Implicit Standard Jacobi Gives High Relative Accuracy on Rank Revealing Decompositions

Froilán M. Dopico[1]

Plamen Koev[2]   Juan M. Molera[1]

[1]Departamento de Matemáticas, Universidad Carlos III de Madrid

[2]Department of Mathematics, North Carolina State University

7th International Workshop on Accurate Solution of Eigenvalue Problems, Dubrovnik, June 9-12, 2008

Abstract (1)

- **INPUT:** Factors X and D of a decomposition $A = XDX^T$ of a symmetric matrix, where X is well-conditioned and D is diagonal, perhaps indefinite.
- We run the standard Jacobi algorithm to compute eigenvalues and eigenvectors but applying the rotations only on X.
- **BASIC STEP:** Compute a plane Jacobi rotation R such that $(R^T AR)_{ij} = 0$, for some $i \neq j$, then

$$XDX^T \longrightarrow (R^T X)D(R^T X)^T.$$

- From a decomposition of A we obtain a decomposition of $R^T AR$. The matrix A is never formed.

Abstract (1)

- **INPUT:** Factors X and D of a decomposition $A = XDX^T$ of a symmetric matrix, where X is well-conditioned and D is diagonal, perhaps indefinite.
- We run the standard Jacobi algorithm to compute eigenvalues and eigenvectors but applying the rotations only on X.
- **BASIC STEP:** Compute a plane Jacobi rotation R such that $(R^T AR)_{ij} = 0$, for some $i \neq j$, then

$$XDX^T \longrightarrow (R^T X)D(R^T X)^T.$$

- From a decomposition of A we obtain a decomposition of $R^T AR$. The matrix A is never formed.

Abstract (1)

- **INPUT:** Factors X and D of a decomposition $A = XDX^T$ of a symmetric matrix, where X is well-conditioned and D is diagonal, perhaps indefinite.
- We run the standard Jacobi algorithm to compute eigenvalues and eigenvectors but applying the rotations only on X.
- **BASIC STEP:** Compute a plane Jacobi rotation R such that $(R^T AR)_{ij} = 0$, for some $i \neq j$, then

$$XDX^T \longrightarrow (R^T X)D(R^T X)^T.$$

- From a decomposition of A we obtain a decomposition of $R^T AR$. The matrix A is never formed.

Abstract (1)

- **INPUT:** Factors X and D of a decomposition $A = XDX^T$ of a symmetric matrix, where X is well-conditioned and D is diagonal, perhaps indefinite.
- We run the standard Jacobi algorithm to compute eigenvalues and eigenvectors but applying the rotations only on X.
- **BASIC STEP:** Compute a plane Jacobi rotation R such that $(R^T AR)_{ij} = 0$, for some $i \neq j$, then

$$XDX^T \longrightarrow (R^T X)D(R^T X)^T.$$

- From a decomposition of A we obtain a decomposition of $R^T AR$. The matrix A is never formed.

Abstract (2)

- Algorithm stops when the off diagonal part of $A_f = X_f D X_f^T$ is small enough.
- **OUTPUT:**
  1. The eigenvalues of A are the computed diagonal entries of $X_f D X_f^T$.
  2. Eigenvectors are the columns of $R_1 R_2 \cdots R_f$
- Let ϵ be the unit roundoff. The **errors** in computed eigenvalues and eigenvectors are

$$\frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} \leq O(\epsilon\kappa(X)) \quad \text{and} \quad \theta(v_i, \hat{v}_i) \leq \frac{O(\epsilon\kappa(X))}{\min_{j\neq i} \left|\frac{\lambda_i - \lambda_j}{\lambda_i}\right|} \quad \text{for all} \quad i,$$

for any condition number of A, i.e., of D. ($\kappa(X) = \|X\|_2 \|X^{-1}\|_2$)

Abstract (2)

- Algorithm stops when the off diagonal part of $A_f = X_f D X_f^T$ is small enough.
- **OUTPUT:**
  1. The eigenvalues of A are the computed diagonal entries of $X_f D X_f^T$.
  2. Eigenvectors are the columns of $R_1 R_2 \cdots R_f$
- Let ϵ be the unit roundoff. The **errors** in computed eigenvalues and eigenvectors are

$$\frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} \leq O(\epsilon\kappa(X)) \quad \text{and} \quad \theta(v_i, \hat{v}_i) \leq \frac{O(\epsilon\kappa(X))}{\min_{j\neq i} \left|\frac{\lambda_i - \lambda_j}{\lambda_i}\right|} \quad \text{for all} \quad i,$$

for any condition number of A, i.e., of D. ($\kappa(X) = \|X\|_2 \|X^{-1}\|_2$)

Abstract (2)

- Algorithm stops when the off diagonal part of $A_f = X_f D X_f^T$ is small enough.
- **OUTPUT:**
  1. The eigenvalues of A are the computed diagonal entries of $X_f D X_f^T$.
  2. Eigenvectors are the columns of $R_1 R_2 \cdots R_f$
- Let ϵ be the unit roundoff. The **errors** in computed eigenvalues and eigenvectors are

$$\frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} \leq O(\epsilon \kappa(X)) \quad \text{and} \quad \theta(v_i, \hat{v}_i) \leq \frac{O(\epsilon\kappa(X))}{\min_{j \neq i} \left| \frac{\lambda_i - \lambda_j}{\lambda_i} \right|} \quad \text{for all} \quad i,$$

for any condition number of A, i.e., of D. ($\kappa(X) = \|X\|_2 \|X^{-1}\|_2$)

Abstract (2)

- Algorithm stops when the off diagonal part of $A_f = X_f D X_f^T$ is small enough.
- **OUTPUT:**
  1. The eigenvalues of A are the computed diagonal entries of $X_f D X_f^T$.
  2. Eigenvectors are the columns of $R_1 R_2 \cdots R_f$
- Let ϵ be the unit roundoff. The **errors** in computed eigenvalues and eigenvectors are

$$\frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} \leq O(\epsilon\kappa(X)) \quad \text{and} \quad \theta(v_i, \hat{v}_i) \leq \frac{O(\epsilon\kappa(X))}{\min_{j \neq i} \left| \frac{\lambda_i - \lambda_j}{\lambda_i} \right|} \quad \text{for all} \quad i,$$

for any condition number of A, i.e., of D. ($\kappa(X) = \|X\|_2 \|X^{-1}\|_2$)

Abstract (2)

- Algorithm stops when the off diagonal part of $A_f = X_f D X_f^T$ is small enough.
- **OUTPUT:**
  1. The eigenvalues of A are the computed diagonal entries of $X_f D X_f^T$.
  2. Eigenvectors are the columns of $R_1 R_2 \cdots R_f$
- Let ϵ be the unit roundoff. The **errors** in computed eigenvalues and eigenvectors are

$$\frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} \leq O(\epsilon \kappa(X)) \quad \text{and} \quad \theta(v_i, \hat{v}_i) \leq \frac{O(\epsilon \kappa(X))}{\min_{j \neq i} \left| \frac{\lambda_i - \lambda_j}{\lambda_i} \right|} \quad \text{for all} \quad i,$$

for any condition number of A, i.e., of D. ($\kappa(X) = \|X\|_2 \|X^{-1}\|_2$)

Abstract (3)

- This implicit Jacobi algorithm is mathematically equivalent to the standard one.

- This is the first algorithm that
  1. computes accurate eigenvalues an eigenvectors of symmetric (indefinite) matrices,
  2. respects and preserves the symmetry of the problem, and
  3. uses only orthogonal transformations.

Abstract (3)

- This implicit Jacobi algorithm is mathematically equivalent to the standard one.

- This is the first algorithm that
  1. computes accurate eigenvalues an eigenvectors of symmetric (indefinite) matrices,
  2. respects and preserves the symmetry of the problem, and
  3. uses only orthogonal transformations.

Outline

1. **Why is the Implicit Jacobi algorithm interesting?**
2. **Why does Implicit Jacobi compute accurate eigenvalues and eigenvectors?**
3. **The rigorous roundoff error result**
4. **Singular matrices $A = XDX^T$**
5. **Numerical Experiments**
6. **Conclusions**

Outline

1. **Why is the Implicit Jacobi algorithm interesting?**
2. Why does Implicit Jacobi compute accurate eigenvalues and eigenvectors?
3. The rigorous roundoff error result
4. Singular matrices $A = XDX^T$
5. Numerical Experiments
6. Conclusions

Accurate eigencomputations for symmetric matrices

- In the last twenty years an intensive research effort has been made to compute eigenvalues and eigenvectors of $n \times n$ symmetric matrices **to high relative accuracy (hra)**.
- Given $A = A^T \in \mathbb{R}^{n \times n}$, we will say that an algorithm computes **all** its **eigenvalues and eigenvectors** to **hra** if the computed eigenvalues and eigenvectors satisfy

$$|\widehat{\lambda}_i - \lambda_i| = O(\epsilon)\,|\lambda_i| \quad \text{and} \quad \theta(v_i, \hat{v}_i) \leq \frac{O(\epsilon)}{\min\limits_{j \neq i} \left| \frac{\lambda_i - \lambda_j}{\lambda_i} \right|} \quad \text{for all} \quad i$$

  and, in addition,
  1. the cost is $O(n^3)$ flops,
  2. and extra precision is not used.
- HRA is only possible for **special types of matrices**.

Accurate eigencomputations for symmetric matrices

- In the last twenty years an intensive research effort has been made to compute eigenvalues and eigenvectors of $n \times n$ symmetric matrices **to high relative accuracy (hra)**.
- Given $A = A^T \in \mathbb{R}^{n \times n}$, we will say that an algorithm computes **all** its **eigenvalues and eigenvectors** to **hra** if the computed eigenvalues and eigenvectors satisfy

$$|\widehat{\lambda}_i - \lambda_i| = O(\epsilon)\,|\lambda_i| \quad \text{and} \quad \theta(v_i, \hat{v}_i) \leq \frac{O(\epsilon)}{\min\limits_{j \neq i} \left| \frac{\lambda_i - \lambda_j}{\lambda_i} \right|} \quad \text{for all} \quad i$$

  and, in addition,

  1. the cost is $O(n^3)$ flops,
  2. and extra precision is not used.

- HRA is only possible for **special types of matrices**.

Accurate eigencomputations for symmetric matrices

- In the last twenty years an intensive research effort has been made to compute eigenvalues and eigenvectors of $n \times n$ symmetric matrices **to high relative accuracy (hra)**.
- Given $A = A^T \in \mathbb{R}^{n \times n}$, we will say that an algorithm computes **all** its **eigenvalues and eigenvectors** to **hra** if the computed eigenvalues and eigenvectors satisfy

$$|\widehat{\lambda}_i - \lambda_i| = O(\epsilon)\,|\lambda_i| \quad \text{and} \quad \theta(v_i, \hat{v}_i) \leq \frac{O(\epsilon)}{\min\limits_{j \neq i} \left| \frac{\lambda_i - \lambda_j}{\lambda_i} \right|} \quad \text{for all} \quad i$$

  and, in addition,

  1. the cost is $O(n^3)$ flops,
  2. and extra precision is not used.

- HRA is only possible for **special types of matrices**.

