17

Singular Values and Singular Value Inequalities

Roy Mathias
University of Birmingham

17.1 Definitions and Characterizations

Singular values and the singular value decomposition are defined in Chapter 5.6. Additional information on computation of the singular value decomposition can be found in Chapter 45. A brief history of the singular value decomposition and early references can be found in [HJ91, Chap. 3].

Throughout this chapter, $q = \min\{m, n\}$, and if $A \in \mathbb{C}^{n\times n}$ has real eigenvalues, then they are ordered $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$.

Definitions:

For $A \in \mathbb{C}^{m\times n}$, define the **singular value vector** $\mathrm{sv}(A) = (\sigma_1(A), \ldots, \sigma_q(A))$.

For $A \in \mathbb{C}^{m\times n}$, define $r_1(A) \geq \cdots \geq r_m(A)$ and $c_1(A) \geq \cdots \geq c_n(A)$ to be the ordered Euclidean row and column lengths of A, that is, the square roots of the ordered diagonal entries of AA^* and A^*A.

For $A \in \mathbb{C}^{m\times n}$ define $|A|_{pd} = (A^*A)^{1/2}$. This is called the **spectral absolute value** of A. (This is also called the **absolute value**, but the latter term will not be used in this chapter due to potential confusion with the entry-wise absolute value of A, denoted $|A|$.)

A **polar decomposition** or **polar form** of the matrix $A \in \mathbb{C}^{m\times n}$ with $m \geq n$ is a factorization $A = UP$, where $P \in \mathbb{C}^{n\times n}$ is positive semidefinite and $U \in \mathbb{C}^{m\times n}$ satisfies $U^*U = I_n$.

Facts:

The following facts can be found in most books on matrix theory, for example [HJ91, Chap. 3] or [Bha97].

1. Take $A \in \mathbb{C}^{m\times n}$, and set
$$B = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}.$$
Then $\sigma_i(A) = \sigma_i(B)$ for $i = 1, \ldots, q$ and $\sigma_i(B) = 0$ for $i > q$. We may choose the zero blocks in B to ensure that B is square. In this way we can often generalize results on the singular values of square matrices to rectangular matrices. For simplicity of exposition, in this chapter we will sometimes state a result for square matrices rather than the more general result for rectangular matrices.
2. (Unitary invariance) Take $A \in \mathbb{C}^{m\times n}$. Then for any unitary $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$,
$$\sigma_i(A) = \sigma_i(UAV), \quad i = 1, 2, \ldots, q.$$
3. Take $A, B \in \mathbb{C}^{m\times n}$. There are unitary matrices $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ such that $A = UBV$ if and only if $\sigma_i(A) = \sigma_i(B)$, $i = 1, 2, \ldots, q$.
4. Let $A \in \mathbb{C}^{m\times n}$. Then $\sigma_i^2(A) = \lambda_i(AA^*) = \lambda_i(A^*A)$ for $i = 1, 2, \ldots, q$.
5. Let $A \in \mathbb{C}^{m\times n}$. Let $\mathcal{S}_i$ denote the set of subspaces of $\mathbb{C}^n$ of dimension i. Then for $i = 1, 2, \ldots, q$,
$$\sigma_i(A) = \min_{\mathcal{X}\in\mathcal{S}_{n-i+1}} \max_{\mathbf{x}\in\mathcal{X}, \|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2 = \min_{\mathcal{Y}\in\mathcal{S}_{i-1}} \max_{\mathbf{x}\perp\mathcal{Y}, \|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2,$$
$$\sigma_i(A) = \max_{\mathcal{X}\in\mathcal{S}_{i}} \min_{\mathbf{x}\in\mathcal{X}, \|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2 = \max_{\mathcal{Y}\in\mathcal{S}_{n-i}} \min_{\mathbf{x}\perp\mathcal{Y}, \|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2.$$
6. Let $A \in \mathbb{C}^{m\times n}$and define the Hermitian matrix
$$J = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix} \in \mathbb{C}^{m+n, m+n}.$$
The eigenvalues of J are $\pm\sigma_1(A), \ldots, \pm\sigma_q(A)$ together with $|m - n|$ zeros. The matrix J is called the Jordan–Wielandt matrix. Its use allows one to deduce singular value results from results for eigenvalues of Hermitian matrices.
7. Take $m \geq n$ and $A \in \mathbb{C}^{m\times n}$. Let $A = UP$ be a polar decomposition of A. Then $\sigma_i(A) = \lambda_i(P)$, $i = 1, 2, \ldots, q$.
8. Let $A \in \mathbb{C}^{m\times n}$ and $1 \leq k \leq q$. Then
$$\sum_{i=1}^{k} \sigma_i(A) = \max\{\text{Re tr } U^*AV : U \in \mathbb{C}^{m\times k}, V \in \mathbb{C}^{n\times k,} U^*U = V^*V = I_k\},$$
$$\prod_{i=1}^{k} \sigma_i(A) = \max\{|\det U^*AV| : U \in \mathbb{C}^{m\times k}, V \in \mathbb{C}^{n\times k},\ U^*U = V^*V = I_k\}.$$
If $m = n$, then
$$\sum_{i=1}^{n} \sigma_i(A) = \max\left\{\sum_{i=1}^{n} |(U^*AU)_{ii}| : U \in \mathbb{C}^{n\times n},\ U^*U = I_n\right\}.$$
We cannot replace the n by a general $k \in \{1, \ldots, n\}$.

9. Let $A \in \mathbb{C}^{m\times n}$. A yields
   (a) $\sigma_i(A^T) = \sigma_i(A^*) = \sigma_i(\bar{A}) = \sigma_i(A)$, for $i = 1, 2, \ldots, q$.
   (b) Let $k = \text{rank}(A)$. Then $\sigma_i(A^\dagger) = \sigma_{k-i+1}^{-1}(A)$ for $i = 1, \ldots, k$, and $\sigma_i(A^\dagger) = 0$ for $i = k+1, \ldots, q$. In particular, if $m = n$ and A is invertible, then

$$\sigma_i(A^{-1}) = \sigma_{n-i+1}^{-1}(A), \quad i = 1, \ldots, n.$$

   (c) For any $j \in \mathbb{N}$

$$\sigma_i((A^*A)^j) = \sigma_i^{2j}(A), \quad i = 1, \ldots, q;$$
$$\sigma_i((A^*A)^j A^*) = \sigma_i(A(A^*A)^j) = \sigma_i^{2j+1}(A) \quad i = 1, \ldots, q.$$

10. Let UP be a polar decomposition of $A \in \mathbb{C}^{m\times n}$ $(m \geq n)$. The positive semidefinite factor P is uniquely determined and is equal to $|A|_{pd}$. The factor U is uniquely determined if A has rank n. If A has singular value decomposition $A = U_1 \Sigma U_2^*$ $(U_1 \in \mathbb{C}^{m\times n}, U_2 \in \mathbb{C}^{n\times n})$, then $P = U_2 \Sigma U_2^*$, and U may be taken to be $U_1 U_2^*$.
11. Take $A, U \in \mathbb{C}^{n\times n}$ with U unitary. Then $A = U|A|_{pd}$ if and only if $A = |A^*|_{pd} U$.

