Topics in Applied Linear Algebra - Part II

April 23, 2013

Some Preliminary Remarks

The purpose of these notes is to provide a guide through the material for the second part of the graduate module HM802 *Topics in Applied Linear Algebra*. The material presented is by and large classical and the exposition draws heavily on the texts [1, 2, 3]. The proofs of a number of results are included in full; where proofs are omitted, the student is encouraged to either provide the details themselves or to consult the above references to find a proof. Typically, if a proof is requested as an exercise, the student should at least attempt to prove the result themselves before consulting the texts. In addition to the exercises provided here, sample problems will be circulated and these should be worked through during the period between the end of the classes and the examination. Finally, I am confident that the notes contain omissions and typos, and I will be grateful if you communicate these to me as you find them!

Chapter 1

Norms, Spectral Radii and Eigenvalues

1.1 Definitions, Examples and Elementary Results

It is very important for many applications to be able to describe the size of vectors and matrices. For instance, in numerical analysis, we often need to be able to quantify the size of errors that arise in the course of running the algorithm. Similarly to study the rate at which iterative methods converge, some means of measuring the distance to the final, steady state vector is required. The same issue is of central important in stability theory and applications in control.

Norms can be used to measure the length or size of vectors or matrices. In general, there are a variety of norms that can be defined on a vector space and the general definition given below tries to capture the essential features of our notions of length and size.

Definition 1.1.1 *A norm $\|\cdot\|$ on a vector space V is a function from V to $\mathbb{R}$ satisfying the following properties:*

(i) $\|x\| \geq 0$ *for all* $x \in V$*;*

(ii) $\|x\| = 0$ *if and only if* $x = 0$*;*

(iii) $\|\lambda x\| = |\lambda|\|x\|$ *for all* $x \in V$ *and* $\lambda \in \mathbb{C}$*;*

(iv) $\|x + y\| \leq \|x\| + \|y\|$ *for all* x, y *in* V*.*

In this course, we shall only be concerned with norms on finite dimensional vector spaces. In fact, the norms we study will be defined on either the spaces $\mathbb{R}^n$, $\mathbb{C}^n$ or the matrix spaces $\mathbb{R}^{n\times n}$, $\mathbb{C}^{n\times n}$.

Before discussing some well-known examples of norms, we note the following simple fact.

Lemma 1.1.1 *Let* $\|\cdot\|$ *be a norm on a vector space* V*. Then for any* $x, y \in V$,

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

Proof:
As $x = y + (x - y)$, the triangle inequality implies that

$$\|x\| \leq \|y\| + \|x - y\|$$

from which we see that

$$\|x\| - \|y\| \leq \|x - y\|. \tag{1.1}$$

A similar calculation using $y = x + (y - x)$ yields

$$\|y\| - \|x\| \leq \|x - y\|. \tag{1.2}$$

Combining (1.1), (1.2) yields the result.

Example 1.1.1 *The most familiar norm on* $\mathbb{C}^n$ *(or* $\mathbb{R}^n$*) is the Euclidean or* l_2 *norm given by*

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}. \tag{1.3}$$

Two other well-known norms are the infinity norm

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|, \tag{1.4}$$

and the l_1 *norm*

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|. \tag{1.5}$$

The three norms introduced above are all special cases of a general family of norms given by

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}. \tag{1.6}$$

Proving that the l_2, l_1 *and* l_∞ *norms satisfy the axioms of a norm is relatively straightforward with the triangle inequality the only non-trivial exercise in each case. For the general* l_p *norm, proving the triangle inequality is a little more involved and the standard method of proof relies on using Hölder's inequality or properties of convex functions.*

Exercise: Verify that the l_1 and l_∞ norms all satisfy the axioms of a norm.

Inner Products

The l_2 norm has a special property; it arises naturally from the usual 'dot-product'

$$\langle z, w\rangle = w^* z = \sum_{i=1}^{n} \bar{w}_i z_i.$$

This is an example of an inner product, the general definition of which is given below.

Definition 1.1.2 *An inner product $\langle \cdot, \cdot \rangle$ on a vector space V is a mapping from $V \times V$ to $\mathbb{C}$ with the properties:*

(i) $\langle z + v, w\rangle = \langle z, w\rangle + \langle v, w\rangle$;

(ii) $\langle cz, w\rangle = c\langle z, w\rangle$;

(iii) $\langle w, z\rangle = \overline{\langle z, w\rangle}$;

(iv) $\langle z, z\rangle \geq 0$ *and* $\langle z, z\rangle = 0$ *if and only if* $z = 0$.

The following fundamental, and well-known, fact about inner products is known as the *Cauchy-Schwartz* inequality.

Proposition 1.1.1 *Let $\langle \cdot, \cdot \rangle$ be an inner product on V. Then for any x, y in V,*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle. \tag{1.7}$$

Moreover, the above inequality is strict unless y and x are linearly dependent.

Proof:
If y is zero, the result is trivial so we assume that $y \neq 0$. Then for any real number t

$$\langle x + ty, x + ty\rangle \geq 0.$$

Expanding this expression shows that

$$t^2\langle y, y\rangle + 2t\mathrm{Re}(\langle x, y\rangle) + \langle x, x\rangle \geq 0$$

for all $t \in \mathbb{R}$. As the above quadratic expression is non-negative for all real t, we must have

$$4(\mathrm{Re}(\langle x, y\rangle))^2 - 4\langle y, y\rangle\langle x, x\rangle \leq 0.$$

To see that (1.7) holds, note that the above inequality must also be true if we replace y with $\langle x, y\rangle y$. This implies that

$$4(\mathrm{Re}(\overline{\langle x, y\rangle}\langle x, y\rangle))^2 \leq 4|\langle x, y\rangle|^2\langle y, y\rangle\langle x, x\rangle.$$

This immediately yields (1.7).
Finally, note that the inequality is strict unless for some t,

$$\langle x + ty, x + ty\rangle = 0.$$

This implies that $x + ty = 0$ and hence x, y are linearly dependent.

There is a norm on V naturally associated with any inner product $\langle \cdot, \cdot \rangle$, given by $\|x\| = \sqrt{\langle x, x \rangle}$.

Exercise: Prove that $\|x\| = \sqrt{\langle x, x \rangle}$ satisfies the axioms of a norm on V.

Exercise: Show that for any two norms $\| \cdot \|_a$, $\| \cdot \|_b$, $\|x\| = \max\{\|x\|_a, \|x\|_b\}$ is also a norm.

Exercise: Show that for any norm $\| \cdot \|$, and non-singular linear operator or matrix T, $\|x\|_T = \|Tx\|$ is a norm.

Exercise: Show that the mapping f on $\mathbb{R}^3$ defined by $f(x_1, x_2, x_3) = \sqrt{(2x_1)^2 + (x_1 - 3x_2)^2 + (3x_1 + x_2 - 2x_3)^2}$ is a norm.

Continuity

We have introduced a norm $\| \cdot \|$ as a function from V to $\mathbb{R}$. For convenience, we shall write $\mathbb{F}$ to denote either the base field $\mathbb{R}$ or $\mathbb{C}$. Also recall that we can naturally map any finite n-dimensional vector space isomorphically to $\mathbb{F}^n$ by fixing a basis.
The next result notes that any norm on $\mathbb{F}^n$ is *continuous* with respect to the simple l_∞ norm above. This elementary fact has important consequences and significantly simplifies the study of convergence for iterative processes and dynamical systems in finite dimensions.

Proposition 1.1.2 *Let $\| \cdot \|$ be any norm on $\mathbb{F}^n$. Then $\| \cdot \|$ is Lipschitz continuous with respect to the infinity norm $\| \cdot \|_\infty$.*

Proof: Let $x, y \in \mathbb{F}^n$ be given. It follows from Lemma 1.1.1 that

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

Next, apply the triangle inequality and the homogeneity property of norms to conclude that

$$\|x - y\| \leq \sum_{i=1}^{n} |x_i - y_i| \|e_i\|$$

where e_i is the ith standard basis vector in $\mathbb{F}^n$. If we combine the previous two inequalities, we see immediately that

$$|\|x\| - \|y\|| \leq K \|x - y\|_\infty$$

where $K = \sum_{i=1}^{n} \|e_i\|$. This completes the proof.

