Diagonal scalings for improving the accuracy of computed eigenvalues of arbitrary pencils

Froilán M. Dopico

joint work with **María C. Quintana** (Aalto University, Finland), and **Paul Van Dooren** (UC Louvain, Belgium)

Depto de Matemáticas, Universidad Carlos III de Madrid, Spain

Part of “Proyecto de I+D+i PID2019-106362GB-I00 financiado por MCIN/AEI/10.13039/501100011033”

Online Seminar in Linear Algebra and Operator Theory (OSELOT)
April 21, 2022

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN

AGENCIA ESTATAL DE INVESTIGACIÓN

uc3m | Universidad Carlos III de Madrid

- It is a well-known pre-processing of a non-normal matrix A before computing its eigenvalues with the QR algorithm, or with any other backward stable algorithm, for improving the accuracy of the computed eigenvalues.
- It is performed by default by command `eig` in MATLAB.
- Given

$$A \in \mathbb{C}^{n\times n} \longmapsto B = D^{-1}AD$$

- with D positive diagonal matrix whose entries are integer powers of 2, (no rounding errors in computing B) and
- $\|\text{col}_i(B)\|_2 \approx \|\text{row}_i(B)\|_2, \quad i = 1, 2, \ldots, n.$
- **The eigenvalue algorithm is applied to B!!**
- LAPACK uses $\|\cdot\|_1$, we will use $\|\cdot\|_2$ throughout the talk for vectors and $\|\cdot\|_F$ for matrices. These choices have better "theoretical" properties and better numerical properties (James, Langou, Lowery, 2014).

- It is a well-known pre-processing of a non-normal matrix A before computing its eigenvalues with the QR algorithm, or with any other backward stable algorithm, for improving the accuracy of the computed eigenvalues.
- It is performed by default by command `eig` in MATLAB.
- Given

$$A \in \mathbb{C}^{n\times n} \longmapsto B = D^{-1}AD$$

- with D positive diagonal matrix whose entries are integer powers of 2, (no rounding errors in computing B) and
- $\|\text{col}_i(B)\|_2 \approx \|\text{row}_i(B)\|_2, \quad i = 1, 2, \ldots, n.$
- **The eigenvalue algorithm is applied to B!!**
- LAPACK uses $\|\cdot\|_1$, we will use $\|\cdot\|_2$ throughout the talk for vectors and $\|\cdot\|_F$ for matrices. These choices have better "theoretical" properties and better numerical properties (James, Langou, Lowery, 2014).

Preliminaries: Balancing matrices for computing eigenvalues (I)

- It is a well-known pre-processing of a non-normal matrix A before computing its eigenvalues with the QR algorithm, or with any other backward stable algorithm, for improving the accuracy of the computed eigenvalues.
- It is performed by default by command `eig` in MATLAB.
- Given

$$\boxed{A \in \mathbb{C}^{n\times n} \longmapsto B = D^{-1}AD}$$

- with D positive diagonal matrix whose entries are integer powers of 2, (no rounding errors in computing B) and
- $\|\text{col}_i(B)\|_2 \approx \|\text{row}_i(B)\|_2, \quad i = 1, 2, \ldots, n.$
- **The eigenvalue algorithm is applied to B!!**
- LAPACK uses $\|\cdot\|_1$, we will use $\|\cdot\|_2$ throughout the talk for vectors and $\|\cdot\|_F$ for matrices. These choices have better "theoretical" properties and better numerical properties (James, Langou, Lowery, 2014).

- It is a well-known pre-processing of a non-normal matrix A before computing its eigenvalues with the QR algorithm, or with any other backward stable algorithm, for improving the accuracy of the computed eigenvalues.
- It is performed by default by command `eig` in MATLAB.
- Given

$$A \in \mathbb{C}^{n \times n} \longmapsto B = D^{-1} A D$$

- with D positive diagonal matrix whose entries are integer powers of 2, (no rounding errors in computing B) and
- $\|\text{col}_i(B)\|_2 \approx \|\text{row}_i(B)\|_2, \quad i = 1, 2, \ldots, n.$
- **The eigenvalue algorithm is applied to B!!**
- LAPACK uses $\|\cdot\|_1$, we will use $\|\cdot\|_2$ throughout the talk for vectors and $\|\cdot\|_F$ for matrices. These choices have better "theoretical" properties and better numerical properties (James, Langou, Lowery, 2014).

Preliminaries: Balancing matrices for computing eigenvalues (I)

- It is a well-known pre-processing of a non-normal matrix A before computing its eigenvalues with the QR algorithm, or with any other backward stable algorithm, for improving the accuracy of the computed eigenvalues.
- It is performed by default by command `eig` in MATLAB.
- Given

$$\boxed{A \in \mathbb{C}^{n\times n} \longmapsto B = D^{-1}AD}$$

- with D positive diagonal matrix whose entries are integer powers of 2, (no rounding errors in computing B) and
- $\|\mathrm{col}_i(B)\|_2 \approx \|\mathrm{row}_i(B)\|_2, \quad i = 1, 2, \ldots, n.$
- **The eigenvalue algorithm is applied to B!!**
- LAPACK uses $\|\cdot\|_1$, we will use $\|\cdot\|_2$ throughout the talk for vectors and $\|\cdot\|_F$ for matrices. These choices have better "theoretical" properties and better numerical properties (James, Langou, Lowery, 2014).

- It is a well-known pre-processing of a non-normal matrix A before computing its eigenvalues with the QR algorithm, or with any other backward stable algorithm, for improving the accuracy of the computed eigenvalues.
- It is performed by default by command `eig` in MATLAB.
- Given

$$\boxed{A \in \mathbb{C}^{n\times n} \longmapsto B = D^{-1} A D}$$

- with D positive diagonal matrix whose entries are integer powers of 2, (no rounding errors in computing B) and
- $\|\text{col}_i(B)\|_2 \approx \|\text{row}_i(B)\|_2, \quad i = 1, 2, \ldots, n.$
- **The eigenvalue algorithm is applied to B!!**
- LAPACK uses $\|\cdot\|_1$, we will use $\|\cdot\|_2$ throughout the talk for vectors and $\|\cdot\|_F$ for matrices. These choices have better "theoretical" properties and better numerical properties (James, Langou, Lowery, 2014).

- It is a well-known pre-processing of a non-normal matrix A before computing its eigenvalues with the QR algorithm, or with any other backward stable algorithm, for improving the accuracy of the computed eigenvalues.
- It is performed by default by command `eig` in MATLAB.
- Given

$$\boxed{A \in \mathbb{C}^{n\times n} \longmapsto B = D^{-1} A D}$$

- with D positive diagonal matrix whose entries are integer powers of 2, (no rounding errors in computing B) and
- $\|\text{col}_i(B)\|_2 \approx \|\text{row}_i(B)\|_2, \quad i = 1, 2, \ldots, n.$
- **The eigenvalue algorithm is applied to B!!**
- LAPACK uses $\|\cdot\|_1$, we will use $\|\cdot\|_2$ throughout the talk for vectors and $\|\cdot\|_F$ for matrices. These choices have better “theoretical” properties and better numerical properties (James, Langou, Lowery, 2014).

Preliminaries: Balancing matrices for computing eigenvalues (II)

- The basic algorithm for balancing is a cyclic iterative procedure proposed by Osborne (1960) and by Parlett-Reinsch (1969, D_{ii} integer power of 2) that starts with $D = I_n$, updates one diagonal entry of D and one row and one column of A in each step making their norms equal:

$$f = \sqrt{\frac{\|\mathrm{row}_i(A)\|_2}{\|\mathrm{col}_i(A)\|_2}}$$

$$d_{ii} \longleftarrow f \cdot d_{ii}$$

$$\mathrm{col}_i(A) \longleftarrow f \cdot \mathrm{col}_i(A)$$

$$\mathrm{row}_i(A) \longleftarrow \mathrm{row}_i(A)/f$$

- If A is irreducible, then the algorithm converges to $B = D^{-1}AD$ such that

$$\|\mathrm{row}_i(B)\|_2 = \|\mathrm{col}_i(B)\|_2, \quad i = 1, \ldots, n \iff \|B\|_F = \inf_{D \text{ diagonal}} \|D^{-1}AD\|_F$$

- These equalities become approximate if the entries of D are restricted to be integer powers of 2,
- in this case, the process converges in general quickly and costs $O(n^2)$ flops, which is negligible with respect to $O(n^3)$ cost of QR.

- The basic algorithm for balancing is a cyclic iterative procedure proposed by Osborne (1960) and by Parlett-Reinsch (1969, D_{ii} integer power of 2) that starts with $D = I_n$, updates one diagonal entry of D and one row and one column of A in each step making their norms equal:

$$f = \sqrt{\frac{\|\text{row}_i(A)\|_2}{\|\text{col}_i(A)\|_2}}$$

$$d_{ii} \longleftarrow f \cdot d_{ii}$$

$$\text{col}_i(A) \longleftarrow f \cdot \text{col}_i(A)$$

$$\text{row}_i(A) \longleftarrow \text{row}_i(A)/f$$

- If A is irreducible, then the algorithm converges to $B = D^{-1}AD$ such that

$$\|\text{row}_i(B)\|_2 = \|\text{col}_i(B)\|_2, \quad i = 1, \dots, n \iff \|B\|_F = \inf_{D \text{ diagonal}} \|D^{-1}AD\|_F$$

- These equalities become approximate if the entries of D are restricted to be integer powers of 2,
- in this case, the process converges in general quickly and costs $O(n^2)$ flops, which is negligible with respect to $O(n^3)$ cost of QR.

- The basic algorithm for balancing is a cyclic iterative procedure proposed by Osborne (1960) and by Parlett-Reinsch (1969, D_{ii} integer power of 2) that starts with $D = I_n$, updates one diagonal entry of D and one row and one column of A in each step making their norms equal:

$$f = \sqrt{\frac{\|\text{row}_i(A)\|_2}{\|\text{col}_i(A)\|_2}}$$

$$d_{ii} \longleftarrow f \cdot d_{ii}$$

$$\text{col}_i(A) \longleftarrow f \cdot \text{col}_i(A)$$

$$\text{row}_i(A) \longleftarrow \text{row}_i(A)/f$$

- If A is irreducible, then the algorithm converges to $B = D^{-1}AD$ such that

$$\|\text{row}_i(B)\|_2 = \|\text{col}_i(B)\|_2, \quad i = 1, \ldots, n \iff \|B\|_F = \inf_{D \text{ diagonal}} \|D^{-1}AD\|_F$$

- These equalities become approximate if the entries of D are restricted to be integer powers of 2,
- in this case, the process converges in general quickly and costs $O(n^2)$ flops, which is negligible with respect to $O(n^3)$ cost of QR.

- The basic algorithm for balancing is a cyclic iterative procedure proposed by Osborne (1960) and by Parlett-Reinsch (1969, D_{ii} integer power of 2) that starts with $D = I_n$, updates one diagonal entry of D and one row and one column of A in each step making their norms equal:

$$f = \sqrt{\frac{\|\mathrm{row}_i(A)\|_2}{\|\mathrm{col}_i(A)\|_2}}$$

$$d_{ii} \longleftarrow f \cdot d_{ii}$$

$$\mathrm{col}_i(A) \longleftarrow f \cdot \mathrm{col}_i(A)$$

$$\mathrm{row}_i(A) \longleftarrow \mathrm{row}_i(A)/f$$

- If A is irreducible, then the algorithm converges to $B = D^{-1}AD$ such that

$$\|\mathrm{row}_i(B)\|_2 = \|\mathrm{col}_i(B)\|_2, \quad i = 1, \ldots, n \iff \|B\|_F = \inf_{D \text{ diagonal}} \|D^{-1}AD\|_F$$

- These equalities become approximate if the entries of D are restricted to be integer powers of 2,
- in this case, the process converges in general quickly and costs $O(n^2)$ flops, which is negligible with respect to $O(n^3)$ cost of QR.

Why does balancing improve often the accuracy of computed e-values?

- The computed eigenvalues $\widehat{\lambda}_A$ of $A \in \mathbb{C}^{n \times n}$, via a backward stable algorithm, are the exact ones of $A + E$ with $\|E\|_F = O(\varepsilon)\|A\|_F$, with $\varepsilon \approx 10^{-16}$ the unit roundoff of the computer.
- Thus, up to $O(\varepsilon^2)$,

$$|\widehat{\lambda}_A - \lambda| \leq O(\varepsilon)\, \frac{\|y\|_2\, \|x\|_2}{|y^* x|}\, \|A\|_F\,,$$

where, λ is a exact simple eigenvalue of A, $Ax = \lambda x$ and $y^* A = \lambda y^*$, and $\frac{\|y\|_2\, \|x\|_2}{|y^* x|}$ is the Wilkinson-eigenvalue condition number.
- If we compute instead the eigenvalues $\widehat{\lambda}_B$ of $B = D^{-1}AD$ with the same exact eigenvalues as A, then

$$|\widehat{\lambda}_B - \lambda| \leq O(\varepsilon)\, \frac{\|Dy\|_2\, \|D^{-1}x\|_2}{|y^* x|}\, \|B\|_F\,.$$

- If $\|B\|_F < \|A\|_F$, one of the factors in the error bound decreases, but what happens with the other one?, i.e., with the eigenvalue condition number?

Why does balancing improve often the accuracy of computed e-values?

- The computed eigenvalues $\widehat{\lambda}_A$ of $A \in \mathbb{C}^{n\times n}$, via a backward stable algorithm, are the exact ones of $A+E$ with $\|E\|_F = O(\varepsilon)\|A\|_F$, with $\varepsilon \approx 10^{-16}$ the unit roundoff of the computer.
- Thus, up to $O(\varepsilon^2)$,

$$|\widehat{\lambda}_A - \lambda| \leq O(\varepsilon)\, \frac{\|y\|_2\, \|x\|_2}{|y^* x|}\, \|A\|_F\,,$$

  where, λ is a exact simple eigenvalue of A, $Ax = \lambda x$ and $y^* A = \lambda y^*$, and $\frac{\|y\|_2\, \|x\|_2}{|y^* x|}$ is the Wilkinson-eigenvalue condition number.
- If we compute instead the eigenvalues $\widehat{\lambda}_B$ of $B = D^{-1}AD$ with the same exact eigenvalues as A, then

$$|\widehat{\lambda}_B - \lambda| \leq O(\varepsilon)\, \frac{\|Dy\|_2\, \|D^{-1}x\|_2}{|y^* x|}\, \|B\|_F\,.$$

- If $\|B\|_F < \|A\|_F$, one of the factors in the error bound decreases, but what happens with the other one?, i.e., with the eigenvalue condition number?

Why does balancing improve often the accuracy of computed e-values?

- The computed eigenvalues $\widehat{\lambda}_A$ of $A \in \mathbb{C}^{n\times n}$, via a backward stable algorithm, are the exact ones of $A+E$ with $\|E\|_F = O(\varepsilon)\|A\|_F$, with $\varepsilon \approx 10^{-16}$ the unit roundoff of the computer.