Examples:

1. Take

$$A = \begin{bmatrix} 11 & -3 & -5 & 1 \\ 1 & -5 & -3 & 11 \\ -5 & 1 & 11 & -3 \\ -3 & 11 & 1 & -5 \end{bmatrix}.$$

The singular value decomposition of A is $A = U\Sigma V^*$, where $\Sigma = \text{diag}(20, 12, 8, 4)$, and

$$U = \frac{1}{2}\begin{bmatrix} -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad V = \frac{1}{2}\begin{bmatrix} -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}.$$

The singular values of A are 20, 12, 8, 4. Let Q denote the permutation matrix that takes (x_1, x_2, x_3, x_4) to (x_1, x_4, x_3, x_2). Let $P = |A|_{pd} = QA$. The polar decomposition of A is $A = QP$. (To see this, note that a permutation matrix is unitary and that P is positive definite by Geršchgorin's theorem.) Note also that $|A|_{pd} \neq |A^*|_{pd} = AQ$.

17.2 Singular Values of Special Matrices

In this section, we present some matrices where the singular values (or some of the singular values) are known, and facts about the singular values of certain structured matrices.

Facts:

The following results can be obtained by straightforward computations if no specific reference is given.

1. Let $D = \text{diag}(\alpha_1, \ldots, \alpha_n)$, where the α_i are integers, and let H_1 and H_2 be Hadamard matrices. (See Chapter 32.2.) Then the matrix $H_1 D H_2$ has integer entries and has integer singular values $n|\alpha_1|, \ldots, n|\alpha_n|$.

2. (2×2 *matrix*) Take $A \in \mathbb{C}^{2\times 2}$. Set $D = |\det(A)|^2$, $N = \|A\|_F^2$. The singular values of A are

$$\sqrt{\frac{N \pm \sqrt{N^2 - 4D}}{2}}.$$

3. Let $X \in \mathbb{C}^{m\times n}$ have singular values $\sigma_1 \geq \cdots \geq \sigma_q$ $(q = \min\{m, n\})$. Set

$$A = \begin{bmatrix} I & 2X \\ 0 & I \end{bmatrix} \in \mathbb{C}^{m+n,m+n}.$$

The $m + n$ singular values of A are

$$\sigma_1 + \sqrt{\sigma_1^2 + 1}, \ldots, \sigma_q + \sqrt{\sigma_q^2 + 1}, 1, \ldots, 1, \sqrt{\sigma_q^2 + 1} - \sigma_q, \ldots, \sqrt{\sigma_1^2 + 1} - \sigma_1.$$

4. [HJ91, Theorem 4.2.15] Let $A \in \mathbb{C}^{m_1\times n_1}$ and $B \in \mathbb{C}^{m_2\times n_2}$ have rank m and n. The nonzero singular values of $A \otimes B$ are $\sigma_i(A)\sigma_j(B)$, $i = 1, \ldots, m$, $j = 1, \ldots, n$.
5. Let $A \in \mathbb{C}^{n\times n}$ be normal with eigenvalues $\lambda_1, \ldots, \lambda_n$, and let p be a polynomial. Then the singular values of $p(A)$ are $|p(\lambda_k)|$, $k = 1, \ldots, n$. In particular, if A is a circulant with first row $a_0, \ldots, a_{n-1}$, then A has singular values

$$\left| \sum_{j=0}^{n-1} a_j e^{-2\pi i jk/n} \right|, \quad k = 1, \ldots, n.$$

6. Take $A \in \mathbb{C}^{n\times n}$ and nonzero $\mathbf{x} \in \mathbb{C}^n$. If $A\mathbf{x} = \lambda\mathbf{x}$ and $\mathbf{x}^*A = \lambda\mathbf{x}^*$, then $|\lambda|$ is a singular value of A. In particular, if A is doubly stochastic, then $\sigma_1(A) = 1$.
7. [Kit95] Let A be the companion matrix corresponding to the monic polynomial $p(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1t + a_0$. Set $N = 1 + \sum_{i=0}^{n-1} |a_i|^2$. The n singular values of A are

$$\sqrt{\frac{N + \sqrt{N^2 - 4|a_0|^2}}{2}}, 1, \ldots, 1, \sqrt{\frac{N - \sqrt{N^2 - 4|a_0|^2}}{2}}.$$

8. [Hig96, p. 167] Take $s, c \in \mathbb{R}$ such that $s^2 + c^2 = 1$. The matrix

$$A = \text{diag}(1, s, \ldots, s^{n-1}) \begin{bmatrix} 1 & -c & -c & \cdots & -c \\ & 1 & -c & \cdots & -c \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & -c \\ & & & & 1 \end{bmatrix}$$

is called a Kahan matrix. If c and s are positive, then $\sigma_{n-1}(A) = s^{n-2}\sqrt{1 + c}$.
9. [GE95, Lemma 3.1] Take $0 = d_1 < d_2 < \cdots < d_n$ and $0 \neq z_i \in \mathbb{C}$. Let

$$A = \begin{bmatrix} z_1 & & & \\ z_2 & d_2 & & \\ \vdots & & \ddots & \\ z_n & & & d_n \end{bmatrix}.$$

The singular values of A satisfy the equation

$$f(t) = 1 + \sum_{i=1}^{n} \frac{|z_i|^2}{d_i^2 - t^2} = 0$$

and exactly one lies in each of the intervals $(d_1, d_2), \ldots, (d_{n-1}, d_n), (d_n, d_n + \|z\|_2)$. Let $\sigma_i = \sigma_i(A)$. The left and right ith singular vectors of A are $\mathbf{u}/\|\mathbf{u}\|_2$ and $\mathbf{v}/\|\mathbf{v}\|_2$ respectively, where

$$\mathbf{u} = \left[\frac{z_1}{d_1^2 - \sigma_i^2}, \cdots, \frac{z_n}{d_n^2 - \sigma_i^2}\right]^T \quad \text{and} \quad \mathbf{v} = \left[-1, \frac{d_2 z_2}{d_2^2 - \sigma_i^2}, \cdots, \frac{d_n z_n}{d_n^2 - \sigma_i^2}\right]^T.$$

10. (*Bidiagonal*) Take

$$B = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ & \alpha_2 & \ddots & \\ & & \ddots & \beta_{n-1} \\ & & & \alpha_n \end{bmatrix} \in \mathbb{C}^{n \times n}.$$

If all the α_i and β_i are nonzero, then B is called an *unreduced bidiagonal matrix* and

(a) The singular values of B are distinct.

(b) The singular values of B depend only on the moduli of $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_{n-1}$.

(c) The largest singular value of B is a strictly increasing function of the modulus of each of the α_i and β_i.

(d) The smallest singular value of B is a strictly increasing function of the modulus of each of the α_i and a strictly decreasing function of the modulus of each of the β_i.

(e) (High relative accuracy) Take $\tau > 1$ and multiply one of the entries of B by τ to give $\hat{B}$. Then $\tau^{-1}\sigma_i(B) \leq \sigma_i(\hat{B}) \leq \tau\sigma_i(B)$.

11. [HJ85, Sec. 4.4, prob. 26] Let $A \in \mathbb{C}^{n \times n}$ be skew-symmetric (and possibly complex). The nonzero singular values of A occur in pairs.

17.3 Unitarily Invariant Norms

Throughout this section, $q = \min\{m, n\}$.

Definitions:

A vector norm $\|\cdot\|$ on $\mathbb{C}^{m \times n}$ is **unitarily invariant** (*u.i.*) if $\|A\| = \|UAV\|$ for any unitary $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ and any $A \in \mathbb{C}^{m \times n}$.

$\|\cdot\|_{UI}$ is used to denote a **general unitarily invariant norm**.

A function $g : \mathbb{R}^n \to \mathbb{R}_0^+$ is a **permutation invariant absolute norm** if it is a norm, and in addition $g(x_1, \ldots, x_n) = g(|x_1|, \ldots, |x_n|)$ and $g(\mathbf{x}) = g(P\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ and all permutation matrices $P \in \mathbb{R}^{n \times n}$. (Many authors call a permutation invariant absolute norm a symmetric gauge function.)

