

An orthogonal high relative accuracy algorithm for the symmetric eigenproblem

Froilán M. Dopico

**Departamento de Matemáticas
Universidad Carlos III de Madrid
Spain**

*Joint work with Juan M. Molera and Julio Moro from
Universidad Carlos III*

Summary on high relative accuracy (HRA) (I)

Standard algorithms (QR, Divide and Conquer,...) for the SYMMETRIC eigenproblem or for SVD are normwise backward stable. Therefore eigenvalues (sing. values) and eigenvectors (sing. vectors) are computed with RELATIVE errors:

$$\frac{|\lambda_i - \hat{\lambda}_i|}{|\lambda_i|} = \frac{O(\epsilon)\|A\|}{|\lambda_i|} \quad \theta(v_i, \hat{v}_i) = \frac{O(\epsilon)\|A\|}{\min_{j \neq i} |\lambda_i - \lambda_j|}$$

This implies LARGE relative errors for tiny eigenvalues

$$\frac{\|A\|}{\lambda_i} \gtrsim \frac{1}{\epsilon}$$

Notation: ϵ machine precision. Hats denote computed quantities. $\|A\|$ spectral norm.

In 1990 Demmel and Kahan showed that things may be different for SVD of BIDIAGONAL matrices using zero shift bidiagonal QR

$$\frac{|\sigma_i - \hat{\sigma}_i|}{\sigma_i} = O(\epsilon)$$

$$\theta(u_i, \hat{u}_i) \text{ and } \theta(v_i, \hat{v}_i) = \frac{O(\epsilon)}{\min_{j \neq i} \frac{|\sigma_i - \sigma_j|}{\sigma_i}}$$

Also relevant **dqds** by Fernando and Parlett (1994).

Summary on high relative accuracy (HRA) (II)

Afterwards Demmel and Kahan's several works on HRA **spectral** computations for SPECIAL CLASSES of matrices.

SVD and Symmetric Positive Definite Matrices:

- Graded Matrices: Demmel, Drmač, Mathias, Veselić...
- Acyclic(Bidiagonal): Demmel, Gragg, Kahan, Parlett...
- Scaled-Cauchy, Vandermonde, Poly. Vandermonde, Diag. Dominant M-matrices: Demmel, Koev...
- Others...

Symmetric Indefinite Matrices:

- Certain well-scalable matrices: Slapničar, Veselić, J-orthogonal (Hyperbolic) algorithm.
- Scaled diagonally dominant: Barlow, Demmel.
- Any symmetric matrix allowing to compute a high accuracy SVD: Dopico, Molera, Moro. This talk.

Nonsymmetric Matrices: Nothing yet.

REMARK: Before the algorithm presented in this talk
HRA SVD computations much more developed than
HRA eigendecompositions of indefinite matrices.

Unified approach

Many classes of matrices and many algorithms.

Demmel, Gu, Eisenstat, Slapničar, Veselić, Drmač (1999):
propose a unified ALGORITHMIC and THEORETICAL
approach for all **possible (so far) HRA SVD** computations:

Algorithm to compute SVD of general A .

STEP 1: Compute a **rank revealing decomposition (RRD)**

$A = XDY^T$, where D is a nonsingular diagonal matrix
and X, Y are well conditioned matrices.

STEP 2: Apply a Jacobi-type algorithm of Demmel et al. to
the computed RRD to compute the SVD of $A = U\Sigma V^T$.

**IF THE RRD IS COMPUTED (GECF) WITH SMALL
FORWARD ERRORS THEN THE SVD IS COMPUTED
WITH HRA.**

For **SYMMETRIC INDEFINITE** eigencomputations a
similar approach is possible using

- Symmetric RRDs, $A = XDX^T$, for **STEP 1**.
- J-orthogonal Hyperbolic Jacobi on XDX^T for **STEP 2**.

PROBLEMS: Little is known about **SYMMETRIC RRDs**.
Hyperbolic rotations.

