Part 3 Norms and norm inequalities

The study of norms has connections to many pure and applied areas. We will focus on approximation problems and norm inequalities in matrix spaces.

1 $\mathcal{S}$-invariant norms

Definition 1.1 *A* **norm** $\|\cdot\|$ *on a vector space* $\mathbf{V}$ *is a function from* V *to* $\mathbf{R}$ *such that*

(a) $\|v\| \geq 0$ *for all* $v \in V$*, where* $\|v\| = 0$ *if and only if* $v = 0$*.*

(b) $|\gamma v\| = |\gamma|\|v\|$ *for all* $\gamma \in \mathbf{F}$ *and* $v \in V$*.*

(c) $\|u + v\| \leq \|u\| + \|v\|$ *for all* $u, v \in V$*.*

If the condition " $\|v\| = 0$ *if and only if* $v = 0$ *" is relaxed, we have a* **semi-norm***.*

Definition 1.2 *Let* $\mathcal{S}$ *be a set of operators acting on the vector space* $\mathbf{V}$*. A norm* $\|\cdot\|$ *on* $\mathbf{V}$ *is* $\mathcal{S}$**-invariant** *if* $\|Av\| = \|v\|$ *for all* $A \in \mathcal{S}$ *and* $v \in \mathbf{V}$*.*

Example 1.3 *Here are several commonly used* $\mathcal{S}$*-invariant norms on vectors or matrices.*

1. **Absolute norms** *on* $\mathbf{F}^n$*, i.e.,* $\|Dv\| = \|v\|$ *for all diagonal unitary / diagonal orthogonal matrices* D*.*

2. **Permutationally invariant norms** *on* $\mathbf{F}^n$*, i.e.,* $\|Pv\| = \|v\|$ *for all permutation matrices* P*.*

3. **Symmetric norms** *(also known as* **symmetric gauge functions***) on* $\mathbf{R}^n$*, i.e, absolute permutationally invariant norms.*

4. **Unitarily invariant norms** *on* $m \times n$ *matrices, i.e.,* $\|UAV\| = \|A\|$ *for all unitary matrices* U *and* V*.*

5. **Unitary similarity invariant norms** *on* M_n *or* H_n*, i.e.,* $\|U^*AU\| = \|A\|$ *for all unitary matrices* U*.*

6. **Unitary congruence invariant norms** *on* M_n*, symmetric matrices, or skew-symmetric matrices, i.e.,* $\|U^tAU\| = \|A\|$ *for all unitary matrices* U*.*

Theorem 1.4 *Suppose* $\mathbf{V}$ *is an inner product space and* $\mathcal{S}$ *is a group of unitary operators.*

(a) *For* $\gamma \in \mathbf{V}$*, the* $\mathcal{S}(\gamma)$*-radius defined by*

$$r_\gamma(v) = \max\{|\langle v, g(\gamma)\rangle| : g \in \mathcal{S}\}$$

is an $\mathcal{S}$*-invariant semi-norm. It is a norm if and only if* $\mathcal{S}(\gamma) = \{g(\gamma) : g \in \mathcal{S}\}$ *spans* $\mathbf{V}$*.*

(b) *For every $\mathcal{S}$-invariant norm $\|\cdot\|$ on $\mathbf{V}$, there is a compact subset $\mathcal{E}$ of $\mathbf{V}$ such that*

$$\|v\| = \max\{r_\gamma(v) : \gamma \in \mathcal{E}\}.$$

(c) *For any $u, v \in \mathbf{V}$, $\|u\| \le \|v\|$ for all $\mathcal{S}$-invariant norms if and only if $r_\gamma(u) \le r_\gamma(v)$ for all $\gamma \in \mathbf{V}$ (or a suitable collection of γ).*

Remark 1.5 *The theorem shows that $\mathcal{S}(\gamma)$-radii are the building blocks of $\mathcal{S}$-norms if $\mathcal{S}$ is a group of unitary operators.*