Accurate eigencomputations for symmetric matrices

- In the last twenty years an intensive research effort has been made to compute eigenvalues and eigenvectors of $n \times n$ symmetric matrices **to high relative accuracy (hra)**.
- Given $A = A^T \in \mathbb{R}^{n\times n}$, we will say that an algorithm computes **all** its **eigenvalues and eigenvectors** to **hra** if the computed eigenvalues and eigenvectors satisfy

$$|\widehat{\lambda}_i - \lambda_i| = O(\epsilon)\,|\lambda_i| \quad \text{and} \quad \theta(v_i, \hat{v}_i) \leq \frac{O(\epsilon)}{\min\limits_{j\neq i} \left|\frac{\lambda_i - \lambda_j}{\lambda_i}\right|} \quad \text{for all} \quad i$$

  and, in addition,
  1. the cost is $O(n^3)$ flops,
  2. and extra precision is not used.
- HRA is only possible for **special types of matrices**.

Accurate eigencomputations for symmetric matrices

- In the last twenty years an intensive research effort has been made to compute eigenvalues and eigenvectors of $n \times n$ symmetric matrices **to high relative accuracy (hra)**.
- Given $A = A^T \in \mathbb{R}^{n\times n}$, we will say that an algorithm computes **all** its **eigenvalues and eigenvectors** to **hra** if the computed eigenvalues and eigenvectors satisfy

$$|\widehat{\lambda}_i - \lambda_i| = O(\epsilon)\,|\lambda_i| \quad \text{and} \quad \theta(v_i, \hat{v}_i) \leq \frac{O(\epsilon)}{\min_{j\neq i} \left|\frac{\lambda_i - \lambda_j}{\lambda_i}\right|} \quad \text{for all} \quad i$$

  and, in addition,

  1. the cost is $O(n^3)$ flops,
  2. and extra precision is not used.

- HRA is only possible for **special types of matrices**.

Accurate eigencomputations for symmetric matrices

- In the last twenty years an intensive research effort has been made to compute eigenvalues and eigenvectors of $n \times n$ symmetric matrices **to high relative accuracy (hra)**.
- Given $A = A^T \in \mathbb{R}^{n \times n}$, we will say that an algorithm computes **all** its **eigenvalues and eigenvectors** to **hra** if the computed eigenvalues and eigenvectors satisfy

$$|\widehat{\lambda}_i - \lambda_i| = O(\epsilon)\,|\lambda_i| \quad \text{and} \quad \theta(v_i, \hat{v}_i) \leq \frac{O(\epsilon)}{\min_{j \neq i} \left| \frac{\lambda_i - \lambda_j}{\lambda_i} \right|} \quad \text{for all} \quad i$$

  and, in addition,
  1. the cost is $O(n^3)$ flops,
  2. and extra precision is not used.
- HRA is only possible for **special types of matrices**.

HRA is not obtained from standard algorithms

EXAMPLE: Symmetric INDEFINITE 100×100 Cauchy matrix A

$$a_{ij} = \frac{1}{x_i + x_j}, \quad \text{with} \quad \begin{cases} x_i = i - 0.5 \;\; for \;\; i = 1:99 \\ x_{100} = -99.5 \end{cases}$$

- $\kappa(A) = 3.5 \cdot 10^{147}$
- **Errors in accurate algorithm (Factorization + Imp. Jacobi)** compared to 200-decimal digits MATLAB's `eig` command

$$\max_i \frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} = 1.2 \cdot 10^{-13} \quad \text{and} \quad \max_i \|\hat{v}_i - v_i\|_2 = 5.7 \cdot 10^{-14}.$$

- **Errors in MATLAB's `eig` function**

$$\max_i \frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} = 1.84 \cdot 10^{132} \quad \text{and} \quad \max_i \|\hat{v}_i - v_i\|_2 = 1.41.$$

HRA is not obtained from standard algorithms

EXAMPLE: Symmetric INDEFINITE 100×100 Cauchy matrix A

$$a_{ij} = \frac{1}{x_i + x_j}, \quad \text{with} \quad \begin{cases} x_i = i - 0.5 \quad for \quad i = 1:99 \\ x_{100} = -99.5 \end{cases}$$

- $\kappa(A) = 3.5 \cdot 10^{147}$
- **Errors in accurate algorithm (Factorization + Imp. Jacobi)** compared to 200-decimal digits MATLAB's `eig` command

$$\max_i \frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} = 1.2 \cdot 10^{-13} \quad \text{and} \quad \max_i \|\hat{v}_i - v_i\|_2 = 5.7 \cdot 10^{-14}.$$

- **Errors in MATLAB's `eig` function**

$$\max_i \frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} = 1.84 \cdot 10^{132} \quad \text{and} \quad \max_i \|\hat{v}_i - v_i\|_2 = 1.41.$$

HRA is not obtained from standard algorithms

EXAMPLE: Symmetric INDEFINITE 100×100 Cauchy matrix A

$$a_{ij} = \frac{1}{x_i + x_j}, \quad \text{with} \quad \begin{cases} x_i = i - 0.5 \quad for \quad i = 1:99 \\ x_{100} = -99.5 \end{cases}$$

- $\kappa(A) = 3.5 \cdot 10^{147}$
- **Errors in accurate algorithm (Factorization + Imp. Jacobi)** compared to 200-decimal digits MATLAB's `eig` command

$$\max_i \frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} = 1.2 \cdot 10^{-13} \quad \text{and} \quad \max_i \|\hat{v}_i - v_i\|_2 = 5.7 \cdot 10^{-14}.$$

- **Errors in MATLAB's `eig` function**

$$\max_i \frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} = 1.84 \cdot 10^{132} \quad \text{and} \quad \max_i \|\hat{v}_i - v_i\|_2 = 1.41.$$

HRA is not obtained from standard algorithms

EXAMPLE: Symmetric INDEFINITE 100×100 Cauchy matrix A

$$a_{ij} = \frac{1}{x_i + x_j}, \quad \text{with} \quad \begin{cases} x_i = i - 0.5 \quad for \quad i = 1:99 \\ x_{100} = -99.5 \end{cases}$$

- $\kappa(A) = 3.5 \cdot 10^{147}$
- **Errors in accurate algorithm (Factorization + Imp. Jacobi)** compared to 200-decimal digits MATLAB's `eig` command

$$\max_i \frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} = 1.2 \cdot 10^{-13} \quad \text{and} \quad \max_i \|\hat{v}_i - v_i\|_2 = 5.7 \cdot 10^{-14}.$$

- **Errors in MATLAB's `eig` function**

$$\max_i \frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} = 1.84 \cdot 10^{132} \quad \text{and} \quad \max_i \|\hat{v}_i - v_i\|_2 = 1.41.$$

Selected references for HRA algorithms for symmetric eigenproblems (SVDs)

- Demmel-Kahan (1990), Barlow-Demmel (1990), Demmel-Veselić (1992), Demmel-Gragg (1993), Demmel (1999)
- Veselić-Slapničar (1992, 93, 03)
- Fernando-Parlett (1994)
- Drmač (1998, 99), Drmač-Veselić (2008)
- Demmel, Gu, Eisenstat, Slapničar, Veselić, Drmač (1999)
- Demmel-Koev (2001, 04, 06), Koev (2005, 07)
- D-Molera-Moro(03),D-Koev(06,07),Peláez-Moro(06),D-Molera(08)
- Ye (2008)
- It has motivated Spectral Relative Perturbation Theory (Eisenstat, Ipsen, R.C. Li, Mathias,Truhar)
- Improved Convergence analysis of Jacobi Algorithms (Drmač, Hari, Matejas).
- Application to MRRR $O(n^2)$-algorithm by Dhillon and Parlett.
- Analysis of block Jacobi methods (Hari, Drmač, Singer)...

Selected references for HRA algorithms for symmetric eigenproblems (SVDs)

- Demmel-Kahan (1990), Barlow-Demmel (1990), Demmel-Veselić (1992), Demmel-Gragg (1993), Demmel (1999)
- Veselić-Slapničar (1992, 93, 03)
- Fernando-Parlett (1994)
- Drmač (1998, 99), Drmač-Veselić (2008)
- Demmel, Gu, Eisenstat, Slapničar, Veselić, Drmač (1999)
- Demmel-Koev (2001, 04, 06), Koev (2005, 07)
- D-Molera-Moro(03),D-Koev(06,07),Peláez-Moro(06),D-Molera(08)
- Ye (2008)
- It has motivated Spectral Relative Perturbation Theory (Eisenstat, Ipsen, R.C. Li, Mathias,Truhar)
- Improved Convergence analysis of Jacobi Algorithms (Drmač, Hari, Matejas).
- Application to MRRR $O(n^2)$-algorithm by Dhillon and Parlett.
- Analysis of block Jacobi methods (Hari, Drmač, Singer)...

Selected references for HRA algorithms for symmetric eigenproblems (SVDs)

- Demmel-Kahan (1990), Barlow-Demmel (1990), Demmel-Veselić (1992), Demmel-Gragg (1993), Demmel (1999)
- Veselić-Slapničar (1992, 93, 03)
- Fernando-Parlett (1994)
- Drmač (1998, 99), Drmač-Veselić (2008)
- Demmel, Gu, Eisenstat, Slapničar, Veselić, Drmač (1999)
- Demmel-Koev (2001, 04, 06), Koev (2005, 07)
- D-Molera-Moro(03),D-Koev(06,07),Peláez-Moro(06),D-Molera(08)
- Ye (2008)
- It has motivated Spectral Relative Perturbation Theory (Eisenstat, Ipsen, R.C. Li, Mathias,Truhar)
- Improved Convergence analysis of Jacobi Algorithms (Drmač, Hari, Matejas).
- Application to MRRR $O(n^2)$-algorithm by Dhillon and Parlett.
- Analysis of block Jacobi methods (Hari, Drmač, Singer)...

Selected references for HRA algorithms for symmetric eigenproblems (SVDs)

- Demmel-Kahan (1990), Barlow-Demmel (1990), Demmel-Veselić (1992), Demmel-Gragg (1993), Demmel (1999)
- Veselić-Slapničar (1992, 93, 03)
- Fernando-Parlett (1994)
- Drmač (1998, 99), Drmač-Veselić (2008)
- Demmel, Gu, Eisenstat, Slapničar, Veselić, Drmač (1999)
- Demmel-Koev (2001, 04, 06), Koev (2005, 07)
- D-Molera-Moro(03),D-Koev(06,07),Peláez-Moro(06),D-Molera(08)
- Ye (2008)
- It has motivated Spectral Relative Perturbation Theory (Eisenstat, Ipsen, R.C. Li, Mathias,Truhar)
- Improved Convergence analysis of Jacobi Algorithms (Drmač, Hari, Matejas).
- Application to MRRR $O(n^2)$-algorithm by Dhillon and Parlett.
- Analysis of block Jacobi methods (Hari, Drmač, Singer)...

Selected references for HRA algorithms for symmetric eigenproblems (SVDs)

- Demmel-Kahan (1990), Barlow-Demmel (1990), Demmel-Veselić (1992), Demmel-Gragg (1993), Demmel (1999)
- Veselić-Slapničar (1992, 93, 03)
- Fernando-Parlett (1994)
- Drmač (1998, 99), Drmač-Veselić (2008)
- Demmel, Gu, Eisenstat, Slapničar, Veselić, Drmač (1999)
- Demmel-Koev (2001, 04, 06), Koev (2005, 07)
- D-Molera-Moro(03),D-Koev(06,07),Peláez-Moro(06),D-Molera(08)
- Ye (2008)
- It has motivated Spectral Relative Perturbation Theory (Eisenstat, Ipsen, R.C. Li, Mathias,Truhar)
- Improved Convergence analysis of Jacobi Algorithms (Drmač, Hari, Matejas).
- Application to MRRR $O(n^2)$-algorithm by Dhillon and Parlett.
- Analysis of block Jacobi methods (Hari, Drmač, Singer)...

