Basic Linear Algebra Notation and Operations

Leigh J. Little SUNY Brockport September 4, 2020

Contents

1 Introduction

Linear algebra is an important fundamental topic in mathematics and it is important for scientists to be familiar with the basic notation and operations associated with linear algebra. The purpose of this document is to familarize the reader with these operations, with a primary emphasis on the mechanics of linear algebra calculations.

2 Linear Algebra Quantities

2.1 Scalars

Scalars are constants. π and e are 2 well-known examples of scalars. Scalars can be added, subtracted, multiplied and (most of the time) divided. Other operations, such as the trigonometric, logarithmic and exponential functions can also be applied to scalars. Scalars can be either real or complex. In this document, scalars will be indicated by greek letters.

2.2 Vectors

A vector is a set of scalars arranged into a row or column. Examples of row vectors are:

$$w = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}; \qquad v = \begin{pmatrix} -3 & 4 & 0 \end{pmatrix}.$$

while examples of column vectors are

$$q = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}; \qquad r = \begin{pmatrix} 3 \\ 9 \\ -2 \end{pmatrix}.$$

By convention, vectors are indicated by lower case letters.

Each scalar in the vector is called an *element* or a *component* of the vector and every vector has an associated length (or size). The length of a vector is equal to the number of components (the vectors above all have a length of 3).

In the case of the vectors w and v, we can emphasize that they are row vectors by indicating their *dimensions* as 1 x 3 (*i.e.,* 1 row by 3 columns). Similarly, we can emphasize q and r are column vectors by indicating their dimensions as 3 x 1 (*i.e.,* 3 rows by 1 column). Unless otherwise specified, vectors are assumed to be *column* vectors.

Subscripts can be used to reference a specific element of a vector. For example, r_2 is the value of the element in the second position of r, so $r_2 = 9$. Similarly, $v_3 = 0$.

In general, a column vector x of length n can be written as

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

or in index notation,

$$x = x_i, \quad \text{for} \quad i = 1, \ldots, n.$$

The subscripts allow each element to be referenced individually if needed.

2.2.1 The Transpose Operation

A basic operation on vectors that is frequently performed is called a *transpose.* The transpose changes a row vector to a column vector (and vice-versa). The symbol for transposition is a superscript T. Thus, if $y = x^T$, then

$$y = x^T = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}.$$

If the vector elements are also complex, the *Hermitian* transpose is used instead. The Hermitian transpose, indicated by a superscript H, takes the complex conjugate of the elements in addition to performing a transpose, thus, if $y = x^H$, then

$$y = x^H = \overline{(x^T)} = (\overline{x})^T = \begin{pmatrix} \overline{x}_1 & \overline{x}_2 & \cdots & \overline{x}_n \end{pmatrix}.$$

where $\overline{x}_i$ represents the complex conjugate of x_i.

2.3 Matrices

A matrix is a collection of scalars arranged in a rectangular grid. Examples of matrices are:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}; \qquad B = \begin{pmatrix} 3 & -2 \\ 0 & 5 \\ -1 & 9 \end{pmatrix}; \qquad C = \begin{pmatrix} 5 & 2 & -1 \\ 0 & 8 & -7 \end{pmatrix}.$$

By convention, matrix variables are indicated by capital letters. Just as with vectors, matrices have an associated size (or dimension). In the above example, A is called a 3×3 matrix because it has 3 rows and 3 columns. Similarly, B is 3×2 and C is 2×3. If a matrix has the same number of rows and columns it is called a *square* matrix, otherwise it is *rectangular.*

2.3.1 Indexing of Matrices

The elements of a matrix are indexed by a double subscript according to their row and column location. Thus, in the above example, $A_{2,3} = 6$ and $A_{3,1} = 7$. As with vectors, the first element of a matrix is indexed as $A_{1,1}$ (not $A_{0,0}$ as it would be in some programming languages). In general, if A is an $n \times m$ matrix, then

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \cdots & \cdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix}$$

or in index notation,

$$A = a_{i,j}, \quad \text{for } i = 1, \ldots, n, \; j = 1, \ldots, m,$$

where $a_{i,j}$ is the element in row i and column j.

