Matrix Algebra — A Minimal Introduction

James H. Steiger

Department of Psychology and Human Development
Vanderbilt University

Matrix Algebra — A Minimal Introduction

Definition of a Linear Combination

Definition

- Suppose you have two predictors, x_1 and x_2
- The variable $\hat{y} = b_1 x_1 + b_2 x_2$ is said to be a *linear combination* of x_1 and x_2
- b_1 and b_2 are the *linear weights* which, in a sense, define a particular linear combination

Linear Combinations in Regression

Linear Combinations in Regression

- As we just saw, a linear model for y is a linear combination of one or more predictor variables, plus an intercept and an error term
- Statistical laws that generally apply to linear combinations must then also apply to linear models

Linear Combinations in Regression

Linear Combinations in Regression

- Regression models often contain *many* predictors, so we might well profit by a notation that allows us to talk about linear combinations with any number of predictors
- Matrix algebra provides mathematical tools and notation for discussing linear models compactly

Definition of A Matrix

Definition

- A *matrix* is defined as an ordered array of numbers, of dimensions p,q.

Example

Below is a matrix **A** of dimensions 3×3.

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 4 \\ 7 & 1 & 8 \\ 6 & 6 & 0 \end{pmatrix}$$

Matrix Notation

Notation

- My standard notation for a matrix $\mathbf{A}$ of order p, q will be:

$$ {}_p\mathbf{A}_q $$

- Note that in my notation, matrices and vectors are in boldface

Matrix Notation

Elements of a Matrix

Matrix Elements

- The individual numbers in a matrix are its *elements*
- We use the following notation to indicate that "**A** is a matrix with elements a_{ij} in the i, jth position"

$$\mathbf{A} = \{a_{ij}\}$$

Matrix Notation

Subscript Notation

Subscript Notation

- When we refer to element a_{ij}, the *first* subscript will refer to the *row position* of the elements in the array
- The *second* subscript (regardless of which letter is used in this position) will refer to the column position.
- Hence, a typical matrix ${}_p\mathbf{A}_q$ will be of the form:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1q} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2q} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3q} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & a_{p3} & \cdots & a_{pq} \end{pmatrix}$$

Types of Matrices

On several subsequent slides, we will define a number of types of matrices that are referred to frequently in practice.

Types of Matrices

Rectangular Matrix

Rectangular Matrix

For any ${}_p\mathbf{A}_q$, if $p \neq q$, $\mathbf{A}$ is a *rectangular* matrix

Example

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}$$

Types of Matrices

Square Matrix

Square Matrix

For any ${}_p\mathbf{A}_q$, if $p = q$, $\mathbf{A}$ is a *square* matrix

Example

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

Types of Matrices

Lower Triangular Matrix

Lower Triangular Matrix

For any square matrix $\mathbf{A}$, $\mathbf{A}$ is *lower triangular* if $a_{ij} = 0$ for $i < j$

Example

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 10 \end{pmatrix}$$

Types of Matrices

Upper Triangular Matrix

Upper Triangular Matrix

For any square matrix $\mathbf{A}$, $\mathbf{A}$ is *upper triangular* if $a_{ij} = 0$ for $i > j$

Example

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \\ 0 & 0 & 0 & 10 \end{pmatrix}$$

Types of Matrices

Diagonal Matrix

Diagonal Matrix

For any square matrix $\mathbf{A}$, $\mathbf{A}$ is a *diagonal matrix* if $a_{ij} = 0$ for $i \neq j$

Example

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 7 \end{pmatrix}$$

Types of Matrices

Scalar Matrix

Scalar Matrix

For any diagonal matrix $\mathbf{A}$, if all diagonal elements are equal, $\mathbf{A}$ is a *scalar* matrix

Example

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Types of Matrices

Identity Matrix

Identity Matrix

For any scalar matrix $\mathbf{A}$, if all diagonal elements are 1, $\mathbf{A}$ is an *identity* matrix

Example

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Types of Matrices

Symmetric Matrix

Square Matrix

A square matrix $\mathbf{A}$ is *symmetric* if $a_{ji} = a_{ij} \ \forall \ i, j$

Example

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 4 \\ 3 & 4 & 2 \end{bmatrix}$$

Types of Matrices

Null Matrix

Null Matrix

For any ${}_p\mathbf{A}_q$, $\mathbf{A}$ is a *null* matrix if all elements of $\mathbf{A}$ are 0.

Example

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Types of Matrices

Row Vector

Row Vector

- A *row vector* is a matrix with only one row
- It is common to identify row vectors in matrix notation with lower-case boldface and a "prime" symbol, like this

$$\mathbf{a}'$$

Types of Matrices

Column Vector

Column Vector

- A *column vector* is a matrix with only one column
- It is common to identify column vectors in matrix notation with lower-case boldface, but without the "prime" symbol.

