Perturbation Theory for Simultaneous Bases of Singular Subspaces

and

Numerical Applications

Froilán M. Dopico

Departamento de Matemáticas

Universidad Carlos III de Madrid

Joint work with **Julio Moro** ***from Universidad Carlos III***

Setting the Problem

Let A and $\widetilde{A}$ be complex $m \times n$ ($m \geq n$) matrices with partitioned singular value decompositions (SVD):

$$A = \begin{pmatrix} U_1 & U_2 & U_3 \end{pmatrix} \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1 & V_2 \end{pmatrix}^*,$$

$$\widetilde{A} = \begin{pmatrix} \widetilde{U}_1 & \widetilde{U}_2 & \widetilde{U}_3 \end{pmatrix} \begin{pmatrix} \widetilde{\Sigma}_1 & 0 \\ 0 & \widetilde{\Sigma}_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \widetilde{V}_1 & \widetilde{V}_2 \end{pmatrix}^*$$

- where Σ_1 and $\widetilde{\Sigma}_1$ are $k \times k$ matrices,
- $(\ )^*$ denotes the conjugate transpose matrix,
- No special order is assumed on the singular values.

How to bound the variation of left/right singular vectors ($U_1, \widetilde{U}_1$ and $V_1, \widetilde{V}_1$) using the difference $A - \widetilde{A}$

Well-known answer: $\sin\Theta$ theorems for subspaces

Let the following *residuals and gaps* be defined:

$$R = A\widetilde{V}_1 - \widetilde{U}_1\widetilde{\Sigma}_1 = (A - \widetilde{A})\widetilde{V}_1$$

$$S = A^*\widetilde{U}_1 - \widetilde{V}_1\widetilde{\Sigma}_1 = (A^* - \widetilde{A}^*)\widetilde{U}_1.$$

$$\text{gap} = \min_{\widetilde{\mu}\in\sigma(\widetilde{\Sigma}_1),\,\mu\in\sigma_{\text{ext}}(\Sigma_2)} |\widetilde{\mu} - \mu|$$

where, for any matrix B, $\sigma(B)$ denotes the set of its singular values and $\sigma_{\text{ext}}(\Sigma_2) \equiv \sigma(\Sigma_2) \cup \{0\}$ if $m > n$ and $\sigma_{\text{ext}}(\Sigma_2) \equiv \sigma(\Sigma_2)$ if $m = n$.

Theorem in Frobenius norm (Wedin, 1972):

If $\text{gap} > 0$ *then*

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

$$\leq \frac{\sqrt{\|R\|_F^2 + \|S\|_F^2}}{\text{gap}} \leq \frac{\sqrt{2}\|A - \widetilde{A}\|_F}{\text{gap}}$$

$\sin\Theta(V_1, \widetilde{V}_1)$ *are the singular values of* $\widetilde{V}_2^* V_1$.

A drawback of $\sin\Theta$ theorems

When $m > n$ the left singular subspaces may be much more sensible that the right singular subspaces: the zero appearing in $\sigma_{\text{ext}}(\Sigma_2) \equiv \sigma(\Sigma_2) \cup \{0\}$ only affects to left singular subspaces.

$$\text{gap} = \min\left\{ \min_{\widetilde{\mu}\in\sigma(\widetilde{\Sigma}_1),\mu\in\sigma(\Sigma_2)} |\widetilde{\mu} - \mu| \, , \, \sigma_{\min}(\widetilde{\Sigma}_1) \right\}$$

EXAMPLE:

$$A = \begin{pmatrix} \epsilon & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \epsilon & 0 \\ 0 & 1 \end{pmatrix}$$

$$\widetilde{A} = \begin{pmatrix} \epsilon & 0 \\ 0 & 1 \\ \epsilon & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \\ \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} \sqrt{2}\epsilon & 0 \\ 0 & 1 \end{pmatrix}$$

$\sin\Theta$ **bound:**

$$\sqrt{\|\sin\Theta(U_1, \widetilde{U}_1)\|_F^2 + \|\sin\Theta(V_1, \widetilde{V}_1)\|_F^2} \leq \frac{\sqrt{2}\epsilon}{\sqrt{2}\epsilon} = 1$$

