From: "Feedback Systems: Input-Output Properties" by Desoer and Vidyasagar. Academic Press, 1975.

II NORMS

The purpose of this chapter is to supply us with a number of tools that we shall use repeatedly in the sequel. It is important for us to think of norms as a yardstick with which we measure the size of vectors in $\mathbb{R}^n$, or of real-valued function, or of vector-valued functions. Also we use norms to measure the "gain" of linear operators. In Section 1, we introduce the general definition of norm together with many examples that we shall repeatedly use in later chapters. In Section 2, we define equivalent norms and prove that all norms in finite-dimensional spaces are equivalent.

Section 3 establishes some inclusion relations between some normed spaces that will be frequently encountered later. A geometric interpretation of norms is given in Section 4; it gives an intuitive insight into the concept of norm and the way of visualizing the difference between norms. Section 5 defines the concept of induced norms of linear maps, and two examples of convolution maps are worked out in Section 6. In Section 7, the relation between norms and spectral radius is developed in detail. Finally, in Section 8, the concept of measure of a matrix, which is closely related to the concept of norm, is developed. It is used to obtain lower and upper bounds on solution of differential equations and to obtain an existence and uniqueness result for equilibrium points.

1 Norms: definitions and examples

1 Let E be a *linear space* over the *field* $\mathbb{K}$. (Typically, $\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$). The zero vector in E is denoted by 0. We say that the function $\rho: E \to \mathbb{R}_+$ is a norm on E iff

(i) $x \in E$ and $x \neq 0 \Rightarrow \rho(x) > 0$
(ii) $\rho(\alpha x) = |\alpha| \rho(x)$, $\forall \alpha \in \mathbb{K}$, $\forall x \in E$
(iii) $\rho(x + y) \leq \rho(x) + \rho(y)$, $\forall x, y \in E$ (triangle inequality)

Remark Given a linear space E, there may be many possible norms on E. However, given the linear space E and a norm ρ on E, the pair (E, ρ) is called a normed space.

Example 1 Let the linear space E be $\mathbb{C}^n$. More precisely, $x \in \mathbb{C}^n$ means that $x = (x_1, x_2, \ldots, x_n)$ with $x_i = \mathbb{C}$, $\forall i$. We shall repeatedly use the following norms on E:

2 $$\|x\|_1 \triangleq \sum_{i=1}^{n} |x_i|$$

3 $$\|x\|_p \triangleq \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}, \qquad \text{where } 1 \leq p < \infty$$

4 $$\|x\|_\infty \triangleq \max_i |x_i|$$

$\|x\|_2$ is called the Euclidean norm of x.

Exercise 1 Consider $\mathbb{R}^n$ and define norms as above. For $\mathbb{R}^2$ draw the sets $\{x \mid \|x\|_p = 1\}$ for $p = 1, 2, 4, \infty$.

Example 2 Let E be the space of infinite sequences of complex numbers: $x = (\xi_1, \xi_2, \ldots)$ with $\xi_i \in \mathbb{C}$ for $i = 1, 2, \ldots$. Frequently used norms on appropriate proper subsets of E are given as follows:

5 $$\|x\|_1 \triangleq \sum_{i=1}^{\infty} |\xi_i|$$

6 $$\|x\|_p \triangleq \left(\sum_{i=1}^{\infty} |\xi_i|^p \right)^{1/p}, \qquad 1 \leq p < \infty$$

7 $$\|x\|_\infty \triangleq \sup_{i \geq 1} |\xi_i|$$

The corresponding normed spaces are called, respectively, l^1, l^p, l^∞.

We used above the expression "on appropriate subsets of E" because the norms defined by (5), (6), and (7) have the property that (calling $\rho(x)$ any such norm)

$$M \triangleq \{x \in E \mid \rho(x) < \infty\} \text{ is a linear subspace of } E$$

Indeed, it follows directly from Minkowski's inequality that if $x_1, x_2 \in M$, then $x_1 + x_2 \in M$ because $\rho(x_1) + \rho(x_2) \geq \rho(x_1 + x_2)$.

Example 3 Let $E = \{f: \mathbb{R} \to \mathbb{R} \mid f \text{ locally (Lebesgue) integrable}\}$. Frequently used norms on appropriate proper subsets of E are as follows:

8 $$\|x\|_1 \triangleq \int |x(t)|\, dt$$

9 $$\|x\|_p \triangleq \left(\int |x(t)|^p\, dt\right)^{1/p}, \qquad 1 \leq p < \infty$$

10 $$\|x\|_\infty \triangleq \operatorname*{ess\,sup}_{t \in \mathbb{R}} |x(t)| = \inf\{a \in \mathbb{R} \mid \mu[\{t \mid |x(t)| > a\}] = 0\}$$

where $\mu[A]$ denotes the Lebesgue measure of the set A.† The corresponding normed spaces are called, respectively, L^1, L^p, L^∞.

More generally, for a given $w: \mathbb{R} \to \mathbb{R}$ positive, continuous, and bounded on $\mathbb{R}$, we can define a norm as follows:

11 $$\|x\|_p' \triangleq \left(\int w(t)|x(t)|^p\, dt\right)^{1/p}, \qquad 1 \leq p < \infty$$

and the corresponding expression for $\|x\|_\infty$.

Example 4 Let $E = \{f: \mathbb{R} \to \mathbb{R}^n, f = (f_1, f_2, \ldots, f_n) \mid \text{all } f_i\text{'s locally (Lebesgue) integrable}\}$. Let $|\cdot|$ denote *any norm on* $\mathbb{R}^n$, then we define

12 $$\|f\|_p \triangleq \left(\int |f(t)|^p\, dt\right)^{1/p}, \qquad 1 \leq p < \infty$$

13 $$\|f\|_\infty \triangleq \operatorname*{ess\,sup}_{t \in \mathbb{R}} |f(t)|$$

Exercise 2 For the linear space E defined in Example 2, give an example of some $x \in E$ for which $\|x\|_1 = \infty$ but $\|x\|_\infty = 1$. Give some $x \in E$ for which $\|x\|_\infty = \infty$. (These examples justify the locution "*on appropriate proper subsets of E*".)

Note: It is a standard, though messy, exercise in functional analysis to verify that each of the functions (2) to (13) satisfy the axioms of the norm.

† In our later work we shall abuse notation and write $\sup|x(t)|$ instead of $\operatorname{ess\,sup}|x(t)|$.

Example 5 Let $E = \mathbb{C}^{n \times n}$, the set of all $n \times n$ matrices with elements in $\mathbb{C}$. E is a linear space. The following are norms on $\mathbb{C}^{n \times n}$.

14 $$\|A\|_a \triangleq \max_{i,j} |a_{ij}|$$

15 $$\|A\|_b \triangleq \sum_{i,j=1}^{n} |a_{ij}|$$

16 $$\|A\|_s \triangleq \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right)^{1/2}$$

17 $$\|A\|_1 \triangleq \max_j \sum_{i=1}^{n} |a_{ij}| \qquad \text{(column sums)}$$

18 $$\|A\|_2 \triangleq \max_i [\lambda_i(A^*A)]^{1/2} \qquad \text{(where } \lambda_i(M) \text{ denotes the } i\text{th eigenvalue of } M)$$

19 $$\|A\|_\infty \triangleq \max_i \sum_{j=1}^{n} |a_{ij}| \qquad \text{(row sums)}$$

In this example we consider matrices as elements of a linear space. In a later section, we shall consider matrices as representations of linear maps and shall relate the matrix norms to the vector norms of the domain and range spaces.

Exercise 3 Let $\|\cdot\|$ denote a norm on R^n. Let $\phi(\cdot)$ be continuous and map $[0, T]$ into $\mathbb{R}^n$. By examining the Riemann sums, show that

20 $$\left\| \int_0^T \phi(t)\, dt \right\|_\infty \leq \int_0^T \|\phi(t)\|_\infty\, dt$$

In fact, as we shall see later, inequality (20) holds for *any* norm in $\mathbb{R}^n$. Furthermore, it also holds for any $\phi: [0, T]$ into $\mathbb{R}^n$, which is Lebesgue integrable.

2 Equivalent norms

1 Let E be a linear space. Let $\|\cdot\|_a$ and $\|\cdot\|_b$ be two norms on E. The norms $\|\cdot\|_a$ and $\|\cdot\|_b$ are said to be *equivalent* iff, there exist two positive numbers m_l and m_u such that

2 $$m_l\|x\|_a \leq \|x\|_b \leq m_u\|x\|_a, \qquad \forall x \in E$$

It is crucial to note that the *same* m_l and m_u must work for *all* $x \in E$ in (2). The relation expressed by (2) between the two norms is an *equivalence* relation (check reflexivity, symmetry, transitivity).