Matrix Operations

Matrix Operations

- Two matrix operations, addition and subtraction, are essentially the same as their familiar scalar equivalents
- But multiplication and division are rather different!
  - There is only a limited notion of division in matrix algebra, and
  - Matrix multiplication shares some properties with scalar multiplication, but in other ways is dramatically different
- We will try to keep reminding you where you need to be careful!

Matrix Addition

Matrix Addition

- For two matrices **A** and **B** to be *conformable* for addition or subtraction, they must have the same numbers of rows and columns
- To add two matrices, simply add the corresponding elements together

Example

$$\begin{pmatrix} 1 & 4 & 1 \\ 1 & 3 & 3 \\ 3 & 0 & 5 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 2 \\ 3 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 6 & 2 \\ 3 & 5 & 5 \\ 6 & 2 & 5 \end{pmatrix}$$

Matrix Subtraction

Matrix Subtraction

- Subtracting matrices works like addition
- You simply subtract corresponding elements

Example

$$\begin{pmatrix} 1 & 4 & 1 \\ 1 & 3 & 3 \\ 3 & 0 & 5 \end{pmatrix} - \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 2 \\ 3 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 0 \\ -1 & 1 & 1 \\ 0 & -2 & 5 \end{pmatrix}$$

Properties of Matrix Addition

Matrix addition has some important mathematical properties, which, fortunately, mimic those of scalar addition and subtraction. Consequently, there is little "negative transfer" involved in generalizing from the scalar to the matrix operations.

Properties of Matrix Addition

For matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, properties include:

- *Associativity*. $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$
- *Commutativity*. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$

Scalar Multiple

Scalar Multiple

- When we multiply a matrix by a scalar, we are computing a *scalar multiple*, not to be confused with a scalar product, which we will learn about subsequently
- To compute a scalar multiple, simply multiply every element of the matrix by the scalar

Example

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

Properties of Scalar Multiplication

For matrices $\mathbf{A}$ and $\mathbf{B}$, and scalars a and b, scalar multiplication has the following mathematical properties:

Properties of Scalar Multiplication

- $(a+b)\mathbf{A} = a\mathbf{A} + b\mathbf{A}$
- $a(\mathbf{A}+\mathbf{B}) = a\mathbf{A} + a\mathbf{B}$
- $a(b\mathbf{A}) = (ab)\mathbf{A}$
- $a\mathbf{A} = \mathbf{A}a$

Scalar Product

Scalar Product

- Given a row vector $\mathbf{x}'$ and a column vector $\mathbf{y}$ having q elements each
- The *scalar product* $\mathbf{x}'\mathbf{y}$ is a scalar equal to the sum of cross-products of the elements of $\mathbf{x}'$ and $\mathbf{y}$.

Example

If $\mathbf{x}' = \begin{pmatrix} 1 & 2 & 1 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}$ then $\mathbf{x}'\mathbf{y} = (1)(2) + (2)(3) + (1)(2) = 10$

Matrix Transposition

Transposing a matrix is an operation which plays a very important role in multivariate statistical theory. The operation, in essence, switches the rows and columns of a matrix.

Matrix Transposition

Let ${}_p\mathbf{A}_q = \{a_{ij}\}$. Then the *transpose* of $\mathbf{A}$, denoted $\mathbf{A}'$, is defined as

$$ {}_q\mathbf{A}'_p = \{a_{ji}\} $$

Example

If ${}_2\mathbf{A}_3 = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, then ${}_3\mathbf{A}'_2 = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$

Matrix Transposition

Some Key Properties

Key Properties of Matrix Transposition

- $(\mathbf{A}')' = \mathbf{A}$
- $(c\mathbf{A})' = c\mathbf{A}'$
- $(\mathbf{A}+\mathbf{B})' = \mathbf{A}' + \mathbf{B}'$
- $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$
- A square matrix $\mathbf{A}$ is symmetric if and only if $\mathbf{A} = \mathbf{A}'$

Matrix Multiplication

Conformability

Conformability

- Matrix multiplication is an operation with properties quite different from its scalar counterpart.
- *order matters* in matrix multiplication.

