Review of Mathematics for Social Scientists

SOC 512 & CSSS 505

authored by Laina Mercer and Jessica Godwin

Aaron Osgood-Zimmerman

Department of Statistics
University of Washington

January 14, 2021

Outline

- Matrix Algebra
  - Definitions, notation
  - Matrix Arithmetic
  - Determinants - existence of an inverse
  - Linear equations
  - Least Squares and Regression with matrices

Motivation

Matrix algebra provides concise notation and rules for manipulating matrices (arrays of numbers). Many types of data are naturally stored in arrays.

Matrix algebra will be important for computing linear regression estimates.

The applications of Linear Algebra are quite rich and crop up all over the place. Matrices (vector, and arrays) serve as a fundamental building block for it all.

Motivation

Example dataframe:

```
  region years       u5m     lower     upper
1    All 80-84 0.1691030 0.1573394 0.1815566
2    All 85-89 0.1603335 0.1490694 0.1722763
3    All 90-94 0.1208087 0.1079371 0.1349829
4  tanga 80-84 0.1810487 0.1369700 0.2354425
5  tanga 85-89 0.2230574 0.1677716 0.2902086
```

Caveat!

Although I will primarily be focusing on the mathematical/operational manipulation of matrices - since that is likely what you will primarily need to use - all of these operations have geometric interpretations that I find immensely helpful to understand.

If you're interetested in gaining a deeper intuition for matrices, vectors, and operations on them, I strongly encourage you to watch the 3blue1brown youtube video series: Essence of Linear Algebra. He covers way more than I have time for, and with much more beautiful graphics than I am capable of producing...

Essence of Linear Algebra Link

Definitions & Notation

What is a matrix?

A *matrix* is an array of number is a rectangular form.
Examples:

$$A = \begin{bmatrix} 1 & 2 & 6 & 4 \\ 5 & 8 & 12 & 8 \\ 4 & 3 & 2 & 1 \end{bmatrix} B = \begin{bmatrix} 4 & 3 & 2 \\ 1 & 2 & 4 \end{bmatrix}$$

where A is a 3×4 matrix and B is a 2×3 matrix. Note: matrix dimensions are always listed as rows × columns.

Definitions & Notation

What is a matrix?

In mathematical notation, a matrix is written

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}$$

Where x_{ij} is the value in the ith row and the jth column of matrix X.

We refer to row i of a matrix by $X_{i,*}$
We refer to column j of a matrix by $X_{*,j}$

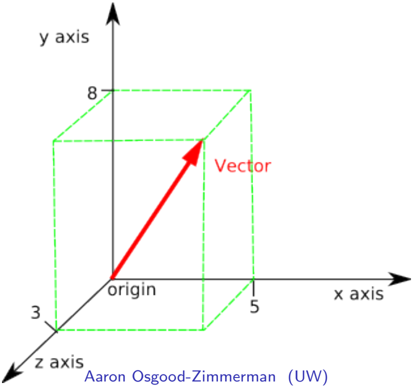
Definitions & Notation

Special Matrices

A *vector* is a matrix that has n rows and 1 column (or 1 row and n columns). Every row and column of a matrix is a vector.

Examples:

$$\begin{bmatrix} 1 & 2 & 6 & 4 \end{bmatrix} \text{ or } \begin{bmatrix} 5 \\ 8 \\ 3 \end{bmatrix}$$



A *square* matrix has the same number of rows and columns.
Example:

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

A *symmetric* matrix has elements such that $x_{ij} = x_{ji}$.
Example:

$$\begin{bmatrix} 1 & 4 & 5 \\ 4 & 2 & 3 \\ 5 & 3 & 7 \end{bmatrix}$$

A symmetric matrix must also be a square matrix.

A *diagonal* matrix is a matrix that is zero everywhere except on the diagonal. Where the diagonal is defined as all elements for which the row number is equal to the column number $\{(1,1),(2,2),(3,3),...\}$.

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

A special case of a diagonal matrix is the *identity* matrix. Its diagonal elements are all ones.

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

Clearly, the identity matrix (or any other diagonal matrix) is also symmetric.

Matrix Equality: Two matrices A, B are equal if and only if, for all elements, each $a_{ij} = b_{ij}$. (Note: this means they must have the same dimensions.)

