{ii} + f_i}{1 - \|G\|_2}}.$$

Theorem 12 $\sqrt{\dfrac{D_{ii} - \|E\|_\infty}{1 + \|G\|_2}} \le \underline{\zeta}_i^M, \overline{\zeta}_i^M \le \sqrt{\dfrac{D_{ii} + \|E\|_\infty}{1 - \|G\|_2}}.$

Proposed 8 bounds for GSV

If B^TB is nonsingular, $\exists LL^T = B^TB$ (Cholesky).

$\sigma_i(A,B)^2 = \lambda_i(A^TA, B^TB) = \lambda_i(L^{-1}A^TAL^{-T})$

$= \lambda_i((AL^{-T})^TAL^{-T}) = \sigma_i(AL^{-T})^2$, i.e. $\sigma_i(A,B) = \sigma_i(AL^{-T})$.

$\Rightarrow$ Apply 4 bounds for SV to AL^{-T}

Let $r^* := \mathrm{rank}(B)$. If A^TA is nonsingular, $\exists LL^T = A^TA$.

$\sigma_i(A,B)^2 = 1/\lambda_{r^*-i+1}(B^TB, A^TA) = 1/\lambda_{r^*-i+1}(L^{-1}B^TBL^{-T})$

$= 1/\sigma_{r^*-i+1}(BL^{-T})^2$, i.e. $\sigma_i(A,B) = 1/\sigma_{r^*-i+1}(BL^{-T})$.

$\Rightarrow$ Apply 4 bounds for SV to BL^{-T} and take reciprocals

Numerical examples for SV

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

`svd` is used for numerical SVD. `eig` is used for numerical ED.

$$\text{the radius} := \frac{\text{the upper bound} - \text{the lower bound}}{2}$$

M1 – M4:	The proposed 1st – 4th bounds
O:	Oishi's bound
R1 and R2 :	Rump's 1st and 2nd bounds
V:	VERSOFT function `versingval`

Example 1 [computing times (sec)]

```
A = randn(m,n);
```

m	n	M1	M2	M3	M4	O	R1	R2	V
1000	300	0.7469	0.7657	0.7132	0.4450	2.0967	2.1410	0.4438	21228
3000	300	1.4742	1.4923	1.3803	0.9191	28.986	29.162	0.9212	MO
10000	300	4.1867	4.2379	3.8856	2.5624	MO	MO	2.5846	MO
1000	1000	16.927	17.537	16.507	7.7877	16.929	17.604	7.6970	22996
3000	3000	465.48	484.82	458.83	193.69	467.38	485.96	193.88	MO
300	1000	0.7690	0.7912	0.6999	0.4381	2.1021	2.1554	4.8422	21260
300	3000	1.6009	1.6287	1.3783	0.9170	29.184	29.222	112.20	MO
300	10000	5.6943	5.7541	3.8787	2.5856	MO	MO	MO	MO

MO: memory over

Example 2 [obtained maximum radii]

```
A = gallery('randsvd',[1000,10],cnd);
```

cnd	M1	M2	M3	M4	O	R1	R2	V
1e+0	3.3e–14	7.6e–14	3.2e–14	2.9e–14	3.0e–13	3.5e–13	2.9e–14	4.4e–14
1e+4	5.2e–14	4.4e–14	4.7e–14	1.5e–12	3.3e–13	2.4e–13	2.0e–10	2.0e–14
1e+8	3.2e–14	3.7e–14	3.1e–14	1.2e–8	3.0e–13	2.1e–13	fail1	2.2e–14
1e+12	2.9e–14	3.4e–14	3.0e–14	1.0e–7	3.0e–13	2.0e–13	fail1	2.0e–14
1e+16	3.1e–14	3.8e–14	3.2e–14	1.1e–7	3.3e–13	2.2e–13	fail1	NaN

fail1: $\sqrt{\dfrac{D_{ii} - \|E\|_2}{1 + \|G\|_2}}$ could not be verified to be real for some i

Numerical examples for GSV

`svd` is used for numerical SVD. `eig` is used for numerical GED.

M1A – M4A: Extension of the proposed 1st – 4th bounds for nonsingular B^TB

M1B – M4B: Extension of the proposed 1st – 4th bounds for nonsingular A^TA

Example 1 [computing times (sec)]

`A = randn(m,n); B = randn(p,n);`

m	n	p	M1A	M2A	M3A	M4A	M1B	M2B	M3B	M4B
3000	300	1000	1.70	1.69	1.58	1.04	1.05	1.03	1.01	0.75
1000	300	3000	1.05	1.02	0.97	0.69	1.70	1.69	1.56	1.04
3000	300	3000	1.83	1.83	1.71	1.17	1.86	1.83	1.72	1.17
10000	300	3000	4.78	4.87	4.45	2.82	2.25	2.25	2.13	1.60
3000	300	10000	2.26	2.26	2.13	1.60	4.73	4.77	4.45	2.85
10000	300	10000	5.20	5.27	4.87	3.27	5.20	5.21	4.89	3.35
1000	1000	1000	18.2	18.2	17.2	9.73	18.1	18.1	17.1	9.74
3000	3000	3000	482	492	466	245	481	493	467	245
3000	1000	300	–	–	–	–	3.96	3.33	3.84	8.74
3000	3000	300	–	–	–	–	33.5	25.2	33.1	188
10000	3000	300	–	–	–	–	72.8	64.4	72.2	220

Example 2 [obtained maximum radii]

```
A = gallery('randsvd',[1000,10],cndA);

B = randn(1000,10);
```

cndA	M1A	M2A	M3A	M4A	M1B	M2B	M3B	M4B
1e+0	1.6e–14	5.2e–15	4.0e–15	3.9e–15	1.4e–14	7.0e–15	3.7e–15	6.3e–15
1e+2	7.2e–15	3.7e–15	3.6e–15	3.3e–15	2.7e–9	4.3e–12	4.2e–12	8.4e–12
1e+4	6.2e–15	3.5e–15	3.6e–15	3.7e–14	1.2e–3	9.7e–9	9.7e–9	2.0e–8
1e+6	6.1e–15	3.8e–15	3.6e–15	7.1e–12	7.6e–5	2.1e–4	2.1e–4	3.8e–4
1e+8	5.6e–15	3.6e–15	3.4e–15	2.9e–10	fail2	fail2	fail2	fail2

fail2: nonsingularity of $A^T A$ could not be verified

Example 3 [obtained maximum radii]

```
A = randn(1000,10);

B = gallery('randsvd',[1000,10],cndB);
```

cndB	M1A	M2A	M3A	M4A	M1B	M2B	M3B	M4B
1e+0	1.2e–11	5.1e–12	3.9e–12	3.8e–12	1.9e–11	7.6e–12	4.3e–12	7.1e–12
1e+2	2.6e–6	3.2e–7	3.2e–7	3.2e–7	4.2e–8	7.6e–9	2.7e–9	2.6e–8
1e+4	1.6e+0	1.5e–1	1.5e–1	1.5e–1	3.1e–4	7.0e–5	3.0e–5	1.8e–2
1e+6	9.6e+5	1.3e+5	1.3e+5	1.3e+5	1.9e+0	6.7e–1	1.7e–1	1.5e+4
1e+8	fail3	fail3	fail3	fail3	7.9e+4	7.2e+3	1.7e+3	2.6e+7

fail3: nonsingularity of B^TB could not be verified