Matrix Algebra

ERSH 8320

Today's Lecture

- Matrix Algebra.
- Matrix Algebra.
- Matrix Algebra.

Matrix Introduction

- Imagine you are interested in studying the cultural differences in student achievement in high school.
- Your study gets IRB approval and you work hard to get parental approval.
- You go into school and collect data using many different instruments.
- You then input the data into your favorite stat package (or MS Excel).
- How do you think the data is stored?

Definitions

- We begin this class with some general definitions (from dictionary.com):
  - Matrix:
    1. A rectangular array of numeric or algebraic quantities subject to mathematical operations.
    2. The substrate on or within which a fungus grows.
  - Algebra:
    1. A branch of mathematics in which symbols, usually letters of the alphabet, represent numbers or members of a specified set and are used to represent quantities and to express general relationships that hold for all members of the set.
    2. A set together with a pair of binary operations defined on the set. Usually, the set and the operations include an identity element, and the operations are commutative or associative.

Matrices

- Away from the definition, a matrix is simply a rectangular way of storing data.
- Matrices can have unlimited dimensions, however for our purposes, all matrices will be in two dimensions:
  - Rows
  - Columns
- Matrices are symbolized by **boldface** font in text.

$$\mathbf{A} = \begin{bmatrix} 4 & 7 & 5 \\ 6 & 6 & 3 \end{bmatrix}$$

MATLAB

- To help demonstrate the topics we will discuss today, I will be showing examples in MATLAB.
- MATLAB (Matrix Laboratory) is a scientific computing package typically used in other technical fields.
- For many statistical applications MATLAB is very useful.
- SPSS and SAS both have matrix computing capabilities, but (in my opinion) neither are as efficient, as user friendly, or as flexible as MATLAB.
  - It is better to leave most of the statistical computing to the computer scientists.

Matrix Elements

- A matrix is composed of a set of elements, each denoted it's row and column position within the matrix.
- For a matrix **A** of size $r \times c$, each element is denoted by:

$$a_{ij}$$

  - The first subscript is the index for the rows: $i = 1, \ldots, r$.
  - The second subscript is the index for the columns: $j = 1, \ldots, c$.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1c} \\ a_{21} & a_{22} & \ldots & a_{1c} \\ \vdots & \vdots & \vdots & \vdots \\ a_{r1} & a_{r2} & \ldots & a_{rc} \end{bmatrix}$$

Transpose

- The transpose of a matrix is simply the switching of the indices for rows and columns.
- An element a_{ij} in the original matrix (in the i^{th} row and j^{th} column) would be a_{ji} in the transposed matrix (in the j^{th} row and the i^{th} column).
- If the original matrix was of size $i \times j$ the transposed matrix would be of size $j \times i$.

$$\mathbf{A} = \begin{bmatrix} 4 & 7 & 5 \\ 6 & 6 & 3 \end{bmatrix}$$

$$\mathbf{A}' = \begin{bmatrix} 4 & 6 \\ 7 & 6 \\ 5 & 3 \end{bmatrix}$$

Types of Matrices

- Square Matrix: A matrix that as the same number of rows and columns.
  - Correlation and covariance matrices are examples of square matrices.
- Diagonal Matrix: A diagonal matrix is a square matrix with non-zero elements down the diagonal and zero values for the off-diagonal elements.

$$\mathbf{A} = \begin{bmatrix} 2.759 & 0 & 0 \\ 0 & 1.643 & 0 \\ 0 & 0 & 0.879 \end{bmatrix}$$

- Symmetric Matrix: A symmetric matrix is a square matrix where $a_{ij} = a_{ji}$ for all elements in i and j.
  - Correlation/covariance and distance matrices are examples of symmetric matrices.

Vectors

- A vector is a matrix where one dimension is equal to size one.
  - Column vector: A column vector is a matrix of size $r \times 1$.
  - Row vector: A row vector is a matrix of size $1 \times c$.
- Vectors allow for geometric representations of matrices.
- The Pearson correlation coefficient is a function of the angle between vectors.
- Much of the statistical theory given in this course (and ANOVA-type courses) can be conceptualized by projections of vectors (think of the dependent variable Y as a column vector).