Key unifying idea: Rank Revealing Decompositions (RRD) (Demmel et al. 1999)

We restrict to symmetric RRDs of $A = A^T \in \mathbb{R}^{n\times n}$.

- Compute first an **accurate** RRD

$$A = XDX^T,$$

X is **well-conditioned** and D is **diagonal and nonsingular**.

Remark: Accuracy is only possible for special types of matrices through structured implementations of Gaussian elimination with complete pivoting (**GECP**), or variations of GECP.

- Compute eigenvalues and eigenvectors with **hra** from the **factors** X **and** D with a **Jacobi-type** method.

Key unifying idea: Rank Revealing Decompositions (RRD) (Demmel et al. 1999)

We restrict to symmetric RRDs of $A = A^T \in \mathbb{R}^{n \times n}$.

- Compute first an **accurate** RRD

$$A = XDX^T,$$

X is **well-conditioned** and D is **diagonal and nonsingular**.

Remark: **Accuracy** is only possible for special types of matrices through structured implementations of Gaussian elimination with complete pivoting (**GECP**), or variations of GECP.

- Compute eigenvalues and eigenvectors with **hra** from the **factors** X **and** D with a **Jacobi-type** method.

Key unifying idea: Rank Revealing Decompositions (RRD) (Demmel et al. 1999)

We restrict to symmetric RRDs of $A = A^T \in \mathbb{R}^{n\times n}$.

- Compute first an **accurate** RRD

$$A = XDX^T,$$

  X is **well-conditioned** and D is **diagonal and nonsingular**.

  Remark: **Accuracy** is only possible for special types of matrices through structured implementations of Gaussian elimination with complete pivoting (**GECP**), or variations of GECP.

- Compute eigenvalues and eigenvectors with **hra** from the **factors** X **and** D with a **Jacobi-type** method.

Key unifying idea: Rank Revealing Decompositions (RRD) (Demmel et al. 1999)

We restrict to symmetric RRDs of $A = A^T \in \mathbb{R}^{n\times n}$.

- Compute first an **accurate** RRD

$$A = XDX^T,$$

  X is **well-conditioned** and D is **diagonal and nonsingular**.

  Remark: **Accuracy** is only possible for special types of matrices through structured implementations of Gaussian elimination with complete pivoting (**GECP**), or variations of GECP.

- Compute eigenvalues and eigenvectors with **hra** from the **factors** X **and** D with a **Jacobi-type** method.

Classes of symmetric matrices with accurate RRDs algorithms

1. Well Scaled Symmetric Positive Definite (Demmel and Veselić).
2. Scaled diagonally dominant (Barlow and Demmel)
3. Symmetric Cauchy and Scaled-Cauchy (D and Koev).
4. Symmetric Vandermonde (D and Koev).
5. Symmetric Totally nonnegative (D and Koev).
6. Symmetric Graded Matrices (D and Molera).
7. Symmetric DSTU and TSC (Peláez and Moro).
8. Symmetric diagonally dominant M-matrices (Demmel and Koev), (Peña).
9. Symmetric diagonally dominant (Ye)....

A symmetric RRD determines accurately its eigenvalues and eigenvectors (I): multiplicative perturbations

Theorem (D., Koev (2006))

Let $A = A^T \in \mathbb{R}^{n \times n}$ *and* $A = XDX^T$ *be an RRD of* A*, where* $X \in \mathbb{R}^{n \times r}$*,* $n \geq r$*, and* $D = \mathrm{diag}(d_1, \ldots, d_r) \in \mathbb{R}^{r \times r}$*. Let* $\widehat{X}$ *and* $\widehat{D} = \mathrm{diag}(\widehat{d}_1, \ldots, \widehat{d}_r)$ *be perturbations of* X *and* D*, respectively, that satisfy*

$$\frac{\|\widehat{X} - X\|_2}{\|X\|_2} \leq \delta \quad \textit{and} \quad \frac{|\widehat{d}_i - d_i|}{|d_i|} \leq \delta \quad \textit{for} \;\; i = 1, \ldots, r,$$

where $\delta < 1$*. Then*

$$\widehat{X}\widehat{D}\widehat{X}^T = (I + F)A(I + F)^T,$$

with $\|F\|_2 \leq (2\delta + \delta^2)\kappa(X)$*.*

A symmetric RRD determines accurately its eigenvalues and eigenvectors (II): multiplicative perturbation theory

Theorem (Eisenstat, Ipsen (1995) and R. C. Li (2000))

Let $A = A^T \in \mathbb{R}^{n \times n}$ *and* $\widetilde{A} = (I + F)A(I + F)^T \in \mathbb{R}^{n \times n}$, *where* $\|F\|_2 < 1$. *Let* $\lambda_1 \geq \cdots \geq \lambda_n$ *and* $\widetilde{\lambda}_1 \geq \cdots \geq \widetilde{\lambda}_n$ *be, respectively, the eigenvalues of* A *and* $\widetilde{A}$. *Then*

- $$|\widetilde{\lambda}_i - \lambda_i| \leq (2\,\|F\|_2 + \|F\|_2^2)\ |\lambda_i|, \quad \text{for } i = 1, \ldots, n$$
- *For the corresponding eigenvectors,* v_i *and* $\widetilde{v}_i$,

$$\frac{1}{2} \sin\, 2\theta(v_i, \widetilde{v}_i) \leq \frac{2}{\min\limits_{j \neq i} \left| \frac{\lambda_i - \lambda_j}{\lambda_i} \right|} \cdot \frac{1 + \|F\|_2}{1 - \|F\|_2} (2\,\|F\|_2 + \|F\|_2^2)$$

A symmetric RRD determines accurately its eigenvalues and eigenvectors (III): Final Result

Corollary (D., Koev (2006))

Let $A = A^T = XDX^T$ *be an RRD. Let* $\widehat{X}$ *and* $\widehat{D} = \mathrm{diag}(\widehat{d}_1, \ldots, \widehat{d}_r)$ *be perturbations of* X *and* D *such that*

$$\|\widehat{X} - X\|_2 \leq \delta \, \|X\|_2 \quad \text{and} \quad |\widehat{d}_i - d_i| \leq \delta \, |d_i| \quad \text{for } \ i = 1, \ldots, r,$$

where $\delta < 1$. *Then, for all* i, *the e-values,* $\widehat{\lambda}_i$, *and e-vectors,* $\widehat{v}_i$, *of* $\widehat{X}\widehat{D}\widehat{X}^T$ *satisfy*

$$\left| \frac{\lambda_i - \widehat{\lambda}_i}{\lambda_i} \right| \leq \kappa(X) \left(4\delta + 2\delta^2 + \kappa(X) \left(2\delta + \delta^2\right)^2 \right) \approx 4\, \delta \, \kappa(X) + O(\delta^2)$$

$$\frac{1}{2} \sin \, 2\theta(v_i, \widehat{v}_i) \leq \frac{8\, \delta \, \kappa(X) + O(\delta^2)}{\min\limits_{j \neq i} \left| \frac{\lambda_i - \lambda_j}{\lambda_i} \right|}$$

Accurate e-values and e-vectors from X and D (1): Positive definite case

Algorithm (Demmel, Veselić (1992))

Given RRD $A = XDX^T$ **positive definite**:

1. Compute SVD of
$$X\sqrt{D} = U\Sigma V^T$$
with **one-sided Jacobi on the left.**
2. The spectral decomposition is
$$A = X\sqrt{D}(X\sqrt{D})^T = U\Sigma^2 U^T.$$

Accurate e-values and e-vectors from X and D (2)

Comments on Algorithm by Demmel and Veselić

Fully satisfactory algorithm because:

- The symmetry is preserved.
- Only orthogonal transformations are used.

Remarks

- If the Jacobi rotations are applied on $X\sqrt{D}$ from the **right** then the algorithm is faster but it is not possible to prove that the error bounds are small.
- If the rotations are applied on $X\sqrt{D}$ on the **left** then it is mathematically equivalent to apply the standard Jacobi algorithm to XDX^T.

Accurate e-values and e-vectors from X and D (2)

Comments on Algorithm by Demmel and Veselić

Fully satisfactory algorithm because:

- The symmetry is preserved.
- Only orthogonal transformations are used.

Remarks

- If the Jacobi rotations are applied on $X\sqrt{D}$ from the **right** then the algorithm is faster but it is not possible to prove that the error bounds are small.
- If the rotations are applied on $X\sqrt{D}$ on the **left** then it is mathematically equivalent to apply the standard Jacobi algorithm to XDX^T.

Accurate e-values and e-vectors from X and D (3): General case

Hyperbolic Algorithm (Veselić (1993), Slapničar (1992, 2003))

Given RRD $A = XDX^T$ **possibly indefinite**:

1. Write
$$A = X\sqrt{|D|}\, J \left(X\sqrt{|D|}\right)^T,$$
with $J = \text{diag}\{\pm 1\}$.
2. Compute **Hyperbolic SVD** of
$$X\sqrt{|D|} = U\Sigma H^T \text{ where } U^T U = I,\ H^T J H = J$$
with **hyperbolic one-sided Jacobi** on the right.
3. The spectral decomposition is
$$A = U\,(\Sigma^2\, J)\, U^T$$

Accurate e-values and e-vectors from X and D (4)

Comments on Hyperbolic Algorithm

Not fully satisfactory algorithm because:

- **Hyperbolic rotations are used.**
- Symmetric matrices are diagonalizable by orthogonal similarity.
- It is not possible to prove that the error bounds are small.
- It works well in practice.

Accurate e-values and e-vectors from X and D (5): General case

SSVD Algorithm (D, Molera, Moro (2003), D, Molera (2008))

Given RRD $A = XDX^T$ **possibly indefinite**:

1. Compute SVD of $A = U\Sigma V^T$ from RRD using a **nonsymmetric** algorithm by Demmel et al. (1999) that uses one-sided Jacobi.
2. Compute eigenvalues and eigenvectors from SVD by using $A = A^T$.

Comments on SSVD Algorithm

Not fully satisfactory algorithm because:

- **The symmetry is not respected.** (It allows us flexibility by using nonsymmetric RRDs).
- HRA error bounds are perfect for eigenvalues and eigenvectors,
- but to get accurate e-vectors requires a delicate process.

Accurate e-values and e-vectors from X and D (5): General case

SSVD Algorithm (D, Molera, Moro (2003), D, Molera (2008))

Given RRD $A = XDX^T$ **possibly indefinite**:

1. Compute SVD of $A = U\Sigma V^T$ from RRD using a **nonsymmetric** algorithm by Demmel et al. (1999) that uses one-sided Jacobi.
2. Compute eigenvalues and eigenvectors from SVD by using $A = A^T$.

Comments on SSVD Algorithm

Not fully satisfactory algorithm because:

- **The symmetry is not respected.** (It allows us flexibility by using nonsymmetric RRDs).
- HRA error bounds are perfect for eigenvalues and eigenvectors,
- but to get accurate e-vectors requires a delicate process.

Our Goal

To prove that the standard Jacobi algorithm **implicitly** applied on the factor X of a given RRD

$$A = XDX^T$$

possibly indefinite:

1. computes the eigenvalues and eigenvectors of A to high relative accuracy.
2. Note that it preserves the symmetry of the problem, and
3. uses only orthogonal transformations.

