PS 2010: 12. Matrix Algebra 1

Qing Chang

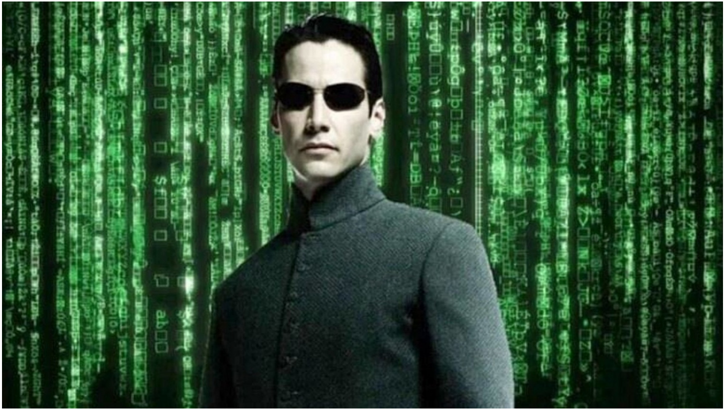
Department of Political Science
University of Pittsburgh

Fall 2023

- Homework 6 due next week

Agenda

- Matrix
- OLS estimator in matrix form



Definition: A rectangular array of numbers is called a matrix

Matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

- The horizontal arrays of a matrix are called its rows and the vertical arrays are called its columns.
- A matrix having m rows and n columns is said to have the order m $\times$ n
- A matrix having only one column is called a column vector; and a matrix with only one row is called a row vector.
- Matrix can also be represented as $[a_{ij}]$

Special Matrices

- Zero-matrix: a matrix in which each entry is zero

$$\mathbf{0}_{2\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

- Square matrix: a matrix having the number of rows equal to the number of columns
- Symmetric matrix: one in which $a_{ij} = a_{ji}$ for all i and j

$$\mathbf{A} = \begin{bmatrix} 1 & .5 & 2 \\ .5 & 1 & .75 \\ 2 & .75 & 1 \end{bmatrix}$$

- A diagonal matrix is one in which the only non-zero entries appear along the main diagonal from upper-left to lower-right

$$\Omega = \begin{bmatrix} 6.8 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & .0 & 2.5 \end{bmatrix}$$

Special Matrices

- A scalar matrix is one in which the same non-zero element appears along the main diagonal.

$$\mathbf{\Sigma} = \begin{bmatrix} .47 & 0 & 0 \\ 0 & .47 & 0 \\ 0 & .0 & .47 \end{bmatrix}$$

- The identity matrix I_n matrix is an $n \times n$ with $1'$ s along the main diagonal and 0's off the diagonal:

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- A triangular matrix is one that has only zeros above or below the main diagonal. If the zeros are below, the matrix is upper triangular.

$$\boldsymbol{\Psi} = \begin{bmatrix} .97 & .75 & .69 \\ 0 & .82 & .52 \\ 0 & 0 & .32 \end{bmatrix}$$

- A partitioned matrix is one that is divided into submatrices

$$\mathbf{Z} = \left[\begin{array}{cc|cc} 1.5 & .5 & 0 & 0 \\ .5 & 1.5 & 0 & 0 \\ \hline 0 & 0 & .5 & .75 \\ 0 & 0 & .75 & .5 \end{array}\right] = \left[\begin{array}{c|c} \mathbf{Z}_{11} & \mathbf{Z}_{12} \\ \hline \mathbf{Z}_{21} & \mathbf{Z}_{22} \end{array}\right]$$

Transpose of a Matrix

The transpose of an $m \times n(\text{ m by n})$ matrix is an $n \times m$ matrix whose (i, j) entry is the original matrix's (j, i) entry:

$$\mathbf{X} = \begin{bmatrix} 0 & 1 & 0 \\ .5 & 0 & .5 \\ 0 & 1 & 0 \end{bmatrix}, \mathbf{X}' = \begin{bmatrix} 0 & .5 & 0 \\ 1 & 0 & 1 \\ 0 & .5 & 0 \end{bmatrix}$$

