On the Convergence of Block Kogbetliantz Methods for Calculating the SVD

7th International Workshop on Accurate Solution of Eigenvalue Problems – IWASEP7

Zvonimir Bujanović,
Zlatko Drmač

Department of Mathematics, University of Zagreb

June 9th, 2008.

Outline

1. Motivation
2. The Kogbetliantz Algorithm
3. The block Kogbetliantz algorithm

Goals: Accurate SVD's, EVD's

We are developing sharp high precision tools for numerical linear algebra. Our recent efforts have been focused to high accuracy computation of the SVD and its use as a kernel routine for accurate and efficient computation of

- Matrix spectral decomposition of symmetric matrices, i.e. solving $Hv_i = \lambda_i v_i$, $HMv_i = \lambda_i v_i$, $Hv_i = \lambda M v_i$ where H and M are symmetric positive definite matrices.
- The (P,Q)SVD decompositions, $Af(B) = U\Sigma V^T$, U, V orthogonal, Σ diagonal. Here $f(B) \in \{I, B, B^T, B^{-1}\}$.

Our goal is reliable mathematical software that computes e.g. $\tilde{\lambda}_i \approx \lambda_i$ with an error bound

$$\frac{|\tilde{\lambda}_i - \lambda_i|}{|\lambda_i|} \le C(H, M)\varepsilon, \quad \varepsilon = \text{roundoff unit}$$

where $C(H, M)$ estimates how well is the spectrum determined as function of the entries of H and M.

Jacobi SVD: $A = U\Sigma V^T$

We follow the Jacobi's idea:
$A^{(0)} = A$; $V^{[0]} := I$
$A^{(k+1)} = A^{(k)} V^{(k)}$, $V^{[k+1]} := V^{[k]} V^{(k)}$, $k = 0, 1, 2, \ldots$
Under certain conditions, as $k \longrightarrow \infty$:

$$A^{(k)} \longrightarrow U\Sigma, \quad ((A^{(k)})^T A^{(k)} \longrightarrow \Lambda = \Sigma^T \Sigma)$$
$$V^{(0)} V^{(1)} \ldots V^{(k)} \longrightarrow V$$

- works with two dense arrays $A^{(k)}$ and $V^{[k]}$; busy memory traffic
- high flop count $n(n-1)/2$ * *(one dot product and two plane rotations)* $\approx 3mn^2 + 2n^3$ flops per sweep. Many sweeps ($> 5, 6, 10$) needed for numerical convergence. mn^2 flops only to check convergence.
- all flops on BLAS 1 level
- destroys any initial zero structure

Recent improvement by Drmač and Veselić ...

Jacobi SVD: $A = U\Sigma V^T$

We follow the Jacobi's idea:
$A^{(0)} = A$; $V^{[0]} := I$
$A^{(k+1)} = A^{(k)}V^{(k)}$, $V^{[k+1]} := V^{[k]}V^{(k)}$, $k = 0, 1, 2, \ldots$
Under certain conditions, as $k \longrightarrow \infty$:

$$A^{(k)} \longrightarrow U\Sigma, \quad ((A^{(k)})^T A^{(k)} \longrightarrow \Lambda = \Sigma^T\Sigma)$$
$$V^{(0)}V^{(1)} \ldots V^{(k)} \longrightarrow V$$

• works with two dense arrays $A^{(k)}$ and $V^{[k]}$; busy memory traffic
• high flop count $n(n-1)/2$ * *(one dot product and two plane rotations)* $\approx 3mn^2 + 2n^3$ flops per sweep. Many sweeps ($> 5, 6, 10$) needed for numerical convergence. mn^2 flops only to check convergence.
• all flops on BLAS 1 level
• destroys any initial zero structure

Recent improvement by Drmač and Veselić ...

Jacobi ++ (Drmač, Veselić 06.)

- $(\Pi A)P = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$; $\rho = \mathrm{rank}(R)$; BLAS 3
  - $R(1:\rho, 1:n)^T = Q_1 \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$; BLAS 3
  - $X = R_1^T = \begin{pmatrix} \blacksquare & 0 \\ \blacksquare & \blacksquare \end{pmatrix}$;
    - ★ $X_\infty = U_x \Sigma = X \underbrace{\langle J_1 J_2 \cdots J_\infty \rangle}_{V_x}$ BLAS 1
    - $V_x = R_1^{-T}(X_\infty)$ BLAS 3
- $U = \Pi^T Q \begin{pmatrix} U_x & 0 \\ 0 & I_{m-\rho} \end{pmatrix}$; $V = PQ_1 \begin{pmatrix} V_x & 0 \\ 0 & I_{n-\rho} \end{pmatrix}$ BLAS 3
- if $\rho = n$, $Q_1 V_x = R^{-1} X_\infty$ BLAS 3

For optimal performance, we need a BLAS 3 Jacobi SVD for triangular matrices. ▸ Accuracy ▸ Efficiency

Jacobi ++ (Drmač, Veselić 06.)

- $(\Pi A)P = Q\begin{pmatrix} R \\ 0 \end{pmatrix}$; $\rho = \mathrm{rank}(R)$; BLAS 3
  - $R(1:\rho, 1:n)^T = Q_1\begin{pmatrix} R_1 \\ 0 \end{pmatrix}$; BLAS 3
  - $X = R_1^T = \begin{pmatrix} \blacksquare & 0 \\ \blacksquare & \blacksquare \end{pmatrix}$;
    - ★ $X_\infty = U_x\Sigma = X\underbrace{\langle J_1 J_2 \cdots J_\infty \rangle}_{V_x}$ BLAS 1
    - $V_x = R_1^{-T}(X_\infty)$ BLAS 3
- $U = \Pi^T Q\begin{pmatrix} U_x & 0 \\ 0 & I_{m-\rho} \end{pmatrix}$; $V = PQ_1\begin{pmatrix} V_x & 0 \\ 0 & I_{n-\rho} \end{pmatrix}$ BLAS 3
- if $\rho = n$, $Q_1 V_x = R^{-1} X_\infty$ BLAS 3

For optimal performance, we need a BLAS 3 Jacobi SVD for triangular matrices. ▸ Accuracy ▸ Efficiency

Jacobi ++ (Drmač, Veselić 06.)

- $(\Pi A)P = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$; $\rho = \mathrm{rank}(R)$; BLAS 3
  - $R(1:\rho, 1:n)^T = Q_1 \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$; BLAS 3
  - $X = R_1^T = \begin{pmatrix} \blacksquare & 0 \\ \blacksquare & \blacksquare \end{pmatrix}$;
    - ★ $X_\infty = U_x\Sigma = X\underbrace{\langle J_1 J_2 \cdots J_\infty \rangle}_{V_x}$ BLAS 1
    - $V_x = R_1^{-T}(X_\infty)$ BLAS 3
- $U = \Pi^T Q \begin{pmatrix} U_x & 0 \\ 0 & I_{m-\rho} \end{pmatrix}$; $V = PQ_1 \begin{pmatrix} V_x & 0 \\ 0 & I_{n-\rho} \end{pmatrix}$ BLAS 3
- if $\rho = n$, $Q_1 V_x = R^{-1} X_\infty$ BLAS 3

For optimal performance, we need a BLAS 3 Jacobi SVD for triangular matrices. Accuracy Efficiency

Jacobi ++ (Drmač, Veselić 06.)

- $(\Pi A)P = Q\begin{pmatrix} R \\ 0 \end{pmatrix}$; $\rho = \mathrm{rank}(R)$; BLAS 3
  - $R(1:\rho, 1:n)^T = Q_1 \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$; BLAS 3
  - $X = R_1^T = \begin{pmatrix} \blacksquare & 0 \\ \blacksquare & \blacksquare \end{pmatrix}$;
    - ★ $X_\infty \equiv U_x \Sigma = X \underbrace{\langle J_1 J_2 \cdots J_\infty \rangle}_{V_x}$ BLAS 1
  - $V_x = R_1^{-T}(X_\infty)$ BLAS 3
- $U = \Pi^T Q \begin{pmatrix} U_x & 0 \\ 0 & I_{m-\rho} \end{pmatrix}$; $V = PQ_1 \begin{pmatrix} V_x & 0 \\ 0 & I_{n-\rho} \end{pmatrix}$ BLAS 3
- if $\rho = n$, $Q_1 V_x = R^{-1} X_\infty$ BLAS 3

For optimal performance, we need a BLAS 3 Jacobi SVD for triangular matrices. ▸ Accuracy ▸ Efficiency

Jacobi ++ (Drmač, Veselić 06.)

