A DISCUSSION OF VECTOR AND MATRIX NORMS IN FINITE DIMENSIONAL SPACES

HORACE E. WILLIAMS
Vanderbilt School of Engineering, Nashville, Tennessee

In numerical work, one customarily encounters problems dealing with vector and matrix computations. Familiar examples of this are: the solution of linear equations, including matrix inversions; the finding of eigenvalues and eigenvectors corresponding to a matrix; the solution of differential and integral equations; the solution of linear programming models; and the solution of correlation-regression models. In error analyses of such processes, the choice of a good basis of comparison is a crucial feature. It is desirable if not essential to have a single measure which gives an evaluation of the size of a vector or matrix. Such measures are referred to as vector or matrix norms. In the majority of cases one may assume a positive definite inner product space of finite dimension. The discussion will be limited to this case, although several of the concepts discussed need not necessarily be restricted in this way.

VECTOR NORMS

The norm of a vector X with elements $(x_1, x_2, \ldots, x_n)$ from a space of dimension n will be denoted by $||X||$. Three desirable properties of a useful vector norm are as follows:

$$||X|| > 0 \quad \text{for } X \neq 0, \quad ||X|| = 0 \quad \text{for } X = 0. \tag{1}$$
$$||kX|| = |k|\,||X||, \text{ where k is a complex constant.} \tag{2}$$
$$||X + Y|| \leqslant ||X|| + ||Y||. \tag{3}$$

Two other useful properties follow from the above three properties.

$$||X - Y|| \geqslant ||X|| - ||Y||. \tag{4}$$
$$||X - Y|| \geqslant ||Y|| - ||X||. \tag{5}$$

To derive (4), consider the relations,

$||X|| = ||X - Y + Y|| \leqslant ||X - Y|| + ||Y||$, and thus
$||X|| - ||Y|| \leqslant ||X - Y||$.

In an analogous fashion we may obtain (5) by considering

$||Y|| = ||-Y|| = ||X - Y - X|| \leqslant ||X - Y|| + ||-X||$, thus giving

$||Y|| - ||-X|| \leqslant ||X - Y||$,
$||Y|| - ||X|| \leqslant ||X - Y||$.

We define three vector norms which are commonly used in numerical analysis.[1]

$$||X||_1 = \sum_{i=1}^{n} |x_i|. \tag{6}$$

$||X||_2 = (\sum_{i=1}^{n} x_i^2)^{1/2} = (X \cdot X)^{1/2}$, where $X \cdot X$ is the inner product or dot product of X with itself. (7)

$$||X||_\infty = \max_i |x_i|. \tag{8}$$

The norms so defined may be shown to possess properties (1)-(3).

MATRIX NORMS

The norm of a matrix A with elements

$$\begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{Vmatrix}$$

of dimension n^2 is denoted by $||A||$. Desirable properties generally imposed on these are

$$||A|| > 0 \quad \text{if } A \neq 0, \; ||A|| = 0 \quad \text{for } A = 0. \tag{9}$$
$$||kA|| = |k|\,||A|| \quad \text{for k a complex constant.} \tag{10}$$
$$||A + B|| \leqslant ||A|| + ||B||, \text{ where A and B are both matrices.} \tag{11}$$
$$||AB|| \leqslant ||A||\,||B||. \tag{12}$$

We now present three commonly encountered definitions for matrix norms. These norms each possess properties (9)-(12) as may be shown in a straightforward manner.

$$||A||_1 = \max_j \sum_{i=1}^{n} |a_{ij}|. \tag{6'}$$

$||A||_2 = \sqrt{t_1}$, where t_1 is the maximum eigenvalue of AA^T. (7')

$$||A||_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|. \tag{8'}$$

SUBORDINATION AND COMPATIBILITY

A matrix norm $||A||$ is said to be subordinate to a vector norm $||X||$ if and only if it satisfies the relationship

$||A|| = \max \dfrac{||AX||}{||X||}$ where X ranges over all vectors in the given domain for which $X \neq 0$. (13)

We shall subsequently demonstrate that $||A||_1$ is a matrix norm subordinate to $||X||_1$, $||A||_2$ is subordinate to $||X||_2$, and $||A||_\infty$ is subordinate to $||X||_\infty$.

By definition (13) it follows that for a matrix norm which is subordinate to a given vector norm

$||A|| \geqslant \dfrac{||AX||}{||X||}$ for any $X \neq 0$. Thus

$$||A||\,||X|| \geqslant ||AX||. \tag{14}$$

A given matrix norm and a given vector norm which satisfy this property are often called compatible norms. If X = 0 the case is trivial of course.

[1] These definitions may be generalized by

$$||X||_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p} \text{ for } p \geqslant 1.$$

An interesting exercise for adept students is to show that $||X_\infty|| = \lim_{p \to \infty} ||X||_p$. This may be accomplished by writing a polynomial expansion for $||X||_p$ and showing that the largest term (in magnitude) dominates the resulting series.

