15

Matrix Perturbation Theory

Ren-Cang Li
University of Texas at Arlington

There is a vast amount of material in matrix (operator) perturbation theory. Related books that are worth mentioning are [SS90], [Par98], [Bha96], [Bau85], and [Kat70]. In this chapter, we attempt to include the most fundamental results up to date, except those for linear systems and least squares problems for which the reader is referred to Section 38.1 and Section 39.6.

Throughout this chapter, $\|\cdot\|_{\mathrm{UI}}$ denotes a general unitarily invariant norm. Two commonly used ones are the spectral norm $\|\cdot\|_2$ and the Frobenius norm $\|\cdot\|_{\mathrm{F}}$.

15.1 Eigenvalue Problems

The reader is referred to Sections 4.3, 14.1, and 14.2 for more information on eigenvalues and their locations.

Definitions:

Let $A \in \mathbb{C}^{n\times n}$. A scalar–vector pair $(\lambda, \mathbf{x}) \in \mathbb{C} \times \mathbb{C}^n$ is an **eigenpair** of A if $\mathbf{x} \neq 0$ and $A\mathbf{x} = \lambda\mathbf{x}$. A vector–scalar–vector triplet $(\mathbf{y}, \lambda, \mathbf{x}) \in \mathbb{C}^n \times \mathbb{C} \times \mathbb{C}^n$ is an **eigentriplet** if $\mathbf{x} \neq 0$, $\mathbf{y} \neq 0$, and $A\mathbf{x} = \lambda\mathbf{x}$, $\mathbf{y}^*A = \lambda\mathbf{y}^*$. The quantity

$$\mathrm{cond}(\lambda) = \frac{\|\mathbf{x}\|_2\|\mathbf{y}\|_2}{|\mathbf{y}^*\mathbf{x}|}$$

is the **individual condition number** for λ, where $(\mathbf{y}, \lambda, \mathbf{x}) \in \mathbb{C}^n \times \mathbb{C} \times \mathbb{C}^n$ is an eigentriplet.

Let $\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, the multiset of A's eigenvalues, and set

$$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n), \quad \Lambda_\tau = \mathrm{diag}(\lambda_{\tau(1)}, \lambda_{\tau(2)}, \ldots, \lambda_{\tau(n)}),$$

where τ is a **permutation** of $\{1, 2, \ldots, n\}$. For real Λ, i.e., all λ_j's are real,

$$\Lambda^{\uparrow} = \operatorname{diag}(\lambda_1^{\uparrow}, \lambda_2^{\uparrow}, \ldots, \lambda_n^{\uparrow}).$$

$\Lambda^{\uparrow}$ is in fact a Λ_τ for which the permutation τ makes $\lambda_{\tau(j)} = \lambda_j^{\uparrow}$ for all j.

Given two square matrices A_1 and A_2, the **separation** $\operatorname{sep}(A_1, A_2)$ between A_1 and A_2 is defined as [SS90, p. 231]

$$\operatorname{sep}(A_1, A_2) = \inf_{\|X\|_2 = 1} \|XA_1 - A_2 X\|_2.$$

A is perturbed to $\widetilde{A} = A + \Delta A$. *The same notation is adopted for* $\widetilde{A}$, *except all symbols with tildes.*

Let $X, Y \in \mathbb{C}^{n\times k}$ with $\operatorname{rank}(X) = \operatorname{rank}(Y) = k$. The **canonical angles** between their column spaces are $\theta_i = \arccos \sigma_i$, where $\{\sigma_i\}_{i=1}^k$ are the singular values of $(Y^*Y)^{-1/2} Y^* X (X^*X)^{-1/2}$. Define the **canonical angle matrix** between X and Y as

$$\Theta(X, Y) = \operatorname{diag}(\theta_1, \theta_2, \ldots, \theta_k).$$

For $k = 1$, i.e., $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ (both nonzero), we use $\angle(\mathbf{x}, \mathbf{y})$, instead, to denote the canonical angle between the two vectors.

Facts:

1. [SS90, p. 168] (**Elsner**) $\max_i \min_j |\widetilde{\lambda}_i - \lambda_j| \le (\|A\|_2 + \|\widetilde{A}\|_2)^{1-1/n} \|\Delta A\|_2^{1/n}$.
2. [SS90, p. 170] (**Elsner**) There exists a permutation τ of $\{1, 2, \ldots, n\}$ such that

$$\|\Lambda - \widetilde{\Lambda}_\tau\|_2 \le 2 \left\lfloor \frac{n}{2} \right\rfloor (\|A\|_2 + \|\widetilde{A}\|_2)^{1-1/n} \|\Delta A\|_2^{1/n}.$$

3. [SS90, p. 183] Let $(\mathbf{y}, \mu, \mathbf{x})$ be an eigentriplet of A. ΔA changes μ to $\mu + \Delta\mu$ with

$$\Delta\mu = \frac{\mathbf{y}^*(\Delta A)\mathbf{x}}{\mathbf{y}^*\mathbf{x}} + O(\|\Delta A\|_2^2),$$

and $|\Delta\mu| \le \operatorname{cond}(\mu)\|\Delta A\|_2 + O(\|\Delta A\|_2^2)$.

4. [SS90, p. 205] If A and $A + \Delta A$ are Hermitian, then

$$\|\Lambda^{\uparrow} - \widetilde{\Lambda}^{\uparrow}\|_{\text{UI}} \le \|\Delta A\|_{\text{UI}}.$$

5. [Bha96, p. 165] (**Hoffman–Wielandt**) If A and $A + \Delta A$ are normal, then there exists a permutation τ of $\{1, 2, \ldots, n\}$ such that $\|\Lambda - \widetilde{\Lambda}_\tau\|_{\text{F}} \le \|\Delta A\|_{\text{F}}$.
6. [Sun96] If A is normal, then there exists a permutation τ of $\{1, 2, \ldots, n\}$ such that $\|\Lambda - \widetilde{\Lambda}_\tau\|_{\text{F}} \le \sqrt{n}\|\Delta A\|_{\text{F}}$.
7. [SS90, p. 192] (**Bauer–Fike**) If A is diagonalizable and $A = X\Lambda X^{-1}$ is its eigendecomposition, then

$$\max_i \min_j |\widetilde{\lambda}_i - \lambda_j| \le \|X^{-1}(\Delta A)X\|_p \le \kappa_p(X)\|\Delta A\|_p.$$

8. [BKL97] Suppose both A and $\widetilde{A}$ are diagonalizable and have eigendecompositions $A = X\Lambda X^{-1}$ and $\widetilde{A} = \widetilde{X}\widetilde{\Lambda}\widetilde{X}^{-1}$.

   (a) There exists a permutation τ of $\{1, 2, \ldots, n\}$ such that

   $\|\Lambda - \widetilde{\Lambda}_\tau\|_{\text{F}} \le \sqrt{\kappa_2(X)\kappa_2(\widetilde{X})}\,\|\Delta A\|_{\text{F}}$.

   (b) $\|\Lambda^{\uparrow} - \widetilde{\Lambda}^{\uparrow}\|_{\text{UI}} \le \sqrt{\kappa_2(X)\kappa_2(\widetilde{X})}\|\Delta A\|_{\text{UI}}$ for real Λ and $\widetilde{\Lambda}$.

9. [KPJ82] Let residuals $\mathbf{r} = A\widetilde{\mathbf{x}} - \widetilde{\mu}\widetilde{\mathbf{x}}$ and $\mathbf{s}^* = \widetilde{\mathbf{y}}^* A - \widetilde{\mu}\widetilde{\mathbf{y}}^*$, where $\|\widetilde{\mathbf{x}}\|_2 = \|\widetilde{\mathbf{y}}\|_2 = 1$, and let $\varepsilon = \max\{\|\mathbf{r}\|_2, \|\mathbf{s}\|_2\}$. The smallest error matrix ΔA in the 2-norm, for which $(\widetilde{\mathbf{y}}, \widetilde{\mu}, \widetilde{\mathbf{x}})$ is an exact eigentriplet of $\widetilde{A} = A + \Delta A$, satisfies $\|\Delta A\|_2 = \varepsilon$, and $|\widetilde{\mu} - \mu| \leq \mathrm{cond}(\widetilde{\mu})\,\varepsilon + O(\varepsilon^2)$ for some $\mu \in \sigma(A)$.
10. [KPJ82], [DK70],[Par98, pp. 73, 244] Suppose A is Hermitian, and let residual $\mathbf{r} = A\widetilde{\mathbf{x}} - \widetilde{\mu}\widetilde{\mathbf{x}}$ and $\|\widetilde{\mathbf{x}}\|_2 = 1$.

   (a) The smallest Hermitian error matrix ΔA (in the 2-norm), for which $(\widetilde{\mu}, \widetilde{\mathbf{x}})$ is an exact eigenpair of $\widetilde{A} = A + \Delta A$, satisfies $\|\Delta A\|_2 = \|\mathbf{r}\|_2$.

   (b) $|\widetilde{\mu} - \mu| \leq \|\mathbf{r}\|_2$ for some eigenvalue μ of A.

   (c) Let μ be the closest eigenvalue in $\sigma(A)$ to $\widetilde{\mu}$ and $\mathbf{x}$ be its associated eigenvector with $\|\mathbf{x}\|_2 = 1$, and let $\eta = \min_{\mu \neq \lambda \in \sigma(A)} |\widetilde{\mu} - \lambda|$. If $\eta > 0$, then

$$|\widetilde{\mu} - \mu| \leq \frac{\|\mathbf{r}\|_2^2}{\eta}, \quad \sin\angle(\widetilde{\mathbf{x}}, \mathbf{x}) \leq \frac{\|\mathbf{r}\|_2}{\eta}.$$

11. Let A be Hermitian, $X \in \mathbb{C}^{n\times k}$ have full column rank, and $M \in \mathbb{C}^{k\times k}$ be Hermitian having eigenvalues $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$. Set

$$R = AX - XM.$$

   There exist k eigenvalues $\lambda_{i_1} \leq \lambda_{i_2} \leq \cdots \leq \lambda_{i_k}$ of A such that the following inequalities hold. Note that subset $\{\lambda_{i_j}\}_{j=1}^k$ may be different at different occurrences.