- $(\Pi A)P = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$; $\rho = \mathrm{rank}(R)$; BLAS 3
  - $R(1:\rho, 1:n)^T = Q_1 \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$; BLAS 3
  - $X = R_1^T = \begin{pmatrix} \blacksquare & 0 \\ \blacksquare & \blacksquare \end{pmatrix}$;
    - ★ $X_\infty \equiv U_x \Sigma = X \underbrace{\langle J_1 J_2 \cdots J_\infty \rangle}_{V_x}$ BLAS 1
    - $V_x = R_1^{-T}(X_\infty)$ BLAS 3
- $U = \Pi^T Q \begin{pmatrix} U_x & 0 \\ 0 & I_{m-\rho} \end{pmatrix}$; $V = PQ_1 \begin{pmatrix} V_x & 0 \\ 0 & I_{n-\rho} \end{pmatrix}$ BLAS 3
- if $\rho = n$, $Q_1 V_x = R^{-1} X_\infty$ BLAS 3

For optimal performance, we need a BLAS 3 Jacobi SVD for triangular matrices. ▸ Accuracy ▸ Efficiency

Jacobi ++ (Drmač, Veselić 06.)

- $(\Pi A)P = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$; $\rho = \mathrm{rank}(R)$; BLAS 3
  - $R(1:\rho, 1:n)^T = Q_1 \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$; BLAS 3
  - $X = R_1^T = \begin{pmatrix} \blacksquare & 0 \\ \blacksquare & \blacksquare \end{pmatrix}$;
    - ★ $X_\infty \equiv U_x \Sigma = X \underbrace{\langle J_1 J_2 \cdots J_\infty \rangle}_{V_x}$ BLAS 1
    - $V_x = R_1^{-T}(X_\infty)$ BLAS 3
- $U = \Pi^T Q \begin{pmatrix} U_x & 0 \\ 0 & I_{m-\rho} \end{pmatrix}$; $V = PQ_1 \begin{pmatrix} V_x & 0 \\ 0 & I_{n-\rho} \end{pmatrix}$ BLAS 3
- if $\rho = n$, $Q_1 V_x = R^{-1} X_\infty$ BLAS 3

For optimal performance, we need a BLAS 3 Jacobi SVD for triangular matrices. ▸ Accuracy ▸ Efficiency

Jacobi ++ (Drmač, Veselić 06.)

- $(\Pi A)P = Q\begin{pmatrix} R \\ 0 \end{pmatrix}$; $\rho = \mathrm{rank}(R)$; BLAS 3
  - $R(1:\rho, 1:n)^T = Q_1\begin{pmatrix} R_1 \\ 0 \end{pmatrix}$; BLAS 3
  - $X = R_1^T = \begin{pmatrix} \blacksquare & 0 \\ \blacksquare & \blacksquare \end{pmatrix}$;
    - ★ $X_\infty \equiv U_x\Sigma = X\underbrace{\langle J_1 J_2 \cdots J_\infty \rangle}_{V_x}$ BLAS 1
    - $V_x = R_1^{-T}(X_\infty)$ BLAS 3
- $U = \Pi^T Q\begin{pmatrix} U_x & 0 \\ 0 & I_{m-\rho} \end{pmatrix}$; $V = PQ_1\begin{pmatrix} V_x & 0 \\ 0 & I_{n-\rho} \end{pmatrix}$ BLAS 3
- if $\rho = n$, $Q_1 V_x = R^{-1} X_\infty$ BLAS 3

For optimal performance, we need a BLAS 3 Jacobi SVD for triangular matrices. ▸ Accuracy ▸ Efficiency

The non-blocked version

Kogbetliantz (1955.):
for $A \in \mathbb{C}^{n\times n}$ generate $(A^{(k)})_k$ such that $A^{(0)} = A$ and

$$A^{(k)} = \begin{bmatrix} \bullet & \bullet & \bullet & {\color{red}\bullet} & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ {\color{red}\bullet} & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \end{bmatrix} \longrightarrow A^{(k+1)} = \begin{bmatrix} \bullet & \bullet & \bullet & 0 & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \end{bmatrix}$$

Transformation via plane rotations: $A^{(k+1)} = \left(U^{(k)}\right)^{\star} A^{(k)} V^{(k)}$:

$$U_{2x2}^{(k)} = \begin{bmatrix} \cos\phi^{(k)} & e^{i\tilde{\phi}^{(k)}} \sin\phi^{(k)} \\ -e^{i\tilde{\phi}^{(k)}} \sin\phi^{(k)} & \cos\phi^{(k)} \end{bmatrix}, \quad V_{2x2}^{(k)} = \begin{bmatrix} \cos\psi^{(k)} & e^{i\tilde{\psi}^{(k)}} \sin\psi^{(k)} \\ -e^{i\tilde{\psi}^{(k)}} \sin\psi^{(k)} & \cos\psi^{(k)} \end{bmatrix}$$

Convergence results for row/column cyclic schemes

- Forsythe and Henrici (1960.):
  Globally convergent if all $\phi^{(k)}, \psi^{(k)} < \frac{\pi}{2} - \epsilon$.
  Not always attainable; instead use relaxation

$$A^{(k+1)} = \begin{bmatrix} \bullet & \bullet & \bullet & < \tau & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ < \tau & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \end{bmatrix}$$

- Fernando (1989.):
  Globally convergent if all $\phi^{(k)} < \frac{\pi}{2} - \epsilon$ or if all $\psi^{(k)} < \frac{\pi}{2} - \epsilon$. Always attainable.
- Paige and van Dooren (1986.), Hari (1986.):
  Ultimate quadratic convergence.

Convergence results for row/column cyclic schemes

- Forsythe and Henrici (1960.):
  Globally convergent if all $\phi^{(k)}, \psi^{(k)} < \frac{\pi}{2} - \epsilon$.
  Not always attainable; instead use relaxation

$$A^{(k+1)} = \begin{bmatrix} \bullet & \bullet & \bullet & < \tau & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ < \tau & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \end{bmatrix}$$

- Fernando (1989.):
  Globally convergent if all $\phi^{(k)} < \frac{\pi}{2} - \epsilon$ or if all $\psi^{(k)} < \frac{\pi}{2} - \epsilon$. Always attainable.
- Paige and van Dooren (1986.), Hari (1986.):
  Ultimate quadratic convergence.

Convergence results for row/column cyclic schemes

- Forsythe and Henrici (1960.):
  Globally convergent if all $\phi^{(k)}, \psi^{(k)} < \frac{\pi}{2} - \epsilon$.
  Not always attainable; instead use relaxation

$$A^{(k+1)} = \begin{bmatrix} \bullet & \bullet & \bullet & < \tau & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ < \tau & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \end{bmatrix}$$

- Fernando (1989.):
  Globally convergent if all $\phi^{(k)} < \frac{\pi}{2} - \epsilon$ or if all $\psi^{(k)} < \frac{\pi}{2} - \epsilon$. Always attainable.
- Paige and van Dooren (1986.), Hari (1986.):
  Ultimate quadratic convergence.

Efficiency (1955. ... 2008.)