Matrices can also be viewed as collections of row or column vectors. For example, the matrix B above can be viewed as the column vectors

$$\begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}; \quad \begin{pmatrix} -2 \\ 5 \\ 9 \end{pmatrix}.$$

placed side-by-side or the row vectors

$$\begin{pmatrix} 3 & -2 \end{pmatrix}$$
$$\begin{pmatrix} 0 & 5 \end{pmatrix}$$
$$\begin{pmatrix} -1 & 9 \end{pmatrix}$$

stacked on top of each other.

To refer to an entire row or column of a matrix, the $*$ notation can be used:

$$\begin{aligned} b_{i,*} &= \text{row } i \text{ of the matrix } B, \\ b_{*,j} &= \text{column } j \text{ of the matrix } B. \end{aligned}$$

For the above example,

$$b_{*,2} = \begin{pmatrix} -2 \\ 5 \\ 9 \end{pmatrix}$$

and

$$b_{1,*} = \begin{pmatrix} 3 & -2 \end{pmatrix}.$$

Using the $*$ notation, the matrix B could also be written as

$$B = \begin{pmatrix} b_{1,*} \\ b_{2,*} \\ b_{3,*} \end{pmatrix}.$$

This is known as partitioning a matrix by rows. Similarly,

$$B = \begin{pmatrix} b_{*,1} & b_{*,2} \end{pmatrix}.$$

is a column partitioning of B.

2.3.2 Matrix Transpose Operation

Matrices can also be transposed. For some matrix A, the transpose operation A^T interchanges the rows and columns of a matrix (or equivalently, exchanges the row and column indicies of the matrix). In index notation, the transpose is written

$$a^T_{i,j} = a_{j,i}.$$

The transposes of the matrices A, B and C above are

$$A^T = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}; \qquad B^T = \begin{pmatrix} 3 & 0 & -1 \\ -2 & 5 & 9 \end{pmatrix}; \qquad C^T = \begin{pmatrix} 5 & 0 \\ 2 & 8 \\ -1 & -7 \end{pmatrix}.$$

As in the case of vectors, if the elements of a matrix are complex, the Hermitian transpose is used instead.

2.3.3 Symmetric Matrices

A square matrix is *symmetric* if

$$A = A^T.$$

For example, the matrix

$$E = \begin{pmatrix} 3 & -2 & 6 \\ -2 & 1 & 4 \\ 6 & 4 & -5 \end{pmatrix}$$

is symmetric.

2.3.4 Matrix Diagonal

The *diagonal* of a matrix consists of those elements whose row and column entries are the same.

$$\text{diag}(A) = \{a_{1,1}, a_{2,2}, \ldots, a_{k,k}\}$$

where $k = \min(n, m)$. For the matrices A, B and C defined above

$$\begin{aligned} \text{diag}(A) &= \{1, 5, 9\} \\ \text{diag}(B) &= \{3, 5\} \\ \text{diag}(C) &= \{5, 8\}. \end{aligned}$$

Depending on the context, the diagonal of a matrix might instead refer to the matrix consisting just of the diagonal elements; for example,

$$\text{diag}(A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 9 \end{pmatrix}; \qquad \text{diag}(B) = \begin{pmatrix} 3 & 0 \\ 0 & 5 \\ 0 & 0 \end{pmatrix}; \qquad \text{diag}(C) = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 8 & 0 \end{pmatrix}.$$

3 Linear Algebra Operations

3.1 Vector-Vector Operations

This section describes the mathematical operations that can be performed on vectors.

3.1.1 Vector Addition

Two vectors, x and y can be added or subtracted only of they have the same length and orientation (*i.e.*, are both row or column vectors). The operation is performed *componentwise*, thus $z = x + y$ means

$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

or more compactly,

$$z_i = x_i + y_i \quad \text{for} \quad i = 1, \ldots, n.$$

A similar definition holds for subtraction.

3.1.2 Multiplying a Vector by a Scalar

A vector x can be multipied by a scalar α. This is also done componentwise

$$\alpha x = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}.$$

or using index notation,

$$\alpha x = \alpha x_i, \quad \text{for} \quad i = 1, \ldots, n.$$

An expression like

$$z = \alpha x + \beta y$$

would be computed as

$$z_i = \alpha x_i + \beta y_i \quad \text{for} \quad i = 1, \ldots, n.$$

3.1.3 Examples

Suppose q, r, w and v are defined as in Section 2.2. Some examples of the above operations are given below.