We can now use the previous proposition and the fact that the unit ball in the ∞-norm is a compact set to prove the following key fact.

Theorem 1.1.1 *Let $\|\cdot\|_a$, $\|\cdot\|_b$ be two norms on $\mathbb{F}^n$. There exist positive constants m and M such that*

$$m\|x\|_a \leq \|x\|_b \leq M\|x\|_a \tag{1.8}$$

for all $x \in \mathbb{F}^n$.

Proof:
If $x = 0$, then the result is trivial. Let B denote the unit ball in the ∞-norm,

$$B := \{x \in \mathbb{F}^n \mid \|x\|_\infty = 1\}.$$

Then B is a compact set and moreover, we know that $\|x\|_a > 0$, $\|x\|_b > 0$ for all $x \in B$. Consider the function

$$\phi(x) = \frac{\|x\|_b}{\|x\|_a}$$

defined for $x \in B$. From the above observations, ϕ is well-defined. Moreover, Proposition 1.1.2 implies that ϕ as the quotient of 2 continuous functions is continuous. Hence ϕ is lower and upper bounded on B and attains these bounds. Formally, there exist $m > 0$, $M > 0$ with

$$m \leq \phi(x) \leq M$$

for all $x \in B$. Rewriting this, we see that

$$m\|x\|_a \leq \|x\|_b \leq M\|x\|_a$$

for all $x \in B$. The result follows from the observation that for any non-zero x, $\frac{x}{\|x\|_\infty} \in B$.

Theorem 1.1.1 shows that any two norms on $\mathbb{F}^n$ are *equivalent.* This has important implications for the study of convergence in $\mathbb{F}^n$. As usual, a sequence x_n in $\mathbb{F}^n$ converges to a point $x \in \mathbb{F}^n$ with respect to a norm $\|\cdot\|$ if $\|x_n - x\| \to 0$ as $n \to \infty$. It follows readily from Theorem 1.1.1 that the choice of norm is not important in determining the convergence of a sequence. Formally, given two norms $\|\cdot\|_a$, $\|\cdot\|_b$, x_n converges to x with respect to $\|\cdot\|_a$ if and only if x_n converges to x with respect to $\|\cdot\|_b$.

Exercise: Find constants m and M such that the inequalities

$$m\|x\|_2 \leq \|x\|_1 \leq M\|x\|_2$$

hold and are tight.

Exercise: Find constants m and M such that the inequalities

$$m\|x\|_\infty \leq \|x\|_1 \leq M\|x\|_\infty$$

hold and are tight.

1.2 Dual Norms, Absolute Norms and Monotone Norms

In this section, we describe 3 important classes of norms on $\mathbb{F}^n$: *dual norms*; *absolute norms*; *monotone norms*. We begin with the definition of a dual norm.

Definition 1.2.1 *Let $\|\cdot\|$ be a norm on $\mathbb{F}^n$. Then the dual norm $\|\cdot\|^D$ is defined by*

$$\|y\|^D = \max_{\|x\|=1} Re(y^* x). \tag{1.9}$$

A few points are worth noting here. First of all, as the set $\{x \mid \|x\| = 1\}$ is compact the maximum above is indeed well-defined and attained at some x with $\|x\| = 1$. Secondly, it can be readily verified that an alternative, equivalent way of characterising $\|\cdot\|^D$ is

$$\|y\|^D = \max_{\|x\|=1} |y^* x|. \tag{1.10}$$

Exercise: Verify the previous assertion.

We shall use this second characterisation widely. As we have called $\|\cdot\|$ a dual *norm*, it is reasonable to verify that it is indeed a norm.
It is obvious that $\|y\|^D \geq 0$ for all y and moreover $\|y\|^D = 0$ if and only if $y^* x = 0$ for all x with $\|x\| = 1$. This is in turn equivalent to $y^* x = 0$ for all $x \in \mathbb{F}^n$. Choosing $x = e_i$ for $1 \leq i \leq n$ where e_i denotes the ith standard basis vector shows that this is equivalent to $y = 0$. Homogeneity is easily verified and the triangle inequality follows from

$$\begin{aligned}
\|y+z\|^D &= \max_{\|x\|=1} |(y+z)^* x| \\
&\leq \max_{\|x\|=1} (|y^* x| + |z^* x|) \\
&\leq \max_{\|x\|=1} |y^* x| + \max_{\|x\|=1} |z^* x| \\
&= \|y\|^D + \|z\|^D.
\end{aligned}$$

The following result follows readily from the definition.

Lemma 1.2.1 *For any norm $\|\cdot\|$ on $\mathbb{F}^n$,*

$$\begin{aligned}
|y^* x| &\leq \|x\| \|y\|^D \\
|y^* x| &\leq \|y\| \|x\|^D
\end{aligned}$$

for all $x, y \in \mathbb{F}^n$.

Proof: If either x or y is zero, the inequalities are trivial. Let $y \neq 0$, $x \neq 0$ be given. Then it follows from the definition that

$$\|y\|^D \geq \left| y^* \left(\frac{x}{\|x\|} \right) \right|.$$

Rearranging this we see that

$$|y^* x| \leq \|x\| \|y\|^D.$$

The second inequality is now immediate as $|x^* y| = |y^* x|$.

Let's look at some examples of dual norms.

Example 1.2.1 *First note the following simple calculation.*

$$|y^* x| = |\sum_{i=1}^{n} \bar{y}_i x_i| \leq \|x\|_\infty \|y\|_1. \tag{1.11}$$

It follows immediately that

$$\|y\|_\infty^D = \max_{\|x\|_\infty = 1} |y^* x| \leq \|y\|_1.$$

Now consider x with $x_i = \frac{y_i}{|y_i|}$ (for any $y \neq 0$) and it is easily seen that $\|x\|_\infty = 1$ and $|y^* x| = \|y\|_1$ (this equality is of course immediate if $y = 0$). It follows that $\| \cdot \|_\infty^D = \| \cdot \|_1$.

Exercise: Show that $\| \cdot \|_1^D = \| \cdot \|_\infty$.

Recall the Cauchy-Schwartz inequality for the Euclidean norm

$$|y^* x| \leq \|x\|_2 \|y\|_2.$$

It follows easily that for any y, $\|y\|_2^D \leq \|y\|_2$. If $y = 0$, then it is trivial that $\|y\|_2^D = \|y\|_2$. If $y \neq 0$, choose $x = \frac{y}{\|y\|_2}$ to see that $\|y\|_2^D = \|y\|_2$. Thus the Euclidean norm is *self-dual*. In fact, it is the only such norm on $\mathbb{F}^n$. See Theorem 5.4.16 of [1].
We saw above that $(\| \cdot \|_\infty^D)^D = \| \cdot \|_\infty$. This is true in general and we shall show this in the duality theorem below. For details of the proof of this, see Theorem 5.5.14 of [1].

Theorem 1.2.1 *Let $\| \cdot \|$ be a norm on $\mathbb{F}^n$. Then*

$$(\| \cdot \|^D)^D = \| \cdot \|.$$

Absolute and Monotone Norms

The l_p norms introduced above are examples of *absolute* norms. If we write $|x|$ for the vector $(|x_1|, \ldots, |x_n|)$ then the definition of an absolute norm is as follows.

Definition 1.2.2 *A norm $\|\cdot\|$ is* absolute *if*

$$\|x\| = \||x|\|$$

for all $x \in \mathbb{F}^n$.

Essentially, an absolute norm only depends on the absolute values of the entries of a vector. The l_p norms are also monotone in the following sense.

Definition 1.2.3 *A norm $\|\cdot\|$ is* monotone *if*

$$|x| \leq |y| \Rightarrow \|x\| \leq \|y\|$$

for all $x, y \in \mathbb{F}^n$.