Kogbetliantz on the performance of his new algorithm:

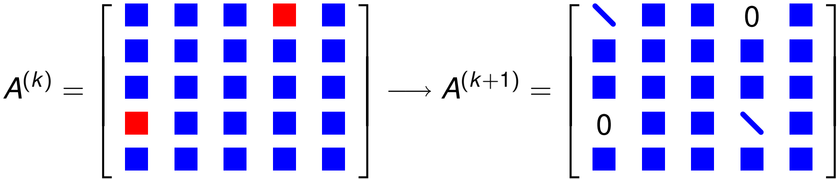
The results of numerical computations performed with the aid of IBMs new electronic data processing machine type 701 agree completely with the conclusion of this study of convergence. The type 701 performs with extreme rapidity: a 32×32 *matrix is diagonalized in* 19 *minutes and the numerical results are printed in four minutes, the total time being* 23 *minutes.*

For better performance

- faster machines ✓
- preconditioning: apply to a diagonally dominant triangular matrix (as in Jacobi++) ✓
- block transformations for higher $\frac{\text{flop}}{\text{memory reference}}$ ratio

The blocked version

Partition $A \in \mathbb{C}^{n\times n}$ into blocks $\mathcal{A}_{ij} \in \mathbb{C}^{n_i \times n_j}$, $\sum n_i = n$.
Generate $(A^{(k)})_k$ such that $A^{(0)} = A$ and



Transformation via block rotations: $A^{(k+1)} = \left(U^{(k)}\right)^\star A^{(k)} V^{(k)}$:

$$U_{2x2}^{(k)} = \begin{array}{c} \\ n_i \\ n_j \end{array} \begin{array}{c} \begin{array}{cc} n_i & n_j \end{array} \\ \begin{bmatrix} \mathcal{U}_{1,1}^{(k)} & \mathcal{U}_{1,2}^{(k)} \\ \mathcal{U}_{2,1}^{(k)} & \mathcal{U}_{2,2}^{(k)} \end{bmatrix} \end{array} ; \quad V_{2x2}^{(k)} = \begin{array}{c} \\ n_i \\ n_j \end{array} \begin{array}{c} \begin{array}{cc} n_i & n_j \end{array} \\ \begin{bmatrix} \mathcal{V}_{1,1}^{(k)} & \mathcal{V}_{1,2}^{(k)} \\ \mathcal{V}_{2,1}^{(k)} & \mathcal{V}_{2,2}^{(k)} \end{bmatrix} \end{array} ;$$

The plan

Our goal: to establish conditions for the global convergence of cyclic block–Kogbetliantz methods.

Steps in the process:

- the off–diagonal converges to zero;
- $U^{(k)}$, $V^{(k)}$ converge to identity;
- the diagonal converges to a specific permutation of singular values;
- columns of $\prod U^{(k)}, \prod V^{(k)}$ converge to left/right singular vectors.

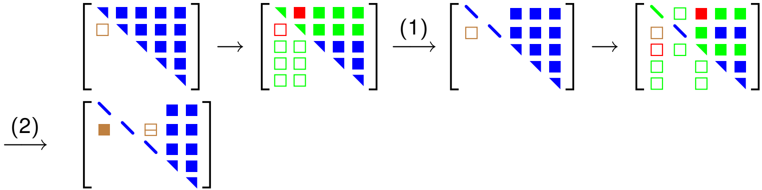
Convergence of the block Jacobi algorithm in the general case is recently proved by Drmač, we use similar techniques...

Tools and techniques

- Block-versions of standard 2×2 techniques (Forsythe and Henrici, Fernando)
- New block transformations to guarantee uniformly bounded angles (Drmač)
- Tracking the diagonals of the iterates (Schur–Horn theorem in the Hermitian case, similarly for the SVD)
- Infinite products for singular vectors convergence in the case of simple singular values
- Canonical angles between subspaces for singular subspaces convergence in the case of multiple singular values

Annihilation of the first row

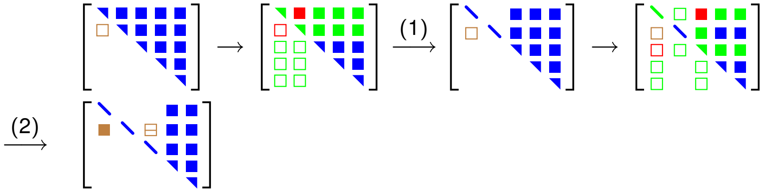
The most efficient implementation of the Kogbetliantz algortihm works on a triangular matrix, obtained e.g. by QR factorization.
Keep track of ■.



Note: ■ = ⊟ · $\mathcal{V}_{2,1}^{(2)}$.

Annihilation of the first row

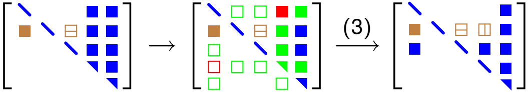
The most efficient implementation of the Kogbetliantz algortihm works on a triangular matrix, obtained e.g. by QR factorization.
Keep track of ■.



Note: ■ $= ⊟ \cdot \mathcal{V}_{2,1}^{(2)}$.

Annihilation of the first row

Note: $\blacksquare = \boxminus \cdot \mathcal{V}_{2,1}^{(2)}$.



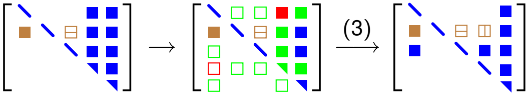
We have: $\blacksquare = \boxminus \cdot \mathcal{V}_{2,1}^{(2)} \cdot \mathcal{V}_{1,1}^{(3)} + \text{◫} \cdot \mathcal{V}_{2,1}^{(3)}$.

Finally:

$$\blacksquare = \begin{bmatrix} \boxminus & \text{◫} & \boxtimes \end{bmatrix} \cdot \begin{bmatrix} \mathcal{V}_{2,1}^{(2)} \cdot \mathcal{V}_{1,1}^{(3)} \cdot \mathcal{V}_{1,1}^{(4)} \\ \mathcal{V}_{2,1}^{(3)} \cdot \mathcal{V}_{1,1}^{(4)} \\ \mathcal{V}_{2,1}^{(4)} \end{bmatrix} = \begin{bmatrix} \boxminus & \text{◫} & \boxtimes \end{bmatrix} \cdot \mathcal{V}$$

Annihilation of the first row

Note: $\blacksquare = \boxminus \cdot \mathcal{V}^{(2)}_{2,1}$.



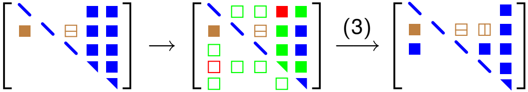
We have: $\blacksquare = \boxminus \cdot \mathcal{V}^{(2)}_{2,1} \cdot \mathcal{V}^{(3)}_{1,1} + \boxbar \cdot \mathcal{V}^{(3)}_{2,1}$.

Finally:

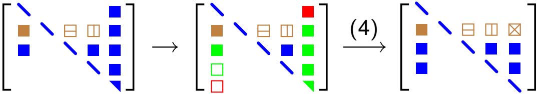
$$\blacksquare = \begin{bmatrix} \boxminus & \boxbar & \boxtimes \end{bmatrix} \cdot \begin{bmatrix} \mathcal{V}^{(2)}_{2,1} \cdot \mathcal{V}^{(3)}_{1,1} \cdot \mathcal{V}^{(4)}_{1,1} \\ \mathcal{V}^{(3)}_{2,1} \cdot \mathcal{V}^{(4)}_{1,1} \\ \mathcal{V}^{(4)}_{2,1} \end{bmatrix} = \begin{bmatrix} \boxminus & \boxbar & \boxtimes \end{bmatrix} \cdot \mathcal{V}$$

Annihilation of the first row

Note: $\blacksquare = \boxminus \cdot \mathcal{V}_{2,1}^{(2)}$.



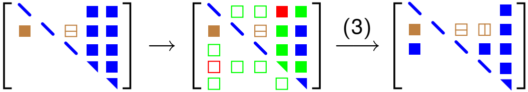
We have: $\blacksquare = \boxminus \cdot \mathcal{V}_{2,1}^{(2)} \cdot \mathcal{V}_{1,1}^{(3)} + \boxbar \cdot \mathcal{V}_{2,1}^{(3)}$.



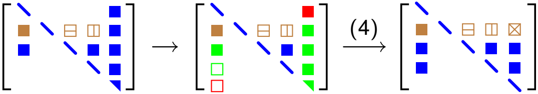
Finally:

$$\blacksquare = \begin{bmatrix} \boxminus & \boxbar & \boxtimes \end{bmatrix} \cdot \begin{bmatrix} \mathcal{V}_{2,1}^{(2)} \cdot \mathcal{V}_{1,1}^{(3)} \cdot \mathcal{V}_{1,1}^{(4)} \\ \mathcal{V}_{2,1}^{(3)} \cdot \mathcal{V}_{1,1}^{(4)} \\ \mathcal{V}_{2,1}^{(4)} \end{bmatrix} = \begin{bmatrix} \boxminus & \boxbar & \boxtimes \end{bmatrix} \cdot \mathcal{V}$$

Note: $\blacksquare = \boxminus \cdot \mathcal{V}_{2,1}^{(2)}$.