If two norms are equivalent, then sequences that converge in terms of one norm converge in terms of the other. The same holds for continuity and boundedness. In short, equivalent norms define identical topologies. In applications, some norms are preferable because they give sharper results as we shall see below.

Exercise 1 Let

$$A = \begin{bmatrix} 0.9 & 10^4 \\ 0 & 0.9 \end{bmatrix}.$$

Calculate the norm of A using the six norms of Example 5 of Section 1. Prove that $A^k \to 0$ as $k \to \infty$. (Any comment on your results?)

3 **Theorem** All norms on $\mathbb{C}^n$ are equivalent.

Proof (a) First we show that *any* norm $\|\cdot\|$ in $\mathbb{C}^n$ is a continuous function in the sense that whenever $x \in \mathbb{C}^n$ tends to $\bar{x} \in \mathbb{C}^n$, then the real number $\|x\|$ tends to the real number $\|\bar{x}\|$. Call x_i, $(\bar{x}_i)$, the ith component of x, $(\bar{x})$, with respect to some basis $\{e_1, e_2, \ldots, e_n\}$. From advanced calculus, if $x \to \bar{x}$, then $x_i \to \bar{x}_i$ for $i = 1, 2, \ldots, n$. Now for any x, $\bar{x} \in \mathbb{C}^n$, using the axioms of the norm

$$0 \le \big|\|x\| - \|\bar{x}\|\big| \le \|x - \bar{x}\| = \left\|\sum_{i=1}^{n} (x_i - \bar{x}_i)e_i\right\| \le \sum_{i=1}^{n} |x_i - \bar{x}_i|\, \|e_i\|$$

Hence $x \to \bar{x}$ implies that $\|x\| \to \|\bar{x}\|$.

(b) Let

$$S_\infty = \{x \in \mathbb{C}^n \,|\, \|x\|_\infty = 1\}$$

Clearly S_∞ is a bounded set. Furthermore, it is closed, since it is the inverse image of the closed set $\{1\}$ under the continuous map $\|\cdot\|_\infty$. Let $\|\cdot\|$ be an arbitrary norm on $\mathbb{C}^n$. Now the continuous function $z \mapsto \|z\|$, restricted to the closed and bounded set S_∞, reaches its minimum and maximum at, say, z_m and z_M, respectively. Thus,

$$0 < \|z_m\| \le \|z\| \le \|z_M\|, \qquad \forall z \in S_\infty$$

Let x be any point in $\mathbb{C}^n$ with $x \ne 0$, then $x/\|x\|_\infty \in S_\infty$. Therefore,

$$\|z_m\| \le \left\|\frac{x}{\|x\|_\infty}\right\| \le \|z_M\|, \qquad \forall x \in \mathbb{C}^n, \quad x \ne 0$$

Consequently,

$$\|z_m\|\, \|x\|_\infty \le \|x\| \le \|z_M\|\, \|x\|_\infty, \qquad \forall x \in \mathbb{C}^n$$

Hence the norms $\|\cdot\|$ and $\|\cdot\|_\infty$ are equivalent. The same holds for any other pair of norms, say, $\|\cdot\|$ and $\|\cdot\|'$, by transitivity. ∎

4 **Notation** From now on we shall encounter norms on $\mathbb{R}^n$ or $\mathbb{C}^n$, norms on function spaces, and induced norms of linear operators. Some authors use $\|\cdot\|$ to denote the first and $|\|\cdot\||$ to denote the last two. From now on we shall use $|\cdot|$ for norms on $\mathbb{R}^n$ or $\mathbb{C}^n$ and $\|\cdot\|$ for norms on function spaces or for induced norms of linear operators. We shall also use $|\cdot|$ to denote the absolute value of a number in $\mathbb{R}$ or $\mathbb{C}$.

5 **Remark** In infinite-dimensional spaces, norms are not necessarily equivalent. For example, in the space of infinite sequences (Section 1, Example 2) if

$$x_1 = (1, 0, 0, \ldots),\ x_2 = (1, 2^{-1}, 0, \ldots),\ x_3 = (1, 2^{-1}, 3^{-1}, \ldots), \ldots$$

then

$$\|x_k\|_\infty = 1, \qquad \forall k \in \mathbb{Z}_+$$

and the sequence $(x_k)_1^\infty$ converges. However,

$$\|x_k\|_1 \to \infty, \qquad \text{as} \quad k \to \infty$$

6 **Example** Let E be the space of sequences in $\mathbb{C}^n$; i.e., $x \in E$ iff $x = (\xi_1, \xi_2, \ldots)$ with $\xi_i \in \mathbb{C}^n$ for $i = 1, 2, \ldots$. Let $|\cdot|$ denote *any* norm on $\mathbb{C}^n$. Thus, $|\xi_i|$ denotes the nonnegative number equal to the norm of the $\mathbb{C}^n$ vector ξ_i. Then on some proper subsets of E, we can define norms

7 $$\|x\|_1 \triangleq \sum_{i=1}^{\infty} |\xi_i|$$

8 $$\|x\|_p \triangleq \left(\sum_{i=1}^{\infty} |\xi_i|^p\right)^{1/p}, \qquad 1 \le p < \infty$$

9 $$\|x\|_\infty \triangleq \sup_{i \ge 1} |\xi_i|$$

The corresponding normed spaces are denoted, respectively, by l_n^1, l_n^p, l_n^∞. Verify that if we set up the definition of l_n^p using two different norms on $\mathbb{C}^n$, we end up with the same proper subsets of E.

3 Relations between normed spaces

In discrete-time systems, the input is completely specified by the corresponding sequence of values. Thus, we consider the space of sequences with values in $\mathbb{R}$ (or $\mathbb{C}$). Associated with the norms defined in (1.5) to (1.7), we have the normed spaces l^1, l^p, and l^∞.

1 **Theorem** We have the strict inclusions $l^1 \subset l^p \subset l^\infty$, for any integer $p \in (1, \infty)$.

Proof Let $x = (\xi_1, \xi_2, \ldots)$, where $\xi_i \in \mathbb{C}$, $\forall i \in \mathbb{Z}_+$. Consider any integer $p \in [1, \infty)$. If $x \in l^p$, then $\sum_{i=1}^{\infty} |\xi_i|^p < \infty$. Therefore, for any integer k, $|\xi_k|^p \le \sum_{i=1}^{\infty} |\xi_i|^p < \infty$, hence $x \in l^\infty$. Thus, $l^p \subset l^\infty$ for $p \in [1, \infty)$.

Now for any integer $N \ge 0$ and any integer $p \ge 1$,

$$\sum_{i=1}^{N} |\xi_i|^p \le \left(\sum_{i=1}^{N} |\xi_i| \right)^p \le (\|x\|_1)^p$$

Hence, as $N \to \infty$, if $x \in l^1$, we have $\|x\|_p \le \|x\|_1 < \infty$. Thus, $l^1 \subset l^p$. ∎

2 **Exercise 1** Use Example (2.6) to verify that

3 $$l_n^1 \subset l_n^p \subset l_n^\infty, \qquad \text{for integers} \quad p \in (1, \infty)$$

and show that these inclusion relations are strict.

In continuous-time systems, the inputs and outputs are functions of time t, and t is usually restricted to $t \ge 0$. For simplicity, we shall use Lebesgue integration results. (Any reader uncomfortable with Lebesgue theory may take all the functions to be piecewise continuous; see Appendix A.) We say that $f: \mathbb{R}_+ \to \mathbb{R}$ is *locally integrable* iff f is integrable over any bounded interval (i.e., intervals such as $[a, b]$ with $0 \le a \le b < \infty$).

For any fixed $p \in [1, \infty)$, we say that $f: \mathbb{R}_+ \to \mathbb{R}$ belongs to L^p iff f is locally integrable and $\int_0^\infty |f(t)|^p \, dt < \infty$. We write

4 $$\|f\|_p = \left(\int_0^\infty |f(t)|^p \, dt \right)^{1/p}$$

We say that $f: \mathbb{R}_+ \to \mathbb{R}$ belongs to L^∞ iff $\text{ess sup}_{t \ge 0} |f(t)| < \infty$. We write

5 $$\|f\|_\infty = \operatorname*{ess\,sup}_{t \ge 0} |f(t)|$$

By essential supremum we mean

$$\operatorname*{ess\,sup}_{t \ge 0} |f(t)| = \inf\{a \mid |f(t)| \le a \text{ almost everywhere}\}$$

that is, $|f(t)| \le a$ except for a set of measure zero, and the ess sup is the smallest number which has that property. From now on we shall write sup for ess sup.

It is understood once and for all that elements of L^p are equivalence classes in the sense that if f and $g \in L^p$ but

$$\|f - g\|_p = 0$$

the functions f and g (which may be different as functions) are considered to be the same element of L^p. In the same spirit, we write $\sup|x(t)|$ instead of ess $\sup|x(t)|$.