   (a) [Par98, pp. 253–260], [SS90, Remark 4.16, p. 207] (**Kahan–Cao–Xie–Li**)

$$\max_{1\leq j\leq k} |\mu_j - \lambda_{i_j}| \leq \frac{\|R\|_2}{\sigma_{\min}(X)},$$
$$\sqrt{\sum_{j=1}^{k} (\mu_j - \lambda_{i_j})^2} \leq \frac{\|R\|_F}{\sigma_{\min}(X)}.$$

   (b) [SS90, pp. 254–257], [Sun91] If $X^*X = I$ and $M = X^*AX$, and if all but k of A's eigenvalues differ from every one of M's by at least $\eta > 0$ and $\varepsilon_F = \|R\|_F/\eta < 1$, then

$$\sqrt{\sum_{j=1}^{k} (\mu_k - \lambda_{i_j})^2} \leq \frac{\|R\|_F^2}{\eta\sqrt{1 - \varepsilon_F^2}}.$$

   (c) [SS90, pp. 254–257], [Sun91] If $X^*X = I$ and $M = X^*AX$, and there is a number $\eta > 0$ such that either all but k of A's eigenvalues lie outside the open interval $(\mu_1 - \eta, \mu_k + \eta)$ or all but k of A's eigenvalues lie inside the closed interval $[\mu_\ell + \eta, \mu_{\ell+1} - \eta]$ for some $1 \leq \ell \leq k - 1$, and $\varepsilon = \|R\|_2/\eta < 1$, then

$$\max_{1\leq j\leq k} |\mu_j - \lambda_{i_j}| \leq \frac{\|R\|_2^2}{\eta\sqrt{1 - \varepsilon^2}}.$$

12. [DK70] Let A be Hermitian and have decomposition

$$\begin{bmatrix} X_1^* \\ X_2^* \end{bmatrix} A [X_1 \ X_2] = \begin{bmatrix} A_1 & \\ & A_2 \end{bmatrix},$$

where $[X_1\ \ X_2]$ is unitary and $X_1 \in \mathbb{C}^{n\times k}$. Let $Q \in \mathbb{C}^{n\times k}$ have orthonormal columns and for a $k \times k$ Hermitian matrix M set

$$R = AQ - QM.$$

Let $\eta = \min |\mu - \nu|$ over all $\mu \in \sigma(M)$ and $\nu \in \sigma(A_2)$. If $\eta > 0$, then $\| \sin \Theta(X_1, Q)\|_{\mathrm{F}} \leq \dfrac{\|R\|_{\mathrm{F}}}{\eta}$.

13. [LL05] Let

$$A = \begin{bmatrix} M & E^* \\ E & H \end{bmatrix}, \quad \widetilde{A} = \begin{bmatrix} M & 0 \\ 0 & H \end{bmatrix}$$

be Hermitian, and set $\eta = \min |\mu - \nu|$ over all $\mu \in \sigma(M)$ and $\nu \in \sigma(H)$. Then

$$\max_{1\leq j\leq n} |\lambda_j^{\uparrow} - \widetilde{\lambda}_j^{\uparrow}| \leq \frac{2\|E\|_2^2}{\eta + \sqrt{\eta^2 + 4\|E\|_2^2}}.$$

14. [SS90, p. 230] Let $[X_1\ \ Y_2]$ be unitary and $X_1 \in \mathbb{C}^{n\times k}$, and let

$$\begin{bmatrix} X_1^* \\ Y_2^* \end{bmatrix} A[X_1\ \ Y_2] = \begin{bmatrix} A_1 & G \\ E & A_2 \end{bmatrix}.$$

Assume that $\sigma(A_1) \bigcap \sigma(A_2) = \emptyset$, and set $\eta = \mathrm{sep}(A_1, A_2)$. If $\|G\|_2\|E\|_2 < \eta^2/4$, then there is a unique $W \in \mathbb{C}^{(n-k)\times k}$, satisfying $\|W\|_2 \leq 2\|E\|_2/\eta$, such that $[\widetilde{X}_1\ \ \widetilde{Y}_2]$ is unitary and

$$\begin{bmatrix} \widetilde{X}_1^* \\ \widetilde{Y}_2^* \end{bmatrix} A[\widetilde{X}_1\ \ \widetilde{Y}_2] = \begin{bmatrix} \widetilde{A}_1 & \widetilde{G} \\ 0 & \widetilde{A}_2 \end{bmatrix},$$

where

$$\begin{aligned}
\widetilde{X}_1 &= (X_1 + Y_2 W)(I + W^*W)^{-1/2}, \\
\widetilde{Y}_2 &= (Y_2 - X_1 W^*)(I + WW^*)^{-1/2}, \\
\widetilde{A}_1 &= (I + W^*W)^{1/2}(A_1 + GW)(I + W^*W)^{-1/2}, \\
\widetilde{A}_2 &= (I + WW^*)^{-1/2}(A_2 - WG)(I + WW^*)^{1/2}.
\end{aligned}$$

Thus, $\| \tan \Theta(X_1, \widetilde{X}_1)\|_2 < \dfrac{2\|E\|_2}{\eta}$.

Examples:

1. Bounds on $\|\Lambda - \widetilde{\Lambda}_\tau\|_{\mathrm{UI}}$ are, in fact, bounds on $\lambda_j - \lambda_{\tau(j)}$ in disguise, only more convenient and concise. For example, for $\|\cdot\|_{\mathrm{UI}} = \|\cdot\|_2$ (spectral norm), $\|\Lambda - \widetilde{\Lambda}_\tau\|_2 = \max_j |\lambda_j - \lambda_{\tau(j)}|$, and for $\|\cdot\|_{\mathrm{UI}} = \|\cdot\|_{\mathrm{F}}$ (Frobenius norm), $\|\Lambda - \widetilde{\Lambda}_\tau\|_{\mathrm{F}} = \left[\sum_{j=1}^n |\lambda_j - \lambda_{\tau(j)}|^2\right]^{1/2}$.
2. Let $A,\ \widetilde{A} \in \mathbb{C}^{n\times n}$ as follows, where $\varepsilon > 0$.

$$A = \begin{bmatrix} \mu & 1 & & \\ & \mu & \ddots & \\ & & \ddots & 1 \\ & & & \mu \end{bmatrix}, \quad \widetilde{A} = \begin{bmatrix} \mu & 1 & & \\ & \mu & \ddots & \\ & & \ddots & 1 \\ \varepsilon & & & \mu \end{bmatrix}.$$

It can be seen that $\sigma(A) = \{\mu, \ldots, \mu\}$ (repeated n times) and the characteristic polynomial $\det(tI - \widetilde{A}) = (t - \mu)^n - \varepsilon$, which gives $\sigma(\widetilde{A}) = \{\mu + \varepsilon^{1/n} e^{2ij\pi/n}, 0 \leq j \leq n-1\}$. Thus,

$|\widetilde{\lambda} - \mu| = \varepsilon^{1/n} = \|\Delta A\|_2^{1/n}$. This shows that the fractional power $\|\Delta A\|_2^{1/n}$ in Facts 1 and 2 cannot be removed in general.

3. Consider

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 4.001 \end{bmatrix} \quad \text{is perturbed by} \quad \Delta A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.001 & 0 & 0 \end{bmatrix}.$$

A's eigenvalues are easily read off, and

$$\lambda_1 = 1,\ \mathbf{x}_1 = [1, 0, 0]^T,\ \mathbf{y}_1 = [0.8285, -0.5523, 0.0920]^T,$$

$$\lambda_2 = 4,\ \mathbf{x}_2 = [0.5547, 0.8321, 0]^T,\ \mathbf{y}_2 = [0, 0.0002, -1.0000]^T,$$

$$\lambda_3 = 4.001,\ \mathbf{x}_3 = [0.5547, 0.8321, 0.0002]^T,\ \mathbf{y}_3 = [0, 0, 1]^T.$$

On the other hand, $\widetilde{A}$'s eigenvalues computed by MATLAB's `eig` are $\widetilde{\lambda}_1 = 1.0001$, $\widetilde{\lambda}_2 = 3.9427$, $\widetilde{\lambda}_3 = 4.0582$. The following table gives $|\widetilde{\lambda}_j - \lambda_j|$ with upper bounds up to the 1st order by Fact 3.

j	$\text{cond}(\lambda_j)$	$\text{cond}(\lambda_j)\|\Delta A\|_2$	$|\widetilde{\lambda}_j - \lambda_j|$
1	1.2070	0.0012	0.0001
2	$6.0 \cdot 10^3$	6.0	0.057
3	$6.0 \cdot 10^3$	6.0	0.057

We see that $\text{cond}(\lambda_j)\|\Delta A\|_2$ gives a fairly good error bound for $j = 1$, but dramatically worse for $j = 2, 3$. There are two reasons for this: One is in the choice of ΔA and the other is that ΔA's order of magnitude is too big for the first order bound $\text{cond}(\lambda_j)\|\Delta A\|_2$ to be effective for $j = 2, 3$. Note that ΔA has the same order of magnitude as the difference between λ_2 and λ_3 and that is too big usually. For better understanding of this first order error bound, the reader may play with this example with $\Delta A = \varepsilon \frac{\mathbf{y}_j \mathbf{x}_j^*}{\|\mathbf{y}_j\|_2 \|\mathbf{x}_j^*\|_2}$ for various tiny parameters ε.

4. Let $\Sigma = \text{diag}(c_1, c_2, \dots, c_k)$ and $\Gamma = \text{diag}(s_1, s_2, \dots, s_k)$, where $c_j, s_j \geq 0$ and $c_j^2 + s_j^2 = 1$ for all j. The canonical angles between

$$X = Q \begin{bmatrix} I_k \\ 0 \\ 0 \end{bmatrix} V^*, \quad Y = Q \begin{bmatrix} \Sigma \\ \Gamma \\ 0 \end{bmatrix} U^*$$

are $\theta_j = \arccos c_j$, $j = 1, 2, \dots, k$, where Q, U, V are unitary. On the other hand, every pair of $X, Y \in \mathbb{C}^{n \times k}$ with $2k \leq n$ and $X^*X = Y^*Y = I_k$, having canonical angles $\arccos c_j$, can be represented this way [SS90, p. 40].

5. Fact 13 is most useful when $\|E\|_2$ is tiny and the computation of A's eigenvalues is then decoupled into two smaller ones. In eigenvalue computations, we often seek unitary $[X_1\ X_2]$ such that

$$\begin{bmatrix} X_1^* \\ X_2^* \end{bmatrix} A [X_1\ X_2] = \begin{bmatrix} M & E^* \\ E & H \end{bmatrix}, \quad \begin{bmatrix} X_1^* \\ X_2^* \end{bmatrix} \widetilde{A} [X_1\ X_2] = \begin{bmatrix} M & 0 \\ 0 & H \end{bmatrix},$$

and $\|E\|_2$ is tiny. Since a unitarily similarity transformation does not alter eigenvalues, Fact 13 still applies.

