Math for Political Scientists Workshop

Day Three: Linear Algebra I

Yang Yang

Department of Political Science
Penn State

08/21/2024

Review Time

Exponential and logarithm function

- Natural exponential function
- Natural logarithm function
- Their derivatives

Introduction to Probability

- Random event; outcome; sample space
- Simple event and compound event
- Independence; mutually exclusive; collectively exhaustive
- Conditional probability
- Joint probability; union probability
- Combination and permutation
- n factorial

Linear Algebra

What is Algebra?

Algebra is the mathematical study of **systems of equations**.

The simplest possible system of equation is the **linear equation**:

For example:

- Linear equations with two variables

$$x_1 + 2x_2 = 3 \tag{1}$$

$$2x_1 - 3x_2 = 8 \tag{2}$$

What is a linear equation?

Standard Form of a Linear Equation with Multiple Variables:

$$a_1x_1 + a_2x_2 + a_3x_3 + ... + a_nx_n + b = 0$$

where:

- The letters $a_1, ...a_n$ stand for coefficients, which are constants.
- The letters $x_1, ...x_n$ stand for variables.
- The letter b stands for a constant.

How to tell it?

- Simply, their graphs are straight lines.
- Formally, the highest power of each variable is 1.[1]

Eg., $x^2 - 9 = y$ is not a linear equation.

[1] Yet, do not confuse it with the OLS assumption of linearity in parameters.

From Linear Equations to Matrix

The general linear system of m equations in n unknowns can be written:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n &= b_2 \\
\vdots \qquad\quad \vdots \qquad\quad \dots \qquad\quad &\vdots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \dots + a_{mn}x_n &= b_m
\end{aligned}
$$

Three ways of solving such systems:

- Substitution
- Elimination of variables
- Matrix methods.

- It is a rectangular array of numbers, which arranges numbers in rows and columns.

$$\boldsymbol{A}_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \ldots & \ldots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

- It is very convenient and powerful method to represent and manipulate data/numbers in a structured way. (statistical inference and machine learning)

We can represent the linear systems with n variables and m equations:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n &= b_2 \\
\vdots \qquad\quad \vdots \qquad\quad \ldots \qquad\quad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots + a_{mn}x_n &= b_m
\end{aligned}$$

simply as:

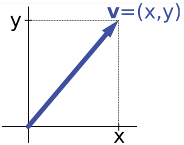
$$\boldsymbol{Ax} = \boldsymbol{b}$$

$$\boldsymbol{A}_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \ldots & \ldots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \boldsymbol{x}_{n\times 1} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \boldsymbol{b}_{m\times 1} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

- Scalars are real numbers, which are vector components.
- A vector is a list of scalars, arranged either in a row (referred to as a row vector) or a column (known as a column vector).

$$\begin{bmatrix}1, 2, 3\end{bmatrix}, \begin{bmatrix}1\\2\\3\end{bmatrix}$$

- A vector will either be a lowercase letter in a **bold** font, such as **x** or **a**, or a lowercase letter with an arrow over it, such as $\vec{v}$.
- The dimension of a vector is the number of components in the vector.



- The square matrix is a type of matrix whose number of rows(noted as m) and columns(noted as n) are the same(m=n).
- 3×3 Square Matrix

$$\boldsymbol{A}_{3\times 3} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{32} \end{bmatrix}$$

- The identity matrix is a type of square matrix in which all the elements of the principal diagonal are ones and all other elements are zeros.
- The identity matrix is often denoted by I_n or simply by I.

$$I_1 = [\,1\,],\ I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ \dots,\ I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

- It is the operation of adding two matrices by adding (subtracting) the corresponding entries together.(Entrywise operation)
- Two matrices must have an equal number of rows and columns (same dimensions).

$$\begin{bmatrix} 1 & 3 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 7 & 5 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1+0 & 3+0 \\ 1+7 & 0+5 \\ 1+2 & 2+1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 8 & 5 \\ 3 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 7 & 5 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1-0 & 3-0 \\ 1-7 & 0-5 \\ 1-2 & 2-1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ -6 & -5 \\ -1 & 1 \end{bmatrix}$$

Multiplying a matrix by a scalar:

Formally, **C**=r**A**

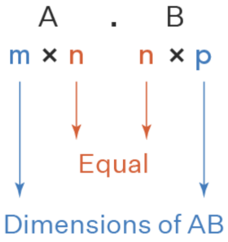
- It is defined as $c_{i,j} = r \times a_{i,j}$, where $1 \leq i \leq m$ and $1 \leq j \leq n$.

$$2 \cdot \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 2.1 & 2.2 & 2.3 \\ 2.4 & 2.5 & 2.6 \\ 2.7 & 2.8 & 2.9 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \\ 14 & 16 & 18 \end{bmatrix}$$

Matrix multiplication is between two matrices, or between a matrix and a vector, or a vector and a matrix.

- If A and B can only be multiplied, when the number of columns in A match the number of rows in B, i.e., $A_{m \times n} B_{n \times p} = C_{m \times p}$.



- Otherwise, they cannot be multiplied. Eg. $B_{n \times p} A_{m \times n}$ does not exist, unless p=m.