6 **Exercise 2** Let $f(t) = \sin t$ for $t \ge 0$. Let $g(t) = 1$ for $t > 0$ and $g(0) = 2$. Show that

$$\sup_{t \ge 0} |f(t)| = 1, \qquad \sup_{t \ge 0} |g(t)| = 2, \qquad \text{and} \quad \|f\|_\infty = \|g\|_\infty$$

Let p be fixed but $p \in [1, \infty]$. It is well known that the linear space L^p together with the norm $\|\cdot\|_p$ is complete (i.e., every Cauchy sequence in L^p converges to an element in L^p; furthermore, this element is unique). Complete normed spaces are called Banach spaces.

7 **Fact** If $f: \mathbb{R}_+ \to \mathbb{R}$ and $f \in L^1 \cap L^\infty$, then $f \in L^p$ for $p \in [1, \infty]$.

Proof Since $f \in L^1$, the Lebesgue measure of the set I, where $I \triangleq \{t \mid |f(t)| \ge 1\}$, is finite. This fact together with $f \in L^\infty \Rightarrow \int_I |f|^p \, dt < \infty$. Now if I^c denotes the complement of I,

$$\infty > \int_{I^c} |f| \, dt \ge \int_{I^c} |f|^p \, dt, \qquad \text{for all} \quad p \in [1, \infty)$$

The conclusion follows from these two observations. ∎

The relation among L^1, L^2, and L^∞ is shown in the Venn diagram of Fig. II.1. The diagram illustrates that $f \in L^1 \cap L^\infty \Rightarrow f \in L^2$.

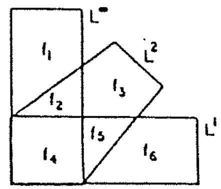

Figure II.1

8 **Exercise 3** Consider functions mapping $\mathbb{R}_+$ into R and defined by

$$f_1: t \mapsto 1; \qquad f_2: t \mapsto \frac{1}{1+t}; \qquad f_3: t \mapsto \frac{1}{1+t} \frac{1+t^{1/4}}{t^{1/4}};$$

$$f_4: t \mapsto \varepsilon^{-t}; \qquad f_5: t \mapsto \frac{1}{1+t^2} \frac{1+t^{1/4}}{t^{1/4}}; \qquad f_6: t \mapsto \frac{1}{1+t^2} \frac{1+t^{1/2}}{t^{1/2}}$$

Show that these functions satisfy the inclusion relations shown on the Venn diagram in Fig. II.1.

9 **Exercise 4** Let $[a, b]$ be a bounded interval (i.e., a and b are finite). Show that $L^1[a, b] \supset L^2[a, b] \supset L^\infty[a, b]$, where

$$L^p[a, b] = \left\{ f: [a, b] \to \mathbb{R} \;\middle|\; \int_a^b |f(t)|^p \, dt < \infty \right\}$$

10 **Exercise 5** Let $1 \le p \le \infty$ and f_p: $\mathbb{R}_+ \to \mathbb{R}$ be defined by

$$f_p\colon t \mapsto \left[\frac{1}{(t)^{1/2}[1+\log(t)]}\right]^{2/p}$$

Show that if $p' \in [1, \infty)$ and if $p' \neq p$, then $f_p \notin L^{p'}$, but $f_p \in L^p$.

11 **Remark** In later applications, we shall consider functions from $\mathbb{R}_+$ into $\mathbb{R}^n$. In this case, *all* above definitions hold, except that $|f(t)|$ is interpreted as the chosen norm of the *vector* $f(t) \in \mathbb{R}^n$; one norm in $\mathbb{R}^n$ is chosen once and for all for the whole development. Since all norms in $\mathbb{R}^n$ are equivalent, norms in $\mathbb{R}^n$ are selected for numerical convenience and for obtaining best bounds.

12 **Notation** Sometimes we consider functions on $\mathbb{R}$ (or $\mathbb{R}_+$) to $\mathbb{R}$ (or $\mathbb{R}^n$). To keep track of these cases we use the symbol

$$L_n^p(\mathbb{R}_+) \triangleq \left\{ f\colon \mathbb{R}_+ \to \mathbb{R}^n \,\Big|\, \int_0^\infty |f(t)|^p\,dt < \infty \right\}.$$

and a similar definition for $L_n^p(\mathbb{R})$. In case $n = 1$, we drop the subscript n. Once it is understood that we consider $\mathbb{R}$ (or $\mathbb{R}_+$) exclusively, we then use the abbreviation L_n^p.

4 Geometric interpretation of norms

In order to obtain a more intuitive interpretation of norms, we show below that to every norm is associated a certain type of convex set. Conversely, to any such convex set is associated a norm. We start by defining some concepts.

1 Let E be a linear space, over $\mathbb{R}$ or $\mathbb{C}$. A set K in E is said to be convex iff $x, y \in K \Rightarrow \lambda x + (1-\lambda)y \in K$, $\forall \lambda \in [0, 1]$. A set K in E is said to be **balanced** (équilibré) iff $x \in K \Rightarrow \alpha x \in K$, $\forall |\alpha| \le 1$. A set K in E is said to be **absorbing** iff $x \in E \Rightarrow \exists \lambda(x) \ge 0$ such that $x \in \lambda(x)K$. The set $\{y | y = \lambda x;\ \lambda \in \mathbb{R}_+\}$ is called the **ray** Ox.

2 **Theorem** Let E be a linear space.

(A) If $N: E \to \mathbb{R}_+$ is a norm on E, then the unit ball associated with $N(\cdot)$, namely, the set

$$B = \{x \in E | N(x) \le 1\}$$

is *convex*, *balanced*, *absorbing* and B intersects every ray Ox on *a finite interval* (i.e., $\forall x \in E$, $\exists \lambda_N < \infty$ such that $Ox \cap B = \{y | y = \lambda x, 0 \le \lambda \le \lambda_N\}$).

(B) If the set $K \subset E$ is convex, balanced, absorbing and if K intersects every ray Ox on a finite interval, then the function $p_K\colon E \to \mathbb{R}_+$ defined by

3 $$p_K(x) \triangleq \inf\{\lambda | \lambda > 0 \text{ and } x \in \lambda K\}$$

is a norm on E.

4 **Remark** The ray condition is indispensable; to wit, let $K \subset \mathbb{R}^2$ and $K = \{(x, y) | x \in \mathbb{R},\ |y| \le 1\}$. Clearly K is convex, balanced, and absorbing, but p_K is not a norm [$p_K(x_0) = 0$ for $x_0 = (1, 0)$].

Proof (A) By assumption N is a norm.

(i) B is convex because $x, y \in B$ imply $N(x) \le 1$, $N(y) \le 1$. Now $\forall \lambda \in [0, 1]$.

$$N[\lambda x + (1-\lambda)y] \le N(\lambda x) + N[(1-\lambda)y] = \lambda N(x) + (1-\lambda)N(y) \le 1$$

Hence $\lambda x + (1-\lambda)y$ is in B, for all $\lambda \in [0, 1]$.

(ii) B is balanced because $x \in B \Rightarrow N(x) \le 1$. Hence $\forall |\alpha| \le 1$, $N(\alpha x) = |\alpha| N(x) \le |\alpha| \le 1$; i.e., $\alpha x \in B$ for all $|\alpha| \le 1$.

(iii) B is absorbing because for any $x \in E$, $N(x) < \infty$ and $x/N(x) \in B$; i.e., $x \in N(x)B$.

(iv) $Ox \cap B = \{\lambda x | \lambda \ge 0$ and $N(\lambda x) \le 1\}$. Now $N(\lambda x) \le 1 \Leftrightarrow |\lambda| \le 1/N(x)$; hence $\lambda x \in Ox \cap B$ iff $0 \le \lambda \le 1/N(x)$.

(B) By assumption K has the four properties; we have to show that the function p_K satisfies the axioms of a norm.

(i) Obviously by (3), $p_K(0) = 0$. Next we show that $x \neq 0 \Rightarrow p_K(x) > 0$ by contradiction. Suppose $x_0 \neq 0$ and $p_K(x_0) = 0 = \inf\{\lambda | \lambda > 0,\ x_0 \in \lambda K\}$. Hence $\forall \lambda > 0$, $x_0 \in \lambda K$; i.e., $(1/\lambda)x_0 \in K$. Hence the whole ray Ox_0 is contained in K.

(ii) Since K is absorbing, $p_K(x) < \infty$ for any $x \in E$.