Multiplicative backward errors

Theorem (Demmel et al.): *If the computed RRD satisfies*

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

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

Then $(I + E)A(I + F) = \hat{U}\hat{\Sigma}\hat{V}^T$

with $\|E\| = O(\epsilon\kappa(X))$ *and* $\|F\| = O(\epsilon\kappa(Y))$.

And $|\sigma_i - \hat{\sigma}_i| = O(\epsilon \max(\kappa(X), \kappa(Y)))\sigma_i$,

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

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

Our goal

GIVEN A COMPUTED SVD OF A SYM. INDEFINITE MATRIX WITH SMALL MULTIPLICATIVE BACKWARD ERROR, WE WANT TO COMPUTE ITS EIGENVALUES AND EIGENVECTORS WITH HRA.

Improvements:

Many classes of matrices.

Orthogonal rotations.

Spectral Decompositions from SVDs (I)

EASY CASE

$A = A^T = U\Sigma V^T \in \mathbb{R}^{n \times n}$ with $\sigma_1 > \dots > \sigma_n$:

$\lambda_i = (v_i^T u_i)\sigma_i$ and v_i eigenvectors for $i = 1 : n$.

EQUAL SINGULAR VALUES CASE

$A = A^T = U\Sigma V^T \in \mathbb{R}^{n \times n}$ with $\sigma_1 > \dots > \sigma_p$ and

$$\Sigma = \text{diag}[\sigma_1 I_{m_1}, \dots, \sigma_p I_{m_p}]$$

$$m_1 + \dots + m_p = n$$

$$U = [\mathcal{U}_1 \mid \mathcal{U}_2 \mid \dots \mid \mathcal{U}_p]$$

$$V = [\mathcal{V}_1 \mid \mathcal{V}_2 \mid \dots \mid \mathcal{V}_p]$$

with $\mathcal{U}_i, \mathcal{V}_i \in \mathbb{R}^{n \times m_i}$ and $\mathcal{V}_i^T \mathcal{U}_j = 0$ for $i \neq j$.

Then

$$V^T U = \text{diag}[\mathcal{V}_1^T \mathcal{U}_1, \dots, \mathcal{V}_p^T \mathcal{U}_p]$$

Each diagonal block $\mathcal{V}_i^T \mathcal{U}_i \in \mathbb{R}^{m_i \times m_i}$ is orthogonal.

$$V^T A V = \text{diag}[\sigma_1 \mathcal{V}_1^T \mathcal{U}_1, \dots, \sigma_p \mathcal{V}_p^T \mathcal{U}_p]$$

Each $\mathcal{V}_i^T \mathcal{U}_i$ is orthogonal and symmetric.

Spectral Decompositions from SVDs (II)

EQUAL SINGULAR VALUES CASE (continued)

$$V^T AV = \text{diag}[\sigma_1 \mathcal{V}_1^T \mathcal{U}_1, \dots, \sigma_p \mathcal{V}_p^T \mathcal{U}_p]$$

Each $\mathcal{V}_i^T \mathcal{U}_i$ is orthogonal and symmetric.

Number of positive/negative eigenvalues equal to $\pm\sigma_i$:

$$m_i^+ + m_i^- = m_i \quad \text{and} \quad m_i^+ - m_i^- = \text{trace}(\mathcal{V}_i^T \mathcal{U}_i)$$

$$m_i^\pm = \frac{m_i \pm \text{trace}(\mathcal{V}_i^T \mathcal{U}_i)}{2}$$

Computing Eigenvectors: For $i = 1 : p$, we compute an orthogonal diagonalization of

$$\mathcal{V}_i^T \mathcal{U}_i = \mathcal{W}_i J_i \mathcal{W}_i^T \quad \text{with} \quad J_i = \text{diag}[I_{m_i^+}, -I_{m_i^-}]$$

and $\mathcal{W}_i = [\mathcal{W}_i^+ \mid \mathcal{W}_i^-] \in \mathbb{R}^{m_i \times m_i}$.