Matrix Transpose: The *transpose* of a matrix is found by interchanging the corresponding rows and columns of a matrix. The first row becomes the first column, the second row becomes the second clump, etc. The dimensions are then switched and the element a_{ij} becomes the element a_{ji}. The transposed matrix is often denoted A^t (or A'). You can find the transpose of a matrix in `R` by using the `t()` function.

$$A = \begin{bmatrix} 1 & 2 & 6 \\ 3 & 5 & 9 \end{bmatrix} \quad A^t = \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 6 & 9 \end{bmatrix}$$

Matrix Arithmetic

Addition & Subtraction

Two matrices can be added or subtracted only if their dimensions are the same (both rows and columns). The corresponding elements are then added or subtracted.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}$$

Example:

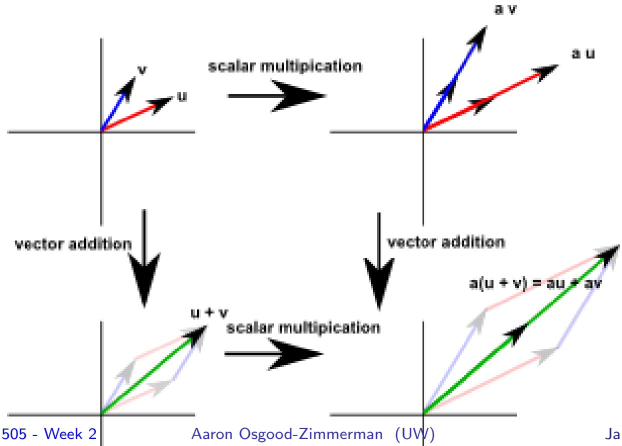
$$\begin{bmatrix} 1 & 2 & 6 \\ 3 & 5 & 9 \end{bmatrix} - \begin{bmatrix} 1 & 3 & 8 \\ 6 & 9 & 6 \end{bmatrix} = \begin{bmatrix} 0 & -1 & -2 \\ -3 & -4 & 3 \end{bmatrix}$$

Matrix Arithmetic

Multiplication

To multiply a matrix by a *scalar* (a constant value), multiply each element by that number. Example:

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 6 & 9 & 6 \end{bmatrix} \quad 3A = \begin{bmatrix} 3 & 9 & 24 \\ 18 & 27 & 18 \end{bmatrix}$$



There are a few different ways to multiply vectors. Let

$$u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

The *dot product* of u and v is

$$\begin{aligned} u \cdot v &= \sum_{i=1}^{3} u_i \times v_u \\ &= u_1 \times v_1 + u_2 \times v_2 + u_3 \times v_3 \end{aligned}$$

Matrix Arithmetic

Multiplication Examples

Two matrices can be multiplied only if the number of columns of the first matrix equals the number of rows of the second matrix (the inside numbers). The resulting matrix has dimensions: number of rows of first matrix by number of columns of second matrix (the outside numbers).

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

A is (2×3); B is (3×2).
$A \cdot B$ is $(2 \times 3 \cdot 3 \times 2)$ results in 2×2.
$B \cdot A$ is $(3 \times 2 \cdot 2 \times 3)$ results in 3×3.

We find each element $(ab)_{ij}$by summing the element-wise products of the ith row of A and the jth column of B. This is the *dot product* of the ith row of A and the jth column of B

Matrix Arithmetic

Multiplication Examples

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

We find each element $(ab)_{ij}$ by summing the element-wise products of the ith row of A and the jth column of B. This is also called the dot product of the ith row of A and the jth column of B

$$A \cdot B = \begin{bmatrix} (ab)_{11} & (ab)_{12} \\ (ab)_{21} & (ab)_{22} \end{bmatrix}$$

$$A \cdot B = \begin{bmatrix} A_{1,*} \cdot B_{*,1} & A_{1,*} \cdot B_{*,2} \\ A_{2,*} \cdot B_{*,1} & A_{2,*} \cdot B_{*,2} \end{bmatrix}$$

$$A \cdot B = \begin{bmatrix} a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + a_{13} \cdot b_{31} & a_{11} \cdot b_{12} + a_{12} \cdot b_{22} + a_{13} \cdot b_{32} \\ a_{21} \cdot b_{11} + a_{22} \cdot b_{21} + a_{23} \cdot b_{31} & a_{21} \cdot b_{12} + a_{22} \cdot b_{22} + a_{23} \cdot b_{32} \end{bmatrix}$$

