RECENT RESULTS ON MATRIX PERTURBATION THEORY

Froilán M. Dopico

Departamento de Matemáticas

Universidad Carlos III de Madrid

Main Goals of the Course

- To show recent results on Matrix Perturbation Theory.
- Results with applications to error analysis of accurate algorithms for spectral problems.
- These results can be easily proved assuming standard knowledge on Matrix Theory.
- To present a variety of techniques.

Outline of the Course

- Multiplicative perturbation theory for eigenvalues of Hermitian matrices and singular values.
- Multiplicative perturbation theory for invariant subspaces of Hermitian matrices and singular subspaces.
- Perturbations through factors.
- Perturbation Theory for factorizations of LU type.
- Perturbation of the Perron Root of nonnegative matrices.

One of the most famous motivations (I)

Definition (Singular Value Decomposition-SVD): Let $A \in \mathbb{C}^{m\times n}$, with $m \geq n$. Then we can write

$$A = U\Sigma V^*,$$

where $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ are unitary matrices, and $\Sigma \in \mathbb{R}^{m\times n}$ is diagonal. The diagonal elements of Σ,

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0,$$

are called the singular values of A.

One of the most famous motivations (II)

The singular values computed by the most frequently used algorithms (Golub-Kahan Bidiagonalization+QR, Divide and Conquer, etc), are the exact singular values of:

$$A + E, \quad \text{with} \quad \|E\|_2 = O(\epsilon)\|A\|_2,$$

where ϵ is the unit roundoff and $\|A\|_2$ is the spectral norm. (Additive BACKWARD error.)

Theorem (Weyl): Let $\sigma_1 \geq \ldots \geq \sigma_n$ be the singular values of A and $\widehat{\sigma}_1 \geq \ldots \geq \widehat{\sigma}_n$ be the singular values of $A + E$, then

$$|\sigma_i - \widehat{\sigma}_i| \leq \|E\|_2 \quad 1 \leq i \leq n.$$

One of the most famous motivations (III)

$$\left|\frac{\sigma_i - \widehat{\sigma}_i}{\sigma_i}\right| \le \frac{\|E\|_2}{\sigma_i} = O(\epsilon)\frac{\sigma_1}{\sigma_i}, \quad 1 \le i \le n.$$

Good if $\sigma_i \approx \sigma_1$. Bad if $\sigma_i \ll \sigma_1$.

Disastrous if $\sigma_i \le \epsilon\,\sigma_1$.

Other Algorithms for certain classes of matrices are able to compute all the singular values with small relative errors

$$\left|\frac{\sigma_i - \widehat{\sigma}_i}{\sigma_i}\right| = O(\epsilon), \quad 1 \le i \le n,$$

because they produce other types of backward errors.

Multiplicative Perturbations

Let $A \in \mathbb{C}^{m \times n}$, with $m \geq n$. A multiplicative perturbation of A is

$$\tilde{A} = D_1 A D_2 \in \mathbb{C}^{m \times n},$$

where D_1 and D_2 are nonsingular matrices.

If $A = A^* \in \mathbb{C}^{n \times n}$, then $\tilde{A} = D^* A D$.

Interesting case: D_1, D_2 close to the identity,

$$E_1 = D_1 - I \quad \text{and} \quad \|E_1\|_2 \ll 1$$
$$E_2 = D_2 - I \quad \text{and} \quad \|E_2\|_2 \ll 1$$

These are called SMALL MULTIPLICATIVE PERTURBATIONS

Small multiplicative perturbations imply small additive perturbations but the opposite is not true

Let $\widetilde{A} = (I + E_1)A(I + E_2)$, and denote $\max\{\|E_1\|_2, \|E_2\|_2\} = \eta$ then

$$\widetilde{A} = A + E_1\,A + A\,E_2 + E_1\,A\,E_2 \equiv A + F,$$

with $\|F\|_2 = (2\eta + \eta^2)\,\|A\|_2$.

Let $\widetilde{A} = A + F$ with $\|F\|_2 \leq \beta\|A\|_2$, then

$$\widetilde{A} = A(I + A^{-1}\,F) \equiv A(I + E),$$

with $\|E\|_2 \leq \beta\,\|A\|_2\,\|A^{-1}\|_2 = \beta\,\kappa_2(A)$.

Change of the singular values under Multiplicative Perturbations

Theorem 1 (Eisenstat-Ipsen 1995): Let

$$\widetilde{A} = (I + E_1)A(I + E_2)$$

with $\widetilde{A}, A \in \mathbb{C}^{m\times n}$ $(m \geq n)$. Let us assume that

$$\eta = \max\{\|E_1\|_2, \|E_2\|_2\} < 1.$$

Let $\sigma_1 \geq \ldots \geq \sigma_n$ and $\widetilde{\sigma}_1 \geq \ldots \geq \widetilde{\sigma}_n$ be, respectively, the singular values of A and $\widetilde{A}$. Then

$$|\widetilde{\sigma}_i - \sigma_i| \leq (2\,\eta + \eta^2)\,\sigma_i, \qquad 1 \leq i \leq n.$$

Preliminaries for the Proof of Theorem 1 (I)

Theorem (Courant-Fisher): Let $B = B^* \in \mathbb{C}^{n \times n}$ with eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$. Then

$$\lambda_k = \max_{\dim(S)=k} \min_{0 \neq y \in S} \frac{y^* B y}{y^* y},$$

with $S \subset \mathbb{C}^n$ subspace.

Corollary: Let $A \in \mathbb{C}^{m \times n}$ $(m \geq n)$, with singular values $\sigma_1 \geq \ldots \geq \sigma_n$. Then

$$\sigma_k = \max_{\dim(S)=k} \min_{0 \neq y \in S} \frac{\|Ay\|_2}{\|y\|_2}.$$

Preliminaries for the Proof of Theorem 1 (II)

Corollary: Let A, B, and AB three matrices, and let us denote by $\sigma_i(A), \sigma_i(B)$, and $\sigma_i(AB)$ the corresponding singular values ordered in non-increasing order. Then

$\sigma_i(AB) \leq \|A\|_2 \, \sigma_i(B)$ and $\sigma_i(AB) \leq \|B\|_2 \, \sigma_i(A)$.

NOW THE PROOF OF THEOREM 1 FOLLOWS EASILY

Change of the eigenvalues of Hermitian matrices under Multiplicative Perturbations

Theorem 2 (Eisenstat-Ipsen 1995):
Let $A = A^* \in \mathbb{C}^{n\times n}$, and

$$\widetilde{A} = D^*AD = (I+E^*)A(I+E)$$

with $D, E \in \mathbb{C}^{n\times n}$ and D nonsingular. Let $\lambda_1 \geq \ldots \geq \lambda_n$ and $\widetilde{\lambda}_1 \geq \ldots \geq \widetilde{\lambda}_n$ be, respectively, the eigenvalues of A and $\widetilde{A}$. Then

$$|\widetilde{\lambda}_i - \lambda_i| \leq \|D^*D - I\|_2 \, |\lambda_i| \leq (2\,\|E\|_2 + \|E\|_2^2)\, |\lambda_i|,$$

for $1 \leq i \leq n$.