1) $q + r = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} + \begin{pmatrix} 3 \\ 9 \\ -2 \end{pmatrix} = \begin{pmatrix} 4+3 \\ 5+9 \\ 6-2 \end{pmatrix} = \begin{pmatrix} 7 \\ 14 \\ 4 \end{pmatrix}$

2) $q + w =$ undefined

3) $w^T = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}^T = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

4) $q + w^T = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}^T = \begin{pmatrix} 4+1 \\ 5+2 \\ 6+3 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix}$

5) $3q - 4r = 3 \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} - 4 \begin{pmatrix} 3 \\ 9 \\ -2 \end{pmatrix} = \begin{pmatrix} 12 \\ 15 \\ 18 \end{pmatrix} + \begin{pmatrix} -12 \\ -36 \\ 8 \end{pmatrix} = \begin{pmatrix} 12-12 \\ 15-36 \\ 18-8 \end{pmatrix} = \begin{pmatrix} 0 \\ -21 \\ 8 \end{pmatrix}$

3.1.4 Vector Multiplication

Multiplication of vectors is much more complicated. In order for two vectors to be multiplied, the "inner dimensions" must match. For two column vectors x and y of length n, the vector product

$$z = x \cdot y$$

is not defined. This is because

$$\begin{array}{ccc} x \cdot y & = & (n \times 1) \cdot (n \times 1). \\ & & \quad\ \ \uparrow \quad \uparrow \end{array}$$

The inner dimensions (1 and n) do not match, so this quantity has no meaning.

3.1.5 The Dot Product

There are two basic vector product operations that can be performed. The first is the *dot* product. The dot product of two column vectors x and y of length n is the scalar quantity $x^T y$ and is defined as

$$x^T y = \sum_{i=1}^{n} x_i \, y_i.$$

In other words, the dot product is the sum of the componentwise product of the vector elements. If x and y are allowed to be complex, then the inner product is defined as $x^H y$. Note that x and y must have the same length in order to compute their dot product.

As an example, suppose that x and y are defined as

$$x = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}; \qquad y = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}.$$

Then

$$x^T y = \begin{pmatrix} 3 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} = 3(2) + 0(1) + (-1)(-2) = 8.$$

It is instructive to examine the dot product in more detail:

$$\begin{aligned} x^T y \quad &= \quad (1 \times n) \cdot (n \times 1) \\ & \qquad\quad\;\; \uparrow \quad\; \uparrow \end{aligned}$$

Here, the inner dimensions match, so the operation is permitted. Further, if we discard the inner dimensions, the remaining dimensions are 1×1 (*i.e.*, a scalar). The dot product is also called the *scalar product*

3.1.6 The Outer Product

The second type of vector product, the *outer product*, is less common. Let x be a vector of lenth n and y be a vector of length m. The outer product of x and y is defined to be xy^T. Again, examine this product in more detail.

$$\begin{aligned} xy^T \quad &= \quad (n \times 1) \cdot (1 \times m). \\ & \qquad\quad\;\; \uparrow \quad\; \uparrow \end{aligned}$$

The inner dimensions (1 and 1) match. The remaining dimensions are $n \times m$, hence, an outer product results in a matrix. Also, note that the length of the vectors x and y does not have to be the same.

Let $A = xy^T$. Then the elements of the outer product matrix A are denoted by

$$a_{i,j} = x_i \, y_j, \quad \text{for } i = 1, \ldots, n, \; j = 1, \ldots, m.$$

Finally, as with the dot product, if the vectors are complex then the outer product is defined as xy^H. For x and y as defined in Section 3.1.2,

$$xy^T = \begin{pmatrix} 3(2) & 3(1) & 3(-2) \\ 0(2) & 0(1) & 0(-2) \\ 1(2) & 1(1) & 1(-2) \end{pmatrix} = \begin{pmatrix} 6 & 3 & -6 \\ 0 & 0 & 0 \\ 2 & 1 & -2 \end{pmatrix}.$$

3.2 Matrix-Vector Products

Because matrices and vectors have different dimensions, they cannot be added or subtracted. They can, however, be multiplied.