An equivalent definition for (13) is to say that matrix norm $||A||$ is subordinate to vector norm $||X||$ if

$$||A|| = \max ||AX|| \text{ for all X such that } ||X|| = 1. \quad (13')$$

To demonstrate subordination then we wish to choose a vector X' such that $||AX'|| = ||A||$ and further that $||AX'|| \geqslant ||AX||$ for any other choice of X, where $||X'|| = ||X|| = 1$.

1. Subordination of $||A||_1$.

Choose X' such that $x_i' = \begin{cases} 0 \text{ if } i \neq j_o \\ 1 \text{ if } i = j_o \end{cases}$ where j_o is the subscript corresponding to

$$\max_j \sum_{i=1}^{n} |a_{ij}| = \sum_{i=1}^{n} |a_{ij_o}|.$$

$$||AX'||_1 = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij} x_j' \right| = \sum_{i=1}^{n} |a_{ij_o}| = ||A||_1.$$

For any other choice of X ($||X||_1 = 1$) we have

$$||AX||_1 = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \leq \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}||x_j|$$

$$\leq \sum_{j=1}^{n} \left(|x_j| \sum_{i=1}^{n} |a_{ij}|\right) \leq \sum_{j=1}^{n} |x_j| \sum_{i=1}^{n} |a_{ij_o}|$$

$$\leq ||X||_1 \, ||A||_1 = ||A||_1 = ||AX'||_1.$$

Thus $||AX||_1 \leqslant ||AX'||_1$.

2. Subordination of $||A||_2$.

Before proceeding to a selection of X' for this case it is expedient to recall some basic results from linear algebra. For a given n x n matrix B and a n x l vector Y, the scalars t_k that satisfy the equations $BY = tY$ are called the eigenvalues of B. The vectors Y_k that satisfy $BY_k = t_kY_k$ are called eigenvectors. If B is a positive definite or positive semi-definite symmetric matrix, the eigenvalues are all real and non-negative; one may choose a basis of eigenvectors for the space; and any two eigenvectors corresponding to distinct eigenvalues are orthogonal. The matrix B^TB, where B^T denotes the transpose of B, is positive semi-definite symmetric, and the inner product $(BY{:}BY) = (B^TBY \cdot Y)$.

For a matrix A let t_1 represent the maximum eigenvalue of A^TA and let U_1 represent an eigenvector corresponding to t_1. We choose $X' = \frac{U_1}{||U_1||_2}$; thus $||X||_2 = 1$. It then follows that

$$||AX'||_2^2 = (AX' \cdot AX') = (A^TAX' \cdot X') = (t_1 X' \cdot X')$$

$$= t_1 (X' \cdot X') = t_1 ||X'||_2^2 = t_1.$$

$$||AX'||_2 = \sqrt{t_1} = ||A||_2.$$

Let us choose an orthonormal basis of eigenvectors and let us designate these by V_i, where V_i corresponds to the eigenvalue t_i. Thus

$$||V_i||_2 = 1.$$

Any vector X ($||X||_2 = 1$) may be written as a linear combination of the V_i's since these form a basis for the space. Thus $X = \sum_{i=1}^{n} c_i V_i$ with $\sum_{i=1}^{n} c_i^2 = 1$. We may now state the results,

$$||AX||_2^2 = (AX \cdot AX) = (A^TAX \cdot X) = (A^TA \sum_{i=1}^{n} c_i V_i \cdot \sum_{i=1}^{n} c_i V_i)$$

$$= (\sum_{i=1}^{n} c_i t_i V_i \cdot \sum_{i=1}^{n} c_i V_i).$$

Using the orthogonality property of the eigenvectors we obtain,

$$= (\sum_{i=1}^{n} t_i c_i^2) \leq t_1 \sum_{i=1}^{n} c_i^2 = t_1$$

$$||AX||_2 \leq \sqrt{t_1} = ||A||_2 = ||AX'||_2.$$

3. Subordination of $||A||_\infty$.

Select X' such that $x_j' = \begin{cases} 0 \text{ if } a_{i_oj} = 0, \\ 1 \text{ if } a_{i_oj} > 0 \\ -1 \text{ if } a_{i_oj} < 0 \end{cases}$ where i_o is the index corresponding to $\max_i \sum_{j=1}^{n} |a_{ij}| = \sum_{i=1}^{n} |a_{i_oj}|$.

$$||AX'||_\infty = \max_i \left| \sum_{j=1}^{n} a_{ij} x_j' \right| \leq \max_i \sum_{j=1}^{n} |a_{ij}||x_j'|.$$