The **Ky Fan k norms** of $A \in \mathbb{C}^{m \times n}$ are

$$\|A\|_{K,k} = \sum_{i=1}^{k} \sigma_i(A), \quad k = 1, 2, \ldots, q.$$

The **Schatten-p norms** of $A \in \mathbb{C}^{m \times n}$ are

$$\|A\|_{S,p} = \left(\sum_{i=1}^{q} \sigma_i^p(A)\right)^{1/p} = \left(\operatorname{tr} |A|_{pd}^p\right)^{1/p} \quad 0 \leq p < \infty$$

$$\|A\|_{S,\infty} = \sigma_1(A).$$

The **trace norm** of $A \in \mathbb{C}^{m\times n}$ is

$$\|A\|_{\mathrm{tr}} = \sum_{i=1}^{q} \sigma_i(A) = \|A\|_{K,q} = \|A\|_{S,1} = \mathrm{tr}\,|A|_{pd}.$$

Other norms discussed in this section, such as the **spectral norm** $\|\cdot\|_2$ ($\|A\|_2 = \sigma_1(A) = \max_{\mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$) and the **Frobenius norm** $\|\cdot\|_F$ ($\|A\|_F = (\sum_{i=1}^{q} \sigma_i^2(A))^{1/2} = (\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2)^{1/2}$), are defined in Section 7.1. and discussed extensively in Chapter 37.

Warning: There is potential for considerable confusion. For example, $\|A\|_2 = \|A\|_{K,1} = \|A\|_{S,\infty}$, while $\|\cdot\|_\infty \neq \|\cdot\|_{S,\infty}$ (unless $m = 1$), and generally $\|A\|_2$, $\|A\|_{S,2}$ and $\|A\|_{K,2}$ are all different, as are $\|A\|_1$, $\|A\|_{S,1}$ and $\|A\|_{K,1}$. Nevertheless, many authors use $\|\cdot\|_k$ for $\|\cdot\|_{K,k}$ and $\|\cdot\|_p$ for $\|\cdot\|_{S,p}$.

Facts:

The following standard facts can be found in many texts, e.g., [HJ91, §3.5] and [Bha97, Chap. IV].

1. Let $\|\cdot\|$ be a norm on $\mathbb{C}^{m\times n}$. It is unitarily invariant if and only if there is a permutation invariant absolute norm g on $\mathbb{R}^q$ such that $\|A\| = g(\sigma_1(A), \ldots, \sigma_q(A))$ for all $A \in \mathbb{C}^{m\times n}$.
2. Let $\|\cdot\|$ be a unitarily invariant norm on $\mathbb{C}^{m\times n}$, and let g be the corresponding permutation invariant absolute norm g. Then the dual norms (see Chapter 37) satisfy $\|A\|^D = g^D(\sigma_1(A), \ldots, \sigma_q(A))$.
3. [HJ91, Prob. 3.5.18] The spectral norm and trace norm are duals, while the Frobenius norm is self dual. The dual of $\|\cdot\|_{S,p}$ is $\|\cdot\|_{S,\tilde{p}}$, where $1/p + 1/\tilde{p} = 1$ and

$$\|A\|_{K,k}^D = \max\left\{\|A\|_2, \frac{\|A\|_{tr}}{k}\right\}, \quad k = 1, \ldots, q.$$

4. For any $A \in \mathbb{C}^{m\times n}$, $q^{-1/2}\|A\|_F \leq \|A\|_2 \leq \|A\|_F$.
5. If $\|\cdot\|$ is a u.i. norm on $\mathbb{C}^{m\times n}$, then $N(A) = \|A^*A\|^{1/2}$ is a u.i. norm on $\mathbb{C}^{n\times n}$. A norm that arises in this way is called a *Q-norm.*
6. Let $A, B \in \mathbb{C}^{m\times n}$ be given. The following are equivalent

   (a) $\|A\|_{UI} \leq \|B\|_{UI}$ for all unitarily invariant norms $\|\cdot\|_{UI}$.

   (b) $\|A\|_{K,k} \leq \|B\|_{K,k}$ for $k = 1, 2, \ldots, q$.

   (c) $(\sigma_1(A), \ldots, \sigma_q(A)) \preceq_w (\sigma_1(B), \ldots, \sigma_q(B))$. ($\preceq_w$ is defined in Preliminaries)

   The equivalence of the first two conditions is Fan's Dominance Theorem.
7. The Ky–Fan-k norms can be represented in terms of an extremal problem involving the spectral norm and the trace norm. Take $A \in \mathbb{C}^{m\times n}$. Then

$$\|A\|_{K,k} = \min\{\|X\|_{\mathrm{tr}} + k\|Y\|_2 : X + Y = A\} \quad k = 1, \ldots, q.$$

8. [HJ91, Theorem 3.3.14] Take $A, B \in \mathbb{C}^{m\times n}$. Then

$$|\mathrm{tr} AB^*| \leq \sum_{i=1}^{q} \sigma_i(A)\sigma_i(B).$$

   This is an important result in developing the theory of unitarily invariant norms.

Examples:

1. The matrix A in Example 1 of Section 17.1 has singular values 20, 12, 8, and 4. So

$$\|A\|_2 = 20, \quad \|A\|_F = \sqrt{624}, \quad \|A\|_{tr} = 44;$$
$$\|A\|_{K,1} = 20, \quad \|A\|_{K,2} = 32, \quad \|A\|_{K,3} = 40, \quad \|A\|_{K,4} = 44;$$
$$\|A\|_{S,1} = 44, \quad \|A\|_{S,2} = \sqrt{624}, \quad \|A\|_{S,3} = \sqrt[3]{10304} = 21.7605, \quad \|A\|_{S,\infty} = 20.$$

17.4 Inequalities

Throughout this section, $q = \min\{m, n\}$ and if $A \in \mathbb{C}^{m\times n}$ has real eigenvalues, then they are ordered $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$.

Definitions:

Pinching is defined recursively. If

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in \mathbb{C}^{m\times n}, \quad B = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \in \mathbb{C}^{m\times n},$$

then B is a pinching of A. (Note that we do not require the A_{ii} to be square.) Furthermore, any pinching of B is a pinching of A.

For positive α, β, define the **measure of relative separation** $\chi(\alpha, \beta) = |\sqrt{\alpha/\beta} - \sqrt{\beta/\alpha}|$.

Facts:

The following facts can be found in standard references, for example [HJ91, Chap. 3], unless another reference is given.

1. (*Submatrices*) Take $A \in \mathbb{C}^{m\times n}$ and let B denote A with one of its rows or columns deleted. Then $\sigma_{i+1}(A) \leq \sigma_i(B) \leq \sigma_i(A), \quad i = 1, \dots, q-1$.
2. Take $A \in \mathbb{C}^{m\times n}$ and let B be A with a row *and* a column deleted. Then
$$\sigma_{i+2}(A) \leq \sigma_i(B) \leq \sigma_i(A), \quad i = 1, \dots, q-2.$$
The $i+2$ cannot be replaced by $i+1$. (Example 2)
3. Take $A \in \mathbb{C}^{m\times n}$ and let B be an $(m-k)\times(n-l)$ submatrix of A. Then
$$\sigma_{i+k+l}(A) \leq \sigma_i(B) \leq \sigma_i(A), \quad i = 1, \dots, q-(k+l).$$
4. Take $A \in \mathbb{C}^{m\times n}$ and let B be A with some of its rows and/or columns set to zero. Then $\sigma_i(B) \leq \sigma_i(A), \quad i = 1, \dots, q$.
5. Let B be a pinching of A. Then $\mathrm{sv}(B) \preceq_w \mathrm{sv}(A)$. The inequalities $\prod_{i=1}^k \sigma_i(B) \leq \prod_{i=1}^k \sigma_i(A)$ and $\sigma_k(B) \leq \sigma_k(A)$ are not necessarily true for $k > 1$. (Example 1)
6. (*Singular values of* $A+B$) Let $A, B \in \mathbb{C}^{m\times n}$.