Suppose x is a column vector of length m and A is an $n \times m$ matrix. Then the matrix-vector product is denoted by Ax. Examining the quantities more closely,

$$\begin{aligned} Ax \quad &= \quad (n \times m) \cdot (m \times 1). \\ & \qquad\quad\;\; \uparrow \qquad \uparrow \end{aligned}$$

Thus, the result of a matrix-vector product is another column vector of length n. The elements of this vector (let $z = Ax$) are defined by

$$z_i = \sum_{k=1}^{m} a_{i,k}\, x_k \quad \text{for } \; i = 1, \ldots, n.$$

This more easily remembered in terms of dot products as

$$z_i = \text{dot product of } A_{i,*} \text{ and } x.$$

where $A_{i,*} =$ row i of A.

Another version of a matrix-vector product can occur when the vector is multiplied from the left. If y is assumed to be a column vector of length n, then the only way a matrix-vector product can be formed is $u^T = y^T A$ since

$$\begin{array}{rcl} y^T A & = & (1 \times n) \cdot (n \times m). \\ & & \qquad\;\; \uparrow \quad\;\; \uparrow \end{array}$$

The result of this operation is another row vector of length m. Again,

$$u_j^T = \sum_{k=1}^{n} y_k\, a_{k,j} \quad \text{for } \; j = 1, \ldots, m$$

or

$$u_j^T = \text{dot product of } y \text{ and } A_{*,j}$$

where $A_{*,j} =$ column j of A.

There is a third type of matrix-vector product called the *weighted scalar product*. This is defined as $\alpha = y^T A x$. It can be thought of as either the matrix vector product $z = Ax$ followed by the dot product $y^T z$ or the matrix-vector product $u^T = y^T A$ followed by the dot product $u^T x$.

For example, define x and y as in section 3.1.5 and A as in Section 2.3. Then

$$\begin{array}{rcl} Ax & = & \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \quad = \quad \begin{pmatrix} 1(3) + 2(0) + 3(1) \\ 4(3) + 5(0) + 6(1) \\ 7(3) + 8(0) + 9(1) \end{pmatrix} \\ & = & \begin{pmatrix} 6 \\ 18 \\ 30 \end{pmatrix} \\ y^T A x & = & \begin{pmatrix} -8 & -7 & -6 \end{pmatrix} \begin{pmatrix} 6 \\ 18 \\ 30 \end{pmatrix} \quad = \quad -8(6) + (-7)(18) + (-6)(30) \quad = -354 \end{array}$$

3.3 Matrix-Matrix Operations

3.3.1 Matrix Addition

Matrices can be added or subtracted if they have the same dimensions. Like vectors, the sum is performed componentwise. For matrices A and B of dimension $n \times m$,

$$C \;\; = \;\; A + B$$

or in index notation,

$$c_{i,j} \;\; = \;\; a_{i,j} \; + \; b_{i,j}, \quad \text{for } \; i = 1, \ldots, n, \; j = 1, \ldots, m.$$

Matrices can be multiplied by scalars and an expression such as

$$E = 3A - 6B$$

would be computed in a straightfoward way.

3.3.2 Matrix-Matrix Products

Matrix-matrix products are most simply understood in terms of dot products. Suppose A is an $n \times k$ matrix and B is an $k \times m$ matrix. Then the matrix product $C = AB$ is a matrix of size $n \times m$ whose elements are given by

$$c_{i,j} = \sum_{l=1}^{k} a_{i,l}\, b_{l,j}, \quad \text{for } i = 1, \ldots, n, \; j = 1, \ldots, m.$$

In terms of dot products,

$$c_{i,j} = A_{i,*}\, B_{*,j}, \quad \text{for } i = 1, \ldots, n, \; j = 1, \ldots, m$$

where $A_{i,*}$ = row i of A and $B_{*,j}$ = column j of B. As an example, take A and B as defined in Section 2.3 and let $C = AB$. Then

$$\begin{aligned} C &= \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 0 & 5 \\ 1 & 9 \end{pmatrix} \\ &= \begin{pmatrix} 6 & 35 \\ 18 & 71 \\ 30 & 107 \end{pmatrix}. \end{aligned}$$

The elements $c_{2,1}$ and $c_{3,2}$ are computed from

$$\begin{aligned} c_{2,1} &= \begin{pmatrix} 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} &= 4(3) + 5(0) + 6(1) &= 18 \\ c_{3,2} &= \begin{pmatrix} 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} -2 \\ 5 \\ 9 \end{pmatrix} &= 7(-2) + 8(5) + 9(9) &= 107. \end{aligned}$$

The other elements of C are computed in a similar way.