6. [LL05] Consider the 2×2 Hermitian matrix

$$A = \begin{bmatrix} \alpha & \varepsilon \\ \varepsilon & \beta \end{bmatrix},$$

where $\alpha > \beta$ and $\varepsilon > 0$. It has two eigenvalues

$$\lambda_\pm = \frac{\alpha + \beta \pm \sqrt{(\alpha - \beta)^2 + 4\varepsilon^2}}{2},$$

and

$$0 < \begin{Bmatrix} \lambda_+ - \alpha \\ \beta - \lambda_- \end{Bmatrix} = \frac{2\varepsilon^2}{(\alpha - \beta) + \sqrt{(\alpha - \beta)^2 + 4\varepsilon^2}}.$$

The inequalities in Fact 13 become equalities for the example.

15.2 Singular Value Problems

The reader is referred to Section 5.6, Chapters 17 and 45 for more information on singular value decompositions.

Definitions:

$B \in \mathbb{C}^{m\times n}$ has a (first standard form) SVD $B = U\Sigma V^*$, where $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ are unitary, and $\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots) \in \mathbb{R}^{m\times n}$ is leading diagonal (σ_j starts in the top-left corner) with all $\sigma_j \geq 0$.

Let $\operatorname{SV}(B) = \{\sigma_1, \sigma_2, \ldots, \sigma_{\min\{m,n\}}\}$, the set of B's singular values, and $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$, and let $\operatorname{SV}_{\text{ext}}(B) = \operatorname{SV}(B)$ unless $m > n$ for which $\operatorname{SV}_{\text{ext}}(B) = \operatorname{SV}(B) \bigcup \{0, \ldots, 0\}$ (additional $m - n$ zeros).

A vector–scalar–vector triplet $(\mathbf{u}, \sigma, \mathbf{v}) \in \mathbb{C}^m \times \mathbb{R} \times \mathbb{C}^n$ is a **singular-triplet** if $\mathbf{u} \neq 0, \mathbf{v} \neq 0, \sigma \geq 0$, and $B\mathbf{v} = \sigma\mathbf{u}$, $B^*\mathbf{u} = \sigma\mathbf{v}$.

B is perturbed to $\widetilde{B} = B + \Delta B$. *The same notation is adopted for $\widetilde{B}$, except all symbols with tildes.*

Facts:

1. [SS90, p. 204] (**Mirsky**) $\|\Sigma - \widetilde{\Sigma}\|_{\text{UI}} \leq \|\Delta B\|_{\text{UI}}$.
2. Let residuals $\mathbf{r} = B\widetilde{\mathbf{v}} - \widetilde{\mu}\widetilde{\mathbf{u}}$ and $\mathbf{s} = B^*\widetilde{\mathbf{u}} - \widetilde{\mu}\widetilde{\mathbf{v}}$, and $\|\widetilde{\mathbf{v}}\|_2 = \|\widetilde{\mathbf{u}}\|_2 = 1$.
   (a) [Sun98] The smallest error matrix ΔB (in the 2-norm), for which $(\widetilde{\mathbf{u}}, \widetilde{\mu}, \widetilde{\mathbf{v}})$ is an exact singular-triplet of $\widetilde{B} = B + \Delta B$, satisfies $\|\Delta B\|_2 = \varepsilon$, where $\varepsilon = \max\{\|\mathbf{r}\|_2, \|\mathbf{s}\|_2\}$.
   (b) $|\widetilde{\mu} - \mu| \leq \varepsilon$ for some singular value μ of B.
   (c) Let μ be the closest singular value in $\operatorname{SV}_{\text{ext}}(B)$ to $\widetilde{\mu}$ and $(\mathbf{u}, \sigma, \mathbf{v})$ be the associated singular-triplet with $\|\mathbf{u}\|_2 = \|\mathbf{v}\|_2 = 1$, and let $\eta = \min|\widetilde{\mu} - \sigma|$ over all $\sigma \in \operatorname{SV}_{\text{ext}}(B)$ and $\sigma \neq \mu$. If $\eta > 0$, then $|\widetilde{\mu} - \mu| \leq \varepsilon^2/\eta$, and [SS90, p. 260]

$$\sqrt{\sin^2 \angle(\widetilde{\mathbf{u}}, \mathbf{u}) + \sin^2 \angle(\widetilde{\mathbf{v}}, \mathbf{v})} \leq \frac{\sqrt{\|\mathbf{r}\|_2^2 + \|\mathbf{s}\|_2^2}}{\eta}.$$

3. [LL05] Let

$$B = \begin{bmatrix} B_1 & F \\ E & B_2 \end{bmatrix} \in \mathbb{C}^{m\times n}, \quad \widetilde{B} = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix},$$

where $B_1 \in \mathbb{C}^{k\times k}$, and set $\eta = \min|\mu - \nu|$ over all $\mu \in \operatorname{SV}(B_1)$ and $\nu \in \operatorname{SV}_{\text{ext}}(B_2)$, and $\varepsilon = \max\{\|E\|_2, \|F\|_2\}$. Then

$$\max_j |\sigma_j - \widetilde{\sigma}_j| \leq \frac{2\varepsilon^2}{\eta + \sqrt{\eta^2 + 4\varepsilon^2}}.$$

4. [SS90, p. 260] (**Wedin**) Let $B,\ \widetilde{B} \in \mathbb{C}^{m\times n}$ $(m \geq n)$ have decompositions

$$\begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix} B[V_1\ \ V_2] = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}, \quad \begin{bmatrix} \widetilde{U}_1^* \\ \widetilde{U}_2^* \end{bmatrix} \widetilde{B}[\widetilde{V}_1\ \ \widetilde{V}_2] = \begin{bmatrix} \widetilde{B}_1 & 0 \\ 0 & \widetilde{B}_2 \end{bmatrix},$$

where $[U_1\ \ U_2]$, $[V_1\ \ V_2]$, $[\widetilde{U}_1\ \ \widetilde{U}_2]$, and $[\widetilde{V}_1\ \ \widetilde{V}_2]$ are unitary, and $U_1,\ \widetilde{U}_1 \in \mathbb{C}^{m\times k}$, $V_1,\ \widetilde{V}_1 \in \mathbb{C}^{n\times k}$. Set

$$R = B\widetilde{V}_1 - \widetilde{U}_1\widetilde{B}_1, \quad S = B^*\widetilde{U}_1 - \widetilde{V}_1\widetilde{B}_1.$$

If $\text{SV}(\widetilde{B}_1) \bigcap \text{SV}_{\text{ext}}(B_2) = \emptyset$, then

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

where $\eta = \min|\widetilde{\mu} - \nu|$ over all $\widetilde{\mu} \in \text{SV}(\widetilde{B}_1)$ and $\nu \in \text{SV}_{\text{ext}}(B_2)$.

Examples:

1. Let

$$B = \begin{bmatrix} 3\cdot 10^{-3} & 1 \\ 2 & 4\cdot 10^{-3} \end{bmatrix}, \quad \widetilde{B} = \begin{bmatrix} & 1 \\ 2 & \end{bmatrix} = [\mathbf{e}_2\ \mathbf{e}_1] \begin{bmatrix} 2 & \\ & 1 \end{bmatrix} \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \end{bmatrix}.$$

Then $\sigma_1 = 2.000012$, $\sigma_2 = 0.999988$, and $\widetilde{\sigma}_1 = 2$, $\widetilde{\sigma}_2 = 1$. Fact 1 gives

$$\max_{1\leq j\leq 2}|\sigma_j - \widetilde{\sigma}_j| \leq 4\cdot 10^{-3}, \quad \sqrt{\sum_{j=1}^{2}|\sigma_j - \widetilde{\sigma}_j|^2} \leq 5\cdot 10^{-3}.$$

2. Let B be as in the previous example, and let $\widetilde{\mathbf{v}} = \mathbf{e}_1$, $\widetilde{\mathbf{u}} = \mathbf{e}_2$, $\widetilde{\mu} = 2$. Then $\mathbf{r} = B\widetilde{\mathbf{v}} - \widetilde{\mu}\widetilde{\mathbf{u}} = 3\cdot 10^{-3}\mathbf{e}_1$ and $\mathbf{s} = B^*\widetilde{\mathbf{u}} - \widetilde{\mu}\widetilde{\mathbf{v}} = 4\cdot 10^{-3}\mathbf{e}_2$. Fact 2 applies. Note that, without calculating $\text{SV}(B)$, one may bound η needed for Fact 2(c) from below as follows. Since B has two singular values that are near 1 and $\widetilde{\mu} = 2$, respectively, with errors no bigger than $4\cdot 10^{-3}$, then $\eta \geq 2-(1+4\cdot 10^{-3}) = 1-4\cdot 10^{-3}$.
3. Let B and $\widetilde{B}$ be as in Example 1. Fact 3 gives $\max_{1\leq j\leq 2}|\sigma_j - \widetilde{\sigma}_j| \leq 1.6\cdot 10^{-5}$, a much better bound than by Fact 1.
4. Let B and $\widetilde{B}$ be as in Example 1. Note $\widetilde{B}$'s SVD there. Apply Fact 4 with $k = 1$ to give a similar bound as by Fact 2(c).
5. Since unitary transformations do not change singular values, Fact 3 applies to $B,\ \widetilde{B} \in \mathbb{C}^{m\times n}$ having decompositions

$$\begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix} B[V_1\ \ V_2] = \begin{bmatrix} B_1 & F \\ E & B_2 \end{bmatrix}, \quad \begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix} \widetilde{B}[V_1\ \ V_2] = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix},$$

where $[U_1\ \ U_2]$ and $[V_1\ \ V_2]$ are unitary and $U_1 \in \mathbb{C}^{m\times k}$, $V_1 \in \mathbb{C}^{n\times k}$.

15.3 Polar Decomposition

The reader is referred to Chapter 17.1 for definition and for more information on polar decompositions.

Definitions:

$B \in \mathbb{F}^{m\times n}$ is perturbed to $\widetilde{B} = B + \Delta B$, and their polar decompositions are

$$B = QH, \quad \widetilde{B} = \widetilde{Q}\widetilde{H} = (Q + \Delta Q)(H + \Delta H),$$

where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. ΔB *is restricted to* $\mathbb{F}$ *for* $B \in \mathbb{F}$.