As noted above, the l_p norms are both monotone and absolute. This is no coincidence as the two properties are equivalent.

Proposition 1.2.1 *Let $\|\cdot\|$ be a norm on $\mathbb{F}^n$. Then $\|\cdot\|$ is monotone if and only if it is absolute.*

Exercise: Is the norm $\|x\| = |x_1 - x_2| + |2x_2|$ a monotone norm on $\mathbb{R}^2$?

Exercise: Give an example of a monotone norm $\|\cdot\|$ and a non-singular matrix T such that $\|x\|_T = \|Tx\|$ is *not* monotone.

1.3 Matrix Norms

It is of course possible to define norms on the space of matrices $\mathbb{F}^{n\times n}$. Matrix spaces are more than just vector spaces however; it is possible to multiply matrices. The triangle inequality describes how a norm behaves with respect to vector addition. When discussing norms on matrices, it is natural to seek a family of norms that behave "well" with regard to the operation of matrix multiplication. With this in mind, we introduce the concept of a matrix norm.

Definition 1.3.1 *A mapping $\|\cdot\| : \mathbb{F}^{n\times n} \to \mathbb{R}_+$ is a* matrix norm *if it satisfies the axioms necessary to be a norm and in addition:*

$$\|AB\| \leq \|A\|\|B\|$$

for all A, B in $\mathbb{F}^{n\times n}$.

Matrix norms are particularly useful when analysing convergence of matrix powers in dynamical systems or iterative methods.

Exercise: Show that $\|I\| \geq 1$ for any matrix norm where I is the identity matrix.

Exercise: Show that $\|A^{-1}\| \geq \frac{1}{\|A\|}$ for any matrix norm and any non-singular matrix A.

Exercise: Verify that $\|A\|_1 = \sum_{i,j} |a_{ij}|$ and $\|A\|_2 = \sqrt{\sum_{i,j} |a_{ij}|^2}$ define matrix norms on $\mathbb{C}^{n\times n}$.

Exercise: Show by example that $\|A\|_\infty = \max_{i,j} |a_{ij}|$ is *not* a matrix norm.

There is a natural way to associate a matrix norm on $\mathbb{F}^{n\times n}$ with a given norm $\|\cdot\|$ on $\mathbb{F}^n$.

Definition 1.3.2 *Let $\|\cdot\|$ be a norm on $\mathbb{F}^n$. For any $A \in \mathbb{F}^{n\times n}$ define*

$$\|A\|^I = \max_{\|x\|=1} \|Ax\|. \tag{1.12}$$

Exercise: Show that $\|\cdot\|^I$ defines a matrix norm on $\mathbb{F}^{n\times n}$ and that

$$\|Ax\| \leq \|A\|^I \|x\| \tag{1.13}$$

for all $x \in \mathbb{F}^{n\times n}$.

The norm $\|\cdot\|^I$ is said to be *induced* by the vector norm $\|\cdot\|$. The property (1.13) means that $\|\cdot\|^I$ is *compatible* with $\|\cdot\|$.
We can work out direct formulae for the matrix norms induced by $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\infty$.

Example 1.3.1 *Consider the l_1 norm on $\mathbb{C}^n$. Let $A \in \mathbb{C}^{n\times n}$ be given. Then for any non-zero $x \in \mathbb{C}^n$,*

$$\begin{aligned}
\|Ax\|_1 &= \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right| \\
&\leq \sum_{i=1}^n \sum_{j=1}^n |a_{ij}||x_j| \\
&= \sum_{j=1}^n \left(\sum_{i=1}^n |a_{ij}| \right) |x_j| \\
&\leq \max_{1\leq j\leq n} \left(\sum_{i=1}^n |a_{ij}| \right) \sum_{j=1}^n |x_j|.
\end{aligned}$$

So

$$\|Ax\|_1 \leq \max_{1\leq j\leq n}\left(\sum_{i=1}^{n}|a_{ij}|\right)\|x\|_1$$

for any non-zero x*. It is easy to check that if the* k*th column of* A *gives rise to the maximal column sum* $\sum_{i=1}^n |a_{ik}| = \max_{1\leq j\leq n}(\sum_{i=1}^n |a_{ij}|)$*, then*

$$\|Ae_k\|_1 = \max_{1\leq j\leq n}\left(\sum_{i=1}^{n}|a_{ij}|\right)\|e_k\|_1.$$

Thus the matrix norm induced by the l_1 *norm is given by the maximal column sum.*

$$\|A\|_1^I = \max_{1\leq j\leq n}\left(\sum_{i=1}^{n}|a_{ij}|\right).$$

Exercise: Show that $\|A\|_\infty^I = \max_{1\leq i\leq n}\left(\sum_{j=1}^n |a_{ij}|\right)$.

Exercise: Show that $\|A\|_2^I = \max\{\sqrt{\lambda} : \lambda \in \sigma(A^*A)\}$. For this reason, $\|A\|_2^I$ is referred to as the *spectral norm. Hint:* As A^*A is Hermitian, there exists a unitary matrix U with $U^*(A^*A)U$ diagonal. Moreover, the l_2 norm is unitarily invariant.

1.3.1 Matrix Norms, Spectral Radii and Convergence

A major advantage of using matrix norms is that they are closely related to the spectral properties of a matrix and to the convergence of powers of the matrix. We first note the following simple fact concerning the spectral radius $\rho(A)$ of a matrix A.

Lemma 1.3.1 *Let* $\|\cdot\|$ *be a matrix norm on* $\mathbb{C}^{n\times n}$*. Then for any* $A \in \mathbb{C}^{n\times n}$

$$\rho(A) \leq \|A\|. \tag{1.14}$$

Proof: Let λ be an eigenvalue of A with $|\lambda| = \rho(A)$ and let $x \in \mathbb{C}^n$ satisfy $Ax = \lambda x$, $x \neq 0$. Consider the matrix $X = [xx\ldots x]$ all of whose columns are equal to x. Then $AX = \lambda X$ and hence as $\|\cdot\|$ is a matrix norm

$$\rho(A)\|X\| = \|AX\| \leq \|A\|\|X\|.$$

It follows that $\|A\| \geq \rho(A)$ as claimed.

The last result shows that the spectral radius of a matrix A is a lower bound for $\|A\|$ for any matrix norm. We next note that it is possible to find a matrix norm arbitrarily close to this lower bound.

Lemma 1.3.2 *Let* $A \in \mathbb{C}^{n\times n}$ *and* $\epsilon > 0$ *be given. There exists a matrix norm* $\|\cdot\|$ *on* $\mathbb{C}^{n\times n}$ *such that* $\rho(A) \leq \|A\| \leq \rho(A) + \epsilon$.

Proof: This follows quite easily from Schur's triangularisation theorem. There exists a unitary matrix U such that $T = U^*AU$ is upper triangular. If we apply a diagonal similarity to T, DTD^{-1} where $D = \text{diag}(k, k^2, \ldots, k^n)$, then

$$DTD^{-1} = \begin{pmatrix} \lambda_1 & t_{12}/k & t_{13}/k^2 & \cdots & t_{1n}/k^{n-1} \\ 0 & \lambda_2 & t_{23}/k & \cdots & t_{2n}/k^{n-2} \\ 0 & 0 & \lambda_3 & \cdots & t_{3n}/k^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Take $\| \cdot \|_1$ to be the maximum row sum norm (which is an induced matrix norm). Then we can select k large enough to ensure that $\|DTD^{-1}\|_\infty^I \leq \rho(A) + \epsilon$. Now simply define the norm

$$\|B\| = \|DU^*BUD^{-1}\|_\infty^I = \|(DU^*)B(DU^*)^{-1}\|_\infty^I$$

and the result follows immediately.

We can now illustrate the connection between matrix norms and the convergence of powers of a matrix $A \in \mathbb{C}^{n\times n}$. This has important applications in the analysis of iterative processes and dynamical systems. In the following result, $A^k \to 0$ as $k \to \infty$ means that each entry $a_{ij}^{(k)}$ of the matrix A^k tends to zero as $k \to \infty$.