Matrix-matrix products can also be viewed as collections of matrix-vector products. For example, partition the matrices B and C into columns as

$$B = \begin{pmatrix} b_{*,1} & b_{*,2} \end{pmatrix}$$

and

$$C = \begin{pmatrix} c_{*,1} & c_{*,2} \end{pmatrix}.$$

Then $C = AB$ can be viewed as a collection of column vectors, each of which is the result of the matrix-vector product of A with a column of B.

$$\begin{aligned} AB &= A \begin{pmatrix} b_{*,1} & b_{*,2} \end{pmatrix} \\ &= \begin{pmatrix} Ab_{*,1} & Ab_{*,2} \end{pmatrix} \\ &= \begin{pmatrix} c_{*,1} & c_{*,2} \end{pmatrix} \\ &= C \end{aligned}$$

For example, the first column of C is given by

$$\begin{aligned} c_{*,1} &= A\, b_{*,1} \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 6 \\ 18 \\ 30 \end{pmatrix}. \end{aligned}$$

An analogous process can be defined in terms of a row based partitioning of A.

3.4 Important Relations

Finally, some other important linear algebra relationships are:

1) Matrix multiplication is not a commutative operation thus,

$$AB \neq BA.$$

For example, suppose A is an $n \times k$ matrix and B is an $k \times m$ matrix. Then the product AB is an $n \times m$ matrix because

$$AB = (n \times k) \cdot (k \times m) = (n \times m),$$

but the product BA is does not exist. Examining the dimensions of the matrices, we see that

$$BA = (k \times m) \cdot (n \times k).$$

The inner dimensions don't match so this product is undefined.

2) The transpose of a product of vectors/matrices is the product of the individual transposes in reverse order, *i.e.*

$$(AB)^T = B^T A^T.$$

More relationships will be given later.

4 Norms

The need to compare two scalars arises frequently in mathematics and programming. Normally when we want to determine which of two scalars is *larger*, we mean *larger* in the *absolute value sense.* For example, if $a = -36$ and $b = 4$, then

$$a < b$$

in the number line sense, but

$$|a| > |b|$$

in the absolute value sense.

Similarly, the need to compare the magnidute of two vectors or matrices also arises, but the nature of the comparison is less clear. For example, suppose x and y are defined to be

$$x = \begin{pmatrix} 3 \\ 1 \end{pmatrix}; \qquad y = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

Which vector is larger? You could argue that x is larger since it has the largest element ($x_1 = 3$). However, you could also argue that x and y are the same size since the sum of their elements is the same (4).

We need a way to measure the magnidute of vectors and matrices so that comparisons can be made. In a sense, we need to generalize the concept of absolute value to vectors and matrices. The quantity that does this is a *norm.* A norm is an operation that is performed on a vector or matrix that reduces its elements to a scalar value that is representative of its magnitude. We use double vertical bars ($\| * \|$)to represent a norm.

A norm is a generalization of the concept of absolute value thus it should have the same properties as absolute value. In order to be a norm, an operation must satisfy the following criteria (x and y are vectors or matrices):

i) $\|x\| > 0$

ii) $\|x\| = 0 \quad \text{iff} \quad x = 0$

iii) $\|\alpha x\| = |\alpha| \|x\| \quad$ for any scalar α

iv) $\|x + y\| \leq \|x\| + \|y\| \quad$ (triangle inequality)

In addition, if either of x or y is a matrix, norms are also required to satisfy the compatibility condition

v) $\|xy\| \leq \|x\| \cdot \|y\| \quad$ (assuming the product makes sense)

4.1 Vector Norms

The most commonly used vector norms are given below.