Our Goal

To prove that the standard Jacobi algorithm **implicitly** applied on the factor X of a given RRD

$$A = XDX^T$$

possibly indefinite:

1. computes the eigenvalues and eigenvectors of A to high relative accuracy.
2. Note that it preserves the symmetry of the problem, and
3. uses only orthogonal transformations.

Our Goal

To prove that the standard Jacobi algorithm **implicitly** applied on the factor X of a given RRD

$$A = XDX^T$$

possibly indefinite:

1. computes the eigenvalues and eigenvectors of A to high relative accuracy.
2. Note that it preserves the symmetry of the problem, and
3. uses only orthogonal transformations.

Our Goal

To prove that the standard Jacobi algorithm **implicitly** applied on the factor X of a given RRD

$$A = XDX^T$$

possibly indefinite:

1. computes the eigenvalues and eigenvectors of A to high relative accuracy.
2. Note that it preserves the symmetry of the problem, and
3. uses only orthogonal transformations.

Outline

Notation for Jacobi rotation ($c^2 + s^2 = 1$)

$$
R(i,j,c,s) \quad = \quad \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array}
\begin{array}{c} \begin{array}{ccccccc} & & i & & j & & \end{array} \\
\begin{bmatrix}
1 & & & & & & \\
& \ddots & & & & & \\
& & c & & -s & & \\
& & & \ddots & & & \\
& & s & & c & & \\
& & & & & \ddots & \\
& & & & & & 1
\end{bmatrix}
\end{array}
$$

Implicit Jacobi for square factors

INPUT: $X \in \mathbb{R}^{n\times n}$ nonsingular and $D \in \mathbb{R}^{n\times n}$ diag. and nonsingular
OUTPUT: e-values, λ_i, and matrix of e-vectors, U, of $A = XDX^T$

$U = I_n$
repeat
  for $i < j$
    compute a_{ii}, a_{ij}, a_{jj} of $A = XDX^T$ and $T = \left[\begin{smallmatrix} c & -s \\ s & c \end{smallmatrix}\right]$, such that

$$T^T \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ij} & a_{jj} \end{bmatrix} T = \begin{bmatrix} \mu_1 & \\ & \mu_2 \end{bmatrix}$$

    $X = R(i,j,c,s)^T\, X$
    $U = U\, R(i,j,c,s)$
  endfor
until convergence $\left(\frac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}} \leq \text{tol} = O(\epsilon) \quad \text{for all } i > j \right)$
compute $\lambda_k = a_{kk}$ for $k = 1, 2, \ldots, n$.

Jacobi rotations on X preserve accurate e-values and e-vectors

Lemma (Small multiplicative backward errors of Jacobi rotations)

Let R_i be **exact** *Jacobi rotations and $\widehat{R}_i$ their floating point approximations. Then*

1. $$\widehat{X}_N \equiv \texttt{fl}(\widehat{R}_N^T \cdots \widehat{R}_1^T X) = (I + F) R_N^T \cdots R_1^T X,$$

   where $\|F\|_2 = O(N\,\epsilon\,\kappa(X))$, and

2. $$\widehat{X}_N D \widehat{X}_N^T = (I+F)(R_1 \cdots R_N)^T X D X^T (R_1 \cdots R_N)(I+F)^T$$

Jacobi rotations on X preserve accurate e-values and e-vectors

Lemma (Small multiplicative backward errors of Jacobi rotations)

Let R_i be **exact** *Jacobi rotations and $\widehat{R}_i$ their floating point approximations. Then*

1. $$\widehat{X}_N \equiv \texttt{fl}(\widehat{R}_N^T \cdots \widehat{R}_1^T X) = (I+F) R_N^T \cdots R_1^T X,$$

   where $\|F\|_2 = O(N\,\epsilon\,\kappa(X))$, and

2. $$\widehat{X}_N D \widehat{X}_N^T = (I+F)(R_1 \cdots R_N)^T X D X^T (R_1 \cdots R_N)(I+F)^T$$

Jacobi rotations on X preserve accurate e-values and e-vectors

Lemma (Small multiplicative backward errors of Jacobi rotations)

Let R_i be **exact** *Jacobi rotations and $\widehat{R}_i$ their floating point approximations. Then*

1. $$\widehat{X}_N \equiv \texttt{fl}(\widehat{R}_N^T \cdots \widehat{R}_1^T X) = (I+F) R_N^T \cdots R_1^T X,$$

   where $\|F\|_2 = O(N\,\epsilon\,\kappa(X))$, and

2. $$\widehat{X}_N D \widehat{X}_N^T = (I+F)(R_1 \cdots R_N)^T X D X^T (R_1 \cdots R_N)(I+F)^T$$

Proof of Rounding Errors in Jacobi rotations

Proof.

Let $U^T = R_N^T \cdots R_1^T$.

- $\mathtt{fl}(\widehat{R}_N^T \cdots \widehat{R}_1^T X) = R_N^T \cdots R_1^T (X + E)$ with $\|E\|_2 = O(N\epsilon \|X\|_2)$.
- $\mathtt{fl}(\widehat{R}_N^T \cdots \widehat{R}_1^T X) = U^T (I + EX^{-1}) X = (I + U^T E X^{-1} U) U^T X$.
- $\|U^T E X^{-1} U\|_2 = \|E X^{-1}\|_2 = O(N\epsilon\kappa(X))$.

$\square$

Proof of Rounding Errors in Jacobi rotations

Proof.

Let $U^T = R_N^T \cdots R_1^T$.

- $\mathtt{fl}(\widehat{R}_N^T \cdots \widehat{R}_1^T X) = R_N^T \cdots R_1^T (X + E)$ with $\|E\|_2 = O(N\epsilon\|X\|_2)$.
- $\mathtt{fl}(\widehat{R}_N^T \cdots \widehat{R}_1^T X) = U^T(I + EX^{-1})X = (I + U^T E X^{-1} U) U^T X$.
- $\|U^T E X^{-1} U\|_2 = \|EX^{-1}\|_2 = O(N\epsilon\kappa(X))$.

□

Proof of Rounding Errors in Jacobi rotations

Proof.

Let $U^T = R_N^T \cdots R_1^T$.

- $\mathtt{fl}(\widehat{R}_N^T \cdots \widehat{R}_1^T X) = R_N^T \cdots R_1^T (X + E)$ with $\|E\|_2 = O(N\epsilon \|X\|_2)$.
- $\mathtt{fl}(\widehat{R}_N^T \cdots \widehat{R}_1^T X) = U^T (I + EX^{-1}) X = (I + U^T E X^{-1} U) U^T X$.
- $\|U^T E X^{-1} U\|_2 = \|EX^{-1}\|_2 = O(N\epsilon\kappa(X))$.

□

Implicit Jacobi for square factors

INPUT: $X \in \mathbb{R}^{n\times n}$ nonsingular and $D \in \mathbb{R}^{n\times n}$ diag. and nonsingular
OUTPUT: e-values, λ_i, and matrix of e-vectors, U, of $A = XDX^T$

$U = I_n$
repeat
  for $i < j$
    compute a_{ii}, a_{ij}, a_{jj} of $A = XDX^T$ and $T = \left[\begin{smallmatrix} c & -s \\ s & c \end{smallmatrix}\right]$, such that

$$T^T \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ij} & a_{jj} \end{bmatrix} T = \begin{bmatrix} \mu_1 & \\ & \mu_2 \end{bmatrix}$$

    $X = R(i,j,c,s)^T\, X$
    $U = U\, R(i,j,c,s)$
  endfor
until convergence $\left(\frac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}} \leq \text{tol} = O(\epsilon) \quad \text{for all } i > j \right)$
compute $\lambda_k = a_{kk}$ for $k = 1, 2, \ldots, n$. $\longrightarrow$ IS THIS ACCURATE???

Implicit Jacobi for square factors

INPUT: $X \in \mathbb{R}^{n\times n}$ nonsingular and $D \in \mathbb{R}^{n\times n}$ diag. and nonsingular
OUTPUT: e-values, λ_i, and matrix of e-vectors, U, of $A = XDX^T$

$U = I_n$
repeat
  for $i < j$
    compute a_{ii}, a_{ij}, a_{jj} of $A = XDX^T$ and $T = \left[\begin{smallmatrix} c & -s \\ s & c \end{smallmatrix}\right]$, such that

$$T^T \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ij} & a_{jj} \end{bmatrix} T = \begin{bmatrix} \mu_1 & \\ & \mu_2 \end{bmatrix}$$

    $X = R(i,j,c,s)^T\, X$
    $U = U\, R(i,j,c,s)$
  endfor
until convergence $\left(\dfrac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}} \leq \text{tol} = O(\epsilon) \quad \text{for all } i > j \right)$
compute $\lambda_k = a_{kk}$ for $k = 1, 2, \ldots, n.$ $\longrightarrow$ IS THIS ACCURATE???

Errors on diagonal entries of almost diagonal RRDs (I)

Given $X \in \mathbb{R}^{n\times n}$ nonsingular and $D = \mathrm{diag}(d_1, \ldots, d_n) \in \mathbb{R}^{n\times n}$ diagonal and nonsingular:

- Assume that $A = XDX^T$ satisfies $\frac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}} = O(\epsilon)$ for all $i > j$.
- $a_{ii} = \sum_{k=1}^{n} x_{ik}^2 d_k$
- $$\left| \frac{\mathtt{fl}(a_{ii}) - a_{ii}}{a_{ii}} \right| \leq \frac{(n+1)\epsilon}{1-(n+1)\epsilon} \, \frac{\sum_{k=1}^{n} x_{ik}^2 |d_k|}{\left| \sum_{k=1}^{n} x_{ik}^2 d_k \right|}$$

Errors on diagonal entries of almost diagonal RRDs (I)

Given $X \in \mathbb{R}^{n\times n}$ nonsingular and $D = \operatorname{diag}(d_1, \ldots, d_n) \in \mathbb{R}^{n\times n}$ diagonal and nonsingular:

- Assume that $A = XDX^T$ satisfies $\dfrac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}} = O(\epsilon)$ for all $i > j$.
- $a_{ii} = \displaystyle\sum_{k=1}^{n} x_{ik}^2 d_k$
-

$$\left|\frac{\mathtt{fl}(a_{ii}) - a_{ii}}{a_{ii}}\right| \leq \frac{(n+1)\epsilon}{1-(n+1)\epsilon}\,\frac{\displaystyle\sum_{k=1}^{n} x_{ik}^2 |d_k|}{\left|\displaystyle\sum_{k=1}^{n} x_{ik}^2 d_k\right|}$$

Errors on diagonal entries of almost diagonal RRDs (I)

Given $X \in \mathbb{R}^{n\times n}$ nonsingular and $D = \operatorname{diag}(d_1, \dots, d_n) \in \mathbb{R}^{n\times n}$ diagonal and nonsingular:

- Assume that $A = XDX^T$ satisfies $\dfrac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}} = O(\epsilon)$ for all $i > j$.
- $a_{ii} = \displaystyle\sum_{k=1}^{n} x_{ik}^2 d_k$
- $$\left|\frac{\mathtt{fl}(a_{ii}) - a_{ii}}{a_{ii}}\right| \leq \frac{(n+1)\epsilon}{1-(n+1)\epsilon} \; \frac{\displaystyle\sum_{k=1}^{n} x_{ik}^2 |d_k|}{\left|\displaystyle\sum_{k=1}^{n} x_{ik}^2 d_k\right|}$$