Theorem 1.3.1 *Let $A \in \mathbb{C}^{n\times n}$ be given. Then $A^k \to 0$ as $k \to \infty$ if and only if $\rho(A) < 1$.*

Proof: First of all, suppose that $A^k \to 0$ as $k \to \infty$. Let $x \neq 0$ be an eigenvector of A with eigenvalue λ. Then $A^k x = \lambda^k x \to 0$ as $k \to \infty$. This can only happen if $|\lambda| < 1$. Hence $\rho(A) < 1$. For the converse, if $\rho(A) < 1$, it follows form Lemma 1.3.2 that there exists a matrix norm $\| \cdot \|$ on $\mathbb{C}^{n\times n}$ such that $\|A\| < 1$. Using the submultiplicative property of the matrix norm, this implies that

$$\|A^k\| \leq \|A\|^k \to 0$$

as $k \to \infty$. To complete the proof, remember that all norms on $\mathbb{C}^{n\times n}$ are equivalent so that the infinity norm (as a vector norm) of A must also converge to 0.

The next result describes an asymptotic relationship between $\rho(A)$ and $\|A\|$ for any matrix norm $\| \cdot \|$.

Theorem 1.3.2 *Let $A \in \mathbb{C}^{n\times n}$ be given and let $\| \cdot \|$ be a matrix norm on $\mathbb{C}^{n\times n}$. Then*

$$\rho(A) = lim_{k\to\infty} \|A^k\|^{1/k}.$$

Proof: Consider the sequence $\|A^k\|^{1/k}$. We know that $\rho(A^k) = \rho(A)^k$. Also $\rho(A^k) \leq \|A^k\|$. Putting these together we see that $\rho(A) \leq \|A^k\|^{1/k}$ for all k. Next note that for any $\epsilon > 0$, the matrix $\frac{A}{\rho(A)+\epsilon}$ has spectral radius strictly less than one. This implies that

$$(\frac{A}{\rho(A)+\epsilon})^k \to 0$$

as $k \to \infty$. In particular, for k sufficiently large

$$\|A^k\| \leq (\rho(A)+\epsilon)^k$$

for all k. Taking the kth root of both sides yields $\|A^k\|^{1/k} \leq (\rho(A)+\epsilon)$ for sufficiently large k. As $\epsilon > 0$ was arbitrary, the result now follows.

Exercise: Are there are any real values of a for which the matrix

$$A = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}$$

is convergent (meaning $A^k \to 0$ as $k \to \infty$).
A natural next step is to consider series and, in particular power series of matrices. The following result is important in this context.

Theorem 1.3.3 *Let $A \in \mathbb{C}^{n \times n}$ be given. Assume there is some matrix norm such that the series*

$$\sum_{i=0}^{\infty} |a_k| \|A\|^k$$

converges. Then the series

$$\sum_{i=0}^{\infty} a_k A^k$$

converges in $\mathbb{C}^{n \times n}$.

The above result implies immediately that the matrix exponential $e^A = \sum_{i=0}^{\infty} \frac{A^i}{i!}$ converges for all $A \in \mathbb{C}^{n \times n}$. It also has the following consequence for matrix invertibility.

Corollary 1.3.1 *Let $A \in \mathbb{C}^{n \times n}$ be given. Assume $\rho(A) < 1$. Then $I - A$ is invertible and moreover*

$$(I-A)^{-1} = \sum_{i=0}^{\infty} A^i. \tag{1.15}$$

Proof: As $\rho(A) < 1$, there exists a matrix norm $\|\cdot\|$ with $\|A\| < 1$. This implies that the series in (1.15) converges (Theorem 1.3.3). Consider the product

$$(I - A)\sum_{i=0}^{N} A^i = I - A^{N+1}.$$

As $\rho(A) < 1$, this converges to I as $N \to \infty$. The result now follows.

If we substitute $I - A$ for A in the above result, we see that a matrix A is invertible if $\rho(I - A) < 1$ (alternatively if there is some matrix norm such that $\|I - A\| < 1$). We can use this to derive the following simple condition for invertibility (known as the Levy-Desplanques theorem).

Theorem 1.3.4 *Let* $A \in \mathbb{C}^{n\times n}$ *be given. Suppose that*

$$|a_{ii}| > \sum_{j\neq i} |a_{ij}|$$

for $1 \leq i \leq n$. *Then* A *is invertible.*

Proof: Let $D = \mathrm{diag}(a_{11}, \ldots, a_{nn})$ and consider the matrix $\hat{A} = D^{-1}A$. It is easy to check that the maximal row-sum norm (which is an induced norm and hence a matrix norm) of $I - \hat{A}$ is less than 1. It follows from the observation before the statement of the theorem that $\hat{A}$ is invertible. Hence $A = D\hat{A}$ is also invertible.

Matrices satisfying the conditions of Theorem 1.3.4 are said to *strictly diagonally dominant.*

Exercise: Consider

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 1/2 & 1 \end{pmatrix}.$$

Use Corollary 1.3.1 to compute A^{-1}.

1.3.2 Condition Numbers and Iterative Processes

Matrix norms play an important role in the analysis of numerical algorithms for the solution of linear systems $Ax = b$ where $A \in \mathbb{C}^{n\times n}$, $b \in \mathbb{C}^n$. In many applications, the entries of A and b will not be known exactly and norms can be used to quantify the effect of errors in A or b on the accuracy of the solution.

For instance, suppose the vector b is not known exactly so that what we are given is $b + \delta b$ where δb represents the error in b. We then solve the equation and find

$$A(x + \delta x) = b + \delta b$$

where δx represents the error arising in the solution through the error in b. Clearly,

$$A\delta x = \delta b$$

from which it follows that if $\|\cdot\|$ is a norm on $\mathbb{C}^n$ and $\|\cdot\|^I$ is the corresponding induced norm then

$$\|\delta x\| \leq \|A^{-1}\|^I \|\delta b\|.$$

If we are interested in the *relative error* $\frac{\|\delta x\|}{\|x\|}$, a simple calculation shows that

$$\frac{\|\delta x\|}{\|x\|} \leq \|A\|_I \|A^{-1}\|_I \frac{\|\delta b\|}{\|b\|}.$$

The quantity $\|A\|_I \|A^{-1}\|_I$ is known as the *condition number* of A and is widely used to describe the sensitivity of a linear system to errors in the right hand side.

A second, related application is to the study of iterative methods for solving systems of the form $Ax = b$. The general setup is as follows. Choose any matrix B that is non-singular. We then take the highly imaginative step of adding and subtracting Bx to the equation to get

$$Ax + Bx - Bx = b$$

or

$$Bx = Bx - Ax + b.$$

A solution of the above equation is a fixed point of the iterative process defined by

$$x_{n+1} = (I - B^{-1}A)x_n + B^{-1}b.$$

This observation is central to many iterative methods, which solve the linear system $Ax = b$ by iterating the above system for an appropriate choice of B. Of course, it is vital that the process will converge. We can deal with this question simply by noting that the iterations can be rewritten in the form

$$\begin{pmatrix} 1 \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ B^{-1}b & (I - B^{-1}A) \end{pmatrix} \begin{pmatrix} 1 \\ x_n \end{pmatrix}.$$

It is not difficult to see that the powers of the block matrix above will converge provided the powers of $I - B^{-1}A$ converge to zero and we know that this will happen if and only if $\rho(I - B^{-1}A) < 1$.

Chapter 2

Positive Definite Matrices

We next consider the class of *positive definite* and *positive semi-definite matrices*. These matrices are important in applications ranging from optimisation to dynamics and control. We also discuss the *Singular Value Decomposition*: a fundamental result for general matrices that is crucial in many application areas including data mining and information retrieval.

2.1 Definition and Elementary Results

We begin with the definition.