Denote the singular values of B and $\widetilde{B}$ as $\sigma_1 \geq \sigma_2 \geq \cdots$ and $\widetilde{\sigma}_1 \geq \widetilde{\sigma}_2 \geq \cdots$, respectively.

The **condition numbers for the polar factors** in the Frobenius norm are defined as

$$\mathrm{cond}_{\mathbb{F}}(X) = \lim_{\delta \to 0} \sup_{\|\Delta B\|_{\mathrm{F}} \leq \delta} \frac{\|\Delta X\|_{\mathrm{F}}}{\delta}, \quad \text{for } X = H \text{ or } Q.$$

B is **multiplicatively perturbed** to $\widetilde{B}$ if $\widetilde{B} = D_{\mathrm{L}}^* B D_{\mathrm{R}}$ for some $D_{\mathrm{L}} \in \mathbb{F}^{m\times m}$ and $D_{\mathrm{R}} \in \mathbb{F}^{n\times n}$.

B is said to be **graded** if it can be scaled as $B = GS$ such that G is "well-behaved" (i.e., $\kappa_2(G)$ is of modest magnitude), where S is a scaling matrix, often diagonal but not required so for the facts below. Interesting cases are when $\kappa_2(G) \ll \kappa_2(B)$.

Facts:

1. [CG00] The condition numbers $\mathrm{cond}_{\mathbb{F}}(Q)$ and $\mathrm{cond}_{\mathbb{F}}(H)$ are tabulated as follows, where $\kappa_2(B) = \sigma_1/\sigma_n$.

		$\mathbb{R}$	$\mathbb{C}$
Factor Q	$m = n$	$2/(\sigma_{n-1} + \sigma_n)$	$1/\sigma_n$
	$m > n$	$1/\sigma_n$	$1/\sigma_n$
Factor H	$m \geq n$	$\dfrac{\sqrt{2(1+\kappa_2(B)^2)}}{1+\kappa_2(B)}$	

2. [Kit86] $\|\Delta H\|_{\mathrm{F}} \leq \sqrt{2}\|\Delta B\|_{\mathrm{F}}$.
3. [Li95] If $m = n$ and $\mathrm{rank}(B) = n$, then

$$\|\Delta Q\|_{\mathrm{UI}} \leq \frac{2}{\sigma_n + \widetilde{\sigma}_n}\|\Delta B\|_{\mathrm{UI}}.$$

4. [Li95], [LS02] If $\mathrm{rank}(B) = n$, then

$$\|\Delta Q\|_{\mathrm{UI}} \leq \left(\frac{2}{\sigma_n + \widetilde{\sigma}_n} + \frac{1}{\max\{\sigma_n, \widetilde{\sigma}_n\}}\right)\|\Delta B\|_{\mathrm{UI}},$$

$$\|\Delta Q\|_{\mathrm{F}} \leq \frac{2}{\sigma_n + \widetilde{\sigma}_n}\|\Delta B\|_{\mathrm{F}}.$$

5. [Mat93] If $B \in \mathbb{R}^{n\times n}$, $\mathrm{rank}(B) = n$, and $\|\Delta B\|_2 < \sigma_n$, then

$$\|\Delta Q\|_{\mathrm{UI}} \leq -\frac{2\|\Delta B\|_{\mathrm{UI}}}{|||\Delta B|||_2} \ln\left(1 - \frac{|||\Delta B|||_2}{\sigma_n + \sigma_{n-1}}\right),$$

where $|||\cdot|||_2$ is the **Ky Fan 2-norm**, i.e., the sum of the first two largest singular values. (See Chapter 17.3.)

6. [LS02] If $B \in \mathbb{R}^{n\times n}$, $\mathrm{rank}(B) = n$, and $\|\Delta B\|_2 < \sigma_n + \widetilde{\sigma}_n$, then

$$\|\Delta Q\|_{\mathrm{F}} \leq \frac{4}{\sigma_{n-1} + \sigma_n + \widetilde{\sigma}_{n-1} + \widetilde{\sigma}_n}\|\Delta B\|_{\mathrm{F}}.$$

7. [Li97] Let B and $\widetilde{B} = D_{\mathrm{L}}^* B D_{\mathrm{R}}$ having full column rank. Then

$$\|\Delta Q\|_{\mathrm{F}} \leq \sqrt{\|I - D_{\mathrm{L}}^{-1}\|_{\mathrm{F}}^2 + \|D_{\mathrm{L}} - I\|_{\mathrm{F}}^2} + \sqrt{\|I - D_{\mathrm{R}}^{-1}\|_{\mathrm{F}}^2 + \|D_{\mathrm{R}} - I\|_{\mathrm{F}}^2}.$$

8. [Li97], [Li05] Let $B = GS$ and $\widetilde{B} = \widetilde{G}S$ and assume that G and B have full column rank. If $\|\Delta G\|_2\|G^{\dagger}\|_2 < 1$, then

$$\|\Delta Q\|_{\mathrm{F}} \leq \gamma \, \|G^{\dagger}\|_2 \|\Delta G\|_{\mathrm{F}},$$
$$\|(\Delta H)S^{-1}\|_{\mathrm{F}} \leq \left(\gamma \, \|G^{\dagger}\|_2 \|G\|_2 + 1\right) \|\Delta G\|_{\mathrm{F}},$$

where $\gamma = \sqrt{1 + \left(1 - \|G^{\dagger}\|_2 \|\Delta G\|_2\right)^{-2}}$.

Examples:

1. Take both B and $\widetilde{B}$ to have orthonormal columns to see that some of the inequalities above on ΔQ are attainable.
2. Let

$$B = \frac{1}{\sqrt{2}} \begin{bmatrix} 2.01 & 502 \\ -1.99 & -498 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 10^{-2} & 2 \\ 2 & 5 \cdot 10^2 \end{bmatrix}$$

and

$$\widetilde{B} = \begin{bmatrix} 1.4213 & 3.5497 \cdot 10^2 \\ -1.4071 & -3.5214 \cdot 10^2 \end{bmatrix}$$

obtained by rounding each entry of B to have five significant decimal digits. $B = QH$ can be read off above and $\widetilde{B} = \widetilde{Q}\widetilde{H}$ can be computed by $\widetilde{Q} = \widetilde{U}\widetilde{V}^*$ and $\widetilde{H} = \widetilde{V}\widetilde{\Sigma}\widetilde{V}^*$, where $\widetilde{B}$'s SVD is $\widetilde{U}\widetilde{\Sigma}\widetilde{V}^*$. One has

$$\mathrm{SV}(B) = \{5.00 \cdot 10^2, 2.00 \cdot 10^{-3}\}, \ \mathrm{SV}(\widetilde{B}) = \{5.00 \cdot 10^2, 2.04 \cdot 10^{-3}\}$$

and

$\|\Delta B\|_2$	$\|\Delta B\|_{\mathrm{F}}$	$\|\Delta Q\|_2$	$\|\Delta Q\|_{\mathrm{F}}$	$\|\Delta H\|_2$	$\|\Delta H\|_{\mathrm{F}}$
$3 \cdot 10^{-3}$	$3 \cdot 10^{-3}$	$2 \cdot 10^{-6}$	$3 \cdot 10^{-6}$	$2 \cdot 10^{-3}$	$2 \cdot 10^{-3}$

Fact 2 gives $\|\Delta H\|_{\mathrm{F}} \leq 3 \cdot 10^{-3}$ and Fact 6 gives $\|\Delta Q\|_{\mathrm{F}} \leq 10^{-5}$.
3. [Li97] and [Li05] have examples on the use of inequalities in Facts 7 and 8.

15.4 Generalized Eigenvalue Problems

The reader is referred to Section 43.1 for more information on generalized eigenvalue problems.

Definitions:

Let $A, B \in \mathbb{C}^{m \times n}$. A **matrix pencil** is a family of matrices $A - \lambda B$, parameterized by a (complex) number λ. The associated **generalized eigenvalue problem** is to find the nontrivial solutions of the equations

$$A\mathbf{x} = \lambda B\mathbf{x} \quad \text{and/or} \quad \mathbf{y}^* A = \lambda \mathbf{y}^* B,$$

where $\mathbf{x} \in \mathbb{C}^n$, $\mathbf{y} \in \mathbb{C}^m$, and $\lambda \in \mathbb{C}$, and $(\mathbf{x}, \lambda, \mathbf{y})$ is called a **generalized eigentriplet**.

$A - \lambda B$ is **regular** if $m = n$ and $\det(A - \lambda B) \neq 0$ for some $\lambda \in \mathbb{C}$.

In what follows, all pencils in question are assumed regular.

An eigenvalue λ is conveniently represented by a nonzero number pair, so-called a **generalized eigenvalue** $\langle \alpha, \beta \rangle$, interpreted as $\lambda = \alpha/\beta$. $\beta = 0$ corresponds to eigenvalue infinity.

A **generalized eigenpair** of $A - \lambda B$ refers to $(\langle\alpha, \beta\rangle, \mathbf{x})$ such that $\beta A\mathbf{x} = \alpha B\mathbf{x}$, where $\mathbf{x} \neq 0$ and $|\alpha|^2 + |\beta|^2 > 0$. A **generalized eigentriplet** of $A - \lambda B$ refers to $(\mathbf{y}, \langle\alpha, \beta\rangle, \mathbf{x})$ such that $\beta A\mathbf{x} = \alpha B\mathbf{x}$ and $\beta\mathbf{y}^* A = \alpha\mathbf{y}^* B$, where $\mathbf{x} \neq 0$, $\mathbf{y} \neq 0$, and $|\alpha|^2 + |\beta|^2 > 0$. The quantity

$$\text{cond}(\langle\alpha, \beta\rangle) = \frac{\|\mathbf{x}\|_2\|\mathbf{y}\|_2}{\sqrt{|\mathbf{y}^* A\mathbf{x}|^2 + |\mathbf{y}^* B\mathbf{x}|^2}}$$

is the **individual condition number** for the generalized eigenvalue $\langle\alpha, \beta\rangle$, where $(\mathbf{y}, \langle\alpha, \beta\rangle, \mathbf{x})$ is a generalized eigentriplet of $A - \lambda B$.

$A - \lambda B$ is perturbed to $\widetilde{A} - \lambda\widetilde{B} = (A + \Delta A) - \lambda(B + \Delta B)$.