- Thus, up to $O(\varepsilon^2)$,

$$|\widehat{\lambda}_A - \lambda| \leq O(\varepsilon)\, \frac{\|y\|_2\, \|x\|_2}{|y^* x|}\, \|A\|_F \,,$$

where, λ is a exact simple eigenvalue of A, $Ax = \lambda x$ and $y^*A = \lambda y^*$, and $\frac{\|y\|_2\, \|x\|_2}{|y^* x|}$ is the Wilkinson-eigenvalue condition number.

- If we compute instead the eigenvalues $\widehat{\lambda}_B$ of $B = D^{-1}AD$ with the same exact eigenvalues as A, then

$$|\widehat{\lambda}_B - \lambda| \leq O(\varepsilon)\, \frac{\|Dy\|_2\, \|D^{-1}x\|_2}{|y^* x|}\, \|B\|_F \,.$$

- If $\|B\|_F < \|A\|_F$, one of the factors in the error bound decreases, but what happens with the other one?, i.e., with the eigenvalue condition number?

Why does balancing improve often the accuracy of computed e-values?

- The computed eigenvalues $\widehat{\lambda}_A$ of $A \in \mathbb{C}^{n\times n}$, via a backward stable algorithm, are the exact ones of $A + E$ with $\|E\|_F = O(\varepsilon)\|A\|_F$, with $\varepsilon \approx 10^{-16}$ the unit roundoff of the computer.
- Thus, up to $O(\varepsilon^2)$,

$$|\widehat{\lambda}_A - \lambda| \leq O(\varepsilon)\, \frac{\|y\|_2\, \|x\|_2}{|y^* x|}\, \|A\|_F\,,$$

  where, λ is a exact simple eigenvalue of A, $Ax = \lambda x$ and $y^* A = \lambda y^*$, and $\frac{\|y\|_2\, \|x\|_2}{|y^* x|}$ is the Wilkinson-eigenvalue condition number.
- If we compute instead the eigenvalues $\widehat{\lambda}_B$ of $B = D^{-1}AD$ with the same exact eigenvalues as A, then

$$|\widehat{\lambda}_B - \lambda| \leq O(\varepsilon)\, \frac{\|Dy\|_2\, \|D^{-1}x\|_2}{|y^* x|}\, \|B\|_F\,.$$

- If $\|B\|_F < \|A\|_F$, one of the factors in the error bound decreases, but what happens with the other one?, i.e., with the eigenvalue condition number?

Why does balancing improve often the accuracy of comput e-values? (2)

- The condition number also "often decreases".
- An explanation for this is that if $A \in \mathbb{C}^{n\times n}$ is diagonalizable with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then

$$\|\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\|_F = \min_{S \text{ invertible}} \|S^{-1}AS\|_F$$

  is attained at

$$S^{-1}AS = U \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) U^*,$$

  with U an arbitrary unitary matrix, which have the smallest possible eigenvalue condition numbers all equal to 1.
- Thus, balancing will likely improve the eigenvalue condition numbers, since it solves the same minimization problem but restricted to diagonal invertible matrices.
- It is known that there are matrices for which "balancing" yields larger errors of computed eigenvalues than "no-balancing". See for instance Watkins (2006).

Why does balancing improve often the accuracy of comput e-values? (2)

- The condition number also “often decreases”.
- An explanation for this is that if $A \in \mathbb{C}^{n\times n}$ is diagonalizable with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then

$$\|\operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)\|_F = \min_{S \text{ invertible}} \|S^{-1}AS\|_F$$

  is attained at

$$S^{-1}AS = U \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) U^*,$$

  with U an arbitrary unitary matrix, which have the smallest possible eigenvalue condition numbers all equal to 1.
- Thus, balancing will likely improve the eigenvalue condition numbers, since it solves the same minimization problem but restricted to diagonal invertible matrices.
- It is known that there are matrices for which “balancing” yields larger errors of computed eigenvalues than “no-balancing”. See for instance Watkins (2006).

Why does balancing improve often the accuracy of comput e-values? (2)

- The condition number also "often decreases".
- An explanation for this is that if $A \in \mathbb{C}^{n \times n}$ is diagonalizable with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then

$$\| \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \|_F = \min_{S \text{ invertible}} \|S^{-1} A S\|_F$$

  is attained at

$$S^{-1} A S = U \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) U^*,$$

  with U an arbitrary unitary matrix, which have the smallest possible eigenvalue condition numbers all equal to 1.
- Thus, balancing will likely improve the eigenvalue condition numbers, since it solves the same minimization problem but restricted to diagonal invertible matrices.
- It is known that there are matrices for which "balancing" yields larger errors of computed eigenvalues than "no-balancing". See for instance Watkins (2006).

Why does balancing improve often the accuracy of comput e-values? (2)

- The condition number also "often decreases".
- An explanation for this is that if $A \in \mathbb{C}^{n \times n}$ is diagonalizable with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then

$$\| \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \|_F = \min_{S \text{ invertible}} \| S^{-1} A S \|_F$$

  is attained at

$$S^{-1} A S = U \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) U^*,$$

  with U an arbitrary unitary matrix, which have the smallest possible eigenvalue condition numbers all equal to 1.
- Thus, balancing will likely improve the eigenvalue condition numbers, since it solves the same minimization problem but restricted to diagonal invertible matrices.
- It is known that there are matrices for which "balancing" yields larger errors of computed eigenvalues than "no-balancing". See for instance Watkins (2006).

Outline

1. Previous results for balancing regular pencils
2. Connecting the problem to the Sinkhorn-Knopp algorithm
3. Diagonal scalings of rectangular pencils
4. Regularized scaling methods for pencils
5. Conclusions

Outline

Some comments on regular matrix pencils

We consider first $\lambda B - A$, with $A, B \in \mathbb{C}^{n\times n}$ and $\det(\lambda B - A)$ not identically zero, i.e., the pencil is regular or the generalized eigenvalue problem is regular.

- Note that the eigenvalues are now invariant under strict equivalence

$$\lambda B - A \longrightarrow T_\ell(\lambda B - A)T_r = \lambda T_\ell B T_r - T_\ell B T_r,$$

  with $T_\ell, T_r \in \mathbb{C}^{n\times n}$ invertible (and different to each other).

- Thus, balancing or diagonally scaling a pencil, is

$$\boxed{\lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}}$$

  with D_ℓ, D_r nonsingular positive diagonal matrices whose entries are integer powers of 2 before applying the QZ algorithm.

- Observe that the purpose of such diagonal scalings cannot be simply to decrease the norms of $D_\ell B D_r$ and of $D_\ell A D_r$ since these norms can be made arbitrarily small without changing the eigenvalues, just by multiplying the pencil by a small number.

Some comments on regular matrix pencils

We consider first $\lambda B - A$, with $A, B \in \mathbb{C}^{n\times n}$ and $\det(\lambda B - A)$ not identically zero, i.e., the pencil is regular or the generalized eigenvalue problem is regular.

- Note that the eigenvalues are now invariant under strict equivalence

$$\lambda B - A \longrightarrow T_\ell(\lambda B - A)T_r = \lambda T_\ell B T_r - T_\ell B T_r,$$

with $T_\ell, T_r \in \mathbb{C}^{n\times n}$ invertible (and different to each other).

- Thus, balancing or diagonally scaling a pencil, is

$$\lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

with D_ℓ, D_r nonsingular positive diagonal matrices whose entries are integer powers of 2 before applying the QZ algorithm.

- Observe that the purpose of such diagonal scalings cannot be simply to decrease the norms of $D_\ell B D_r$ and of $D_\ell A D_r$ since these norms can be made arbitrarily small without changing the eigenvalues, just by multiplying the pencil by a small number.

Some comments on regular matrix pencils

We consider first $\lambda B - A$, with $A, B \in \mathbb{C}^{n\times n}$ and $\det(\lambda B - A)$ not identically zero, i.e., the pencil is regular or the generalized eigenvalue problem is regular.

- Note that the eigenvalues are now invariant under strict equivalence

$$\lambda B - A \longrightarrow T_\ell(\lambda B - A)T_r = \lambda T_\ell B T_r - T_\ell B T_r,$$

with $T_\ell, T_r \in \mathbb{C}^{n\times n}$ invertible (and different to each other).

- Thus, balancing or diagonally scaling a pencil, is

$$\boxed{\lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}}$$

with D_ℓ, D_r nonsingular positive diagonal matrices whose entries are integer powers of 2 before applying the QZ algorithm.

- Observe that the purpose of such diagonal scalings cannot be simply to decrease the norms of $D_\ell B D_r$ and of $D_\ell A D_r$ since these norms can be made arbitrarily small without changing the eigenvalues, just by multiplying the pencil by a small number.

We consider first $\lambda B - A$, with $A, B \in \mathbb{C}^{n \times n}$ and $\det(\lambda B - A)$ not identically zero, i.e., the pencil is regular or the generalized eigenvalue problem is regular.

- Note that the eigenvalues are now invariant under strict equivalence

$$\lambda B - A \longrightarrow T_\ell(\lambda B - A)T_r = \lambda T_\ell B T_r - T_\ell B T_r,$$

  with $T_\ell, T_r \in \mathbb{C}^{n \times n}$ invertible (and different to each other).

- Thus, balancing or diagonally scaling a pencil, is

$$\boxed{\lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}}$$

  with D_ℓ, D_r nonsingular positive diagonal matrices whose entries are integer powers of 2 before applying the QZ algorithm.

- Observe that the purpose of such diagonal scalings cannot be simply to decrease the norms of $D_\ell B D_r$ and of $D_\ell A D_r$ since these norms can be made arbitrarily small without changing the eigenvalues, just by multiplying the pencil by a small number.

Available methods for "balancing" or "diagonally scaling" regular pencils

$$A, B \in \mathbb{C}^{n\times n}. \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

- Ward's Method (1981) tries to get $\widetilde{A}$ and $\widetilde{B}$ so that the magnitude of each of their elements is as close to 1 as possible.
  - Available in LAPACK.
  - It is well known that it can severely deteriorate the accuracy of the computed eigenvalues of some pencils with entries of strongly varying order of magnitude (Kressner (2004), Lemonnier-Van Dooren (2006)).
- Lemonnier-Van Dooren's (LVD) method (2006) gets

$$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 \approx \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 \approx \gamma^2,$$

  for $i, j = 1, \ldots, n$, and some constant γ whose value is irrelevant.
- MATLAB does not include any built-in option for scaling pencils.

Available methods for “balancing” or “diagonally scaling” regular pencils

$$A, B \in \mathbb{C}^{n\times n}. \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

- Ward’s Method (1981) tries to get $\widetilde{A}$ and $\widetilde{B}$ so that the magnitude of each of their elements is as close to 1 as possible.
  - Available in LAPACK.
  - It is well known that it can severely deteriorate the accuracy of the computed eigenvalues of some pencils with entries of strongly varying order of magnitude (Kressner (2004), Lemonnier-Van Dooren (2006)).
- Lemonnier-Van Dooren’s (LVD) method (2006) gets

$$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 \approx \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 \approx \gamma^2,$$

for $i, j = 1, \ldots, n$, and some constant γ whose value is irrelevant.

- MATLAB does not include any built-in option for scaling pencils.

$$A, B \in \mathbb{C}^{n\times n}. \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

- Ward's Method (1981) tries to get $\widetilde{A}$ and $\widetilde{B}$ so that the magnitude of each of their elements is as close to 1 as possible.
  - Available in LAPACK.
  - It is well known that it can severely deteriorate the accuracy of the computed eigenvalues of some pencils with entries of strongly varying order of magnitude (Kressner (2004), Lemonnier-Van Dooren (2006)).
- Lemonnier-Van Dooren's (LVD) method (2006) gets

$$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 \approx \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 \approx \gamma^2,$$

  for $i, j = 1, \ldots, n$, and some constant γ whose value is irrelevant.
- MATLAB does not include any built-in option for scaling pencils.

$$A, B \in \mathbb{C}^{n\times n}. \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

- Ward's Method (1981) tries to get $\widetilde{A}$ and $\widetilde{B}$ so that the magnitude of each of their elements is as close to 1 as possible.
  - Available in LAPACK.
  - It is well known that it can severely deteriorate the accuracy of the computed eigenvalues of some pencils with entries of strongly varying order of magnitude (Kressner (2004), Lemonnier-Van Dooren (2006)).
- Lemonnier-Van Dooren's (LVD) method (2006) gets

$$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 \approx \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 \approx \gamma^2,$$

  for $i, j = 1, \ldots, n$, and some constant γ whose value is irrelevant.
- MATLAB does not include any built-in option for scaling pencils.

- Due to the fact that pencils may have "infinite eigenvalues" and that eigenvalues are computed by QZ as ratios of numbers, the "right way" to study the eigenvalue condition numbers of pencils is via the homogeneous formulation (Stewart-Sun, 1990), i.e. $A, B \in \mathbb{C}^{n\times n}$,

$$\lambda B - A \longleftrightarrow \alpha B - \beta A\,,$$

- where each eigenvalue $(\alpha,\beta) \neq (0,0)$ satisfying $\det(\alpha B - \beta A) = 0$ becomes a line $\langle \alpha, \beta \rangle$ through the origin in $\mathbb{C}^2$ ($\lambda = \alpha/\beta$ and $\infty \longleftrightarrow \langle 1, 0\rangle$)
- and the diference between two eigenvalues is measured in terms of the chordal metric:

$$\chi(\langle\alpha,\beta\rangle,\langle\gamma,\delta\rangle) = \frac{|\alpha\delta - \beta\gamma|}{\sqrt{|\alpha|^2+|\beta|^2}\,\sqrt{|\gamma|^2+|\delta|^2}}$$

Also

$$\sin\theta(\langle\alpha,\beta\rangle,\langle\gamma,\delta\rangle) = \chi(\langle\alpha,\beta\rangle,\langle\gamma,\delta\rangle)$$

- Due to the fact that pencils may have "infinite eigenvalues" and that eigenvalues are computed by QZ as ratios of numbers, the "right way" to study the eigenvalue condition numbers of pencils is via the homogeneous formulation (Stewart-Sun, 1990), i.e. $A, B \in \mathbb{C}^{n\times n}$,

$$\lambda B - A \longleftrightarrow \alpha B - \beta A\,,$$

- where each eigenvalue $(\alpha, \beta) \neq (0,0)$ satisfying $\det(\alpha B - \beta A) = 0$ becomes a line $\langle \alpha, \beta \rangle$ through the origin in $\mathbb{C}^2$ ($\lambda = \alpha/\beta$ and $\infty \longleftrightarrow \langle 1, 0\rangle$)
- and the diference between two eigenvalues is measured in terms of the chordal metric:

$$\chi(\langle \alpha, \beta\rangle, \langle \gamma, \delta\rangle) = \frac{|\alpha\delta - \beta\gamma|}{\sqrt{|\alpha|^2 + |\beta|^2}\sqrt{|\gamma|^2 + |\delta|^2}}$$

Also

$$\sin\theta(\langle \alpha, \beta\rangle, \langle \gamma, \delta\rangle) = \chi(\langle \alpha, \beta\rangle, \langle \gamma, \delta\rangle)$$

- Due to the fact that pencils may have "infinite eigenvalues" and that eigenvalues are computed by QZ as ratios of numbers, the "right way" to study the eigenvalue condition numbers of pencils is via the homogeneous formulation (Stewart-Sun, 1990), i.e. $A, B \in \mathbb{C}^{n\times n}$,

$$\lambda B - A \longleftrightarrow \alpha B - \beta A\,,$$

- where each eigenvalue $(\alpha, \beta) \neq (0,0)$ satisfying $\det(\alpha B - \beta A) = 0$ becomes a line $\langle \alpha, \beta \rangle$ through the origin in $\mathbb{C}^2$ ($\lambda = \alpha/\beta$ and $\infty \longleftrightarrow \langle 1, 0 \rangle$)
- and the diference between two eigenvalues is measured in terms of the chordal metric:

$$\chi(\langle \alpha, \beta \rangle, \langle \gamma, \delta \rangle) = \frac{|\alpha\delta - \beta\gamma|}{\sqrt{|\alpha|^2 + |\beta|^2}\sqrt{|\gamma|^2 + |\delta|^2}}$$

Also

$$\sin\theta(\langle \alpha, \beta \rangle, \langle \gamma, \delta \rangle) = \chi(\langle \alpha, \beta \rangle, \langle \gamma, \delta \rangle)$$

Why does LVD method improve often the accuracy of computed eigenvalues? (2)

- The eigenvalues $\langle\widehat{\alpha},\widehat{\beta}\rangle$ of $\alpha B - \beta A$ computed via a backward stable algorithm as QZ are the exact ones of $\alpha(B+F) - \beta(A+E)$ with $\|F\|_F = O(\varepsilon)\|B\|_F$ and $\|E\|_F = O(\varepsilon)\|A\|_F$.