For $C_{m\times p}$=$A_{m\times n}B_{n\times p}$,

each element in C is

$$c_{i,j} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$$

$$\begin{aligned}
AB &= \begin{bmatrix} 1 & -2 \\ 0 & 5 \\ 4 & 3 \end{bmatrix} \times \begin{bmatrix} 3 & 1 & 4 \\ -1 & 2 & 5 \end{bmatrix} \\
&= \begin{bmatrix} (1\times 3)+(-2\times -1) & (1\times 1)+(-2\times 2) & (1\times 4)+(-2\times 5) \\ (0\times 3)+(\ \ 5\times -1) & (0\times 1)+(\ \ 5\times 2) & (0\times 4)+(\ \ 5\times 5) \\ (4\times 3)+(\ \ 3\times -1) & (4\times 1)+(\ \ 3\times 2) & (4\times 4)+(\ \ 3\times 5) \end{bmatrix} \\
&= \begin{bmatrix} 3+2 & 1-\ \ 4 & 4-10 \\ 0-5 & 0+10 & 0+25 \\ 12-3 & 4+\ \ 6 & 16+15 \end{bmatrix} \\
&= \begin{bmatrix} 5 & -3 & -6 \\ -5 & 10 & 25 \\ 9 & 10 & 31 \end{bmatrix}.
\end{aligned}$$

Exercise:

$$BA = \begin{bmatrix} 3 & 1 & 4 \\ -1 & 2 & 5 \end{bmatrix} \times \begin{bmatrix} 1 & -2 \\ 0 & 5 \\ 4 & 3 \end{bmatrix}$$

Answer:

$$
\begin{aligned}
BA &= \begin{bmatrix} 3 & 1 & 4 \\ -1 & 2 & 5 \end{bmatrix} \times \begin{bmatrix} 1 & -2 \\ 0 & 5 \\ 4 & 3 \end{bmatrix} \\
&= \begin{bmatrix} (\ \ 3 \times 1) + (1 \times 0) + (4 \times 4) & (\ \ 3 \times -2) + (1 \times 5) + (4 \times 3) \\ (-1 \times 1) + (2 \times 0) + (5 \times 4) & (-1 \times -2) + (2 \times 5) + (5 \times 3) \end{bmatrix} \\
&= \begin{bmatrix} 3 + 0 + 16 & -6 + \ \ 5 + 12 \\ -1 + 0 + 20 & 2 + 10 + 15 \end{bmatrix} \\
&= \begin{bmatrix} 19 & 11 \\ 19 & 27 \end{bmatrix}.
\end{aligned}
$$

- We just need to view a vector as a column matrix.
- Matrix-vector multiplication is defined only for the case when the number of columns in **A** equals the number of rows in x.[2]

The general formula for a matrix-vector product is:

$$\mathbf{Ax} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}.$$

[2] If **A** is an $m \times n$ matrix (i.e., with n columns), then the product **A**x is defined for $n \times 1$ vectors x.

Exercise:

$$\boldsymbol{A}_{2\times 3} = \begin{bmatrix} 1 & -1 & 2 \\ 0 & -3 & 1 \end{bmatrix}, x = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

Then, what is $\mathbf{A}x$?

Answer:

$$\mathbf{Ax} = \begin{bmatrix} 1 & -1 & 2 \\ 0 & -3 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \cdot 1 - 1 \cdot 1 + 0 \cdot 2 \\ 2 \cdot 0 - 1 \cdot 3 + 0 \cdot 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ -3 \end{bmatrix}.$$

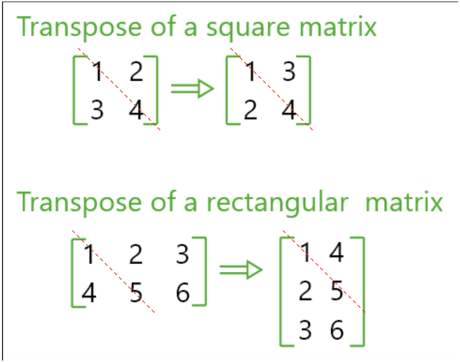
For any $n \times n$ square matrix A and the $n \times n$ identity matrix I:

$$AI = IA = A$$

Matrix Operations
Matrix Transposition

The transpose of a matrix is an operator which flips a matrix over its diagonal; that is, it switches the row and column indices of the matrix $\mathbf{A}$ by producing another matrix, often denoted by $\mathbf{A}^T$ or $\mathbf{A}'$.

- If $\mathbf{A}$ is a $m \times n$ matrix, then $\mathbf{A}^T$ is $n \times m$ matrix.



Exercise:

Find the transpose of the following matrix:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$$

Matrix Operations

Matrix Transposition

Answer:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \\ 4 & 8 & 12 \end{bmatrix}^{\mathrm{T}}$$

Let A be a $n \times n$ square matrix. The $n \times n$ matrix B is an **inverse** for A if $AB = BA = I_n$.