Let $\sigma(A, B) = \{\langle\alpha_1, \beta_1\rangle, \langle\alpha_2, \beta_2\rangle, \ldots, \langle\alpha_n, \beta_n\rangle\}$ be the set of the generalized eigenvalues of $A - \lambda B$, and set $Z = [A,\ B] \in \mathbb{C}^{2n\times n}$.

$A - \lambda B$ is **diagonalizable** if it is equivalent to a diagonal pencil, i.e., there are nonsingular $X, Y \in \mathbb{C}^{n\times n}$ such that $Y^* AX = \Lambda$, $Y^* BX = \Omega$, where $\Lambda = \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and $\Omega = \text{diag}(\beta_1, \beta_2, \ldots, \beta_n)$.

$A - \lambda B$ is a **definite pencil** if both A and B are Hermitian and

$$\gamma(A, B) = \min_{\mathbf{x}\in\mathbb{C}^n, \|\mathbf{x}\|_2=1} |\mathbf{x}^* A\mathbf{x} + i\mathbf{x}^* B\mathbf{x}| > 0.$$

The same notation is adopted for $\widetilde{A} - \lambda\widetilde{B}$, except all symbols with tildes.

The **chordal distance** between two nonzero pairs $\langle\alpha, \beta\rangle$ and $\langle\widetilde{\alpha}, \widetilde{\beta}\rangle$ is

$$\chi\left(\langle\alpha, \beta\rangle, \langle\widetilde{\alpha}, \widetilde{\beta}\rangle\right) = \frac{|\widetilde{\beta}\alpha - \widetilde{\alpha}\beta|}{\sqrt{|\alpha|^2 + |\beta|^2}\sqrt{|\widetilde{\alpha}|^2 + |\widetilde{\beta}|^2}}.$$

Facts:

1. [SS90, p. 293] Let $(\mathbf{y}, \langle\alpha, \beta\rangle, \mathbf{x})$ be a generalized eigentriplet of $A - \lambda B$. $[\Delta A,\ \Delta B]$ changes $\langle\alpha, \beta\rangle = \langle\mathbf{y}^* A\mathbf{x}, \mathbf{y}^* B\mathbf{x}\rangle$ to

$$\langle\widetilde{\alpha}, \widetilde{\beta}\rangle = \langle\alpha, \beta\rangle + \langle\mathbf{y}^*(\Delta A)\mathbf{x}, \mathbf{y}^*(\Delta B)\mathbf{x}\rangle + O(\varepsilon^2),$$

   where $\varepsilon = \|[\Delta A,\ \Delta B]\|_2$, and $\chi\left(\langle\alpha, \beta\rangle, \langle\widetilde{\alpha}, \widetilde{\beta}\rangle\right) \leq \text{cond}(\langle\alpha, \beta\rangle)\,\varepsilon + O(\varepsilon^2)$.
2. [SS90, p. 301], [Li88] If $A - \lambda B$ is diagonalizable, then

$$\max_i \min_j \chi\left(\langle\alpha_i, \beta_i\rangle, \langle\widetilde{\alpha}_j, \widetilde{\beta}_j\rangle\right) \leq \kappa_2(X)\|\sin\Theta(Z^*, \widetilde{Z}^*)\|_2.$$

3. [Li94, Lemma 3.3] (**Sun**)

$$\|\sin\Theta(Z^*, \widetilde{Z}^*)\|_{\text{UI}} \leq \frac{\|Z - \widetilde{Z}\|_{\text{UI}}}{\max\{\sigma_{\min}(Z), \sigma_{\min}(\widetilde{Z})\}},$$

   where $\sigma_{\min}(Z)$ is Z's smallest singular value.
4. The quantity $\gamma(A, B)$ is the minimum distance of the numerical range $W(A + iB)$ to the origin for definite pencil $A - \lambda B$.
5. [SS90, p. 316] Suppose $A - \lambda B$ is a definite pencil. If $\widetilde{A}$ and $\widetilde{B}$ are Hermitian and $\|[\Delta A,\ \Delta B]\|_2 < \gamma(A, B)$, then $\widetilde{A} - \lambda\widetilde{B}$ is also a definite pencil and there exists a permutation τ of $\{1, 2, \ldots, n\}$ such that

$$\max_{1\leq j\leq n} \chi\left(\langle\alpha_j, \beta_j\rangle, \langle\widetilde{\alpha}_{\tau(j)}, \widetilde{\beta}_{\tau(j)}\rangle\right) \leq \frac{\|[\Delta A,\ \Delta B]\|_2}{\gamma(A, B)}.$$

6. [SS90, p. 318] Definite pencil $A - \lambda B$ is always diagonalizable: $X^* AX = \Lambda$ and $X^* BX = \Omega$, and with real spectra. Facts 7 and 10 apply.

7. [Li03] Suppose $A - \lambda B$ and $\widetilde{A} - \lambda \widetilde{B}$ are diagonalizable with real spectra, i.e.,

$$Y^* A X = \Lambda,\ Y^* B X = \Omega \quad \text{and} \quad \widetilde{Y}^* \widetilde{A} \widetilde{X} = \widetilde{\Lambda},\ \widetilde{Y}^* \widetilde{B} \widetilde{X} = \widetilde{\Omega},$$

and all $\langle \alpha_j, \beta_j \rangle$ and all $\langle \widetilde{\alpha}_j, \widetilde{\beta}_j \rangle$ are real. Then the follow statements hold, where

$$\Xi = \operatorname{diag}(\chi(\langle \alpha_1, \beta_1 \rangle, \langle \widetilde{\alpha}_{\tau(1)}, \widetilde{\beta}_{\tau(1)} \rangle), \dots, \chi(\langle \alpha_n, \beta_n \rangle, \langle \widetilde{\alpha}_{\tau(n)}, \widetilde{\beta}_{\tau(n)} \rangle))$$

for some permutation τ of $\{1, 2, \dots, n\}$ (possibly depending on the norm being used). In all cases, the constant factor $\pi/2$ can be replaced by 1 for the 2-norm and the Frobenius norm.

   (a) $\|\Xi\|_{\text{UI}} \le \frac{\pi}{2} \sqrt{\kappa_2(X) \kappa_2(\widetilde{X})} \|\sin \Theta(Z^*, \widetilde{Z}^*)\|_{\text{UI}}$.

   (b) If all $|\alpha_j|^2 + |\beta_j|^2 = |\widetilde{\alpha}_j|^2 + |\widetilde{\beta}_j|^2 = 1$ in their eigendecompositions, then $\|\Xi\|_{\text{UI}} \le \frac{\pi}{2} \sqrt{\|X\|_2 \|Y^*\|_2 \|\widetilde{X}\|_2 \|\widetilde{Y}^*\|_2} \|[\Delta A,\ \Delta B]\|_{\text{UI}}$.

8. Let residuals $\mathbf{r} = \widetilde{\beta} A \widetilde{\mathbf{x}} - \widetilde{\alpha} B \widetilde{\mathbf{x}}$ and $\mathbf{s}^* = \widetilde{\beta} \widetilde{\mathbf{y}}^* A - \widetilde{\alpha} \widetilde{\mathbf{y}}^* B$, where $\|\widetilde{\mathbf{x}}\|_2 = \|\widetilde{\mathbf{y}}\|_2 = 1$. The smallest error matrix $[\Delta A,\ \Delta B]$ in the 2-norm, for which $(\widetilde{\mathbf{y}}, \langle \widetilde{\alpha}, \widetilde{\beta} \rangle, \widetilde{\mathbf{x}})$ is an exact generalized eigentriplet of $\widetilde{A} - \lambda \widetilde{B}$, satisfies $\|[\Delta A,\ \Delta B]\|_2 = \varepsilon$, where $\varepsilon = \max\{\|\mathbf{r}\|_2, \|\mathbf{s}\|_2\}$, and $\chi\left(\langle \alpha, \beta \rangle, \langle \widetilde{\alpha}, \widetilde{\beta} \rangle\right) \le \operatorname{cond}(\langle \widetilde{\alpha}, \widetilde{\beta} \rangle)\, \varepsilon + O(\varepsilon^2)$ for some $\langle \alpha, \beta \rangle \in \sigma(A, B)$.
9. [BDD00, p. 128] Suppose A and B are Hermitian and B is positive definite, and let residual $\mathbf{r} = A\widetilde{\mathbf{x}} - \widetilde{\mu} B \widetilde{\mathbf{x}}$ and $\|\widetilde{\mathbf{x}}\|_2 = 1$.

   (a) For some eigenvalue μ of $A - \lambda B$,

$$|\widetilde{\mu} - \mu| \le \frac{\|\mathbf{r}\|_{B^{-1}}}{\|\widetilde{\mathbf{x}}\|_B} \le \|B^{-1}\|_2 \|\mathbf{r}\|_2,$$

   where $\|\mathbf{z}\|_M = \sqrt{\mathbf{z}^* M \mathbf{z}}$.