Matrix Arithmetic

Multiplication Examples

Examples:

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 6 & 9 & 6 \\ 2 & 1 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 9 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}$$

$$A \cdot B = \begin{bmatrix} 1 \cdot 3 + 3 \cdot 2 + 8 \cdot 3 & 1 \cdot 9 + 3 \cdot 1 + 8 \cdot 2 \\ 6 \cdot 3 + 9 \cdot 2 + 6 \cdot 3 & 6 \cdot 9 + 9 \cdot 1 + 6 \cdot 2 \\ 2 \cdot 3 + 1 \cdot 2 + 3 \cdot 3 & 2 \cdot 9 + 1 \cdot 1 + 3 \cdot 2 \end{bmatrix} = \begin{bmatrix} 33 & 28 \\ 54 & 75 \\ 17 & 25 \end{bmatrix}$$

Note: $A \cdot B$ is not necessarily equal to $B \cdot A$. For matrix multiplication, order matters. In this case $B \cdot A$ cannot be computed as the dimensions are not compatible $(3 \times 2 \cdot 3 \times 3)$.

Matrix Inverse: The inverse of a number is its reciprocal; a number multiplied by its inverse equals 1. $(4 \cdot 1/4 = 1)$

The inverse of a matrix A is the matrix A^{-1} that satisfies $A \cdot A^{-1} = I$. Where I is the identity matrix (ones along the diagonal and the rest are zeros).

Remember that matrix multiplication is not just multiplying pairs of elements, so we can't just find the reciprocal of each element. So how do we find the inverse? How do we know if the inverse exists?

The *determinant* is a measure, in a sense, of the ‘volume’ of the matrix.
For a 2×2 matrix,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

the determinant is $D(A) = a \cdot d - b \cdot c$.

If the determinant is zero, there is no ‘volume’ to the matrix and no inverse exists. If the determinant is nonzero then the inverse exists.

Matrix Arithmetic

Determinant Example

$$A = \begin{bmatrix} 4 & 12 \\ 3 & 6 \end{bmatrix}$$

$D = 4 \cdot 6 - 12 \cdot 3 = -12$. Inverse exists.

$$A = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$$

$D = 2 \cdot 2 - 4 \cdot 1 = 0$. Inverse does not exist. Matrix is called *singular*.

Once we know the inverse exists, we can find it.
For a 2×2 matrix,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$A^{-1} = \frac{1}{D(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

where again $D(A) = a \cdot d - b \cdot c$. Example:

$$A = \begin{bmatrix} 4 & 12 \\ 3 & 6 \end{bmatrix}, \quad A^{-1} = \frac{1}{-12} \begin{bmatrix} 6 & -12 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} -1/2 & 1 \\ 1/4 & -1/3 \end{bmatrix}$$

Note: Finding the inverse for higher dimensions involves more comp;located formulas and is usually solved by a math software. (`solve()` in `R`)

Linear Equations

Let's go back to thinking about systems of two equations:

$$ax + by = g$$

$$cx + dy = f$$

Previously we solved this system by eliminating the y variable, solving for x, and then substituting back in for y.

No we can write this system in matrix notation:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \end{bmatrix}, \quad w = \begin{bmatrix} g \\ f \end{bmatrix}$$

Solving our system of equations is the same as solving for z in the matrix equation:

$$A \cdot z = w$$

Linear Equations
Examples

Solving our system of equations is the same as solving for z in the matrix equation:

$$A \cdot z = w$$

So how do we solve for z? First, left-multiply the equation by A^{-1}:

$$A^{-1} \cdot A \cdot z = A^{-1} \cdot w$$

By definition $A^{-1} \cdot A = I$. Thus,

$$I \cdot z = A^{-1} \cdot w \text{ or } z = A^{-1} \cdot w.$$

So we can find $z = (x, y)$, the solution to our system, by finding $z = A^{-1} \cdot w$.

Linear Equations

Examples

$$2x + y = 1$$

$$4x + 3y = 8$$

$$A = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \end{bmatrix}, \quad w = \begin{bmatrix} 1 \\ 8 \end{bmatrix}$$

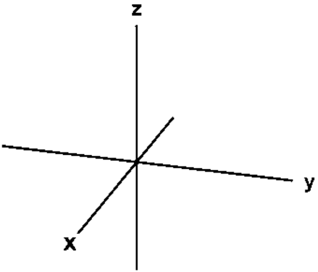
$$A^{-1} = \frac{1}{2 \cdot 3 - 4 \cdot 1} \begin{bmatrix} 3 & -1 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} 3/2 & -1/2 \\ -2 & 1 \end{bmatrix}$$

Geometry of matrices

When we perform matrix multiplication, we are projecting a vector (or each column vector of the second matrix) into a new space defined by the columns of the transformation matrix.