- The matrix B is uniquely determined by A, and is called the inverse of A, denoted by A^{-1}.
- Only a square matrix has an inverse.[3]

[3] Let A be an $k \times n$ matrix. The $n \times k$ matrix B is a **right inverse** for A if AB=I. The $n \times k$ matrix C is a **left inverse** for A if CA=I. But, neither B or C is an inverse for A.

Matrix in R

- Creating a matrix

```r
mat1<- matrix(c(2,4,6,8,10,12), nrow = 3,ncol =2, byrow=FALSE)
print(mat1)
```

```
##      [,1] [,2]
## [1,]    2    8
## [2,]    4   10
## [3,]    6   12
```

```r
mat2 <- matrix(c(2,4,6,8,10,12), nrow = 3,ncol=2, byrow=TRUE)
print(mat2)
```

```
##      [,1] [,2]
## [1,]    2    4
## [2,]    6    8
## [3,]   10   12
```

Matrix in R

- Creating a matrix

```r
# By default = by column = (byrow=FALSE)
mat1<- matrix(c(2,4,6,8,10,12), nrow = 3,ncol =2, byrow=FALSE)
print(mat1)
```

```
##      [,1] [,2]
## [1,]    2    8
## [2,]    4   10
## [3,]    6   12
```

```r
mat3 <- matrix(c(2,4,6,8,10,12), nrow = 3,ncol=2)
print(mat3)
```

```
##      [,1] [,2]
## [1,]    2    8
## [2,]    4   10
## [3,]    6   12
```

Matrix in R

Hadamard product

- When a matrix is multiplied with another matrix, the element-wise multiplication of two matrices take place. All the corresponding elements of both matrices will be multiplied under the condition that both matrices will be of the same dimension.

```
# R program for Hadamard product (Not matrix multiplication!)
# Creating matrices
m <- matrix(1:8, nrow=2)
n <- matrix(8:15, nrow=2)

# Multiplying matrices
print(m*n)
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    8   30   60   98
## [2,]   18   44   78  120
```

Matrix in R

Matrix Multiplication

- The Operator $\% * \%$ is used for matrix multiplication satisfying the condition that the number of columns in the first matrix is equal to the number of rows in second.

```r
# R program for matrix multiplication
# Creating matrices
m <- matrix(1:8, nrow=2)
n <- matrix(8:15, nrow=4)

# Multiplying matrices using %*% operator
print(m %*% n)
```

```
##      [,1] [,2]
## [1,]  162  226
## [2,]  200  280
```

OLS in Matrix Form

Suppose we are interested in a study where Y is our dependent variable and X is a list of k independent variables $\{X_1, X_2, ..., X_k\}$.

Our statistical model will essentially look something like the following:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ \vdots \\ Y_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} 1 & X_{11} & X_{21} & \dots & X_{k1} \\ 1 & X_{12} & X_{22} & \dots & X_{k2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & X_{1n} & X_{2n} & \dots & X_{kn} \end{bmatrix}_{n\times k} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_n \end{bmatrix}_{k\times 1} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \vdots \\ \epsilon_n \end{bmatrix}_{n\times 1}$$

This can be rewritten more simply as:

$$y = X\beta + \varepsilon$$

Our goal is to obtain estimates of the population parameters in the β vector.

Preview: OLS in Matrix Form

- Let X be an $n \times k$ matrix where we have observations on k independent variables for n units of analysis (such as, countries, individual politicians, or rebel groups). Since our model will usually contain a constant term, one of the columns in the X matrix will contain only ones. This column should be treated exactly the same as any other column in the X matrix.
- Let y be an $n \times 1$ vector of observations on the dependent variable.
- Let ε be an $n \times 1$ vector of disturbances or errors.
- Let β be an $k \times 1$ vector of unknown population parameters that we want to estimate.

Preview: OLS in Matrix Form

We want to solve for β in the following equation (assuming that ε is infinitesimally small and approaches 0):

$$y = X\beta + \varepsilon$$
$$\rightarrow y = X\beta$$

But we can't divide y directly by the matrix X, we need to multiply by their inverse instead. However, we can't invert a matrix that isn't square, so we start by multiply both sides of the equation by X^T.

$$X^T y = X^T X\beta$$

Now that we have a square matrix on the right side, we can multiply by the inverse of $X^T X$ to isolate β :

$$(X^T X)^{-1} X^T y = (X^T X)^{-1} X^T X\beta$$
$$(X^T X)^{-1} X^T y = I\beta$$
$$(X^T X)^{-1} X^T y = \beta$$

Preview: OLS in Matrix Form

An Important Note: The previous page might be helpful for you to understand the naive idea of getting the (multivariate) OLS estimates (assuming ε as zero) before you start PLSC 502 and 503.

However, once you learnt PLSC 502 and 503, you should know that we get the OLS estimate $\beta = (X^T X)^{-1} X^T y$ by minimizing the residual sum of squares in a linear regression model.

Wrap-up Time

- Scalar
- Vector
- Matrix
- Square Matrix
- Identity Matrix
- Matrix Addition (Subtraction)
- A scalar $\times$ a matrix
- A matrix $\times$ a matrix
- Matrix transposition
- Matrix inversion

More Exercises