If $\mathcal{W} = \text{diag}[\mathcal{W}_1, \dots, \mathcal{W}_p]$,

$$(V\mathcal{W})^T A (V\mathcal{W}) = \text{diag}[\sigma_1 J_1, \dots, \sigma_p J_p]$$

The columns of

$$Q_i^+ = \mathcal{V}_i \mathcal{W}_i^+ \quad \text{and} \quad Q_i^- = \mathcal{V}_i \mathcal{W}_i^-$$

are eigenvectors of σ_i and $-\sigma_i$ of A .

Spectral Decompositions from SVDs (III)

CLUSTERS OF CLOSE SINGULAR VALUES CASE

Starting point

$$(I + E)A(I + F) = \hat{U}\hat{\Sigma}\hat{V}^T$$
$$\|E\| = O(\epsilon\kappa(X)) \text{ and } \|F\| = O(\epsilon\kappa(Y)).$$

Then $|\sigma_l - \hat{\sigma}_l| = O(\epsilon \max(\kappa(X), \kappa(Y)))\sigma_l, \quad l = 1 : n.$

We have to include two computed contiguous singular values in the same cluster whenever

$$\frac{|\hat{\sigma}_j - \hat{\sigma}_{j+1}|}{\hat{\sigma}_j} \lesssim \epsilon \max(\kappa(X), \kappa(Y))$$

Therefore we are forced to consider

$$A = A^T = U\Sigma V^T$$

$$\Sigma = \text{diag}[\Sigma_1, \dots, \Sigma_k]$$

$$\Sigma_i \in \mathbb{R}^{n_i \times n_i} \text{ cluster and } n_1 + \dots + n_p = n$$

$$U = [U_1 \mid U_2 \mid \dots \mid U_k]$$

$$V = [V_1 \mid V_2 \mid \dots \mid V_k]$$

with $U_i, V_i \in \mathbb{R}^{n \times n_i}$ and $V_i^T U_j = 0$ for $i \neq j$.

Spectral Decompositions from SVDs (IV)

Starting point

$$(I + E)A(I + F) = \hat{U}\hat{\Sigma}\hat{V}^T$$
$$\|E\| = O(\epsilon\kappa(X)) \text{ and } \|F\| = O(\epsilon\kappa(Y)).$$

There exist exact orthogonal matrices $P_i, i = 1 : k$, s.t.

$$\sqrt{\|U_i P_i - \hat{U}_i\|_F^2 + \|V_i P_i - \hat{V}_i\|_F^2} \leq \frac{O(\epsilon \max(\kappa(X), \kappa(Y)))}{\text{relgap}(\Sigma_i, \Sigma_{\bar{i}})}$$

Problem: We only have good approximations to $U_i P_i$ and $V_i P_i$, with $i = 1 : k$.

But, it can be proved

Number of positive/negative eigenvalues with absolute value in $\Sigma_i, \quad 1 : k$:

$$n_i^\pm = \frac{n_i \pm \text{trace}(V_i^T U_i)}{2} = \frac{n_i \pm \text{trace}((V_i P_i)^T U_i P_i)}{2}$$

Spectral Decompositions from SVDs (V)

For the eigenvectors

$$(V_i P_i)^T U_i P_i = W_i J_i W_i^T \text{ with } J_i = \text{diag}[I_{n_i^+}, -I_{n_i^-}]$$

and $W_i = [W_i^+ | W_i^-] \in \mathbb{R}^{n_i \times n_i}$.

The columns of

$$Q_i^+ = V_i P_i W_i^+ \text{ and } Q_i^- = V_i P_i W_i^-$$

are ORTHONORMAL BASES OF THE INVARIANT SUBSPACES CORRESPONDING TO THE POSITIVE AND NEGATIVE EIGENVALUES WITH ABS. VALUES IN Σ_i .