Matrix Multiplication

Conformability

Conformability

- That is, the matrix product $\mathbf{AB}$ need not be the same as the matrix product $\mathbf{BA}$.
- Indeed, the matrix product $\mathbf{AB}$ might be well-defined, while the product $\mathbf{BA}$ might not exist.
- When we compute the product $\mathbf{AB}$, we say that $\mathbf{A}$ is *post-multiplied* by $\mathbf{B}$, or that $\mathbf{B}$ is *premultiplied* by $\mathbf{A}$

Matrix Multiplication

Dimension of a Matrix Product

If two or more matrices are conformable, there is a strict rule for determining the dimension of their product

Matrix Multiplication — Dimensions of a Product

- The product ${}_p\mathbf{A}_q\mathbf{B}_r$ will be of dimension $p \times r$
- More generally, the product of any number of conformable matrices will have the number of rows in the leftmost matrix, and the number of columns in the rightmost matrix.
- For example, the product ${}_p\mathbf{A}_q\mathbf{B}_r\mathbf{C}_s$ will be of dimensionality $p \times s$

Matrix Multiplication

Three Approaches

Three Approaches

- Matrix multiplication might well be described as the key operation in matrix algebra
- What makes matrix multiplication particularly interesting is that there are numerous lenses through which it may be viewed
- We shall examine 3 ways of "looking at" matrix multiplication.
- All of them rely on *matrix partitioning*, which we'll examine briefly in the next 2 slides

Matrix Multiplication

The Row by Column Approach

Partitioning a Matrix into Rows

- Any $p \times q$ matrix $\mathbf{A}$ may be partitioned into as a set of p rows
- For example, the 2×3 matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \end{pmatrix}$$

  may be thought of as two rows, $(1 \;\; 2 \;\; 3)$ and $(3 \;\; 3 \;\; 3)$ stacked on top of each other
- We have a notation for this. We write

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \end{pmatrix}$$

Matrix Multiplication

The Row by Column Approach

Partitioning a Matrix into Columns

- We can also view any $p \times q$ matrix as a set of q columns, joined side-by-side
- For example, for the 2×3 matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \end{pmatrix}$$

we can write

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{pmatrix}$$

where, for example,

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

Matrix Multiplication

The Row by Column Approach

The Row by Column Approach

- Suppose you wish to multiply the two matrices **A** and **B**, where

$$\mathbf{A} = \begin{pmatrix} 2 & 7 \\ 3 & 5 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix}$$

- You know that the product, $\mathbf{C} = \mathbf{AB}$, will be a 2×3 matrix
- Partition **A** into 2 rows, and **B** into 3 columns.
- Element $c_{i,j}$ is the scalar product of row i of **A** with column j of **B**

Matrix Multiplication

The Row by Column Approach

Again suppose you wish to compute the product $\mathbf{C} = \mathbf{AB}$ using the matrices from the preceding slide.

Example

Compute $c_{1,1}$.

$$\begin{pmatrix} 2 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix}$$

Taking the product of the row 1 of **A** and column 1 of **B**, we obtain $(2)(1) + (7)(2) = 16$

Matrix Multiplication

The Row by Column Approach

Again suppose you wish to compute the product $\mathbf{C} = \mathbf{AB}$ using the matrices from the preceding slide.

Example

Compute $c_{2,3}$.

$$\begin{pmatrix} 2 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix}$$

Taking the product of the row 2 of $\mathbf{A}$ and column 3 of $\mathbf{B}$, we obtain $(3)(1) + (5)(3) = 18$

Matrix Multiplication

Linear Combination of Columns Approach

Linear Combination of Columns

- When you post-multiply a matrix $\mathbf{A}$ by a matrix $\mathbf{B}$, each column of $\mathbf{B}$ generates, in effect, a column of the product $\mathbf{AB}$
- Each column of $\mathbf{B}$ contains a set of linear weights
- These linear weights are applied to the columns of $\mathbf{A}$ to produce a single column of numbers.

Matrix Multiplication

Linear Combination of Columns Approach

Linear Combination of Columns

- Consider the product

$$\begin{pmatrix} 2 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix}$$

- The first column of the product is produced by applying the linear weights 1 and 2 to the columns of the first matrix
- The result is

$$1 \begin{pmatrix} 2 \\ 3 \end{pmatrix} + 2 \begin{pmatrix} 7 \\ 5 \end{pmatrix} = \begin{pmatrix} 16 \\ 13 \end{pmatrix}$$

Matrix Multiplication

Linear Combination of Columns Approach

Linear Combination of Columns

- Consider once again the product

$$\begin{pmatrix} 2 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix}$$

- The second column of the product is produced by applying the linear weights 2 and 2 to the columns of the first matrix
- The result is

$$2\begin{pmatrix} 2 \\ 3 \end{pmatrix} + 2\begin{pmatrix} 7 \\ 5 \end{pmatrix} = \begin{pmatrix} 18 \\ 16 \end{pmatrix}$$

Matrix Multiplication

Linear Combination of Rows Approach

Linear Combination of Rows

- When you pre-multiply a matrix **B** by a matrix **A**, each row of **A** generates, in effect, a rowof the product **AB**
- Each row of **A** contains a set of linear weights
- These linear weights are applied to the rows of **B** to produce a single row vector of numbers.