Proof of Theorem 2 (I)

Sylvester's Inertia Theorem implies that 0 is the ith eigenvalue of

$$D^*(A-\lambda_i I)D = \tilde{A}-\lambda_i D^*D = (\tilde{A}-\lambda_i I)+\lambda_i(I-D^*D).$$

Weyl's Theorem will be used: Let $C = C^* \in \mathbb{C}^{n\times n}$ and $F = F^* \in \mathbb{C}^{n\times n}$. Let $\alpha_1 \geq \ldots \geq \alpha_n$ be the eigenvalues of C and $\tilde{\alpha}_1 \geq \ldots \geq \tilde{\alpha}_n$ be the eigenvalues of $C + F$, then

$$|\alpha_k - \tilde{\alpha}_k| \leq \|F\|_2, \quad 1 \leq k \leq n.$$

Proof of Theorem 2 (II)

Sylvester's Inertia Theorem implies that 0 is the ith eigenvalue of

$$D^*(A - \lambda_i I)D = (\widetilde{A} - \lambda_i I) + \lambda_i (I - D^* D).$$

Apply Weyl's Theorem to the ith eigenvalue of

$$(\widetilde{A} - \lambda_i I) \quad (\to \quad C)$$
$$(\widetilde{A} - \lambda_i I) + \lambda_i (I - D^* D) \quad (\to \quad C + F)$$

$$|(\widetilde{\lambda}_i - \lambda_i) - 0| = |\widetilde{\lambda}_i - \lambda_i| \leq |\lambda_i| \, \|D^* D - I\|_2.$$

Change of Invariant Subspaces: Preliminaries

Definition: The subspace $\mathcal{X} \subset \mathbb{C}^n$ is an invariant subspace of $A \in \mathbb{C}^{n\times n}$ if

$$A\mathcal{X} \subset \mathcal{X}, \text{ where } A\mathcal{X} = \{Az : z \in \mathcal{X}\}.$$

Theorem: Let $\mathcal{X}$ be an invariant subspace of $A \in \mathbb{C}^{n\times n}$, and let the columns of $X \in \mathbb{C}^{n\times p}$ form a basis of $\mathcal{X}$. Then there is a unique nonsingular matrix $L \in \mathbb{C}^{p\times p}$, such that

$$AX = XL,$$

and the eigenvalues of L are eigenvalues of A.

One dimensional invariant subspaces: X correspond to an eigenvector v and L correspond to an eigenvalue λ.

Change of Invariant Subspaces: Canonical Angles (I)

- Since we are concerned with variations of invariant subspaces of matrices when they are perturbed, we need a way to measure the differences between two subspaces.

Let $\mathcal{X}$ and $\mathcal{Y}$ be subspaces of the same dimension p. Let $X, Y \in \mathbb{C}^{n\times p}$ be orthonormal bases for $\mathcal{X}$ and $\mathcal{Y}$. Then

$$\|X^*Y\|_2 \leq \|X\|_2 \, \|Y\|_2 = 1.$$

- The singular values of X^*Y lie in $[0, 1]$.
- They can be considered as cosines of angles.
- They are independent of the choice of orthonormal bases for $\mathcal{X}$ and $\mathcal{Y}$.

Change of Invariant Subspaces: Canonical Angles (II)

Definition: Let $X, Y \in \mathbb{C}^{n\times p}$ be orthonormal bases of the subspaces of $\mathbb{C}^n$ $\mathcal{X}$ and $\mathcal{Y}$. Then the canonical angles between $\mathcal{X}$ and $\mathcal{Y}$ are

$$\theta_i = \cos^{-1}\gamma_i, \quad 1 \le i \le p, \quad 0 \le \theta_i \le \frac{\pi}{2}$$

where γ_i are the singular values of X^*Y. We will denote

$$\Theta(\mathcal{X},\mathcal{Y}) \equiv \Theta(X,Y) \equiv \text{diag}(\theta_1,\ldots,\theta_p), \text{ with } \theta_1 \ge \ldots \ge \theta_p.$$

Theorem: $\mathcal{X} = \mathcal{Y}$ if and only if $\Theta(\mathcal{X},\mathcal{Y}) = 0_{p\times p}$.

Theorem: Let $X_\perp$ be an orthonormal basis of the orthogonal complement of $\mathcal{X}$. Then the singular values of $X_\perp^* Y$ are the sines of the canonical angles between $\mathcal{X}$ and $\mathcal{Y}$.

Change of Invariant Subspaces: Canonical Angles (III)

Theorem: Let $\mathcal{X}$ and $\mathcal{Y}$ be two subspaces of $\mathbb{C}^n$ of dimension p. Let $P_\mathcal{X}$ and $P_\mathcal{Y}$ be the matrices of the orthogonal projections onto $\mathcal{X}$ and $\mathcal{Y}$, then

$$\|P_\mathcal{X} - P_\mathcal{Y}\|_2 = \|\sin\Theta(\mathcal{X},\mathcal{Y})\|_2 = \sin\theta_1,$$

moreover the singular values of $P_\mathcal{X} - P_\mathcal{Y}$ are

$$\{\sin\theta_1, \sin\theta_1, \sin\theta_2, \sin\theta_2, \ldots\}.$$

Therefore,

$$\begin{aligned} \mathrm{dist}_2(\mathcal{X},\mathcal{Y}) &= \|P_\mathcal{X} - P_\mathcal{Y}\|_2 \\ \mathrm{dist}_F(\mathcal{X},\mathcal{Y}) &= \|P_\mathcal{X} - P_\mathcal{Y}\|_F, \end{aligned}$$

are distances in the set of p-dimensional subspaces of $\mathbb{C}^n$, with a well defined geometrical meaning.

Change of Invariant Subspaces of Hermitian Matrices (I)

Let $A, \widetilde{A} \in \mathbb{C}^{n\times n}$ be two Hermitian matrices with unitary eigendecompositions

$$A = [\ \overbrace{U_1}^{p} \quad \overbrace{U_2}^{n-p}\ ] \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix} \begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix}$$

$$A + E \equiv \widetilde{A} = [\ \overbrace{\widetilde{U}_1}^{p} \quad \overbrace{\widetilde{U}_2}^{n-p}\ ] \begin{bmatrix} \widetilde{\Lambda}_1 & 0 \\ 0 & \widetilde{\Lambda}_2 \end{bmatrix} \begin{bmatrix} \widetilde{U}_1^* \\ \widetilde{U}_2^* \end{bmatrix}$$

Definition: The residual is $R = \widetilde{A}U_1 - U_1\Lambda_1 = (\widetilde{A} - A)U_1$.