Definition 2.1.1 *A Hermitian matrix* $P \in \mathbb{C}^{n \times n}$ *is* positive semi-definite *if* $x^*Px \geq 0$ *for all* $x \in \mathbb{C}^n$*. If* $x^*PX > 0$ *for all non-zero* $x \in \mathbb{C}^n$*, then* P *is* positive definite*.*

Note that as P is Hermitian by assumption, we know that x^*Px is a real number for all $x \in \mathbb{C}^n$. It is straightforward to prove the following fact.

Lemma 2.1.1 *If* P *is positive semi-definite (positive definite) then all the eigenvalues of* P *are non-negative (positive).*

Exercise: Prove Lemma 2.1.1.

In fact, using the fact that a Hermitian matrix can be unitarily diagonalised, we can show the converse also. Putting these facts together yields the next result.

Proposition 2.1.1 *A Hermitian matrix* $P \in \mathbb{C}^{n \times n}$ *is positive semi-definite (positive definite) if and only if all of its eigenvalues are nonnegative (positive).*

Exercise: Give an example of a non-Hermitian matrix P, all of whose eigenvalues are positive, for which x^TPx is not positive for all non-zero x.

It follows immediately from Proposition 2.1.1 that the trace and determinant of any positive semi-definite (positive definite) matrix must be non-negative (positive). Recall that given a set $\{i_1, \ldots, i_p\}$ of indices in $\{1, \ldots, n\}$, the corresponding principal submatrix of A is the $p \times p$ matrix consisting of the rows and columns of A indexed by $i_1, \ldots, i_p$.

Exercise: Show that any principal submatrix of a positive semi-definite (positive definite) matrix is itself positive semi-definite (positive definite).

Exercise: Show that is P is positive semi-definite and $p_{ii} = 0$ for some i, then $p_{ij} = 0$ for $1 \leq j \leq n$.

Another immediate consequence of the previous proposition is the following.

Corollary 2.1.1 *Let $P \in \mathbb{C}^{n \times n}$ be positive semi-definite (positive definite). Then P^k is positive semi-definite (positive definite) for $k = 1, 2, 3, 4, \ldots$.*

There are several equivalent characterisations of positive semi-definite and positive definite matrices. We first note the following simple characterisation of positive definiteness. In the statement of the proposition, we denote by P_i the principal submatrix of P formed from the first i rows and columns of P.

Proposition 2.1.2 *A Hermitian matrix $P \in \mathbb{C}^{n \times n}$ is positive definite if and only if $det(P_i) > 0$ for $1 \leq i \leq n$.*

Proof: If P is positive definite, then all of its principal submatrices are again positive definite and hence have positive determinant. Thus $\det(P_i) > 0$ for $1 \leq i \leq n$ as claimed. For the converse, we proceed by induction and make use of the interlacing inequalities for the eigenvalues of bordered Hermitian matrices (see Theorem 4.3.8 of [1]). The argument is as follows. The single eigenvalue of P_1 is positive as it is equal to the determinant of P_1. Suppose all the eigenvalues of P_{k-1} are positive where $2 \leq k \leq n$. The eigenvalues of P_k must interlace with those of P_{k-1}. Hence there can be at most 1 negative eigenvalue of P_k. However, if P_k has exactly 1 negative eigenvalue, this would contradict $\det(P_k) > 0$. Hence all eigenvalues of P_k are positive. By induction, we conclude that all eigenvalues of P are positive.

In analogy with nonnegative real numbers, every positive semi-definite matrix possesses a unique positive semi-definite square root.

Proposition 2.1.3 *Let P be a positive semi-definite matrix. There exists a unique positive semi-definite matrix Q satisfying $Q^2 = P$. Moreover, Q can be written as a polynomial in P and $PQ = QP$.*

We shall write $P^{1/2}$ for the square root Q appearing in Proposition 2.1.3. The previous proposition yields the following corollary.

Corollary 2.1.2 *A Hermitian matrix $P \in \mathbb{C}^{n\times n}$ is positive semi-definite if and only if there is some C in $\mathbb{C}^{n\times n}$ with $P = C^*C$. Furthermore, P is positive definite if and only if C is non-singular.*

For positive definite matrices, it is possible to say more about the matrix C appearing in the above corollary. The following result is known as the Cholesky decomposition (it follows readily by applying the QR decomposition to the matrix C in the last corollary).

Proposition 2.1.4 *A matrix P in $\mathbb{C}^{n\times n}$ is positive definite if and only if there exists some non-singular lower triangular matrix L with positive diagonal elements such that $P = LL^*$.*

The final results of this section connect positive semi-definite matrices with so-called *Gram matrices*. Given a set $\{v_1, \dots, v_k\}$ of vectors in $\mathbb{C}^n$ and an inner product $\langle \cdot, \cdot \rangle$, the associated gram matrix G if given by $g_{ij} = \langle v_j, v_i \rangle$.

Proposition 2.1.5 *Let $v_1, \dots, v_k$ be vectors in $\mathbb{C}^n$ and let $\langle \cdot, \cdot \rangle$ be an inner product. The Gram matrix G, $g_{ij} = \langle v_j, v_i \rangle$ is positive semi-definite. Moreover, G is positive definite if and only if the vector $v_1, \dots, v_k$ are linearly independent.*

Exercise: Prove Proposition 2.1.5.

Conversely, if $P \in \mathbb{C}^{n\times n}$ is positive semi-definite, then P is a Gram matrix. This follows easily from the fact that P can be written in the form $P = CC^*$ for some matrix $C \in \mathbb{C}^{n\times n}$.

2.2 Singular Value Decomposition and Applications

We now discuss two general results concerning decompositions of matrices, both of which rely on some of the elementary facts concerning positive semi-definite matrices established in the previous section.

For any matrix $A \in \mathbb{C}^{m\times n}$, the matrices A^*A and AA^* are positive semi-definite matrices in $\mathbb{C}^{n\times n}$ and $\mathbb{C}^{m\times m}$ respectively. Suppose $m \geq n$ for now. Then the eigenvalues of A^*A are all nonnegative and can be written as $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_n^2 \geq 0$. The quantities $\sigma_1, \dots, \sigma_n$ are known as the *singular values* of A. In the case where $n > m$, we simply choose the m square roots of the eigenvalues of AA^* instead. Before presenting the singular value decomposition of A, let's first note some very simple facts about the singular values that follow immediately from their definition.

Using the variational characterisation of the eigenvalues of A^*A (see Theorem 4.2.2 of [1]), it is easy to see that the largest singular value satisfies

$$\begin{aligned} \sigma_1^2 &= \max_{\|x\|=1} x^*A^*Ax \\ &= \max_{\|x\|=1} \|Ax\|_2^2. \end{aligned}$$

As the maximum is attained (at a corresponding eigenvector), it follows that

$$\sigma_1 = \max_{\|x\|=1} \|Ax\|_2. \tag{2.1}$$

Similarly, it can be shown that the smallest singular value satisfies

$$\sigma_n = \min_{\|x\|=1} \|Ax\|_2. \tag{2.2}$$

If the matrix A is square ($m = n$), the smallest singular value σ_n measures how close A is to the set of singular matrices.

Lemma 2.2.1 *Let* $A \in \mathbb{C}^{n\times n}$ *have singular values* $\sigma_1 \geq \cdots \geq \sigma_n$. *Then:*

(i) if $A + E$ *is singular for some matrix* $E \in \mathbb{C}^{n\times n}$, *then* $\|E\|_2^I \geq \sigma_n(A)$;

(ii) there exists some matrix $E \in \mathbb{C}^{n\times n}$ *with* $\|E\|_2^I = \sigma_n(A)$ *and* $A + E$ *singular.*

Exercise: Use the fact that for any $x \neq 0$, $\|Ax\|_2 \geq \sigma_n\|x\|_2$ to prove (i).

Exercise: Show that there exist vector x, y in $\mathbb{C}^{n\times n}$ with $\|x\|_2 = \|y\|_2 = 1$ and $Ax = \sigma_n y$. Use this to show (ii).

Exercise: Given

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$$

find a singular matrix that is closest to A in the spectral norm.