Solving the drawback

$$A = \begin{pmatrix} \epsilon & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \epsilon & 0 \\ 0 & 1 \end{pmatrix}$$

$$\widetilde{A} = \begin{pmatrix} \epsilon & 0 \\ 0 & 1 \\ \epsilon & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \\ \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} \sqrt{2}\epsilon & 0 \\ 0 & 1 \end{pmatrix}$$

FIRST WAY: Use $\sin\Theta$ theorem on U_2 and V_2,

$$\sqrt{\|\sin\Theta(U_2, \widetilde{U}_2)\|_F^2 + \|\sin\Theta(V_2, \widetilde{V}_2)\|_F^2} \leq \frac{\sqrt{2}\epsilon}{1-\epsilon}$$

Relative $\sin\Theta$ **theorems**

Results for multiplicative perturbations are the most relevant in high relative accuracy algorithms for SVD.

These have been developed by several authors (Eisenstat, Ipsen, Li,...).

Theorem (R.C. Li 1999) *Let* A *and* $\widetilde{A} = D_1^* A D_2$ *where* D_1, D_2 *are nonsingular matrices. Define the relative gap*

$$\text{relgap} = \min_{\mu \in \sigma(\Sigma_1), \widetilde{\mu} \in \sigma_{ext}(\widetilde{\Sigma}_2)} \frac{|\mu - \widetilde{\mu}|}{\widetilde{\mu}}$$

If $\text{relgap} > 0$ *then*

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

$$\sqrt{\|(I - D_1^*)U_1\|_F^2 + \|(I - D_2^*)V_1\|_F^2}$$

$$+ \frac{\sqrt{\|(D_1^* - D_1^{-1})U_1\|_F^2 + \|(D_2^* - D_2^{-1})V_1\|_F^2}}{\text{relgap}}$$

Are these results enough?

A simple example:

$$A = \begin{pmatrix} \epsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{and} \quad \widetilde{A} = \begin{pmatrix} -\epsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

where $\epsilon > 0$ is small enough. $\widetilde{A}$ is a small additive normwise perturbation of A, but not small in a multiplicative sense. So **Wedin's theorem** applies to the singular subspaces associated with ϵ:

$$\|R\|_F = \|S\|_F = 2\epsilon \quad \text{and} \quad \text{gap} = 1 - \epsilon$$

Thus

$$\sqrt{\|\sin\Theta(U_1, \widetilde{U}_1)\|_F^2 + \|\sin\Theta(V_1, \widetilde{V}_1)\|_F^2} \leq \frac{2\sqrt{2}\epsilon}{1-\epsilon}$$

This is right!!, because the left/right singular subspaces of A and $\widetilde{A}$ are equal.

Simple example (continued)

$$A = \begin{pmatrix} \epsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{and} \quad \widetilde{A} = \begin{pmatrix} -\epsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

But the **SIMULTANEOUS** left and right singular vectors of $A, \widetilde{A}$ corresponding to ϵ are

$$u_1 = (1,0,0)^T \quad \text{and} \quad v_1 = (1,0,0)^T$$

$$\widetilde{u}_1 = (1,0,0)^T \quad \text{and} \quad \widetilde{v}_1 = (-1,0,0)^T$$

and, **NO MATTER THE SIZE OF** ϵ **THE PAIR** (u_1, v_1) **IS VERY FAR FROM** $(\widetilde{u}_1, \widetilde{v}_1)$**.**

Huge differences between $u_1^T v_1$ ($u_1 v_1^T$) and $\widetilde{u}_1^T \widetilde{v}_1$ ($\widetilde{u}_1 \widetilde{v}_1^T$).