   (a) $\mathrm{sv}(A+B) \preceq_w \mathrm{sv}(A) + \mathrm{sv}(B)$, or equivalently
$$\sum_{i=1}^k \sigma_i(A+B) \leq \sum_{i=1}^k \sigma_i(A) + \sum_{i=1}^k \sigma_i(B), \quad i = 1, \dots, q.$$

   (b) If $i + j - 1 \leq q$ and $i, j \in \mathbb{N}$, then $\sigma_{i+j-1}(A+B) \leq \sigma_i(A) + \sigma_j(B)$.

(c) We have the weak majorization $|\text{sv}(A+B) - \text{sv}(A)| \preceq_w \text{sv}(B)$ or, equivalently, if $1 \leq i_1 < \cdots < i_k \leq q$, then

$$\sum_{j=1}^{k} |\sigma_{i_j}(A+B) - \sigma_{i_j}(A)| \leq \sum_{j=1}^{k} \sigma_j(B),$$

$$\sum_{i=1}^{k} \sigma_{i_j}(A) - \sum_{j=1}^{k} \sigma_j(B) \leq \sum_{j=1}^{k} \sigma_{i_j}(A+B) \leq \sum_{i=1}^{k} \sigma_{i_j}(A) + \sum_{j=1}^{k} \sigma_j(B).$$

(d) [Tho75] (Thompson's Standard Additive Inequalities) If $1 \leq i_1 < \cdots < i_k \leq q$, $1 \leq i_1 < \cdots < i_k \leq q$ and $i_k + j_k \leq q + k$, then

$$\sum_{s=1}^{k} \sigma_{i_s + j_s - s}(A+B) \leq \sum_{s=1}^{k} \sigma_{i_s}(A) + \sum_{s=1}^{k} \sigma_{j_s}(B).$$

7. (*Singular values of* AB) Take $A, B \in \mathbb{C}^{n \times n}$.

(a) For all $k = 1, 2, \ldots, n$ and all $p > 0$, we have

$$\prod_{i=n}^{i=n-k+1} \sigma_i(A)\sigma_i(B) \leq \prod_{i=n}^{i=n-k+1} \sigma_i(AB),$$

$$\prod_{i=1}^{k} \sigma_i(AB) \leq \prod_{i=1}^{k} \sigma_i(A)\sigma_i(B),$$

$$\sum_{i=1}^{k} \sigma_i^p(AB) \leq \sum_{i=1}^{k} \sigma_i^p(A)\sigma_i^p(B).$$

(b) If $i, j \in \mathbb{N}$ and $i + j - 1 \leq n$, then $\sigma_{i+j-1}(AB) \leq \sigma_i(A)\sigma_j(B)$.

(c) $\sigma_n(A)\sigma_i(B) \leq \sigma_i(AB) \leq \sigma_1(A)\sigma_i(B)$, $i = 1, 2, \ldots, n$.

(d) [LM99] Take $1 \leq j_1 < \cdots < j_k \leq n$. If A is invertible and $\sigma_{j_i}(B) > 0$, then $\sigma_{j_i}(AB) > 0$ and

$$\prod_{i=n-k+1}^{n} \sigma_i(A) \leq \prod_{i=1}^{k} \max\left\{ \frac{\sigma_{j_i}(AB)}{\sigma_{j_i}(B)}, \frac{\sigma_{j_i}(B)}{\sigma_{j_i}(AB)} \right\} \leq \prod_{i=1}^{k} \sigma_i(A).$$

(e) [LM99] Take invertible $S, T \in \mathbb{C}^{n \times n}$. Set $\tilde{A} = SAT$. Let the singular values of A and $\tilde{A}$ be $\sigma_1 \geq \cdots \geq \sigma_n$ and $\tilde{\sigma}_1 \geq \cdots \geq \tilde{\sigma}_n$. Then

$$\|\text{diag}(\chi(\sigma_1, \tilde{\sigma}_1), , \ldots, \chi(\sigma_n, \tilde{\sigma}_n))\|_{UI} \leq \frac{1}{2}\left(\|S^* - S^{-1}\|_{UI} + \|T^* - T^{-1}\|_{UI} \right).$$

(f) [TT73] (Thompson's Standard Multiplicative Inequalities) Take $1 \leq i_1 < \cdots < i_m \leq n$ and $1 \leq j_1 < \cdots < j_m \leq n$. If $i_m + j_m \leq m + n$, then

$$\prod_{s=1}^{m} \sigma_{i_s + j_s - s}(AB) \leq \prod_{s=1}^{m} \sigma_{i_s}(A) \prod_{s=1}^{m} \sigma_{j_s}(B).$$

8. [Bha97, §IX.1] Take $A, B \in \mathbb{C}^{n \times n}$.

(a) If AB is normal, then

$$\prod_{i=1}^{k} \sigma_i(AB) \leq \prod_{i=1}^{k} \sigma_i(BA), \quad k = 1, \ldots, q,$$

and, consequently, $\text{sv}(AB) \preceq_w \text{sv}(BA)$, and $\|AB\|_{UI} \leq \|BA\|_{UI}$.

(b) If AB is Hermitian, then $\text{sv}(AB) \preceq_w \text{sv}(H(BA))$ and $\|AB\|_{UI} \leq \|H(BA)\|_{UI}$, where $H(X) = (X + X^*)/2$.

9. (*Term-wise singular value inequalities*) [Zha02, p. 28] Take $A, B \in \mathbb{C}^{m\times n}$. Then

$$2\sigma_i(AB^*) \leq \sigma_i(A^*A + B^*B), \quad i = 1, \ldots, q$$

and, more generally, if $p, \tilde{p} > 0$ and $1/p + 1/\tilde{p} = 1$, then

$$\sigma_i(AB^*) \leq \sigma_i\left(\frac{(A^*A)^{p/2}}{p} + \frac{(B^*B)^{\tilde{p}/2}}{\tilde{p}}\right) = \sigma_i\left(\frac{|A|^p_{pd}}{p} + \frac{|B|^{\tilde{p}}_{pd}}{\tilde{p}}\right).$$

The inequalities $2\sigma_1(A^*B) \leq \sigma_1(A^*A + B^*B)$ and $\sigma_1(A + B) \leq \sigma_1(|A|_{pd} + |B|_{pd})$ are not true in general (Example 3), but we do have

$$\|A^*B\|^2_{UI} \leq \|A^*A\|_{UI}\|B^*B\|_{UI}.$$

10. [Bha97, Prop. III.5.1] Take $A \in \mathbb{C}^{n\times n}$. Then $\lambda_i(A + A^*) \leq 2\sigma_i(A)$, $i = 1, 2, \ldots, n$.

11. [LM02] (*Block triangular matrices*) Let $A = \begin{bmatrix} R & 0 \\ S & T \end{bmatrix} \in \mathbb{C}^{n\times n}$ ($R \in \mathbb{C}^{p\times p}$) have singular values $\alpha_1 \geq \cdots \geq \alpha_n$. Let $k = \min\{p, n-p\}$. Then

(a) If $\sigma_{\min}(R) \geq \sigma_{\max}(T)$, then

$$\begin{aligned} \sigma_i(R) &\leq \alpha_i, \quad i = 1, \ldots, p \\ \alpha_i &\leq \sigma_{i-p}(T), \quad i = p+1, \ldots, n. \end{aligned}$$

(b) $(\sigma_1(S), \ldots, \sigma_k(S)) \preceq_w (\alpha_1 - \alpha_n, \cdots, \alpha_k - \alpha_{n-k+1})$.