- Thus, using Stewart-Sun homogeneous condition numbers, up to $O(\varepsilon^2)$,

$$\chi(\langle\widehat{\alpha},\widehat{\beta}\rangle,\langle\alpha,\beta\rangle) \leq O(\varepsilon)\,\frac{\|y\|_2\,\|x\|_2}{\sqrt{|y^*Ax|^2+|y^*Bx|^2}}\,\sqrt{\|A\|_F^2+\|B\|_F^2},$$

where $\langle\alpha,\beta\rangle$ is an exact simple eigenvalue of $\alpha B - \beta A$, $(\alpha B - \beta A)x = 0$ and $y^*(\alpha B - \beta A) = 0$.

- Observe invariance under multiplying the pencil by a number.

- The eigenvalues $\langle\widehat{\alpha},\widehat{\beta}\rangle$ of $\alpha B - \beta A$ computed via a backward stable algorithm as QZ are the exact ones of $\alpha(B+F) - \beta(A+E)$ with $\|F\|_F = O(\varepsilon)\|B\|_F$ and $\|E\|_F = O(\varepsilon)\|A\|_F$.

- Thus, using Stewart-Sun homogeneous condition numbers, up to $O(\varepsilon^2)$,

$$\chi(\langle\widehat{\alpha},\widehat{\beta}\rangle,\langle\alpha,\beta\rangle) \leq O(\varepsilon)\,\frac{\|y\|_2\,\|x\|_2}{\sqrt{|y^*Ax|^2+|y^*Bx|^2}}\sqrt{\|A\|_F^2+\|B\|_F^2},$$

where $\langle\alpha,\beta\rangle$ is an exact simple eigenvalue of $\alpha B - \beta A$, $(\alpha B - \beta A)x = 0$ and $y^*(\alpha B - \beta A) = 0$.

- Observe invariance under multiplying the pencil by a number.

- The eigenvalues $\langle\widehat{\alpha},\widehat{\beta}\rangle$ of $\alpha B-\beta A$ computed via a backward stable algorithm as QZ are the exact ones of $\alpha(B+F)-\beta(A+E)$ with $\|F\|_F=O(\varepsilon)\|B\|_F$ and $\|E\|_F=O(\varepsilon)\|A\|_F$.

- Thus, using Stewart-Sun homogeneous condition numbers, up to $O(\varepsilon^2)$,

$$\chi(\langle\widehat{\alpha},\widehat{\beta}\rangle,\langle\alpha,\beta\rangle)\leq O(\varepsilon)\,\frac{\|y\|_2\,\|x\|_2}{\sqrt{|y^*Ax|^2+|y^*Bx|^2}}\,\sqrt{\|A\|_F^2+\|B\|_F^2},$$

where $\langle\alpha,\beta\rangle$ is an exact simple eigenvalue of $\alpha B-\beta A$, $(\alpha B-\beta A)x=0$ and $y^*(\alpha B-\beta A)=0$.

- Observe invariance under multiplying the pencil by a number.

- Lemonnier-Van Dooren (2006) proved that in the diagonalizable case the pencils

$$\alpha \widetilde{B} - \beta \widetilde{A} := \alpha\, T_\ell B T_r - \beta\, T_\ell A T_r$$

that solve the minimization problem

$$\min_{\det T_\ell \,.\, \det T_r = 1} \|T_\ell A T_r\|_F^2 + \|T_\ell B T_r\|_F^2,$$

where T_ℓ and T_r are arbitrary nonsingular matrices, satisfy

$$\frac{\|\widetilde{y}\|_2\, \|\widetilde{x}\|_2}{\sqrt{|\widetilde{y}^* \widetilde{A} \widetilde{x}|^2 + |\widetilde{y}^* \widetilde{B} \widetilde{x}|^2}} \sqrt{\|\widetilde{A}\|_F^2 + \|\widetilde{B}\|_F^2} \leq \sqrt{n}.$$

$\sqrt{n} \longrightarrow \sqrt{2}$ in the spectral norm.

- The minimizers are the so-called standardized normal pencils satisfying

$$\alpha \hat{B} - \beta \hat{A} = U_\ell (\alpha \Lambda_B - \beta \Lambda_A) U_r, \quad |\Lambda_B|^2 + |\Lambda_A|^2 = c^2 I_n,$$

where Λ_B and Λ_A are diagonal and U_ℓ, U_r are unitary.

Why does LVD method improve often the accuracy of computed eigenvalues? (3)

- Lemonnier-Van Dooren (2006) proved that in the diagonalizable case the pencils

$$\alpha \widetilde{B} - \beta \widetilde{A} := \alpha\, T_\ell B T_r - \beta\, T_\ell A T_r$$

that solve the minimization problem

$$\min_{\det T_\ell \,.\, \det T_r = 1} \|T_\ell A T_r\|_F^2 + \|T_\ell B T_r\|_F^2,$$

where T_ℓ and T_r are arbitrary nonsingular matrices, satisfy

$$\frac{\|\widetilde{y}\|_2 \, \|\widetilde{x}\|_2}{\sqrt{|\widetilde{y}^* \widetilde{A} \widetilde{x}|^2 + |\widetilde{y}^* \widetilde{B} \widetilde{x}|^2}} \sqrt{\|\widetilde{A}\|_F^2 + \|\widetilde{B}\|_F^2} \leq \sqrt{n}.$$

$\sqrt{n} \longrightarrow \sqrt{2}$ in the spectral norm.

- The minimizers are the so-called standardized normal pencils satisfying

$$\alpha \hat{B} - \beta \hat{A} = U_\ell (\alpha \Lambda_B - \beta \Lambda_A) U_r, \quad |\Lambda_B|^2 + |\Lambda_A|^2 = c^2 I_n,$$

where Λ_B and Λ_A are diagonal and U_ℓ, U_r are unitary.

Why does LVD method improve often the accuracy of computed eigenvalues? (4)

- Moreover, Lemonnier-Van Dooren proved that the same minimization over **positive diagonal matrices** D_ℓ, D_r

$$\min_{\det D_\ell \,.\, \det D_r = 1} \|D_\ell A D_r\|_F^2 + \|D_\ell B D_r\|_F^2,$$

has as solution a pencil $\widetilde{A} = D_\ell A D_r$ and $\widetilde{B} = D_\ell B D_r$ such that

$$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 = \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 = \gamma^2,$$

for $i, j = 1, \dots, n$, and some constant γ.

- Thus, LVD scaling method will likely improve the error of computed eigenvalues measured in the chordal metric, since it solves the same minimization problem that makes the errors $O(\varepsilon)$ but restricted to diagonal invertible matrices.
- These equalities become approximate if the entries of D_ℓ and D_r are restricted to be integer power of 2.

- Moreover, Lemonnier-Van Dooren proved that the same minimization over **positive diagonal matrices** D_ℓ, D_r

$$\min_{\det D_\ell \,.\, \det D_r = 1} \|D_\ell A D_r\|_F^2 + \|D_\ell B D_r\|_F^2,$$

has as solution a pencil $\widetilde{A} = D_\ell A D_r$ and $\widetilde{B} = D_\ell B D_r$ such that

$$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 = \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 = \gamma^2,$$

for $i, j = 1, \ldots, n$, and some constant γ.

- Thus, LVD scaling method will likely improve the error of computed eigenvalues measured in the chordal metric, since it solves the same minimization problem that makes the errors $O(\varepsilon)$ but restricted to diagonal invertible matrices.
- These equalities become approximate if the entries of D_ℓ and D_r are restricted to be integer power of 2.

- Moreover, Lemonnier-Van Dooren proved that the same minimization over **positive diagonal matrices** D_ℓ, D_r

$$\min_{\det D_\ell \,.\, \det D_r = 1} \|D_\ell A D_r\|_F^2 + \|D_\ell B D_r\|_F^2,$$

has as solution a pencil $\widetilde{A} = D_\ell A D_r$ and $\widetilde{B} = D_\ell B D_r$ such that

$$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 = \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 = \gamma^2,$$

for $i, j = 1, \dots, n$, and some constant γ.

- Thus, LVD scaling method will likely improve the error of computed eigenvalues measured in the chordal metric, since it solves the same minimization problem that makes the errors $O(\varepsilon)$ but restricted to diagonal invertible matrices.
- These equalities become approximate if the entries of D_ℓ and D_r are restricted to be integer power of 2.

Lemonnier-Van Dooren algorithm for scaling pencils

- Since the precise value of the constant γ is not relevant in the error bounds, we take it equal to 1 and the LVD algorithm for computing the diagonal scalings starts with $D_\ell = D_r = I_n$ and consist in alternatively updating the diagonal matrices

$$D_r \longleftarrow D_r \, D_{r,up} \quad \text{and} \quad D_\ell \longleftarrow D_{\ell,up} \, D_\ell$$

such that

$$\begin{bmatrix} A \\ B \end{bmatrix} \longleftarrow \begin{bmatrix} A \\ B \end{bmatrix} D_{r,up} \quad \text{and} \quad \begin{bmatrix} A & B \end{bmatrix} \longleftarrow D_{\ell,up} \begin{bmatrix} A & B \end{bmatrix}$$

have column 2-norms all equal to 1 and row 2-norms all equal to 1, respectively.

- If the entries of the diagonal matrices are restricted to be integer powers of 2, the process converges in general quickly and costs $O(n^2)$ flops, which is negligible with respect to $O(n^3)$ of QZ.
- The theoretical conditions of convergence were not analyzed by LVD.

Lemonnier-Van Dooren algorithm for scaling pencils

- Since the precise value of the constant γ is not relevant in the error bounds, we take it equal to 1 and the LVD algorithm for computing the diagonal scalings starts with $D_\ell = D_r = I_n$ and consist in alternatively updating the diagonal matrices

$$D_r \longleftarrow D_r\, D_{r,up} \quad \text{and} \quad D_\ell \longleftarrow D_{\ell,up}\, D_\ell$$

  such that

$$\begin{bmatrix} A \\ B \end{bmatrix} \longleftarrow \begin{bmatrix} A \\ B \end{bmatrix} D_{r,up} \quad \text{and} \quad \begin{bmatrix} A & B \end{bmatrix} \longleftarrow D_{\ell,up} \begin{bmatrix} A & B \end{bmatrix}$$

  have column 2-norms all equal to 1 and row 2-norms all equal to 1, respectively.
- If the entries of the diagonal matrices are restricted to be integer powers of 2, the process converges in general quickly and costs $O(n^2)$ flops, which is negligible with respect to $O(n^3)$ of QZ.
- The theoretical conditions of convergence were not analyzed by LVD.

- Since the precise value of the constant γ is not relevant in the error bounds, we take it equal to 1 and the LVD algorithm for computing the diagonal scalings starts with $D_\ell = D_r = I_n$ and consist in alternatively updating the diagonal matrices

$$D_r \longleftarrow D_r\, D_{r,up} \quad \text{and} \quad D_\ell \longleftarrow D_{\ell,up}\, D_\ell$$

  such that
$$\begin{bmatrix} A \\ B \end{bmatrix} \longleftarrow \begin{bmatrix} A \\ B \end{bmatrix} D_{r,up} \quad \text{and} \quad \begin{bmatrix} A & B \end{bmatrix} \longleftarrow D_{\ell,up} \begin{bmatrix} A & B \end{bmatrix}$$
  have column 2-norms all equal to 1 and row 2-norms all equal to 1, respectively.

- If the entries of the diagonal matrices are restricted to be integer powers of 2, the process converges in general quickly and costs $O(n^2)$ flops, which is negligible with respect to $O(n^3)$ of QZ.

- The theoretical conditions of convergence were not analyzed by LVD.

Some numerical tests

- Next, we illustrate the LVD algorithm with two examples where the "exact" eigevalues are known, either because we construct the test pencils starting from the eigenvalues or because we compute them via MATLAB `vpa` with 64 decimal digits.

- We measure the errors with

$$c = \left\| \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix} \right\|_2, \quad \text{where} \quad c_i = \chi(\langle \widehat{\alpha}_i, \widehat{\beta}_i \rangle, \langle \alpha_i, \beta_i \rangle)$$

- Next, we illustrate the LVD algorithm with two examples where the "exact" eigevalues are known, either because we construct the test pencils starting from the eigenvalues or because we compute them via MATLAB `vpa` with 64 decimal digits.

- We measure the errors with

$$c = \| \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix} \|_2, \quad \text{where} \quad c_i = \chi(\langle \widehat{\alpha}_i, \widehat{\beta}_i \rangle, \langle \alpha_i, \beta_i \rangle)$$

Numerical test 1 illustrates that LVD can be much better than Ward

- 500×500 pencils $\lambda T - TD$, with $T = \texttt{randn}$ and with $T(1, 2:500) = 10^{-k} T(1, 2:500)$ and $T(4:500, 3) = 10^{-k} T(4:500, 3)$, and $D = \text{diag}(\texttt{randi})$.
- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{ward} error QZ applied to Ward scaled pencil.

k	c_{orig}	c_{bal}	c_{ward}
1	2.61e-13	3.40e-15	8.87e-15
3	1.48e-13	7.59e-15	1.91e-14
5	4.13e-13	8.72e-15	4.56e-09
7	7.16e-14	2.27e-15	3.47e-02
9	3.90e-13	3.01e-15	1.05e+00
11	1.34e-13	7.99e-15	1.08e+00

- Ward's method works very badly in this example, but often works well.
- We have not found examples so far where LVD deteriorates the error with respect to original pencil.