Errors on diagonal entries of almost diagonal RRDs (I)

Given $X \in \mathbb{R}^{n\times n}$ nonsingular and $D = \operatorname{diag}(d_1, \ldots, d_n) \in \mathbb{R}^{n\times n}$ diagonal and nonsingular:

- Assume that $A = XDX^T$ satisfies $\frac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}} = O(\epsilon)$ for all $i > j$.
- $a_{ii} = \sum_{k=1}^{n} x_{ik}^2 d_k$
- $$\left| \frac{\mathtt{fl}(a_{ii}) - a_{ii}}{a_{ii}} \right| \leq \frac{(n+1)\epsilon}{1-(n+1)\epsilon} \frac{\sum_{k=1}^{n} x_{ik}^2 |d_k|}{\left| \sum_{k=1}^{n} x_{ik}^2 d_k \right|}$$

Errors on diagonal entries of almost diagonal RRDs (II): EXAMPLE

INPUT: $\kappa(X) = 7.21$

$$XDX^T = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 10^{50} & & \\ & 1 & \\ & & -10^{50} \end{bmatrix} X^T$$

RUNNING IMPLICIT JACOBI UNTIL CONVERGENCE

$$X_f D X_f^T = \begin{bmatrix} 4.79 \cdot 10^{-48} & 5.35 \cdot 10^{-1} & 2.04 \cdot 10^{-47} \\ 3.8 \cdot 10^{-1} & 4.03 \cdot 10^{-2} & 1.64 \\ 2.42 & 1.65 & 5.67 \cdot 10^{-1} \end{bmatrix} \begin{bmatrix} 10^{50} & & \\ & 1 & \\ & & -10^{50} \end{bmatrix} X_f^T$$

$$= \begin{bmatrix} 2.86 \cdot 10^{-1} & -3.16 \cdot 10^{3} & 2.39 \cdot 10^{-3} \\ -3.16 \cdot 10^{3} & -2.53 \cdot 10^{50} & 1.04 \cdot 10^{34} \\ 2.39 \cdot 10^{-3} & 2.08 \cdot 10^{34} & 5.53 \cdot 10^{50} \end{bmatrix}$$

THERE IS NO CANCELLATION

$$2.86 \cdot 10^{-1} = (4.79 \cdot 10^{-48})^2 \times 10^{50} + (5.35 \cdot 10^{-1})^2 \times 1 + (2.04 \cdot 10^{-47})^2 \times (-10^{50})$$

$$= 2.29 \cdot 10^{-45} + 2.86 \cdot 10^{-1} - 4.18 \cdot 10^{-44}$$

Errors on diagonal entries of almost diagonal RRDs (II): EXAMPLE

INPUT: $\kappa(X) = 7.21$

$$XDX^T = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 10^{50} & & \\ & 1 & \\ & & -10^{50} \end{bmatrix} X^T$$

RUNNING IMPLICIT JACOBI UNTIL CONVERGENCE

$$\begin{aligned} X_f D X_f^T &= \begin{bmatrix} 4.79 \cdot 10^{-48} & 5.35 \cdot 10^{-1} & 2.04 \cdot 10^{-47} \\ 3.8 \cdot 10^{-1} & 4.03 \cdot 10^{-2} & 1.64 \\ 2.42 & 1.65 & 5.67 \cdot 10^{-1} \end{bmatrix} \begin{bmatrix} 10^{50} & & \\ & 1 & \\ & & -10^{50} \end{bmatrix} X_f^T \\ &= \begin{bmatrix} 2.86 \cdot 10^{-1} & -3.16 \cdot 10^{3} & 2.39 \cdot 10^{-3} \\ -3.16 \cdot 10^{3} & -2.53 \cdot 10^{50} & 1.04 \cdot 10^{34} \\ 2.39 \cdot 10^{-3} & 2.08 \cdot 10^{34} & 5.53 \cdot 10^{50} \end{bmatrix} \end{aligned}$$

THERE IS NO CANCELLATION

$$\begin{aligned} 2.86 \cdot 10^{-1} &= (4.79 \cdot 10^{-48})^2 \times 10^{50} + (5.35 \cdot 10^{-1})^2 \times 1 + (2.04 \cdot 10^{-47})^2 \times (-10^{50}) \\ &= 2.29 \cdot 10^{-45} + 2.86 \cdot 10^{-1} - 4.18 \cdot 10^{-44} \end{aligned}$$

Errors on diagonal entries of almost diagonal RRDs (II): EXAMPLE

INPUT: $\kappa(X) = 7.21$

$$XDX^T = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 10^{50} & & \\ & 1 & \\ & & -10^{50} \end{bmatrix} X^T$$

RUNNING IMPLICIT JACOBI UNTIL CONVERGENCE

$$X_f D X_f^T = \begin{bmatrix} 4.79 \cdot 10^{-48} & 5.35 \cdot 10^{-1} & 2.04 \cdot 10^{-47} \\ 3.8 \cdot 10^{-1} & 4.03 \cdot 10^{-2} & 1.64 \\ 2.42 & 1.65 & 5.67 \cdot 10^{-1} \end{bmatrix} \begin{bmatrix} 10^{50} & & \\ & 1 & \\ & & -10^{50} \end{bmatrix} X_f^T$$

$$= \begin{bmatrix} 2.86 \cdot 10^{-1} & -3.16 \cdot 10^{3} & 2.39 \cdot 10^{-3} \\ -3.16 \cdot 10^{3} & -2.53 \cdot 10^{50} & 1.04 \cdot 10^{34} \\ 2.39 \cdot 10^{-3} & 2.08 \cdot 10^{34} & 5.53 \cdot 10^{50} \end{bmatrix}$$

THERE IS NO CANCELLATION

$$\begin{aligned} 2.86 \cdot 10^{-1} &= (4.79 \cdot 10^{-48})^2 \times 10^{50} + (5.35 \cdot 10^{-1})^2 \times 1 + (2.04 \cdot 10^{-47})^2 \times (-10^{50}) \\ &= 2.29 \cdot 10^{-45} + 2.86 \cdot 10^{-1} - 4.18 \cdot 10^{-44} \end{aligned}$$

Errors on diagonal entries of almost diagonal RRDs (II): EXAMPLE

INPUT: $\kappa(X) = 7.21$

$$XDX^T = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 10^{50} & & \\ & 1 & \\ & & -10^{50} \end{bmatrix} X^T$$

RUNNING IMPLICIT JACOBI UNTIL CONVERGENCE

$$X_f D X_f^T = \begin{bmatrix} 4.79 \cdot 10^{-48} & 5.35 \cdot 10^{-1} & 2.04 \cdot 10^{-47} \\ 3.8 \cdot 10^{-1} & 4.03 \cdot 10^{-2} & 1.64 \\ 2.42 & 1.65 & 5.67 \cdot 10^{-1} \end{bmatrix} \begin{bmatrix} 10^{50} & & \\ & 1 & \\ & & -10^{50} \end{bmatrix} X_f^T$$

$$= \begin{bmatrix} 2.86 \cdot 10^{-1} & -3.16 \cdot 10^{3} & 2.39 \cdot 10^{-3} \\ -3.16 \cdot 10^{3} & -2.53 \cdot 10^{50} & 1.04 \cdot 10^{34} \\ 2.39 \cdot 10^{-3} & 2.08 \cdot 10^{34} & 5.53 \cdot 10^{50} \end{bmatrix}$$

THERE IS NO CANCELLATION

$$2.86 \cdot 10^{-1} = (4.79 \cdot 10^{-48})^2 \times 10^{50} + (5.35 \cdot 10^{-1})^2 \times 1 + (2.04 \cdot 10^{-47})^2 \times (-10^{50})$$

$$= 2.29 \cdot 10^{-45} + 2.86 \cdot 10^{-1} - 4.18 \cdot 10^{-44}$$

Errors on diagonal entries of almost diagonal RRDs (II): EXAMPLE

INPUT: $\kappa(X) = 7.21$

$$XDX^T = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & 1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 10^{50} & & \\ & 1 & \\ & & -10^{50} \end{bmatrix} X^T$$

RUNNING IMPLICIT JACOBI UNTIL CONVERGENCE

$$\begin{aligned} X_f D X_f^T &= \begin{bmatrix} 4.79 \cdot 10^{-48} & 5.35 \cdot 10^{-1} & 2.04 \cdot 10^{-47} \\ 3.8 \cdot 10^{-1} & 4.03 \cdot 10^{-2} & 1.64 \\ 2.42 & 1.65 & 5.67 \cdot 10^{-1} \end{bmatrix} \begin{bmatrix} 10^{50} & & \\ & 1 & \\ & & -10^{50} \end{bmatrix} X_f^T \\ &= \begin{bmatrix} 2.86 \cdot 10^{-1} & -3.16 \cdot 10^{3} & 2.39 \cdot 10^{-3} \\ -3.16 \cdot 10^{3} & -2.53 \cdot 10^{50} & 1.04 \cdot 10^{34} \\ 2.39 \cdot 10^{-3} & 2.08 \cdot 10^{34} & 5.53 \cdot 10^{50} \end{bmatrix} \end{aligned}$$

THERE IS NO CANCELLATION

$$\begin{aligned} 2.86 \cdot 10^{-1} &= (4.79 \cdot 10^{-48})^2 \times 10^{50} + (5.35 \cdot 10^{-1})^2 \times 1 + (2.04 \cdot 10^{-47})^2 \times (-10^{50}) \\ &= 2.29 \cdot 10^{-45} + 2.86 \cdot 10^{-1} - 4.18 \cdot 10^{-44} \end{aligned}$$

Errors on diagonal entries of almost diagonal RRDs (III): THE MAIN THEOREM

Theorem

Let $X, D \in \mathbb{R}^{n\times n}$ *be nonsingular and* $D = \mathrm{diag}(d_1, \ldots, d_n)$ *be diagonal. If the matrix* $A \equiv XDX^T$ *satisfies* $a_{ii} = \sum_{k=1}^{n} x_{ik}^2 d_k \neq 0$ *for all* i*, and*

$$\frac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}} \leq \delta, \quad \textit{for all } \; i \neq j, \quad \textit{where } \delta \leq \frac{1}{5n}, \textit{ then}$$

$$\frac{\sum_{k=1}^{n} x_{ik}^2 |d_k|}{|a_{ii}|} \leq \frac{\kappa(X)}{1-2n\delta}\left(1 + \frac{2\,n^{5/2}\delta}{1-n\delta} + n^2\left(\frac{n\delta}{1-n\delta}\right)^2\right), \qquad i = 1, \ldots, n.$$

$$\frac{\sum_{k=1}^{n} x_{ik}^2 |d_k|}{|a_{ii}|} \leq \kappa(X)\left(1 + O(n^{5/2}\delta)\right), \qquad i = 1, \ldots, n.$$

Errors on diagonal entries of almost diagonal RRDs (III): THE MAIN THEOREM

Theorem

Let $X, D \in \mathbb{R}^{n\times n}$ *be nonsingular and* $D = \mathrm{diag}(d_1, \ldots, d_n)$ *be diagonal. If the matrix* $A \equiv XDX^T$ *satisfies* $a_{ii} = \sum_{k=1}^{n} x_{ik}^2 d_k \neq 0$ *for all* i*, and*

$$\frac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}} \leq \delta, \quad \text{for all } i \neq j, \quad \text{where } \delta \leq \frac{1}{5n}, \text{ then}$$

$$\frac{\sum_{k=1}^{n} x_{ik}^2 |d_k|}{|a_{ii}|} \leq \frac{\kappa(X)}{1-2n\delta} \left(1 + \frac{2\,n^{5/2}\delta}{1-n\delta} + n^2 \left(\frac{n\delta}{1-n\delta}\right)^2\right), \qquad i = 1, \ldots, n.$$

$$\frac{\sum_{k=1}^{n} x_{ik}^2 |d_k|}{|a_{ii}|} \leq \kappa(X) \left(1 + O(n^{5/2}\delta)\right), \qquad i = 1, \ldots, n.$$