a) 1-norm – Given a column vector z of length n, the 1-norm of the vector is defined as

$$\begin{aligned}\|z\|_1 &= \sum_{i=1}^{n} |z_i| \\ &= |z_1| + |z_2| + \cdots + |z_n|.\end{aligned}$$

b) ∞-norm – Given a column vector z of length n, the ∞-norm of the vector is defined as

$$\begin{aligned}\|z\|_\infty &= \max_{1\le i\le n} |z_i| \\ &= \text{maximum element of } z \text{ in absolute value.}\end{aligned}$$

c) 2-norm – Given a column vector z of length n, the 2-norm of the vector is defined as

$$\begin{aligned}\|z\|_2 &= \sqrt{\sum_{i=1}^{n} |z_i|^2} \\ &= \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2}.\end{aligned}$$

It is easy to demonstrate that

$$\begin{aligned}\|z\|_2 &= \sqrt{z^T z} \\ &= \sqrt{\text{dot product of } z \text{ with itself.}}\end{aligned}$$

This norm is also called the Pythagorean norm, the Euclidian norm or the distance norm since it measures the Euclidian length of the vector.

d) p-norm – Given a column vector z of length n, the p-norm of the vector is defined as

$$\begin{aligned}\|z\|_p &= \sqrt[p]{\sum_{i=1}^{n} |z_i|^p} \\ &= \sqrt[p]{|z_1|^p + |z_2|^p + \cdots + |z_n|^p}.\end{aligned}$$

This norm is not used in practice, but it can be used to prove that

$$\lim_{p\to\infty} \|z\|_p = \|z\|_\infty.$$

As an example, suppose that x and y are defined as

$$x = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}; \qquad y = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}.$$

It can quickly be verified that

$$\begin{aligned}\|x\|_1 &= |3| + |0| + |-1| = 4 \\ \|x\|_2 &= \sqrt{|3|^2 + |0|^2 + |-1|^2} = \sqrt{10} \\ \|x\|_\infty &= \max\left(|3|, |0|, |-1|\right) = 3\end{aligned}$$

and

$$\begin{aligned}
\|y\|_1 &= |2|+|1|+|-2| = 5 \\
\|y\|_2 &= \sqrt{|2|^2+|1|^2+|2|^2} = \sqrt{9} \\
\|y\|_\infty &= \max\left(|2|,|1|,|-2|\right) = 2.
\end{aligned}$$

Thus

$$\begin{aligned}
\|x\|_1 &< \|y\|_1 \\
\|x\|_2 &> \|y\|_2 \\
\|x\|_\infty &> \|y\|_\infty.
\end{aligned}$$

These results indicate that different norms have different values. Also, the outcome of a comparison depends on the particular norm used.

4.2 Matrix Norms

The most commonly used matrix norms are given in the list below.

a) 1-norm – Given an $n \times m$ matrix A, the 1-norm is defined as

$$\|A\|_1 = \max_{1 \le j \le m} \sum_{i=1}^{n} |a_{i,j}|.$$

To compute this norm, add the absolute value of the elements down the columns then take the maximum of the resulting sums.

b) ∞-norm – Given an $n \times m$ matrix A, the ∞-norm is defined as

$$\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{m} |a_{i,j}|.$$

To compute this norm, add the absolute value of the elements across the rows then take the maximum of the resulting sums.

c) Frobenius norm – Given an $n \times m$ matrix A, the Frobenious norm is defined as

$$\|A\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m} |a_{i,j}|^2}.$$

To compute this norm, add the squares of all the elements then take the square root of the sum.

For example, if

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

then

a) The sum of the absolute values down the columns gives the set $\{5, 6, 8\}$. The maximum of these values is 8, so $\|A\|_1 = 8$.

b) The sum of the absolute values across the rows gives the set $\{4, 8, 7\}$. The maximum of these values is 8, so $\|A\|_\infty = 8$.

c) The sum of the squares of all elements of A is

$$1+4+1+9+1+16+1+9+9 = 51,$$

so $\|A\|_F = \sqrt{51}$.

A handy way to remember which way to sum when computing the 1- and ∞- norms is that a 1 looks like a column and an ∞ looks like a row.