Notice that: $\|R\|_F \leq \|\widetilde{A} - A\|_F \, \|U_1\|_2 = \|E\|_F$, i.e., the residual is small if E is small, and it is zero if $E = 0$.

Change of Invariant Subspaces of Hermitian Matrices (II)

Theorem (Davis-Kahan 1970) (Additive Perturbations):

$$\|\sin\Theta(U_1,\widetilde{U}_1)\|_F \le \frac{\|\widetilde{U}_2^*\,(\widetilde{A}-A)\,U_1\|_F}{\min\limits_{\substack{\widetilde{\mu}\in\widetilde{\Lambda}_2\\ \lambda\in\Lambda_1}} |\widetilde{\mu}-\lambda|} \le \frac{\|\widetilde{A}-A\|_F}{\min\limits_{\substack{\widetilde{\mu}\in\widetilde{\Lambda}_2\\ \lambda\in\Lambda_1}} |\widetilde{\mu}-\lambda|}$$

Proof:

$$\begin{aligned} \widetilde{U}_2^*\,R &= \widetilde{U}_2^*\widetilde{A}U_1 - \widetilde{U}_2^*U_1\Lambda_1 \\ &= \widetilde{\Lambda}_2\widetilde{U}_2^*U_1 - \widetilde{U}_2^*U_1\Lambda_1 \end{aligned}$$

Therefore, $(\widetilde{U}_2^*\,R)_{ij} = (\widetilde{\Lambda}_2)_{ii}\,(\widetilde{U}_2^*U_1)_{ij} - (\widetilde{U}_2^*U_1)_{ij}\,(\Lambda_1)_{jj}$

Then

$$(\widetilde{U}_2^*U_1)_{ij} = \frac{(\widetilde{U}_2^*\,R)_{ij}}{(\widetilde{\Lambda}_2)_{ii} - (\Lambda_1)_{jj}}$$

Drawback of Sin Θ Theorems

$$\|\sin\Theta(U_1,\widetilde{U}_1)\|_F \leq \frac{\|\widetilde{U}_2^*(\widetilde{A}-A)U_1\|_F}{\min\limits_{\substack{\widetilde{\mu}\in\widetilde{\Lambda}_2 \\ \lambda\in\Lambda_1}} |\widetilde{\mu}-\lambda|} \leq \frac{\|\widetilde{A}-A\|_F}{\min\limits_{\substack{\widetilde{\mu}\in\widetilde{\Lambda}_2 \\ \lambda\in\Lambda_1}} |\widetilde{\mu}-\lambda|}$$

Remark: The algorithms compute the exact eigenvalues of $\widetilde{A} = A+E$ (the name does not matter), therefore the eigenvalues of A are not known.

Sol. 1. To use the most complicated $\sin 2\Theta$ Theorems.

Sol. 2. To use Weyl's Perturbation Theorem for eigenvalues and

$$\min_{\substack{\widetilde{\mu}\in\widetilde{\Lambda}_2 \\ \lambda\in\Lambda_1}} |\widetilde{\mu}-\lambda| \longrightarrow \min_{\substack{\widetilde{\mu}\in\widetilde{\Lambda}_2 \\ \widetilde{\lambda}\in\widetilde{\Lambda}_1}} |\widetilde{\mu}-\widetilde{\lambda}| - \|E\|_2,$$

IF POSSIBLE.

Change of Invariant Subspaces of Hermitian Matrices (III)

Let $A, \widetilde{A} \in \mathbb{C}^{n\times n}$ be two Hermitian matrices such that

$$A = [\overbrace{U_1}^{p} \quad \overbrace{U_2}^{n-p}] \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix} \begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix}$$

$$D^* A D \equiv \widetilde{A} = [\overbrace{\widetilde{U}_1}^{p} \quad \overbrace{\widetilde{U}_2}^{n-p}] \begin{bmatrix} \widetilde{\Lambda}_1 & 0 \\ 0 & \widetilde{\Lambda}_2 \end{bmatrix} \begin{bmatrix} \widetilde{U}_1^* \\ \widetilde{U}_2^* \end{bmatrix}$$

Theorem (R. C. Li 1998) (Multiplicative Perturbations):

$$\|\sin\Theta(U_1, \widetilde{U}_1)\|_F \le \|(I - D^*)U_1\|_F + \frac{\|(D^* - D^{-1})U_1\|_F}{\displaystyle\min_{\substack{\widetilde{\mu} \in \widetilde{\Lambda}_2 \\ \lambda \in \Lambda_1}} \left|\frac{\widetilde{\mu} - \lambda}{\widetilde{\mu}}\right|}$$

Change of Invariant Subspaces of Hermitian Matrices (IV)

Corollary: If $D = I + F$ and $\|F\|_2 < 1$ then

$$\|\sin\Theta(U_1, \widetilde{U}_1)\|_F \leq \sqrt{p} \left(\|F\|_2 + \frac{2\,\|F\|_2 + \|F\|_2^2}{1 - \|F\|_2} \frac{1}{\min\limits_{\substack{\widetilde{\mu} \in \widetilde{\Lambda}_2 \\ \lambda \in \Lambda_1}} \left| \frac{\widetilde{\mu} - \lambda}{\widetilde{\mu}} \right|} \right)$$

To be compare with:

Theorem (Davis-Kahan 1970) (Additive Perturbations):

$$\|\sin\Theta(U_1, \widetilde{U}_1)\|_F \leq \frac{\|\widetilde{A} - A\|_F}{\min\limits_{\substack{\widetilde{\mu} \in \widetilde{\Lambda}_2 \\ \lambda \in \Lambda_1}} |\widetilde{\mu} - \lambda|}$$

Change of Invariant Subspaces of Hermitian Matrices (V)

Theorem (R. C. Li 1998): Let A and $\widetilde{A} = D^* A D$ be two Hermitian matrices. Then

$$\|\sin\Theta(U_1, \widetilde{U}_1)\|_F \leq \|(I - D^*)U_1\|_F + \frac{\|(D^* - D^{-1})U_1\|_F}{\min\limits_{\substack{\widetilde{\mu} \in \widetilde{\Lambda}_2 \\ \lambda \in \Lambda_1}} \left| \frac{\widetilde{\mu} - \lambda}{\widetilde{\mu}} \right|}$$

