3.5 VECTOR AND MATRIX NORMS

Consider the space,

$$R^2 = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}; x_1, x_2 \in R \right\},$$

our 'usual' two-dimensional plane. If x = $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ is any vector in this space we define its 'usual' 'length' or 'norm' as

$$\|x\| = \sqrt{x_1^2 + x_2^2}$$

We observe that

(i) $\|x\| \geq 0$ for every vector x in R^2,
$\|x\| = 0$ if and only if x is θ;

(ii) $\|\alpha x\| = |\alpha| \|x\|$ for any scalar α; for any vector x.

(iii) $\|x + y\| \leq \|x\| + \|y\|$ for any two vectors x and y.

(The inequality (iii) is usually referred to as the triangle inequality).

We now generalize this idea to define the concept of a norm on C^n or R^n.

The norm can be thought of intuitively as a rule which associates with each vector x in V, a real number $\|x\|$, and more precisely as a function from the collection of vectors to the real numbers, satisfying the following properties:

(i) $\|x\| \geq 0$ for every x $\in$ V and
$\|x\| = 0$ if and only if x = θ;

(ii) $\|\alpha x\| = |\alpha| \|x\|$ for every scalar α and every vector x in V,

(iii) $\|x + y\| \leq \|x\| + \|y\|$ for every x, y in V.

Examples of Vector Norms on C^n and R^n

Let $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ be any vector x in C^n (or R^n)

We can define various norms as follows:

(1) $\|x\|_2 = \left[|x_1|^2 + |x_2|^2 + ... + |x_n|^2\right]^{\frac{1}{2}} = \left[\sum_{i=1}^{n} |x_i|^2\right]^{\frac{1}{2}}$

(2) $\|x\|_1 = |x_1| + |x_2| + ... + |x_n| = \sum_{i=1}^{n} |x_i|$

In general for $1 \leq p < \infty$ we can define,

(3) $\|x\|_p = \left\{\sum_{i=1}^{n} |x_i|^p\right\}^{\frac{1}{p}}$

If we set p = 2 in (3) we get $\|x\|_2$ as in (1) and if we set p = 1 in (3) we get $\|x\|_1$ as in (2).

(4) $\|x\|_\infty = \max.\{|x_1|, |x_2|, ..., |x_n|\}$

All these can be verified to satisfy the above mentioned properties (i), (ii) and (iii) required of a norm. Thus these give several types of norms on C^n and R^n.

Example:

(1) Let $x = \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}$ in R^3

Then

$$\|x\|_1 = 1 + 1 + 2 = 4$$

$$\|x\|_2 = (1+1+4)^{1/2} = \sqrt{6}$$

$$\|x\|_\infty = \max.\{1,1,2\} = 2$$

$$\|x\|_4 = \left(1^4 + 1^4 + 2^4\right)^{\frac{1}{4}} = 18^{\frac{1}{4}}$$

(2) Let $x = \begin{pmatrix} 1 \\ i \\ -2i \end{pmatrix}$ in C^3

Then

$$\|x\|_1 = 1 + 2 + 1 = 4$$

$$\|x\|_2 = \left(1+4+1\right)^{\frac{1}{2}} = \sqrt{6}$$

$$\|x\|_\infty = \max.\{1,2,1\} = 2$$

$$\|x\|_3 = \left(1^3 + 2^3 + 1^3\right)^{\frac{1}{3}} = 10^{\frac{1}{3}}$$

Consider a sequence $\left\{x^{(k)}\right\}_{k=1}^{\infty}$ of vectors in C^n (or R^n)

$$x^{(k)} = \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix}$$

Suppose $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in C^n \text{ (or } R^n\text{)}$

DEFINITION:

We say that the sequence $\{x^{(k)}\}$ of vectors CONVERGES to the vector x as k tends to infinity if the sequence of numbers, $\{x_1^{(k)}\}$ converges to the number x_1; $\{x_2^{(k)}\}$ converges to x_2, …., and $\{x_n^{(k)}\}$ converges to x_n i.e.

As k $\to \infty$; $\quad x_i^{(k)} \to x_i \quad$ for every i=1, 2, …., n.

Example:

Let $x^{(k)} = \begin{pmatrix} i/k \\ 1-2/k \\ \dfrac{1}{k^2+1} \end{pmatrix}$ be a sequence of vectors in R^3.

Let $x = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. Here $x_1^{(k)} = \dfrac{1}{k} \to 0 = x_1$

$$x_2^{(k)} = 1 - \frac{2}{k} \to 1 = x_2$$

$$x_3^{(k)} = \frac{1}{k^2+1} \to 0 = x_3$$

$\therefore x^{(k)}{}_i \to x_i$ for i =1,2,3.