   (b) Let μ be the closest eigenvalue to $\widetilde{\mu}$ among all eigenvalues of $A - \lambda B$ and $\mathbf{x}$ its associated eigenvector with $\|\mathbf{x}\|_2 = 1$, and let $\eta = \min |\widetilde{\mu} - \nu|$ over all other eigenvalues $\nu \ne \mu$ of $A - \lambda B$. If $\eta > 0$, then

$$|\widetilde{\mu} - \mu| \le \frac{1}{\eta} \cdot \left(\frac{\|\mathbf{r}\|_{B^{-1}}}{\|\widetilde{\mathbf{x}}\|_B} \right)^2 \le \|B^{-1}\|_2^2 \frac{\|\mathbf{r}\|_2^2}{\eta},$$
$$\sin \angle(\widetilde{\mathbf{x}}, \mathbf{x}) \le \|B^{-1}\|_2 \sqrt{2\kappa_2(B)}\, \frac{\|\mathbf{r}\|_2}{\eta}.$$

10. [Li94] Suppose $A - \lambda B$ and $\widetilde{A} - \lambda \widetilde{B}$ are diagonalizable and have eigendecompositions

$$\begin{bmatrix} Y_1^* \\ Y_2^* \end{bmatrix} A[X_1,\ X_2] = \begin{bmatrix} \Lambda_1 & \\ & \Lambda_2 \end{bmatrix}, \begin{bmatrix} Y_1^* \\ Y_2^* \end{bmatrix} B[X_1,\ X_2] = \begin{bmatrix} \Omega_1 & \\ & \Omega_2 \end{bmatrix},$$

$$X^{-1} = [W_1,\ W_2]^*,$$

and the same for $\widetilde{A} - \lambda \widetilde{B}$ except all symbols with tildes, where $X_1,\ Y_1,\ W_1 \in \mathbb{C}^{n \times k}$, $\Lambda_1,\ \Omega_1 \in \mathbb{C}^{k \times k}$. Suppose $|\alpha_j|^2 + |\beta_j|^2 = |\widetilde{\alpha}_j|^2 + |\widetilde{\beta}_j|^2 = 1$ for $1 \le j \le n$ in the eigendecompositions, and set $\eta = \min \chi\left(\langle \alpha, \beta \rangle, \langle \widetilde{\alpha}, \widetilde{\beta} \rangle\right)$ taken over all $\langle \alpha, \beta \rangle \in \sigma(\Lambda_1, \Omega_1)$ and $\langle \widetilde{\alpha}, \widetilde{\beta} \rangle \in \sigma(\widetilde{\Lambda}_2, \widetilde{\Omega}_2)$. If $\eta > 0$, then

$$\left\| \sin \Theta(X_1, \widetilde{X}_1) \right\|_{\text{F}} \le \frac{\|X_1^\dagger\|_2 \|\widetilde{W}_2^\dagger\|_2}{\eta} \left\| \widetilde{Y}_2^* (\widetilde{Z} - Z) \begin{bmatrix} X_1 & \\ & X_1 \end{bmatrix} \right\|_{\text{F}}.$$

15.5 Generalized Singular Value Problems

Definitions:

Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{\ell\times n}$. A matrix pair $\{A, B\}$ is an (m,ℓ,n)-**Grassmann matrix pair** if $\operatorname{rank}\left(\begin{bmatrix} A \\ B \end{bmatrix}\right) = n$.

In what follows, all matrix pairs are (m,ℓ,n)-Grassmann matrix pairs.

A pair $\langle\alpha, \beta\rangle$ is a **generalized singular value** of $\{A, B\}$ if

$$\det(\beta^2 A^*A - \alpha^2 B^*B) = 0,\ \langle\alpha,\beta\rangle \neq \langle 0,0\rangle,\ \alpha,\ \beta \geq 0,$$

i.e., $\langle\alpha,\beta\rangle = \langle\sqrt{\mu}, \sqrt{\nu}\rangle$ for some generalized eigenvalue $\langle\mu,\nu\rangle$ of matrix pencil $A^*A - \lambda B^*B$.

Generalized Singular Value Decomposition (GSVD) of $\{A, B\}$:

$$U^*AX = \Sigma_A, \quad V^*BX = \Sigma_B,$$

where $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{\ell\times\ell}$ are unitary, $X \in \mathbb{C}^{n\times n}$ is nonsingular, $\Sigma_A = \operatorname{diag}(\alpha_1, \alpha_2, \cdots)$ is **leading diagonal** (α_j starts in the top left corner), and $\Sigma_B = \operatorname{diag}(\cdots, \beta_{n-1}, \beta_n)$ is **trailing diagonal** (β_j ends in the bottom-right corner), $\alpha_j,\ \beta_j \geq 0$ and $\alpha_j^2 + \beta_j^2 = 1$ for $1 \leq j \leq n$. (Set some $\alpha_j = 0$ and/or some $\beta_j = 0$, if necessary.)

$\{A, B\}$ is perturbed to $\{\widetilde{A}, \widetilde{B}\} = \{A + \Delta A, B + \Delta B\}$.

Let $\operatorname{SV}(A, B) = \{\langle\alpha_1, \beta_1\rangle, \langle\alpha_2, \beta_2\rangle, \ldots, \langle\alpha_n, \beta_n\rangle\}$ be the set of the generalized singular values of $\{A, B\}$, and set $Z = \begin{bmatrix} A \\ B \end{bmatrix} \in \mathbb{C}^{(m+\ell)\times n}$.

The same notation is adopted for $\{\widetilde{A}, \widetilde{B}\}$, except all symbols with tildes.

Facts:

1. If $\{A, B\}$ is an (m,ℓ,n)-Grassmann matrix pair, then $A^*A - \lambda B^*B$ is a definite matrix pencil.
2. [Van76] The GSVD of an (m,ℓ,n)-Grassmann matrix pair $\{A, B\}$ exists.
3. [Li93] There exist permutations τ and ω of $\{1, 2, \ldots, n\}$ such that

$$\max_{1\leq j\leq n} \chi\left(\langle\alpha_i, \beta_i\rangle, \langle\widetilde{\alpha}_{\tau(i)}, \widetilde{\beta}_{\tau(i)}\rangle\right) \leq \|\sin\Theta(Z, \widetilde{Z})\|_2,$$

$$\sqrt{\sum_{j=1}^{n}\left[\chi\left(\langle\alpha_i, \beta_i\rangle, \langle\widetilde{\alpha}_{\omega(i)}, \widetilde{\beta}_{\omega(i)}\rangle\right)\right]^2} \leq \|\sin\Theta(Z, \widetilde{Z})\|_{\mathrm{F}}.$$

4. [Li94, Lemma 3.3] (**Sun**)

$$\|\sin\Theta(Z, \widetilde{Z})\|_{\mathrm{UI}} \leq \frac{\|Z - \widetilde{Z}\|_{\mathrm{UI}}}{\max\{\sigma_{\min}(Z),\ \sigma_{\min}(\widetilde{Z})\}},$$

where $\sigma_{\min}(Z)$ is Z's smallest singular value.

5. [Pai84] If $\alpha_i^2 + \beta_i^2 = \widetilde{\alpha}_i^2 + \widetilde{\beta}_i^2 = 1$ for $i = 1, 2, \ldots, n$, then there exists a permutation ϖ of $\{1, 2, \ldots, n\}$ such that

$$\sqrt{\sum_{j=1}^{n}\left[(\alpha_j - \widetilde{\alpha}_{\varpi(j)})^2 + (\beta_j - \widetilde{\beta}_{\varpi(j)})^2\right]} \leq \min_{Q \text{ unitary}} \|Z_0 - \widetilde{Z}_0 Q\|_{\mathrm{F}},$$

where $Z_0 = Z(Z^*Z)^{-1/2}$ and $\widetilde{Z}_0 = \widetilde{Z}(\widetilde{Z}^*\widetilde{Z})^{-1/2}$.

6. [Li93], [Sun83] Perturbation bounds on generalized singular subspaces (those spanned by one or a few columns of U, V, and X in GSVD) are also available, but it is quite complicated.

15.6 Relative Perturbation Theory for Eigenvalue Problems

Definitions:

Let scalar $\widetilde{\alpha}$ be an approximation to α, and $1 \le p \le \infty$. Define **relative distances** between α and $\widetilde{\alpha}$ as follows. For $|\alpha|^2 + |\widetilde{\alpha}|^2 \neq 0$,

$$\begin{aligned} d(\alpha,\widetilde{\alpha}) &= \left|\frac{\widetilde{\alpha}}{\alpha} - 1\right| = \frac{|\widetilde{\alpha}-\alpha|}{|\alpha|}, && \text{(classical measure)} \\ \varrho_p(\alpha,\widetilde{\alpha}) &= \frac{|\widetilde{\alpha}-\alpha|}{\sqrt[p]{|\alpha|^p + |\widetilde{\alpha}|^p}}, && \text{([Li98])} \\ \zeta(\alpha,\widetilde{\alpha}) &= \frac{|\widetilde{\alpha}-\alpha|}{\sqrt{|\alpha\widetilde{\alpha}|}}, && \text{([BD90], [DV92])} \\ \varsigma(\alpha,\widetilde{\alpha}) &= |\ln(\widetilde{\alpha}/\alpha)|, \quad \text{for } \widetilde{\alpha}\alpha > 0, && \text{([LM99a], [Li99b])} \end{aligned}$$

and $d(0,0) = \varrho_p(0,0) = \zeta(0,0) = \varsigma(0,0) = 0$.

$A \in \mathbb{C}^{n\times n}$ is **multiplicatively perturbed** to $\widetilde{A}$ if $\widetilde{A} = D_{\text{L}}^* A D_{\text{R}}$ for some $D_{\text{L}},\ D_{\text{R}} \in \mathbb{C}^{n\times n}$.

Denote $\sigma(A) = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ and $\sigma(\widetilde{A}) = \{\widetilde{\lambda}_1, \widetilde{\lambda}_2, \dots, \widetilde{\lambda}_n\}$.

$A \in \mathbb{C}^{n\times n}$ is said to be **graded** if it can be scaled as $A = S^* H S$ such that H is "well-behaved" (i.e., $\kappa_2(H)$ is of modest magnitude), where S is a scaling matrix, often diagonal but not required so for the facts below. Interesting cases are when $\kappa_2(H) \ll \kappa_2(A)$.

Facts:

1. [Bar00] $\varrho_p(\cdot\,,\cdot)$ is a metric on $\mathbb{C}$ for $1 \le p \le \infty$.
2. Let $A, \widetilde{A} = D^* A D \in \mathbb{C}^{n\times n}$ be Hermitian, where D is nonsingular.

