NORMS

Norms are a measure of the "size" of objects. We introduce norms to measure the size of vectors and matrices.

Let $\mathcal{V}$ be a vector space. A *norm* on $\mathcal{V}$ is a non-negative real-valued function $N(x)$ defined on $\mathcal{V}$ and satisfying the following properties:
N1. $N(x) = 0$ if and only if $x = 0$.
N2. $N(\alpha x) = |\alpha|\, N(x)$, for all $x \in \mathcal{V}$ and all scalars α.
N3. $N(x + y) \leq N(x) + N(y)$, for all $x, y \in \mathcal{V}$.

Using N, we define the distance between vectors x and y to be $N(x - y)$. From N3, we have

$$N(x - z) \leq N(x - y) + N(y - z), \quad x, y, z \in \mathcal{V}$$

We can also prove the *reverse triangle inequality*:

$$|N(x) - N(y)| \leq N(x - y), \quad x, y \in \mathcal{V}$$

More commonly, we write $\|x\| = N(x)$.

EXAMPLES

Let p be any real number satisfying $1 \leq p < \infty$. For $x \in \mathbb{R}^n$ or $\mathbb{C}^n$, introduce the <u>p-norm</u>:

$$\|x\|_p = \left[\sum_{j=1}^{n} |x_i|^p \right]^{\frac{1}{p}}$$

This is called the p-norm. For $p = 1$, showing N1-N3 is relatively straightforward. For $p = 2$, we have the Euclidean norm introduced earlier and associated with an inner product.

Define the *maximum norm* by

$$\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$$

It can be shown that

$$\lim_{p \to \infty} \left[\sum_{j=1}^{n} |x_i|^p \right]^{\frac{1}{p}} = \max_{i=1,\ldots,n} |x_i|, \qquad x \in \mathbb{C}^n$$

which is the justification for the notation $\|x\|_\infty$.

Proof that

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

is a norm. Our vector space is $\mathcal{V} = \mathbb{R}^n$ or $\mathbb{C}^n$.

N1. Let $x \in \mathcal{V}$. Clearly $\|x\|_1 \geq 0$. Also $\|x\|_1 = 0$ if and only if all components $x_i = 0$, or equivalently, $x = 0$.

N2. Let $x \in \mathcal{V}$, and let α be a scalar. Then

$$\alpha x = [\alpha x_1, ..., \alpha x_n]^T$$

and

$$\|\alpha x\|_1 = \sum_{i=1}^{n} |\alpha x_i| = |\alpha| \sum_{i=1}^{n} |x_i| = |\alpha| \, \|x\|_1$$

N3. Let $x, y \in \mathcal{V}$. Then

$$x + y = [x_1 + y_1, ..., x_n + y_n]^T$$

$$\begin{aligned}\|x+y\|_1 &= \sum_{i=1}^{n} |x_i + y_i| \\ &\le \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} |y_i| = \|x\|_1 + \|y\|_1\end{aligned}$$

The quantity

$$\|x\|_\infty = \max_{1\le i\le n} |x_i|\,, \qquad x \in \mathbb{R}^n \text{ or } \mathbb{C}^n$$

can be shown to be a norm by similar arguments. This particular norm was introduced in Chapter 1; and it is the discrete analogue to the function norm

$$\|f\|_\infty = \max_{a\le t\le b} |f(t)|\,, \qquad f \in C[a,b]$$

The quantity

$$\|x\|_2 = \left[\sum_{j=1}^{n} |x_i|^2\right]^{\frac{1}{2}}, \quad x \in \mathbb{R}^n \text{ or } \mathbb{C}^n$$

is the Euclidean norm studied previously in Section 7.2, where we showed it satisfied the triangle inequality. It is the discrete analogue of the function norm

$$\|f\|_2 = \left[\int_a^b |f(t)|^2\,dt\right]^{\frac{1}{2}}, \qquad f \in C[a,b]$$

CONVERGENCE IN $\mathbb{R}^n$ AND $\mathbb{C}^n$

We say a sequence of vectors $\left\{x^{(1)}, ..., x^{(m)}, ...\right\}$ in $\mathcal{V} = \mathbb{R}^n$ (or $\mathbb{C}^n$) converges to a vector $x \in \mathcal{V}$ if and only if

$$\lim_{m \to \infty} \left\| x - x^{(m)} \right\| = 0$$

What norm am I using? It turns out to not matter. As a lead-in to this, note that

$$\max_{1 \le i \le n} |x_i| \le \sum_{i=1}^{n} |x_i| \le n \max_{1 \le i \le n} |x_i|$$

$$\|x\|_\infty \le \|x\|_1 \le n \|x\|_\infty, \quad x \in \mathcal{V}$$

Therefore

$$\left\| x - x^{(m)} \right\|_\infty \le \left\| x - x^{(m)} \right\|_1 \le n \left\| x - x^{(m)} \right\|_\infty, \quad x \in \mathcal{V}$$

As $m \to \infty$, convergence with the norm $\|\cdot\|_1$ is equivalent to convergence with the norm $\|\cdot\|_\infty$.