Proof:

$$\widetilde{U}_2^* R = \widetilde{U}_2^* \widetilde{A} U_1 - \widetilde{U}_2^* U_1 \Lambda_1 = \widetilde{\Lambda}_2 \widetilde{U}_2^* U_1 - \widetilde{U}_2^* U_1 \Lambda_1$$

$$\|$$

$$\widetilde{U}_2^* R = \widetilde{U}_2^* (\widetilde{A} - A) U_1 = \widetilde{U}_2^* [D^* A D (I - D^{-1}) + (D^* - I) A] U_1$$

$$\widetilde{\Lambda}_2 \widetilde{U}_2^* U_1 - \widetilde{U}_2^* U_1 \Lambda_1 = \widetilde{\Lambda}_2 \widetilde{U}_2^* (I - D^{-1}) U_1 + \widetilde{U}_2^* (D^* - I) U_1 \Lambda_1$$

Let us define $X \equiv \widetilde{U}_2^* U_1 - \widetilde{U}_2^* (I - D^*) U_1$. Then from

Let us define $X \equiv \widetilde{U}_2^* U_1 - \widetilde{U}_2^*(I - D^*)U_1$. Then from

$$\widetilde{\Lambda}_2 \widetilde{U}_2^* U_1 - \widetilde{U}_2^* U_1 \Lambda_1 = \widetilde{\Lambda}_2 \widetilde{U}_2^* (I - D^{-1}) U_1 + \widetilde{U}_2^* (D^* - I) U_1 \Lambda_1$$

we get

$$\widetilde{\Lambda}_2 X - X \Lambda_1 = \widetilde{\Lambda}_2 \widetilde{U}_2^* (D^* - D^{-1}) U_1$$

This means

$$X_{ij} = \frac{(\widetilde{\Lambda}_2)_{ii}\,(\widetilde{U}_2^*(D^* - D^{-1})U_1)_{ij}}{(\widetilde{\Lambda}_2)_{ii} - (\Lambda_1)_{jj}}$$

Combining this with

$$\|\sin\Theta(U_1, \widetilde{U}_1)\|_F = \|\widetilde{U}_2^* U_1\|_F \le \|X\|_F + \|\widetilde{U}_2^*(I - D^*)U_1\|_F,$$

the result follows.

Jordan-Wielandt Matrices: a way to get perturbation results for the SVD from perturbation results for Hermitian matrices

Definition: Let $B \in \mathbb{C}^{m\times n}$, $m \geq n$. The Hermitian matrix

$$C = \begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix},$$

is the Jordan-Wielandt matrix associated with B.

Let us partition the SVD of B as

$$B = \begin{bmatrix} U_B & U_0 \end{bmatrix} \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^*,$$

where $\Sigma \in \mathbb{R}^{n\times n}$, $[U_B\, U_0] \in \mathbb{C}^{m\times m}$, and $V \in \mathbb{C}^{n\times n}$.

Eigendecomposition of Jordan-Wielandt Matrices

Let us define the unitary matrix

$$U_C = \frac{1}{\sqrt{2}} \begin{bmatrix} U_B & U_B & \sqrt{2}\,U_0 \\ V & -V & 0 \end{bmatrix} \in \mathbb{C}^{(m+n)\times(m+n)},$$

then

$$\begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix} = U_C \begin{bmatrix} \Sigma & 0 & 0 \\ 0 & -\Sigma & 0 \\ 0 & 0 & 0 \end{bmatrix} U_C^*,$$

i.e., the eigenvalues of the Jordan-Wielandt matrix are:

1. The n singular values Σ of B.
2. Minus the n singular values Σ of B.
3. $m-n$ additional zeroes,

with eigenvectors the corresponding columns of U_C.

Singular Subspaces

Let us partition the SVD of B as

$$B = \begin{bmatrix} U_B & U_0 \end{bmatrix} \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^*.$$

Definition: A right (left) singular subspace of B is an invariant subspace of B^*B (BB^*).

Definition: A pair of subspaces $\mathcal{U} \in \mathbb{C}^m$ and $\mathcal{V} \in \mathbb{C}^n$ form a pair of left and right singular subspaces of B if they are singular subspaces and

$$Bv \in \mathcal{U} \text{ for all } v \in \mathcal{V} \quad AND \quad B^*u \in \mathcal{V} \text{ for all } u \in \mathcal{U}$$

The columns of U and the columns of V span, respectively, pairs of left and right singular subspaces of B.

Change of singular subspaces (I)

Let $B, \widetilde{B} \in \mathbb{C}^{m\times n}$ be two matrices with SVDs

$$B = [\ \overbrace{U_1}^{p}\ \ \overbrace{U_2}^{n-p}\ \ \overbrace{U_0}^{m-n}\ ] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix} \begin{matrix} \}\ p \\ \}\ n-p \end{matrix}$$

$$B + E \equiv \widetilde{B} = [\ \overbrace{\widetilde{U}_1}^{p}\ \ \overbrace{\widetilde{U}_2}^{n-p}\ \ \overbrace{\widetilde{U}_0}^{m-n}\ ] \begin{bmatrix} \widetilde{\Sigma}_1 & 0 \\ 0 & \widetilde{\Sigma}_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{V}_1^* \\ \widetilde{V}_2^* \end{bmatrix} \begin{matrix} \}\ p \\ \}\ n-p \end{matrix}$$

Definition:
The right residual is $R_R = \widetilde{B}V_1 - U_1\Sigma_1 = (\widetilde{B} - B)V_1$.

The left residual is $R_L = \widetilde{B}^*U_1 - V_1\Sigma_1 = (\widetilde{B} - B)^*U_1$.

Change of singular subspaces (II)

Theorem (Wedin 1972) (Additive Perturbations):

$$\sqrt{\|\sin\Theta(U_1,\widetilde{U}_1)\|_F^2 + \|\sin\Theta(V_1,\widetilde{V}_1)\|_F^2} \leq \frac{\sqrt{\|R_L\|_F^2 + \|R_R\|_F^2}}{\min\limits_{\substack{\widetilde{\mu} \in (\widetilde{\Sigma}_2)_{ext} \\ \lambda \in \Sigma_1}} |\widetilde{\mu} - \lambda|},$$

with $(\widetilde{\Sigma}_2)_{ext} = \widetilde{\Sigma}_2$ if $m = n$, and $(\widetilde{\Sigma}_2)_{ext} = \widetilde{\Sigma}_2 \cup \{0\}$ if $m > n$.

Proof: Consider the invariant subspace of $C = \begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix}$ spanned by the columns of $W_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} U_1 & U_1 \\ V_1 & -V_1 \end{bmatrix}$, associated with the eigen-values $\Sigma_1 \cup (-\Sigma_1)$.

The key idea is to apply Davis-Kahan perturbation Theorem for invariant subspaces of Hermitian matrices to C, $\widetilde{C}$ and W_1, $\widetilde{W}_1$.