Any normal matrix in $\mathbb{C}^{n\times n}$ can be unitarily diagonalised; the singular value decomposition can be viewed as a generalisation of this fact to rectangular matrices.

Theorem 2.2.1 *Let* $A \in \mathbb{C}^{m\times n}$ *be given. There exist unitary matrices* $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$ *and a matrix* $\Sigma \in \mathbb{C}^{m\times n}$ *with* $\sigma_{ij} = 0$ *for* $i \neq j$ *and*

$$A = U\Sigma V^*. \tag{2.3}$$

Furthermore, the diagonal entries $\sigma_{11}, \ldots, \sigma_{qq}$ *of* Σ, *where* $q = \min\{m, n\}$ *are the singular values of* A.

Proof: We consider the case $m \geq n$ (the other case can be readily handled by considering A^* instead of A). So fix some $p \geq 0$ and write $m = n+p$. We prove the result by induction on n. When $n = 1$, A is a $p \times 1$ matrix or a column vector $v \in \mathbb{C}^p$. Put $u_1 = v/\|v\|_2$ and complete an orthonormal basis $u_1, \dots, u_p$. The required decomposition is then given by $U = (u_1, \dots, u_p)$, $V = 1$ and

$$\Sigma = \left(\begin{array}{c} \|v\|_2 \\ 0 \\ \vdots \\ 0 \end{array} \right).$$

We next assume the result is true for all matrices of size $(n-1+p) \times (n-1)$ (or $(m-1) \times (n-1)$). Let σ_1 be the maximum singular value of A. There exists a vector $x_1 \neq 0$ in $\mathbb{C}^n$ with $\|x_1\|_2 = 1$ and $A^*Ax_1 = \sigma_1 x_1$. We can construct an orthonormal basis $x_1, \dots, x_n$ and the corresponding unitary matrix $X = (x_1, \dots, x_n)$. Next set $y_1 = \frac{1}{\sigma_1} A x_1$ in $\mathbb{C}^m$ and construct an orthonormal basis $(y_1, \dots, y_m)$.
The matrices $X = (x_1, \dots, x_n)$ and $Y = (y_1, \dots, y_m)$ are unitary by construction. It can also be checked that the product Y^*AX takes the form

$$\left(\begin{array}{cc} \sigma_1 & 0 \\ 0 & \hat{A} \end{array} \right).$$

Applying the induction hypothesis to $\hat{A}$, we know that there exist unitary matrices $\hat{U} \in \mathbb{C}^{m-1 \times m-1}$ and $\hat{V} \in \mathbb{C}^{n-1 \times n-1}$ such that

$$\hat{A} = \hat{U} \hat{\Sigma} \hat{V}^*$$

where $\hat{\sigma}_{ij} = 0$ for $i \neq j$ and $\hat{\sigma}_{11} \geq \cdots \geq \hat{\sigma}_{n-1,n-1}$. Put

$$U_1 = \left(\begin{array}{cc} 1 & 0 \\ 0 & \hat{U} \end{array} \right)$$

$$V_1 = \left(\begin{array}{cc} 1 & 0 \\ 0 & \hat{V} \end{array} \right).$$

Then setting $U = YU_1$, $V = V_1 X^*$, we see that $A = U \Sigma V^*$ where $\sigma_{ij} = 0$ for $i \neq j$ and $\sigma_{22} \geq \cdots \sigma_{nn}$. It can be easily verified that the diagonal elements σ_{ii} are indeed the singular values of A from which it follows immediately that $\sigma_{11} \geq \sigma_{22}$.

- The non-zero singular values of A are the non-zero eigenvalues of AA^* or A^*A.
- The columns of U are the eigenvectors of the matrix AA^*, while the columns of V are the eigenvectors of the matrix A^*A.

- The columns of U are referred to as the left singular vectors of A, while the columns of V are known as the right singular vectors of A.

Exercise: Compute the singular value decomposition for the 4×2 matrix.

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \\ 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

The next result, known as the *polar decomposition*, follows easily from the singular value decomposition.

Theorem 2.2.2 *Let $A \in \mathbb{C}^{m\times n}$ be given. There exists a positive semi-definite matrix $P \in \mathbb{C}^{m\times m}$ and a matrix $U \in \mathbb{C}^{m\times n}$ with orthonormal rows such that*

$$A = PU$$

Proof: Assume that $m \geq n$. Use the singular value decomposition to write $A = \hat{U}\Sigma V^*$. Now write $\hat{U} = (\hat{U}_1\hat{U}_2)$, where $\hat{U}_1 \in \mathbb{C}^{m\times m}$,

$$\Sigma = \begin{pmatrix} D \\ 0 \end{pmatrix}$$

where D is square of size n. The desired decomposition is given by choosing

$$P = \begin{pmatrix} \hat{U}_1\hat{U}_2 \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{U}_1{}^* \\ \hat{U}_2{}^* \end{pmatrix}$$

and $U = \hat{U}_1 V$.
In the previous result, if A is square ($m = n$), then $A = PU$ with P positive semi-definite and U unitary.

Applications of the Singular Value Decomposition

Given a matrix $A \in \mathbb{C}^{m\times n}$ and a vector $b \in \mathbb{C}^m$, consider the linear system $Ax = b$. In general, it may not be possible to find an exact solution of this system; in this circumstance we can ask for a vector x such that $\|Ax - b\|_2$ is as small as possible - a "least squares solution". The singular value decomposition plays a key role in identifying such solutions.
Let $A \in \mathbb{C}^{m\times n}$ and $b \in \mathbb{C}^m$ be given. Let $A = U\Sigma V^*$ be a singular value decomposition of A with

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$$

where D is a diagonal matrix containing the non-zero singular values of A. Let Σ^+ be the diagonal matrix in $\mathbb{R}^{n\times m}$ (note the different size to Σ) given by

$$\Sigma^+ = \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

Next set $\hat{x} = V\Sigma^+ U^* b$. Then:

(i) $\|A\hat{x} - b\|_2 \leq \|Ay - b\|_2$ for all $y \in \mathbb{C}^n$;

(ii) if $\|Ay - b\|_2 = \|A\hat{x} - b\|_2$ for any $y \in \mathbb{C}^n$, $\|\hat{x}\|_2 \leq \|y\|_2$.

The matrix $V\Sigma^+ U^*$ appearing above is actually the *Moore-Penrose* generalised inverse of A. The vector $\hat{x}$ is a lest squares solution of the system $Ax = b$.

Exercise: Consider

$$A = \begin{pmatrix} -2 & 1 \\ 3 & 1 \\ 1 & -1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Compute a least squares solution of the system $Ax = b$.

A second important application of the singular value decomposition is to the construction of low-rank approximations to a matrix $A \in \mathbb{C}^{n\times n}$. As above let $A = U\Sigma V$ be a singular value decomposition of A and suppose that the singular values of A are $\sigma_1 \geq \sigma_2 \geq \cdots \geq c_n$. The "best" rank k approximation to A (in the sense of being closest in the l_2 norm) is then given by $U\Sigma_k V$ where

$$\Sigma_k = \text{diag}(\sigma_1, \ldots, \sigma_k, 0, \ldots, 0).$$

2.3 The Loewner Order and Some Simple Inequalities

Denote by $PSD(n)$ the set of all positive semi-definite matrices in $\mathbb{C}^{n\times n}$. It is not difficult to verify that $PSD(n)$ satisfies the following conditions:

(i) for any $\alpha \geq 0, \beta \geq 0$ and any $P, Q \in PSD(n)$, $\alpha P_1 + \beta P_2 \in PSD(n)$;

(ii) if $P \in PSD(n)$ and $-P \in PSD(n)$, then $P = 0$;

(iii) the interior of $PSD(n)$ is non-empty.