   (a) [HJ85, p. 224] (**Ostrowski**) There exists t_j, satisfying

   $$\lambda_{\min}(D^*D) \le t_j \le \lambda_{\max}(D^*D),$$

   such that $\widetilde{\lambda}_j^{\uparrow} = t_j \lambda_j^{\uparrow}$ for $j = 1, 2, \dots, n$ and, thus,

   $$\max_{1\le j\le n} d(\lambda_j^{\uparrow}, \widetilde{\lambda}_j^{\uparrow}) \le \|I - D^*D\|_2.$$

   (b) [LM99], [Li98]

   $$\begin{aligned} \left\|\operatorname{diag}\left(\varsigma(\lambda_1^{\uparrow}, \widetilde{\lambda}_1^{\uparrow}), \dots, \varsigma(\lambda_n^{\uparrow}, \widetilde{\lambda}_n^{\uparrow})\right)\right\|_{\text{UI}} &\le \|\ln(D^*D)\|_{\text{UI}}, \\ \left\|\operatorname{diag}\left(\zeta(\lambda_1^{\uparrow}, \widetilde{\lambda}_1^{\uparrow}), \dots, \zeta(\lambda_n^{\uparrow}, \widetilde{\lambda}_n^{\uparrow})\right)\right\|_{\text{UI}} &\le \|D^* - D^{-1}\|_{\text{UI}}. \end{aligned}$$

3. [Li98], [LM99] Let $A = S^* H S$ be a positive semidefinite Hermitian matrix, perturbed to $\widetilde{A} = S^*(H + \Delta H)S$. Suppose H is positive definite and $\|H^{-1/2}(\Delta H)H^{-1/2}\|_2 < 1$, and set

   $$M = H^{1/2} S S^* H^{1/2}, \quad \widetilde{M} = DMD,$$

   where $D = \left[I + H^{-1/2}(\Delta H)H^{-1/2}\right]^{1/2} = D^*$. Then $\sigma(A) = \sigma(M)$ and $\sigma(\widetilde{A}) = \sigma(\widetilde{M})$, and the inequalities in Fact 2 above hold with D here. Note that

$$\|D - D^{-1}\|_{\text{UI}} \leq \frac{\|H^{-1/2}(\Delta H)H^{-1/2}\|_{\text{UI}}}{\sqrt{1 - \|H^{-1/2}(\Delta H)H^{-1/2}\|_2}}$$
$$\leq \frac{\|H^{-1}\|_2}{\sqrt{1 - \|H^{-1}\|_2\|\Delta H\|_2}} \|\Delta H\|_{\text{UI}}.$$

4. [BD90], [VS93] Suppose A and $\widetilde{A}$ are Hermitian, and let $|A| = (A^2)^{1/2}$ be the positive semidefinite square root of A^2. If there exists $0 \leq \delta < 1$ such that

$$|\mathbf{x}^*(\Delta A)\mathbf{x}| \leq \delta \mathbf{x}^*|A|\mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{C}^n,$$

then either $\widetilde{\lambda}_j^{\uparrow} = \lambda_j^{\uparrow} = 0$ or $1 - \delta \leq \widetilde{\lambda}_j^{\uparrow}/\lambda_j^{\uparrow} \leq 1 + \delta$.

5. [Li99a] Let Hermitian A, $\widetilde{A} = D^*AD$ have decompositions

$$\begin{bmatrix} X_1^* \\ X_2^* \end{bmatrix} A[X_1 \;\; X_2] = \begin{bmatrix} A_1 & \\ & A_2 \end{bmatrix}, \quad \begin{bmatrix} \widetilde{X}_1^* \\ \widetilde{X}_2^* \end{bmatrix} \widetilde{A}[\widetilde{X}_1 \;\; \widetilde{X}_2] = \begin{bmatrix} \widetilde{A}_1 & \\ & \widetilde{A}_2 \end{bmatrix},$$

where $[X_1 \;\; X_2]$ and $[\widetilde{X}_1 \;\; \widetilde{X}_2]$ are unitary and $X_1, \widetilde{X}_1 \in \mathbb{C}^{n\times k}$. If $\eta_2 = \min\limits_{\mu\in\sigma(A_1), \widetilde{\mu}\in\sigma(\widetilde{A}_2)} \varrho_2(\mu, \widetilde{\mu}) > 0$, then

$$\|\sin\Theta(X_1, \widetilde{X}_1)\|_{\text{F}} \leq \frac{\sqrt{\|(I - D^{-1})X_1\|_{\text{F}}^2 + \|(I - D^*)X_1\|_{\text{F}}^2}}{\eta_2}.$$

6. [Li99a] Let $A = S^*HS$ be a positive semidefinite Hermitian matrix, perturbed to $\widetilde{A} = S^*(H + \Delta H)S$, having decompositions, in notation, the same as in Fact 5. Let $D = \left[I + H^{-1/2}(\Delta H)H^{-1/2}\right]^{1/2}$. Assume H is positive definite and $\|H^{-1/2}(\Delta H)H^{-1/2}\|_2 < 1$. If $\eta_\zeta = \min\limits_{\mu\in\sigma(A_1), \widetilde{\mu}\in\sigma(\widetilde{A}_2)} \zeta(\mu, \widetilde{\mu}) > 0$, then

$$\|\sin\Theta(X_1, \widetilde{X}_1)\|_{\text{F}} \leq \frac{\|D - D^{-1}\|_{\text{F}}}{\eta_\zeta}.$$

Examples:

1. [DK90], [EI95] Let A be a real symmetric tridiagonal matrix with zero diagonal and off-diagonal entries $b_1, b_2, \ldots, b_{n-1}$. Suppose $\widetilde{A}$ is identical to A except for its off-diagonal entries which change to $\beta_1 b_1, \beta_2 b_2, \ldots, \beta_{n-1}b_{n-1}$, where all β_i are real and supposedly close to 1. Then $\widetilde{A} = DAD$, where $D = \text{diag}(d_1, d_2, \ldots, d_n)$ with

$$d_{2k} = \frac{\beta_1\beta_3\cdots\beta_{2k-1}}{\beta_2\beta_4\cdots\beta_{2k-2}}, \quad d_{2k+1} = \frac{\beta_2\beta_4\cdots\beta_{2k}}{\beta_1\beta_3\cdots\beta_{2k-1}}.$$

Let $\beta = \prod_{j=1}^{n-1} \max\{\beta_j, 1/\beta_j\}$. Then $\beta^{-1}I \leq D^2 \leq \beta I$, and Fact 2 and Fact 5 apply. Now if all $1 - \varepsilon \leq \beta_j \leq 1 + \varepsilon$, then $(1 - \varepsilon)^{n-1} \leq \beta^{-1} \leq \beta \leq (1 + \varepsilon)^{n-1}$.

2. Let $A = SHS$ with $S = \text{diag}(1, 10, 10^2, 10^3)$, and

$$A = \begin{bmatrix} 1 & 1 & & \\ 1 & 10^2 & 10^2 & \\ & 10^2 & 10^4 & 10^4 \\ & & 10^4 & 10^6 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 10^{-1} & & \\ 10^{-1} & 1 & 10^{-1} & \\ & 10^{-1} & 1 & 10^{-1} \\ & & 10^{-1} & 1 \end{bmatrix}.$$

Suppose that each entry A_{ij} of A is perturbed to $A_{ij}(1+\delta_{ij})$ with $|\delta_{ij}| \le \varepsilon$. Then $|(\Delta H)_{ij}| \le \varepsilon |H_{ij}|$ and thus $\|\Delta H\|_2 \le 1.2\varepsilon$. Since $\|H^{-1}\|_2 \le 10/8$, Fact 3 implies

$$\zeta(\lambda_j^{\uparrow}, \widetilde{\lambda}_j^{\uparrow}) \le 1.5\varepsilon/\sqrt{1-1.5\varepsilon} \approx 1.5\varepsilon.$$

15.7 Relative Perturbation Theory for Singular Value Problems

Definitions:

$B \in \mathbb{C}^{m\times n}$ is **multiplicatively perturbed** to $\widetilde{B}$ if $\widetilde{B} = D_{\mathrm{L}}^* B D_{\mathrm{R}}$ for some $D_{\mathrm{L}} \in \mathbb{C}^{m\times m}$ and $D_{\mathrm{R}} \in \mathbb{C}^{n\times n}$.

Denote the singular values of B and $\widetilde{B}$ as

$$\mathrm{SV}(B) = \{\sigma_1, \sigma_2, \dots, \sigma_{\min\{m,n\}}\}, \quad \mathrm{SV}(\widetilde{B}) = \{\widetilde{\sigma}_1, \widetilde{\sigma}_2, \dots, \widetilde{\sigma}_{\min\{m,n\}}\}.$$

B is said to be (highly) **graded** if it can be scaled as $B = GS$ such that G is "well-behaved" (i.e., $\kappa_2(G)$ is of modest magnitude), where S is a scaling matrix, often diagonal but not required so for the facts below. Interesting cases are when $\kappa_2(G) \ll \kappa_2(B)$.

Facts:

1. Let $B,\ \widetilde{B} = D_{\mathrm{L}}^* B D_{\mathrm{R}} \in \mathbb{C}^{m\times n}$, where D_{L} and D_{R} are nonsingular.
   (a) [EI95] For $1 \le j \le n$, $\dfrac{\sigma_j}{\|D_{\mathrm{L}}^{-1}\|_2 \|D_{\mathrm{R}}^{-1}\|_2} \le \widetilde{\sigma}_j \le \sigma_j \|D_{\mathrm{L}}\|_2 \|D_{\mathrm{R}}\|_2$.
   (b) [Li98], [LM99]

$$\begin{aligned} &\|\mathrm{diag}\,(\zeta(\sigma_1, \widetilde{\sigma}_1), \dots, \zeta(\sigma_n, \widetilde{\sigma}_n))\|_{\mathrm{UI}} \\ &\le \frac{1}{2}\|D_{\mathrm{L}}^* - D_{\mathrm{L}}^{-1}\|_{\mathrm{UI}} + \frac{1}{2}\|D_{\mathrm{R}}^* - D_{\mathrm{R}}^{-1}\|_{\mathrm{UI}}. \end{aligned}$$

2. [Li99a] Let $B,\ \widetilde{B} = D_{\mathrm{L}}^* B D_{\mathrm{R}} \in \mathbb{C}^{m\times n}$ $(m \ge n)$ have decompositions

$$\begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix} B [V_1\ \ V_2] = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}, \quad \begin{bmatrix} \widetilde{U}_1^* \\ \widetilde{U}_2^* \end{bmatrix} \widetilde{B} [\widetilde{V}_1\ \ \widetilde{V}_2] = \begin{bmatrix} \widetilde{B}_1 & 0 \\ 0 & \widetilde{B}_2 \end{bmatrix},$$

   where $[U_1\ U_2]$, $[V_1\ V_2]$, $[\widetilde{U}_1\ \widetilde{U}_2]$, and $[\widetilde{V}_1\ \widetilde{V}_2]$ are unitary, and $U_1,\ \widetilde{U}_1 \in \mathbb{C}^{m\times k}$, $V_1,\ \widetilde{V}_1 \in \mathbb{C}^{n\times k}$. Set $\Theta_U = \Theta(U_1, \widetilde{U}_1)$, $\Theta_V = \Theta(V_1, \widetilde{V}_1)$. If $\mathrm{SV}(B_1) \bigcap \mathrm{SV}_{\mathrm{ext}}(\widetilde{B}_2) = \emptyset$, then

$$\begin{aligned} &\sqrt{\|\sin\Theta_U\|_{\mathrm{F}}^2 + \|\sin\Theta_V\|_{\mathrm{F}}^2} \\ &\le \frac{1}{\eta_2}\left[\|(I - D_{\mathrm{L}}^*)U_1\|_{\mathrm{F}}^2 + \|(I - D_{\mathrm{L}}^{-1})U_1\|_{\mathrm{F}}^2\right. \\ &\quad \left. + \|(I - D_{\mathrm{R}}^*)V_1\|_{\mathrm{F}}^2 + \|(I - D_{\mathrm{R}}^{-1})V_1\|_{\mathrm{F}}^2\right]^{1/2}, \end{aligned}$$

   where $\eta_2 = \min \varrho_2(\mu, \widetilde{\mu})$ over all $\mu \in \mathrm{SV}(B_1)$ and $\widetilde{\mu} \in \mathrm{SV}_{\mathrm{ext}}(\widetilde{B}_2)$.
3. [Li98], [Li99a], [LM99] Let $B = GS$ and $\widetilde{B} = \widetilde{G}S$ be two $m \times n$ matrices, where $\mathrm{rank}(G) = n$, and let $\Delta G = \widetilde{G} - G$. Then $\widetilde{B} = DB$, where $D = I + (\Delta G)G^{\uparrow}$. Fact 1 and Fact 2 apply with $D_{\mathrm{L}} = D$ and $D_{\mathrm{R}} = I$. Note that

$$\|D^* - D^{-1}\|_{\mathrm{UI}} \le \left(1 + \frac{1}{1 - \|(\Delta G)G^{\uparrow}\|_2}\right) \frac{\|(\Delta G)G^{\uparrow}\|_{\mathrm{UI}}}{2}.$$