(c) If A is invertible, then

$$(\sigma_1(T^{-1}SR^{-1}, \ldots, \sigma_k(T^{-1}SR^{-1}) \preceq_w \left(\alpha_n^{-1} - \alpha_1^{-1}, \cdots, \alpha_{n-k+1}^{-1} - \alpha_k^{-1}\right),$$

$$(\sigma_1(T^{-1}S), \ldots, \sigma_k(T^{-1}S)) \preceq_w \frac{1}{2}\left(\frac{\alpha_1}{\alpha_n} - \frac{\alpha_n}{\alpha_1}, \cdots, \frac{\alpha_k}{\alpha_{n-k+1}} - \frac{\alpha_{n-k+1}}{\alpha_k}\right).$$

12. [LM02] (*Block positive semidefinite matrices*) Let $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{bmatrix} \in \mathbb{C}^{n\times n}$ be positive definite with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. Assume $A_{11} \in \mathbb{C}^{p\times p}$. Set $k = \min\{p, n-p\}$. Then

$$\prod_{i=1}^{j} \sigma_i^2(A_{12}) \leq \prod_{i=1}^{j} \sigma_i(A_{11})\sigma_i(A_{22}), \quad j = 1, \ldots, k,$$

$$\left(\sigma_1\left(A_{11}^{-1/2}A_{12}\right), \ldots, \sigma_k\left(A_{11}^{-1/2}A_{12}\right)\right) \preceq_w \left(\sqrt{\lambda_1} - \sqrt{\lambda_n}, \ldots, \sqrt{\lambda_k} - \sqrt{\lambda_{n-k+1}}\right),$$

$$\left(\sigma_1\left(A_{11}^{-1}A_{12}\right), \ldots, \sigma_k\left(A_{11}^{-1}A_{12}\right)\right) \preceq_w \frac{1}{2}\left(\chi(\lambda_1, \lambda_n), \ldots, \chi(\lambda_k, \lambda_{n-k+1})\right).$$

If $k = n/2$, then

$$\|A_{12}\|^2_{UI} \leq \|A_{11}\|_{UI}\ \|A_{22}\|_{UI}.$$

13. (*Singular values and eigenvalues*) Let $A \in \mathbb{C}^{n\times n}$. Assume $|\lambda_1(A)| \geq \cdots \geq |\lambda_n(A)|$. Then

(a) $\prod_{i=1}^{k} |\lambda_i(A)| \leq \prod_{i=k}^{k} \sigma_i(A), \quad k = 1, \ldots, n$, with equality for $k = n$.

(b) Fix $p > 0$. Then for $k = 1, 2, \ldots, n$,

$$\sum_{i=1}^{k} |\lambda_i^p(A)| \le \sum_{i=1}^{k} \sigma_i^p(A).$$

Equality holds with $k = n$ if and only if equality holds for all $k = 1, 2, \ldots, n$, if and only if A is normal.

(c) [HJ91, p. 180] (Yamamoto's theorem) $\lim_{k\to\infty}(\sigma_i(A^k))^{1/k} = |\lambda_i(A)|, \quad i = 1, \ldots, n.$

14. [LM01] Let $\lambda_i \in \mathbb{C}$ and $\sigma_i \in \mathbb{R}_0^+$, $i = 1, \ldots, n$ be ordered in nonincreasing absolute value. There is a matrix A with eigenvalues $\lambda_1, \ldots, \lambda_n$ and singular values $\sigma_1, \ldots, \sigma_n$ if and only if

$$\prod_{i=1}^{k} |\lambda_i| \le \prod_{i=1}^{k} \sigma_i, \quad k = 1, \ldots, n, \quad \text{with equality for } k = n.$$

In addition:

(a) The matrix A can be taken to be upper triangular with the eigenvalues on the diagonal in any order.

(b) If the complex entries in $\lambda_1, \ldots, \lambda_n$ occur in conjugate pairs, then A may be taken to be in real Schur form, with the 1×1 and 2×2 blocks on the diagonal in any order.

(c) There is a finite construction of the upper triangular matrix in cases (a) and (b).

(d) If $n > 2$, then A cannot always be taken to be bidiagonal. (Example 5)

15. [Zha02, Chap. 2] (*Singular values of* $A \circ B$) Take $A, B \in \mathbb{C}^{n\times n}$.

(a) $\sigma_i(A \circ B) \le \min\{r_i(A), c_i(B)\} \cdot \sigma_1(B), \quad i = 1, 2, \ldots, n.$

(b) We have the following weak majorizations:

$$\sum_{i=1}^{k} \sigma_i(A \circ B) \le \sum_{i=1}^{k} \min\{r_i(A), c_i(A)\}\sigma_i(B), \quad k = 1, \ldots, n,$$

$$\sum_{i=1}^{k} \sigma_i(A \circ B) \le \sum_{i=1}^{k} \sigma_i(A)\sigma_i(B), \quad k = 1, \ldots, n,$$

$$\prod_{i=1}^{k} \sigma_i^2(A \circ B) \le \prod_{i=1}^{k} \sigma_i((A^*A) \circ (B^*B)), \quad k = 1, \ldots, n.$$

(c) Take $X, Y \in \mathbb{C}^{n\times n}$. If $A = X^*Y$, then we have the weak majorization

$$\sum_{i=1}^{k} \sigma_i(A \circ B) \le \sum_{i=1}^{k} c_i(X)c_i(Y)\sigma_i(B), \quad k = 1, \ldots, n.$$

(d) If B is positive semidefinite with diagonal entries $b_{11} \ge \cdots \ge b_{nn}$, then

$$\sum_{i=1}^{k} \sigma_i(A \circ B) \le \sum_{i=1}^{k} b_{ii}\sigma_i(A), \quad k = 1, \ldots, n.$$

(e) If both A and B are positive definite, then so is $A \circ B$ (Schur product theorem). In this case the singular values of A, B and $A \circ B$ are their eigenvalues and BA has positive eigenvalues and we have the weak multiplicative majorizations

$$\prod_{i=k}^{n} \lambda_i(B)\lambda_i(A) \le \prod_{i=k}^{n} b_{ii}\lambda_i(A) \le \prod_{i=k}^{n} \lambda_i(BA) \le \prod_{i=k}^{n} \lambda_i(A \circ B), \quad k = 1, 2, \ldots, n.$$

The inequalities are still valid if we replace $A \circ B$ by $A \circ B^T$. (Note B^T is not necessarily the same as $B^* = B$.)

16. Let $A \in \mathbb{C}^{m\times n}$. The following are equivalent:

    (a) $\sigma_1(A \circ B) \le \sigma_1(B)$ for all $B \in \mathbb{C}^{m\times n}$.

    (b) $\sum_{i=1}^{k} \sigma_i(A \circ B) \le \sum_{i=1}^{k} \sigma_i(B)$ for all $B \in \mathbb{C}^{m\times n}$ and all $k = 1, \dots, q$.

    (c) There are positive semidefinite $P \in \mathbb{C}^{n\times n}$ and $Q \in \mathbb{C}^{m\times m}$ such that

$$\begin{bmatrix} P & A \\ A^* & Q \end{bmatrix}$$

    is positive semidefinite, and has diagonal entries at most 1.