Properties of Matrix Transposition

- Invertibility Property

$$(\mathbf{X}')' = \mathbf{X}$$

- Additive Property

$$(\mathbf{X} + \mathbf{Y})' = \mathbf{X}' + \mathbf{Y}'$$

- Multiplicative Property

$$(\mathbf{X}\mathbf{Y})' = \mathbf{Y}'\mathbf{X}'$$

- Scalar Multiplication Property

$$(c\mathbf{X})' = c\mathbf{X}'$$

- Inverse Transpose Property

$$(\mathbf{X}^{-1})' = (\mathbf{X}')^{-1}$$

- Symmetric Matrix Property

$$\mathbf{X}' = \mathbf{X}$$

Matrix Addition and Subtraction

- We define matrix addition and subtraction by the addition and subtraction of the corresponding elements of the matrices.
- Thus, we can only add and subtract matrices with the same dimensions.
- The addition (subtraction) of one $m \times n$ matrix by another $m \times n$ matrix produces an $m \times n$ matrix whose elements are the sums (differences) of the corresponding elements from the original matrices.

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

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

Multiplying a Matrix by a Scalar

Multiplying an $m \times n$ matrix by a scalar produces an $m \times n$ matrix whose elements are the products of the scalar and each element of the matrix.

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

Matrix Multiplication

- Let $\mathbf{A}$ be an $m \times n$ matrix and $\mathbf{B}$ be an $n \times p$ matrix. The matrix product $\mathbf{AB}$ is an $m \times p$ matrix whose j^{th} column is $\mathbf{Ab_j}$.
- Note that the inner dimensions of the two matrices must be equal (product conformable), and the outer dimensions determine the dimensions of the product.

$$\mathbf{Wy} = \begin{bmatrix} 0 & 1 & 0 \\ .5 & 0 & .5 \\ 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} 4 \\ 2 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \times 4 + 1 \times 2 + 0 \times 9 \\ .5 \times 4 + 0 \times 2 + .5 \times 9 \\ 0 \times 4 + 1 \times 2 + 0 \times 9 \end{bmatrix} = \begin{bmatrix} 2 \\ 6.5 \\ 2 \end{bmatrix}$$

Matrix Multiplication Properties

If the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ are conformable for multiplication, then

- Associative Property

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

- Distributive Property

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$

- Transpose of a Product

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$$

Linear System

Definition: A linear system of m equations in n unknowns $x_1, x_2, ..., x_n$ is a set of equations of the form

$$
\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\
&a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\
&\vdots \\
&a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m
\end{aligned}
$$

where for $1 \leq i \leq n$, and $1 \leq j \leq m$; a_{ij} , $b_i \in R$

Linear System

We rewrite the above equations in the form $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \text{ and } \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Rank of Matrix

Definition: The number of non-zero rows in the row reduced form of a matrix A is called the rank of A, denoted rank(A).

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

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

Inverse of a Matrix

- An $n \times n$ matrix $\mathbf{A}$ is invertible if there is an $n \times n$ matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{BA} = \mathbf{I_n}$. $\mathbf{B}$ is the inverse of $\mathbf{A}$, and, typically, we write $\mathbf{B}$ as $\mathbf{A}^{-1}$.
- $\left(A^{-1}\right)^{-1} = A$.
- $(AB)^{-1} = B^{-1}A^{-1}$.
- $\left(A^{'}\right)^{-1} = \left(A^{-1}\right)^{'}$
- See R example

Equivalent conditions for Invertibility

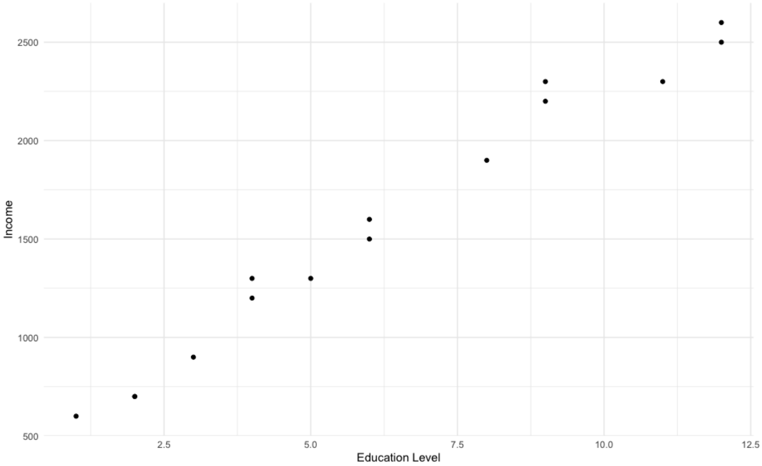
For a $n \times n$ square matrix A, the following statements are equivalent:

