Computing accurate eigenvalues and eigenvectors of symmetric quasi-Cauchy matrices

Froilán M. Dopico

Departamento de Matemáticas

Universidad Carlos III de Madrid
Spain

Setting the Problem

Demmel, SIMAX (1999) Vol. 21, pp. 562-580, gave an algorithm to compute **SVD** of quasi-Cauchy matrices fast and with high relative accuracy:

$$A = D_1 C D_2 \quad D_1, D_2 \quad \text{diagonal matrices}$$

and C is a Cauchy matrix:

$$C_{ij} = \frac{1}{x_i + y_j} \quad i, j = 1 : n,$$

where x_i and y_j are given floating point numbers.

Our Goal:

Compute

EIGENVALUES and EIGENVECTORS

of REAL SYMMETRIC ($D_1 = D_2$, $x_i = y_i$)

quasi-Cauchy matrices with

HIGH RELATIVE ACCURACY.

Overview of Demmel's work (1)

The following **previous result** by Demmel, Gu, Eisenstat, Slapničar, Veselić, Drmač (LAA, 1999) is used:

Let A be a matrix with **rank revealing decomposition (RRD)** $A = XDY^T$*, where D is a nonsingular diagonal matrix and X, Y are well conditioned matrices.*

1. **Let us assume** *that a RRD of A, $\widehat{X}, \widehat{D}, \widehat{Y}$, is computed such that*

$$\|X - \widehat{X}\| = O(\epsilon)\|X\| \quad , \quad \|Y - \widehat{Y}\| = O(\epsilon)\|Y\|$$

$$|D_{ii} - \widehat{D}_{ii}| = O(\epsilon)|D_{ii}| \quad \text{and} \quad \widehat{D} \quad \text{diagonal.}$$

2. *Let us apply the Jacobi-type Alg. 3.1 of Demmel et al. to $\widehat{X}\widehat{D}\widehat{Y}^T$ to obtain the computed SVD of A, $\widehat{U}, \widehat{\Sigma}, \widehat{V}$.*

Then $\boldsymbol{|\sigma_i - \widehat{\sigma}_i| = O(\epsilon \max(\kappa(X), \kappa(Y)))\sigma_i,}$

$$\boldsymbol{\theta(u_i, \widehat{u}_i) \text{ or } \theta(v_i, \widehat{v}_i) = \frac{O(\epsilon \max(\kappa(X), \kappa(Y)))}{relgap_i}}$$

where $relgap_i = \min\left(\min_{j\neq i} \frac{|\sigma_i - \sigma_j|}{\sigma_i}\,,\, 2\right).$

Overview of Demmel's work (2)

Gaussian elimination with complete pivoting (GECP) is used to compute an accurate RRD, $\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{Y}^{\mathbf{T}}$, but **not in the usual way**.

Notice that after having done k steps of GE on A

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ A_{21}U_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} U_{11} & L_{11}^{-1}A_{12} \\ 0 & S^{(k)} \end{bmatrix}$$

where $A_{11} = L_{11}U_{11}$ and the **k-th Schur Complement**

$$S^{(k)} = A_{22} - A_{21}A_{11}^{-1}A_{12} \in \mathbb{R}^{(n-k)\times(n-k)}.$$

- $D_{k+1,k+1} = S_{11}^{(k)}$.
- $X(k+1:n, k+1) = S^{(k)}(:,1)/S_{11}^{(k)}$.
- $Y^T(k+1, k+1:n) = S^{(k)}(1,:)/S_{11}^{(k)}$.
- The pivoting strategy can be introduced in this scheme.

In this case:

$$S_{rs}^{(k)} = S_{rs}^{(k-1)} \frac{(x_r - x_k)(y_s - y_k)}{(x_k + y_s)(x_r + y_k)}.$$

The elements of all the Schur complements are computed with relative accuracy $O(\epsilon)$. Then, the RRD is computed with the required accuracy and cost $\frac{4}{3}n^3$ or $\frac{8}{3}n^3$ flops.

Computing eigenvalues and eigenvectors

Consider the **SYMMETRIC** real case

$$A = DCD \quad \text{where} \quad D \quad \text{diagonal matrix}$$

and C is a symmetric Cauchy matrix:

$$C_{ij} = \frac{1}{x_i + x_j} \quad i, j = 1 : n.$$

First Method: Signed SVD.
(Dopico, Molera, Moro (2000)).

1. Compute SVD, $A = U\Sigma V^T$, using Demmel's algorithm.
2. $$\lambda_i = (v_i^T u_i)\sigma_i,$$ $$q_i = v_i.$$

This method may be generalized when tight clusters of singular values are present.

Symmetric method for eigenvalues and eigenvectors

Problem: GECP on a symmetric matrix does not take advantage of symmetry ($A = XDY^T$ with $X \neq Y$).