Matrix Multiplication

Linear Combination of Rows Approach

Linear Combination of Rows

- Consider the product

$$\begin{pmatrix} 2 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix}$$

- The first row of the product is produced by applying the linear weights 2 and 7 to the rows of the second matrix
- The result is

$$2 \begin{pmatrix} 1 & 2 & 1 \end{pmatrix} + 7 \begin{pmatrix} 2 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 16 & 18 & 23 \end{pmatrix}$$

Matrix Multiplication

Mathematical Properties

The following are some key properties of matrix multiplication:

Mathematical Properties of Matrix Multiplication

- *Associativity*.

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$$

- *Not generally commutative*. That is, often $\mathbf{AB} \neq \mathbf{BA}$.
- *Distributive over addition and subtraction*.

$$\mathbf{C}(\mathbf{A} + \mathbf{B}) = \mathbf{CA} + \mathbf{CB}$$

- Assuming it is conformable, the identity matrix $\mathbf{I}$ functions like the number 1, that is ${}_p\mathbf{A}_q\mathbf{I}_q = \mathbf{A}$, and ${}_p\mathbf{I}_p\mathbf{A}_q = \mathbf{A}$.
- $\mathbf{AB} = \mathbf{0}$ does not necessarily imply that either $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$.

Inverse of a Square Matrix

Definition

Matrix Inverse

- A $p \times p$ matrix has an inverse if and only if it is square and of full rank, i.e., i.e., no column of A is a linear combination of the others.
- If a square matrix $\mathbf{A}$ has an inverse, it is the unique square matrix $\mathbf{A}^{-1}$ such that

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

Inverse of a Square Matrix

Properties

Mathematical Properties of Matrix Inverses

- $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$
- If $\mathbf{A} = \mathbf{A}'$, then $\mathbf{A}^{-1} = (\mathbf{A}^{-1})'$
- The inverse of the product of several invertible square matrices is the product of their inverses in reverse order. For example

$$(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$$

- For nonzero scalar c, $(c\mathbf{A})^{-1} = (1/c)\mathbf{A}^{-1}$
- For diagonal matrix $\mathbf{D}$, $\mathbf{D}^{-1}$ is a diagonal matrix with diagonal elements equal to the reciprocal of the corresponding diagonal elements of $\mathbf{D}$.

Inverse of a Square Matrix

Uses

Systems of Linear Equations

- Suppose you have the two simultaneous equations $2x_1 + x_2 = 5$, and $x_1 + x_2 = 3$.
- These two equations may be expressed in matrix algebra in the form

$$\mathbf{Ax} = \mathbf{b},$$

or

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix} \tag{1}$$

- We wish to solve for x_1 and x_2, which amounts to solving $\mathbf{Ax} = \mathbf{b}$ for $\mathbf{x}$

Inverse of a Square Matrix

Systems of Linear Equations

Solving the System

- To solve $\mathbf{Ax} = \mathbf{b}$ for $\mathbf{x}$, we premultiply both sides of the equation by $\mathbf{A}^{-1}$, obtaining

$$\mathbf{AA}^{-1}\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

- Since $\mathbf{AA}^{-1} = \mathbf{I}$, we end up with

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

Inverse of a Square Matrix

Systems of Linear Equations

Example

In the previous numerical example, we had

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \text{ and } \mathbf{b} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$$

It is easy to see that

$$\mathbf{A}^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

and so

$$\mathbf{x} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Trace of a Matrix

Trace of a Matrix

- The *trace* of a square matrix $\mathbf{A}$, $\mathrm{Tr}(\mathbf{A})$, is the sum of its diagonal elements
- The trace is often employed in matrix algebra to compute the sum of squares of all the elements of a matrix
- Verify for yourself that

$$\mathrm{Tr}(\mathbf{AA}') = \sum_i \sum_j a_{i,j}^2$$

Trace of a Matrix

We may verify that the trace has the following properties:

1. $\text{Tr}(\mathbf{A} + \mathbf{B}) = \text{Tr}(\mathbf{A}) + \text{Tr}(\mathbf{B})$
2. $\text{Tr}(\mathbf{A}) = \text{Tr}(\mathbf{A}')$
3. $\text{Tr}(c\mathbf{A}) = c\,\text{Tr}(\mathbf{A})$
4. $\text{Tr}(\mathbf{A}'\mathbf{B}) = \sum_i \sum_j a_{ij} b_{ij}$
5. $\text{Tr}(\mathbf{E}'\mathbf{E}) = \sum_i \sum_j e_{ij}^2$
6. The "cyclic permutation rule":

$$\text{Tr}(\mathbf{ABC}) = \text{Tr}(\mathbf{CAB}) = \text{Tr}(\mathbf{BCA})$$