This fact happens for any choice of simultaneous left/right singular vectors. **The** $\sin\Theta$ **theorems do not give any information about this fact.**

How to define the magnitude to be bounded

Notice that if

$$A = \begin{pmatrix} U_1 & U_2 & U_3 \end{pmatrix} \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1 & V_2 \end{pmatrix}^*,$$

$$A = \begin{pmatrix} \widehat{U}_1 & \widehat{U}_2 & \widehat{U}_3 \end{pmatrix} \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \widehat{V}_1 & \widehat{V}_2 \end{pmatrix}^*$$

are TWO SVD's OF THE SAME MATRIX A, Σ_1 **is not singular**, and $\sigma(\Sigma_1) \cap \sigma(\Sigma_2) = \phi$ then there exists a unitary $k \times k$ matrix Q such that

$$\widehat{U}_1 = U_1 Q \quad \text{and} \quad \widehat{V}_1 = V_1 Q.$$

This leads us to bound the following quantity

$$\boxed{\min_{W \text{unitary}} \sqrt{\|U_1 W - \widetilde{U}_1\|_F^2 + \|V_1 W - \widetilde{V}_1\|_F^2}}$$

Absolute perturbation theorem for bases

Theorem *Define the "new" gap*

$$\mathrm{gap}_b = \min\left\{ \mathrm{gap}\ ,\ \sigma_{\min}(\widetilde{\Sigma}_1) + \sigma_{\min}(\Sigma_1)\right\}$$

where $\sigma_{\min}(\Sigma_1)$ *and* $\sigma_{\min}(\widetilde{\Sigma}_1)$ *denote the minimum of the singular values of* Σ_1 *and* $\widetilde{\Sigma}_1$*. If* $\mathrm{gap}_b > 0$ *then*

$$\min_{W \text{unitary}} \sqrt{\|U_1 W - \widetilde{U}_1\|_F^2 + \|V_1 W - \widetilde{V}_1\|_F^2} \leq$$

$$\leq \sqrt{2}\frac{\sqrt{\|R\|_F^2 + \|S\|_F^2}}{\mathrm{gap}_b} \leq \frac{2\|A - \widetilde{A}\|_F}{\mathrm{gap}_b}$$

Moreover, the left hand side is minimized for $W = YZ^*$*, where* YSZ^* *is any SVD of* $U_1^* \widetilde{U}_1 + V_1^* \widetilde{V}_1$*, and the equality can be attained.*

Remember that:

$$\mathrm{gap} = \min_{\widetilde{\mu} \in \sigma(\widetilde{\Sigma}_1), \mu \in \sigma_{\text{ext}}(\Sigma_2)} |\widetilde{\mu} - \mu|$$

Remarks

$$\text{gap}_b = \min\left\{ \text{gap}\ ,\ \sigma_{\min}(\widetilde{\Sigma}_1) + \sigma_{\min}(\Sigma_1)\right\}$$

$$\text{gap} = \min_{\widetilde{\mu}\in\sigma(\widetilde{\Sigma}_1),\mu\in\sigma_{\text{ext}}(\Sigma_2)} |\widetilde{\mu} - \mu|$$

- If $m > n$ then $\text{gap}_b = \text{gap}$. The relevant changes appear when $m = n$.
- If $m = n$ then it may happen

$$\text{gap}_b \ll \text{gap}$$

when Σ_1 contains the smallest singular values and these are "very small".

BAD NEWS!!: simultaneous bases can be much more sensitive than singular subspaces for the smallest singular values.

Simple example revisited

$$A = \begin{pmatrix} \epsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{and} \quad \widetilde{A} = \begin{pmatrix} -\epsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\text{gap}_b = 2\epsilon$$

This implies

$$2 = \min_{W \text{unitary}} \sqrt{\|U_1 W - \widetilde{U}_1\|_F^2 + \|V_1 W - \widetilde{V}_1\|_F^2} \leq$$

$$\leq \sqrt{2} \frac{\sqrt{\|R\|_F^2 + \|S\|_F^2}}{\text{gap}_b} = 2$$

Relative perturbation theorem for bases

Theorem *Let* A *and* $\widetilde{A} = D_1^* A D_2$ *where* D_1 *and* D_2 *are nonsingular matrices. Define the "new" relative gap*

$$\mathrm{relgap}_b = \min\left\{ \mathrm{relgap}\,,\ \left(\min_{\mu\in\sigma(\Sigma_1),\widetilde{\mu}\in\sigma(\widetilde{\Sigma}_1)} \frac{\mu+\widetilde{\mu}}{\widetilde{\mu}} \right) \right\}.$$