17. (*Singular values and matrix entries*) Take $A \in \mathbb{C}^{m\times n}$. Then

$$\left(|a_{11}|^2, |a_{12}|^2, \dots, |a_{mn}|^2\right) \preceq \left(\sigma_1^2(A), \dots, \sigma_q^2(A), 0, \dots, 0\right),$$

$$\sum_{i=1}^{q} \sigma_i^p(A) \le \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^p, \quad 0 \le p \le 2,$$

$$\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^p \le \sum_{i=1}^{q} \sigma_i^p(A), \quad 2 \le p < \infty.$$

    If $\sigma_1(A) = |a_{ij}|$, then all the other entries in row i and column j of A are 0.

18. Take $\sigma_1 \ge \cdots \ge \sigma_n \ge 0$ and $\alpha_1 \ge \cdots \ge \alpha_n \ge 0$. Then

$$\exists A \in \mathbb{R}^{n\times n} \text{ s.t. } \sigma_i(A) = \sigma_i \quad \text{and} \quad c_i(A) = \alpha_i \quad \Leftrightarrow \quad \left(\alpha_1^2, \dots, \alpha_n^2\right) \preceq \left(\sigma_1^2, \dots, \sigma_n^2\right).$$

    This statement is still true if we replace $\mathbb{R}^{n\times n}$ by $\mathbb{C}^{n\times n}$ and/or $c_i(\cdot)$ by $r_i(\cdot)$.

19. Take $A \in \mathbb{C}^{n\times n}$. Then

$$\prod_{i=k}^{n} \sigma_i(A) \le \prod_{i=k}^{n} c_i(A), \quad k = 1, 2, \dots, n.$$

    The case $k = 1$ is Hadamard's Inequality: $|\det(A)| \le \prod_{i=1}^{n} c_i(A)$.

20. [Tho77] Take $F = \mathbb{C}$ or $\mathbb{R}$ and $d_1, \dots, d_n \in F$ such that $|d_1| \ge \cdots \ge |d_n|$, and $\sigma_1 \ge \cdots \ge \sigma_n \ge 0$. There is a matrix $A \in F^{n\times n}$ with diagonal entries $d_1, \dots, d_n$ and singular values $\sigma_1, \dots, \sigma_n$ if and only if

$$(|d_1|, \dots, |d_n|) \preceq_w (\sigma_1(A), \dots, \sigma_n(A)) \quad \text{and} \quad \sum_{j=1}^{n-1} |d_j| - |d_n| \le \sum_{j=1}^{n-1} \sigma_j(A) - \sigma_n(A).$$

21. (*Nonnegative matrices*) Take $A = [a_{ij}] \in \mathbb{C}^{m\times n}$.

    (a) If $B = [|a_{ij}|]$, then $\sigma_1(A) \le \sigma_1(B)$.

    (b) If A and B are real and $0 \le a_{ij} \le b_{ij} \;\; \forall \;\; i, j$, then $\sigma_1(A) \le \sigma_1(B)$. The condition $0 \le a_{ij}$ is essential. (Example 4)

    (c) The condition $0 \le b_{ij} \le 1 \;\; \forall \;\; i, j$ does not imply $\sigma_1(A \circ B) \le \sigma_1(A)$. (Example 4)

22. (*Bound on* σ_1) Let $A \in \mathbb{C}^{m\times n}$. Then $\|A\|_2 = \sigma_1(A) \le \sqrt{\|A\|_1 \|A\|_\infty}$.
23. [Zha99] (*Cartesian decomposition*) Let $C = A + iB \in \mathbb{C}^{n\times n}$, where A and B are Hermitian. Let A, B, C have singular values $\alpha_j, \beta_j, \gamma_i$, $j = 1, \dots, n$. Then

$$(\gamma_1, \dots, \gamma_n) \preceq_w \sqrt{2}(|\alpha_1 + i\beta_1|, \dots, |\alpha_n + i\beta_n|) \preceq_w 2(\gamma_1, \dots, \gamma_n).$$

Examples:

1. Take

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

   Then B is a pinching of A, and C is a pinching of both A and B. The matrices A, B, C have singular values $\alpha = (3, 0, 0)$, $\beta = (2, 1, 0)$, and $\gamma = (1, 1, 1)$. As stated in Fact 5, $\gamma \preceq_w \beta \preceq_w \alpha$. In fact, since the matrices are all positive semidefinite, we may replace $\preceq_w$ by $\preceq$. However, it is not true that $\gamma_i \le \alpha_i$ except for $i = 1$. Nor is it true that $|\det(C)| \le |\det(A)|$.

2. The matrices

$$A = \begin{bmatrix} 11 & -3 & -5 & 1 \\ 1 & -5 & -3 & 11 \\ -5 & 1 & 11 & -3 \\ -3 & 11 & 1 & -5 \end{bmatrix}, \quad B = \begin{bmatrix} 11 & -3 & -5 & 1 \\ 1 & -5 & -3 & 11 \\ -5 & 1 & 11 & -3 \end{bmatrix}, \quad C = \begin{bmatrix} 11 & -3 & -5 \\ 1 & -5 & -3 \\ -5 & 1 & 11 \end{bmatrix}$$

   have singular values $\alpha = (20, 12, 8, 4)$, $\beta = (17.9, 10.5, 6.0)$, and $\gamma = (16.7, 6.2, 4.5)$ (to 1 decimal place). The singular values of B interlace those of A ($\alpha_4 \le \beta_3 \le \alpha_3 \le \beta_2 \le \alpha_2 \le \beta_1 \le \alpha_1$), but those of C do not. In particular, $\alpha_3 \not\le \gamma_2$. It is true that $\alpha_{i+2} \le \gamma_i \le \alpha_i$ $(i = 1, 2)$.

3. Take

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

   Then $\|A + B\|_2 = \sigma_1(A + B) = 2 \not\le \sqrt{2} = \sigma_1(|A|_{pd} + |B|_{pd}) = \|\,|A|_{pd} + |B|_{pd}\,\|_2$. Also, $2\sigma_1(A^*B) = 4 \not\le 2 = \sigma_1(A^*A + B^*B)$.

4. Setting entries of a matrix to zero can increase the largest singular value. Take

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

   Then $\sigma_1(A) = \sqrt{2} < (1 + \sqrt{5})/2 = \sigma_1(B)$.

5. A bidiagonal matrix B cannot have eigenvalues 1, 1, 1 and singular values 1/2, 1/2, 4. If B is unreduced bidiagonal, then it cannot have repeated singular values. (See Fact 10, section 17.2.) However, if B were reduced, then it would have a singular value equal to 1.

17.5 Matrix Approximation

Recall that $\|\cdot\|_{UI}$ denotes a general unitarily invariant norm, and that $q = \min\{m, n\}$.

Facts:

The following facts can be found in standard references, for example, [HJ91, Chap. 3], unless another reference is given.