Question: *Can we develop a symmetric algorithm with half operation cost for the factorization?*

Idea:

1. Can we compute an accurate enough *symmetric indefinite decomposition* (SID)?

$$A = GJG^T \quad , \quad J = I_p \oplus -I_{r-p}.$$

2. **If yes**, we can apply the implicit one-sided **J-orthogonal Jacobi method** of Veselić and Slapničar (1992) to obtain eigenvalues and eigenvectors with high relative accuracy.

$$G\, J_0 J_1 \dots J_{M-1} = G_M \quad \text{where}$$

$$G_M^T G_M \approx \Delta \quad \text{diag. positive} \quad \text{and} \quad J_k^T J J_k = J.$$

The eigenvalue and eigenvector matrices are

$$\boldsymbol{J\Delta} \quad \textbf{and} \quad \boldsymbol{G_M \Delta^{-1/2}}.$$

Accurate $A = GJG^T$ (1)

We only describe the **first stage** of the computation using Bunch and Parlett's pivoting strategy:

$\mu_0 = \max_{ij} |a_{ij}|$, $\mu_1 = \max_i |a_{ii}|$, $\alpha \approx 0.64$

```
if μ1 ≥ αμ0
      1 × 1 pivot s.t. |e11| = μ1.
else
      2 × 2 pivot s.t. |e21| = μ0.
end
```

where E is the chosen pivot

$$\Pi A \Pi^T = \begin{bmatrix} E & C^T \\ C & B \end{bmatrix}, \qquad E \;\; s \times s \;\; \text{and} \;\; s = 1, 2.$$

The first step of the *diagonal pivoting method* is

$$\Pi A \Pi^T = \begin{bmatrix} I_s & 0 \\ CE^{-1} & I \end{bmatrix} \begin{bmatrix} E & 0 \\ 0 & S^{(s)} \end{bmatrix} \begin{bmatrix} I_s & E^{-1}C^T \\ 0 & I \end{bmatrix},$$

with the Schur complement

$$S^{(s)} = B - CE^{-1}C^T.$$

Accurate $A = GJG^T$ (2)

$$\Pi A \Pi^T = \begin{bmatrix} I_s & 0 \\ CE^{-1} & I \end{bmatrix} \begin{bmatrix} E & 0 \\ 0 & S^{(s)} \end{bmatrix} \begin{bmatrix} I_s & E^{-1}C^T \\ 0 & I \end{bmatrix}.$$

No problem in the 1×1 pivot case, but in the 2×2 case to obtain the symmetric indefinite decomposition more work has to be done

$$E = Q_s \Lambda_s Q_s^T = (Q_s \sqrt{|\Lambda_s|}) \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} (\sqrt{|\Lambda_s|} Q_s^T).$$

The same scheme is repeated on the Schur complement $S^{(2)}$, and son on...

We are able again to compute the elements of all the Schur complements with relative accuracy $O(\epsilon)$.

- In 1×1-case just as in Demmel's algorithm.
- In 2×2-case

$$S_{rs}^{(k+1)} = S_{rs}^{(k-1)} \frac{(x_r - x_k)(x_s - x_k)}{(x_k + x_s)(x_r + x_k)} \frac{(x_r - x_{k+1})(x_s - x_{k+1})}{(x_{k+1} + x_s)(x_r + x_{k+1})}.$$

Rounding errors in $A = GJG^T$

Using

- Previous result on errors of Schur complements.
- 2×2 pivots fulfill $\kappa(E) \leq 4.6$.
- $|(CE^{-1})_{ij}| \leq 2.78$.

the following theorem can be proven:

Theorem: *Let $\widehat{G}, \widehat{J}$ be the computed factors when the previous algorithm for computing a SID is applied to the symmetric quasi-Cauchy matrix A in floating point arithmetic.* **If the same sequence of pivots (in size and positions)** *is applied in* **exact** *arithmetic, a SID of A is obtained:*

$$A = GJG^T, \qquad \textit{such that}$$

1. $\boldsymbol{\widehat{J} = J}$.
2. $\boldsymbol{\|\widehat{G}(:,j) - G(:,j)\| = O(\epsilon)\|G(:,j)\| \ \ \forall j}$.
3. $\|\widehat{G}D_G - GD_G\| = O(\epsilon)\|GD_G\| \ \ \forall$ *diag. nonsing.* D_G.
4. $\widehat{G}J\widehat{G}^T = (I+E)A(I+E)^T$

$$\text{where} \qquad \|E\|_2 = O(\epsilon) \min_{D_G} \kappa(GD_G).$$

Overall error in $A = GJG^T$ + J-Jacobi

Slapničar (1992) proves that if **implicit J-orthogonal Jacobi** method is applied in floating point arithmetic to the SID $\widehat{G}J\widehat{G}^T$, **the exact eigenvalues and eigenvectors** of

$$(\widehat{G}+\delta\widehat{G})J(\widehat{G}+\delta\widehat{G})^T = (I+\delta\widehat{G}\widehat{G}^\dagger)\widehat{G}J\widehat{G}^T(I+\delta\widehat{G}\widehat{G}^\dagger)^T$$

are obtained, where

$$\|\delta\widehat{G}\widehat{G}^\dagger\|_2 = O(\epsilon)\max_{0\le k\le (M-1)} \kappa_2(\widehat{G}_k\widehat{D}_{\widehat{G}_k}).$$

