

MATRIX ALGEBRA REVIEW FOR STATISTICS

LESSON 1 NOTES

Rev. 2.0

by

Robert A. LaBudde, Ph.D.

Notes prepared for online courses at
The Institute for Statistics Education
at Statistics.com

I. INTRODUCTION

This brief text is meant as a review of vector and matrix algebra for students of statistics who plan to go on to study the methods of multivariate statistics, where the knowledge of notation and basic concepts of matrix algebra is essential to understanding. This text is not meant as a thorough and complete coverage of linear algebra as a field of study in itself.

Vectors and matrices are one- and two-dimensional ordered arrays of numbers with an associated collection of mathematical operations (an ‘algebra’) that they are ‘closed’ under (i.e., the result of the operation is once again a vector or matrix). This contrasts with other data constructs such as ‘lists’ (i.e., ordered collection of things that do not have to be numbers, and may be collections themselves), ‘data frames’ (i.e., two-dimensional arrays with each column having elements all of the same kind (e.g., strings or numbers)), ‘tables’ (multidimensional ordered arrays, with focus on structuring rather than numerical operations) and 3 or higher dimensional arrays.

It is the numerical algebra of operations that makes vectors and matrices useful in statistics.

The notation for vectors and matrices is quite difficult to implement using a word processor not designed for the purpose, so examples will be given mostly using the programming language R, which handles vectors and matrices as native objects. R is the basis of an freeware open source large-scale statistical system which is becoming the de facto standard for statistical computing. Knowing R is not necessary to do the assignments or understand this material, but it is useful here to illustrate numerical examples. Other statistical programming systems, such as JMP in SAS, also allow a wide variety of matrix operations, but lack the useful symbolic representations available with R.

Some engineering programming languages, such as MATLAB[®] and GAUSS[®], are designed specifically to handle matrix operations in a symbolic manner. Also, Texas Instruments calculators, such as the TI-8x and TI-9x series, can manipulate vectors and matrices and perform basic operations. Spreadsheet software, such as Microsoft Excel, can do matrix multiplication and inversion.

II. OBJECTS:

Scalar: A single numeric value (number). Denoted here by a non-bold symbol. E.g., $s = 1.4142$.

Vector: A 1-dimensional n-tuple (ordered collection) of n scalar elements. Denoted here by a lower-case bold symbol. E.g., $\mathbf{v} = [2, 3, 5]$. The i-th element of a vector is denoted by the notation v_i or $\mathbf{v}[i]$. For the previous example, $v_1 = 2$, $v_2 = 3$ and $v_3 = 5$. Note that a 1-tuple vector $\mathbf{u} = [3]$ is *not* the same as the scalar $a = 3$. They are two different classes of objects.

Row vector: A vector orientated horizontally. E.g.,

$$\mathbf{u} = [2 \ 3 \ 5] \quad (\text{II.1})$$

Column vector: A vector orientated vertically, E.g.,

$$\mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} \quad (\text{II.2})$$

Length or order of a vector: The number of elements in a vector. E.g., the order of \mathbf{u} in eq.(II.1) and \mathbf{v} in eq.(II.2) are both 3. Note that the term ‘length’ is frequently used instead as synonymous to ‘norm’ (see Chapter III below), so we will prefer ‘order’ here as the number of elements in a vector. Note, however, that R uses ‘length’ for this purpose.

Matrix: a 2-dimensional ($n \times m$) set of n row vectors of order m, or a set of m column vectors of order n. Denoted here by an upper-case bold symbol. Displayed using in-line notation as $\mathbf{A} = [1, 2; 3, 4; 5, 6]$, or, more commonly, as a two-dimension format

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \quad (\text{II.3})$$

The elements of a matrix \mathbf{A} are denoted algebraically by A_{ij} where i is the row number and j is the column number.

A row vector may be considered equivalent to a matrix of one row with the same elements, and a column vector considered equivalent to a matrix of one column with the same elements.

Order of a matrix: A matrix organized as n rows and m columns is said to be of order n x m.

Square matrix: A matrix is called ‘square’ if its row order equals its column order, i.e., it is an $n \times n$ matrix.

Identity matrix: The identity matrix \mathbf{I} of order $n \times n$ is a square matrix whose elements $I_{ij} = 0$ if $i \neq j$ and 1 if $i = j$. For example, the 3×3 identity matrix is

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{II.4})$$

Zero matrix: The $\mathbf{0}$ matrix of order $n \times n$ is a square matrix whose elements are all 0. For example, the 3×3 zero matrix is

$$\mathbf{0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (\text{II.5})$$

Array: a d -dimensional ($n_1 \times n_2 \times \dots \times n_d$) collection of elements. Denoted here by an upper-case bold italic symbol. The elements of a $3 \times 4 \times 2$ array \mathbf{B} are denoted algebraically by B_{ijk} , where $i = 1$ to 3, $j = 1$ to 4, and $k = 1$ to 2.

Vectors and matrices are examples of one and two dimensional arrays.

III. VECTOR OPERATIONS:

Row vector addition: Row vectors of the same order n are closed under addition and subtraction. If $\mathbf{u} = [u_1, u_2, \dots, u_n]$ and $\mathbf{v} = [v_1, v_2, \dots, v_n]$ are two row vectors of order n , then $\mathbf{u} \pm \mathbf{v}$ is a row vector of order n also, and its elements are $u_i \pm v_i$ for $i = 1, 2, \dots, n$.

Column vector addition: Column vectors of the same order n are also closed under addition and subtraction, with the resultant column vector elements equal to the sum or difference of the addend elements.