2 UI norms, USI norms, and UCI norms

Definition 2.1 *Let $c = (c_1, \dots, c_n)$ be a vector with nonnegative entries arranged in descending order. The c-**norm** (or c-**spectral norm**) on M_n is defined by*

$$\|A\|_c = \sum_{j=1}^{n} c_j s_j(A).$$

*When $c_1 = \cdots = c_k = 1$ and $c_j = 0$ for $j > k$, have the **Ky Fan k-norm** $F_k(A)$.*

Theorem 2.2 *There is a one-one correspondence between symmetric gauge functions on $\mathbf{R}^n$ and UI norms on M_n, namely, each UI norm $\|\cdot\|$ corresponds to a unique symmetric gauge unction Φ such that $\|A\| = \Phi(s(A))$.*

(a) *Every UI norm $\|\cdot\|$ there is a compact set $\mathcal{E}$ of nonnegative vectors with entries arranged in descending order such that*

$$\|A\| = \max\{\|A\|_c : c \in \mathcal{E}\}.$$

(b) *Let $A, B \in M_n$. Then $\|A\| \le \|B\|$ for all UI norms $\|\cdot\|$ if and only if $F_k(A) \le F_k(B)$ for all $k = 1, \dots, n$.*

Theorem 2.3 *There is a one-one correspondence between permutationally invariant norms on $\mathbf{R}^n$ and USI norms on H_n, namely, each USI norm $\|\cdot\|$ corresponds to a permutationally invariant norm Φ such that $\|A\| = \Phi(\lambda(A))$.*

(a) *For every USI norm on H_n, there is a compact set $\mathcal{E}$ of real vectors $c = (c_1, \dots, c_n)$ with entries arranged in descending order such that*

$$\|A\| = \max\{r_c(A) : c \in \mathcal{E}\},$$

where

$$r_c(A) = \max\left\{\sum_{j=1}^{n} c_j \lambda_j(A), -\sum_{j=1}^{n} c_{n-j+1} \lambda_j(A)\right\}$$

*is the c-**numerical radius** of A.*

(b) *Let* $A, B \in H_n$*. Then* $\|A\| \le \|B\|$ *for all USI norms* $\|\cdot\|$ *if and only if* $r_c(A) \le r_c(B)$ *for a dense subset of*

$$\left\{(c_1, \ldots, c_n) : c_1 \ge \cdots \ge c_n, \sum_{j=1}^{n} c_j^2 = 1\right\}.$$

Definition 2.4 *Let* $C \in M_n$*. Define the* C**-numerical radius** *of* A *by*

$$r_C(A) = \max\{|\mathrm{tr}\,(CU^*AU)| : U \text{ is unitary}\},$$

and define the C**-congruence numerical radius** *of* A *by*

$$\tilde{r}_C(A) = \max\{|\mathrm{tr}\,(CU^tAU)| : U \text{ is unitary}\}.$$

Theorem 2.5 (a) *For every USI norm on* M_n*, there is a compact set* $\mathcal{E}$ *of matrices*

$$\|A\| = \max\{r_C(A) : C \in \mathcal{E}\}.$$

(b) *For every UCI norm on* M_n*, there is a compact set* $\mathcal{E}$ *of matrices*

$$\|A\| = \max\{\tilde{r}_C(A) : C \in \mathcal{E}\}.$$

Theorem 2.6 *A norm* $\|\cdot\|$ *on* M_n *is UI if and only if it is USI and UCI.*

3 Best approximations

Theorem 3.1 *Let* $\|\cdot\|$ *be a UI norm on* M_n*.*

(a) $\|A - (A + A^*)/2\| \le \|A - H\|$ *for any* $H \in H_n$*.*

(b) $\|A - (A - A^*)/2\| \le \|A - G\|$ *for any* $G \in iH_n$*.*

(c) *If* $A = UP$ *for a unitary* U *and a positive semi-definite* P*, then* $\|A - U\| \le \|A - V\|$ *for any unitary* $V \in M_n$*.*

Theorem 3.2 *Let* $\|\cdot\|$ *be a UI norm on* $m \times n$ *matrices.*

(a) *If* $A = UDV$*, where* $U, V \in M_n$ *are unitary and* $D = \mathrm{diag}\,(s_1, \ldots, s_n)$ *such that* $s_1 \ge \cdots \ge s_n \ge 0$*, then* $A_k = U\mathrm{diag}\,(s_1, \ldots, s_k, 0, \ldots, 0)V$ *satisfies* $\|A - A_k\| \le \|A - X\|$ *for all* X *with rank at most* k*.*