The above conditions mean that $PSD(n)$ is a convex (i), pointed (ii), solid (iii) cone in the space of Hermitian matrices in $\mathbb{C}^{n\times n}$. For any such cone, it is possible to define a partial order on Hermitian matrices in $\mathbb{C}^{n\times n}$. In our context we define $A \geq B$ if $A - B \in PSD(n)$. The notation $A > B$ means $A - B \in \text{int}(PSD(n))$. This order is known as the Loewner order. It is called a partial order because there exist Hermitian matrices A, B for which neither $A \geq B$ nor $B \geq A$ is true.

Exercise: Give a simple example of two Hermitian matrices A and B for which neither of $A \geq B$ nor $B \geq A$ is true.

We shall give a simple description of when two positive definite matrices satisfy $A \geq B$ below. We first need a useful fact concerning congruence.

Lemma 2.3.1 *Let P be positive definite and Q be Hermitian. Then there exists a nonsingular matrix C such that $C^*PC = I$ and C^*QC is diagonal.*

Proof: We can find a unitary matrix U_1 such that $D_1 = U_1^* P U_1$ is diagonal. As P is positive definite all the diagonal elements of D_1 are positive. Hence if we define $C_1 = U_1 D_1^{-1/2}$, it is readily seen that C_1 is non-singular and $C_1^* P C_1 = I$. The matrix $C_1^* Q C_1$ is Hermitian and hence there exists some unitary matrix U_2 such that $U_2^* C_1^* Q C_1 U_2$ is diagonal. Set $C = C_1 U_2$. As U_2 is unitary, we have that $C^*PC = I$, C^*QC is diagonal as claimed.

Using the above fact, we can now give a simple spectral condition for two positive definite matrices to be related by the Loewner order.

Theorem 2.3.1 *Let P, Q be positive definite matrices. Then $P \geq Q$ if and only if $\rho(QP^{-1}) \leq 1$.*

Proof: Using Lemma 2.3.1, we can choose a non-singular matrix C such that $C^*PC = I$ and $D = C^*QC$ is diagonal. The condition $P \geq Q$ is equivalent to $d_{ii} \leq 1$ for all i. As $C^{-1}P^{-1}C^{*-1} = I$, this is in turn equivalent to the diagonal entries of

$$DI = C^*QCC^{-1}P^{-1}C^{*-1} = C^*QP^{-1}C^{*-1}$$

all being less than or equal to 1. As $D = DI$ is diagonal and C^* is non-singular, this is equivalent to $\rho(QP^{-1}) \leq 1$.

Exercise: Consider

$$P = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}, Q = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}.$$

Is $P \geq Q$? Is $Q \geq P$?

We can now use the above result to derive some simple inequalities for positive definite matrices.

Proposition 2.3.1 *Let P, Q be positive definite matrices and assume $P \geq Q$. Then* *(i)* $det(P) \geq det(Q)$;

(ii) $trace(P) \geq trace(Q)$.

Exercise: Prove Proposition 2.3.1.

We next describe a condition for a block matrix to be positive definite in terms of the so-called *Schur complement.*

Theorem 2.3.2 *Consider the Hermitian matrix*

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{12}^* & P_{22} \end{pmatrix}$$

where P_{11} *and* P_{22} *are Hermitian. The matrix* P *is positive definite if and only if* P_{11} *is positive definite and* $P_{22} - P_{12}^* P_{11}^{-1} P_{12}$ *is positive definite.*

We finish off this section with two more well-known inequalities for positive definite matrices.

Theorem 2.3.3 *Let* $P \in PSD(n)$ *be given. Then*

$$det(P) \leq \prod_{i=1}^{n} p_{ii}.$$

The previous result is known as *Hadamard's inequality.* The next result concerns the situation when the matrix $H(A) = A + A^*$ is positive definite. We shall encounter a generalised version of this condition in relation to stability in the next chapter. The result presented here is known as the Ostrowski-Taussky inequality.

Theorem 2.3.4 *Let* $A \in \mathbb{C}^{n \times n}$ *be given. Suppose that* $H(A) = A + A^*$ *is positive definite. Then*

$$det(H(A)) \leq |det(A)|.$$

Chapter 3

Matrix Stability

3.1 Lyapunov's Theorem and Extensions

The results we discuss here are motivated by the study of the linear time-invariant (LTI) system

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0. \tag{3.1}$$

Here $A \in \mathbb{R}^{n\times n}$ (or $\mathbb{C}^{n\times n}$), $x_0 \in \mathbb{R}^n$. We denote by $x(\cdot, x_0)$ the solution of (3.1) satisfying $x(0) = x_0$. LTI systems of this form play a fundamental role in analysing more complicated dynamical and control systems. In particular, the question of their stability is of paramount importance. A few points are worth recalling.

For any $x_0 \in \mathbb{R}^n$, there exists a solution $x(t, x_0)$ of (3.1) defined for all $t \geq 0$. In fact, it is relatively straightforward to verify that $x(t, x_0) = e^{At}x_0$. Next, note that the origin is an *equilibrium* of (3.1); this means that if $x_0 = 0$, then $x(t, x_0) = 0$ for all $t \geq 0$. If A is non-singular, then the origin is the only equilibrium of (3.1). In all of the situations we consider, the origin is the only equilibrium of interest (for more general systems, this is of course not always the case) and for this reason, we will speak of the stability properties of the system (3.1) rather than (more correctly) of the equilibrium in question.

The system is globally asymptotically stable if it satisfies the following two conditions.

(i) For any $\epsilon > 0$, there exists some $\delta > 0$ such that $\|x_0\| \leq \delta$ implies $\|x(t, x_0)\| \leq \epsilon$ for all $t \geq 0$.

(ii) For any $x_0 \in \mathbb{R}^n$, $x(t, x_0) \to 0$ as $t \to \infty$.

A fundamental result in the theory of LTI systems asserts that (3.1) is globally asymptotically stable if and only if all of the eigenvalues of A have negative real part. Such matrices are said to be *Hurwitz* or *stable.*

The next classical theorem, originally due to A.M. Lyapunov, relates the stability of a matrix A to the existence of a positive definite solution of an associated linear matrix equation.

Theorem 3.1.1 *Let $A \in \mathbb{C}^{n\times n}$ be given. Then A is Hurwitz if and only if for every positive definite $Q \in \mathbb{C}^{n\times n}$, there exists some $P \in \mathbb{C}^{n\times n}$ with*

$$A^*P + PA = -Q < 0. \tag{3.2}$$

Proof: First suppose there exist $Q > 0$, $P > 0$ satisfying (3.2). If some eigenvalue λ of A has $\mathrm{Re}(\lambda) \geq 0$, then for the associated eigenvector $x \neq 0$, we would have

$$\begin{aligned} -x^*Qx &= x^*A^*Px + x^*PAx \\ &= (\bar{\lambda} + \lambda)x^*Px \\ &= 2\mathrm{Re}(\lambda)x^*Px \geq 0 \end{aligned}$$

which contradicts the assumption that Q is positive definite. Thus all eigenvalues of A must have negative real part.
For the converse, suppose that all eigenvalues of A have a negative real part. Consider the linear operator L_A defined on $\mathbb{C}^{n\times n}$ by

$$L_A(X) = A^*X + XA.$$

The eigenvalues of L_A are the n^2 numbers $\lambda_i + \bar{\lambda_j}$ for $1 \leq i, j, \leq n$. It follows that if A is Hurwitz, then L_A is invertible and hence for any choice of $Y \in \mathbb{C}^{n\times n}$, there is a unique X in $\mathbb{C}^{n\times n}$ satisfying

$$L_A(X) = A^*X + XA = Y.$$

Next note that for Y Hermitian, the solution of $A^*X + XA = Y$ must also be Hermitian. This is because $A^*X^* + X^*A = Y^* = Y$ and thus $X = X^*$ as the solution is unique. Finally, note that the matrix

$$P = \int_0^\infty e^{A^*t}Qe^{At}dt$$

satisfies $A^*P + PA = -Q$.

Exercise: Consider

$$A = \begin{pmatrix} -2 & 3 \\ -1 & 1 \end{pmatrix}.$$

Explain why A is Hurwitz. Verify that there exists a positive definite matrix P with $A^*P + PA = -I$.