EQUIVALENCE OF NORMS

Let $N(x)$ and $M(x)$ denote two norms on the vector space $\mathcal{V} = \mathbb{R}^n$ or $\mathbb{C}^n$. Then there are positive constants c_1 and c_2 for which

$$c_1 M(x) \leq N(x) \leq c_2 M(x), \quad x \in \mathcal{V}$$

Note that this says convergence is equivalent in all norms, as

$$c_1 M(x - x^{(m)}) \leq N(x - x^{(m)}) \leq c_2 M(x - x^{(m)})$$

PROOF: How do we shown this result. A proof is given in the text if the following is too brief. We begin by letting $M(x) = \|x\|_\infty$, as that will be sufficient to show the general case. Then we need to show

$$c_1 \|x\|_\infty \leq N(x) \leq c_2 \|x\|_\infty, \quad x \in \mathcal{V}$$

Divide by $\|x\|_\infty$ to obtain

$$c_1 \leq N\left(\frac{x}{\|x\|_\infty}\right) \leq c_2, \quad x \in \mathcal{V}$$

or equivalently, show

$$c_1 \leq N(z) \leq c_2, \quad \text{for all } z \in \mathcal{V},\ \|z\|_\infty = 1$$

Boundedness of $N(x)$. For general x, write

$$x = [x_1, ..., x_n]^T = \sum_{j=1}^{n} x_j e^{(j)}$$

with $e^{(j)}$ the standard *unit vectors*. Then

$$N(x) = N\left(\sum_{j=1}^{n} x_j e^{(j)}\right) \leq \sum_{j=1}^{n} \left|x_j\right| N\left(e^{(j)}\right) \leq c_2 \|x\|_\infty$$

with

$$c_2 = \sum_{j=1}^{n} N\left(e^{(j)}\right)$$

Next we need to show

$$c_1 \leq N(z), \quad \text{for all } z \in \mathcal{V},\ \|z\|_\infty = 1$$

for some $c_1 > 0$.

The above boundedness result also shows that N is a continuous function. In particular,

$$|N(x) - N(y)| \leq N(x - y) \leq c_2 \|x - y\|_\infty$$

Thus $N(x)$ is a continuous function of x.

Introduce the set

$$S = \{z \in \mathcal{V} \mid \|z\|_\infty = 1\}$$

This is called the *unit sphere* relative to the norm $\|\cdot\|_\infty$. It is a closed and bounded set in $\mathcal{V}$, relative to $\|\cdot\|_\infty$; and $N(z)$ is a continuous real-valued function as z varies over this set. Therefore, the maximum and minimum of $N(z)$ over S occurs at points z^* and z_* in S :

$$c_2 \equiv \max_{z \in S} N(z) = N(z^*)$$

$$c_1 \equiv \min_{z \in S} N(z) = N(z_*)$$

Since points in S are nonzero, we have $c_1 \neq 0$. This completes the proof.

MATRIX NORMS

Notice that sets of matrices can be considered as a vector space. For example, consider all 2×2 matrices with real entries:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$$

Then this set of matrices is a vector space $\mathcal{V}$:

$$\alpha \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} = \begin{bmatrix} \alpha a_{1,1} & \alpha a_{1,2} \\ \alpha a_{2,1} & \alpha a_{2,2} \end{bmatrix}$$

$$\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} = \begin{bmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} \end{bmatrix}$$

The dimension of this vector space is 4.

The set of matrices of order $m \times n$ form a vector space of dimension mn. Thus every matrix norm must first be a vector norm.

Let $\mathcal{M}$ denote the set of square matrices of order $n \times n$. Then a matrix norm on $\mathcal{M}$ is a non-negative real-value function $\|A\|$ for which

N1. $\|A\| = 0$ if and only if $A = 0$.
N2. $\|\alpha A\| = |\alpha| \, \|A\|$, for all scalars α and all $A \in \mathcal{M}$.
N3. $\|A + B\| \leq \|A\| + \|B\|$, for all $A, B \in \mathcal{M}$.

In addition, matrices have a sense of multiplication, involving both vectors and other matrices. We would like to have this recognized in defining a sense of size of a matrix. In particular, we require a matrix norm to also satisfy the following.

N4. $\|AB\| \le \|A\| \, \|B\|, \quad A, B \in \mathcal{M}$

For $\mathcal{V} = \mathbb{R}^n$ or $\mathbb{C}^n$, we will have some vector norm $\|x\|_v$ we are using. Then we require a "compatibility" between the matrix norm and the vector norm:
N5.

$$\|Ax\|_v \le \|A\| \, \|x\|_v, \quad x \in \mathcal{V}, \quad A \in \mathcal{M}$$

In defining a matrix norm, we usually begin with the vector norm and then discover how to define the matrix norm so as to maintain compatibility with the vector norm.