(iii) $p_K(\alpha x) = |\alpha| p_K(x)$ because K is balanced. Indeed, $\alpha x \in \lambda K$ for some λ implies $|\alpha| x \in \lambda K$ or $x \in (\lambda/|\alpha|)K$. Hence

$$p_K(x) = (1/|\alpha|)p_K(\alpha x)$$

(iv) $p_K(x + y) \le p_K(x) + p_K(y)$ follows from the convexity of K and the definition of p_K. By definition of p_K, for any $\varepsilon > 0$,

$$\frac{x}{p_K(x) + \varepsilon} \quad \text{and} \quad \frac{y}{p_K(y) + \varepsilon} \in K$$

Since K is convex (we drop the subscript K for simplicity),

$$\left[\frac{p(x) + \varepsilon}{p(x) + p(y) + 2\varepsilon}\right] \frac{x}{p(x) + \varepsilon} + \left[\frac{p(y) + \varepsilon}{p(x) + p(y) + 2\varepsilon}\right] \frac{y}{p(y) + \varepsilon} \in K$$

So, $\forall \varepsilon > 0$, $x + y \subset [p(x) + p(y) + 2\varepsilon]K$; hence $p(x + y) \le p(x) + p(y)$.

::

5 Induced norms of linear maps

5.1 The space of linear maps

Let E be a linear space over the field $\mathbb{K}$, which is $\mathbb{R}$ or $\mathbb{C}$. Let $\mathscr{L}(E, E)$ be the class of all *linear* maps from E into E. $\mathscr{L}(E, E)$ is a linear space if the addition of A and B is defined by

1 $$(A + B)x = Ax + Bx, \qquad \forall x \in E, \quad \forall A, B \in \mathscr{L}(E, E)$$

and the scalar product of α and A is defined by

2 $$(\alpha A)(x) = \alpha(Ax), \qquad \forall \alpha \in \mathbb{K}, \quad \forall x \in E, \quad \forall A \in \mathscr{L}(E, E)$$

In addition, the **product** AB can be defined as the composition of the maps A and B:

3 $$(AB)(x) = A(Bx), \qquad \forall x \in E, \quad \forall A, B \in \mathscr{L}(E, E)$$

Note that the *product is not commutative*. It is easy to check that

(i) multiplication defined by (3) is associative;
(ii) addition and multiplication are distributive;
(iii) multiplication and scalar multiplication commute;
(iv) I, the identity map from E into E, is the unit of the multiplication.

Thus with the three operations defined by (1), (2), and (3), $\mathscr{L}(E, E)$ *is a (noncommutative) algebra with a unit*. (Refer to Appendix D for a discussion of algebras.)

4 **Remark** With obvious but simple modifications, the considerations of this chapter apply to linear maps from a linear space E to a linear space E. In Section 5.2, where we introduce induced norms, it becomes then crucial to distinguish between the norms on the domain and the norms on the range. To avoid cluttering the notation, we restrict ourselves to $\mathscr{L}(E, E)$, i.e., to the case where the domain and the range are the same spaces.

5.2 Induced norms

Let $|\cdot|$ be a norm on E and $A \in \mathscr{L}(E, E)$. Define the function $\|\cdot\|$ from a subset of $\mathscr{L}(E, E)$ into $\mathbb{R}_+$ by

5 $$\|A\| \triangleq \sup_{x \neq 0} |Ax|/|x|$$

6 **Fact** Definition (5) is equivalent to

7 $$\|A\| = \sup_{|z|=1} |Az|$$

8 $\|A\|$ is called the induced norm of the linear map A or *the operator norm induced by the vector norm* $|\cdot|$.

To show that (5) and (7) are equivalent, let $z = x/|x|$, then

$$|Az| = \left|A\left(\frac{x}{|x|}\right)\right| = \left|\frac{1}{|x|}Ax\right| = \frac{1}{|x|}|Ax|$$

∎

To interpret (7) geometrically, let K be the convex, balanced, absorbing set corresponding to the vector norm $|\cdot|$; i.e., the unit ball

$$K \triangleq \{x \in E \,|\, |x| \leq 1\}$$

Let AK denote the image of K under the map A. Then (7) is equivalent to

9 $$\|A\| = \inf\{\lambda \,|\, AK \subset \lambda K\}$$

Roughly speaking, $\|A\|$ is the smallest magnification factor λ for which λK includes AK.

10 **Theorem** Let $|\cdot|$ be any norm on $\mathbb{C}^n$. Let A be a nonsingular linear map from $\mathbb{C}^n$ to $\mathbb{C}^n$. Let $\|\cdot\|$ denote the induced norm on $n \times n$ matrices with elements in $\mathbb{C}$. U.t.c.

(a) There is a constant β which depends only on n and $|\cdot|$ such that

11 $$\|A^{-1}\| \leq \beta \frac{\|A\|^{n-1}}{|\det A|}$$

(b) If Euclidean norms are used, then $\beta \leq 1$.

Proof (a) From the equivalence of norms and (2.3), there are constants α_l and α_u such that

$$\alpha_l \max_{i,j} |a_{ij}| \leq \|A\| \leq \alpha_u \max_{i,j} |a_{ij}|$$

Then (11) follows from Cramer's formula.

(b) By polar decomposition

$$A = UH$$

where U is unitary (i.e., $UU^* = I$) and H is Hermitian. Note that

$$|\det A| = |\det U|\,|\det H| = |\det H|$$

Since A is nonsingular and $H^2 = A^*A$, H is positive definite. Call $h_1, h_2, \ldots, h_n$ its eigenvalues and order them so that $0 < h_1 \leq h_2 \leq \cdots \leq h_n$. By the properties of the (operator) norm induced by the Euclidean norm [see (1.18)],

12 $$\|A\| = \|H\| = h_n \qquad \text{and} \qquad \|A^{-1}\| = \|H^{-1}\| = 1/h_1$$

So

13 $$\|A^{-1}\| = \frac{1}{h_1} \le \frac{(h_n)^{n-1}}{h_1 h_2 \cdots h_n} = \frac{\|A\|^{n-1}}{\det H} = \frac{\|A\|^{n-1}}{|\det A|}$$ ∎

Note that the inequality (13) holds with equal sign if $h_2 = h_n$. In fact, the right-hand side of (13) is equal to the left-hand side times $\prod_{k=2}^{n} (h_n/h_k)$.

5.3 Continuous linear maps

Let $(E, |\cdot|)$ be a normed space over the field $\mathbb{K}$, and let $\|\cdot\|$ be the induced norm on some subspace of $\mathscr{D}(E, E)$. We define

14 $$\mathscr{L}(E, E) \triangleq \{A \in \mathscr{D}(E, E) \,|\, \|A\| < \infty\}$$

Theorem U.t.c. $\forall A,\ B \in \mathscr{L}(E, E), \forall \alpha \in \mathbb{K},\ \forall x \in E$

15 $$|Ax| \le \|A\|\ |x|$$

16 $$\|\alpha A\| = |\alpha|\ \|A\|$$

17 $$\|A + B\| \le \|A\| + \|B\|$$

18 $$\|AB\| \le \|A\|\ \|B\|$$

Proof All these inequalities follow directly from Definition (7). For example,

$$\|AB\| = \sup_{|z|=1} |ABz| \le \|A\| \sup_{|z|=1} |Bz| = \|A\|\ \|B\|$$ ∎

The inequalities above show that $\mathscr{L}(E, E)$ is a normed algebra. It can be shown that if $(E, |\cdot|)$ is complete (i.e., if E is a Banach space), then $\mathscr{L}(E, E)$ is also a *Banach algebra* with the norm (7). It is, in fact, a noncommutative Banach algebra with a unit.

Exercise 1 Let $A \in \mathscr{D}(E, E)$. Show that the following three statements are equivalent:

(i) The linear function A is continuous at $0 \in E$.
(ii) The linear function A is continuous on E.
(iii) The induced norm of A, $\|A\|$, is finite.

This exercise justifies our saying that $\mathscr{L}(E, E)$ is the class of all continuous linear maps from the normed space E into E.

Some authors call the elements of $\mathscr{L}(E, E)$ "bounded" linear maps. This terminology conflicts with the general definition of a bounded map from a normed space into another normed space. What they mean is that $A \in \mathscr{L}(E, E)$, iff $A \in \mathscr{D}(E, E)$, and the restriction of A to the unit ball is a bounded map.

Remark Suppose that on $\mathscr{D}(E, E)$ we define a norm $N(\cdot)$ which, *in addition to the axioms of the norm*, satisfies

19 $$N(AB) \le N(A)N(B)$$

and suppose that we have a vector norm $|\cdot|$ such that

20 $$|Ax| \le N(A)|x|, \qquad \forall x \in E, \quad \forall A \in \mathscr{D}(E, E)$$

Then we say that $N(\cdot)$ is an (operator) norm on $\mathscr{D}(E, E)$ compatible with the (vector) norm $|\cdot|$ on E.