Scalars

- A scalar is a matrix of size 1×1.
- Scalars can be thought of as any single value.
- The difficult concept to get used to is seeing a number as a matrix:

$$A = \left[\begin{array}{c} 2.759 \end{array} \right]$$

Algebraic Operations

- As mentioned in the definition at the beginning of class, algebra is simply a set of math that defines basic operations.
  - Identity
  - Zero
  - Addition
  - Subtraction
  - Multiplication
  - Division
- Matrix algebra is simply the use of these operations with matrices.

Matrix Addition

- Matrix addition is very much like scalar addition, the only constraint is that the two matrices must be of the same size (same number of rows and columns).

- The resulting matrix contains elements that are simply the result of adding two scalars.

Matrix Addition

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix}$$

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \\ a_{31}+b_{31} & a_{32}+b_{32} \\ a_{41}+b_{41} & a_{42}+b_{42} \end{bmatrix}$$

Matrix Subtraction

- Matrix subtraction is identical to matrix addition, with the exception that all elements of the new matrix are the subtracted elements of the previous matrices.
- Again, the only constraint is that the two matrices must be of the same size (same number of rows and columns).
- The resulting matrix contains elements that are simply the result of subtracting two scalars.

Matrix Subtraction

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix}$$

$$\mathbf{A} - \mathbf{B} = \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} \\ a_{21} - b_{21} & a_{22} - b_{22} \\ a_{31} - b_{31} & a_{32} - b_{32} \\ a_{41} - b_{41} & a_{42} - b_{42} \end{bmatrix}$$

Matrix Multiplication

- Unlike matrix addition and subtraction, matrix multiplication is much more complicated.
- Matrix multiplication results in a new matrix that can be of differing size from either of the two original matrices.
- Matrix multiplication is defined only for matrices where the number of columns of the first matrix is equal to the number of rows of the second matrix.
- The resulting matrix as the same number of rows as the first matrix, and the same number of columns as the second matrix.

$$\underset{(r \times \mathbf{c})}{\mathbf{A}} \quad \underset{(\mathbf{c} \times k)}{\mathbf{B}} \quad = \quad \underset{(r \times k)}{\mathbf{C}}$$

Overview

Matrix Algebra

Algebra
- Algebra
- Addition
- Subtraction
- Multiplication
- Identity
- Zero
- “Division”
- Singular Matrices

More Matrices

Wrapping Up

Matrix Multiplication

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} & a_{31}b_{13} + a_{32}b_{23} \\ a_{41}b_{11} + a_{42}b_{21} & a_{41}b_{12} + a_{42}b_{22} & a_{41}b_{13} + a_{42}b_{23} \end{bmatrix}$$

Matrix Multiplication

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} & a_{31}b_{13} + a_{32}b_{23} \\ a_{41}b_{11} + a_{42}b_{21} & a_{41}b_{12} + a_{42}b_{22} & a_{41}b_{13} + a_{42}b_{23} \end{bmatrix}$$

Matrix Multiplication

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} & a_{31}b_{13} + a_{32}b_{23} \\ a_{41}b_{11} + a_{42}b_{21} & a_{41}b_{12} + a_{42}b_{22} & a_{41}b_{13} + a_{42}b_{23} \end{bmatrix}$$

Multiplication and Summation

- Because of the additive nature induced by matrix multiplication, many statistical formulas that use:

$$\sum$$

  can be expressed by matrix notation.

- For instance, consider a single variable X_i, with $i = 1, \ldots, N$ observations.

- Putting the set of observations into the column vector **X**, of size $N \times 1$, we can show that:

$$\sum_{i=1}^{N} X^2 = \mathbf{X}' \, \mathbf{X}$$

Matrix Multiplication by Scalar

- Recall that a scalar is simply a matrix of size (1×1).
- Matrix multiplication by a scalar causes all elements of the matrix to be multiplied by the scalar.
- The resulting matrix has all elements multiplied by the scalar.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \qquad s \times \mathbf{A} = \begin{bmatrix} s \times a_{11} & s \times a_{12} \\ s \times a_{21} & s \times a_{22} \\ s \times a_{31} & s \times a_{32} \\ s \times a_{41} & s \times a_{42} \end{bmatrix}$$