We have: $\blacksquare = \boxminus \cdot \mathcal{V}_{2,1}^{(2)} \cdot \mathcal{V}_{1,1}^{(3)} + \square\!\!\mid \cdot \mathcal{V}_{2,1}^{(3)}$.



Finally:

$$\blacksquare = \begin{bmatrix} \boxminus & \square\!\!\mid & \boxtimes \end{bmatrix} \cdot \begin{bmatrix} \mathcal{V}_{2,1}^{(2)} \cdot \mathcal{V}_{1,1}^{(3)} \cdot \mathcal{V}_{1,1}^{(4)} \\ \mathcal{V}_{2,1}^{(3)} \cdot \mathcal{V}_{1,1}^{(4)} \\ \mathcal{V}_{2,1}^{(4)} \end{bmatrix} = \begin{bmatrix} \boxminus & \square\!\!\mid & \boxtimes \end{bmatrix} \cdot \mathcal{V}$$

Off–norm reduction depends on the angles

Cosine-Sine Decomposition (Stewart, 1977.) of $V_{2x2}^{(k)}$:

$$V_{2x2}^{(k)} = \begin{bmatrix} \mathcal{V}_{1,1}^{(k)} & \mathcal{V}_{1,2}^{(k)} \\ \mathcal{V}_{2,1}^{(k)} & \mathcal{V}_{2,2}^{(k)} \end{bmatrix}$$

$$= \begin{bmatrix} \mathcal{Z}_1^{(k)} & \\ & \mathcal{Z}_2^{(k)} \end{bmatrix} \cdot \left[\begin{array}{cc|c} \cos\Psi^{(k)} & 0 & \sin\Psi^{(k)} \\ 0 & I & 0 \\ \hline -\sin\Psi^{(k)} & 0 & \cos\Psi^{(k)} \end{array}\right] \cdot \begin{bmatrix} \mathcal{W}_1^{(k)} & \\ & \mathcal{W}_2^{(k)} \end{bmatrix}$$

Easy to show:

$$\|\mathcal{V}\|_2 \geq \sigma_{\min}(\mathcal{V}_{1,1}^{(2)}) \cdot \sigma_{\min}(\mathcal{V}_{1,1}^{(3)}) \cdot \sigma_{\min}(\mathcal{V}_{1,1}^{(4)}) = \cos\psi^{(2)} \cdot \cos\psi^{(3)} \cdot \cos\psi^{(4)}$$

Off–norm reduction after one sweep:

$$\mathrm{off}^2(A^{[1]}) \leq \mathrm{off}^2(A) \cdot \left(1 - \cos^2\psi^{(2)} \cdot \cos^2\psi^{(3)} \cdot \ldots \cdot \cos^2\psi^{(N(N-1)/2)}\right) .$$

Off–norm reduction depends on the angles

Cosine-Sine Decomposition (Stewart, 1977.) of $V_{2x2}^{(k)}$:

$$V_{2x2}^{(k)} = \begin{bmatrix} \mathcal{V}_{1,1}^{(k)} & \mathcal{V}_{1,2}^{(k)} \\ \mathcal{V}_{2,1}^{(k)} & \mathcal{V}_{2,2}^{(k)} \end{bmatrix}$$

$$= \begin{bmatrix} \mathcal{Z}_1^{(k)} & \\ & \mathcal{Z}_2^{(k)} \end{bmatrix} \cdot \left[\begin{array}{cc|c} \cos \Psi^{(k)} & 0 & \sin \Psi^{(k)} \\ 0 & I & 0 \\ \hline -\sin \Psi^{(k)} & 0 & \cos \Psi^{(k)} \end{array}\right] \cdot \begin{bmatrix} \mathcal{W}_1^{(k)} & \\ & \mathcal{W}_2^{(k)} \end{bmatrix}$$

Easy to show:

$$\|\mathcal{V}\|_2 \geq \sigma_{\min}(\mathcal{V}_{1,1}^{(2)}) \cdot \sigma_{\min}(\mathcal{V}_{1,1}^{(3)}) \cdot \sigma_{\min}(\mathcal{V}_{1,1}^{(4)}) = \cos\psi^{(2)} \cdot \cos\psi^{(3)} \cdot \cos\psi^{(4)}$$

Off–norm reduction after one sweep:

$$\mathrm{off}^2(A^{[1]}) \leq \mathrm{off}^2(A) \cdot \left(1 - \cos^2\psi^{(2)} \cdot \cos^2\psi^{(3)} \cdot \ldots \cdot \cos^2\psi^{(N(N-1)/2)}\right) .$$

Convergence to the diagonal form

Note: ◥ $\xrightarrow{[1]}$ ◣ $\xrightarrow{[2]}$ ◥ $\xrightarrow{[3]}$ ◣ $\xrightarrow{[4]}$ ◥ $\ldots$

Let:

$$\psi^{[s]} = \max \psi^{(j)}, \quad \phi^{[s]} = \max \phi^{(j)},$$

where j are indices of rotations in the s–th sweep.

Theorem

Let $\epsilon^{[s]} > 0$ such that

$$\psi^{[2s-1]} < \frac{\pi}{2} - \epsilon^{[2s-1]}, \quad \phi^{[2s]} < \frac{\pi}{2} - \epsilon^{[2s]}\ .$$

If $\limsup_s \epsilon^{[s]} > 0$ then $(A^{(k)})_k$ converges to diagonal form.

Convergence to the diagonal form

Note: ◥ $\xrightarrow{[1]}$ ◣ $\xrightarrow{[2]}$ ◥ $\xrightarrow{[3]}$ ◣ $\xrightarrow{[4]}$ ◥ . . .

Let:

$$\psi^{[s]} = \max \psi^{(j)}, \quad \phi^{[s]} = \max \phi^{(j)},$$

where j are indices of rotations in the s–th sweep.

Theorem

Let $\epsilon^{[s]} > 0$ such that

$$\psi^{[2s-1]} < \frac{\pi}{2} - \epsilon^{[2s-1]}, \quad \phi^{[2s]} < \frac{\pi}{2} - \epsilon^{[2s]} .$$

If $\limsup_s \epsilon^{[s]} > 0$ then $(A^{(k)})_k$ converges to diagonal form.

Uniformly bounded cosines (UBC)

Theorem (Drmač, 2008.)

Let $U = \begin{bmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{bmatrix}$ *be a unitary block matrix. There exists a permutation* P *such that* $\hat{U} = UP = \begin{bmatrix} \boxtimes & \boxtimes \\ \boxtimes & \boxtimes \end{bmatrix}$ *and*

$$1 \geq \sigma_{\min}(\boxtimes) = \sigma_{\min}(\boxtimes) \geq f(b, r) > 0 \ ,$$

where b, r *are block dimensions of* $\blacksquare$ *and* $\blacksquare$ *resp.*

$f(b, r) = 3(4^b + 6b - 1)^{-1/2}(r + 1)^{-1/2}$

Take P from the Businger–Golub QRF of

$$\begin{bmatrix} \blacksquare & \blacksquare \end{bmatrix} \cdot P = Q \cdot \begin{bmatrix} \blacktriangleleft & \blacksquare \end{bmatrix} \ .$$

P = "UBC permutation", UP = "UBC transform".

Uniformly bounded cosines (UBC)

Theorem (Drmač, 2008.)

Let $U = \begin{bmatrix} \color{blue}\blacksquare & \color{brown}\blacksquare \\ \color{green}\blacksquare & \color{red}\blacksquare \end{bmatrix}$ *be a unitary block matrix. There exists a permutation* P *such that* $\hat{U} = UP = \begin{bmatrix} \color{blue}\boxtimes & \color{brown}\boxtimes \\ \color{green}\boxtimes & \color{red}\boxtimes \end{bmatrix}$ *and*

$$1 \geq \sigma_{\min}({\color{blue}\boxtimes}) = \sigma_{\min}({\color{red}\boxtimes}) \geq f(b, r) > 0 \ ,$$

where b, r *are block dimensions of* $\color{blue}\blacksquare$ *and* $\color{red}\blacksquare$ *resp.*
$f(b, r) = 3(4^b + 6b - 1)^{-1/2}(r + 1)^{-1/2}$

Take P from the Businger–Golub QRF of

$$\begin{bmatrix} \blacksquare & \blacksquare \end{bmatrix} \cdot P = Q \cdot \begin{bmatrix} \blacktriangle & \blacksquare \end{bmatrix} \ .$$

P = "UBC permutation", UP = "UBC transform".