(b) *If* $A = XDY^*$ *so that* X *is* $m \times k$ *with* $X^*X = I_k$*,* Y *is* $n \times k$ *with* $Y^*Y = I_k$*, and* $D = \mathrm{diag}\,(s_1, \ldots, s_k)$ *with* $s_1 \ge \cdots \ge s_k > 0$*, then for any given* $m \times n$ *matrix* B $\|A - UD^{-1}VB\| \le \|A - X\|$ *for all* $m \times n$ *matrix* X*.*

4 Norm bounds for sum and difference of matrices

Theorem 4.1 *Let $\|\cdot\|$ be a UI norm on M_n. Suppose $A, B \in H_n$ have singular values $a_1 \geq \cdots \geq a_n$ and $b_1 \geq \cdots \geq b_n$. Then*

$$\frac{1}{\sqrt{2}}\|A + iB\| \leq \|\mathrm{diag}\,(a_1 + ib_1, \ldots, a_n + ib_n)\| \leq \sqrt{2}\|A + iB\|.$$

Theorem 4.2 *Suppose $\|\cdot\|$ is a USI norm on H_n. If $A, B \in H_n$ have eigenvalues $a_1 \geq \cdots \geq a_n$ and $b_1 \geq \cdots \geq b_n$, then*

$$\|\mathrm{diag}\,(a_1 - b_1, \ldots, a_n - b_n)\| \leq \|A - B\| \leq \|\mathrm{diag}\,(a_1 - b_n, \ldots, a_n - b_1)\|.$$

Theorem 4.3 *Suppose $\|\cdot\|$ is a UI norm on M_n. If $A, B \in H_n$ have singular values $a_1 \geq \cdots \geq a_n$ and $b_1 \geq \cdots \geq b_n$, then*

$$\|\mathrm{diag}\,(a_1 - b_1, \ldots, a_n - b_n)\| \leq \|A - B\|.$$

Theorem 4.4 *Let $A, B \in M_n$, and let $\|\cdot\|$ be the operator norm. Then*

$$\max\{\|U^*AU - V^*BV\| : U, V \, unitary\} = \min\{\|A + \mu I\| + \|B + \mu I\| : \mu \in \mathbf{C}\}.$$

The minimum on the right hand side will attain at a certain μ_0 with $|\mu_0| \leq \max\{\|A\|, \|B\|\}$.

5 Norm bounds involving product of matrices

Proposition 5.1 *Suppose $A, B \in M_n$ are such that $AB \in H_n$.*

(a) *For any USI norm on H_n, $\|AB\| \leq \|\mathrm{Re}(BA)\|$.*

(b) *For any UI norm on M_n, $\|AB\| \leq \|\mathrm{Re}(BA)\|$.*

Theorem 5.2 *Let $\|\cdot\|$ be a UI norm on M_n. For any matrices A, B, X, we have*

$$\|AXB^*\| \leq \frac{1}{2}\|A^*AX + XB^*B\|.$$

Theorem 5.3 *Let $\|\cdot\|$ be a UI norm. Suppose $A, B \in H_n$ are positive semi-definite. Then*

$$4\|AB\| \leq \|(A + B)^2\|.$$

Actually, we can prove

$$2s_j(A^{3/2}B^{1/2} + A^{1/2}B^{3/2}) \leq s_j(A + B)^2, \quad j = 1, \ldots, n.$$

Consequently,

$$4\|AB\| = 4\|A^{1/2}(A^{1/2}B^{1/2})B^{1/2}\| \leq 2\|A^{3/2}B^{1/2} + A^{1/2}B^{3/2}\| \leq \|(A + B)\|^2.$$