Exercise 2 (i) If $\|\cdot\|$ is the norm induced by $|\cdot|$ and if $N(\cdot)$ is any norm compatible with $|\cdot|$, show that

21 $$\|A\| \le N(A), \qquad \forall A \in \mathscr{L}(E, E)$$

(ii) In the matrix case, show that $A \mapsto \max_{i,j} |a_{ij}|$ obeys the axioms of a norm on the space of matrices but not (19).
(iii) Show that $A \mapsto (\sum_{i,j} |a_{ij}|^2)^{1/2}$ is compatible with the Euclidean norm.
(iv) Refer to the examples of Section 1 and show that, for $p = 1, 2, \infty$, $\|A\|_p$ [defined by (1.17) to (1.19)] is induced by $|x|_p$.

6 Two examples

The two examples that follow are designed to illustrate the fact that the induced norm depends on the vector norm. The results of these examples will be useful later.

Example 1 Let $(E, |\cdot|_\infty) = L^\infty(\mathbb{R}_+) = \{f: \mathbb{R}_+ \to \mathbb{R} \,|\, \|f\|_\infty < \infty\}$. Let H be a linear map defined on E in terms of an integrable function $h: \mathbb{R} \to \mathbb{R}$;

1 $$H: u \mapsto Hu \triangleq h * u, \qquad \forall u \in L^\infty$$

i.e.,

2 $$(Hu)(t) = \int_0^t h(t - \tau)u(\tau)\, d\tau, \qquad \forall t \in \mathbb{R}_+$$

We assume that $\|h\|_1 = \int_0^\infty |h(t)|\, dt < \infty$.

3 **Theorem** U.t.c.

(a) $H: L^\infty \to L^\infty$
(b) $\|H\|$, the induced norm of the linear map H, is given by

$$\|H\|_\infty = \|h\|_1; \qquad \text{i.e.,} \quad \|h * u\|_\infty \le \|h\|_1 \|u\|_\infty, \quad \forall u \in L^\infty$$

and $\|h * u\|_\infty$ can be made arbitrarily close to $\|h\|_1 \|u\|_\infty$ by appropriate choice of u.

Proof We start by calculating the induced norm of H. We drop the subscript ∞ throughout; e.g., $\|u\|$ denotes the L^∞ norm of u: $\mathbb{R}_+ \to \mathbb{R}$. We have three norms in this proof: the absolute value of real numbers, e.g., $|u(t)|$; the norm on L^∞, e.g., $\|u\|$; the induced norm on $\mathscr{L}(E, E)$, namely, $\|H\|$.

4
$$\|H\| = \sup_{\|u\|=1} \|h * u\| = \sup_{\|u\|=1} \sup_{t \geq 0} |(h * u)(t)|$$
$$= \sup_{\|u\|=1} \left[\sup_{t \geq 0} \left| \int_0^t h(t-\tau)u(\tau)\, d\tau \right| \right]$$
$$\leq \sup_{\|u\|=1} \left[\sup_{t \geq 0} \int_0^t |h(t-\tau)|\, |u(\tau)|\, d\tau \right]$$

Since $\|u\| = 1$,

$$\|H\| \leq \sup_{t \geq 0} \int_0^t |h(t-\tau)|\, d\tau \leq \int_0^\infty |h(t')|\, dt'$$

Hence

5
$$\|H\| \leq \int_0^\infty |h(t')|\, dt' = \|h\|_1$$

This inequality shows that H is a continuous linear map from L^∞ into L^∞. Inequality (5) shows also that $\|h\|_1$ is an upper bound on the induced norm of the map H: $L^\infty \to L^\infty$. Let us show that $\|h\|_1$ is the induced norm. Consider a sequence of inputs $u_1, u_2, \ldots$, with $\|u_t\| = 1$, where for $t = 1, 2, 3, \ldots$, we define

$$u_t: \tau \mapsto u_t(\tau) = \operatorname{sgn}[h(t-\tau)], \qquad \tau \in \mathbb{R}_+, \quad t \in \mathbb{Z}_+$$

and we take $h(t) = 0$ for $t < 0$. Consider now the value at time t of the output due to $u_t(\cdot)$

$$(h * u_t)(t) = \int_0^t |h(t-\tau)|\, d\tau \leq \|h * u_t\|_\infty, \qquad t = 1, 2, \ldots$$

where $\|\cdot\|_\infty$ denotes the norm on E. Hence for $t = 1, 2, 3, \ldots$, we have by (5)

$$\int_0^t |h(\tau)|\, d\tau \leq \|h * u_t\|_\infty \leq \|H\| \leq \int_0^\infty |h(\tau)|\, d\tau = \|h\|_1$$

Letting $t \to \infty$, these inequalities imply that $\|H\| = \|h\|_1$. ::

Example 2 Let $E = L^2(\mathbb{R}) = \{f: \mathbb{R} \to \mathbb{R} \mid \|f\|_2 < \infty\}$. Let H be a linear map defined on L^2 by H: $u \mapsto Hu$, where

6
$$(Hu)(t) = \int_{-\infty}^\infty h(t-\tau)u(\tau)\, d\tau, \qquad \forall t \in \mathbb{R}$$

7 **Theorem** If the linear map H is defined by (6), where $h \in L^1$, then

(a) H: $L^2 \to L^2$,

(b) $\|H\|_2$, the induced norm of the linear map $H \in \mathscr{L}(L^2, L^2)$, is given by

8
$$\|H\|_2 = \max_{\omega \in \mathbb{R}} |\hat{h}(j\omega)|$$

Proof Since $h \in L^1$, its Fourier transform $\mathscr{F}(h) = \hat{h}(j\omega)$ is uniformly continuous on $\mathbb{R}$ and $\to 0$ as $|\omega| \to \infty$. (See Appendix B.1.1.) Now, with all integrals below being over $\mathbb{R}$, we have

$$\|Hu\|_2^2 = \|h * u\|_2^2 = \int (h * u)(t)(h * u)(t)\, dt$$
$$= \frac{1}{2\pi} \int \widehat{(h * u)}(j\omega)\widehat{(h * u)}(j\omega)^*\, d\omega \qquad \text{(Parseval)}$$
$$= \frac{1}{2\pi} \int |\hat{h}(j\omega)|^2 |\hat{u}(j\omega)|^2\, d\omega \qquad \text{(convolution theorem)}$$

By Parseval,

$$\|u\|_2 = 1 \Leftrightarrow \|\hat{u}\|_2 = (2\pi)^{1/2} \Leftrightarrow (2\pi)^{-1} \int |\hat{u}(j\omega)|^2\, d\omega = 1$$

Hence, $\forall u$ such that $\|u\|_2 = 1$,

$$\|Hu\|_2^2 \leq \max_{\omega \in \mathbb{R}} [|\hat{h}(j\omega)|^2]$$

Thus, the induced norm satisfies

9
$$\|H\|_2 \leq \max_{\omega \in \mathbb{R}} |\hat{h}(j\omega)|$$

Note that since $\omega \mapsto |\hat{h}(j\omega)|$ is continuous on $\mathbb{R}$ and $\to 0$ as $|\omega| \to \infty$, the maximum exists. We are going to show that $\|H\|_2$ is actually equal to the right-hand side of (9). Observe that for $\lambda > 0$,

$$\mathscr{F}(e^{-\lambda t^2}) = (\pi/\lambda)^{1/2} \exp[-\omega^2/4\lambda]$$

and

$$\mathscr{F}[\exp(-\lambda t^2) \cos \omega_0 t]$$
$$= \tfrac{1}{2}(\pi/\lambda)^{1/2}\{\exp[-(\omega - \omega_0)^2/4\lambda] + \exp -[(\omega + \omega_0)^2/4\lambda]\}$$

As $\lambda \to 0$, this expression tends to $\pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$, where $\delta(\cdot)$ denotes the Dirac delta (generalized) function. Pick ω_0 to be the abscissa of the maximum of $\omega \mapsto |\hat{h}(j\omega)|$; for each λ pick a normalization $n(\lambda)$ such that

$$u_\lambda(t) = n(\lambda)[\exp(-\lambda t^2)] \cos \omega_0 t$$

has unit norm. Since $|\hat{h}(j\omega)|$ is continuous, we see that

$$\|h * u_\lambda\|_2 \to \max_\omega |\hat{h}(j\omega)| \qquad \text{as} \quad \lambda \to 0$$

Consequently, we have shown that

$$\|H\|_2 = \max_\omega |\hat{h}(j\omega)|$$ ::

0 Remark These results can be generalized to the case where u and $h * u$ are vector valued; i.e., map $\mathbb{R}_+$ into $\mathbb{R}^n$. We state the result as an exercise.