Transpose: The transpose of a row vector is the column vector of the same order and same elements, and similarly the transpose of a column vector is a row vector with the same elements. The transpose is denoted by a superscript "t" or by an apostrophe '. E.g., the row vector \mathbf{u} in eq.(II.1) is the transpose of the column vector \mathbf{v} in eq.(II.2), or $\mathbf{u} = \mathbf{v}^t$. Similarly, the column vector \mathbf{v} in eq.(II.2) is the transpose of the row vector \mathbf{u} in eq.(II.1), or $\mathbf{v} = \mathbf{u}^t$. The transpose of the transpose of a vector is the original vector itself: $\mathbf{u} = (\mathbf{u}^t)^t$.

Row vector scalar multiplication: A scalar c times a row vector \mathbf{v} of order n is a row vector of order n (denoted ' $c \mathbf{v}$ ') with elements $c v_i$. Scalar multiplication is distributive across row vector addition and subtraction. I.e.,

$$c(\mathbf{u} \pm \mathbf{v}) = (c\mathbf{u}) \pm (c\mathbf{v}) \quad (\text{III.1})$$

Column vector scalar multiplication: A scalar c times a column vector \mathbf{v} of order n is a column vector of order n (denoted ' $c \mathbf{v}$ ') with elements $c v_i$. Scalar multiplication is distributive across row vector addition and subtraction. I.e.,

$$c(\mathbf{u} \pm \mathbf{v}) = (c\mathbf{u}) \pm (c\mathbf{v}) \quad (\text{III.2})$$

Inner product: The inner product (sometimes called the "dot" product) of a row vector \mathbf{u} and a column vector \mathbf{v} of the same order n is denoted ' $\mathbf{u} \mathbf{v}$ ' or ' $\mathbf{u} \cdot \mathbf{v}$ ' or ' $\mathbf{u} * \mathbf{v}$ ', and is a scalar with the value of the sum of the product of the elements of the two vectors. I.e.,

$$\mathbf{u} \mathbf{v} = \sum_{i=1}^n u_i v_i \quad (\text{III.3})$$

For example, if $\mathbf{u} = [1 \ 2 \ 3]$ and $\mathbf{v} = [4 \ 5 \ 6]^t$, then

$$\mathbf{u} \mathbf{v} = (1)(4) + (2)(5) + (3)(6) = 32 \quad (\text{III.4})$$

Using R to reproduce this example (noting that R does not normally distinguish explicitly the difference between row and column vectors):

```
> u = c(1,2,3) #vector u = [1 2 3]
```

```

> u
[1] 1 2 3
> v = c(4,5,6) #vector v = [4 5 6]
> v
[1] 4 5 6
> u %*% v #Inner product of u and v
      [,1]
[1,]    32

```

Because of the way in which the inner product is defined as a scalar, the inner product can actually be computed between any of 2 row vectors, 2 column vectors, a row and a column vector and a column and a row vector, so long as both vectors have the same order (number of elements).

Norm or magnitude of a vector: The ‘norm’ or ‘magnitude’ (and sometimes ‘length’) of a vector \mathbf{v} , denoted $|\mathbf{v}|$ or $\|\mathbf{v}\|$ is the scalar square root of the inner product of the vector with its transpose. For a row vector \mathbf{u} ,

$$|\mathbf{u}| = \sqrt{(\mathbf{u} \mathbf{u}^t)} \quad (\text{III.5})$$

Similarly, for a column vector \mathbf{v} ,

$$|\mathbf{v}| = \sqrt{(\mathbf{v}^t \mathbf{v})} \quad (\text{III.6})$$

Angle between two vectors: Note that the inner product of two vectors geometrically has the form

$$\mathbf{u} \mathbf{v} = |\mathbf{u}| |\mathbf{v}| \cos(\theta) \quad -\pi/2 \leq \theta \leq \pi/2 \quad (\text{III.7})$$

where

$$\cos(\theta) = \frac{\sum u_i v_i}{(|\mathbf{u}| |\mathbf{v}|)} \quad (\text{III.8})$$

is the angle between the two vectors (geometrically), or more generally a measure similar to a correlation coefficient as the cosine lies between -1 and $+1$.

Outer product: The ‘outer’ or ‘tensor’ product of a column vector \mathbf{u} of order n and a row vector \mathbf{v} of order m is denoted $\mathbf{u} \times \mathbf{v}$ here (but most books just use the notation $\mathbf{u} \mathbf{v}$, with the inner product as $\mathbf{u} \cdot \mathbf{v}$ or inferred by which of \mathbf{u} and \mathbf{v} is the row and which is the column vector), and is an $n \times m$ matrix whose i j -th element is $u_i v_j$. Note that \mathbf{u} and \mathbf{v} do *not* have to be the same order. E.g., for the column vector $\mathbf{u} = [1 \ 2 \ 3]^t$ and the row vector $\mathbf{v} = [4 \ 5]$ the outer product is

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} [4 \ 5] = \begin{bmatrix} 4 & 5 \\ 8 & 10 \\ 12 & 15 \end{bmatrix} \quad (\text{III.9})$$

Note that the outer product $\mathbf{u} \times \mathbf{v}$ of two vectors is a *matrix*.
Using R to illustrate the example:

```
> u = c(1,2,3) #vector u = [1 2 3]
> u
[1] 1 2 3

> v = c(4,5)   #vector v = [4 5]
> v
[1] 4 5

> u %o% v      #outer product of u and v
      [,1] [,2]
[1,]    4    5
[2,]    8   10
[3,]   12   15
```

NOTE: It is extremely cumbersome in text to keep the subtle distinction between row and column vectors. Frequently, they are just referred to as “vectors”, and the distinction between a row vector and a column vector indicated by the context used. Typically column vectors are used more, so a “vector” should be considered a column vector until the context implies otherwise. Judicious use of the transpose operator keeps the notation exact in cases of ambiguity.

From now on, we will refer to both row and column vectors as just “vectors” unless the context requires clarification.