Uniformly bounded cosines (UBC)

Theorem (Drmač, 2008.)

Let $U = \begin{bmatrix} {\color{blue}\blacksquare} & {\color{brown}\blacksquare} \\ {\color{green}\blacksquare} & {\color{red}\blacksquare} \end{bmatrix}$ *be a unitary block matrix. There exists a permutation* P *such that* $\hat{U} = UP = \begin{bmatrix} {\color{blue}\boxtimes} & {\color{brown}\boxtimes} \\ {\color{green}\boxtimes} & {\color{red}\boxtimes} \end{bmatrix}$ *and*

$$1 \geq \sigma_{\min}({\color{blue}\boxtimes}) = \sigma_{\min}({\color{red}\boxtimes}) \geq f(b, r) > 0 \ ,$$

where b, r *are block dimensions of* ${\color{blue}\blacksquare}$ *and* ${\color{red}\blacksquare}$ *resp.*
$f(b, r) = 3(4^b + 6b - 1)^{-1/2}(r + 1)^{-1/2}$

Take P from the Businger–Golub QRF of

$$\begin{bmatrix} {\color{blue}\blacksquare} & {\color{brown}\blacksquare} \end{bmatrix} \cdot P = Q \cdot \begin{bmatrix} \text{◥} & \blacksquare \end{bmatrix} \ .$$

P = "UBC permutation", UP = "UBC transform".

Convergence to the diagonal form

To meet the constraints of the convergence theorem:

- Instead of $U^{(k)}$, $V^{(k)}$ rotate with $U^{(k)}P$, $V^{(k)}P$.

This leads to $A^{(k+1)} = P^\star \cdot \left(\left(U^{(k)}\right)^\star \cdot A^{(k)} \cdot V^{(k)} \right) \cdot P$.
Has the same diagonal as the original $A^{(k+1)}$. Same blocks annihilated.

Forsythe' and Henrici's approach:
Must bound both $\phi^{(k)}$ and $\psi^{(k)}$ at once.

- Instead of $U^{(k)}$ rotate with $U^{(k)}P_1$.
- Instead of $V^{(k)}$ rotate with $V^{(k)}P_2$.

This leads to $A^{(k+1)} = P_1^\star \cdot \left(\left(U^{(k)}\right)^\star \cdot A^{(k)} \cdot V^{(k)} \right) \cdot P_2$.
Not OK – diagonal elements run away from the diagonal.
Can't be achieved even in the non–block case.

Convergence to the diagonal form

To meet the constraints of the convergence theorem:

- Instead of $U^{(k)}$, $V^{(k)}$ rotate with $U^{(k)}P$, $V^{(k)}P$.

This leads to $A^{(k+1)} = P^{\star} \cdot \left(\left(U^{(k)}\right)^{\star} \cdot A^{(k)} \cdot V^{(k)} \right) \cdot P$.
Has the same diagonal as the original $A^{(k+1)}$. Same blocks annihilated.

Forsythe' and Henrici's approach:
Must bound both $\phi^{(k)}$ and $\psi^{(k)}$ at once.

- Instead of $U^{(k)}$ rotate with $U^{(k)}P_1$.
- Instead of $V^{(k)}$ rotate with $V^{(k)}P_2$.

This leads to $A^{(k+1)} = P_1^{\star} \cdot \left(\left(U^{(k)}\right)^{\star} \cdot A^{(k)} \cdot V^{(k)} \right) \cdot P_2$.
Not OK – diagonal elements run away from the diagonal.
Can't be achieved even in the non–block case.

Variants of the algorithm

- Gentleman: permutational similarity after each rotation to keep ◥.

  ▸ Details

  Sufficient to bound only $\psi^{(k)}$.

- Fernando: if $A = $ ■ it is also sufficient to bound only $\psi^{(k)}$ (off–norm reduces after two full sweeps)

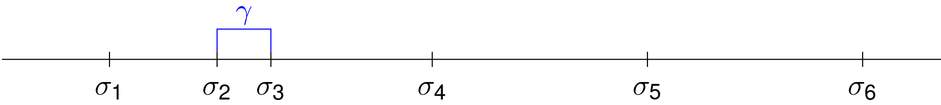
Variants of the algorithm

- Gentleman: permutational similarity after each rotation to keep ◥.
  ▸ Details
  Sufficient to bound only $\psi^{(k)}$.

- Fernando: if $A =$ ■ it is also sufficient to bound only $\psi^{(k)}$ (off–norm reduces after two full sweeps)

Block rotations converge to identity

Let A have simple singular values, $\gamma = \text{gap}(A) := \min_{i \neq j} |\sigma_i - \sigma_j| > 0$.



Block rotations converge to identity

Consider $\tilde{B} = A_{2x2}^{(k-1)}$. as perturbation of $B = \text{diag}(A_{2x2}^{(k-1)})$, $k \gg 1$.

Note:

- $U_{2x2}^{(k)}$ and $V_{2x2}^{(k)}$ have singular vectors of $\tilde{B}$ as columns.
- Wedin's $\Phi - \Theta$ theorem ▸ See statement :
  if sing. values of A are simple then $U_{2x2}^{(k)}$, $V_{2x2}^{(k)}$ are perturbations of the same permutation matrix.
- Application of UBC permutation P moves the $\mathcal{O}(1)$ elements of $U_{2x2}^{(k)}$, $V_{2x2}^{(k)}$ to diagonal blocks. Thus: $U_{2x2}^{(k)}P$, $V_{2x2}^{(k)}P \approx P_1 \oplus P_2$.
- Rotate with

$$U^{(k)} \cdot P \cdot \begin{bmatrix} P_1^\star & \\ & P_2^\star \end{bmatrix} \text{ and } V^{(k)} \cdot P \cdot \begin{bmatrix} P_1^\star & \\ & P_2^\star \end{bmatrix} .$$

These are UBC transforms. $U^{(k)}$ and $V^{(k)}$ converge to I.

Block rotations converge to identity

Consider $\tilde{B} = A_{2x2}^{(k-1)}$. as perturbation of $B = \mathrm{diag}(A_{2x2}^{(k-1)})$, $k \gg 1$.

Note:

- $U_{2x2}^{(k)}$ and $V_{2x2}^{(k)}$ have singular vectors of $\tilde{B}$ as columns.
- Wedin's $\Phi - \Theta$ theorem See statement:
if sing. values of A are simple then $U_{2x2}^{(k)}$, $V_{2x2}^{(k)}$ are perturbations of the same permutation matrix.
- Application of UBC permutation P moves the $\mathcal{O}(1)$ elements of $U_{2x2}^{(k)}$, $V_{2x2}^{(k)}$ to diagonal blocks. Thus: $U_{2x2}^{(k)}P$, $V_{2x2}^{(k)}P \approx P_1 \oplus P_2$.
- Rotate with

$$U^{(k)} \cdot P \cdot \begin{bmatrix} P_1^\star & \\ & P_2^\star \end{bmatrix} \text{ and } V^{(k)} \cdot P \cdot \begin{bmatrix} P_1^\star & \\ & P_2^\star \end{bmatrix} .$$

These are UBC transforms. $U^{(k)}$ and $V^{(k)}$ converge to I.

Block rotations converge to identity

Consider $\tilde{B} = A_{2x2}^{(k-1)}$. as perturbation of $B = \text{diag}(A_{2x2}^{(k-1)})$, $k \gg 1$.