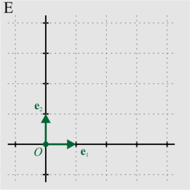
The identity matrix defines original base space, and multiplying by it does not change the space: $Iu = u$

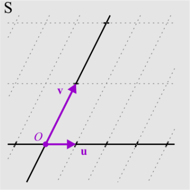
$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



Geometry of matrices

In 2D, a 'change of space', from our base using $e_1 = [1, 0]$ and $e_2 = [0, 1]$ to a space using $u = [1, 0]$ and $v = [1, 2]$ might look like this:





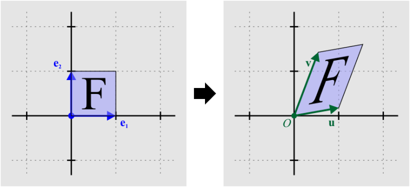
Consider the vector $x = [2, 2]$.

$$\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

Geometry of matrices

As mentioned, the determinant says something about the ‘volume’ of a matrix

It describes how much volume the unit square takes up *after* the transformation to the space represented by the matrix



And if the determinant is zero, this means that the matrix maps vectors to the origin. If multiple vectors all get mapped to the origin, and we want to find the inverse transformation - we can't! How do we know which place to untransform the origin to?

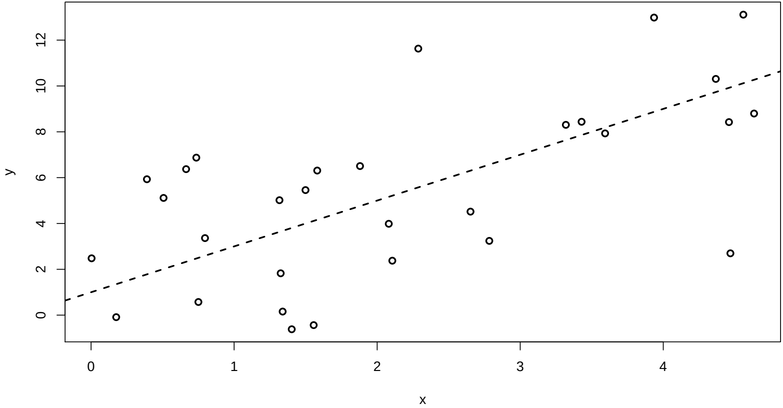
Least Squares

Often we have much more data that pairs of points. We may have a survey where we asked n people the same p questions. We can put that data in a matrix of dimensions $n \times p$, where each row is a person and each column is one of the asked questions.

Before we saw how to put a line through two points ($y = mx + b$). What if we wanted to put a line through many points?

Least Squares

Example



Least Squares

So how do we choose this line?
We can write the equation:

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_p x_p$$

where we have an intercept β_0 and then a slope β_i for each x_i (where $i = 1, .., p$). This equation has to describe the relationship as best it can for all n people we asked. In matrix notation:

$$\begin{bmatrix} y_1 \\ ... \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & ... & x_{1p} \\ 1 & ... & ... & ... \\ 1 & x_{n1} & ... & x_{np} \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ ... \\ \beta_p \end{bmatrix}$$

or $y = X\beta$.

Least Squares

Example

The linear least squares procedure finds the line that minimizes the squared distance between the points and the line.