Errors on diagonal entries of almost diagonal RRDs (III): THE MAIN THEOREM

Theorem

Let $X, D \in \mathbb{R}^{n\times n}$ *be nonsingular and* $D = \mathrm{diag}(d_1, \ldots, d_n)$ *be diagonal. If the matrix* $A \equiv XDX^T$ *satisfies* $a_{ii} = \sum_{k=1}^{n} x_{ik}^2 d_k \neq 0$ *for all* i*, and*

$$\frac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}} \leq \delta, \quad \text{for all } i \neq j, \quad \text{where } \delta \leq \frac{1}{5n}, \text{ then}$$

$$\frac{\sum_{k=1}^{n} x_{ik}^2 |d_k|}{|a_{ii}|} \leq \frac{\kappa(X)}{1-2n\delta}\left(1 + \frac{2\,n^{5/2}\delta}{1-n\delta} + n^2\left(\frac{n\delta}{1-n\delta}\right)^2\right), \qquad i = 1, \ldots, n.$$

$$\frac{\sum_{k=1}^{n} x_{ik}^2 |d_k|}{|a_{ii}|} \leq \kappa(X)\left(1 + O(n^{5/2}\delta)\right), \qquad i = 1, \ldots, n.$$

Errors on diagonal entries of almost diagonal RRDs (IV): Corollary

Corollary

If $A = XDX^T$ *satisfies the stopping criterion then*

$$\left| \frac{\mathtt{fl}(a_{ii}) - a_{ii}}{a_{ii}} \right| \leq (n+1)\,\epsilon\,\kappa(X) + O(\kappa(X)\,\epsilon^2)$$

Key idea in the proof of THE MAIN THEOREM

Proof by contradiction

- $A = XDX^T$ is close to diagonal, then its diagonal entries are close to its eigenvalues.
- Assume
$$\frac{\sum_{k=1}^n x_{ik}^2 |d_k|}{|a_{ii}|} = \frac{\sum_{k=1}^n x_{ik}^2 |d_k|}{|\sum_{k=1}^n x_{ik}^2 d_k|} >> \kappa(X)$$
- Then there are perturbations $\widetilde{d}_k = d_k(1+\delta_k)$, $|\delta_k| < \beta << 1$ such that $(X\widetilde{D}X^T)_{ii} = \sum_{k=1}^n x_{ik}^2 \widetilde{d}_k$, satisfy
$$\frac{|a_{ii} - (X\widetilde{D}X^T)_{ii}|}{|a_{ii}|} >> \beta\kappa(X).$$
- This is in contradiction with an RRD determining accurately its eigenvalues.

Key idea in the proof of THE MAIN THEOREM

Proof by contradiction

- $A = XDX^T$ is close to diagonal, then its diagonal entries are close to its eigenvalues.
- Assume
$$\frac{\sum_{k=1}^{n} x_{ik}^2 |d_k|}{|a_{ii}|} = \frac{\sum_{k=1}^{n} x_{ik}^2 |d_k|}{|\sum_{k=1}^{n} x_{ik}^2 d_k|} >> \kappa(X)$$
- Then there are perturbations $\widetilde{d}_k = d_k(1+\delta_k)$, $|\delta_k| < \beta << 1$ such that $(X\widetilde{D}X^T)_{ii} = \sum_{k=1}^{n} x_{ik}^2 \widetilde{d}_k$, satisfy
$$\frac{|a_{ii} - (X\widetilde{D}X^T)_{ii}|}{|a_{ii}|} >> \beta\kappa(X).$$
- This is in contradiction with an RRD determining accurately its eigenvalues.

Key idea in the proof of THE MAIN THEOREM

Proof by contradiction

- $A = XDX^T$ is close to diagonal, then its diagonal entries are close to its eigenvalues.
- Assume
$$\frac{\sum_{k=1}^n x_{ik}^2 |d_k|}{|a_{ii}|} = \frac{\sum_{k=1}^n x_{ik}^2 |d_k|}{|\sum_{k=1}^n x_{ik}^2 d_k|} >> \kappa(X)$$
- Then there are perturbations $\widetilde{d}_k = d_k(1+\delta_k)$, $|\delta_k| < \beta << 1$ such that $(X\widetilde{D}X^T)_{ii} = \sum_{k=1}^n x_{ik}^2 \widetilde{d}_k$, satisfy
$$\frac{|a_{ii} - (X\widetilde{D}X^T)_{ii}|}{|a_{ii}|} >> \beta\kappa(X).$$
- This is in contradiction with an RRD determining accurately its eigenvalues.

Key idea in the proof of THE MAIN THEOREM

Proof by contradiction

- $A = XDX^T$ is close to diagonal, then its diagonal entries are close to its eigenvalues.
- Assume

$$\frac{\sum_{k=1}^{n} x_{ik}^2 |d_k|}{|a_{ii}|} = \frac{\sum_{k=1}^{n} x_{ik}^2 |d_k|}{|\sum_{k=1}^{n} x_{ik}^2 d_k|} >> \kappa(X)$$

- Then there are perturbations $\widetilde{d}_k = d_k(1+\delta_k)$, $|\delta_k| < \beta << 1$ such that $(X\widetilde{D}X^T)_{ii} = \sum_{k=1}^{n} x_{ik}^2 \widetilde{d}_k$, satisfy

$$\frac{|a_{ii} - (X\widetilde{D}X^T)_{ii}|}{|a_{ii}|} >> \beta\kappa(X).$$

- This is in contradiction with an RRD determining accurately its eigenvalues.

Outline

Implicit Jacobi is multiplicative backward stable

Theorem

Let N be the **number of rotations** *applied by implicit Jacobi on $A = XDX^T$ until convergence, and $\widehat{\Lambda}$ and $\widehat{U}$ be the computed matrices of eigenvalues and eigenvectors. Then there exists an* **exact orthogonal** *matrix $U \in \mathbb{R}^{n \times n}$ such that*

$$U \widehat{\Lambda} U^T = (I + E)\, XDX^T \,(I + E)^T,$$

with

$$\|E\|_F = O(\epsilon\, N\, \kappa(X)) \quad \text{and} \quad \|\widehat{U} - U\|_F = O(N\, \epsilon).$$

Corollary (Forward errors in e-values and e-vectors)

$$\frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} \leq O(\epsilon\, N\, \kappa(X)) \quad \text{and} \quad \theta(v_i, \hat{v}_i) \leq \frac{O(\epsilon\, N\, \kappa(X))}{\min_{j \neq i} \left| \frac{\lambda_i - \lambda_j}{\lambda_i} \right|} \quad \textit{for all} \quad i,$$

Implicit Jacobi is multiplicative backward stable

Theorem

Let N *be the* **number of rotations** *applied by implicit Jacobi on* $A = XDX^T$ *until convergence, and* $\widehat{\Lambda}$ *and* $\widehat{U}$ *be the computed matrices of eigenvalues and eigenvectors. Then there exists an* **exact orthogonal** *matrix* $U \in \mathbb{R}^{n\times n}$ *such that*

$$U\widehat{\Lambda}U^T = (I+E)\, XDX^T \,(I+E)^T,$$

with

$$\|E\|_F = O(\epsilon\, N\, \kappa(X)) \quad \textit{and} \quad \|\widehat{U} - U\|_F = O(N\,\epsilon).$$

Corollary (Forward errors in e-values and e-vectors)

$$\frac{|\hat{\lambda}_i - \lambda_i|}{|\lambda_i|} \leq O(\epsilon\, N\, \kappa(X)) \quad \text{and} \quad \theta(v_i, \hat{v}_i) \leq \frac{O(\epsilon\, N\, \kappa(X))}{\min_{j\neq i} \left|\frac{\lambda_i - \lambda_j}{\lambda_i}\right|} \quad \textit{for all} \quad i,$$

Technical comments

To establish the backward error result, we need to prove that

- The stopping criterion in finite arithmetic on $A = X_f D X_f^T$ gives *exact* information, i.e.,

$$\texttt{fl}\left(\frac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}}\right) \leq \epsilon\,\kappa(X) \Longrightarrow \frac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}} \leq n\,\epsilon\,\kappa(X) + O(\epsilon^2)$$

  for all $i \neq j$, which is the case if there is no cancellation in $fl(a_{ii})$.

- The stopping criterion introduces small multiplicative backward errors, i.e.,

$$\mathrm{diag}(\texttt{fl}(a_{11}), \ldots, \texttt{fl}(a_{nn})) = (I + F) X_f D X_f^T (I + F)^T,$$

  where $\|F\|_F = O(n^2\,\epsilon\,\kappa(X))$.

Technical comments

To establish the backward error result, we need to prove that

- The stopping criterion in finite arithmetic on $A = X_f D X_f^T$ gives *exact* information, i.e.,

$$\texttt{fl}\left(\frac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}}\right) \leq \epsilon\,\kappa(X) \Longrightarrow \frac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}} \leq n\,\epsilon\,\kappa(X) + O(\epsilon^2)$$

  for all $i \neq j$, which is the case if there is no cancellation in $fl(a_{ii})$.

- The stopping criterion introduces small multiplicative backward errors, i.e.,

$$\mathrm{diag}(\texttt{fl}(a_{11}), \ldots, \texttt{fl}(a_{nn})) = (I+F) X_f D X_f^T (I+F)^T,$$

  where $\|F\|_F = O(n^2\,\epsilon\,\kappa(X))$.

Technical comments

To establish the backward error result, we need to prove that

- The stopping criterion in finite arithmetic on $A = X_f D X_f^T$ gives *exact* information, i.e.,

$$\mathtt{fl}\left(\frac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}}\right) \leq \epsilon\,\kappa(X) \implies \frac{|a_{ij}|}{\sqrt{|a_{ii}a_{jj}|}} \leq n\,\epsilon\,\kappa(X) + O(\epsilon^2)$$

  for all $i \neq j$, which is the case if there is no cancellation in $fl(a_{ii})$.

- The stopping criterion introduces small multiplicative backward errors, i.e.,

$$\mathrm{diag}(\mathtt{fl}(a_{11}), \ldots, \mathtt{fl}(a_{nn})) = (I+F) X_f D X_f^T (I+F)^T,$$

  where $\|F\|_F = O(n^2\,\epsilon\,\kappa(X))$.

Outline

Rectangular RRDs

- So far we have considered $A = XDX^T$ with **square and nonsingular** X and D, which excludes singular matrices A.
- If we insist on X being nonsingular, then A is singular if and only if D is singular.
- The zero eigenvalues of A are revealed by the zero diagonal entries of D
- Discarding these entries we get

$$A = XDX^T \in \mathbb{R}^{n\times n} \quad \text{where} \quad X \in \mathbb{R}^{n\times r} \quad D \in \mathbb{R}^{r\times r},$$

  with $n > r$, X with full rank, and D nonsingular.
- Implicit Jacobi converges to an $n \times n$ diagonal matrix with zero entries and cancellation is unavoidable.