If $\mathrm{relgap}_b > 0$ *then*

$$\min_{W\,\mathrm{unitary}} \sqrt{\|U_1 W - \widetilde{U}_1\|_F^2 + \|V_1 W - \widetilde{V}_1\|_F^2} \leq$$

$$\sqrt{2} \quad \sqrt{\|(I - D_1^*)U_1\|_F^2 + \|(I - D_2^*)V_1\|_F^2}$$

$$\left. + \frac{\sqrt{\|(D_1^* - D_1^{-1})U_1\|_F^2 + \|(D_2^* - D_2^{-1})V_1\|_F^2}}{\mathrm{relgap}_b} \right).$$

Moreover the left hand sides is minimized for $W = YZ^*$*, where* YSZ^* *is any SVD of* $U_1^* \widetilde{U}_1 + V_1^* \widetilde{V}_1$.

Remarks and conclusions

$$\mathrm{relgap}_b = \min \left\{ \mathrm{relgap}\,, \quad \min_{\mu\in\sigma(\Sigma_1),\widetilde{\mu}\in\sigma(\widetilde{\Sigma}_1)} \frac{\mu+\widetilde{\mu}}{\widetilde{\mu}} \right) \right\}.$$

$$\frac{\mu+\widetilde{\mu}}{\widetilde{\mu}} \geq 1$$

GOOD NEWS!!

for small multiplicative perturbations the sensitivity of simultaneous bases is similar to that of singular subspaces.

Important changes with respect to singular subspaces sensitivity results for arbitrary additive perturbations, BUT NOT FOR MULTIPLICATIVE PERTURBATIONS.

Sketch of the proofs

STEP 1: Let us define

$$X_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} U_1 \\ V_1 \end{pmatrix} \quad \text{and} \quad \widetilde{X}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} \widetilde{U}_1 \\ \widetilde{V}_1 \end{pmatrix},$$

$$\min_{W \text{unitary}} \sqrt{\|U_1 W - \widetilde{U}_1\|_F^2 + \|V_1 W - \widetilde{V}_1\|_F^2} =$$

$$= \sqrt{2} \min_{W \text{unitary}} \|X_1 W - \widetilde{X}_1\|_F$$

$$= \sqrt{2}\sqrt{\|I - \cos\Theta(X_1, \widetilde{X}_1)\|_F^2 + \|\sin\Theta(X_1, \widetilde{X}_1)\|_F^2}$$

$$\leq 2\|\sin\Theta(X_1, \widetilde{X}_1)\|_F.$$

STEP 2: Apply absolute/relative $\sin\Theta$ theorems to bound the sines of the canonical angles between

$$\text{col}(X_1) \quad \text{and} \quad \text{col}(\widetilde{X}_1)$$

which are invariant subspaces of the Jordan-Wielandt matrices

$$\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & \widetilde{A} \\ \widetilde{A}^* & 0 \end{pmatrix}$$

Numerical Experiments (I)

- MATLAB 5.3
- $A = S * Q$ real 8×8 matrix.
- Q random finite precision orthogonal 8×8 matrix.
- $S = \mathrm{diag}([\mathrm{ones}(1,6), 10^{-j}, 10^{-j}])$.
- $j = 1 : 30$.
- 20 matrices for each j.
- Well-determined TWO-dim. singular subspaces associated with the two smallest singular values.
- Compute $\widetilde{U}\widetilde{\Sigma}\widetilde{V}^T$ with
  1. MATLAB: $A + E = \widetilde{U}\widetilde{\Sigma}\widetilde{V}^T$.
  2. Right-Jacobi: $A(I + F) = \widetilde{U}\widetilde{\Sigma}\widetilde{V}^T$.
- Plot

$$\min_{W \text{unitary}} \sqrt{\|U_1 W - \widetilde{U}_1\|_F^2 + \|V_1 W - \widetilde{V}_1\|_F^2} \quad \text{vs.} \quad j$$

  for singular vectors associated with the smallest singular values.