Numerical test 1 illustrates that LVD can be much better than Ward

- 500×500 pencils $\lambda T - TD$, with T = `randn` and with $T(1, 2:500) = 10^{-k} T(1, 2:500)$ and $T(4:500, 3) = 10^{-k} T(4:500, 3)$, and $D = \text{diag}($`randi`$)$.
- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{ward} error QZ applied to Ward scaled pencil.

k	c_{orig}	c_{bal}	c_{ward}
1	2.61e-13	3.40e-15	8.87e-15
3	1.48e-13	7.59e-15	1.91e-14
5	4.13e-13	8.72e-15	4.56e-09
7	7.16e-14	2.27e-15	3.47e-02
9	3.90e-13	3.01e-15	1.05e+00
11	1.34e-13	7.99e-15	1.08e+00

- Ward's method works very badly in this example, but often works well.
- We have not found examples so far where LVD deteriorates the error with respect to original pencil.

Numerical test 1 illustrates that LVD can be much better than Ward

- 500×500 pencils $\lambda T - TD$, with T = `randn` and with $T(1, 2:500) = 10^{-k} T(1, 2:500)$ and $T(4:500, 3) = 10^{-k} T(4:500, 3)$, and D = diag(`randi`).
- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{ward} error QZ applied to Ward scaled pencil.

k	c_{orig}	c_{bal}	c_{ward}
1	2.61e-13	3.40e-15	8.87e-15
3	1.48e-13	7.59e-15	1.91e-14
5	4.13e-13	8.72e-15	4.56e-09
7	7.16e-14	2.27e-15	3.47e-02
9	3.90e-13	3.01e-15	1.05e+00
11	1.34e-13	7.99e-15	1.08e+00

- Ward's method works very badly in this example, but often works well.
- We have not found examples so far where LVD deteriorates the error with respect to original pencil.

Numerical test 1 illustrates that LVD can be much better than Ward

- 500×500 pencils $\lambda T - TD$, with $T =$ `randn` and with $T(1,2:500) = 10^{-k}T(1,2:500)$ and $T(4:500,3) = 10^{-k}T(4:500,3)$, and $D =$ diag(`randi`).
- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{ward} error QZ applied to Ward scaled pencil.

k	c_{orig}	c_{bal}	c_{ward}
1	2.61e-13	3.40e-15	8.87e-15
3	1.48e-13	7.59e-15	1.91e-14
5	4.13e-13	8.72e-15	4.56e-09
7	7.16e-14	2.27e-15	3.47e-02
9	3.90e-13	3.01e-15	1.05e+00
11	1.34e-13	7.99e-15	1.08e+00

- Ward's method works very badly in this example, but often works well.
- We have not found examples so far where LVD deteriorates the error with respect to original pencil.

Numerical test 1 illustrates that LVD can be much better than Ward

- 500×500 pencils $\lambda T - TD$, with $T = \texttt{randn}$ and with $T(1, 2:500) = 10^{-k} T(1, 2:500)$ and $T(4:500, 3) = 10^{-k} T(4:500, 3)$, and $D = \text{diag}(\texttt{randi})$.
- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{ward} error QZ applied to Ward scaled pencil.

k	c_{orig}	c_{bal}	c_{ward}
1	2.61e-13	3.40e-15	8.87e-15
3	1.48e-13	7.59e-15	1.91e-14
5	4.13e-13	8.72e-15	4.56e-09
7	7.16e-14	2.27e-15	3.47e-02
9	3.90e-13	3.01e-15	1.05e+00
11	1.34e-13	7.99e-15	1.08e+00

- Ward's method works very badly in this example, but often works well.
- We have not found examples so far where LVD deteriorates the error with respect to original pencil.

Numerical test 1 illustrates that LVD can be much better than Ward

- 500×500 pencils $\lambda T - TD$, with $T = \texttt{randn}$ and with $T(1, 2:500) = 10^{-k} T(1, 2:500)$ and $T(4:500, 3) = 10^{-k} T(4:500, 3)$, and $D = \text{diag}(\texttt{randi})$.
- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{ward} error QZ applied to Ward scaled pencil.

k	c_{orig}	c_{bal}	c_{ward}
1	2.61e-13	3.40e-15	8.87e-15
3	1.48e-13	7.59e-15	1.91e-14
5	4.13e-13	8.72e-15	4.56e-09
7	7.16e-14	2.27e-15	3.47e-02
9	3.90e-13	3.01e-15	1.05e+00
11	1.34e-13	7.99e-15	1.08e+00

- Ward's method works very badly in this example, but often works well.
- We have not found examples so far where LVD deteriorates the error with respect to original pencil.

Numerical test 1 illustrates that LVD can be much better than Ward

- 500×500 pencils $\lambda T - TD$, with $T = \texttt{randn}$ and with $T(1, 2:500) = 10^{-k} T(1, 2:500)$ and $T(4:500, 3) = 10^{-k} T(4:500, 3)$, and $D = \mathrm{diag}(\texttt{randi})$.
- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{ward} error QZ applied to Ward scaled pencil.

k	c_{orig}	c_{bal}	c_{ward}
1	2.61e-13	3.40e-15	8.87e-15
3	1.48e-13	7.59e-15	1.91e-14
5	4.13e-13	8.72e-15	4.56e-09
7	7.16e-14	2.27e-15	3.47e-02
9	3.90e-13	3.01e-15	1.05e+00
11	1.34e-13	7.99e-15	1.08e+00

- Ward's method works very badly in this example, but often works well.
- We have not found examples so far where LVD deteriorates the error with respect to original pencil.

Numerical test 2 is related to the dynamic behavior of a nuclear power plant

- 8×8 quadratic eigenvalue problem $Q(\lambda) = \lambda^2 M + \lambda D + K$ describing a simple model for the dynamic behavior of a nuclear plant (Betcke, Higham, Mehrmann, Schröder, Tisseur, NLEVP collection, 2013).
- We solve the problem via QZ applied to first Frobenius companion pencil

$$\lambda B - A = \begin{bmatrix} \lambda M + D & K \\ -I & \lambda I \end{bmatrix}$$

- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{MAT} error command `polyeig` of MATLAB.
- The problem depends on a parameter μ and we have solved it for $\mu = 0.2, 0.5, 0.8, 1.1$ (0.2 is the default value in the toolbox).

c_{orig}	c_{bal}	c_{MAT}
1.5e-05	1.8e-16	1.6e-06
1.6e-05	1.1e-16	4.1e-05
2.0e-04	2.0e-16	1.0e-05
6.3e-05	1.3e-16	1.9e-06

Numerical test 2 is related to the dynamic behavior of a nuclear power plant

- 8×8 quadratic eigenvalue problem $Q(\lambda) = \lambda^2 M + \lambda D + K$ describing a simple model for the dynamic behavior of a nuclear plant (Betcke, Higham, Mehrmann, Schröder, Tisseur, NLEVP collection, 2013).
- We solve the problem via QZ applied to first Frobenius companion pencil

$$\lambda B - A = \begin{bmatrix} \lambda M + D & K \\ -I & \lambda I \end{bmatrix}$$

- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{MAT} error command `polyeig` of MATLAB.
- The problem depends on a parameter μ and we have solved it for $\mu = 0.2, 0.5, 0.8, 1.1$ (0.2 is the default value in the toolbox).

c_{orig}	c_{bal}	c_{MAT}
1.5e-05	1.8e-16	1.6e-06
1.6e-05	1.1e-16	4.1e-05
2.0e-04	2.0e-16	1.0e-05
6.3e-05	1.3e-16	1.9e-06

Numerical test 2 is related to the dynamic behavior of a nuclear power plant

- 8×8 quadratic eigenvalue problem $Q(\lambda) = \lambda^2 M + \lambda D + K$ describing a simple model for the dynamic behavior of a nuclear plant (Betcke, Higham, Mehrmann, Schröder, Tisseur, NLEVP collection, 2013).
- We solve the problem via QZ applied to first Frobenius companion pencil

$$\lambda B - A = \begin{bmatrix} \lambda M + D & K \\ -I & \lambda I \end{bmatrix}$$

- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{MAT} error command `polyeig` of MATLAB.
- The problem depends on a parameter μ and we have solved it for $\mu = 0.2, 0.5, 0.8, 1.1$ (0.2 is the default value in the toolbox).

c_{orig}	c_{bal}	c_{MAT}
1.5e-05	1.8e-16	1.6e-06
1.6e-05	1.1e-16	4.1e-05
2.0e-04	2.0e-16	1.0e-05
6.3e-05	1.3e-16	1.9e-06

Numerical test 2 is related to the dynamic behavior of a nuclear power plant

- 8×8 quadratic eigenvalue problem $Q(\lambda) = \lambda^2 M + \lambda D + K$ describing a simple model for the dynamic behavior of a nuclear plant (Betcke, Higham, Mehrmann, Schröder, Tisseur, NLEVP collection, 2013).
- We solve the problem via QZ applied to first Frobenius companion pencil

$$\lambda B - A = \begin{bmatrix} \lambda M + D & K \\ -I & \lambda I \end{bmatrix}$$

- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{MAT} error command `polyeig` of MATLAB.
- The problem depends on a parameter μ and we have solved it for $\mu = 0.2, 0.5, 0.8, 1.1$ (0.2 is the default value in the toolbox).

c_{orig}	c_{bal}	c_{MAT}
1.5e-05	1.8e-16	1.6e-06
1.6e-05	1.1e-16	4.1e-05
2.0e-04	2.0e-16	1.0e-05
6.3e-05	1.3e-16	1.9e-06

Numerical test 2 is related to the dynamic behavior of a nuclear power plant

- 8×8 quadratic eigenvalue problem $Q(\lambda) = \lambda^2 M + \lambda D + K$ describing a simple model for the dynamic behavior of a nuclear plant (Betcke, Higham, Mehrmann, Schröder, Tisseur, NLEVP collection, 2013).
- We solve the problem via QZ applied to first Frobenius companion pencil

$$\lambda B - A = \begin{bmatrix} \lambda M + D & K \\ -I & \lambda I \end{bmatrix}$$

- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{MAT} error command `polyeig` of MATLAB.
- The problem depends on a parameter μ and we have solved it for $\mu = 0.2, 0.5, 0.8, 1.1$ (0.2 is the default value in the toolbox).

c_{orig}	c_{bal}	c_{MAT}
1.5e-05	1.8e-16	1.6e-06
1.6e-05	1.1e-16	4.1e-05
2.0e-04	2.0e-16	1.0e-05
6.3e-05	1.3e-16	1.9e-06

Numerical test 2 is related to the dynamic behavior of a nuclear power plant

- 8×8 quadratic eigenvalue problem $Q(\lambda) = \lambda^2 M + \lambda D + K$ describing a simple model for the dynamic behavior of a nuclear plant (Betcke, Higham, Mehrmann, Schröder, Tisseur, NLEVP collection, 2013).
- We solve the problem via QZ applied to first Frobenius companion pencil

$$\lambda B - A = \begin{bmatrix} \lambda M + D & K \\ -I & \lambda I \end{bmatrix}$$

- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{MAT} error command `polyeig` of MATLAB.
- The problem depends on a parameter μ and we have solved it for $\mu = 0.2, 0.5, 0.8, 1.1$ (0.2 is the default value in the toolbox).

c_{orig}	c_{bal}	c_{MAT}
1.5e-05	1.8e-16	1.6e-06
1.6e-05	1.1e-16	4.1e-05
2.0e-04	2.0e-16	1.0e-05
6.3e-05	1.3e-16	1.9e-06

Numerical test 2 is related to the dynamic behavior of a nuclear power plant

- 8×8 quadratic eigenvalue problem $Q(\lambda) = \lambda^2 M + \lambda D + K$ describing a simple model for the dynamic behavior of a nuclear plant (Betcke, Higham, Mehrmann, Schröder, Tisseur, NLEVP collection, 2013).
- We solve the problem via QZ applied to first Frobenius companion pencil

$$\lambda B - A = \begin{bmatrix} \lambda M + D & K \\ -I & \lambda I \end{bmatrix}$$

- c_{orig} error QZ applied to original pencil.
- c_{bal} error QZ applied to LVD scaled pencil.
- c_{MAT} error command `polyeig` of MATLAB.
- The problem depends on a parameter μ and we have solved it for $\mu = 0.2, 0.5, 0.8, 1.1$ (0.2 is the default value in the toolbox).

c_{orig}	c_{bal}	c_{MAT}
1.5e-05	1.8e-16	1.6e-06
1.6e-05	1.1e-16	4.1e-05
2.0e-04	2.0e-16	1.0e-05
6.3e-05	1.3e-16	1.9e-06

Goal for the rest of the talk

- To develop a diagonal scaling strategy for arbitrary pencils (regular, singular, square, rectangular) that converges always and quickly.
- As far as we know, this problem has not been considered before for square singular pencils nor for rectangular pencils.
- In the process, we will obtain a much deeper understanding of the Lemonnier-Van Dooren strategy for regular pencils.

- To develop a diagonal scaling strategy for arbitrary pencils (regular, singular, square, rectangular) that converges always and quickly.
- As far as we know, this problem has not been considered before for square singular pencils nor for rectangular pencils.
- In the process, we will obtain a much deeper understanding of the Lemonnier-Van Dooren strategy for regular pencils.

Goal for the rest of the talk

- To develop a diagonal scaling strategy for arbitrary pencils (regular, singular, square, rectangular) that converges always and quickly.
- As far as we know, this problem has not been considered before for square singular pencils nor for rectangular pencils.
- In the process, we will obtain a much deeper understanding of the Lemonnier-Van Dooren strategy for regular pencils.

Outline

A new view of Lemmonier-Van Dooren algorithm for regular pencils

$$A, B \in \mathbb{C}^{n\times n}. \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

such

$$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 = \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 = 1,$$

for $i, j = 1, \ldots, n$.

- Define the nonnegative matrices

$$M := |A|^{\circ 2} + |B|^{\circ 2}, \quad \text{and} \quad \widetilde{M} := |\widetilde{A}|^{\circ 2} + |\widetilde{B}|^{\circ 2} = D_\ell^2 \, M \, D_r^2$$

  where $|X|$ indicates the entry-wise absolute value and $X^{\circ 2}$ indicates the entry-wise square.

- Thus, **the LVD diagonal scaling is equivalent to find positive diagonal matrices that transform the nonnegative matrix M into a doubly stochastic matrix $\widetilde{M}$.**

- There are many results in the literature for diagonally scaling a nonnegative matrix to a matrix with prescribed row and column sums: Kruithof (1937), Sinkhorn-Knopp (1967), Brualdi (1968), Krupp (1979), Rothblum-Schneider (1989), ...

$$A, B \in \mathbb{C}^{n \times n}. \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

such

$$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 = \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 = 1,$$

for $i, j = 1, \ldots, n$.

- Define the nonnegative matrices

$$M := |A|^{\circ 2} + |B|^{\circ 2}, \quad \text{and} \quad \widetilde{M} := |\widetilde{A}|^{\circ 2} + |\widetilde{B}|^{\circ 2} = D_\ell^2 \, M \, D_r^2$$

  where $|X|$ indicates the entry-wise absolute value and $X^{\circ 2}$ indicates the entry-wise square.