Taking into account $\widehat{G}J\widehat{G}^T = (I+E)A(I+E)^T$, we get

$$(\widehat{G}+\delta\widehat{G})J(\widehat{G}+\delta\widehat{G})^T = (I+F)A(I+F)^T,$$

such that

$$\|F\|_2 = O(\epsilon\kappa_2(GC_G)) \quad \text{with} \quad \|(GC_G)(:,j)\|_2 = 1\ \forall j.$$

Well-known results of *multiplicative perturbation theory* (Ipsen, Eisenstat, Li...) yield

$$|\widehat{\boldsymbol{\lambda}}_i - \boldsymbol{\lambda}_i| = \boldsymbol{O(\epsilon\kappa_2(GC_G))}|\boldsymbol{\lambda}_i|,$$

$$\boldsymbol{\theta(q_i,\widehat{q}_i)} = \frac{\boldsymbol{O(\epsilon\kappa_2(GC_G))}}{\mathbf{min}_{j\neq i}\frac{|\boldsymbol{\lambda}_j-\boldsymbol{\lambda}_i|}{|\boldsymbol{\lambda}_i|}}.$$

Numerical Experiment I

Consider the 100×100 Cauchy matrix

$$C_{ij} = \frac{1}{x_i + x_j} \quad , \quad x_i = (-1)^{(i-1)} + 2^{-40} * (i-1),$$

with $\kappa_2(C) = 7.8 * 10^{73}$. We compute spectral decomposition using **Mathematica with 120-decimal digit of precision** and compare with our algorithms implemented in MATLAB ($\epsilon \approx 1.1 * 10^{-16}$).

METHOD	$\max_i \frac{\lvert\lambda_i - \widehat{\lambda}_i\rvert}{\lvert\lambda_i\rvert}$	$\max_i \lVert v_i - \widehat{v}_i \rVert_2$
Signed SVD	$2.8 * 10^{-15}$	$1.7 * 10^{-14}$
J-Orthog.	$1.1 * 10^{-14}$	$1.7 * 10^{-14}$
MATLAB (eig)	$1.1 * 10^{55}$	1.41
Slapničar	$9.7 * 10^{54}$	1.41

Other interesting data:

- $1.9 * 10^{-62} \leq |\lambda| \leq 1.6 * 10^{12}$.
- $\min_i relgap_i(\lambda) = 0.62$; $\min_i relgap_i(\sigma) = 0$!!!.
- #(rel. error $\lambda > 1$) = 64. (Slapničar and MATLAB).
- $\max(\kappa_2(X), \kappa_2(Y)) = 38.7$ (GECP).
- $\kappa_2(GD_G) = 30.5$ (SID).

Numerical Experiment II (1)

- 1400 random symm. quasi-Cauchy matrices of 7 different types and dimensions $10 < n < 100$.
- Use $C_{ij} = 1/(x_i + x_j) \Rightarrow$

$$C^{-1} = D'CD' \quad \text{with} \quad D'_{ii} = \frac{\prod_k (x_k + x_i)}{\prod_{k \neq i}(x_i - x_k)}.$$

- Compare with MATLAB: $\max_i \frac{|\lambda_i - \lambda_i^M|}{\epsilon ||A||_2}$

Signed SVD = 101	J-Jacobi = 264

- The same for the inverse ($\lambda_{inv} = 1/\lambda$):

Signed SVD = 58	J-Jacobi = 313

- Compare with the same algorithm for the inverse:

$$\max_i \frac{|\lambda_i - (\lambda_i^{(-1)})^{-1}|}{\epsilon |\lambda_i|}$$

Signed SVD = 244	J-Jacobi = 932

- Compare with MATLAB:
$\max_i \|v_i - v_i^M\| absgap_i(\sigma \text{ or } \lambda)/\epsilon$

Signed SVD = 47	J-Jacobi = 85

- The same for the inverse:

Signed SVD = 280	J-Jacobi = 370

Numerical Experiment II (2)

- Compare with the same algorithm for the inverse:

$$\max_i \|v_i - v_i^{(-1)}\| relgap_i(\sigma \text{ or } \lambda)/\epsilon$$

Signed SVD = 321	J-Jacobi = 373

Other interesting data:

- $\max \max_i \|v_i^S - v_i^J\| = 3.5 * 10^{-11}$.
- $\max \max_i \frac{|\lambda_i^S - \lambda_i^J|}{|\lambda_i^J|} = 5 * 10^{-14}$.
- $\min \min_i relgap_i(\sigma) = 0$.
- $\min \min_i relgap_i(\lambda) = 6.5 * 10^{-4}$.
- Maximum relative error in eigenvalues

MATLAB = $1.5 * 10^{87}$	Slapničar = $4.4 * 10^{69}$

- Number of matrices with rel. error > 1 in some eigenvalue

MATLAB = 907	Slapničar = 475

- $5.5 * 10^2 < \kappa_2 < 9.9 * 10^{150}$.
- # Jacobi sweeps.

METHOD	MAX	MEAN	MIN
Signed SVD	8	5.3	2
J-Orthog.	10	6.9	3