1. (*Best rank k approximation.*) Let $A \in \mathbb{C}^{m\times n}$ and $1 \le k \le q - 1$. Let $A = U\Sigma V^*$ be a singular value decomposition of A. Let $\tilde{\Sigma}$ be equal to Σ except that $\tilde{\Sigma}_{ii} = 0$ for $i > k$, and let $\tilde{A} = U\tilde{\Sigma}V^*$. Then $\operatorname{rank}(\tilde{A}) \le k$, and

$$\|\Sigma - \tilde{\Sigma}\|_{UI} = \|A - \tilde{A}\|_{UI} = \min\{\|A - B\|_{UI} : \operatorname{rank}(B) \le k\}.$$

In particular, for the spectral norm and the Frobenius norm, we have

$$\sigma_{k+1}(A) = \min\{\|A - B\|_2 : \text{rank}(B) \leq k\},$$
$$\left(\sum_{i=k+1}^{q} \sigma_{k+1}^2(A)\right)^{1/2} = \min\{\|A - B\|_F : \text{rank}(B) \leq k\}.$$

2. [Bha97, p. 276] (*Best unitary approximation*) Take $A, W \in \mathbb{C}^{n\times n}$ with W unitary. Let $A = UP$ be a polar decomposition of A. Then

$$\|A - U\|_{UI} \leq \|A - W\|_{UI} \leq \|A + U\|_{UI}.$$

3. [GV96, §12.4.1] [HJ85, Ex. 7.4.8] (*Orthogonal Procrustes problem*) Let $A, B \in \mathbb{C}^{m\times n}$. Let B^*A have a polar decomposition $B^*A = UP$. Then

$$\|A - BU\|_F = \min\{\|A - BW\|_F : W \in \mathbb{C}^{n\times n}, W^*W = I\}.$$

   This result is not true if $\|\cdot\|_F$ is replaced by $\|\cdot\|_{UI}$ ([Mat93, §4]).
4. [Hig89] (*Best PSD approximation*) Take $A \in \mathbb{C}^{n\times n}$. Set $A_H = (A + A^*)/2$, $B = (A_H + |A_H|)/2)$. Then B is positive semidefinite and is the unique solution to

$$\min\{\|A - X\|_F : X \in \mathbb{C}^{n\times n}, X \in \text{PSD}\}.$$

   There is also a formula for the best PSD approximation in the spectral norm.
5. Let $A, B \in \mathbb{C}^{m\times n}$ have singular value decompositions $A = U_A\Sigma_A V_A^*$ and $B = U_B\Sigma_B V_B^*$. Let $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ be any unitary matrices. Then

$$\|\Sigma_A - \Sigma_B\|_{UI} \leq \|A - UBV^*\|_{UI}.$$

17.6 Characterization of the Eigenvalues of Sums of Hermitian Matrices and Singular Values of Sums and Products of General Matrices

There are necessary and sufficient conditions for three sets of numbers to be the eigenvalues of Hermitian $A, B, C = A + B \in \mathbb{C}^{n\times n}$, or the singular values of $A, B, C = A + B \in \mathbb{C}^{m\times n}$, or the singular values of nonsingular $A,\ B,\ C = AB \in \mathbb{C}^{n\times n}$. The key results in this section were first proved by Klyachko ([Kly98]) and Knutson and Tao ([KT99]). The results presented here are from a survey by Fulton [Ful00]. Bhatia has written an expository paper on the subject ([Bha01]).

Definitions:

The inequalities are in terms of the sets T_r^n of triples (I, J, K) of subsets of $\{1, \ldots, n\}$ of the same cardinality r, defined by the following inductive procedure. Set

$$U_r^n = \left\{(I, J, K) \;\middle|\; \sum_{i\in I} i + \sum_{j\in J} j = \sum_{k\in K} k + r(r+1)/2\right\}.$$

When $r = 1$, set $T_1^n = U_1^n$. In general,

$$T_r^n = \Big\{(I, J, K) \in U_r^n \mid \text{for all } p < r \text{ and all}$$

$$(F, G, H) \text{ in } T_p^r, \sum_{f\in F} i_f + \sum_{g\in G} j_g \leq \sum_{h\in H} k_h + p(p+1)/2\Big\}.$$

In this section, the vectors $\alpha,\ \beta,\ \gamma$ will have real entries ordered in nonincreasing order.

Facts:

The following facts are in [Ful00]:

1. A triple (α, β, γ) of real n-vectors occurs as eigenvalues of Hermitian $A, B, C = A + B \in \mathbb{C}^{n\times n}$ if and only if $\sum \gamma_i = \sum \alpha_i + \sum \beta_i$ and the inequalities
$$\sum_{k\in K} \gamma_k \le \sum_{i\in I} \alpha_i + \sum_{j\in J} \beta_j$$
hold for every (I, J, K) in T_r^n, for all $r < n$. Furthermore, the statement is true if $\mathbb{C}^{n\times n}$ is replaced by $\mathbb{R}^{n\times n}$.
2. Take Hermitian $A, B \in \mathbb{C}^{n\times n}$ (not necessarily PSD). Let the vectors of eigenvalues of A, B, $C = A + B$ be α, β, and γ. Then we have the (nonlinear) inequality
$$\min_{\pi\in S_n} \prod_{i=1}^{n}(\alpha_i + \beta_{\pi(i)}) \le \prod_{i=1}^{n} \gamma_i \le \max_{\pi\in S_n} \prod_{i=1}^{n}(\alpha_i + \beta_{\pi(i)}).$$
3. Fix m, n and set $q = \min\{m, n\}$. For any subset X of $\{1, \ldots, m+n\}$, define $X_q = \{i : i \in X, i \le q\}$ and $X_q' = \{i : i \le q,\ m + n + 1 - i \in X\}$. A triple (α, β, γ) occurs as the singular values of $A, B, C = A + B \in \mathbb{C}^{m\times n}$, if and only if the inequalities
$$\sum_{k\in K_q} \gamma_k - \sum_{k\in K_q'} \gamma_k \le \sum_{i\in I} \alpha_i - \sum_{i\in I_q'} \alpha_i + \sum_{j\in J_q} \beta_j - \sum_{j\in J_q'} \beta_j$$
are satisfied for all (I, J, K) in T_r^{m+n}, for all $r < m+n$. This statement is not true if $\mathbb{C}^{m\times n}$ is replaced by $\mathbb{R}^{m\times n}$. (See Example 1.)
4. A triple of positive real n-vectors (α, β, γ) occurs as the singular values of n by n matrices A,B, $C = AB \in \mathbb{C}^{n\times n}$ if and only if $\gamma_1 \cdots \gamma_n = \alpha_1 \cdots \alpha_n \beta_1 \cdots \beta_n$ and
$$\prod_{k\in K} \gamma_k \le \prod_{i\in I} \alpha_i \cdot \prod_{j\in J} \beta_j$$
for all (I, J, K) in T_r^n, and all $r < n$. This statement is still true if $\mathbb{C}^{n\times n}$ is replaced by $\mathbb{R}^{n\times n}$.

Example:

1. There are $A, B, C = A + B \in \mathbb{C}^{2\times 2}$ with singular values $(1, 1)$, $(1, 0)$, and $(1, 1)$, but there are no $A, B, C = A + B \in \mathbb{R}^{2\times 2}$ with these singular values.
   In the complex case, take $A = \text{diag}(1, 1/2 + (\sqrt{3}/2)i)$, $B = \text{diag}(0, -1)$.
   Now suppose that A and B are real 2×2 matrices such that A and $C = A + B$ both have singular values $(1, 1)$. Then A and C are orthogonal. Consider $BC^T = AC^T - CC^T = AC^T - I$. Because AC^T is real, it has eigenvalues $\alpha, \bar{\alpha}$ and so BC^T has eigenvalues $\alpha - 1, \bar{\alpha} - 1$. Because AC^T is orthogonal, it is normal and, hence, so is BC^T, and so its singular values are $|\alpha - 1|$ and $|\bar{a} - 1|$, which are equal and, in particular, cannot be $(1, 0)$.

17.7 Miscellaneous Results and Generalizations

Throughout this section F can be taken to be either $\mathbb{R}$ or $\mathbb{C}$.