- Thus, **the LVD diagonal scaling is equivalent to find positive diagonal matrices that transform the nonnegative matrix M into a doubly stochastic matrix $\widetilde{M}$.**
- There are many results in the literature for diagonally scaling a nonnegative matrix to a matrix with prescribed row and column sums: Kruithof (1937), Sinkhorn-Knopp (1967), Brualdi (1968), Krupp (1979), Rothblum-Schneider (1989), ...

A new view of Lemmonier-Van Dooren algorithm for regular pencils

$$A, B \in \mathbb{C}^{n\times n}. \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

such

$$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 = \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 = 1,$$

for $i, j = 1, \ldots, n$.

- Define the nonnegative matrices

$$M := |A|^{\circ 2} + |B|^{\circ 2}, \quad \text{and} \quad \widetilde{M} := |\widetilde{A}|^{\circ 2} + |\widetilde{B}|^{\circ 2} = D_\ell^2 \, M \, D_r^2$$

  where $|X|$ indicates the entry-wise absolute value and $X^{\circ 2}$ indicates the entry-wise square.

- Thus, **the LVD diagonal scaling is equivalent to find positive diagonal matrices that transform the nonnegative matrix M into a doubly stochastic matrix $\widetilde{M}$.**

- There are many results in the literature for diagonally scaling a nonnegative matrix to a matrix with prescribed row and column sums: Kruithof (1937), Sinkhorn-Knopp (1967), Brualdi (1968), Krupp (1979), Rothblum-Schneider (1989), ...

A new view of Lemmonier-Van Dooren algorithm for regular pencils

$$A, B \in \mathbb{C}^{n\times n}. \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

such

$$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 = \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 = 1,$$

for $i, j = 1, \ldots, n$.

- Define the nonnegative matrices

$$M := |A|^{\circ 2} + |B|^{\circ 2}, \quad \text{and} \quad \widetilde{M} := |\widetilde{A}|^{\circ 2} + |\widetilde{B}|^{\circ 2} = D_\ell^2 \, M \, D_r^2$$

  where $|X|$ indicates the entry-wise absolute value and $X^{\circ 2}$ indicates the entry-wise square.

- Thus, **the LVD diagonal scaling is equivalent to find positive diagonal matrices that transform the nonnegative matrix M into a doubly stochastic matrix $\widetilde{M}$.**

- There are many results in the literature for diagonally scaling a nonnegative matrix to a matrix with prescribed row and column sums: Kruithof (1937), Sinkhorn-Knopp (1967), Brualdi (1968), Krupp (1979), Rothblum-Schneider (1989), ...

Scaling nonnegative matrices

- The problem of scaling an entrywise nonnegative $m \times n$ matrix M with diagonal transformations
- and prescribed positive vectors r and c for the row and column sums
- consists of finding a matrix of the form

$$S = D_{M,\ell}\, M\, D_{M,r},$$

where $D_{M,\ell} \in \mathbb{R}^{m \times m}$ and $D_{M,r} \in \mathbb{R}^{n \times n}$ are positive diagonal matrices

- such that

$$S\,\mathbf{1}_n = r \quad \text{and} \quad \mathbf{1}_m^T S = c^T,$$

where $\mathbf{1}_\ell := [1, \ldots, 1]^T \in \mathbb{R}^\ell$ for $\ell = n, m$,

- that is, the sum of the entries of the ith row of S is equal to r_i and the sum of the entries of the jth column of S is equal to c_j, for all i, j.
- The doubly stochastic scaling problem corresponds to $m = n$ and $r = c = \mathbf{1}_n$.

Scaling nonnegative matrices

- The problem of scaling an entrywise nonnegative $m \times n$ matrix M with diagonal transformations
- and prescribed positive vectors r and c for the row and column sums
- consists of finding a matrix of the form

$$S = D_{M,\ell}\, M\, D_{M,r},$$

where $D_{M,\ell} \in \mathbb{R}^{m\times m}$ and $D_{M,r} \in \mathbb{R}^{n\times n}$ are positive diagonal matrices

- such that

$$S\,\mathbf{1}_n = r \quad \text{and} \quad \mathbf{1}_m^T S = c^T,$$

where $\mathbf{1}_\ell := [1, \ldots, 1]^T \in \mathbb{R}^\ell$ for $\ell = n, m$,

- that is, the sum of the entries of the ith row of S is equal to r_i and the sum of the entries of the jth column of S is equal to c_j, for all i, j.
- The doubly stochastic scaling problem corresponds to $m = n$ and $r = c = \mathbf{1}_n$.

- The problem of scaling an entrywise nonnegative $m \times n$ matrix M with diagonal transformations
- and prescribed positive vectors r and c for the row and column sums
- consists of finding a matrix of the form

$$S = D_{M,\ell}\, M\, D_{M,r},$$

where $D_{M,\ell} \in \mathbb{R}^{m\times m}$ and $D_{M,r} \in \mathbb{R}^{n\times n}$ are positive diagonal matrices

- such that

$$S\,\mathbf{1}_n = r \quad \text{and} \quad \mathbf{1}_m^T\, S = c^T,$$

where $\mathbf{1}_\ell := [1,\ldots,1]^T \in \mathbb{R}^\ell$ for $\ell = n, m$,

- that is, the sum of the entries of the ith row of S is equal to r_i and the sum of the entries of the jth column of S is equal to c_j, for all i, j.
- The doubly stochastic scaling problem corresponds to $m = n$ and $r = c = \mathbf{1}_n$.

Scaling nonnegative matrices

- The problem of scaling an entrywise nonnegative $m \times n$ matrix M with diagonal transformations
- and prescribed positive vectors r and c for the row and column sums
- consists of finding a matrix of the form

$$S = D_{M,\ell}\, M\, D_{M,r},$$

  where $D_{M,\ell} \in \mathbb{R}^{m\times m}$ and $D_{M,r} \in \mathbb{R}^{n\times n}$ are positive diagonal matrices
- such that

$$S\,\mathbf{1}_n = r \quad \text{and} \quad \mathbf{1}_m^T\, S = c^T,$$

  where $\mathbf{1}_\ell := [1, \ldots, 1]^T \in \mathbb{R}^\ell$ for $\ell = n, m$,
- that is, the sum of the entries of the ith row of S is equal to r_i and the sum of the entries of the jth column of S is equal to c_j, for all i, j.
- The doubly stochastic scaling problem corresponds to $m = n$ and $r = c = \mathbf{1}_n$.

Scaling nonnegative matrices

- The problem of scaling an entrywise nonnegative $m \times n$ matrix M with diagonal transformations
- and prescribed positive vectors r and c for the row and column sums
- consists of finding a matrix of the form

$$S = D_{M,\ell}\, M\, D_{M,r},$$

where $D_{M,\ell} \in \mathbb{R}^{m\times m}$ and $D_{M,r} \in \mathbb{R}^{n\times n}$ are positive diagonal matrices

- such that

$$S\,\mathbf{1}_n = r \quad \text{and} \quad \mathbf{1}_m^T\, S = c^T,$$

where $\mathbf{1}_\ell := [1, \ldots, 1]^T \in \mathbb{R}^\ell$ for $\ell = n, m$,

- that is, the sum of the entries of the ith row of S is equal to r_i and the sum of the entries of the jth column of S is equal to c_j, for all i, j.
- The doubly stochastic scaling problem corresponds to $m = n$ and $r = c = \mathbf{1}_n$.

- The problem of scaling an entrywise nonnegative $m \times n$ matrix M with diagonal transformations
- and prescribed positive vectors r and c for the row and column sums
- consists of finding a matrix of the form

$$S = D_{M,\ell}\, M\, D_{M,r},$$

where $D_{M,\ell} \in \mathbb{R}^{m\times m}$ and $D_{M,r} \in \mathbb{R}^{n\times n}$ are positive diagonal matrices
- such that

$$S\,\mathbf{1}_n = r \quad \text{and} \quad \mathbf{1}_m^T\, S = c^T,$$

where $\mathbf{1}_\ell := [1, \ldots, 1]^T \in \mathbb{R}^\ell$ for $\ell = n, m$,
- that is, the sum of the entries of the ith row of S is equal to r_i and the sum of the entries of the jth column of S is equal to c_j, for all i, j.
- The doubly stochastic scaling problem corresponds to $m = n$ and $r = c = \mathbf{1}_n$.

The Sinkhorn-Knopp-like algorithm for scaling a nonnegative $M \in \mathbb{R}^{m \times n}$

- The problem in the previous slide has a solution if and only if the following algorithm converges

- *Initialize:* $D_{M,\ell} = I_m$ *and* $D_{M,r} = I_n$

  (1) $D_{M,\ell} \longleftarrow D_{\ell,up} D_{M,\ell}$ *and* $M \longleftarrow D_{\ell,up} M$, *such that the updated matrix* M *has row sums equal to* r.

  (2) $D_{M,r} \longleftarrow D_{M,r} D_{r,up}$ *and* $M \longleftarrow M D_{r,up}$, *such that the updated matrix* M *has column sums equal to* c.

  (3) *If the row sums of the matrix* M *obtained in step* (2) *are far from* r, *repeat steps* (1) *and* (2) *with such* M *until an adequate stopping criterion is satisfied.*

- If $m = n$ and $r = c = \mathbf{1}_n$, this reduces to the famous Sinkhorn-Knopp algorithm, which

- applied to the nonnegative matrix $M := |A|^{\circ 2} + |B|^{\circ 2}$ associated to the regular pencil $\lambda B - A$, $A, B \in \mathbb{C}^{n \times n}$, computes matrices $D_\ell = \sqrt{D_{M,\ell}}$ and $D_r = \sqrt{D_{M,r}}$ that scale the pencil according to LVD strategy.

The Sinkhorn-Knopp-like algorithm for scaling a nonnegative $M \in \mathbb{R}^{m\times n}$

- The problem in the previous slide has a solution if and only if the following algorithm converges
- *Initialize:* $D_{M,\ell} = I_m$ *and* $D_{M,r} = I_n$
  (1) $D_{M,\ell} \longleftarrow D_{\ell,up}\, D_{M,\ell}$ *and* $M \longleftarrow D_{\ell,up}\, M$, *such that the updated matrix* M *has row sums equal to* r.
  (2) $D_{M,r} \longleftarrow D_{M,r}\, D_{r,up}$ *and* $M \longleftarrow M\, D_{r,up}$, *such that the updated matrix* M *has column sums equal to* c.
  (3) *If the row sums of the matrix* M *obtained in step* (2) *are far from* r, *repeat steps* (1) *and* (2) *with such* M *until an adequate stopping criterion is satisfied.*
- If $m = n$ and $r = c = \mathbf{1}_n$, this reduces to the famous Sinkhorn-Knopp algorithm, which
- applied to the nonnegative matrix $M := |A|^{\circ 2} + |B|^{\circ 2}$ associated to the regular pencil $\lambda B - A$, $A, B \in \mathbb{C}^{n\times n}$, computes matrices $D_\ell = \sqrt{D_{M,\ell}}$ and $D_r = \sqrt{D_{M,r}}$ that scale the pencil according to LVD strategy.

- The problem in the previous slide has a solution if and only if the following algorithm converges
- *Initialize:* $D_{M,\ell} = I_m$ *and* $D_{M,r} = I_n$
  (1) $D_{M,\ell} \longleftarrow D_{\ell,up}\, D_{M,\ell}$ *and* $M \longleftarrow D_{\ell,up}\, M$, *such that the updated matrix* M *has row sums equal to* r.
  (2) $D_{M,r} \longleftarrow D_{M,r}\, D_{r,up}$ *and* $M \longleftarrow M\, D_{r,up}$, *such that the updated matrix* M *has column sums equal to* c.
  (3) *If the row sums of the matrix* M *obtained in step* (2) *are far from* r, *repeat steps* (1) *and* (2) *with such* M *until an adequate stopping criterion is satisfied.*
- If $m = n$ and $r = c = \mathbf{1}_n$, this reduces to the famous Sinkhorn-Knopp algorithm, which
- applied to the nonnegative matrix $M := |A|^{\circ 2} + |B|^{\circ 2}$ associated to the regular pencil $\lambda B - A$, $A, B \in \mathbb{C}^{n\times n}$, computes matrices $D_\ell = \sqrt{D_{M,\ell}}$ and $D_r = \sqrt{D_{M,r}}$ that scale the pencil according to LVD strategy.

The Sinkhorn-Knopp-like algorithm for scaling a nonnegative $M \in \mathbb{R}^{m \times n}$

- The problem in the previous slide has a solution if and only if the following algorithm converges
- *Initialize:* $D_{M,\ell} = I_m$ *and* $D_{M,r} = I_n$
  - (1) $D_{M,\ell} \longleftarrow D_{\ell,up} D_{M,\ell}$ *and* $M \longleftarrow D_{\ell,up} M$, *such that the updated matrix* M *has row sums equal to* r.
  - (2) $D_{M,r} \longleftarrow D_{M,r} D_{r,up}$ *and* $M \longleftarrow M D_{r,up}$, *such that the updated matrix* M *has column sums equal to* c.
  - (3) *If the row sums of the matrix* M *obtained in step* (2) *are far from* r, *repeat steps* (1) *and* (2) *with such* M *until an adequate stopping criterion is satisfied.*
- If $m = n$ and $r = c = \mathbf{1}_n$, this reduces to the famous Sinkhorn-Knopp algorithm, which
- applied to the nonnegative matrix $M := |A|^{\circ 2} + |B|^{\circ 2}$ associated to the regular pencil $\lambda B - A$, $A, B \in \mathbb{C}^{n \times n}$, computes matrices $D_\ell = \sqrt{D_{M,\ell}}$ and $D_r = \sqrt{D_{M,r}}$ that scale the pencil according to LVD strategy.

- There exist well-known results that guarantee the convergence of the Sinkhorn-Knopp algorithm based on the zero pattern of the matrix.

Theorem (Sinkhorn-Knopp)

If $M \in \mathbb{R}^{n\times n}$ is a nonnegative matrix, then:

- *There exists a doubly stochastic matrix S of the form $S = D_{M,\ell}\, M\, D_{M,r}$, where $D_{M,\ell}$ and $D_{M,r}$ are diagonal matrices with positive main diagonals, if and only if M has total support.*
- *If S exists, then it is unique.*
- *$D_{M,\ell}$ and $D_{M,r}$ are also unique up to a nonnegative scalar multiple if and only if M is fully indecomposable.*

- For other vectors r and c of prescribed row and column sums the available results are not so clear.

- There exist well-known results that guarantee the convergence of the Sinkhorn-Knopp algorithm based on the zero pattern of the matrix.

Theorem (Sinkhorn-Knopp)

If $M \in \mathbb{R}^{n\times n}$ is a nonnegative matrix, then:

- *There exists a doubly stochastic matrix S of the form $S = D_{M,\ell}\, M\, D_{M,r}$, where $D_{M,\ell}$ and $D_{M,r}$ are diagonal matrices with positive main diagonals, if and only if M has total support.*
- *If S exists, then it is unique.*
- *$D_{M,\ell}$ and $D_{M,r}$ are also unique up to a nonnegative scalar multiple if and only if M is fully indecomposable.*

- For other vectors r and c of prescribed row and column sums the available results are not so clear.

Some classical convergence results for Sinkhorn-Knopp algorithm

- There exist well-known results that guarantee the convergence of the Sinkhorn-Knopp algorithm based on the zero pattern of the matrix.