Change of singular subspaces (III)

This means

$$\| \sin \Theta(W_1, \widetilde{W}_1) \|_F \le \frac{\|(\widetilde{C} - C) W_1\|_F}{\min\limits_{\substack{\widetilde{\mu} \in (\widetilde{\Sigma}_2)_{ext} \\ \lambda \in \Sigma_1}} |\widetilde{\mu} - \lambda|}$$

Some direct computations show

$$\| \sin \Theta(W_1, \widetilde{W}_1) \|_F = \sqrt{\| \sin \Theta(U_1, \widetilde{U}_1) \|_F^2 + \| \sin \Theta(V_1, \widetilde{V}_1) \|_F^2},$$

and

$$\|(\widetilde{C} - C) W_1\|_F = \sqrt{\|R_L\|_F^2 + \|R_R\|_F^2}.$$

This proves Wedin's Theorem.

Change of Singular Subspaces (IV): Multiplicative Perturbations

Theorem (R. C. Li 1998) (Multiplicative Perturbations): Let B and $\widetilde{B} = D_1 B D_2$ be two $m \times n$ $(m \geq n)$ matrices, then

$$\sqrt{\|\sin\Theta(U_1,\widetilde{U}_1)\|_F^2 + \|\sin\Theta(V_1,\widetilde{V}_1)\|_F^2} \leq \sqrt{\|(I-D_1)U_1\|_F^2 + \|(I-D_2^*)V_1\|_F^2}$$

$$+\frac{\sqrt{\|(D_1 - D_1^{-*})U_1\|_F^2 + \|(D_2^* - D_2^{-1})V_1\|_F^2}}{\displaystyle\min_{\substack{\widetilde{\mu} \in (\widetilde{\Sigma}_2)_{ext} \\ \lambda \in \Sigma_1}} \frac{|\widetilde{\mu} - \lambda|}{|\widetilde{\mu}|}},$$

Proof: Consider again the invariant subspace of $C = \begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix}$ spanned by the columns of $W_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} U_1 & U_1 \\ V_1 & -V_1 \end{bmatrix}$, associated with the eigenvalues $\Sigma_1 \cup (-\Sigma_1)$.

Change of Singular Subspaces (V): Multiplicative Perturbations

The key idea is to apply R. C. Li perturbation Theorem for invariant subspaces of Hermitian matrices to C, $\widetilde{C}$ and W_1, $\widetilde{W}_1$, taking into account that

$$\begin{bmatrix} 0 & \widetilde{B} \\ \widetilde{B}^* & 0 \end{bmatrix} = \begin{bmatrix} 0 & D_1 B D_2 \\ D_2^* B^* D_1^* & 0 \end{bmatrix} = \begin{bmatrix} D_1 & 0 \\ 0 & D_2^* \end{bmatrix} \begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix} \begin{bmatrix} D_1^* & 0 \\ 0 & D_2 \end{bmatrix}$$

This means

$$\|\sin\Theta(W_1, \widetilde{W}_1)\|_F \leq \left\| \begin{bmatrix} I - D_1 & 0 \\ 0 & I - D_2^* \end{bmatrix} W_1 \right\|_F + \frac{\left\| \begin{bmatrix} D_1 - D_1^{-*} & 0 \\ 0 & D_2^* - D_2^{-1} \end{bmatrix} W_1 \right\|_F}{\displaystyle\min_{\substack{\widetilde{\mu} \in (\widetilde{\Sigma}_2)_{ext} \\ \lambda \in \Sigma_1}} \frac{|\widetilde{\mu} - \lambda|}{|\widetilde{\mu}|}},$$

and the result follows after some easy manipulations.

CONCLUSIONS ON MULTIPLICATIVE PERTURBATION THEORY

- Small multiplicative perturbations of general matrices imply small relative changes of the singular values and small changes of the singular subspaces with respect relative gaps.

- Small multiplicative perturbations of Hermitian matrices imply small relative changes of the eigenvalues and small changes of the invariant subspaces with respect relative gaps.

Rank Revealing Decomposition (RRD)

Definition: An RRD of $G \in \mathbb{C}^{m \times n}$, $m \geq n$, is a factorization

$$G = XDY^*$$

where $D \in \mathbb{C}^{r \times r}$ is diagonal and nonsingular, and $X \in \mathbb{C}^{m \times r}$ and $Y \in \mathbb{C}^{n \times r}$ are well conditioned matrices of full column rank (notice that this implies $r = \text{rank}(G)$).

RRD and Accurate Computation of SVD

Given an RRD computed with the forward errors:

$$|D_{ii} - \widehat{D}_{ii}| = O(\epsilon)|D_{ii}|,$$

$$\|X - \widehat{X}\|_2 = O(\epsilon)\|X\|_2,$$

$$\|Y - \widehat{Y}\|_2 = O(\epsilon)\|Y\|_2,$$

where ϵ is the unit roundoff, Demmel et al. (1999) developed algorithms to compute SVDs with high relative accuracy, i.e.,

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

$$\max\{\theta(v_i, \widehat{v}_i), \theta(u_i, \widehat{u}_i)\} \le \frac{O(\max\{\kappa(X), \kappa(Y)\}\epsilon)}{\min_{j \ne i} \left| \frac{\sigma_i - \sigma_j}{\sigma_i} \right|}$$

Similar bounds hold for singular subspaces.

Perturbations of SVD through factors of RRDs (I)

Theorem (Demmel et al 1999): Let $G = XDY^*$ and $\widetilde{G} = \widetilde{X}\widetilde{D}\widetilde{Y}^*$ be RRDs of the $m \times n$ $(m \geq n)$ matrices G and $\widetilde{G}$. Let $\sigma_1 \geq \ldots \geq \sigma_n$ (resp. $\tilde{\sigma}_1 \geq \ldots \geq \tilde{\sigma}_n$) be the singular values of G (resp. $\widetilde{G}$). Let $u_1, \ldots, u_n$ (resp. $v_1, \ldots, v_n$) and $\tilde{u}_1, \ldots, \tilde{u}_n$ (resp. $\tilde{v}_1, \ldots, \tilde{v}_n$) be the corresponding left (resp. right) singular vectors. Let us assume that

$$\frac{\|\widetilde{X} - X\|_2}{\|X\|_2} \leq \beta, \qquad \frac{\|\widetilde{Y} - Y\|_2}{\|Y\|_2} \leq \beta, \qquad \frac{|\widetilde{D}_{ii} - D_{ii}|}{|D_{ii}|} \leq \beta \quad \text{for all } i,$$

where $0 \leq \beta < 1$. Let $\eta = \beta\,(2 + \beta)\max\{\kappa_2(X), \kappa_2(Y)\}$, then

$$|\sigma_i - \tilde{\sigma}_i| \leq (2\eta + \eta^2)\,|\sigma_i|, \qquad 1 \leq i \leq n,$$

and

$$\sqrt{\sin^2\theta(u_i, \tilde{u}_i) + \sin^2\theta(v_i, \tilde{v}_i)} \leq \sqrt{2}\,\frac{\eta}{1-\eta}\left(1 + \frac{2+\eta}{\min_{j\neq i}\frac{|\sigma_i - \tilde{\sigma}_j|}{|\tilde{\sigma}_j|}}\right), \quad 1 \leq i \leq n,$$

Similar bounds hold for singular subspaces.