$\therefore x^{(k)} \to x$

If $\{x^{(k)}\}$ is a sequence of vectors such that in some norm $\| \ \|$, the sequence of real numbers, $\|x^{(k)} - x\|$ converges to the real number 0 then we say that the sequence of vectors converges to x with respect to this norm. We then write,

$$x^{(k)} \xrightarrow{\| \ \|} x$$

For example consider the sequence,

$$x^{(k)} = \begin{pmatrix} \frac{1}{k} \\ 1-\frac{2}{k} \\ \frac{1}{k^2+1} \end{pmatrix}$$ in R^3 as before and,

$$x = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

We have

$$x^{(k)} - x = \begin{pmatrix} \frac{1}{k} \\ -\frac{2}{k} \\ \frac{1}{k^2+1} \end{pmatrix}$$

Now

$$\left\| x^{(k)} - x \right\|_1 = \frac{1}{k} + \frac{2}{k} + \frac{1}{k^2+1} \to 0$$

$$\therefore\ x^{(k)} \xrightarrow{\|\ \|_1} x$$

Similarly

$$\left\| x^{(k)} - x \right\|_\infty = \max.\left\{ \frac{1}{k}, \frac{2}{k}, \frac{1}{k^2+1} \right\} = \frac{2}{k} \to 0$$

$$\therefore\ x^{(k)} \xrightarrow{\|\ \|_\infty} x$$

$$\left\| x^{(k)} - x \right\|_2 = \left\{ \frac{1}{k^2} + \frac{2}{k^2} + \frac{1}{(k^2+1)^2} \right\}^{\frac{1}{2}} \to 0$$

$$\therefore\ x^{(k)} \xrightarrow{\|\ \|_2} x$$

Also,

$$\left\|x^{(k)}-x\right\|_p=\left\{\frac{1}{k^p}+\left(\frac{2}{k}\right)^p+\frac{1}{\left(k^2+1\right)^p}\right\}^{\frac{1}{p}}\to 0$$

$$\therefore\ x^{(k)}\xrightarrow{\|\ \|_p}x\qquad\forall p;1\le p\le\infty$$

It can be shown that

" IF A SEQUENCE $\left\{x^{(k)}\right\}$OF VECTORS IN C^n (or R^n) CONVERGES TO A VECTOR x IN C^n (or R^n) WITH RESPECT TO ONE VECTOR NORM THEN THE SEQUENCE CONVERGES TO x WITH RESPECT TO ALL VECTOR NORMS AND ALSO THE SEQUENCE CONVERGES TO x ACCORDING TO DEFINITION ON PAGE 113 . CONVERSELY IF A SEQUENCE CONVERGES TO x AS PER DEFINITION ON PAGE 113 THEN IT CONVERGES WITH RESPECT TO ALL VECTOR NORMS".

Thus when we want to check the convergence of a particular sequence of vectors we can choose that norm which is convenient to that sequence.

MATRIX NORMS

Let M be the set of all nxn matrices (real or complex). A matrix norm is a function from the collection of matrices to the real numbers, whose value at any matrix A is denoted by $\|A\|$ having the following properties:

(i) $\|A\|\geq 0$ for all matrices A

$\|A\|=0$ if and only if A = O_n,

(ii) $\|\alpha A\|=|\alpha|\,\|A\|$ for every scalar α and every matrix A,

(iii) $\|A+B\|\le\|A\|+\|B\|$ for all matrices A and B,

(iv) $\|AB\|\le\|A\|\,\|B\|$for all matrices A and B.

Before we give examples of matrix norms we shall see a method of getting a matrix norm starting with a vector norm.

Suppose $\|.\|$ is a vector norm. Then, consider $\dfrac{\|Ax\|}{\|x\|}$ (where A is an nxn matrix); for x $\neq \theta_n$. This given us an idea to by what proportion the matrix A has distorted the length of x. Suppose we take the maximum distortion as we vary x over all vectors. We get

$$\max_{x \neq \theta_n} \frac{\|Ax\|}{\|x\|}$$

a real number. We define

$$\|A\| = \max_{x \neq \theta_n} \frac{\|Ax\|}{\|x\|}$$

We can show this is a matrix norm and this matrix norm is called the matrix norm subordinate to the vector norm $\| \;\|$. We can also show that

$$\|A\| = \max_{x \neq \theta_n} \frac{\|Ax\|}{\|x\|} = \max_{\|x\| = 1} \|Ax\|$$

For example,

$$\|A\|_1 = \max_{\|x\|_1 = 1} \|Ax\|_1$$

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

$$\|\mathrm{A}\|_\infty = \max_{\|x\|_\infty = 1} \|Ax\|_\infty$$

$$\|\mathrm{A}\|_\mathrm{p} = \max_{\|x\|_p = 1} \|Ax\|_p$$

How hard or easy is it to compute these matrix norms? We shall give some idea of computing $\|A\|_1$, $\|A\|_\infty$ and $\|A\|_2$ for a matrix A.

Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots & a_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ a_{n1} & a_{n2} & \ldots\ldots & a_{nn} \end{pmatrix}$$

The sum of the absolute values of the entries in the i^{th} column is called the absolute column sum and is denoted by C_i. We have

$$C_1 = |a_{11}| + |a_{21}| + |a_{31}| + \ldots\ldots + |a_{n1}| = \sum_{i=1}^{n} |a_{i1}|$$

$$C_2 = |a_{12}| + |a_{22}| + |a_{32}| + \ldots\ldots + |a_{n2}| = \sum_{i=1}^{n} |a_{i2}|$$

….. ….. ….. ….. ….. ….. …..

$$C_j = \sum_{i=1}^{n} |a_{ij}| \qquad ; \qquad 1 \le j \le n$$

Thus we have n absolute column sums, C_1 , C_2, ….., C_n.

Let

$$C = \max.\{C_1, C_2, \ldots\ldots, C_n\}$$

This is called the maximum absolute column sum. We can show that,

$$\|A\|_1 = C = \max.\{C_1, \ldots\ldots, C_n\}$$

$$= \max_{1 \le j \le n} \left[\sum_{i=1}^{n} |a_{ij}| \right]$$

For example, if

$$A = \begin{pmatrix} 1 & 2 & -3 \\ -1 & 0 & 1 \\ -3 & 2 & -4 \end{pmatrix},$$

then

$$\left.\begin{aligned} C_1 &= 1 + 1 + 3 = 5; \\ C_2 &= 2 + 0 + 2 = 4; \\ C_3 &= 3 + 1 + 4 = 8 \end{aligned}\right\} \quad \text{and } C = \max.\{5, 4, 8\} = 8$$

$\therefore \|A\|_1 = 8$

Similarly we denote by R_i the sum of the absolute values of the entries in the i^{th} row

$$R_1 = |a_{11}| + |a_{12}| + ... + |a_{1n}| = \sum_{j=1}^{n} a_{1j}$$

$$R_2 = |a_{21}| + |a_{22}| + ... + |a_{2n}| = \sum_{j=1}^{n} a_{2j}$$

....

$$R_i = |a_{i1}| + |a_{i2}| + ... + |a_{in}| = \sum_{j=1}^{n} a_{ij}$$

and define R, the maximum absolute row sum as,

R = max $\{R_1, \ldots., R_n\}$

It can be show that,

$$\|A\|_\infty = R = \max\{R_1, ..., R_n\}$$

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

For example, for the matrix

$$A = \begin{pmatrix} 1 & 2 & -3 \\ -1 & 0 & 1 \\ -3 & 2 & -4 \end{pmatrix},$$ we have

$\left.\begin{array}{l} R_1 = 1 + 2 + 3 = 6; \\ R_2 = 1 + 0 + 1 = 2; \\ R_3 = 3 + 2 + 4 = 9 \end{array}\right\}$ and R = max {6, 2, 9}= 9

$$\therefore \|A\|_\infty = 9$$

The computation of $\|A\|_1$ and $\|A\|_\infty$ for a matrix are thus fairly easy. However, the computation of $\|A\|_2$ is not very easy; but somewhat easier in the case of the Hermitian matrix.

Let A be any nxn matrix; and

$C(\lambda) = (\lambda - \lambda_1)^{a_1} \cdots (\lambda - \lambda_k)^{a_k}$, be its characteristic polynomial, where λ_1, λ_2, ….., λ_k are the distinct characteristic values of A.

Let

$$P = \max .\{|\lambda_1|, |\lambda_2|, \ldots , |\lambda_k|\}$$

This is called the spectral radius of A and is also denoted by $\|A\|_{sp}$

It can be show that for a Hermitian matrix A,

$\|A\|_2$ = P = $\|A\|_{sp}$

For example, for the matrix,

$$A = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{pmatrix}$$

which is Hermitian we found on page 106, the distinct eigenvalues as $\lambda_1 = 2$; $\lambda_2 = 8$

$$\therefore \|A\|_{xp} = P = \max .\{2,8\}$$
$$= 8$$

$$\therefore \|A\|_2 = \|A\|_{sp} = 8$$

If A is any general nxn matrix (not Hermitian) then let B = A^* A. Then $B^* = A^*$ A = B, and hence B is Hermitian and its eigenvalues are real and in fact its eigenvalues are nonnegative. Let the eigenvalues (distinct) of B be μ_1, μ_2, ….., μ_r. Then let

$\mu = \max \{\mu_1, \mu_2, \ldots, \mu_r\}$

We can show that

$$\|A\|_2 = \mu = \max\{\mu_1, \ldots\ldots, \mu_n\}$$

If follows from the matrix norm definition subordinate to a vector norm, that

$$\|A\| = \max_{x \neq \theta_n} \frac{\|Ax\|}{\|x\|}$$

$\therefore$ For any x in C^n or R^n , we have, if $x \neq \theta_n$

$$\frac{\|Ax\|}{\|x\|} \leq \max_{x \neq \theta_n} \frac{\|Ax\|}{\|x\|} = \|A\|$$

and therefore

$$\|Ax\| \leq \|A\|\|x\| \quad \text{for all } x \neq \theta_n$$

But this is obvious for $x = \theta_n$

Thus if $\|A\|$ is a matrix norm subordinate to the vector norm $\|x\|$ then

$$\|Ax\| \leq \|A\|\|x\|$$

for every vector x in C^n (or R^n).