Rectangular RRDs

- So far we have considered $A = XDX^T$ with **square and nonsingular** X and D, which excludes singular matrices A.
- If we insist on X being nonsingular, then A is singular if and only if D is singular.
- The zero eigenvalues of A are revealed by the zero diagonal entries of D
- Discarding these entries we get

$$A = XDX^T \in \mathbb{R}^{n\times n} \quad \text{where} \quad X \in \mathbb{R}^{n\times r} \quad D \in \mathbb{R}^{r\times r},$$

  with $n > r$, X with full rank, and D nonsingular.
- Implicit Jacobi converges to an $n \times n$ diagonal matrix with zero entries and cancellation is unavoidable.

Rectangular RRDs

- So far we have considered $A = XDX^T$ with **square and nonsingular** X and D, which excludes singular matrices A.
- If we insist on X being nonsingular, then A is singular if and only if D is singular.
- The zero eigenvalues of A are revealed by the zero diagonal entries of D
- Discarding these entries we get

$$A = XDX^T \in \mathbb{R}^{n \times n} \quad \text{where} \quad X \in \mathbb{R}^{n \times r} \quad D \in \mathbb{R}^{r \times r},$$

  with $n > r$, X with full rank, and D nonsingular.
- Implicit Jacobi converges to an $n \times n$ diagonal matrix with zero entries and cancellation is unavoidable.

Rectangular RRDs

- So far we have considered $A = XDX^T$ with **square and nonsingular** X and D, which excludes singular matrices A.
- If we insist on X being nonsingular, then A is singular if and only if D is singular.
- The zero eigenvalues of A are revealed by the zero diagonal entries of D
- Discarding these entries we get

$$A = XDX^T \in \mathbb{R}^{n \times n} \quad \text{where} \quad X \in \mathbb{R}^{n \times r} \quad D \in \mathbb{R}^{r \times r},$$

  with $n > r$, X with full rank, and D nonsingular.
- Implicit Jacobi converges to an $n \times n$ diagonal matrix with zero entries and cancellation is unavoidable.

Rectangular RRDs

- So far we have considered $A = XDX^T$ with **square and nonsingular** X and D, which excludes singular matrices A.
- If we insist on X being nonsingular, then A is singular if and only if D is singular.
- The zero eigenvalues of A are revealed by the zero diagonal entries of D
- Discarding these entries we get

$$A = XDX^T \in \mathbb{R}^{n\times n} \quad \text{where} \quad X \in \mathbb{R}^{n\times r} \quad D \in \mathbb{R}^{r\times r},$$

  with $n > r$, X with full rank, and D nonsingular.
- Implicit Jacobi converges to an $n \times n$ diagonal matrix with zero entries and cancellation is unavoidable.

Algorithm for rectangular RRD $A = XDX^T$

$$A = XDX^T \in \mathbb{R}^{n\times n} \quad \text{with} \quad X \in \mathbb{R}^{n\times r}, \quad D \in \mathbb{R}^{r\times r},$$

1. Compute full QR factorization of X

$$Q \begin{bmatrix} R \\ 0 \end{bmatrix} = X \quad \text{where} \quad Q \in \mathbb{R}^{n\times n}, \quad R \in \mathbb{R}^{r\times r}$$

2. Note that

$$A = Q \begin{bmatrix} RDR^T & 0 \\ 0 & 0 \end{bmatrix} Q^T$$

3. Apply Implicit Jacobi on RDR^T (**with factors square and nonsingular**) to compute
   1. Nonzero eigenvalues of A: $\lambda_1, \ldots, \lambda_r$.
   2. Eigenvector matrix of RDR^T: U_R
4. $[Q(:,1:r)U_R \mid Q(:,r+1:n)]$ is the eigenvector matrix of A.

Algorithm for rectangular RRD $A = XDX^T$

$$A = XDX^T \in \mathbb{R}^{n\times n} \quad \text{with} \quad X \in \mathbb{R}^{n\times r}, \quad D \in \mathbb{R}^{r\times r},$$

1. Compute full QR factorization of X

$$Q \begin{bmatrix} R \\ 0 \end{bmatrix} = X \quad \text{where} \quad Q \in \mathbb{R}^{n\times n}, \quad R \in \mathbb{R}^{r\times r}$$

2. Note that

$$A = Q \begin{bmatrix} RDR^T & 0 \\ 0 & 0 \end{bmatrix} Q^T$$

3. Apply Implicit Jacobi on RDR^T (**with factors square and nonsingular**) to compute
   1. Nonzero eigenvalues of A: $\lambda_1, \ldots, \lambda_r$.
   2. Eigenvector matrix of RDR^T: U_R
4. $[Q(:, 1:r)U_R \mid Q(:, r+1:n)]$ is the eigenvector matrix of A.

Algorithm for rectangular RRD $A = XDX^T$

$$A = XDX^T \in \mathbb{R}^{n\times n} \quad \text{with} \quad X \in \mathbb{R}^{n\times r}, \quad D \in \mathbb{R}^{r\times r},$$

1. Compute full QR factorization of X

$$Q \begin{bmatrix} R \\ 0 \end{bmatrix} = X \quad \text{where} \quad Q \in \mathbb{R}^{n\times n}, \quad R \in \mathbb{R}^{r\times r}$$

2. Note that

$$A = Q \begin{bmatrix} RDR^T & 0 \\ 0 & 0 \end{bmatrix} Q^T$$

3. Apply Implicit Jacobi on RDR^T (**with factors square and nonsingular**) to compute
   1. Nonzero eigenvalues of A: $\lambda_1, \ldots, \lambda_r$.
   2. Eigenvector matrix of RDR^T: U_R
4. $[Q(:, 1:r)U_R \mid Q(:, r+1:n)]$ is the eigenvector matrix of A.

Algorithm for rectangular RRD $A = XDX^T$

$$A = XDX^T \in \mathbb{R}^{n\times n} \quad \text{with} \quad X \in \mathbb{R}^{n\times r}, \quad D \in \mathbb{R}^{r\times r},$$

1. Compute full QR factorization of X

$$Q \begin{bmatrix} R \\ 0 \end{bmatrix} = X \quad \text{where} \quad Q \in \mathbb{R}^{n\times n}, \quad R \in \mathbb{R}^{r\times r}$$

2. Note that

$$A = Q \begin{bmatrix} RDR^T & 0 \\ 0 & 0 \end{bmatrix} Q^T$$

3. Apply Implicit Jacobi on RDR^T (**with factors square and nonsingular**) to compute
    1. Nonzero eigenvalues of A: $\lambda_1, \ldots, \lambda_r$.
    2. Eigenvector matrix of RDR^T: U_R
4. $[Q(:, 1:r)U_R \mid Q(:, r+1:n)]$ is the eigenvector matrix of A.

Algorithm for rectangular RRD $A = XDX^T$

$$A = XDX^T \in \mathbb{R}^{n\times n} \quad \text{with} \quad X \in \mathbb{R}^{n\times r}, \quad D \in \mathbb{R}^{r\times r},$$

1. Compute full QR factorization of X

$$Q \begin{bmatrix} R \\ 0 \end{bmatrix} = X \quad \text{where} \quad Q \in \mathbb{R}^{n\times n}, \quad R \in \mathbb{R}^{r\times r}$$

2. Note that

$$A = Q \begin{bmatrix} RDR^T & 0 \\ 0 & 0 \end{bmatrix} Q^T$$

3. Apply Implicit Jacobi on RDR^T (**with factors square and nonsingular**) to compute
   1. Nonzero eigenvalues of A: $\lambda_1, \ldots, \lambda_r$.
   2. Eigenvector matrix of RDR^T: U_R
4. $[Q(:,1:r)U_R \mid Q(:,r+1:n)]$ is the eigenvector matrix of A.

Algorithm for rectangular RRD $A = XDX^T$

$$A = XDX^T \in \mathbb{R}^{n \times n} \quad \text{with} \quad X \in \mathbb{R}^{n \times r}, \quad D \in \mathbb{R}^{r \times r},$$

1. Compute full QR factorization of X

$$Q \begin{bmatrix} R \\ 0 \end{bmatrix} = X \quad \text{where} \quad Q \in \mathbb{R}^{n \times n}, \quad R \in \mathbb{R}^{r \times r}$$

2. Note that

$$A = Q \begin{bmatrix} RDR^T & 0 \\ 0 & 0 \end{bmatrix} Q^T$$

3. Apply Implicit Jacobi on RDR^T (**with factors square and nonsingular**) to compute
   1. Nonzero eigenvalues of A: $\lambda_1, \ldots, \lambda_r$.
   2. Eigenvector matrix of RDR^T: U_R
4. $[Q(:, 1:r)U_R \mid Q(:, r+1:n)]$ is the eigenvector matrix of A.

Outline

Numerical Experiments

- Thousands of numerical experiments confirm the high relative accuracy that we have rigorously proven.
- Traditional Jacobi is **slow**, then Implicit Jacobi is **slow**.
- **Speed is not our main issue**, but we have compared the number of sweeps performed by Implicit Jacobi with respect other high relative accuracy algorithms:
  1. One sided Hyperbolic Jacobi (Slapničar-Veselić): **not** rigorous bounds.
  2. SSVD-l (D-Molera-Moro): **not** rigorous bounds.
  3. SSVD-r (D-Molera-Moro): rigorous bounds.
- We have used `gallery('randsvd',...)` by N. Higham in MATLAB to generate random RRDs with X well-conditioned and D indefinite and **extremely** ill-conditioned.

Numerical Experiments

- Thousands of numerical experiments confirm the high relative accuracy that we have rigorously proven.
- Traditional Jacobi is **slow**, then Implicit Jacobi is **slow**.
- **Speed is not our main issue**, but we have compared the number of sweeps performed by Implicit Jacobi with respect other high relative accuracy algorithms:
  1. One sided Hyperbolic Jacobi (Slapničar-Veselić): **not** rigorous bounds.
  2. SSVD-l (D-Molera-Moro): **not** rigorous bounds.
  3. SSVD-r (D-Molera-Moro): rigorous bounds.
- We have used `gallery('randsvd',...)` by N. Higham in MATLAB to generate random RRDs with X well-conditioned and D indefinite and **extremely** ill-conditioned.

Numerical Experiments

- Thousands of numerical experiments confirm the high relative accuracy that we have rigorously proven.
- Traditional Jacobi is **slow**, then Implicit Jacobi is **slow**.
- **Speed is not our main issue**, but we have compared the number of sweeps performed by Implicit Jacobi with respect other high relative accuracy algorithms:
  1. One sided Hyperbolic Jacobi (Slapničar-Veselić): **not** rigorous bounds.
  2. SSVD-l (D-Molera-Moro): **not** rigorous bounds.
  3. SSVD-r (D-Molera-Moro): rigorous bounds.
- We have used `gallery('randsvd',...)` by N. Higham in MATLAB to generate random RRDs with X well-conditioned and D indefinite and **extremely** ill-conditioned.

Numerical Experiments

- Thousands of numerical experiments confirm the high relative accuracy that we have rigorously proven.
- Traditional Jacobi is **slow**, then Implicit Jacobi is **slow**.
- **Speed is not our main issue**, but we have compared the number of sweeps performed by Implicit Jacobi with respect other high relative accuracy algorithms:
  1. One sided Hyperbolic Jacobi (Slapničar-Veselić): **not** rigorous bounds.
  2. SSVD-l (D-Molera-Moro): **not** rigorous bounds.
  3. SSVD-r (D-Molera-Moro): rigorous bounds.
- We have used `gallery('randsvd',...)` by N. Higham in MATLAB to generate random RRDs with X well-conditioned and D indefinite and **extremely** ill-conditioned.