Perturbations of SVD through factors of RRDs (II)

Proof: Write

$$\widetilde{G} = (I + E_1) G (I + E_2),$$

with

$$\begin{aligned} \|E_1\|_2 &\leq \beta\,\kappa_2(X) \\ \|E_2\|_2 &\leq (2\beta + \beta^2)\,\kappa_2(Y). \end{aligned}$$

Apply multiplicative perturbation Theorems for singular values and singular subspaces.

Perturbations of Eigendecompositions of Hermitian matrices through factors of RRDs

Theorem (FMD and Koev 2005) Let $A = XDX^T$ and $\widetilde{A} = \widetilde{X}\widetilde{D}\widetilde{X}^T$ be RRDs of the symmetric $n \times n$ matrices A and $\widetilde{A}$. Let $\lambda_1 \geq \ldots \geq \lambda_n$ be the eigenvalues of A and $\tilde{\lambda}_1 \geq \ldots \geq \tilde{\lambda}_n$ be the eigenvalues of $\widetilde{A}$. Let $q_1, \ldots, q_n$ and $\tilde{q}_1, \ldots, \tilde{q}_n$ be the corresponding orthonormal eigenvectors. Let us assume that

$$\frac{\|\widetilde{X} - X\|_2}{\|X\|_2} \leq \beta, \qquad \frac{|\widetilde{D}_{ii} - D_{ii}|}{|D_{ii}|} \leq \beta \quad \text{for all } i,$$

where $0 \leq \beta < 1$. Let $\eta = \beta\,(2 + \beta)\,\kappa_2(X)$, then

$$|\lambda_i - \tilde{\lambda}_i| \leq (2\eta + \eta^2)\,|\lambda_i|, \qquad 1 \leq i \leq n,$$

and

$$\sin\theta(q_i, \tilde{q}_i) \leq \frac{\eta}{1-\eta}\left(1 + \frac{2+\eta}{\min_{j\neq i} \frac{|\tilde{\lambda}_i - \lambda_j|}{|\lambda_j|}}\right), \qquad 1 \leq i \leq n.$$

Similar bounds hold for invariant subspaces.

Conclusions on Perturbation through factors of RRDs

- Computing accurate RRDs allows us to compute accurate SVDs.

- Computing accurate RRDs of Hermitian matrices allows us to compute accurate eigendecompositions (FMD, Molera, Moro 2003).

- RRDs are computed in practice through LDU factorizations by using Gaussian elimination with complete pivoting.

- Only possible for special classes of matrices.

- Forward errors: backward errors of LDU + Perturbation Theory of LDU

Perturbation Theory for factorizations of LU type (I)

- Approach based on Series Expansions (FMD, Molera 2005).

- Broad scoped approach that remains valid for all factorizations of LU type: LU, LDU, LDL^T, Cholesky, Block LU, Block LDL^T from diagonal pivoting method…..

- New results for Block LU and Block LDL^T from diagonal pivoting method. Results previously known but improved for the rest of factorizations.

- For the sake of brevity, we will focus on the usual LU factorization, but parallel developments yield results for other factorizations.

We assume L UNIT lower triangular and U upper triangular.

Perturbation Theory for factorizations of LU type (II)

THE PROBLEM: Given a nonsingular square matrix $A = LU$ having LU factorization and a perturbation $\widetilde{A} \equiv A + E = \widetilde{L}\widetilde{U}$, we look for two convergent series of matrices:

$$\widetilde{L} = \sum_{k=0}^{\infty} L_k \quad \text{and} \quad \widetilde{U} = \sum_{k=0}^{\infty} U_k$$

with $L_k = O(\|E\|^k)$, $U_k = O(\|E\|^k)$, $L_0 = L$, $U_0 = U$, as well as for domains and rates of convergence.

MAIN IDEA: PERTURBATION OF THE IDENTITY.

If $I + L^{-1}EU^{-1} \equiv \mathcal{L}\mathcal{U}$ then

$$\widetilde{A} = L(I + L^{-1}EU^{-1})U = (L\mathcal{L})(\mathcal{U}U) = \widetilde{L}\widetilde{U}.$$

Therefore, we can focus on the series of the LU factors of a perturbation of the identity: $I + F = \mathcal{L}\mathcal{U}$.

Perturbation Theory for factorizations of LU type (III)

We will use absolute, or monotone, consistent norms

$$|A| \leq |B| \Rightarrow \|A\| \leq \|B\| \quad \text{and} \quad \|AB\| \leq \|A\|\|B\|.$$

Let us introduce for simplicity a parameter z

$$I + zF = \mathcal{L}(z)\mathcal{U}(z).$$

For $\|zF\| < 1$, the LU factors of $I + zF$

- exist and are unique, because all the leading principal submatrices of $I + zF$ are nonsingular,

- their entries are rational functions of z with nonzero denominators, therefore

$$\mathcal{L}(z) = \sum_{k=0}^{\infty} z^k \mathcal{L}_k \quad \text{and} \quad \mathcal{U}(z) = \sum_{k=0}^{\infty} z^k \mathcal{U}_k,$$

with $\mathcal{L}_0 = \mathcal{U}_0 = I$.

How are the matrices $\mathcal{L}_k$ and $\mathcal{U}_k$ obtained?

Recurrence relations for $\mathcal{L}_k$ and $\mathcal{U}_k$.

$$I + zF = \mathcal{L}(z)\mathcal{U}(z) = \sum_{k=0}^{\infty} \left(\sum_{j=0}^{k} \mathcal{L}_j \mathcal{U}_{k-j} \right) z^k.$$

Consequently

$$\begin{aligned} F &= \mathcal{U}_1 + \mathcal{L}_1 \\ 0 &= \mathcal{U}_k + \mathcal{L}_1 \mathcal{U}_{k-1} + \cdots + \mathcal{L}_{k-1}\mathcal{U}_1 + \mathcal{L}_k \quad \text{for } k \geq 2. \end{aligned}$$

Notice that $\mathcal{L}_k$ is strict lower triangular, i.e., zeroes on the main diagonal, and $\mathcal{U}_k$ is upper triangular, **thus $\mathcal{L}_k$ and $\mathcal{U}_k$ can be obtained from $\mathcal{L}_1, \ldots, \mathcal{L}_{k-1}$ and $\mathcal{U}_1, \ldots, \mathcal{U}_{k-1}$.**

Perturbation Theory for factorizations of LU type (V)

Setting $z = 1$, we can state:

Theorem (FMD, Molera 2005). Let F be an $n \times n$ matrix with $\|F\| < 1$ then:

1. $I + F$ has a unique LU factorization: $I + F = \mathcal{L}\mathcal{U}$.