Roundoff errors in these computations

Using $|\sigma_j - \hat{\sigma}_j| = O(\epsilon \max(\kappa(X), \kappa(Y)))\sigma_j$ and $n_i^+ = n_i, n_i - 1, \dots, 0$, we get

$$|\lambda_j - \hat{\lambda}_j| = O(\epsilon \max(\kappa(X), \kappa(Y)))\lambda_j, \quad j = 1 : n.$$

$$\|Q_i^+ - \hat{Q}_i^+\|_F, \|Q_i^- - \hat{Q}_i^-\|_F \leq \frac{O(\epsilon \max(\kappa(X), \kappa(Y)))}{\text{relgap}^*(\Sigma_i, \Sigma_{\bar{i}})}$$

Algorithm

INPUT: $A = A^T$

OUTPUT: e-values and e-vectors of A

1. Compute a RRD $A = \hat{X}\hat{D}\hat{Y}^T$.
2. Compute SVD of the RRD $\hat{U}\hat{\Sigma}\hat{V}^T = \hat{X}\hat{D}\hat{Y}^T$.
HRA Jacobi-type algor. Demmel et al.
3. Compute e-values, e-vectors from SVD.

Step 1: Choose sing. val. clusters: $\hat{\Sigma}_i, \hat{U}_i, \hat{V}_i, i = 1:k$.

Two contiguous singular values in the same cluster if

$$\frac{|\hat{\sigma}_j - \hat{\sigma}_{j+1}|}{\hat{\sigma}_j} \lesssim \epsilon \max(\kappa(X), \kappa(Y)).$$

Step 2: Number of pos./negative e-values in $\hat{\Sigma}_i, 1 : k$:

$$n_i^\pm = \frac{n_i \pm \text{trace}(\hat{V}_i^T \hat{U}_i)}{2}$$

Assign the signs.

(Step 3: Choose other set clusters for eigenvectors.)

Step 4:

$$\hat{V}_i^T \hat{U}_i = [\hat{W}_i^+ | \hat{W}_i^-] \text{diag}[I_{n_i^+}, -I_{n_i^-}] [\hat{W}_i^+ | \hat{W}_i^-]^T$$

$$\hat{Q}_i^+ = \hat{V}_i \hat{W}_i^+ \text{ and } \hat{Q}_i^- = \hat{V}_i \hat{W}_i^-$$

Orthonormal bases of the invariant subspaces
corresponding to positive/negative eigenvalues in $\hat{\Sigma}_i$.

Numerical Experiments

More than 10^5 very satisfactory tests.

We only show: 1500 matrices $A = D_1 B D_1$ with $n = 500$.

$\kappa(B) \leq 1000$ and $10^2 \leq \kappa(D_1) \leq 10^{10}$.

We compare single precision vs. double precision J-O results.

$A = DBD = XDX^T$ and $\kappa(X) \leq 8000$ and $\langle \kappa(X) \rangle \approx 2000$.

$$\vartheta = \frac{\max_i \left| \frac{\lambda_i - \hat{\lambda}_i}{\lambda_i} \right|}{\kappa(X) \epsilon_s}$$

$$\xi_\sigma = \max_i \frac{\|q_i - \hat{q}_i\|_2 \text{relgap}(\sigma_i)}{\kappa(X) \epsilon_s},$$

$$\xi_\lambda = \max_i \frac{\|q_i - \hat{q}_i\|_2 \text{relgap}(\lambda_i)}{\kappa(X) \epsilon_s}.$$

$n = 500$	ϑ		ξ_σ		ξ_λ	
Method	mean	max	mean	max	mean	max
SSVD	0.22	140	0.076	1.3	0.091	1.3
J - O	0.31	320	—	—	0.090	0.67
JAC	10^{12}	10^{15}	—	—	—	—
QR	10^{10}	10^{13}	—	—	—	—