Then since $|x_j'| \leqslant 1$ we have

$$\leq \max_i \sum_{j=1}^{n} |a_{ij}| = ||A||_\infty.$$ We may also show

$$||AX'||_\infty = \max_i \left| \sum_{j=1}^{n} a_{ij} x_j' \right| \geq \left| \sum_{j=1}^{n} a_{i_oj} x_j' \right|$$

$$\geq \sum_{j=1}^{n} |a_{i_oj}| = ||A||_\infty.$$ We have then

$||A||_\infty \leq ||AX'||_\infty \leq ||A||_\infty$ which implies $||AX'||_\infty = ||A||_\infty$. For any X ($||X||_\infty = 1$), letting $|x_{i_o}| = \max_i |x_i|$, we have

$$||AX|| = \max_i \left| \sum_{j=1}^{n} a_{ij} x_j \right| \leq |x_{i_o}| \max_i \sum_{j=1}^{n} |a_{ij}|$$

$$\leq ||X||_\infty ||A||_\infty = ||A||_\infty = ||AX'||_\infty.$$

BOUNDS FOR VECTOR AND MATRIX NORMS

An interesting and useful set of properties for the vector norms defined by (6), (7), and (8) is that bounds for each may be established in terms of either of the other two. The following chain of inequalities succinctly expresses these relationships. The proofs of them provide a worthwhile set of exercises.

$$||X||_\infty \leq ||X||_2 \leq \sqrt{n}\, ||X||_\infty \leq ||X||_1 \leq \sqrt{n}\, ||X||_2$$
$$\leq n\, ||X||_\infty. \quad (15)$$

Two useful results concerning matrix norms may be easily established. First we consider the equation $AX = sX$ where s is an eigenvalue of A and X is a corresponding eigenvector. For compatible vector and matrix norms we have

$$\|A\| \, \|X\| \geqslant \|AX\| = \|sX\| = |s| \, \|X\|. \quad \text{We obtain}$$
$$\|A\| \geqslant |s|. \tag{16}$$

Statement (16) provides an upper bound for the eigenvalues of a matrix A.

Secondly, consider

$\|A\|_2^2 = t_1$. From (16) above it follows that

$$t_1 \leq \|A^T A\|_\infty \leq \|A^T\|_\infty \|A\|_\infty$$
$$\leq \max_i \Big| \sum_{j=1}^{n} a_{ji} \Big| \max_i \Big| \sum_{j=1}^{n} a_{ij} \Big|$$
$$\leq \max_j \Big| \sum_{i=1}^{n} a_{ij} \Big| \max_i \Big| \sum_{j=1}^{n} a_{ij} \Big| = \|A\|_1 \|A\|_\infty,$$

giving the result

$$\|A\|_2^2 \leqslant \|A\|_1 \, \|A\|_\infty.$$

The Euclidean Norm

Another matrix norm that is often encountered is the so-called Euclidean norm denoted by $\|A\|_E$ and defined by

$$\|A\|_E = \Big(\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \Big)^{1/2} \tag{18}$$

The norm so defined may be shown to be compatible with $\|X\|_2$; however, it may also be shown that $\|A\|_E$ cannot be subordinate to any vector norm. As a consequence this norm is of small value in purely mathematical analyses but is often helpful in practical work and error analyses. For further elucidation of this point, see Wilkinson (1963).

Examples

Let X be a vector with elements $X = (3,-2,1.5,-.8)$. Then one computes

$$\|X\|_1 = |3| + |-2| + |1.5| + |-.8| = 7.3$$
$$\|X\|_2 = ((3)^2 + (-2)^2 + 1.5)^2 + (-.8)^2)^{1/2} = \sqrt{15.89} = 3.99$$
$$\|X\|_\infty = 3.$$

Let A be the matrix A =

$$\begin{Vmatrix} 10 & 0 & 5 & 0 \\ 3 & 8 & 6 & 0 \\ 4 & -6 & 8 & 0 \\ 0 & 0 & 0 & 7 \end{Vmatrix} \quad \begin{matrix} \sum_j |a_{ij}| \\ 15 \\ 17 \\ 18 \\ 7 \end{matrix}$$
$$\sum_i |a_{ij}| = 17 \quad 14 \quad 19 \quad 7$$

The resulting norms are

$\|A\|_1 = \max(17,14,19,7) = 19$

$\|A\|_2 = \sqrt{225} = 15$. The eigenvalues of $A^T A$ are 25, 49, 100, 225.

$\|A\|_\infty = \max(15,17,18,7) = 18$

$$\|A\|_E = [(10)^2 + (5)^2 + (3)^2 + (8)^2 + (6)^2 + (4)^2 + (-6)^2 + (8)^2 + (7)^2]^{1/2} = \sqrt{399} = 19.99$$

An unpopular feature of $\|A\|_2$ is illustrated in working an example of the type above. Computing $\|A\|_2$ is a considerably more complicated numerical task than computing any of the other three.

Literature Cited

Wilkinson, J. H. 1963. Rounding Errors in Algebraic Processes. Prentice-Hall, Englewood Cliffs, N.J. p. 81.