Exercise 1 Let $u: \mathbb{R}_+ \to \mathbb{R}^n$ and let u be locally integrable. Let H be the matrix impulse response; so $H: \mathbb{R}_+ \to \mathbb{R}^{n \times n}$. Assume throughout that the elements of H, namely, the $h_{ij}(\cdot)$'s, are in L^1, for $i, j = 1, 2, \ldots, n$. Denote by $\bar{H}$ the linear operator defined by

$$\bar{H}: u \mapsto \bar{H}u \triangleq H * u$$

where

$$(H * u)(t) = \int_0^\infty H(t - \tau)u(\tau)\, d\tau$$

Establish the following induced norms:

(a) $u \in L_n^\infty(\mathbb{R}_+)$: $\|u\|_\infty = \max_i \sup_{t \geq 0} |u_i(t)|$; show that the induced norm of $\bar{H}$ is

11 $$\|\bar{H}\|_\infty = \max_i \int_0^\infty \sum_{j=1}^n |h_{ij}(\tau)|\, d\tau \qquad \text{(row sum)}$$

(b) $u \in L_n^2(\mathbb{R}_+)$: $\|u\|_2^2 = \int_0^\infty \sum_{i=1}^n |u_i(t)|^2\, dt$; show that the induced norm of $\bar{H}$ is

12 $$\|\bar{H}\|_2 = (\lambda_{max})^{1/2}$$

where

13 $$\lambda_{max} = \max_\omega \max_i \lambda_i[\hat{H}(j\omega)^* \hat{H}(j\omega)]$$

where $\lambda_i(M)$ denotes the (necessarily real) ith eigenvalue of the Hermitian matrix M.

(c) $u \in L_n^1(\mathbb{R}_+)$: $\|u\|_1 = \int_0^\infty \sum_{i=1}^n |u_i(t)|\, dt$; show that the induced norm of $\bar{H}$ is

14 $$\|\bar{H}\|_1 = \max_j \int_0^\infty \sum_{i=1}^n |h_{ij}(t)|\, dt \qquad \text{(column sum)}$$

7 Norms and spectral radius

Let $A \in \mathbb{C}^{n \times n}$ and

1 $$r(A) \triangleq \max_i |\lambda_i(A)|$$

where $r(A)$ is called the **spectral radius** *of* A.

2 Exercise 1 Show that for *any* induced norm, $r(A) \leq \|A\|$.

For the given matrix A, we may try to find a norm on $\mathbb{C}^n$ that gives the "minimum" induced norm for A. In fact, we shall prove the following theorem.

3 Theorem Let $\mathcal{N}$ denote the set of all norms on $\mathbb{C}^n$; then for any $A \in \mathbb{C}^{n \times n}$,

4 $$\inf_{|\cdot| \in \mathcal{N}} \left[\sup_{x \in \mathbb{C}^n} (|Ax|/|x|)\right] = r(A)$$

By the definition of the infimum, (4) is equivalent to inequality (2) and the statement that for any $\varepsilon > 0$ and any $A \in \mathbb{C}^{n \times n}$, there is a (vector) norm on $\mathbb{C}^n$ such that the corresponding induced norm satisfies

5 $$\|A\| \leq r(A) + \varepsilon$$

Proof Let J be the Jordan form of A; then there is a *nonsingular* matrix P s.t.

$$J = PAP^{-1} = \Lambda + U$$

where Λ is diagonal and all the elements of U are either 0 or 1 with all nonzero elements located on the diagonal *above* the main diagonal. For some small $\delta > 0$, define the nonsingular matrix D by

$$D \triangleq \text{diag}[1, \delta^{-1}, \delta^{-2}, \ldots, \delta^{-(n-1)}]$$

Then

$$DJD^{-1} = \Lambda + \delta U$$

Define a norm $|\cdot|$ on $\mathbb{C}^n$ by

$$x \mapsto |DPx|_2$$

where $|\cdot|_2$ is the Euclidean norm on $\mathbb{C}^n$. The corresponding induced norm is

$$\|A\| = \max_{|x|=1} |Ax| = \max_{|DPx|_2=1} |DPAx|_2$$

With $z \triangleq DPx$, we have $DPAx = DPAP^{-1}D^{-1}z = (\Lambda + \delta U)z$; consequently,

$$\begin{aligned}\|A\|^2 &= \max_{|z|_2=1} \langle(\Lambda + \delta U)z|(\Lambda + \delta U)z\rangle \\ &\leq \max_{|z|_2=1} \{|\Lambda z|_2^2 + 2\delta|\langle Uz|\Lambda z\rangle| + \delta^2|Uz|_2^2\} \\ &\leq r(A)^2 + 2\delta r(A) + \delta^2\end{aligned}$$

or

$$\|A\| \leq r(A) + \delta$$

Therefore, by choosing $\delta \leq \varepsilon$, we achieve (5). ::

The trouble with this result is that the vector norm must be specially tailored for the matrix A. It is, however, useful in applications where the problem involves only one matrix; for in that case the norm can be tailored to that matrix.

Exercise 2 Let N be a norm on $\mathbb{C}^n$. Show that $x \mapsto N(Px)$ is a norm on $\mathbb{C}^n$ if and only if P is a *nonsingular* matrix $\in \mathbb{C}^{n\times n}$. Show that if P is singular, then $\rho: x \mapsto N(Px)$ satisfies axioms (ii) and (iii) of the norms [see (II.1.1)], but (i) is replaced by

(i′) $\rho(x) \geq 0, \qquad \forall x \in \mathbb{C}^n$.

(In such a case, ρ is called a seminorm.)

Exercise 3 (contraction mapping theorem) The following is a statement of the contraction mapping theorem in a form particularly useful for applications. The proof is essentially the same as the usual textbook case.

Let (P, d) be a metric space and $(\mathscr{B}, \|\cdot\|)$ be a Banach space. Let $p_0 \in P$ and $x_0 \in \mathscr{B}$. Consider the closed balls:

$$B_P = \{p \in P | d(p - p_0) \leq r_P\}$$
$$B_{\mathscr{B}} = \{x \in \mathscr{B} | \|x - x_0\| \leq r_{\mathscr{B}}\}$$

If

(i) $f: B_P \times B_{\mathscr{B}} \to \mathscr{B}$ and f is continuous in $B_P \times B_{\mathscr{B}}$;

(ii) there is some $k < 1$ such that

$$\|f(p, x) - f(p, x')\| \leq k\|x - x'\|, \qquad \forall p \in B_P, \quad \forall x, x' \in B_{\mathscr{B}}$$

(iii) $\|f(p, x_0) - x_0\| \leq (1 - k)r_{\mathscr{B}}, \qquad \forall p \in B_P$

then

(a) $\forall p \in B_P$, the iteration scheme

$$x_{n+1}(p) = f[p, x_n(p)], \qquad \text{with} \quad x_0(p) = x_0$$

converges to a *unique continuous* function $\bar{x}: B_P \to B_{\mathscr{B}}$ such that

$$\bar{x}(p) = f[p, \bar{x}(p)], \qquad \forall p \in B_P$$

(b) the convergence is uniform in p over B_P; i.e., for the given k and the given B_P,

$$\|\bar{x}(p) - x_n(p)\| \leq k^n r_{\mathscr{B}}, \qquad \forall p \in B_P, \quad \forall n \in \mathbb{Z}_+$$

(*Hint*: Be sure to verify that $x_n(p) \in B_{\mathscr{B}}$ for all $n \geq 1$, and all $p \in B_P$.)

In general, any function such as $f(p, \cdot)$, which maps a set $B_{\mathscr{B}}$ into a Banach space $\mathscr{B}$ and which satisfies an inequality like (ii) above, is called a contraction over $B_{\mathscr{B}}$.

Exercise 4 Let A be a continuous linear map from a Banach space E into itself. We say that $\lambda \in \mathbb{C}$ is a spectral value of A iff $A - \lambda I$ does not have a continuous inverse on E. Call the spectrum of A, $\mathrm{Sp}(A)$, the set of all spectral values of A, and the spectral radius of A, the positive number

$$r(A) \triangleq \sup_{\lambda \in \mathrm{Sp}(A)} |\lambda|$$

(a) Show that

$$(A - \lambda I)^{-1} = -\sum_{k=0}^{\infty} \lambda^{-(k+1)}A^k$$

and that the series converges absolutely [in $\mathscr{L}(E; E)$] for $|\lambda| > \|A\|$.

(b) Show that $(A - \lambda I)^{-1} \in \mathscr{L}(E; E)$ for $|\lambda| > \mathrm{Sp}(A)$.

Exercise 5 Let $A, B \in \mathscr{L}(E; E)$, E a Banach space. Show that it is *not* generally true that

$$r(AB) \leq \|A\| r(B), \qquad r(AB) \leq r(A) r(B)$$

(*Hint*: Consider 2×2 matrices, one of which is nil potent.)