$$\beta = \left(X^t \cdot X\right)^{-1} X^t y$$

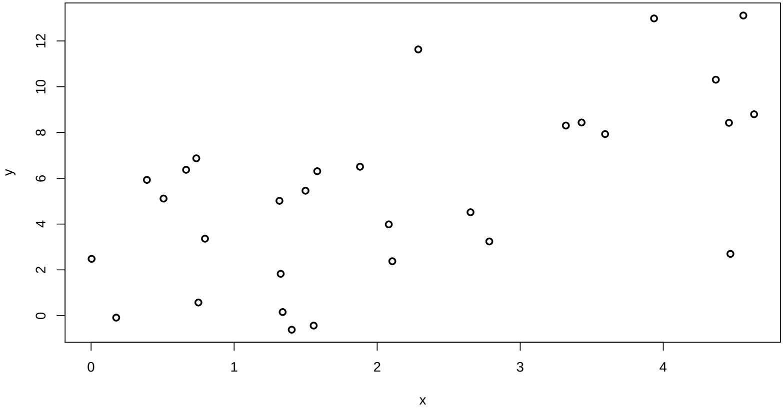
To see this note that y is $n \times 1$, X is $n \times (p+1)$, and β is $(p+1) \times 1$:

$$\begin{aligned}
y &= X\beta \\
X^t y &= X^t X\beta \\
\left(X^t X\right)^{-1} X^t y &= \left(X^t X\right)^{-1} X^t X\beta \\
\left(X^t X\right)^{-1} X^t y &= I \cdot \beta \\
\beta &= \left(X^t \cdot X\right)^{-1} X^t y
\end{aligned}$$

Least Squares
with matrices in R

```
set.seed(1985)
beta_0<-1
beta_1<-2
n<-30
x<-runif(n,0,5)
y<-rnorm(n,mean=beta_1*x+beta_0,sd=3)
plot(x,y)
```

Least Squares



Least Squares

with matrices in R

The function `solve()` provides the inverse and `t()` calculates the matrix transpose. The symbols $\% * \%$ tell R to do the matrix multiplication described in previous slides.

```
X.mat<-matrix(c(rep(1,n),x),ncol=2)
Beta.mat<-solve(t(X.mat)%*%(X.mat))%*%t(X.mat)%*%y
```

Here are the first few rows of our X model matrix and the β coefficients we have calculated.

```
X.mat[1:2,]
      [,1]     [,2]
[1,]    1 3.319174
[2,]    1 1.325468
Beta.mat
         [,1]
[1,] 1.960837
[2,] 1.590737
```

Least Squares

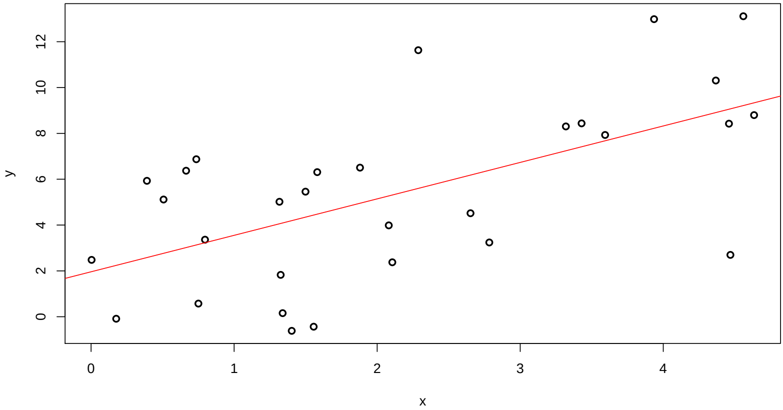


Figure: Our data with the fitted line $y = 1.59x + 196$.

Least Squares

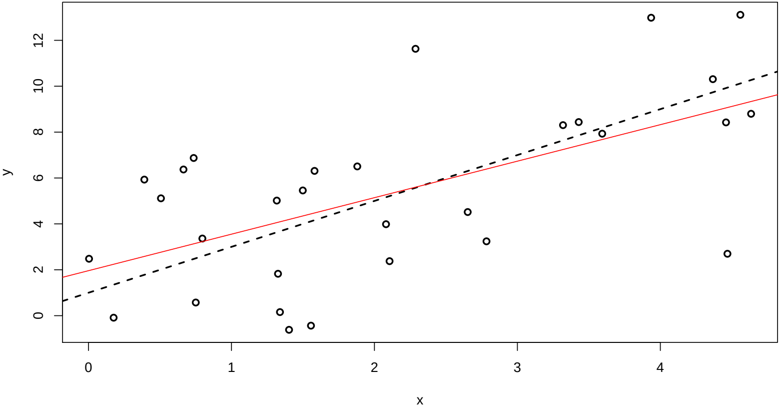


Figure: Our data with the fitted line $y = 1.96 + 1.59x$ and the true line $y = 1 + 2x$.

The End

Please bring questions to Jan 14 class, or to office hours.