2.

$$\mathcal{L} = \sum_{k=0}^{\infty} \mathcal{L}_k \quad and \quad \mathcal{U} = \sum_{k=0}^{\infty} \mathcal{U}_k,$$

where $\mathcal{L}_0 = I$, $\mathcal{U}_0 = I$, $\mathcal{L}_1 = F_L$, $\mathcal{U}_1 = F_U$ and for $k \geq 2$:

$$\begin{aligned} \mathcal{L}_k &= \left(-\mathcal{L}_1 \mathcal{U}_{k-1} - \mathcal{L}_2 \mathcal{U}_{k-2} \cdots - \mathcal{L}_{k-1} \mathcal{U}_1\right)_L, \\ \mathcal{U}_k &= \left(-\mathcal{L}_1 \mathcal{U}_{k-1} - \mathcal{L}_2 \mathcal{U}_{k-2} \cdots - \mathcal{L}_{k-1} \mathcal{U}_1\right)_U. \end{aligned}$$

3. $L_k = O(\|F\|^k)$ and $U_k = O(\|F\|^k)$.

Here $(\cdot)_L$ stands for the strict lower triangular part and $(\cdot)_U$ stands for the upper triangular part.

A few terms:

$$\begin{aligned}
\mathcal{L}_1 + \mathcal{U}_1 &= F \\
\mathcal{L}_2 + \mathcal{U}_2 &= -F_L F_U \\
\mathcal{L}_3 + \mathcal{U}_3 &= F_L (F_L F_U)_U + (F_L F_U)_L F_U \\
\mathcal{L}_4 + \mathcal{U}_4 &= -F_L (F_L (F_L F_U)_U)_U - F_L ((F_L F_U)_L F_U)_U \\
&\quad -(F_L F_U)_L (F_L F_U)_U - (F_L (F_L F_U)_U)_L F_U \\
&\quad -((F_L F_U)_L F_U)_L F_U
\end{aligned}$$

Theorem (FMD, Molera 2005):

$$|\mathcal{L}_k + \mathcal{U}_k| \leq |F|^k \quad \text{for} \quad k \geq 1.$$

Therefore $|\mathcal{L}_k| \leq \left(|F|^k\right)_L$ and $|\mathcal{U}_k| \leq \left(|F|^k\right)_U$

Idea of the proof for $k = 4$.

Using $|C_L D_U| \leq |C|_L |D|_U$

$$
\begin{aligned}
|\mathcal{L}_4 + \mathcal{U}_4| &\leq |F|_L(|F|_L(|F|_L|F|_U)_U)_U + |F|_L((|F|_L|F|_U)_L|F|_U)_U \\
&\quad +(|F|_L|F|_U)_L(|F|_L|F|_U)_U + (|F|_L(|F|_L|F|_U)_U)_L|F|_U \\
&\quad +((|F|_L|F|_U)_L|F|_U)_L|F|_U \\
&\leq |F|_L|F|_L|F|_L|F|_U + |F|_L(|F|_L|F|_U)_L|F|_U \\
&\quad +|F|_L|F|_U|F|_L|F|_U + |F|_L(|F|_L|F|_U)_U|F|_U + |F|_L|F|_U|F|_U|F|_U \\
&= |F|_L|F|_L|F|_L|F|_U + |F|_L|F|_L|F|_U|F|_U \\
&\quad +|F|_L|F|_U|F|_L|F|_U + |F|_L|F|_U|F|_U|F|_U \\
&\leq (|F|_L + |F|_U)^4 \\
&= |F|^4
\end{aligned}
$$

Perturbation Theory for factorizations of LU type (VIII)

Proving strict componentwise bounds (Sun 1992) Given $A = LU$ and E such that $\|L^{-1}EU^{-1}\| < 1$, then $\widetilde{A} = A + E = \widetilde{L}\widetilde{U}$ has a unique LU factorization. Denote $F = L^{-1}EU^{-1}$.

$$\widetilde{A} = L(I+F)U = (L\mathcal{L})(\mathcal{U}U) = \underbrace{\left(L\sum_{k=0}^{\infty}\mathcal{L}_k\right)}_{\widetilde{L}}\underbrace{\left(\sum_{k=0}^{\infty}\mathcal{U}_k\right)U}_{\widetilde{U}}.$$

Then

$$|\widetilde{U} - U| = \left|\left(\sum_{k=1}^{\infty}\mathcal{U}_k\right)U\right| \leq \left(\sum_{k=1}^{\infty}|\mathcal{U}_k|\right)|U|$$

$$\leq \left(\sum_{k=1}^{\infty}\left(|F|^k\right)_U\right)|U| \leq \left(|F|(I-|F|)^{-1}\right)_U |U|.$$

Similarly

$$|\widetilde{L} - L| \leq |L|\left(|F|(I-|F|)^{-1}\right)_L.$$

Perturbation Theory for factorizations of LU type (IX)

Strict Normwise Bounds (Barrlund 1991) From the series expansion, if $\|L^{-1}EU^{-1}\| < 1$ then

$$\max\left\{\frac{\|\widetilde{L}-L\|}{\|L\|}, \frac{\|\widetilde{U}-U\|}{\|U\|}\right\} \leq \frac{\|L^{-1}EU^{-1}\|}{1-\|L^{-1}EU^{-1}\|}.$$

These bounds can sistematically overstimate the errors (Chang-Paige 1996).

First order NORMWISE better error bounds (Stewart 1997) Using again $F = L^{-1}EU^{-1}$,

$$\widetilde{L}-L = L\left(L^{-1}EU^{-1}\right)_L + \|L\|\,O(\|F\|^2) = (LS)\left((LS)^{-1}EU^{-1}\right)_L + \|L\|\,O(\|F\|^2),$$

where $\mathcal{S}$ is the set of diagonal matrices. Then,

$$\frac{\|\widetilde{L}-L\|}{\|L\|} \leq \min_{S\in\mathcal{S}} \kappa(LS)\frac{\|U^{-1}\|\|E\|}{\|L\|} + O(\|F\|^2) \leq \left(\min_{S\in\mathcal{S}} \kappa(LS)\right)\kappa(U)\frac{\|E\|}{\|A\|} + O(\|F\|^2).$$

A similar argument for U implies that

$$\frac{\|\widetilde{U}-U\|}{\|U\|} \leq \min_{S\in\mathcal{S}} \kappa(SU)\frac{\|L^{-1}\|\|E\|}{\|U\|} + O(\|F\|^2) \leq \left(\min_{S\in\mathcal{S}} \kappa(SU)\right)\kappa(L)\frac{\|E\|}{\|A\|} + O(\|F\|^2),$$

with $\kappa(B) = \|B\|\|B^{-1}\|$ is the condition number in $\|\cdot\|$.