Theorem (Sinkhorn-Knopp)

If $M \in \mathbb{R}^{n \times n}$ is a nonnegative matrix, then:

- *There exists a doubly stochastic matrix S of the form $S = D_{M,\ell}\, M\, D_{M,r}$, where $D_{M,\ell}$ and $D_{M,r}$ are diagonal matrices with positive main diagonals, if and only if M has total support.*
- *If S exists, then it is unique.*
- *$D_{M,\ell}$ and $D_{M,r}$ are also unique up to a nonnegative scalar multiple if and only if M is fully indecomposable.*

- For other vectors r and c of prescribed row and column sums the available results are not so clear.

Reminder: total support and fully indecomposability

Definition

- The sequence $m_{1,\sigma(1)}, m_{2,\sigma(2)}, \ldots, m_{n,\sigma(n)}$, where σ is a permutation of $\{1, 2, \cdots, n\}$, is called a diagonal of the matrix $M \in \mathbb{R}^{n\times n}$.
- A nonnegative $M \in \mathbb{R}^{n\times n}$ is said to have total support if every positive element of M lies on a positive diagonal.

Definition

A nonnegative matrix $M \in \mathbb{R}^{n\times n}$ is said to be fully indecomposable if there do not exist permutation matrices P_ℓ and P_r such that $P_\ell M P_r$ can be partitioned as

$$P_\ell M P_r = \begin{bmatrix} M_{11} & M_{12} \\ 0 & M_{22} \end{bmatrix},$$

where M_{11} and M_{22} are square matrices.

Remark (Brualdi, 1980)

A fully indecomposable matrix has total support.

Definition

- The sequence $m_{1,\sigma(1)}, m_{2,\sigma(2)}, \ldots, m_{n,\sigma(n)}$, where σ is a permutation of $\{1, 2, \cdots, n\}$, is called a diagonal of the matrix $M \in \mathbb{R}^{n\times n}$.
- A nonnegative $M \in \mathbb{R}^{n\times n}$ is said to have total support if every positive element of M lies on a positive diagonal.

Definition

A nonnegative matrix $M \in \mathbb{R}^{n\times n}$ is said to be fully indecomposable if there do not exist permutation matrices P_ℓ and P_r such that $P_\ell M P_r$ can be partitioned as

$$P_\ell M P_r = \begin{bmatrix} M_{11} & M_{12} \\ 0 & M_{22} \end{bmatrix},$$

where M_{11} and M_{22} are square matrices.

Remark (Brualdi, 1980)

A fully indecomposable matrix has total support.

Reminder: total support and fully indecomposability

Definition

- The sequence $m_{1,\sigma(1)}, m_{2,\sigma(2)}, \ldots, m_{n,\sigma(n)}$, where σ is a permutation of $\{1, 2, \cdots, n\}$, is called a diagonal of the matrix $M \in \mathbb{R}^{n\times n}$.
- A nonnegative $M \in \mathbb{R}^{n\times n}$ is said to have total support if every positive element of M lies on a positive diagonal.

Definition

A nonnegative matrix $M \in \mathbb{R}^{n\times n}$ is said to be fully indecomposable if there do not exist permutation matrices P_ℓ and P_r such that $P_\ell M P_r$ can be partitioned as

$$P_\ell M P_r = \begin{bmatrix} M_{11} & M_{12} \\ 0 & M_{22} \end{bmatrix},$$

where M_{11} and M_{22} are square matrices.

Remark (Brualdi, 1980)

A fully indecomposable matrix has total support.

Example of a matrix without total support

The following matrix has NOT total support

$$M = \begin{bmatrix} 8 & 30 \\ 2 & 0 \end{bmatrix}$$

Thus, it cannot be diagonally scaled to a doubly stochastic matrix:

$$\begin{bmatrix} s_1 & \\ & s_2 \end{bmatrix} \begin{bmatrix} 8 & 30 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} t_1 & \\ & t_2 \end{bmatrix} \text{ is doubly stochastic } \iff \begin{cases} s_1 t_1 = 0, \\ s_1 t_2 = 1/30, \\ s_2 t_1 = 1/2 \end{cases}$$

and the Sinkhorn-Knopp algorithm does not converge in exact arithmetic, but, in numerical practice,

$$\begin{bmatrix} 2^{-6} & \\ & 2^{-1} \end{bmatrix} \begin{bmatrix} 8 & 30 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & \\ & 2 \end{bmatrix} = \begin{bmatrix} 0.1250 & 0.9375 \\ 1.0000 & 0 \end{bmatrix}$$

is a good result with diagonal scalings whose entries are integer powers of 2.

Example of a matrix without total support

The following matrix has NOT total support

$$M = \begin{bmatrix} 8 & 30 \\ 2 & 0 \end{bmatrix}$$

Thus, it cannot be diagonally scaled to a doubly stochastic matrix:

$$\begin{bmatrix} s_1 & \\ & s_2 \end{bmatrix} \begin{bmatrix} 8 & 30 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} t_1 & \\ & t_2 \end{bmatrix} \text{ is doubly stochastic} \iff \begin{cases} s_1 t_1 = 0, \\ s_1 t_2 = 1/30, \\ s_2 t_1 = 1/2 \end{cases}$$

and the Sinkhorn-Knopp algorithm does not converge in exact arithmetic,
but, in numerical practice,

$$\begin{bmatrix} 2^{-6} & \\ & 2^{-1} \end{bmatrix} \begin{bmatrix} 8 & 30 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & \\ & 2 \end{bmatrix} = \begin{bmatrix} 0.1250 & 0.9375 \\ 1.0000 & 0 \end{bmatrix}$$

is a good result with diagonal scalings whose entries are integer powers of 2.

Example of a matrix without total support

The following matrix has NOT total support

$$M = \begin{bmatrix} 8 & 30 \\ 2 & 0 \end{bmatrix}$$

Thus, it cannot be diagonally scaled to a doubly stochastic matrix:

$$\begin{bmatrix} s_1 & \\ & s_2 \end{bmatrix} \begin{bmatrix} 8 & 30 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} t_1 & \\ & t_2 \end{bmatrix} \text{ is doubly stochastic} \iff \begin{cases} s_1 t_1 = 0, \\ s_1 t_2 = 1/30, \\ s_2 t_1 = 1/2 \end{cases}$$

and the Sinkhorn-Knopp algorithm does not converge in exact arithmetic, but, in numerical practice,

$$\begin{bmatrix} 2^{-6} & \\ & 2^{-1} \end{bmatrix} \begin{bmatrix} 8 & 30 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & \\ & 2 \end{bmatrix} = \begin{bmatrix} 0.1250 & 0.9375 \\ 1.0000 & 0 \end{bmatrix}$$

is a good result with diagonal scalings whose entries are integer powers of 2.

Some practical comments

- When a pencil $\lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r$ is diagonally scaled with the goal of improving the accuracy of its computed eigenvalues,
- it is essential that the elements of the diagonal matrices are integer powers of 2, because in this way the scaling does not produce any rounding errors and the eigenvalues are preserved exactly. Otherwise, the rounding errors would spoil any potential improvement in the accuracy of the computed eigenvalues.
- This implies in practice that we do not need to stop the Sinkhorn-Knopp algorithm with a very stringent criterion, i.e., we do not need to converge to an (almost) stochastic matrix.
- In our experience, this means that for regular pencils the Sinkhorn-Knopp algorithm on $M := |A|^{\circ 2} + |B|^{\circ 2}$ converges "always in practice" with a relaxed stopping criterion (for instance, row sums and column sums equal up to a factor 2), even in situations for which the conditions in previous theorem are not satisfied.
- This is no longer true for SINGULAR square pencils, in particular, when they are sparse.

Some practical comments

- When a pencil $\lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r$ is diagonally scaled with the goal of improving the accuracy of its computed eigenvalues,
- it is essential that the elements of the diagonal matrices are integer powers of 2, because in this way the scaling does not produce any rounding errors and the eigenvalues are preserved exactly. Otherwise, the rounding errors would spoil any potential improvement in the accuracy of the computed eigenvalues.
- This implies in practice that we do not need to stop the Sinkhorn-Knopp algorithm with a very stringent criterion, i.e., we do not need to converge to an (almost) stochastic matrix.
- In our experience, this means that for regular pencils the Sinkhorn-Knopp algorithm on $M := |A|^{\circ 2} + |B|^{\circ 2}$ converges "always in practice" with a relaxed stopping criterion (for instance, row sums and column sums equal up to a factor 2), even in situations for which the conditions in previous theorem are not satisfied.
- This is no longer true for SINGULAR square pencils, in particular, when they are sparse.

Some practical comments

- When a pencil $\lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r$ is diagonally scaled with the goal of improving the accuracy of its computed eigenvalues,
- it is essential that the elements of the diagonal matrices are integer powers of 2, because in this way the scaling does not produce any rounding errors and the eigenvalues are preserved exactly. Otherwise, the rounding errors would spoil any potential improvement in the accuracy of the computed eigenvalues.
- This implies in practice that we do not need to stop the Sinkhorn-Knopp algorithm with a very stringent criterion, i.e., we do not need to converge to an (almost) stochastic matrix.
- In our experience, this means that for regular pencils the Sinkhorn-Knopp algorithm on $M := |A|^{\circ 2} + |B|^{\circ 2}$ converges "always in practice" with a relaxed stopping criterion (for instance, row sums and column sums equal up to a factor 2), even in situations for which the conditions in previous theorem are not satisfied.
- This is no longer true for SINGULAR square pencils, in particular, when they are sparse.

Some practical comments

- When a pencil $\lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r$ is diagonally scaled with the goal of improving the accuracy of its computed eigenvalues,
- it is essential that the elements of the diagonal matrices are integer powers of 2, because in this way the scaling does not produce any rounding errors and the eigenvalues are preserved exactly. Otherwise, the rounding errors would spoil any potential improvement in the accuracy of the computed eigenvalues.
- This implies in practice that we do not need to stop the Sinkhorn-Knopp algorithm with a very stringent criterion, i.e., we do not need to converge to an (almost) stochastic matrix.
- In our experience, this means that for regular pencils the Sinkhorn-Knopp algorithm on $M := |A|^{\circ 2} + |B|^{\circ 2}$ converges "always in practice" with a relaxed stopping criterion (for instance, row sums and column sums equal up to a factor 2), even in situations for which the conditions in previous theorem are not satisfied.
- This is no longer true for SINGULAR square pencils, in particular, when they are sparse.

Some practical comments

- When a pencil $\lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r$ is diagonally scaled with the goal of improving the accuracy of its computed eigenvalues,
- it is essential that the elements of the diagonal matrices are integer powers of 2, because in this way the scaling does not produce any rounding errors and the eigenvalues are preserved exactly. Otherwise, the rounding errors would spoil any potential improvement in the accuracy of the computed eigenvalues.
- This implies in practice that we do not need to stop the Sinkhorn-Knopp algorithm with a very stringent criterion, i.e., we do not need to converge to an (almost) stochastic matrix.
- In our experience, this means that for regular pencils the Sinkhorn-Knopp algorithm on $M := |A|^{\circ 2} + |B|^{\circ 2}$ converges "always in practice" with a relaxed stopping criterion (for instance, row sums and column sums equal up to a factor 2), even in situations for which the conditions in previous theorem are not satisfied.
- This is no longer true for SINGULAR square pencils, in particular, when they are sparse.

Outline

- A first idea is to extend LVD method with obvious restrictions as follows
- $$A, B \in \mathbb{C}^{m\times n}. \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

  such that

  $$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 = m, \quad \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 = n,$$

  for $i = 1, \dots, m$ and $j = 1, \dots, n$,
- which can be solved by the Sinkhorn-Knopp-like algorithm applied to $M = |A|^{\circ 2} + |B|^{\circ 2} \in \mathbb{R}^{m\times n}$ with $r = n\mathbf{1}_m$ and $c = m\mathbf{1}_n$.
- This is equivalent to solving the following minimization problem over positive diagonal matrices

  $$\inf_{\det D_\ell^2 = c_\ell, \det D_r^2 = c_r} \left(\|D_\ell A D_r\|_F^2 + \|D_\ell B D_r\|_F^2 \right)$$
- This strategy works well in practice for dense rectangular pencils but we have observed that the Sinkhorn-Knopp-like algorithm (with a relaxed stopping criterion) may not converge for some sparse rectangular pencils, for which this scaling problem does not have solution.

A first approach for scaling rectangular pencils

- A first idea is to extend LVD method with obvious restrictions as follows
- $$A, B \in \mathbb{C}^{m\times n}. \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

  such that

  $$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 = m, \qquad \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 = n,$$

  for $i = 1, \ldots, m$ and $j = 1, \ldots, n$,
- which can be solved by the Sinkhorn-Knopp-like algorithm applied to $M = |A|^{\circ 2} + |B|^{\circ 2} \in \mathbb{R}^{m\times n}$ with $r = n\mathbf{1}_m$ and $c = m\mathbf{1}_n$.
- This is equivalent to solving the following minimization problem over positive diagonal matrices

  $$\inf_{\det D_\ell^2 = c_\ell, \det D_r^2 = c_r} \left(\|D_\ell A D_r\|_F^2 + \|D_\ell B D_r\|_F^2 \right)$$

- This strategy works well in practice for dense rectangular pencils but we have observed that the Sinkhorn-Knopp-like algorithm (with a relaxed stopping criterion) may not converge for some sparse rectangular pencils, for which this scaling problem does not have solution.

- A first idea is to extend LVD method with obvious restrictions as follows
- $$A, B \in \mathbb{C}^{m\times n}. \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

  such that

  $$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 = m, \quad \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 = n,$$

  for $i = 1, \ldots, m$ and $j = 1, \ldots, n$,
- which can be solved by the Sinkhorn-Knopp-like algorithm applied to $M = |A|^{\circ 2} + |B|^{\circ 2} \in \mathbb{R}^{m\times n}$ with $r = n\mathbf{1}_m$ and $c = m\mathbf{1}_n$.
- This is equivalent to solving the following minimization problem over positive diagonal matrices

  $$\inf_{\det D_\ell^2 = c_\ell, \det D_r^2 = c_r} \left(\|D_\ell A D_r\|_F^2 + \|D_\ell B D_r\|_F^2 \right)$$
- This strategy works well in practice for dense rectangular pencils but we have observed that the Sinkhorn-Knopp-like algorithm (with a relaxed stopping criterion) may not converge for some sparse rectangular pencils, for which this scaling problem does not have solution.

A first approach for scaling rectangular pencils

- A first idea is to extend LVD method with obvious restrictions as follows
- $$A, B \in \mathbb{C}^{m\times n}. \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

  such that

  $$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 = m, \qquad \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 = n,$$

  for $i = 1, \ldots, m$ and $j = 1, \ldots, n$,
- which can be solved by the Sinkhorn-Knopp-like algorithm applied to $M = |A|^{\circ 2} + |B|^{\circ 2} \in \mathbb{R}^{m\times n}$ with $r = n\mathbf{1}_m$ and $c = m\mathbf{1}_n$.
- This is equivalent to solving the following minimization problem over positive diagonal matrices

  $$\inf_{\det D_\ell^2 = c_\ell, \det D_r^2 = c_r} \left(\|D_\ell A D_r\|_F^2 + \|D_\ell B D_r\|_F^2 \right)$$
- This strategy works well in practice for dense rectangular pencils but we have observed that the Sinkhorn-Knopp-like algorithm (with a relaxed stopping criterion) may not converge for some sparse rectangular pencils, for which this scaling problem does not have solution.