Numerical Experiments

- Thousands of numerical experiments confirm the high relative accuracy that we have rigorously proven.
- Traditional Jacobi is **slow**, then Implicit Jacobi is **slow**.
- **Speed is not our main issue**, but we have compared the number of sweeps performed by Implicit Jacobi with respect other high relative accuracy algorithms:
  1. One sided Hyperbolic Jacobi (Slapničar-Veselić): **not** rigorous bounds.
  2. SSVD-I (D-Molera-Moro): **not** rigorous bounds.
  3. SSVD-r (D-Molera-Moro): rigorous bounds.
- We have used `gallery('randsvd',...)` by N. Higham in MATLAB to generate random RRDs with X well-conditioned and D indefinite and **extremely** ill-conditioned.

Numerical Experiments

- Thousands of numerical experiments confirm the high relative accuracy that we have rigorously proven.
- Traditional Jacobi is **slow**, then Implicit Jacobi is **slow**.
- **Speed is not our main issue**, but we have compared the number of sweeps performed by Implicit Jacobi with respect other high relative accuracy algorithms:
  1. One sided Hyperbolic Jacobi (Slapničar-Veselić): **not** rigorous bounds.
  2. SSVD-I (D-Molera-Moro): **not** rigorous bounds.
  3. SSVD-r (D-Molera-Moro): rigorous bounds.
- We have used `gallery('randsvd',...)` by N. Higham in MATLAB to generate random RRDs with X well-conditioned and D indefinite and **extremely** ill-conditioned.

Numerical Experiments

- Thousands of numerical experiments confirm the high relative accuracy that we have rigorously proven.
- Traditional Jacobi is **slow**, then Implicit Jacobi is **slow**.
- **Speed is not our main issue**, but we have compared the number of sweeps performed by Implicit Jacobi with respect other high relative accuracy algorithms:
  1. One sided Hyperbolic Jacobi (Slapničar-Veselić): **not** rigorous bounds.
  2. SSVD-l (D-Molera-Moro): **not** rigorous bounds.
  3. SSVD-r (D-Molera-Moro): rigorous bounds.
- We have used `gallery('randsvd',...)` by N. Higham in MATLAB to generate random RRDs with X well-conditioned and D indefinite and **extremely** ill-conditioned.

Number of sweeps: Increasing $\kappa(D)$ (I)

In all of these tests $\kappa(X) = 30$ and X, D are 100×100.

D has one entry with magnitude 1 and the rest $1/\kappa(D)$

$\kappa(D)$	Imp. Jac.	Hyp. Jac.	SSVD-l	SSVD-r
10^{10}	10	10.8	10	13
10^{30}	10	10.6	9.8	13.2
10^{50}	10.8	10.8	10	14
10^{70}	11	11	10.2	13.6
10^{90}	10.8	10.6	10	13.8
10^{110}	11	10.4	10	14.8

Number of sweeps: Increasing $\kappa(D)$ (II)

In all of these tests $\kappa(X) = 30$ and X, D are 100×100.

D has entries with magnitudes geometrically distributed

$\kappa(D)$	Imp. Jac.	Hyp. Jac.	SSVD-l	SSVD-r
10^{10}	16	9	6.2	27.2
10^{30}	24.8	9	4.8	39.6
10^{50}	32.4	9	4.4	47.2
10^{70}	35.8	9.4	4.4	52.6
10^{90}	40	9	4	57
10^{110}	43.2	9	3	59.6

$$|d_i| = \kappa(D)^{\frac{i-1}{n-1}}, \quad i = 1, \ldots, n$$

Number of sweeps: Increasing the dimension of the RRD (I)

In all of these tests $\kappa(X) = 100$, $\kappa(D) = 10^{40}$, and X, D are $n \times n$.

D has one entry with magnitude 1 and the rest $1/\kappa(D)$

n	Imp. Jac.	Hyp. Jac.	SSVD-l	SSVD-r
100	11	11.4	10.2	15.8
500	13	13.4	14	18
1000	13	14	15	19
2000	14	15	16	20

Number of sweeps: Increasing the dimension of the RRD (II)

In all of these tests $\kappa(X) = 100$, $\kappa(D) = 10^{40}$, and X, D are $n \times n$.

D has entries with magnitudes geometrically distributed

n	Imp. Jac.	Hyp. Jac.	SSVD-l	SSVD-r
100	28.8	10	4.6	44.6
500	46	11	6	87
1000	58	11	7	> 100
2000	68	11	7	> 100

Numerical Experiments: Conclusions

- The comparison of the performance of the available high relative accuracy algorithms for symmetric indefinite RRDs depends heavily on the distribution of the eigenvalues
- The new Implicit Jacobi is the fastest algorithm with guaranteed errors bounds (the other one is SSVD-r).
- The new Implicit Jacobi may be considerably slower than Hyperbolic Jacobi and SSVD-I, both with errors not rigorously bounded.
- The fastest one is SSVD-I that can benefit from new fast and accurate Jacobi SVD algorithm by Drmač and Veselić (2008).

Numerical Experiments: Conclusions

- The comparison of the performance of the available high relative accuracy algorithms for symmetric indefinite RRDs depends heavily on the distribution of the eigenvalues
- The new Implicit Jacobi is the fastest algorithm with guaranteed errors bounds (the other one is SSVD-r).
- The new Implicit Jacobi may be considerably slower than Hyperbolic Jacobi and SSVD-I, both with errors not rigorously bounded.
- The fastest one is SSVD-I that can benefit from new fast and accurate Jacobi SVD algorithm by Drmač and Veselić (2008).

Numerical Experiments: Conclusions

- The comparison of the performance of the available high relative accuracy algorithms for symmetric indefinite RRDs depends heavily on the distribution of the eigenvalues
- The new Implicit Jacobi is the fastest algorithm with guaranteed errors bounds (the other one is SSVD-r).
- The new Implicit Jacobi may be considerably slower than Hyperbolic Jacobi and SSVD-I, both with errors not rigorously bounded.
- The fastest one is SSVD-I that can benefit from new fast and accurate Jacobi SVD algorithm by Drmač and Veselić (2008).

Numerical Experiments: Conclusions

- The comparison of the performance of the available high relative accuracy algorithms for symmetric indefinite RRDs depends heavily on the distribution of the eigenvalues
- The new Implicit Jacobi is the fastest algorithm with guaranteed errors bounds (the other one is SSVD-r).
- The new Implicit Jacobi may be considerably slower than Hyperbolic Jacobi and SSVD-l, both with errors not rigorously bounded.
- The fastest one is SSVD-l that can benefit from new fast and accurate Jacobi SVD algorithm by Drmač and Veselić (2008).

Outline

Conclusions

- The implicit Jacobi algorithm on symmetric rank revealing factorizations

$$A = XDX^T$$

is the first algorithm that:

1. computes the eigenvalues and eigenvectors of A to high relative accuracy,
2. preserves the symmetry, and
3. uses only orthogonal transformations.

- In addition, the error bounds are rigorously proven, and are the **best possible ones** from the sensitivity of the problem.
- The implicit Jacobi algorithm is **very simple and natural**.
- The implicit Jacobi algorithm is **backward stable** in a strong **multiplicative** sense.
- More research to speed up the algorithm is needed.

Conclusions

- The implicit Jacobi algorithm on symmetric rank revealing factorizations
$$A = XDX^T$$
is the first algorithm that:
  1. computes the eigenvalues and eigenvectors of A to high relative accuracy,
  2. preserves the symmetry, and
  3. uses only orthogonal transformations.
- In addition, the error bounds are rigorously proven, and are the **best possible ones** from the sensitivity of the problem.
- The implicit Jacobi algorithm is **very simple and natural**.
- The implicit Jacobi algorithm is **backward stable** in a strong **multiplicative** sense.
- More research to speed up the algorithm is needed.

Conclusions

- The implicit Jacobi algorithm on symmetric rank revealing factorizations

$$A = XDX^T$$

is the first algorithm that:

1. computes the eigenvalues and eigenvectors of A to high relative accuracy,
2. preserves the symmetry, and
3. uses only orthogonal transformations.

- In addition, the error bounds are rigorously proven, and are the **best possible ones** from the sensitivity of the problem.
- The implicit Jacobi algorithm is **very simple and natural**.
- The implicit Jacobi algorithm is **backward stable** in a strong **multiplicative** sense.
- More research to speed up the algorithm is needed.

Conclusions

- The implicit Jacobi algorithm on symmetric rank revealing factorizations

$$A = XDX^T$$

is the first algorithm that:

1. computes the eigenvalues and eigenvectors of A to high relative accuracy,
2. preserves the symmetry, and
3. uses only orthogonal transformations.

- In addition, the error bounds are rigorously proven, and are the **best possible ones** from the sensitivity of the problem.
- The implicit Jacobi algorithm is **very simple and natural**.
- The implicit Jacobi algorithm is **backward stable** in a strong **multiplicative** sense.
- More research to speed up the algorithm is needed.

Conclusions

- The implicit Jacobi algorithm on symmetric rank revealing factorizations

$$A = XDX^T$$

is the first algorithm that:

1. computes the eigenvalues and eigenvectors of A to high relative accuracy,
2. preserves the symmetry, and
3. uses only orthogonal transformations.

- In addition, the error bounds are rigorously proven, and are the **best possible ones** from the sensitivity of the problem.
- The implicit Jacobi algorithm is **very simple and natural**.
- The implicit Jacobi algorithm is **backward stable** in a strong **multiplicative** sense.
- More research to speed up the algorithm is needed.

Conclusions

- The implicit Jacobi algorithm on symmetric rank revealing factorizations

$$A = XDX^T$$

  is the first algorithm that:

  1. computes the eigenvalues and eigenvectors of A to high relative accuracy,
  2. preserves the symmetry, and
  3. uses only orthogonal transformations.

- In addition, the error bounds are rigorously proven, and are the **best possible ones** from the sensitivity of the problem.
- The implicit Jacobi algorithm is **very simple and natural**.
- The implicit Jacobi algorithm is **backward stable** in a strong **multiplicative** sense.
- More research to speed up the algorithm is needed.

Conclusions

- The implicit Jacobi algorithm on symmetric rank revealing factorizations
$$A = XDX^T$$
is the first algorithm that:
  1. computes the eigenvalues and eigenvectors of A to high relative accuracy,
  2. preserves the symmetry, and
  3. uses only orthogonal transformations.
- In addition, the error bounds are rigorously proven, and are the **best possible ones** from the sensitivity of the problem.
- The implicit Jacobi algorithm is **very simple and natural**.
- The implicit Jacobi algorithm is **backward stable** in a strong **multiplicative** sense.
- More research to speed up the algorithm is needed.

Conclusions

- The implicit Jacobi algorithm on symmetric rank revealing factorizations

$$A = XDX^T$$

  is the first algorithm that:

  1. computes the eigenvalues and eigenvectors of A to high relative accuracy,
  2. preserves the symmetry, and
  3. uses only orthogonal transformations.

- In addition, the error bounds are rigorously proven, and are the **best possible ones** from the sensitivity of the problem.
- The implicit Jacobi algorithm is **very simple and natural**.
- The implicit Jacobi algorithm is **backward stable** in a strong **multiplicative** sense.
- More research to speed up the algorithm is needed.