6 Additional results and problems

Definition 6.1 *For $p \geq 1$, define the Schatten p-norm on M_n by*

$$\|A\|_p = \left\{ \sum_{j=1}^{n} s_j(A)^p \right\}^{1/p}.$$

Theorem 6.2 *Let $A, B \in H_n$.*

(a) *If $1 \leq p \leq 2$, then*

$$2^{2/p-1}\|A+iB\|_p^2 \geq \|A\|_p^2 + \|B\|_p^2 \geq 2^{1-2/p}\|A+iB\|_p^2.$$

(b) *If $2 \leq p \leq \infty$, then*

$$2^{2/p-1}\|A+iB\|_p^2 \leq \|A\|_p^2 + \|B\|_p^2 \leq 2^{1-2/p}\|A+iB\|_p^2.$$

Theorem 6.3 *Let $A, B \in H_n$ be such that A is positive semidefinite.*

(a) *If $1 \leq p \leq 2$, then*

$$\|A+iB\|_p^2 \geq \|A\|_p^2 + 2^{1-2/p}\|B\|_p^2.$$

(b) *If $2 \leq p \leq \infty$, then*

$$\|A+iB\|_p^2 \leq \|A\|_p^2 + 2^{1-2/p}\|B\|_p^2.$$

(c) $\|A\|_1^2 + \|B\|_1^2 \leq \|A+iB\|_1^2$.

In addition, suppose B is also positive semidefinite.

(d) *If $1 \leq p \leq 2$, then*

$$\|A+iB\|_p^2 \geq \|A\|_p^2 + \|B\|_p^2.$$

(e) *If $2 \leq p \leq \infty$, then*

$$\|A+iB\|_p^2 \leq \|A\|_p^2 + \|B\|_p^2.$$

Problem 6.4 *Can one prove (d) and (e) without the assumption that B is also positive semidefinite?*

Theorem 6.5 (Böttcher and Wenzel) *Let $A, B \in M_n$. Then*

$$\|AB - BA\|_2 \leq \sqrt{2}\|A\|_2\|B\|_2.$$

Problem 6.6 *Characterize A and B so that the equality holds.*

It is known that if the equality holds, then A and B have rank at most 2, $A \in \{X \in M_n : BX = XB\}^{\perp}$ and $B \in \{X \in M_n : AX = XA\}^{\perp}$.

Problem 6.7 *Prove or disprove that*

$$\|AB - BA\|_p \leq 2^{1/\min(p,q)}\|A\|_p\|B\|_p.$$

7 Exercises

1. Show that every symmetric gauge function is Schur convex monotone.

2. Show that there exist a UI norm and $A, B \in M_n$ with singular values $a_1 \geq \cdots \geq a_n$ and $b_1 \geq \cdots \geq b_n$ such that

$$\|A - B\| > \|\mathrm{diag}\,(a_1 - b_n, \ldots, a_n - b_1)\|.$$

3. Suppose $\|\cdot\|$ is a UI norm on M_n. Let $A, B \in M_n$ have singular values $a_1 \geq \cdots \geq a_n$ and $b_1 \geq \cdots \geq b_n$, and $C = AB$ have singular values $c_1 \geq \cdots \geq c_n$. Prove the following.
(a) For any $r > 0$,

$$(c_1^r, \ldots, c_n^r) \prec_{log} (a_1^r b_1^r, \ldots, a_n^r b_n^r).$$

(b) For any $p, q \geq 1$ such that $1/p + 1/q = 1$,

$$\|AB\| \leq \|\,|A|^p\,\|^{1/p} \|\,|B|^q\,\|^{1/q}.$$

(c) For any $p, q > 0$ and $1/p + 1/q = 1/r$,

$$\|\,|AB|^r\,\|^{1/r} \leq \|\,|A|^p\,\|^{1/p} \|\,|B|^q\,\|^{1/q}.$$

For $p = q = 2$, we have the Cauchy-Schwarz inequality for UI norms.

4. For any USI norm $\|\cdot\|$ on M_n, and $A = (A_{ij})_{1 \leq i,j \leq 2} \in M_n$ with $A_{11} \in M_m$ and $A_{22} \in M_{n-m}$, show that

$$\|A_{11} \oplus A_{22}\| \leq \|A\|.$$

If $\|\cdot\|$ is UI, then $\|A_{11} \oplus 0_{n-m}\| \leq \|A\|$. What about USI norms?