A first approach for scaling rectangular pencils

- A first idea is to extend LVD method with obvious restrictions as follows
- $$A, B \in \mathbb{C}^{m\times n}. \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A}$$

  such that

  $$\|\mathrm{col}_j(\widetilde{A})\|_2^2 + \|\mathrm{col}_j(\widetilde{B})\|_2^2 = m, \qquad \|\mathrm{row}_i(\widetilde{A})\|_2^2 + \|\mathrm{row}_i(\widetilde{B})\|_2^2 = n,$$

  for $i = 1, \ldots, m$ and $j = 1, \ldots, n$,
- which can be solved by the Sinkhorn-Knopp-like algorithm applied to $M = |A|^{\circ 2} + |B|^{\circ 2} \in \mathbb{R}^{m\times n}$ with $r = n\mathbf{1}_m$ and $c = m\mathbf{1}_n$.
- This is equivalent to solving the following minimization problem over positive diagonal matrices

  $$\inf_{\det D_\ell^2 = c_\ell, \det D_r^2 = c_r} \left(\|D_\ell A D_r\|_F^2 + \|D_\ell B D_r\|_F^2\right)$$

- This strategy works well in practice for dense rectangular pencils but we have observed that the Sinkhorn-Knopp-like algorithm (with a relaxed stopping criterion) may not converge for some sparse rectangular pencils, for which this scaling problem does not have solution.

Numerical test 3 illustrates Sinkhorn-Knopp-like for DENSE rectangular pencils

- 150×450 pencils with 149 eigenvalues, matrices $M = |A|^{\circ 2} + |B|^{\circ 2}$ very badly scaled in a nontrivial way (i.e., not constructed multiplying by diagonals).
- We apply the staircase algorithm to the original pencil (c_{orig}) and to the scaled one via Sinkhorn-Knopp-like algorithm (c_{bal}).
- We measure the "scaling" of M with $q_S(M) := \max \left\{ \frac{\max_i r_i(M)}{\min_i r_i(M)}, \frac{\max_i c_i(M)}{\min_i c_i(M)} \right\}$.

c_{orig}	c_{bal}	$q_S(M_{orig})$	$q_S(M_{scal})$	steps
9.96e-15	9.96e-15	2.29e+00	2.29e+00	2
1.95e-14	1.08e-14	4.94e+03	7.77e+00	4
2.62e-13	1.06e-14	1.22e+08	9.66e+00	7
2.27e-12	1.29e-14	4.32e+11	1.06e+01	9
5.61e-09	1.97e-13	1.36e+16	1.17e+01	12
1.51e-05	1.20e-13	8.19e+23	1.07e+01	14
6.03e-05	1.08e-12	3.51e+22	1.27e+01	21
5.49e-02	1.72e-11	2.39e+29	1.17e+01	16
9.76e-02	8.40e-12	1.24e+31	1.32e+01	24

- 150×450 pencils with 149 eigenvalues, matrices $M = |A|^{\circ 2} + |B|^{\circ 2}$ very badly scaled in a nontrivial way (i.e., not constructed multiplying by diagonals).
- We apply the staircase algorithm to the original pencil (c_{orig}) and to the scaled one via Sinkhorn-Knopp-like algorithm (c_{bal}).
- We measure the "scaling" of M with $q_S(M) := \max\left\{\frac{\max_i r_i(M)}{\min_i r_i(M)}, \frac{\max_i c_i(M)}{\min_i c_i(M)}\right\}$.

c_{orig}	c_{bal}	$q_S(M_{orig})$	$q_S(M_{scal})$	steps
9.96e-15	9.96e-15	2.29e+00	2.29e+00	2
1.95e-14	1.08e-14	4.94e+03	7.77e+00	4
2.62e-13	1.06e-14	1.22e+08	9.66e+00	7
2.27e-12	1.29e-14	4.32e+11	1.06e+01	9
5.61e-09	1.97e-13	1.36e+16	1.17e+01	12
1.51e-05	1.20e-13	8.19e+23	1.07e+01	14
6.03e-05	1.08e-12	3.51e+22	1.27e+01	21
5.49e-02	1.72e-11	2.39e+29	1.17e+01	16
9.76e-02	8.40e-12	1.24e+31	1.32e+01	24

Numerical test 3 illustrates Sinkhorn-Knopp-like for DENSE rectangular pencils

- 150×450 pencils with 149 eigenvalues, matrices $M = |A|^{\circ 2} + |B|^{\circ 2}$ very badly scaled in a nontrivial way (i.e., not constructed multiplying by diagonals).
- We apply the staircase algorithm to the original pencil (c_{orig}) and to the scaled one via Sinkhorn-Knopp-like algorithm (c_{bal}).
- We measure the "scaling" of M with $q_S(M) := \max \left\{ \frac{\max_i r_i(M)}{\min_i r_i(M)} , \frac{\max_i c_i(M)}{\min_i c_i(M)} \right\}$.

c_{orig}	c_{bal}	$q_S(M_{orig})$	$q_S(M_{scal})$	steps
9.96e-15	9.96e-15	2.29e+00	2.29e+00	2
1.95e-14	1.08e-14	4.94e+03	7.77e+00	4
2.62e-13	1.06e-14	1.22e+08	9.66e+00	7
2.27e-12	1.29e-14	4.32e+11	1.06e+01	9
5.61e-09	1.97e-13	1.36e+16	1.17e+01	12
1.51e-05	1.20e-13	8.19e+23	1.07e+01	14
6.03e-05	1.08e-12	3.51e+22	1.27e+01	21
5.49e-02	1.72e-11	2.39e+29	1.17e+01	16
9.76e-02	8.40e-12	1.24e+31	1.32e+01	24

Numerical test 3 illustrates Sinkhorn-Knopp-like for DENSE rectangular pencils

- 150×450 pencils with 149 eigenvalues, matrices $M = |A|^{\circ 2} + |B|^{\circ 2}$ very badly scaled in a nontrivial way (i.e., not constructed multiplying by diagonals).
- We apply the staircase algorithm to the original pencil (c_{orig}) and to the scaled one via Sinkhorn-Knopp-like algorithm (c_{bal}).
- We measure the "scaling" of M with $q_S(M) := \max \left\{ \frac{\max_i r_i(M)}{\min_i r_i(M)}, \frac{\max_i c_i(M)}{\min_i c_i(M)} \right\}$.

c_{orig}	c_{bal}	$q_S(M_{orig})$	$q_S(M_{scal})$	steps
9.96e-15	9.96e-15	2.29e+00	2.29e+00	2
1.95e-14	1.08e-14	4.94e+03	7.77e+00	4
2.62e-13	1.06e-14	1.22e+08	9.66e+00	7
2.27e-12	1.29e-14	4.32e+11	1.06e+01	9
5.61e-09	1.97e-13	1.36e+16	1.17e+01	12
1.51e-05	1.20e-13	8.19e+23	1.07e+01	14
6.03e-05	1.08e-12	3.51e+22	1.27e+01	21
5.49e-02	1.72e-11	2.39e+29	1.17e+01	16
9.76e-02	8.40e-12	1.24e+31	1.32e+01	24

Outline

- We regularize the scaling problem

$$A, B \in \mathbb{C}^{m\times n}, \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A},$$

- by considering the following constrained minimization problem over positive diagonal matrices D_ℓ and D_r:

$$\inf_{\det D_\ell^2 \det D_r^2 = c} 2(\|D_\ell A D_r\|_F^2 + \|D_\ell B D_r\|_F^2) + \alpha^2 \left(\frac{1}{m^2} \|D_\ell\|_F^4 + \frac{1}{n^2} \|D_r\|_F^4 \right),$$

for some real number $c > 0$ and a regularization parameter $\alpha \neq 0$, where the regularization term and the constraint penalize solutions with ill-conditioned D_ℓ and D_r.

- It can be proved that this minimization problem has always a unique solution $(\widetilde{D}_\ell, \widetilde{D}_r)$ that can be easily computed with the Sinkhorn-Knopp algorithm.

What to do when the Sinkhorn-Knopp-like algorithm on M does not converge?

- We regularize the scaling problem

$$A, B \in \mathbb{C}^{m\times n}, \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A},$$

- by considering the following constrained minimization problem over positive diagonal matrices D_ℓ and D_r:

$$\inf_{\det D_\ell^2 \det D_r^2 = c} 2(\|D_\ell A D_r\|_F^2 + \|D_\ell B D_r\|_F^2) + \alpha^2 \left(\frac{1}{m^2} \|D_\ell\|_F^4 + \frac{1}{n^2} \|D_r\|_F^4 \right),$$

for some real number $c > 0$ and a regularization parameter $\alpha \neq 0$, where the regularization term and the constraint penalize solutions with ill-conditioned D_ℓ and D_r.

- It can be proved that this minimization problem has always a unique solution $(\widetilde{D}_\ell, \widetilde{D}_r)$ that can be easily computed with the Sinkhorn-Knopp algorithm.

- We regularize the scaling problem

$$A, B \in \mathbb{C}^{m\times n}, \qquad \lambda B - A \longrightarrow \lambda D_\ell B D_r - D_\ell A D_r =: \lambda \widetilde{B} - \widetilde{A},$$

- by considering the following constrained minimization problem over positive diagonal matrices D_ℓ and D_r:

$$\inf_{\det D_\ell^2 \det D_r^2 = c} 2(\|D_\ell A D_r\|_F^2 + \|D_\ell B D_r\|_F^2) + \alpha^2 \left(\frac{1}{m^2} \|D_\ell\|_F^4 + \frac{1}{n^2} \|D_r\|_F^4 \right),$$

for some real number $c > 0$ and a regularization parameter $\alpha \neq 0$, where the regularization term and the constraint penalize solutions with ill-conditioned D_ℓ and D_r.

- It can be proved that this minimization problem has always a unique solution $(\widetilde{D}_\ell, \widetilde{D}_r)$ that can be easily computed with the Sinkhorn-Knopp algorithm.

Regularization ≡ Sinkhorn-Knopp on an extended matrix

- If $M = |A|^{\circ 2} + |B|^{\circ 2}$, then we define

$$M_\alpha^{\circ 2} = \begin{bmatrix} \frac{\alpha^2}{m^2}\mathbf{1}_m\mathbf{1}_m^T & M \\ M^T & \frac{\alpha^2}{n^2}\mathbf{1}_n\mathbf{1}_n^T \end{bmatrix}.$$

- We have proved that if $\alpha \neq 0$ and $M \neq 0$, then
- the nonnegative matrix $M_\alpha^{\circ 2}$ is fully indecomposable,
- it can be always diagonally scaled (multiplying by a unique diagonal matrix on the left and on the right) to have any prescribed common positive vector v for the row and column sums, and
- that the regularized minimization problem in the previous slide has as unique solution the unique diagonal matrices $(\widetilde{D}_\ell, \widetilde{D}_r)$ such that the matrix

$$\begin{bmatrix} \widetilde{D}_\ell^2 & 0 \\ 0 & \widetilde{D}_r^2 \end{bmatrix} M_\alpha^{\circ 2} \begin{bmatrix} \widetilde{D}_\ell^2 & 0 \\ 0 & \widetilde{D}_r^2 \end{bmatrix}$$

is (a scalar multiple of) a doubly stochastic matrix.

- Therefore, $(\widetilde{D}_\ell, \widetilde{D}_r)$ can be computed, by applying the Sinkhorn-Knopp algorithm to $M_\alpha^{\circ 2}$ which, in this case, always converges.

- If $M = |A|^{\circ 2} + |B|^{\circ 2}$, then we define

$$M_\alpha^{\circ 2} = \begin{bmatrix} \frac{\alpha^2}{m^2} \mathbf{1}_m \mathbf{1}_m^T & M \\ M^T & \frac{\alpha^2}{n^2} \mathbf{1}_n \mathbf{1}_n^T \end{bmatrix}.$$

- We have proved that if $\alpha \neq 0$ and $M \neq 0$, then
- the nonnegative matrix $M_\alpha^{\circ 2}$ is fully indecomposable,
- it can be always diagonally scaled (multiplying by a unique diagonal matrix on the left and on the right) to have any prescribed common positive vector v for the row and column sums, and
- that the regularized minimization problem in the previous slide has as unique solution the unique diagonal matrices $(\widetilde{D}_\ell, \widetilde{D}_r)$ such that the matrix

$$\begin{bmatrix} \widetilde{D}_\ell^2 & 0 \\ 0 & \widetilde{D}_r^2 \end{bmatrix} M_\alpha^{\circ 2} \begin{bmatrix} \widetilde{D}_\ell^2 & 0 \\ 0 & \widetilde{D}_r^2 \end{bmatrix}$$

is (a scalar multiple of) a doubly stochastic matrix.

- Therefore, $(\widetilde{D}_\ell, \widetilde{D}_r)$ can be computed, by applying the Sinkhorn-Knopp algorithm to $M_\alpha^{\circ 2}$ which, in this case, always converges.

- If $M = |A|^{\circ 2} + |B|^{\circ 2}$, then we define

$$M_\alpha^{\circ 2} = \begin{bmatrix} \frac{\alpha^2}{m^2} \mathbf{1}_m \mathbf{1}_m^T & M \\ M^T & \frac{\alpha^2}{n^2} \mathbf{1}_n \mathbf{1}_n^T \end{bmatrix}.$$

- We have proved that if $\alpha \neq 0$ and $M \neq 0$, then
- the nonnegative matrix $M_\alpha^{\circ 2}$ is fully indecomposable,
- it can be always diagonally scaled (multiplying by a unique diagonal matrix on the left and on the right) to have any prescribed common positive vector v for the row and column sums, and
- that the regularized minimization problem in the previous slide has as unique solution the unique diagonal matrices $(\widetilde{D}_\ell, \widetilde{D}_r)$ such that the matrix

$$\begin{bmatrix} \widetilde{D}_\ell^2 & 0 \\ 0 & \widetilde{D}_r^2 \end{bmatrix} M_\alpha^{\circ 2} \begin{bmatrix} \widetilde{D}_\ell^2 & 0 \\ 0 & \widetilde{D}_r^2 \end{bmatrix}$$

is (a scalar multiple of) a doubly stochastic matrix.

- Therefore, $(\widetilde{D}_\ell, \widetilde{D}_r)$ can be computed, by applying the Sinkhorn-Knopp algorithm to $M_\alpha^{\circ 2}$ which, in this case, always converges.