- A is full rank means rank (A) $=$ n.
- A is invertiable means A is full rank

Ordinary Least Squares (OLS)

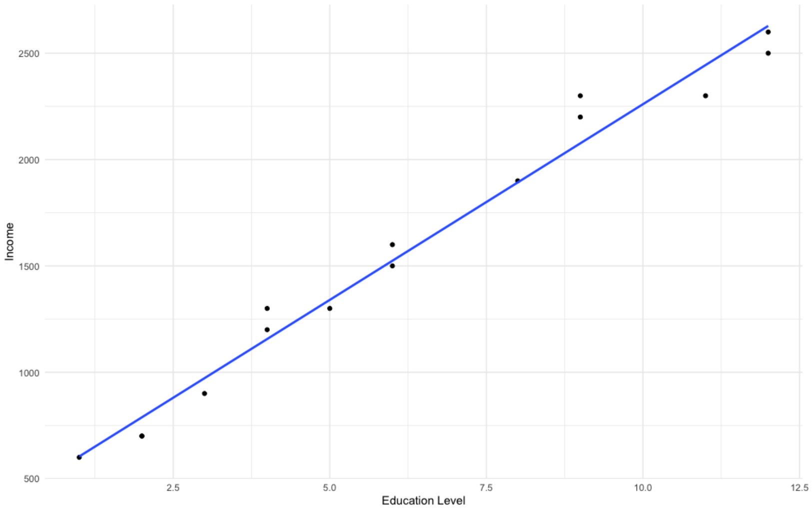
A Linear Model

Suppose we have data on education level and income and want to know their relationship.



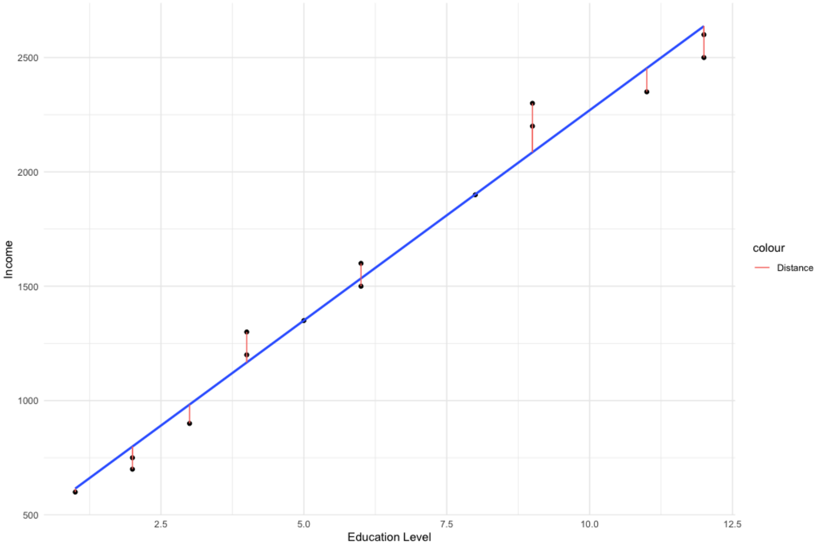
A Linear Model

The goal is to find a line that best describe the linear relationship



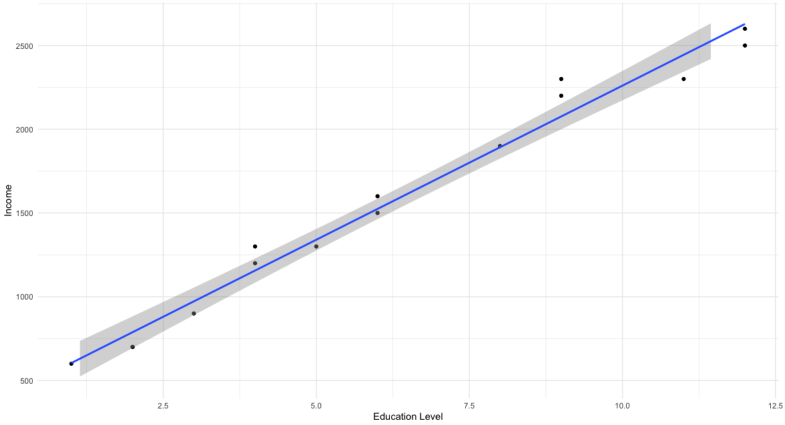
A Linear Model

We minimize the distance between line and data points



A Linear Model

We also want to describe the uncertainty



OLS

Let's formalize our idea a little bit

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ \vdots \\ Y_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix}_{n\times k} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_k \end{bmatrix}_{k\times 1} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ \vdots \\ e_n \end{bmatrix}_{n\times 1}$$