8 The measure of a matrix

Let $|\cdot|$ be some norm on $\mathbb{C}^n$. Let $\|\cdot\|$ denote the corresponding induced norm on the $n \times n$ matrices in $\mathbb{C}^{n\times n}$. The function $\|\cdot\|: \mathbb{C}^{n\times n} \to \mathbb{R}_+$ is a convex function; therefore, at every point $X \in \mathbb{C}^{n\times n}$, $\|\cdot\|$ has a *one-sided directional* derivative in any direction $A \in \mathbb{C}^{n\times n}$; i.e., the limit

$$\lim_{\theta \searrow 0} (\|X + \theta A\| - \|X\|)/\theta$$

exists for all X and all A.

The one-sided directional derivative of $\|\cdot\|$ at $I \in \mathbb{C}^{n\times n}$ in the direction A is called the measure of the matrix A and is denoted by $\mu(A)$. Thus,

1 $$\mu(A) \triangleq \lim_{\theta \searrow 0}(\|I + \theta A\| - 1)/\theta$$

We shall prove the existence of this limit in Lemma (4) below.

The concept of the measure of a matrix appears naturally in the study of differential equations. Consider $\dot{x}(t) = A(t)x(t)$, where $A(\cdot)$ is continuous. Let us bound the one-sided derivative in the positive direction of $t \mapsto |\phi(t)|$, where $\phi(\cdot)$ is a solution of the differential equation:

2 $$\begin{aligned} D^+|\phi(t)| &\triangleq \lim_{\theta \searrow 0}[|\phi(t + \theta)| - |\phi(t)|]/\theta \\ &= \lim_{\theta \searrow 0}[|\phi(t) + \theta A(t)\phi(t)| - |\phi(t)|]/\theta \\ &\le \lim_{\theta \searrow 0}[\|I + \theta A(t)\|\,|\phi(t)| - |\phi(t)|]/\theta \end{aligned}$$

Hence by (1)

3 $$D^+|\phi(t)| \le \mu(A(t))|\phi(t)|$$

Inequality (3) is sharper than the standard one

$$D^+|\phi(t)| \le \|A(t)\|\,|\phi(t)|$$

Indeed, $\mu(A(t))$ may be negative, whereas $\|A(t)\|$ is always nonnegative.

4 **Lemma** For any $A \in \mathbb{C}^{n\times n}$, $\mu(A)$ is well defined.

Proof Call $f(\theta)$ the ratio in (1); we show that $f(\theta)$ decreases as $\theta \searrow 0$ and that it is bounded below. Hence as θ decreases to 0, $f(\theta)$ tends to a well-defined limit. Let $k \in (0, 1)$.

$$\begin{aligned} k\theta f(k\theta) &= \|I + k\theta A\| - 1 = \|k(I + \theta A) + (1 - k)I\| - 1 \\ &\le k\|I + \theta A\| + 1 - k - 1 = k(\|I + \theta A\| - 1) \\ &= k\theta f(\theta) \end{aligned}$$

So for $k \in (0, 1)$, $f(k\theta) \le f(\theta)$; i.e., f is decreasing as θ decreases. Now $f(\theta) \ge -\|A\|$ because, for $\theta > 0$,

$$\theta f(\theta) = \|I + \theta A\| - 1 \ge 1 - \theta\|A\| - 1 = -\theta\|A\|$$

Hence the conclusion follows. ∎

Properties of μ

For convenience, we list a number of properties of μ in the following theorem.

5 **Theorem** Let $A, B \in \mathbb{C}^{n\times n}$, and μ be given by (1). Then

6 (a) $\mu(I) = 1, \quad \mu(-I) = -1, \quad \mu(0) = 0$

(*Note*: $\mu(A) = 0$ *does not imply that* $A = 0$.)

7 (b) $-\|A\| \le -\mu(-A) \le \mu(A) \le \|A\|$

8 (c) $\mu(cA) = c\mu(A), \quad \forall c \ge 0$

9 (d) $\mu(A + cI) = \mu(A) + c, \quad \forall c \in \mathbb{R}$

10 (e) $\max[\mu(A) - \mu(-B), -\mu(-A) + \mu(B)] \le \mu(A + B) \le \mu(A) + \mu(B)$

(f) $\mu: \mathbb{C}^{n\times n} \to \mathbb{R}$ is convex on $\mathbb{C}^{n\times n}$

11 $\mu[\lambda A + (1 - \lambda)B] \le \lambda\mu(A) + (1 - \lambda)\mu(B), \quad \forall\lambda \in [0, 1]$

12 (g) $|\mu(A) - \mu(B)| \le |\mu(A - B)| \le \|A - B\|$

$|\mu(A) - \mu(B)| \le |\mu(B - A)| \le \|A - B\|$

13 (h) $-\mu(-A) \le \operatorname{Re}\lambda_i(A) \le \mu(A), \quad \forall i \in \{1, 2, \ldots, n\}$

14 (i) $-\mu(-A)|x| \le |Ax|$ and $-\mu(A)|x| \le |Ax|, \quad \forall x \in C^n$

(*Note*: $|Ax|$ *is not bounded above* by $\mu(A)|x|$.)

(j) Let $|\cdot|$ be a norm on $\mathbb{C}^n$ and $P \in \mathbb{C}^{n\times n}$ be nonsingular. Call μ_P the measure defined in terms of the induced norm corresponding to the vector norm $|\cdot|_P$ defined by $x \mapsto |x|_P = |Px|$. Then

15 $$\mu_P(A) = \mu(PAP^{-1})$$

(k) If A is nonsingular,

16 $$-\mu(-A) \le (\|A^{-1}\|)^{-1} \le \|A\|$$

Proof (a) Immediate from Definition (1).

(b) The triangle inequality and $\theta > 0$ give

$$-\|A\| = \frac{1 - \theta\|A\| - 1}{\theta} \le \frac{\|I + \theta A\| - 1}{\theta} \le \frac{1 + \theta\|A\| - 1}{\theta} = \|A\|$$

and

$$-\|A\| = \frac{1 - \theta\|A\| - 1}{\theta} = \frac{1 - \theta\|-A\| - 1}{\theta} \le -\frac{\|I + \theta(-A)\| - 1}{\theta}$$

Finally, the relation between $\mu(A)$ and $-\mu(-A)$ follows from

$$0 = \|I + \theta(A - A)\| - 1 \le (\|I + 2\theta A\| - 1 + \|I - 2\theta A\| - 1)/2$$

(c) (8) is true for $c = 0$ in view of (6). For $c > 0$, observe that

$$(\|I + \theta cA\| - 1)/\theta = c(\|I + c\theta A\| - 1)/c\theta$$

Since $c > 0$, as $\theta \searrow 0$, so does $c\theta$.

(d) Consider

$$\frac{\|I + \theta(A + cI)\| - 1}{\theta} = \frac{(1 + c\theta)\|I + [\theta/(1 + c\theta)]A\| - 1}{\theta}$$

$$= \frac{\|I + [\theta/(1 + c\theta)]A\| - 1}{\theta/(1 + c\theta)} + c$$

Finally, $\forall c \in \mathbb{R}$, as $\theta \searrow 0$, $\theta/(1 + c\theta) \searrow 0$.

(e) For the second inequality

$$\|I + \theta(A + B)\| - 1 = \tfrac{1}{2}(\|I + 2\theta A + I + 2\theta B\| - 2)$$
$$\leq \tfrac{1}{2}(\|I + 2\theta A\| - 1) + \tfrac{1}{2}(\|I + 2\theta B\| - 1)$$

With $\theta \searrow 0$, we conclude $\mu(A + B) \leq \mu(A) + \mu(B)$.

The other inequality in (10) follows from the one just proved:

3 $$\mu(A) = \mu(A + B - B) \leq \mu(-B) + \mu(A + B)$$

9 $$\mu(B) = \mu(A + B - A) \leq \mu(-A) + \mu(A + B)$$

(f) Convexity is immediate from (8) and the second inequality (10).

(g) The second inequality (12) follows from (7). To obtain the first, replace B by $-B$ in (18) and obtain

$$\mu(A) - \mu(B) \leq \mu(A - B)$$

Replacing $-A$ by A in (19), we obtain

$$\mu(B) - \mu(A) \leq \mu(B - A)$$

and (12) follows.