Example: Consider the vector norm

$$\|x\|_\infty = \max_{1 \le i \le n} |x_i|\,, \quad x \in \mathbb{R}^n$$

Then

$$Ax = \left[\sum_{j=1}^{n} a_{1,j} x_j, ..., \sum_{j=1}^{n} a_{n,j} x_j\right]^T$$

$$\|Ax\|_\infty = \max_{1 \le i \le n} \left|\sum_{j=1}^{n} a_{i,j} x_j\right|$$

We want to have this be bounded by $\|A\| \, \|x\|_\infty$ for some choice of $\|A\|$. Then

$$\begin{aligned} \|Ax\|_\infty &\le \max_{1 \le i \le n} \sum_{j=1}^{n} \left|a_{i,j}\right| \left|x_j\right| \\ &\le \left[\max_{1 \le i \le n} \sum_{j=1}^{n} \left|a_{i,j}\right|\right] \|x\|_\infty \end{aligned}$$

Define

$$\|A\| = \max_{1 \le i \le n} \sum_{j=1}^{n} \left| a_{i,j} \right|$$

Is this a norm? We know the property N5 is satisfied since we constructed it that way. We can also show this definition satisfies properties N1-N4.

There is a more general way to proceed in getting the norm, and the above is an example of this process. When given a vector norm $\|\cdot\|_v$, we want to have the matrix norm to satisfy

$$\|Ax\|_v \le \|A\| \, \|x\|_v \, , \quad x \in \mathcal{V}, \quad A \in \mathcal{M}$$

To accomplish this, we define

$$\|A\| = \sup_{x \ne 0} \frac{\|Ax\|_v}{\|x\|_v}$$

This is called the *operator matrix norm* associated with the given vector norm $\|\cdot\|_v$.

SOME OPERATOR NORMS

Introduce

$$\sigma(A) = \{\lambda \mid \lambda \text{ an eigenvalue of } A\}$$

$$r_\sigma(A) = \max_{\lambda \in \sigma(A)} |\lambda|$$

The set $\sigma(A)$ is called the *spectrum* of A; and the number $r_\sigma(A)$ is called the *spectral radius* of A.

Vector norm	Operator matrix norm
$\|x\|_\infty$	$\|A\|_\infty \equiv \max_{1\le i\le n} \sum_{j=1}^{n} \left|a_{i,j}\right|$
$\|x\|_1$	$\|A\|_1 \equiv \max_{1\le j\le n} \sum_{i=1}^{n} \left|a_{i,j}\right|$
$\|x\|_2$	$\|A\|_2 \equiv [r_\sigma(AA^*)]^{\frac{1}{2}} = [r_\sigma(A^*A)]^{\frac{1}{2}}$

$\|A\|_\infty$ is called the *row norm*; and $\|A\|_1$ is called the *column norm*. The quantity $\|A\|_2$ is often hard to compute, but it is bounded by the *Frobenius norm*:

$$F(A) \equiv \left[\sum_{i=1}^{n} \sum_{j=1}^{n} \left| a_{i,j} \right|^2 \right]^{\frac{1}{2}}$$

Properties: (a) For the identity matrix $A = I$, we have

$$\|I\| = \sup_{x \neq 0} \frac{\|Ix\|_v}{\|x\|_v} = \sup_{x \neq 0} \frac{\|x\|_v}{\|x\|_v} = 1$$

(b) Let λ be an eigenvalue of A, with $x \neq 0$ an associated eigenvector. Then

$$|\lambda| \, \|x\|_v = \|\lambda x\|_v = \|Ax\|_v \leq \|A\| \, \|x\|_v$$

$$|\lambda| \leq \|A\|$$

This proves

$$r_\sigma(A) \leq \|A\|$$

for any operator matrix norm.

(c) Let A be a real symmetric matrix or a complex Hermitian matrix. Then

$$A^* A = AA^* = A^2$$

For its eigenvalues, let $u_1, ..., u_n$ be an orthogonal set of eigenvectors of A, corresponding to the eigenvalues $\lambda_1, ..., \lambda_n$. Then

$$A^2 u_i = A(Au_i) = A(\lambda_i u_i) = \lambda_i^2 u_i$$

Thus when A is Hermitian,

$$\|A\|_2 = [r_\sigma(AA^*)]^{\frac{1}{2}} = \left[r_\sigma(A^2)\right]^{\frac{1}{2}} = r_\sigma(A)$$

Thus $\|A\|_2$ may be easier to compute when A is symmetric or Hermitian.

There are other properties that are explored in the assigned problems.

THEOREM

Let $\epsilon > 0$ be a given small number and let A be a square matrix. Then there is a special vector norm $\|\cdot\|_v$ and an associated matrix operator matrix norm $\|\cdot\|_\epsilon$ for which

$$\|A\|_\epsilon \leq r_\sigma(A) + \epsilon$$

This not an easy theorem to prove, and I give a reference to another text for a proof. It shows that $r_\sigma(A)$ is close to *some* operator matrix norm for A, since then

$$r_\sigma(A) \leq \|A\|_\epsilon \leq r_\sigma(A) + \epsilon$$

This proves the result that $r_\sigma(A) < 1$ if and only if $\|A\| < 1$ for some operator matrix norm.