Which can be simplified as

$$y = X\beta + e$$

Our job is to find $\hat{\beta}$

$$y = X\hat{\beta} + \epsilon$$

OLS

- Our goal is to minimize the sum of distance
- One way is to minimize the sum of the squared distance
- That is $\epsilon' \epsilon$
- We have $\epsilon = y - X\hat{\beta}$

$$\begin{bmatrix} \epsilon_1 & \epsilon_2 & \dots & \epsilon_n \end{bmatrix}_{1\times n} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}_{n\times 1} = [\epsilon_1 \times \epsilon_1 + \epsilon_2 \times \epsilon_2 + \ldots + \epsilon_n \times \epsilon_n]_{1\times 1}$$

OLS

$$
\begin{aligned}
\epsilon'\epsilon &= (y - X\hat{\beta})'(y - X\hat{\beta}) \\
&= y'y - \hat{\beta}'X'y - y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} \\
&= y'y - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta}
\end{aligned}
$$

Where $\hat{\beta}'X'y$ and $y'X\hat{\beta}$ are scalars. Therefore, an inverse of a scalar is itself.

So, the goal is to find $\hat{\beta}$ that minimize the above equation

$$\epsilon'\epsilon = y'y - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta}$$

OLS

If a and b are $K \times 1$ vectors.

$$\frac{\partial a'b}{\partial b} = \frac{\partial b'a}{\partial b} = a$$

$$\frac{\partial b'Ab}{\partial b} = 2Ab = 2b'A$$

Therefore,

$$\frac{\partial 2\beta'X'y}{\partial b} = \frac{\partial 2\beta'(X'y)}{\partial b} = 2X'y$$

$$\frac{\partial \beta'X'X\beta}{\partial b} = \frac{\partial \beta'A\beta}{\partial b} = 2A\beta = 2X'X\beta$$

OLS Estimator

$$\frac{\partial \epsilon' \epsilon}{\partial \hat{\beta}} = -2X'y + 2X'X\hat{\beta} = 0$$

Then

$$X'X\hat{\beta} = X'y$$

Finally

$$\hat{\beta} = (X'X)^{-1}X'y$$

Assumptions

- We make no assumptions to derive OLS estimator
- However, we need make some assumptions to guarantee OLS is BLUE
- And allow us to do inference

Gauss-Markov Assumptions

- **Assumption 1:** Linear relationship between dependent and independent variables

$$y = X\beta + e$$

- **Assumption 2:** No perfect multicollinearity in independent variables

$$rank(X) = k$$

- **Assumption 3:** Strict exogeneity or zero conditional mean assumption

$$E(e \mid X) = 0$$

- **Assumption 4:** Spherical error variance
  - **Homoskedasticity:** $E(e'e \mid X) = \sigma^2$
  - **No autocorrelation:** $E(e_i e_j \mid X) = 0$ for i, j = 1,2,3,...; $i \neq j$

Unbiaseness

- The Gauss-Markov Assumptions guarantee OLS estimator is BLUE
- Unbiasedness:

$$\begin{aligned}
\hat{\beta} &= (X'X)^{-1}X'y \\
\hat{\beta} &= (X'X)^{-1}X'(X\beta + e) \\
\hat{\beta} &= (X'X)^{-1}X'X\beta + (X'X)^{-1}X'e \\
\hat{\beta} &= \beta + (X'X)^{-1}X'e \\
E[\hat{\beta}] &= E[\beta + (X'X)^{-1}X'e] \\
&= \beta + (X'X)^{-1}X'E[e]
\end{aligned}$$

- $E[e] = 0$ by assumption
- Therefore, we have $E[\hat{\beta}] = \beta$