Summary on Nonnegative Matrices

Definition: A matrix $A \in \mathbb{C}^{n\times n}$ is reducible if there exists an $n \times n$ permutation matrix P such that

$$P^T A P = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

with B and D square matrices. A matrix $A \in \mathbb{C}^{n\times n}$ is irreducible if it is not reducible.

Definition: The spectral radius of a matrix $A \in \mathbb{C}^{n\times n}$ is the real number

$$\rho(A) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\}$$

Definition: A matrix A is nonnegative (resp. positive) if $a_{ij} \geq 0$ (resp. $a_{ij} > 0$) for all i, j.

Perron's Theorems for Nonnegative Matrices

Theorem: Let $A \in \mathbb{R}^{n\times n}$ be a positive matrix. Then

1. $\rho(A)$ is an eigenvalue of A.
2. There is $x \in \mathbb{R}^n$ with $x_i > 0$ for $1 \leq i \leq n$ such that $Ax = \rho(A)x$.
3. $\rho(A)$ is an algebraically simple eigenvalue of A.
4. If $\lambda \neq \rho(A)$ is an eigenvalue of A then $|\lambda| < \rho(A)$.

Theorem: Let $A \in \mathbb{R}^{n\times n}$ be a nonnegative irreducible matrix. Then

1. $\rho(A)$ is an eigenvalue of A.
2. There is $x \in \mathbb{R}^n$ with $x_i > 0$ for $1 \leq i \leq n$ such that $Ax = \rho(A)x$.
3. $\rho(A)$ is a algebraically simple eigenvalue of A.

Theorem: Let $A \in \mathbb{R}^{n\times n}$ be a nonnegative matrix. Then

1. $\rho(A)$ is an eigenvalue of A.
2. There is $0 \neq x \in \mathbb{R}^n$ with $x_i \geq 0$ for $i = 1:n$ s. t. $Ax = \rho(A)x$.

Perturbation of the Perron Root (I)

Definition: Let $A \in \mathbb{R}^{n\times n}$ be a nonnegative matrix. Then $\rho(A)$ is called the Perron root of A.

Lemma: Let $A \in \mathbb{R}^{n\times n}$ be nonnegative and irreducible, and $E \in \mathbb{R}^{n\times n}$ be nonnegative. Then

$$\rho(A) \leq \rho(A+E).$$

Proof: Let $A(t) = A + tE$ with $0 \leq t \leq 1$. Then
1. $A(t)$ is nonnegative and irreducible,
2. there are left and right eigenventors, $y(t)$ and $x(t)$, of $\rho(A(t))$ with positive components, so

$$\rho(A(t_1)) = \rho(A(t)) + (t_1 - t)\frac{y(t)^* E x(t)}{y(t)^* x(t)} + O((t_1 - t)^2),$$

then

$$\frac{d\rho(A(t))}{dt} \geq 0, \qquad 0 \leq t \leq 1,$$

and the result follows.

Perturbation of the Perron Root (II)

Theorem (Elsner et al. 1993): Let $A \in \mathbb{R}^{n\times n}$ be a nonnegative IRREDUCIBLE matrix, and E be a real matrix such that

$$|E| \leq \epsilon A, \text{ where } \epsilon < 1.$$

Then $A + E$ is also nonnegative and IRREDUCIBLE, and

$$\frac{|\rho(A+E) - \rho(A)|}{\rho(A)} \leq \epsilon.$$

Proof:

$$a_{ij} + e_{ij} \geq (1-\epsilon)a_{ij}.$$

Thus, if $a_{ij} > 0$ then $a_{ij} + e_{ij} \geq (1-\epsilon)a_{ij} > 0$, and if $a_{ij} = 0$ then $a_{ij} + e_{ij} = (1-\epsilon)a_{ij} = 0$. As a consequence $A + E$ and $A - \epsilon A$ are nonnegative and irreducible.

From previous Lemma

$$(1-\epsilon)\rho(A) = \rho(A - \epsilon A) \leq \rho(A+E) \leq \rho(A + \epsilon A) = (1+\epsilon)\rho(A),$$

and the result follows.

Perturbation of the Perron Root (III)

Theorem: Let $A \in \mathbb{R}^{n \times n}$ be a nonnegative matrix, and E be a real matrix such that

$$|E| \leq \epsilon A, \quad \text{where } \epsilon < 1.$$

Then $A + E$ is also nonnegative, and

$$\frac{|\rho(A+E) - \rho(A)|}{\rho(A)} \leq \epsilon.$$

Proof: Obviously $A + E$ is nonnegative. Let us define A_θ to be the matrix obtained by setting the zero entries of A equal to $\theta > 0$. Therefore A_θ and $A_\theta + E$ are positive matrices, in particular they are irreducible, and $|E| \leq \epsilon A_\theta$. By using previous Theorem:

$$\frac{|\rho(A_\theta + E) - \rho(A_\theta)|}{\rho(A_\theta)} \leq \epsilon,$$

the results follows by taking the limit $\theta \rightarrow 0$, and from the continuity of eigenvalues.

Selected Bibliography

- J.W. Demmel, *“Applied Numerical Linear Algebra”,* SIAM 1997.
- J. Demmel, M. Gu, S. Eisenstat, I. Slapnicar, K. Veselic, Z. Drmac, *“Computing the singular value decomposition with high relative accuracy”,* Linear Algebra Appl., 299 (1999), 21-80.
- F.M. Dopico and J.M. Molera, *“Perturbation theory for factorizations of LU type through series expansions”,* SIAM J. Matrix Anal. Appl., to appear.
- S. Eisenstat and I.C.F. Ipsen, *“Relative perturbation techniques for singular value problems”,* SIAM J. Numer. Anal., 32 (1995), 1972-1988.
- L. Elsner, I. Koltracht, M. Neumann, and D. Xiao, *“On accurate computations of the Perron root”,* SIAM J. Matrix Anal. Appl., 14 (1993), 456-467.
- I.C.F. Ipsen, *“Relative perturbation results for matrix eigenvalues and singular values”,* Acta Numerica (1998), 151-201.
- I.C.F. Ipsen, *“An overview of relative sin Theta theorems for invariant subspaces of complex matrices”,* J. Comput. Appl. Math., 123 (2000),131-153.
- R.-C. Li, *“Relative perturbation theory: (I). Eigenvalue and Singular value variations”,* SIAM J. Matrix Anal. Appl., 19 (1998), 956-982.
- R.-C. Li, *“Relative perturbation theory: (II). Eigenspace and Singular subspace variations”,* SIAM J. Matrix Anal. Appl., 20 (1998), 471-492.