Note:

- $U_{2x2}^{(k)}$ and $V_{2x2}^{(k)}$ have singular vectors of $\tilde{B}$ as columns.
- Wedin's $\Phi - \Theta$ theorem ▸ See statement:
  if sing. values of A are simple then $U_{2x2}^{(k)}$, $V_{2x2}^{(k)}$ are perturbations of the same permutation matrix.
- Application of UBC permutation P moves the $\mathcal{O}(1)$ elements of $U_{2x2}^{(k)}$, $V_{2x2}^{(k)}$ to diagonal blocks. Thus: $U_{2x2}^{(k)}P$, $V_{2x2}^{(k)}P \approx P_1 \oplus P_2$.
- Rotate with

$$U^{(k)} \cdot P \cdot \begin{bmatrix} P_1^\star & \\ & P_2^\star \end{bmatrix} \text{ and } V^{(k)} \cdot P \cdot \begin{bmatrix} P_1^\star & \\ & P_2^\star \end{bmatrix} .$$

These are UBC transforms. $U^{(k)}$ and $V^{(k)}$ converge to I.

Block rotations converge to identity

Consider $\tilde{B} = A_{2x2}^{(k-1)}$. as perturbation of $B = \text{diag}(A_{2x2}^{(k-1)})$, $k \gg 1$.

Note:

- $U_{2x2}^{(k)}$ and $V_{2x2}^{(k)}$ have singular vectors of $\tilde{B}$ as columns.
- Wedin's $\Phi - \Theta$ theorem ▸ See statement:
  if sing. values of A are simple then $U_{2x2}^{(k)}$, $V_{2x2}^{(k)}$ are perturbations of the same permutation matrix.
- Application of UBC permutation P moves the $\mathcal{O}(1)$ elements of $U_{2x2}^{(k)}$, $V_{2x2}^{(k)}$ to diagonal blocks. Thus: $U_{2x2}^{(k)}P$, $V_{2x2}^{(k)}P \approx P_1 \oplus P_2$.
- Rotate with

$$U^{(k)} \cdot P \cdot \begin{bmatrix} P_1^\star & \\ & P_2^\star \end{bmatrix} \text{ and } V^{(k)} \cdot P \cdot \begin{bmatrix} P_1^\star & \\ & P_2^\star \end{bmatrix} .$$

These are UBC transforms. $U^{(k)}$ and $V^{(k)}$ converge to I.

Convergence of the diagonal

Let $d^{(k)}$ be the diagonal of $A^{(k)}$.
Note:

$$\underbrace{\begin{bmatrix} \diagdown & \\ & \diagdown \end{bmatrix}}_{\text{part of } d^{(k)}} = \left(U^{(k)}_{2x2}\right)^\star \cdot \underbrace{A^{(k-1)}_{2x2}}_{\text{part of } d^{(k-1)}} \cdot V^{(k)}_{2x2}$$

Thus:

$$d^{(k-1)} = \left(U^{(k)} \circ \overline{V}^{(k)}\right) \cdot d^{(k)} .$$

Here: $\circ$ = Hadamard product, $\overline{V}$ = conjugate, no transposing.

Convergence of the diagonal

Let $d^{(k)}$ be the diagonal of $A^{(k)}$.
Note:

$$\underbrace{A^{(k-1)}_{2x2}}_{\text{part of } d^{(k-1)}} = U^{(k)}_{2x2} \cdot \underbrace{\begin{bmatrix} \diagdown & \\ & \diagdown \end{bmatrix}}_{\text{part of } d^{(k)}} \cdot \left(V^{(k)}_{2x2}\right)^{\star}$$

Thus:

$$d^{(k-1)} = \left(U^{(k)} \circ \overline{V}^{(k)}\right) \cdot d^{(k)} \,.$$

Here: $\circ =$ Hadamard product, $\overline{V} =$ conjugate, no transposing.

Convergence of the diagonal

Let $d^{(k)}$ be the diagonal of $A^{(k)}$.
Note:

$$\underbrace{A_{2x2}^{(k-1)}}_{\text{part of } d^{(k-1)}} = U_{2x2}^{(k)} \cdot \underbrace{\begin{bmatrix} \diagdown & \\ & \diagdown \end{bmatrix}}_{\text{part of } d^{(k)}} \cdot \left(V_{2x2}^{(k)}\right)^{\star}$$

Thus:

$$d^{(k-1)} = \left(U^{(k)} \circ \overline{V}^{(k)}\right) \cdot d^{(k)} \ .$$

Here: $\circ =$ Hadamard product, $\overline{V} =$ conjugate, no transposing.

Convergence of the diagonal

We have shown:

$$d^{(k-1)} = \left(U^{(k)} \circ \overline{V}^{(k)}\right) \cdot d^{(k)} \,.$$

If

$$\left\| |d^{(k-1)}| - \sigma_\alpha \right\|_2 \le \gamma/3 \,,$$

then

$$\begin{aligned} \left\| |d^{(k)}| - \sigma_\alpha \right\|_2 &\le \left\| |d^{(k)}| - |d^{(k-1)}| \right\|_2 + \left\| |d^{(k-1)}| - \sigma_\alpha \right\|_2 \\ &\le \left\| I - U^{(k)} \circ \overline{V}^{(k)} \right\|_F \cdot \|d^{(k)}\|_2 + \left\| |d^{(k-1)}| - \sigma_\alpha \right\|_2 \\ &< 2\gamma/3 \end{aligned}$$

Convergence of the diagonal

We have shown:

$$d^{(k-1)} = \left(U^{(k)} \circ \overline{V}^{(k)} \right) \cdot d^{(k)} \ .$$

If

$$\left\| |d^{(k-1)}| - \sigma_\alpha \right\|_2 \leq \gamma/3 \ ,$$

then

$$\begin{aligned}
\left\| |d^{(k)}| - \sigma_\alpha \right\|_2 &\leq \left\| |d^{(k)}| - |d^{(k-1)}| \right\|_2 + \left\| |d^{(k-1)}| - \sigma_\alpha \right\|_2 \\
&\leq \left\| I - U^{(k)} \circ \overline{V}^{(k)} \right\|_F \cdot \|d^{(k)}\|_2 + \left\| |d^{(k-1)}| - \sigma_\alpha \right\|_2 \\
&< 2\gamma/3
\end{aligned}$$

Convergence of the diagonal

We have shown:

$$d^{(k-1)} = \left(U^{(k)} \circ \overline{V}^{(k)}\right) \cdot d^{(k)} \ .$$

If

$$\left\| |d^{(k-1)}| - \sigma_\alpha \right\|_2 \leq \gamma/3 \ ,$$

then

$$\begin{aligned}
\left\| |d^{(k)}| - \sigma_\alpha \right\|_2 &\leq \left\| |d^{(k)}| - |d^{(k-1)}| \right\|_2 + \left\| |d^{(k-1)}| - \sigma_\alpha \right\|_2 \\
&\leq \left\| I - U^{(k)} \circ \overline{V}^{(k)} \right\|_F \cdot \|d^{(k)}\|_2 + \left\| |d^{(k-1)}| - \sigma_\alpha \right\|_2 \\
&< 2\gamma/3
\end{aligned}$$

Convergence of the diagonal

One can show that

$$\left\| I - U^{(k)} \circ \overline{V}^{(k)} \right\|_F \leq c_1 \cdot \mathrm{off}^2(A^{(k)}) \ , c_1 = c_1(n, \gamma)$$

and the ultimate quadratic convergence of the process:

$$\mathrm{off}(A^{(k+s)}) \leq c_2 \cdot \mathrm{off}^2(A^{(k)}) \ , c_2 = c_2(n, \gamma) \ .$$

Thus:

Convergence of the diagonal

$$
\begin{aligned}
\|d^{(k+p)} - d^{(k)}\|_2 &\leq \sum_{i=1}^{p} \|d^{(k+i)} - d^{(k+i-1)}\|_2 \\
&\leq \sum_{i=1}^{p} \|I - U^{(k+i)} \circ \overline{V}^{(k+i)}\|_F \cdot \|d^{(k+i)}\|_2 \\
&\leq \|A\|_F \cdot \sum_{i=1}^{p} c_1 \cdot \operatorname{off}^2(A^{(k+i)}) \\
&\leq c_1 \|A\|_F \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \operatorname{off}^2(A^{(i+k\cdot s)}) \\
&\leq c_1 \|A\|_F \cdot \frac{1}{c_2^2} \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \left(c_2 \operatorname{off}(A^{(i)}) \right)^{2^k+1} \\
&\leq \frac{c_1 \|A\|_F \cdot s}{c_2^2} \cdot \frac{c_2 \operatorname{off}(A^{(k)})}{1 - c_2 \operatorname{off}(A^{(k)})} \xrightarrow{k} 0 \,.
\end{aligned}
$$