Return to the real case and the LTI system (3.1) for a moment. If A is Hurwitz and P is a positive definite matrix satisfying the equation of the theorem, the quadratic form

$$V(x) = x^TPx$$

defines a *Lyapunov function* for (3.1). This means that V is positive definite and its time-derivative along the trajectories of (3.1) is negative everywhere except at the origin.
In applications, it is often of interest to know when two or more matrices possess a *common* solution to the Lyapunov inequality. For two matrices, A_1, A_2, this amounts to asking when there exists a positive definite P such that

$$A_i^* P + PA_i < 0, \quad i \in \{1, 2\}. \tag{3.3}$$

This question arises in the stability of so-called *switching systems* that play an important role in many control applications.
In the context of *robust stability* the following question arises. Given two matrices A_1, A_2, when is the convex hull

$$\mathrm{Co}(A_1, A_2) := \{\alpha A_1 + (1-\alpha)A_2 : 0 \leq \alpha \leq 1\}$$

Hurwitz stable (meaning that every matrix in $\mathrm{Co}(A_1, A_2)$ is Hurwitz stable)? The next result provides some elementary relations between these two questions.

Proposition 3.1.1 *Let A_1, A_2 be Hurwitz matrices in $\mathbb{C}^{n\times n}$. Consider the following statements:*

(i) there exists a positive definite P satisfying (3.3);

(ii) $Co(A_1, A_2)$ is Hurwitz stable;

(iii) $\sigma(A_1^{-1}A_2) \cap (-\infty, 0] = \emptyset$.

Then (i) implies (ii) implies (iii).

Exercise: Prove Proposition 3.1.1.

Exercise: Show that if there exists a P satisfying (3.3), then $\mathrm{Co}(A_1^{-1}, A_2)$ is Hurwitz stable.

Exercise: Construct an example of two matrices A_1, A_2 for which $\mathrm{Co}(A_1, A_2)$ is Hurwitz stable but for which there is no positive definite P satisfying (3.3).

The following fact follows readily from Theorem 3.1.1.

Corollary 3.1.1 *Let $A \in \mathbb{C}^{n\times n}$ be given. The following are equivalent:*

(i) for every $x \in \mathbb{C}^n$, there exists a positive definite P_x such that $Re(x^* P_x Ax) < 0$;

(ii) there exists a positive definite P such that $Re(x^* PAx) < 0$ for all $x \in \mathbb{C}^n$.

Theorem 3.1.1 has been extended to the case where the right hand side of the Lyapunov equation is positive definite but the solution P is not necessarily so. First we need to introduce the concept of *inertia* for a matrix in $\mathbb{C}^{n\times n}$. The inertia of A is the triple $(i_-(A), i_0(A), i_+(A))$ where $i_-(A)$ is the number of eigenvalues of A with negative real part, $i_0(A)$ the number of eigenvalues with zero real part, and $i_+(A)$ the number of eigenvalues of A with positive real part.
The next result is known as the *general inertia theorem.*

Theorem 3.1.2 *Let* $A \in \mathbb{C}^{n\times n}$ *be given. There exists a positive definite matrix* Q *and a Hermitian* P *such that*

$$A^*P + PA = Q$$

if and only if $i_0(A) = 0$. *When this happens,* Q *and* P *have the same inertia.*

3.2 Positivity and Stability

We next consider a class of matrices that play a central role in describing systems that are positive in the sense that their state variables may only take nonnegative values. Formally, we say that the LTI system (3.1) is positive if $x_0 \in \mathbb{R}^n_+$ implies $x(t, x_0) \in \mathbb{R}^n_+$ for all $t \geq 0$.

- It is easy to see that (3.1) will be positive if and only if the matrix e^{At} is nonnegative for all $t \geq 0$.
- This in turn can be shown to be equivalent to the requirement that the off-diagonal elements a_{ij}, $i \neq j$ of A are nonnegative.
- Another way of stating this is that $A = N - kI$ where N is a nonnegative matrix and $k \geq 0$. Matrices of this type are known as *Metzler matrices.*

Exercise: Show that if A is Metzler, then e^{At} is nonnegative for all $t \geq 0$.
Hint: If A and B commute, then $e^{(A+B)} = e^A e^B$.

In this section, we will use the notation $v \gg 0$, where $v \in \mathbb{R}^n$, to denote that all components of the vector v are positive $v_i > 0$ for all i.
There are a (very) wide variety of equivalent conditions for a Metzler matrix to be Hurwitz. We recall a small number of these here.
Remark:
With applications to dynamics in mind, we have considered matrices that are Metzler and Hurwitz. In the mathematics literature, it is common to consider matrices with non-positive off-diagonal entries (known as Z-matrices) and all eigenvalues with positive real part (positively stable matrices). A positively stable Z-matrix is known as an *M-matrix.*

Theorem 3.2.1 *Let $A \in \mathbb{R}^{n\times n}$ be Metzler. The following are equivalent:*

(i) *A is Hurwitz;*

(ii) *there exists some $v \gg 0$ such that $Av \ll 0$;*

(iii) *there exists some $w \gg 0$ such that $A^T w \ll 0$;*

(iv) *there exists a diagonal positive definite matrix D with $A^T D + DA$ negative definite;*

(v) *DA is Hurwitz for any positive definite diagonal matrix D;*

(vi) *for any Metzler matrix B with $b_{ij} \le a_{ij}$ for all i, j, B is Hurwitz;*

(vii) *A^{-1} is entrywise non-positive;*

(viii) *all principal minors of $-A$ are positive;*

(ix) *for every $x \neq 0$ in $\mathbb{R}^n$, there exists some index i with $x_i(Ax)_i < 0$.*

For Metzler Hurwitz matrices, the situation with regard to stability of convex hulls is considerably simpler than for the general case.

Proposition 3.2.1 *Let A_1, A_2 be Metzler Hurwitz matrices in $\mathbb{R}^{n\times n}$. The convex hull $Co(A_1, A_2)$ is Hurwitz stable if and only if $\sigma(A_1^{-1}A_2) \cap (-\infty, 0] = \emptyset$.*

Metzler matrices and their properties are used in analysing a variety of system models arising in ecology and theoretical biology. In particular, nonlinear generalisations of positive systems, known as monotone systems, play a key role in many such applications. We note the following simple result for ecological systems described by Lotka-Volterra equations here. We consider a state vector $x \in \mathbb{R}^n_+$ in which x_i represents the population of species i, r_i represents the intrinsic growth (or decay) rate of species i and a_{ij} is used to describe inter-species interaction. For more details on this type of model and its relation to the results discussed here, see [4].
We say that a Lotka-Volterra system is *uniformly bounded* if there is some $K > 0$ such that $\text{limsup}_{t\to\infty} x_i(t, x_0) \le K$ for $1 \le i \le n$ and all $x_0 \in \mathbb{R}^n_+$.

Theorem 3.2.2 *Consider the Lotka-Volterra system*

$$\dot{x}_i = x_i(r_i + \sum_{j=1}^{n} a_{ij}x_j). \tag{3.4}$$

Suppose that A is Metzler (the system is then said to be mutualistic) and there exists an equilibrium e of (3.4) in $int(\mathbb{R}^n_+)$. Then the following are equivalent.

(i) (3.4) are uniformly bounded.

(ii) A is Hurwitz.

(iii) For any $x_0 \in int(\mathbb{R}^n_+)$, $x(t, x_0) \to e$ as $t \to \infty$.

Bibliography

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

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

[3] J. Stoer and R. Bulirsch Introduction to Numerical Analysis *Springer-Verlag*, 1980.

[4] J. Hofbauer and K. Sigmund Evolutionary Games and Population Dynamics *Cambridge University Press*, 1998.

[5] A. Berman, M. Neumann and R. Stern Nonnegative Matrices in Dynamic Systems *Wiley*, 1989.

[6] A. Berman and R. Plemmons Nonnegative Matrices in the Mathematical Sciences *SIAM Classics in Applied Mathematics*, 1994.