Numerical Experiments (II)

- Compare REAL and COMPLEX matrices which do not deserve high relative accuracy. Thus only MATLAB is used.
- $R = U * S * V'$ REAL 50×50 matrix.
- $C = U_c * S * V_c'$ COMPLEX 50×50 matrix.
- $S = \mathrm{diag}([3 * \mathrm{rand}(1, 49) + 1, 10^{-j}])$.
- U, V random finite precision REAL orthogonal.
- U_c, V_c random finite precision COMPLEX unitary.
- $j = 1 : 0.5 : 25$.
- 4 real matrices and 4 complex matrices for each j.
- Well-determined ONE-dim. singular subspaces associated with the least singular value.
- Plot

$$\min_{W \,\mathrm{unitary}} \sqrt{\|U_1 W - \widetilde{U}_1\|_F^2 + \|V_1 W - \widetilde{V}_1\|_F^2} \quad \text{vs.} \quad j$$

for singular vectors associated with the least singular value in REAL and COMPLEX case.

Absolute perturbation theorem in real case

Theorem *Let A and $\widetilde{A}$ be REAL and* $\boxed{m = n}$.
Define

$$\sigma(\Sigma_1) = \{\sigma_1 \geq \ldots \geq \sigma_k\}$$

$$\mathrm{Rgap}_b = \min \left\{ \min_{\lambda \in \sigma(\Sigma_1), \mu \in \sigma(\Sigma_2)} |\lambda - \mu| \; , \; \sigma_k + \sigma_{k-1} \right\}$$

IF

$$\|A - \widetilde{A}\|_2 < \frac{1}{2} \min \left\{ \min_{\lambda \in \sigma(\Sigma_1), \mu \in \sigma(\Sigma_2)} |\lambda - \mu| \; , \; 2\sigma_k \right\}$$

THEN

$$\min_{W \text{orthogonal}} \sqrt{\|U_1 W - \widetilde{U}_1\|_F^2 + \|V_1 W - \widetilde{V}_1\|_F^2} \leq$$

$$\leq -\frac{\|A - \widetilde{A}\|_F}{\|A - \widetilde{A}\|_2} \ln \left(1 - \frac{2\|A - \widetilde{A}\|_2}{\mathrm{Rgap}_b} \right)$$

$$= \frac{2\|A - \widetilde{A}\|_F}{\mathrm{Rgap}_b} + O(\|A - \widetilde{A}\|_2^2)$$

One-dimensional real case

Moreover, if Σ_1 *is* 1×1 *then*

$$\mathrm{Rgap}_b = \min_{\mu \in \sigma(\Sigma_2)} |\Sigma_1 - \mu|$$

and

$$\min_{w \in \{-1,1\}} \sqrt{\|U_1 w - \widetilde{U}_1\|_2^2 + \|V_1 w - \widetilde{V}_1\|_2^2} \leq$$

$$\leq -\ln\left(1 - \frac{2\|A - \widetilde{A}\|_2}{\mathrm{Rgap}_b}\right)$$

Remarks:

- For numerical algorithms only the one dimensional case is important. If $k > 1$

  $$2\sigma_k \approx \sigma_k + \sigma_{k-1}$$

  if Σ_1 is a "cluster" of singular values.

- The experiments show that the restriction on the size of the perturbation cannot be removed.

- Similar to perturbations results for the unitary polar factor (Kenney, Laub, Barrlund, Mathias, Li...).

Ideas of the proof

1. Derivatives of orthogonal projectors on invariant subspaces of Jordan-Wielandt matrices are used.
2. For the one-dimensional case: Consider the Hermitian matrices $B + tF$, t real parameter, and let $x_1(t), ..., x_n(t)$ be their orthonormal basis of eigenvectors.

$$x_k(t) = x_k + \sum_{i \neq k}^{n} \frac{x_i^* F x_k}{\lambda_k - \lambda_i} x_i + O(t^2)$$

3. **Fundamental Fact:**

$$[u_1^*\,, -v_1^*] \begin{bmatrix} 0 & E \\ E^* & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ v_1 \end{bmatrix} \equiv 0$$

   IN THE REAL CASE, BUT NOT IN THE COMPLEX CASE.

and more...

Other results have been obtained:

- Bounds in arbitrary unitarily invariant norms in the absolute setting.
- Bounds with other relgaps in the relative setting.
- There are also changes in the case $m > n$.

Future work

- Relative results for ADDITIVE perturbations.
- Relative results for arbitrary unitarily invariant norms.