Convergence of the diagonal

$$
\begin{aligned}
\|d^{(k+p)} - d^{(k)}\|_2 &\leq \sum_{i=1}^{p} \|d^{(k+i)} - d^{(k+i-1)}\|_2 \\
&\leq \sum_{i=1}^{p} \|I - U^{(k+i)} \circ \overline{V}^{(k+i)}\|_F \cdot \|d^{(k+i)}\|_2 \\
&\leq \|A\|_F \cdot \sum_{i=1}^{p} c_1 \cdot \mathrm{off}^2(A^{(k+i)}) \\
&\leq c_1 \|A\|_F \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \mathrm{off}^2(A^{(i+k\cdot s)}) \\
&\leq c_1 \|A\|_F \cdot \frac{1}{c_2^2} \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \left(c_2\, \mathrm{off}(A^{(i)})\right)^{2^k+1} \\
&\leq \frac{c_1 \|A\|_F \cdot s}{c_2^2} \cdot \frac{c_2\, \mathrm{off}(A^{(k)})}{1 - c_2\, \mathrm{off}(A^{(k)})} \xrightarrow{k} 0 \,.
\end{aligned}
$$

Convergence of the diagonal

$$
\begin{aligned}
\|d^{(k+p)} - d^{(k)}\|_2 &\leq \sum_{i=1}^{p} \|d^{(k+i)} - d^{(k+i-1)}\|_2 \\
&\leq \sum_{i=1}^{p} \|I - U^{(k+i)} \circ \overline{V}^{(k+i)}\|_F \cdot \|d^{(k+i)}\|_2 \\
&\leq \|A\|_F \cdot \sum_{i=1}^{p} c_1 \cdot \text{off}^2(A^{(k+i)}) \\
&\leq c_1 \|A\|_F \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \text{off}^2(A^{(i+k\cdot s)}) \\
&\leq c_1 \|A\|_F \cdot \frac{1}{c_2^2} \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \left(c_2 \, \text{off}(A^{(i)})\right)^{2^k+1} \\
&\leq \frac{c_1 \|A\|_F \cdot s}{c_2^2} \cdot \frac{c_2 \, \text{off}(A^{(k)})}{1 - c_2 \, \text{off}(A^{(k)})} \xrightarrow{k} 0 \, .
\end{aligned}
$$

Convergence of the diagonal

$$
\begin{aligned}
\|d^{(k+p)} - d^{(k)}\|_2 &\leq \sum_{i=1}^{p} \|d^{(k+i)} - d^{(k+i-1)}\|_2 \\
&\leq \sum_{i=1}^{p} \|I - U^{(k+i)} \circ \overline{V}^{(k+i)}\|_F \cdot \|d^{(k+i)}\|_2 \\
&\leq \|A\|_F \cdot \sum_{i=1}^{p} c_1 \cdot \mathrm{off}^2(A^{(k+i)}) \\
&\leq c_1 \|A\|_F \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \mathrm{off}^2(A^{(i+k\cdot s)}) \\
&\leq c_1 \|A\|_F \cdot \frac{1}{c_2^2} \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \left(c_2 \,\mathrm{off}(A^{(i)})\right)^{2^k+1} \\
&\leq \frac{c_1 \|A\|_F \cdot s}{c_2^2} \cdot \frac{c_2 \,\mathrm{off}(A^{(k)})}{1 - c_2 \,\mathrm{off}(A^{(k)})} \xrightarrow{k} 0 \,.
\end{aligned}
$$

Convergence of the diagonal

$$
\begin{aligned}
\|d^{(k+p)} - d^{(k)}\|_2 &\leq \sum_{i=1}^{p} \|d^{(k+i)} - d^{(k+i-1)}\|_2 \\
&\leq \sum_{i=1}^{p} \|I - U^{(k+i)} \circ \overline{V}^{(k+i)}\|_F \cdot \|d^{(k+i)}\|_2 \\
&\leq \|A\|_F \cdot \sum_{i=1}^{p} c_1 \cdot \operatorname{off}^2(A^{(k+i)}) \\
&\leq c_1 \|A\|_F \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \operatorname{off}^2(A^{(i+k \cdot s)}) \\
&\leq c_1 \|A\|_F \cdot \frac{1}{c_2^2} \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \left(c_2 \operatorname{off}(A^{(i)}) \right)^{2^k+1} \\
&\leq \frac{c_1 \|A\|_F \cdot s}{c_2^2} \cdot \frac{c_2 \operatorname{off}(A^{(k)})}{1 - c_2 \operatorname{off}(A^{(k)})} \xrightarrow{k} 0 \, .
\end{aligned}
$$

Convergence of the diagonal

$$
\begin{aligned}
\|d^{(k+p)} - d^{(k)}\|_2 &\leq \sum_{i=1}^{p} \|d^{(k+i)} - d^{(k+i-1)}\|_2 \\
&\leq \sum_{i=1}^{p} \|I - U^{(k+i)} \circ \overline{V}^{(k+i)}\|_F \cdot \|d^{(k+i)}\|_2 \\
&\leq \|A\|_F \cdot \sum_{i=1}^{p} c_1 \cdot \mathrm{off}^2(A^{(k+i)}) \\
&\leq c_1 \|A\|_F \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \mathrm{off}^2(A^{(i+k\cdot s)}) \\
&\leq c_1 \|A\|_F \cdot \frac{1}{c_2^2} \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \left(c_2\, \mathrm{off}(A^{(i)})\right)^{2^k+1} \\
&\leq \frac{c_1 \|A\|_F \cdot s}{c_2^2} \cdot \frac{c_2\, \mathrm{off}(A^{(k)})}{1 - c_2\, \mathrm{off}(A^{(k)})} \xrightarrow{k} 0 \,.
\end{aligned}
$$

Convergence of the diagonal

$$
\begin{aligned}
\|d^{(k+p)} - d^{(k)}\|_2 &\le \sum_{i=1}^{p} \|d^{(k+i)} - d^{(k+i-1)}\|_2 \\
&\le \sum_{i=1}^{p} \|I - U^{(k+i)} \circ \overline{V}^{(k+i)}\|_F \cdot \|d^{(k+i)}\|_2 \\
&\le \|A\|_F \cdot \sum_{i=1}^{p} c_1 \cdot \mathrm{off}^2(A^{(k+i)}) \\
&\le c_1 \|A\|_F \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \mathrm{off}^2(A^{(i+k\cdot s)}) \\
&\le c_1 \|A\|_F \cdot \frac{1}{c_2^2} \cdot \sum_{i=k}^{k+s-1} \sum_{k=0}^{\infty} \left(c_2\, \mathrm{off}(A^{(i)}) \right)^{2^k+1} \\
&\le \frac{c_1 \|A\|_F \cdot s}{c_2^2} \cdot \frac{c_2\, \mathrm{off}(A^{(k)})}{1 - c_2\, \mathrm{off}(A^{(k)})} \xrightarrow{k} 0 \,.
\end{aligned}
$$

Products of rotations

Use the same trick to show $\sum_{k=1}^{\infty} \|E^{(k)}\|_F < \infty$, $E^{(k)} = I - U^{(k)}$.

Thus:

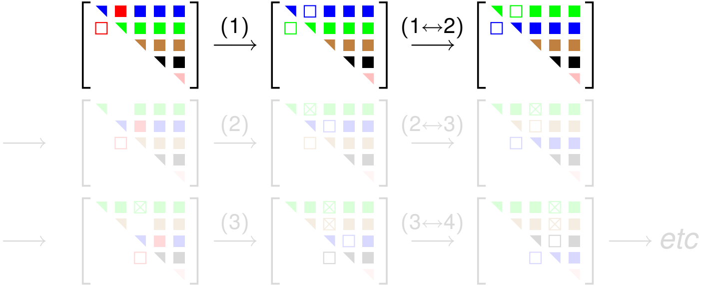
- $U^{(1)} \cdot U^{(2)} \cdot \ldots \cdot U^{(k)} \cdot \ldots$ converges;
- $V^{(1)} \cdot V^{(2)} \cdot \ldots \cdot V^{(k)} \cdot \ldots$ converges.