Identity Matrix

- The identity matrix is defined as a matrix that when multiplied with another matrix produces that original matrix:

$$\mathbf{A}\,\mathbf{I} = \mathbf{A}$$

$$\mathbf{I}\,\mathbf{A} = \mathbf{A}$$

- The identity matrix is simply a square matrix that has all off-diagonal elements equal to zero, and all diagonal elements equal to one.

$$\mathbf{I}_{(3\times 3)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Zero Matrix

- The zero matrix is defined as a matrix that when multiplied with another matrix produces the matrix:

$$\mathbf{A}\,\mathbf{0} = \mathbf{0}$$

$$\mathbf{0}\,\mathbf{A} = \mathbf{0}$$

- The zero matrix is simply a square matrix that has all elements equal to zero.

$$\mathbf{0}_{(3\times 3)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Matrix "Division": The Inverse

- Recall from basic math that:

$$\frac{a}{b} = \frac{1}{b}a = b^{-1}a$$

- And that:

$$\frac{a}{a} = 1$$

- Matrix inverses are just like division in basic math.

The Inverse

- For a square matrix, an inverse matrix is simply the matrix that when pre-multiplied with another matrix produces the identity matrix:

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

- Matrix inverse calculation is complicated and unnecessary since computers are much more efficient at finding inverses of matrices.
- One point of emphasis: just like in regular division, division by zero is undefined.
- By analogy - not all matrices can be inverted.

Singular Matrices

- A matrix that cannot be inverted is called a *singular* matrix.
- In statistics, common causes of singular matrices are found by linear dependence among the rows or columns of a square matrix.
- Linear dependence can be cause by combinations of variables, or by variables with extreme correlations (either near 1.00 or -1.00).

Advanced Matrix Functions/Operations

- We end our matrix discussion with some advanced topics.
- Some of these topics go beyond the scope of the book, but may be encountered in your course as “self-sufficient statisticians.”
- All of these topics are related to applied regression analyses.

Matrix Determinants

- A square matrix can be characterized by a scalar value called a determinant.

$$\det \mathbf{A} = |\mathbf{A}|$$

- Much like the matrix inverse, calculation of the determinant is very complicated and tedious, and is best left to computers.
- What can be learned from determinants is if a matrix is singular.
- Matrices with determinants that are greater than zero are said to be "positive definite," a byproduct of which is that a positive matrix is non-singular.

Matrix Orthogonality

- A square matrix (**A**) is said to be *orthogonal* if:

$$\mathbf{A}\mathbf{A}' = \mathbf{A}'\mathbf{A} = \mathbf{I}$$

- Orthogonal matrices are characterized by two properties:
  1. The product of all row vector multiples is the zero matrix (perpendicular vectors).
  2. For each row vector, the sum of all elements is one.

Eigenvalues and Eigenvectors

- A square matrix can be decomposed into a set of eigenvalues and eigenvectors.

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x}$$

- From a statistical standpoint:
  - Principal components are comprised of linear combination of a set of variables weighed by the eigenvectors.
  - The eigenvalues represent the proportion of variance accounted for by specific principal components.
  - Each principal component is orthogonal to the next, producing a set of uncorrelated variables that may be used for regression purposes.

Spectral Decompositions

- Imagine that a matrix **A** is of size $k \times k$.
- **A** then has:
  - k eigenvalues: λ_i, $i = 1, \ldots, k$.
  - k eigenvectors: $\mathbf{e}_i$, $i = 1, \ldots, k$ (each of size $k \times 1$.
- **A** can be expressed by:

$$\mathbf{A} = \sum_{i=1}^{k} \lambda_i \mathbf{e}_i \mathbf{e}_i'$$

- This expression is called the *Spectral Decomposition*, where **A** is decomposed into k parts.
- One can find $\mathbf{A}^{-1}$ by taking $\frac{1}{\lambda}$ in the spectral decomposition.

Final Thought

- Matrices allow for complex mathematical formulas to be displayed concisely.
- You may encounter matrix functions as you learn about statistical techniques.

- This lecture may have very little influence on your day-to-day life, but it sets the stage for the next: the General Linear Model.

Next Time

- Matrices help us get things we are familiar with.
- The General Linear Model.