Definitions:

Let $\mathcal{X}$, $\mathcal{Y}$ be subspaces of $\mathbb{C}^r$ of dimension m and n. The **principal angles** $0 \le \theta_1 \le \cdots \le \theta_q \le \pi/2$ between $\mathcal{X}$ and $\mathcal{Y}$ and **principal vectors** $\mathbf{u}_1, \ldots, \mathbf{u}_q$ and $\mathbf{v}_1, \ldots, \mathbf{v}_q$ are defined inductively:
$$\cos(\theta_1) = \max\{|\mathbf{x}^*\mathbf{y}| : \mathbf{x} \in \mathcal{X}, \max_{\mathbf{y}\in\mathcal{Y}}, \|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1\}.$$

Let $\mathbf{u}_1$ and $\mathbf{v}_1$ be a pair of maximizing vectors. For $k = 2, \ldots, q$,

$$\cos(\theta_k) = \max\{|\mathbf{x}^*\mathbf{y}| : \mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}, \|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1, \quad \mathbf{x}^*\mathbf{u}_i = \mathbf{y}^*\mathbf{v}_i = 0, \quad i = 1, \ldots, k-1\}.$$

Let $\mathbf{u}_k$ and $\mathbf{v}_k$ be a pair of maximizing vectors. (Principal angles are also called **canonical angles**, and the cosines of the principal angles are called **canonical correlations**.)

Facts:

1. (*Principal Angles*) Let $\mathcal{X}, \mathcal{Y}$ be subspaces of $\mathbb{C}^r$ of dimension m and n.
   (a) [BG73] The principal vectors obtained by the process above are not necessarily unique, but the principal angles are unique (and, hence, independent of the chosen principal vectors).
   (b) Let $m = n \leq r/2$ and X, Y be matrices whose columns form orthonormal bases for the subspaces $\mathcal{X}$ and $\mathcal{Y}$, respectively.
      i. The singular values of X^*Y are the cosines of the principal angles between the subspaces $\mathcal{X}$ and $\mathcal{Y}$.
      ii. There are unitary matrices $U \in \mathbb{C}^{r\times r}$ and V_X and $V_Y \in \mathbb{C}^{n\times n}$ such that

$$UXV_X = \begin{bmatrix} I_n \\ 0_n \\ 0_{r-n,n} \end{bmatrix}, \quad UYV_Y = \begin{bmatrix} \Gamma \\ \Sigma \\ 0_{r-n,n} \end{bmatrix},$$

      where Γ and Σ are nonnegative diagonal matrices. Their diagonal entries are the cosines and sines respectively of the principal angles between $\mathcal{X}$ and $\mathcal{Y}$.
   (c) [QZL05] Take $m = n$. For any permutation invariant absolute norm g on $\mathbb{R}^m$,

$$g(\sin(\theta_1), \ldots, \sin(\theta_m)), \quad g(2\sin(\theta_1/2), \ldots, 2\sin(\theta_m/2)), \quad \text{and} \quad g(\theta_1, \ldots, \theta_m)$$

   are metrics on the set of subspaces of dimension n of $\mathbb{C}^{r\times r}$.
2. [GV96, Theorem 2.6.2] (*CS decomposition*) Let $W \in F^{n\times n}$ be unitary. Take a positive integer l such that $2l \leq n$. Then there are unitary matrices $U_{11}, V_{11} \in F^{l\times l}$ and $U_{22}, V_{22} \in F^{(n-l)\times(n-l)}$ such that

$$\begin{bmatrix} U_{11} & 0 \\ 0 & U_{22} \end{bmatrix} W \begin{bmatrix} V_{11} & 0 \\ 0 & V_{22} \end{bmatrix} = \begin{bmatrix} \Gamma & -\Sigma & 0 \\ \Sigma & \Gamma & 0 \\ 0 & 0 & I_{n-2l} \end{bmatrix},$$

   where $\Gamma = \text{diag}(\gamma_1, \ldots, \gamma_l)$ and $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_l)$ are nonnegative and $\Gamma^2 + \Sigma^2 = I$.
3. [GV96, Theorem 8.7.4] (*Generalized singular value decomposition*) Take $A \in F^{p\times n}$ and $B \in F^{m\times n}$ with $p \geq n$. Then there is an invertible $X \in F^{n\times n}$, unitary $U \in F^{p\times p}$ and $V \in F^{m\times m}$, and nonnegative diagonal matrices $\Sigma_A \in \mathbb{R}^{n\times n}$ and $\Sigma_B \in \mathbb{R}^{q\times q}$ ($q = \min\{m, n\}$) such that $A = U\Sigma_A X$ and $B = V\Sigma_B X$.

References

[And94] T. Ando. Majorization and inequalitites in matrix theory. *Lin. Alg. Appl.*, 199:17–67, 1994.
[Bha97] R. Bhatia. *Matrix Analysis*. Springer-Verlag, New York, 1997.
[Bha01] R. Bhatia. Linear algebra to quantum cohomology: the story of Alfred Horn's inequalities. *Amer. Math. Monthly*, 108(4):289–318, 2001.
[BG73] A. Björk and G. Golub. Numerical methods for computing angles between linear subspaces. *Math. Comp.*, 27:579–594, 1973.

[Ful00] W. Fulton. Eigenvalues, invariant factors, highest weights, and Schurbert calculus. *Bull. Am. Math. Soc.*, 37:255–269, 2000.

[GV96] G.H. Golub and C.F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 3rd ed., 1996.

[GE95] Ming Gu and Stanley Eisenstat. A divide-and-conquer algorithm for the bidiagonal SVD. *SIAM J. Matrix Anal. Appl.*, 16:72–92, 1995.

[Hig96] N.J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, Philadelphia, 1996.

[Hig89] N.J. Higham. Matrix nearness problems and applications. In M.J.C. Gover and S. Barnett, Eds., *Applications of Matrix Theory*, pp. 1–27. Oxford University Press, U.K. 1989.

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

[HJ91] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.

[Kit95] F. Kittaneh. Singular values of companion matrices and bounds on zeros of polynomials. *SIAM. J. Matrix Anal. Appl.*, 16(1):330–340, 1995.

[Kly98] A. A. Klyachko. Stable bundles, representation theory and Hermitian operators. *Selecta Math.*, 4(3):419–445, 1998.

[KT99] A. Knutson and T. Tao. The honeycomb model of $GL_n(C)$ tensor products i: proof of the saturation conjecture. *J. Am. Math. Soc.*, 12(4):1055–1090, 1999.

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

[LM01] C.-K. Li and R. Mathias. Construction of matrices with prescribed singular values and eigenvalues. *BIT*, 41(1):115–126, 2001.

[LM02] C.-K. Li and R. Mathias. Inequalities on singular values of block triangular matrices. *SIAM J. Matrix Anal. Appl.*, 24:126–131, 2002.

[MO79] A.W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, London, 1979.

[Mat93] R. Mathias. Perturbation bounds for the polar decomposition. *SIAM J. Matrix Anal. Appl.*, 14(2):588–597, 1993.

[QZL05] Li Qiu, Yanxia Zhang, and Chi-Kwong Li. Unitarily invariant metrics on the Grassmann space. *SIAM J. Matrix Anal. Appl.*, 27(2):507–531, 2006.

[Tho75] R.C. Thompson. Singular value inequalities for matrix sums and minors. *Lin. Alg. Appl.*, 11(3):251–269, 1975.

[Tho77] R.C. Thompson. Singular values, diagonal elements, and convexity. *SIAM J. Appl. Math.*, 32(1):39–63, 1977.

[TT73] R.C. Thompson and S. Therianos. On the singular values of a matrix product-I, II, III. *Scripta Math.*, 29:99–123, 1973.

[Zha99] X. Zhan. Norm inequalities for Cartesian decompositions. *Lin. Alg. Appl.*, 286(1–3):297–301, 1999.

[Zha02] X. Zhan. *Matrix Inequalities*. Springer-Verlag, Berlin, Heidelberg, 2002. (Lecture Notes in Mathematics 1790.)