Then $A = (\prod_k U^{(k)}) \cdot (\lim_k A^{(k)}) \cdot (\prod_k V^{(k)})^\star$.

Concluding remarks

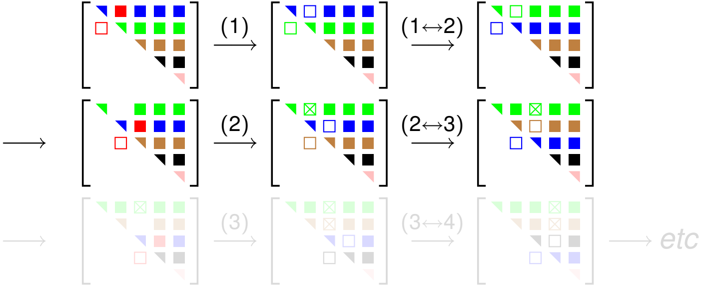
- We have theoretical foundations for block versions of the classical Jacobi and the Kogbetliantz SVD algorithms. The convergence is well understood in the general case of multiple singular values and eigenvalues.
- For both algorithms, we have an efficient preconditioner, based on a rank revealing QR factorization. This preconditioning improves the convergence considerably, and, with proper pivot strategies, higher order convergence may kick in early in the process.
- Our UBC transformations can be implemented as fast scaled block transformations proposed by Hari. This is work in progress and we plan to have new fully BLAS 3 Jacobi++/Kogbetliantz–type SVD software in a near future.

Gentleman's trick



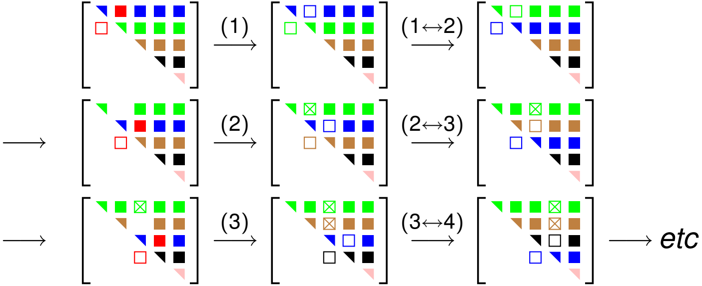
Back

Gentleman's trick



Back

Gentleman's trick



Back

Wedin's $\Phi - \Theta$ theorem

Theorem (Wedin's $\Phi - \Theta$ theorem)

Let B and $\tilde{B}$ have SVDs:

$$B = [\, U_1 \;\; U_2 \,] \cdot \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \cdot [\, V_1 \;\; V_2 \,]^\star \text{ and } \tilde{B} = [\, \tilde{U}_1 \;\; \tilde{U}_2 \,] \cdot \begin{bmatrix} \tilde{\Sigma}_1 & 0 \\ 0 & \tilde{\Sigma}_2 \end{bmatrix} \cdot [\, \tilde{V}_1 \;\; \tilde{V}_2 \,]^\star \; .$$

If $\delta > 0$ is such that $\min |\sigma(\tilde{\Sigma}_1) - \sigma(\Sigma_2)| \geq \delta$ and $\min \sigma(\tilde{\Sigma}_1) \geq \delta$, then

$$\sqrt{\|\sin\Phi\|_F^2 + \|\sin\Theta\|_F^2} \leq \frac{\sqrt{\|R\|_F^2 + \|S\|_F^2}}{\delta} \; .$$

Here $R = B\tilde{V}_1 - \tilde{U}_1\tilde{\Sigma}_1$ and $S = B^\star\tilde{U}_1 - \tilde{V}_1\tilde{\Sigma}_1$.
$\Phi = \measuredangle(\mathcal{R}(U_1), \mathcal{R}(\tilde{U}_1))$, $\Theta = \measuredangle(\mathcal{R}(V_1), \mathcal{R}(\tilde{V}_1))$.

Take $B = \text{diag}(A_{2x2}^{(k-1)})$ and $\tilde{B} = A_{2x2}^{(k-1)}$ as its perturbation.

Back

Wedin's Φ – Θ theorem

Theorem (Wedin's Φ – Θ theorem)

Let B and $\tilde{B}$ have SVDs:

$$B = [\, U_1 \;\; U_2 \,] \cdot \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \cdot [\, V_1 \;\; V_2 \,]^\star \;\; \textit{and} \;\; \tilde{B} = [\, \tilde{U}_1 \;\; \tilde{U}_2 \,] \cdot \begin{bmatrix} \tilde{\Sigma}_1 & 0 \\ 0 & \tilde{\Sigma}_2 \end{bmatrix} \cdot [\, \tilde{V}_1 \;\; \tilde{V}_2 \,]^\star \; .$$

If $\delta > 0$ is such that $\min |\sigma(\tilde{\Sigma}_1) - \sigma(\Sigma_2)| \geq \delta$ and $\min \sigma(\tilde{\Sigma}_1) \geq \delta$, then

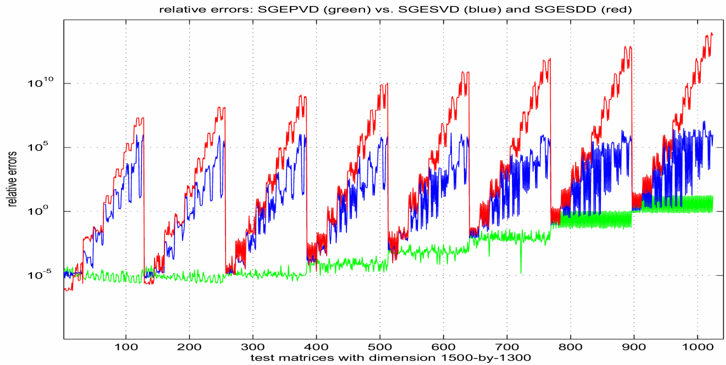
$$\sqrt{\|\sin \Phi\|_F^2 + \|\sin \Theta\|_F^2} \leq \frac{\sqrt{\|R\|_F^2 + \|S\|_F^2}}{\delta} \; .$$

Here $R = B\tilde{V}_1 - \tilde{U}_1\tilde{\Sigma}_1$ and $S = B^\star \tilde{U}_1 - \tilde{V}_1\tilde{\Sigma}_1$.
$\Phi = \measuredangle(\mathcal{R}(U_1), \mathcal{R}(\tilde{U}_1))$, $\Theta = \measuredangle(\mathcal{R}(V_1), \mathcal{R}(\tilde{V}_1))$.

Take $B = \mathrm{diag}(A_{2x2}^{(k-1)})$ and $\tilde{B} = A_{2x2}^{(k-1)}$ as its perturbation.

Back

Relative accuracy: example

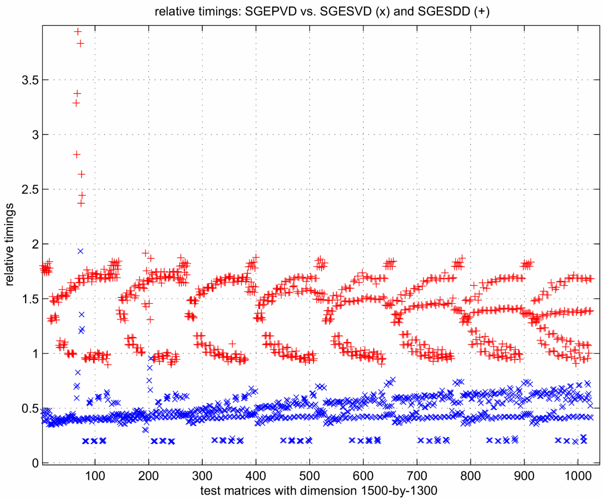


$A = BD \in \mathbb{R}^{1500\times 1300}$, $\|B(:,i)\| = 1$, $D = \mathrm{diag}$

$\kappa(B) = 10^k$, $k = 1:8$; $\kappa(D) \in \{10^2, 10^4, ..., 10^{16}\}$

Accuracy: $\max_i \frac{|\delta\sigma_i|}{\sigma_i}$ of Jacobi, QR, DC ▸ Back

Improved efficiency



Timings: Jacobi/SGESVD, Jacobi/SGESDD ▸ Back