(h) Let $e \in \mathbb{C}^n$ be a normalized eigenvector of A associated with the eigenvalue λ_i; so $Ae = \lambda_i e$ and $\|I + \theta(-A)\| \geq |e - \theta\lambda_i e|$. Therefore,

$$-\frac{\|I + \theta(-A)\| - 1}{\theta} \leq -\frac{|e - \theta\lambda_i e| - 1}{\theta} = -\frac{|1 - \theta\lambda_i| - 1}{\theta}$$

As $\theta \searrow 0$, the right-hand side tends to $\operatorname{Re} \lambda_i$ and the left to $-\mu(-A)$. To obtain the second inequality, consider

$$\frac{\|I + \theta A\| - 1}{\theta} \geq \frac{|e + \theta\lambda_i e| - 1}{\theta} = \frac{|1 + \theta\lambda_i| - 1}{\theta}$$

Again as $\theta \searrow 0$, we obtain the second inequality.

(i) For $\theta > 0$

$$|Ax| = |x - (x - \theta Ax)|/\theta = |x - (I - \theta A)x|/\theta \leq (|x| - \|I - \theta A\|\,|x|)/\theta$$
$$= -[(\|I + \theta(-A)\| - 1)|x|]/\theta \underset{\theta \searrow 0}{\rightarrow} -\mu(-A)|x|$$

where in the last step we noted that the left-hand side is independent of θ. The second inequality (i) follows from the derivation above by changing A into $-A$ and noting that $|Ax| = |-Ax|$.

(j) $$\|I + \theta A\|_P \triangleq \sup_{|x|_P = 1} |x + \theta Ax|_P \neq$$
$$= \sup_{|Px| = 1} |Px + \theta PAP^{-1}Px| \triangleq \|I + \theta PAP^{-1}\|$$

(k) By (14)

$$-\mu(-A)|x| \leq |Ax| \quad \text{and} \quad \inf_{|x|=1} |Ax| = (\|A^{-1}\|)^{-1}$$

the first inequality follows. For the second, take the norm of $AA^{-1} = I$. ∎

24 **Theorem** Let $x = (x_1, x_2, \ldots, x_n) \in \mathbb{C}^n$ and $A = (a_{ij}) \in \mathbb{C}^{n \times n}$; then

$\lvert x\rvert_1 = \sum_i^n \lvert x_i\rvert$	$\lVert A\rVert_1 = \max_j \sum_i \lvert a_{ij}\rvert$ (column sum)	$\mu_1(A) = \max_j \left[\operatorname{Re}(a_{jj}) + \sum_{i \neq j} \lvert a_{ij}\rvert\right]$
$\lvert x\rvert_2 = \left(\sum_i^n \lvert x_i\rvert^2\right)^{1/2}$	$\lVert A\rVert_2 = \left(\max_i \lambda_i(A^*A)\right)^{1/2}$	$\mu_2(A) = \max_i [\lambda_i(A + A^*)/2]$
$\lvert x\rvert_\infty = \max_i \lvert x_i\rvert$	$\lVert A\rVert_\infty = \max_i \sum_j \lvert a_{ij}\rvert$ (row sum)	$\mu_\infty(A) = \max_i \left[\operatorname{Re}(a_{ii}) + \sum_{j \neq i} \lvert a_{ij}\rvert\right]$

Proof The calculation of $\mu_i(A)$ ($i = 1, 2, \infty$) is a simple exercise using the definition. (A^* denotes, as usual, the complex conjugate of A.) ∎

Exercise 1 Show that if $A \neq 0$ and if A is skew Hermitian, or skew symmetric, then $\mu_2(A) = 0$.

Exercise 2 Let $A \in \mathbb{R}^{n \times n}$ be a singular, positive semidefinite, symmetric matrix. Show that $\mu_2(-A) = 0$.

26 Comments The formulas of Theorem (24) show that $\mu(A)$ is easy to calculate for $p = 1, \infty$ or to estimate for $p = 2$. Also $\mu(A)$ may actually be smaller than the corresponding $\|A\|$. Also, $\mu(A)$ may be negative.

As applications of the notion of measure we give two facts. The first one gives *upper* and *lower* bounds on solutions of linear differential equations, and the second gives a sufficient condition for the existence and uniqueness of operating points in circuit theory.

27 **Theorem** Let $t \mapsto A(t)$ be a regulated function from $\mathbb{R}_+$ to $\mathbb{C}^{n\times n}$. Then the solution of

28 $$\dot{x}(t) = A(t)x(t)$$

satisfies the inequalities

29 $$|x(t_0)|\exp\left\{-\int_{t_0}^{t}\mu[-A(t')]\,dt'\right\} \le |x(t)| \le |x(t_0)|\exp\int_{t_0}^{t}\mu[A(t')]\,dt'$$

Proof Let $x(t)$ denote any nonzero solution of (28), let $n(t) = |x(t)|$ and $D^+n(t)$ denote the right-hand derivative of $n(\cdot)$ at t; then by definition, for all $t \notin D$, where D is the at most countable set of points at which $A(\cdot)$ is discontinuous,

30 $$D^+n(t) = \lim_{\theta\searrow 0}[n(t+\theta) - n(t)]/\theta = \lim_{\theta\searrow 0}[|x(t) + \theta A(t)x(t)| - |x(t)|]/\theta$$

But

$$|x(t) + \theta A(t)x(t)| \le \|I + \theta A(t)\|\,|x(t)|$$

Inserting this inequality in (30) and using the definition of μ, we obtain

31 $$D^+n(t) \le \mu[A(t)]n(t)$$

Since $n(t) > 0$ for all t, (31) becomes

$$D^+n(t)/n(t) \le \mu[A(t)]$$

and the second inequality (29) follows by integration. The first inequality is proved in a similar manner. ::

In many applications one has to find the operating point (or equilibrium point) for a system described by

32 $$\dot{x} = f(x) + u$$

where u, $x \in \mathbb{R}^n$ and $f\colon \mathbb{R}^n \to \mathbb{R}^n$. The problem is to find an $x \in \mathbb{R}^n$ such that $f(x) = -u$. It is important to know whether this equation has solutions for any $u \in \mathbb{R}^n$ and whether such solutions are unique. The theorem below gives sufficient conditions for this to be the case. $Df(x)$ denotes the derivative of f at x (equivalently, the Jacobian matrix of f at x).

33 **Theorem** Let $f\colon \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable. U.t.c., if there exists a function $m\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that

34 $$\mu[Df(x)] \le -m(|x|) < 0, \qquad \forall x \in \mathbb{R}^n$$

where $m(\alpha) > 0$ for all $\alpha > 0$ and

35 $$\int_0^\infty m(\alpha)\,d\alpha = \infty$$

then $x \mapsto f(x)$ is a C^1 diffeomorphism of $\mathbb{R}^n$ onto itself [equivalently, f is a continuously differentiable bijection of $\mathbb{R}^n$ *onto* $\mathbb{R}^n$ with a continuously differentiable inverse]; thus, for any $u \in \mathbb{R}^n$, (32) has a solution x which depends in a C^1 fashion on x.

Proof By Palais' theorem our claim will be established if we show that $\det[Df(x)] \ne 0$, $\forall x \in \mathbb{R}^n$ and that $|x| \to \infty$ implies that $|f(x)| \to \infty$. The first condition follows from the observation that $\forall y \in \mathbb{R}^n$ with $y \ne 0$,

$$|Df(x)y| = |-Df(x)y| \ge -\mu[Df(x)]|y| \ge m(|x|)|y| > 0$$

where we used (14) and assumption (34).

The second requirement holds true because by using Taylor's theorem about 0, the origin of $\mathbb{C}^n$, we obtain

$$f(x) = f(0) + \left[\int_0^1 Df(\lambda x)\,d\lambda\right]x$$

So

$$\begin{aligned}
|f(x)| &\ge \left|\left[\int_0^1 Df(\lambda x)\,d\lambda\right]x\right| - |f(0)| \\
&\ge -\mu\left\{\int_0^1 Df(\lambda x)\,d\lambda\right\}|x| - |f(0)| && \text{by (14)} \\
&\ge -\int_0^1 \mu[Df(\lambda x)]\,d\lambda\,|x| - |f(0)| && \text{by (11)} \\
&\ge \int_0^1 m(\lambda|x|)\,d\lambda\,|x| - |f(0)| && \text{by (34)} \\
&= \int_0^{|x|} m(\alpha)\,d\alpha - |f(0)|
\end{aligned}$$

and the second requirement holds as a consequence of (35). ::

36 **Remark** The theorem still holds if (34) is replaced by

37 $$\mu[-Df(x)] \le -m(|x|) < 0, \qquad \forall x \in \mathbb{R}^n$$

Notes and references

The discussion of norms and equivalent norms is standard [Yos.2, Die.1, Edw.1, Hou.1]. The examples may be illuminating. The calculations of induced norms were introduced in the engineering literature by Sandberg, in his many papers starting in 1963, and by Zames at about the same time. (See reference list.) The notion of measure of a matrix is due to Dahlquist [Dah.1]. Inequality (8.29) can be found in Coppel's book [Cop.1]. For Theorem (8.32) see the paper by Desoer and Haneda [Des.11].