Examples:

1. [BD90], [DK90], [EI95] B is a real bidiagonal matrix with diagonal entries $a_1, a_2, \ldots, a_n$ and off-diagonal (the one above the diagonal) entries are $b_1, b_2, \ldots, b_{n-1}$. $\widetilde{B}$ is the same as B, except for its diagonal entries, which change to $\alpha_1 a_1, \alpha_2 a_2, \ldots, \alpha_n a_n$, and its off-diagonal entries, which change to $\beta_1 b_1, \beta_2 b_2, \ldots, \beta_{n-1} b_{n-1}$. Then $\widetilde{B} = D_{\mathrm{L}}^* B D_{\mathrm{R}}$ with

$$D_{\mathrm{L}} = \operatorname{diag}\left(\alpha_1, \frac{\alpha_1\alpha_2}{\beta_1}, \frac{\alpha_1\alpha_2\alpha_3}{\beta_1\beta_2}, \ldots\right),$$
$$D_{\mathrm{R}} = \operatorname{diag}\left(1, \frac{\beta_1}{\alpha_1}, \frac{\beta_1\beta_2}{\alpha_1\alpha_2}, \ldots\right).$$

Let $\alpha = \prod_{j=1}^{n} \max\{\alpha_j, 1/\alpha_j\}$ and $\beta = \prod_{j=1}^{n-1} \max\{\beta_j, 1/\beta_j\}$. Then

$$(\alpha\beta)^{-1} \le \left(\|D_{\mathrm{L}}^{-1}\|_2 \|D_{\mathrm{R}}^{-1}\|_2\right)^{-1} \le \|D_{\mathrm{L}}\|_2 \|D_{\mathrm{R}}\|_2 \le \alpha\beta,$$

and Fact 1 and Fact 2 apply. Now if all $1 - \varepsilon \le \alpha_i, \beta_j \le 1 + \varepsilon$, then $(1-\varepsilon)^{2n-1} \le (\alpha\beta)^{-1} \le (\alpha\beta) \le (1+\varepsilon)^{2n-1}$.

2. Consider block partitioned matrices

$$B = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix},$$
$$\widetilde{B} = \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix} = B \begin{bmatrix} I & -B_{11}^{-1} B_{12} \\ 0 & I \end{bmatrix} = B D_{\mathrm{R}}.$$

By Fact 2, $\zeta(\sigma_j, \widetilde{\sigma}_j) \le \frac{1}{2} \|B_{11}^{-1} B_{12}\|_2$. Interesting cases are when $\|B_{11}^{-1} B_{12}\|_2$ is tiny enough to be treated as zero and so $\mathrm{SV}(\widetilde{B})$ approximates $\mathrm{SV}(B)$ well. This situation occurs in computing the SVD of a bidiagonal matrix.

Author Note: Supported in part by the National Science Foundation under Grant No. DMS-0510664.

References

[BDD00] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst (Eds). *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*. SIAM, Philadelphia, 2000.

[BD90] J. Barlow and J. Demmel. Computing accurate eigensystems of scaled diagonally dominant matrices. *SIAM J. Numer. Anal.*, 27:762–791, 1990.

[Bar00] A. Barrlund. The p-relative distance is a metric. *SIAM J. Matrix Anal. Appl.*, 21(2):699–702, 2000.

[Bau85] H. Baumgärtel. *Analytical Perturbation Theory for Matrices and Operators*. Birkhäuser, Basel, 1985.

[Bha96] R. Bhatia. *Matrix Analysis*. Graduate Texts in Mathematics, Vol. 169. Springer, New York, 1996.

[BKL97] R. Bhatia, F. Kittaneh, and R.-C. Li. Some inequalities for commutators and an application to spectral variation. II. *Lin. Multilin. Alg.*, 43(1-3):207–220, 1997.

[CG00] F. Chatelin and S. Gratton. On the condition numbers associated with the polar factorization of a matrix. *Numer. Lin. Alg. Appl.*, 7:337–354, 2000.

[DK70] C. Davis and W. Kahan. The rotation of eigenvectors by a perturbation. III. *SIAM J. Numer. Anal.*, 7:1–46, 1970.

[DK90] J. Demmel and W. Kahan. Accurate singular values of bidiagonal matrices. *SIAM J. Sci. Statist. Comput.*, 11:873–912, 1990.

[DV92] J. Demmel and K. Veselić. Jacobi's method is more accurate than QR. *SIAM J. Matrix Anal. Appl.*, 13:1204–1245, 1992.

[EI95] S.C. Eisenstat and I.C.F. Ipsen. Relative perturbation techniques for singular value problems. *SIAM J. Numer. Anal.*, 32:1972–1988, 1995.

[HJ85] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[KPJ82] W. Kahan, B.N. Parlett, and E. Jiang. Residual bounds on approximate eigensystems of nonnormal matrices. *SIAM J. Numer. Anal.*, 19:470–484, 1982.

[Kat70] T. Kato. *Perturbation Theory for Linear Operators*, 2nd ed., Springer-Verlag, Berlin, 1970.

[Kit86] F. Kittaneh. Inequalities for the schatten p-norm. III. *Commun. Math. Phys.*, 104:307–310, 1986.

[LL05] Chi-Kwong Li and Ren-Cang Li. A note on eigenvalues of perturbed Hermitian matrices. *Lin. Alg. Appl.*, 395:183–190, 2005.

[LM99] Chi-Kwong Li and R. Mathias. The Lidskii–Mirsky–Wielandt theorem — additive and multiplicative versions. *Numer. Math.*, 81:377–413, 1999.

[Li88] Ren-Cang Li. A converse to the Bauer-Fike type theorem. *Lin. Alg. Appl.*, 109:167–178, 1988.

[Li93] Ren-Cang Li. Bounds on perturbations of generalized singular values and of associated subspaces. *SIAM J. Matrix Anal. Appl.*, 14:195–234, 1993.

[Li94] Ren-Cang Li. On perturbations of matrix pencils with real spectra. *Math. Comp.*, 62:231–265, 1994.

[Li95] Ren-Cang Li. New perturbation bounds for the unitary polar factor. *SIAM J. Matrix Anal. Appl.*, 16:327–332, 1995.

[Li97] Ren-Cang Li. Relative perturbation bounds for the unitary polar factor. *BIT*, 37:67–75, 1997.

[Li98] Ren-Cang Li. Relative perturbation theory: I. Eigenvalue and singular value variations. *SIAM J. Matrix Anal. Appl.*, 19:956–982, 1998.

[Li99a] Ren-Cang Li. Relative perturbation theory: II. Eigenspace and singular subspace variations. *SIAM J. Matrix Anal. Appl.*, 20:471–492, 1999.

[Li99b] Ren-Cang Li. A bound on the solution to a structured Sylvester equation with an application to relative perturbation theory. *SIAM J. Matrix Anal. Appl.*, 21:440–445, 1999.

[Li03] Ren-Cang Li. On perturbations of matrix pencils with real spectra, a revisit. *Math. Comp.*, 72:715–728, 2003.

[Li05] Ren-Cang Li. Relative perturbation bounds for positive polar factors of graded matrices. *SIAM J. Matrix Anal. Appl.*, 27:424–433, 2005.

[LS02] W. Li and W. Sun. Perturbation bounds for unitary and subunitary polar factors. *SIAM J. Matrix Anal. Appl.*, 23:1183–1193, 2002.

[Mat93] R. Mathias. Perturbation bounds for the polar decomposition. *SIAM J. Matrix Anal. Appl.*, 14:588–597, 1993.

[Pai84] C.C. Paige. A note on a result of Sun Ji-Guang: sensitivity of the CS and GSV decompositions. *SIAM J. Numer. Anal.*, 21:186–191, 1984.

[Par98] B.N. Parlett. *The Symmetric Eigenvalue Problem*. SIAM, Philadelphia, 1998.

[SS90] G.W. Stewart and Ji-Guang Sun. *Matrix Perturbation Theory*. Academic Press, Boston, 1990.

[Sun83] Ji-Guang Sun. Perturbation analysis for the generalized singular value decomposition. *SIAM J. Numer. Anal.*, 20:611–625, 1983.

[Sun91] Ji-Guang Sun. Eigenvalues of Rayleigh quotient matrices. *Numer. Math.*, 59:603–614, 1991.

[Sun96] Ji-Guang Sun. On the variation of the spectrum of a normal matrix. *Lin. Alg. Appl.*, 246:215–223, 1996.

[Sun98] Ji-Guang Sun. Stability and accuracy, perturbation analysis of algebraic eigenproblems. Technical Report UMINF 98-07, Department of Computer Science, Umeå Univeristy, Sweden, 1998.

[Van76] C.F. Van Loan. Generalizing the singular value decomposition. *SIAM J. Numer. Anal.*, 13:76–83, 1976.

[VS93] Krešimir Veselić and Ivan Slapničar. Floating-point perturbations of Hermitian matrices. *Lin. Alg. Appl.*, 195:81–116, 1993.