Regularization ≡ Sinkhorn-Knopp on an extended matrix

- If $M = |A|^{\circ 2} + |B|^{\circ 2}$, then we define

$$M_\alpha^{\circ 2} = \begin{bmatrix} \frac{\alpha^2}{m^2}\mathbf{1}_m\mathbf{1}_m^T & M \\ M^T & \frac{\alpha^2}{n^2}\mathbf{1}_n\mathbf{1}_n^T \end{bmatrix}.$$

- We have proved that if $\alpha \neq 0$ and $M \neq 0$, then
- the nonnegative matrix $M_\alpha^{\circ 2}$ is fully indecomposable,
- it can be always diagonally scaled (multiplying by a unique diagonal matrix on the left and on the right) to have any prescribed common positive vector v for the row and column sums, and
- that the regularized minimization problem in the previous slide has as unique solution the unique diagonal matrices $(\widetilde{D}_\ell, \widetilde{D}_r)$ such that the matrix

$$\begin{bmatrix} \widetilde{D}_\ell^2 & 0 \\ 0 & \widetilde{D}_r^2 \end{bmatrix} M_\alpha^{\circ 2} \begin{bmatrix} \widetilde{D}_\ell^2 & 0 \\ 0 & \widetilde{D}_r^2 \end{bmatrix}$$

is (a scalar multiple of) a doubly stochastic matrix.

- Therefore, $(\widetilde{D}_\ell, \widetilde{D}_r)$ can be computed, by applying the Sinkhorn-Knopp algorithm to $M_\alpha^{\circ 2}$ which, in this case, always converges.

Regularization ≡ Sinkhorn-Knopp on an extended matrix

- If $M = |A|^{\circ 2} + |B|^{\circ 2}$, then we define
$$M_\alpha^{\circ 2} = \begin{bmatrix} \frac{\alpha^2}{m^2}\mathbf{1}_m\mathbf{1}_m^T & M \\ M^T & \frac{\alpha^2}{n^2}\mathbf{1}_n\mathbf{1}_n^T \end{bmatrix}.$$
- We have proved that if $\alpha \neq 0$ and $M \neq 0$, then
- the nonnegative matrix $M_\alpha^{\circ 2}$ is fully indecomposable,
- it can be always diagonally scaled (multiplying by a unique diagonal matrix on the left and on the right) to have any prescribed common positive vector v for the row and column sums, and
- that the regularized minimization problem in the previous slide has as unique solution the unique diagonal matrices $(\widetilde{D}_\ell, \widetilde{D}_r)$ such that the matrix
$$\begin{bmatrix} \widetilde{D}_\ell^2 & 0 \\ 0 & \widetilde{D}_r^2 \end{bmatrix} M_\alpha^{\circ 2} \begin{bmatrix} \widetilde{D}_\ell^2 & 0 \\ 0 & \widetilde{D}_r^2 \end{bmatrix}$$
is (a scalar multiple of) a doubly stochastic matrix.
- Therefore, $(\widetilde{D}_\ell, \widetilde{D}_r)$ can be computed, by applying the Sinkhorn-Knopp algorithm to $M_\alpha^{\circ 2}$ which, in this case, always converges.

- If $M = |A|^{\circ 2} + |B|^{\circ 2}$, then we define

$$M_\alpha^{\circ 2} = \begin{bmatrix} \frac{\alpha^2}{m^2}\mathbf{1}_m\mathbf{1}_m^T & M \\ M^T & \frac{\alpha^2}{n^2}\mathbf{1}_n\mathbf{1}_n^T \end{bmatrix}.$$

- We have proved that if $\alpha \neq 0$ and $M \neq 0$, then
- the nonnegative matrix $M_\alpha^{\circ 2}$ is fully indecomposable,
- it can be always diagonally scaled (multiplying by a unique diagonal matrix on the left and on the right) to have any prescribed common positive vector v for the row and column sums, and
- that the regularized minimization problem in the previous slide has as unique solution the unique diagonal matrices $(\widetilde{D}_\ell, \widetilde{D}_r)$ such that the matrix

$$\begin{bmatrix} \widetilde{D}_\ell^2 & 0 \\ 0 & \widetilde{D}_r^2 \end{bmatrix} M_\alpha^{\circ 2} \begin{bmatrix} \widetilde{D}_\ell^2 & 0 \\ 0 & \widetilde{D}_r^2 \end{bmatrix}$$

  is (a scalar multiple of) a doubly stochastic matrix.
- Therefore, $(\widetilde{D}_\ell, \widetilde{D}_r)$ can be computed, by applying the Sinkhorn-Knopp algorithm to $M_\alpha^{\circ 2}$ which, in this case, always converges.

Additional comments on regularization

- In the case of rectangular pencils, we have observed that scalings with row sums of M closer to each other and with column sums of M closer to each other are obtained by scaling $M_\alpha^{\circ 2}$ with the Sinkhorn-Knopp-like algorithm with
$$v := \begin{bmatrix} n\mathbf{1}_m \\ m\mathbf{1}_n \end{bmatrix}$$
as prescribed common vector for the row and column sums, instead of with $v = \mathbf{1}_{m+n}$.
- The selection of the regularization parameter is always an issue in any regularization method.
- In our case, we recommend to try first "Sinkhorn-Knopp-like" (with relaxed stopping criterion) directly on M (with prescribed equal row sums and equal column sums) and if it does not converge in $\approx \max\{m, n\}/10$ iterations move to the regularized method with $\alpha \approx 0.5$ (assuming $\|M\|_F \approx 1$).
- We have not observed that the use of very small values of α have a relevant impact on the accuracy of the computed eigenvalues.

Additional comments on regularization

- In the case of rectangular pencils, we have observed that scalings with row sums of M closer to each other and with column sums of M closer to each other are obtained by scaling $M_\alpha^{\circ 2}$ with the Sinkhorn-Knopp-like algorithm with

$$v := \begin{bmatrix} n\mathbf{1}_m \\ m\mathbf{1}_n \end{bmatrix}$$

as prescribed common vector for the row and column sums, instead of with $v = \mathbf{1}_{m+n}$.

- The selection of the regularization parameter is always an issue in any regularization method.
- In our case, we recommend to try first "Sinkhorn-Knopp-like" (with relaxed stopping criterion) directly on M (with prescribed equal row sums and equal column sums) and if it does not converge in $\approx \max\{m, n\}/10$ iterations move to the regularized method with $\alpha \approx 0.5$ (assuming $\|M\|_F \approx 1$).
- We have not observed that the use of very small values of α have a relevant impact on the accuracy of the computed eigenvalues.

Additional comments on regularization

- In the case of rectangular pencils, we have observed that scalings with row sums of M closer to each other and with column sums of M closer to each other are obtained by scaling $M_\alpha^{\circ 2}$ with the Sinkhorn-Knopp-like algorithm with
$$v := \begin{bmatrix} n\mathbf{1}_m \\ m\mathbf{1}_n \end{bmatrix}$$
as prescribed common vector for the row and column sums, instead of with $v = \mathbf{1}_{m+n}$.
- The selection of the regularization parameter is always an issue in any regularization method.
- In our case, we recommend to try first "Sinkhorn-Knopp-like" (with relaxed stopping criterion) directly on M (with prescribed equal row sums and equal column sums) and if it does not converge in $\approx \max\{m,n\}/10$ iterations move to the regularized method with $\alpha \approx 0.5$ (assuming $\|M\|_F \approx 1$).
- We have not observed that the use of very small values of α have a relevant impact on the accuracy of the computed eigenvalues.

Additional comments on regularization

- In the case of rectangular pencils, we have observed that scalings with row sums of M closer to each other and with column sums of M closer to each other are obtained by scaling $M_\alpha^{\circ 2}$ with the Sinkhorn-Knopp-like algorithm with
$$v := \begin{bmatrix} n\mathbf{1}_m \\ m\mathbf{1}_n \end{bmatrix}$$
as prescribed common vector for the row and column sums, instead of with $v = \mathbf{1}_{m+n}$.
- The selection of the regularization parameter is always an issue in any regularization method.
- In our case, we recommend to try first "Sinkhorn-Knopp-like" (with relaxed stopping criterion) directly on M (with prescribed equal row sums and equal column sums) and if it does not converge in $\approx \max\{m,n\}/10$ iterations move to the regularized method with $\alpha \approx 0.5$ (assuming $\|M\|_F \approx 1$).
- We have not observed that the use of very small values of α have a relevant impact on the accuracy of the computed eigenvalues.

Numerical test 4 illustrates regularized method for SPARSE rectangular pencils

- 700×450 pencils with 148 eigenvalues, matrices $M = |A|^{\circ 2} + |B|^{\circ 2}$ very badly scaled in a nontrivial way (i.e., not constructed multiplying by diagonals) and "sparse".
- We apply the staircase algorithm to the original pencil (c_{orig}) and to the scaled one via the regularized algorithm (c_{bal}) with $\alpha = 0.5$.
- (The direct un-regularized method on M did not converge and produced diagonal scaling matrices with zero entries due to underflows.)
- We measure the "scaling" of M with $q_S(M) := \max\left\{\frac{\max_i r_i(M)}{\min_i r_i(M)}, \frac{\max_i c_i(M)}{\min_i c_i(M)}\right\}$.

c_{orig}	c_{bal}	$q_S(M_{orig})$	$q_S(M_{scal})$	steps
1.43e-14	1.26e-14	5.54e+01	3.27e+01	7
1.73e-14	1.39e-14	4.64e+06	9.30e+03	13
2.81e-13	3.75e-14	3.10e+11	1.80e+06	26
1.77e-11	1.98e-14	5.14e+19	1.42e+10	32
2.42e-06	6.23e-14	5.87e+28	1.09e+13	46
2.42e-02	1.15e-10	4.53e+29	1.11e+18	46

Numerical test 4 illustrates regularized method for SPARSE rectangular pencils

- 700×450 pencils with 148 eigenvalues, matrices $M = |A|^{\circ 2} + |B|^{\circ 2}$ very badly scaled in a nontrivial way (i.e., not constructed multiplying by diagonals) and "sparse".
- We apply the staircase algorithm to the original pencil (c_{orig}) and to the scaled one via the regularized algorithm (c_{bal}) with $\alpha = 0.5$.
- (The direct un-regularized method on M did not converge and produced diagonal scaling matrices with zero entries due to underflows.)
- We measure the "scaling" of M with $q_S(M) := \max\left\{\frac{\max_i r_i(M)}{\min_i r_i(M)}, \frac{\max_i c_i(M)}{\min_i c_i(M)}\right\}$.

c_{orig}	c_{bal}	$q_S(M_{orig})$	$q_S(M_{scal})$	steps
1.43e-14	1.26e-14	5.54e+01	3.27e+01	7
1.73e-14	1.39e-14	4.64e+06	9.30e+03	13
2.81e-13	3.75e-14	3.10e+11	1.80e+06	26
1.77e-11	1.98e-14	5.14e+19	1.42e+10	32
2.42e-06	6.23e-14	5.87e+28	1.09e+13	46
2.42e-02	1.15e-10	4.53e+29	1.11e+18	46

- 700×450 pencils with 148 eigenvalues, matrices $M = |A|^{\circ 2} + |B|^{\circ 2}$ very badly scaled in a nontrivial way (i.e., not constructed multiplying by diagonals) and "sparse".
- We apply the staircase algorithm to the original pencil (c_{orig}) and to the scaled one via the regularized algorithm (c_{bal}) with $\alpha = 0.5$.
- (The direct un-regularized method on M did not converge and produced diagonal scaling matrices with zero entries due to underflows.)
- We measure the "scaling" of M with $q_S(M) := \max \left\{ \frac{\max_i r_i(M)}{\min_i r_i(M)}, \frac{\max_i c_i(M)}{\min_i c_i(M)} \right\}$.

c_{orig}	c_{bal}	$q_S(M_{orig})$	$q_S(M_{scal})$	steps
1.43e-14	1.26e-14	5.54e+01	3.27e+01	7
1.73e-14	1.39e-14	4.64e+06	9.30e+03	13
2.81e-13	3.75e-14	3.10e+11	1.80e+06	26
1.77e-11	1.98e-14	5.14e+19	1.42e+10	32
2.42e-06	6.23e-14	5.87e+28	1.09e+13	46
2.42e-02	1.15e-10	4.53e+29	1.11e+18	46

Numerical test 4 illustrates regularized method for SPARSE rectangular pencils

- 700×450 pencils with 148 eigenvalues, matrices $M = |A|^{\circ 2} + |B|^{\circ 2}$ very badly scaled in a nontrivial way (i.e., not constructed multiplying by diagonals) and "sparse".
- We apply the staircase algorithm to the original pencil (c_{orig}) and to the scaled one via the regularized algorithm (c_{bal}) with $\alpha = 0.5$.
- (The direct un-regularized method on M did not converge and produced diagonal scaling matrices with zero entries due to underflows.)
- We measure the "scaling" of M with $q_S(M) := \max\left\{ \frac{\max_i r_i(M)}{\min_i r_i(M)}, \frac{\max_i c_i(M)}{\min_i c_i(M)} \right\}$.

c_{orig}	c_{bal}	$q_S(M_{orig})$	$q_S(M_{scal})$	steps
1.43e-14	1.26e-14	5.54e+01	3.27e+01	7
1.73e-14	1.39e-14	4.64e+06	9.30e+03	13
2.81e-13	3.75e-14	3.10e+11	1.80e+06	26
1.77e-11	1.98e-14	5.14e+19	1.42e+10	32
2.42e-06	6.23e-14	5.87e+28	1.09e+13	46
2.42e-02	1.15e-10	4.53e+29	1.11e+18	46

- 700×450 pencils with 148 eigenvalues, matrices $M = |A|^{\circ 2} + |B|^{\circ 2}$ very badly scaled in a nontrivial way (i.e., not constructed multiplying by diagonals) and "sparse".
- We apply the staircase algorithm to the original pencil (c_{orig}) and to the scaled one via the regularized algorithm (c_{bal}) with $\alpha = 0.5$.
- (The direct un-regularized method on M did not converge and produced diagonal scaling matrices with zero entries due to underflows.)
- We measure the "scaling" of M with $q_S(M) := \max \left\{ \frac{\max_i r_i(M)}{\min_i r_i(M)}, \frac{\max_i c_i(M)}{\min_i c_i(M)} \right\}$.

c_{orig}	c_{bal}	$q_S(M_{orig})$	$q_S(M_{scal})$	steps
1.43e-14	1.26e-14	5.54e+01	3.27e+01	7
1.73e-14	1.39e-14	4.64e+06	9.30e+03	13
2.81e-13	3.75e-14	3.10e+11	1.80e+06	26
1.77e-11	1.98e-14	5.14e+19	1.42e+10	32
2.42e-06	6.23e-14	5.87e+28	1.09e+13	46
2.42e-02	1.15e-10	4.53e+29	1.11e+18	46

Outline

Conclusions

- We have developed new scaling algorithms for both regular and singular pencils.
- We have revised and analyzed in detail previous scaling algorithms for pencils.
- The considered algorithms are based on applying the Sinkhorn-Knopp-like algorithm to certain nonnegative matrices easily constructed from the pencil.
- A regularization guarantees to get always a unique and bounded scaling, though very often the un-regularized algorithm works well in practice.
- Extensive numerical experiments confirm that the proposed algorithms very often improve significantly the accuracy of computed eigenvalues of arbitrary pencils.
- The scaling algorithms have a computational cost that is much smaller than the cost of the subsequent generalized eigenvalue algorithm as a consequence of using a stopping criterion compatible with computing diagonal scalings whose diagonal entries are integer powers of 2.