Matrix Approaches to Simple Linear Regression Analysis

AU STAT-615

Emil Hvitfeldt

2021-02-24

Matrices

Matrix algebra is widely used in Mathematics and Statistics alike

It is more or less required to do multiple linear regression as it allows us to express large systems of equations and data in a compact way

Example

We have 2 variables of data, income and age

$$\begin{bmatrix} 16,000 \\ 16,000 \\ 16,000 \end{bmatrix} \begin{bmatrix} 23 \\ 47 \\ 35 \end{bmatrix}$$

Example

These columns can also be seen as being composed of rows (of observations)

$$\begin{bmatrix} 16,000 \\ 16,000 \\ 16,000 \end{bmatrix} \begin{bmatrix} 23 \\ 47 \\ 35 \end{bmatrix}$$

Notation

Usual notation for a matrix

$$\mathbf{A} = [a_{ij}] \qquad i = 1, 2; \quad j = 1, 2, 3$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

We call this a 2-by-3 matrix or more generally n-by-m matrix when the matrix has

- n rows
- m columns

Square Matrix

A matrix is called **square** if the number of rows equals the number of columns

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

$$\begin{bmatrix} 4 & 7 \\ 3 & 9 \end{bmatrix}$$

Vector

A vector can be thought of as a matrix with 1 column

A column vector

$$\mathbf{A} = \begin{bmatrix} 4 \\ 7 \\ 10 \end{bmatrix}$$

or a row vector

$$\mathbf{A} = \begin{bmatrix} 15 & 25 & 50 \end{bmatrix}$$

Transpose

A transpose of matrix $\mathbf{A}$ is denotes as $\mathbf{A}^T$ or $\mathbf{A}'$.

$$\mathbf{A} = \begin{bmatrix} 4 & 7 \\ 3 & 9 \\ 1 & 2 \end{bmatrix} \rightarrow \mathbf{A}^T = \begin{bmatrix} 4 & 3 & 1 \\ 7 & 9 & 2 \end{bmatrix}$$

Can be seen as flipping the row and column indices. Or flipping over the diagonal

Equality of matrices

Let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$

then we say that $\mathbf{A} = \mathbf{B}$ if and only if

$$a_{11} = b_{11}, \quad a_{12} = b_{12}, \quad a_{21} = b_{21}, \quad a_{22} = b_{22}$$

Regression Example

In regression analysis on of the basic matrices is $\mathbf{Y}$, that contains thee n observations on $\mathbf{Y}$

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

Regression Example

Another basic matrix we use in regression is matrix $\mathbf{X}$

In simplee linear regression this matrix is

$$\mathbf{X} = \begin{bmatrix} 1 & Y_1 \\ 1 & Y_2 \\ \vdots & \vdots \\ 1 & Y_n \end{bmatrix}$$

- columns of 1s

- n observations of predictor variables

Matrix addition & Subtraction

Let $\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}$

Then we have that

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 1+1 & 4+2 \\ 2+2 & 5+3 \\ 3+3 & 6+4 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 4 & 8 \\ 6 & 10 \end{bmatrix}$$

and

$$\mathbf{A} - \mathbf{B} = \begin{bmatrix} 1-1 & 4-2 \\ 2-2 & 5-3 \\ 3-3 & 6-4 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 0 & 2 \\ 0 & 2 \end{bmatrix}$$

Regression example

$$Y_i = E\{Y_i\} + \varepsilon_i \quad i = 1, \dots, n$$

This can be written in matrix form when

$$\mathbf{E}\{\mathbf{Y_i}\} = \begin{bmatrix} E\{Y_1\} \\ E\{Y_2\} \\ \vdots \\ E\{Y_n\} \end{bmatrix}, \qquad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Regression example

So we get that

$$\mathbf{Y} = \mathbf{E}\{\mathbf{Y_i}\} + \boldsymbol{\varepsilon}$$

can written for

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} E\{Y_1\} \\ E\{Y_2\} \\ \vdots \\ E\{Y_n\} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} = \begin{bmatrix} E\{Y_1\} + \varepsilon_1 \\ E\{Y_2\} + \varepsilon_2 \\ \vdots \\ E\{Y_n\} + \varepsilon_n \end{bmatrix}$$

Matrix Multiplication

Matrix by scalar

let $\mathbf{A} = \begin{bmatrix} 4 & 7 \\ 3 & 9 \end{bmatrix}$, then

$$4 \cdot \mathbf{A} = \begin{bmatrix} 4 \cdot 4 & 4 \cdot 7 \\ 4 \cdot 3 & 4 \cdot 9 \end{bmatrix} = \begin{bmatrix} 16 & 28 \\ 12 & 36 \end{bmatrix}$$

In general we have that for a scalar k and matrix $\mathbf{A}$

$$k \cdot \mathbf{A} = \mathbf{A} \cdot k$$

Matrix Multiplication

matrix by matrix

let $\mathbf{A} = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix}$ and let $\mathbf{A} = \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix}$

then

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix} = \begin{bmatrix} 2 \cdot 4 + 5 \cdot 5 & 2 \cdot 6 + 5 \cdot 8 \\ 4 \cdot 4 + 1 \cdot 5 & 4 \cdot 6 + 1 \cdot 8 \end{bmatrix} = \begin{bmatrix} 33 & 52 \\ 21 & 32 \end{bmatrix}$$

Matrix Multiplication

matrix by matrix

let $\mathbf{A} = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix}$ and let $\mathbf{A} = \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix}$

then

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix} = \begin{bmatrix} 2 \cdot 4 + 5 \cdot 5 & 2 \cdot 6 + 5 \cdot 8 \\ 4 \cdot 4 + 1 \cdot 5 & 4 \cdot 6 + 1 \cdot 8 \end{bmatrix} = \begin{bmatrix} 33 & 52 \\ 21 & 32 \end{bmatrix}$$

Matrix Multiplication

matrix by matrix

let $\mathbf{A} = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix}$ and let $\mathbf{A} = \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix}$

then

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix} = \begin{bmatrix} 2 \cdot 4 + 5 \cdot 5 & 2 \cdot 6 + 5 \cdot 8 \\ 4 \cdot 4 + 1 \cdot 5 & 4 \cdot 6 + 1 \cdot 8 \end{bmatrix} = \begin{bmatrix} 33 & 52 \\ 21 & 32 \end{bmatrix}$$

Matrix Multiplication

matrix by matrix

let $\mathbf{A} = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix}$ and let $\mathbf{A} = \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix}$

then

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix} = \begin{bmatrix} 2 \cdot 4 + 5 \cdot 5 & 2 \cdot 6 + 5 \cdot 8 \\ 4 \cdot 4 + 1 \cdot 5 & 4 \cdot 6 + 1 \cdot 8 \end{bmatrix} = \begin{bmatrix} 33 & 52 \\ 21 & 32 \end{bmatrix}$$

Matrix Multiplication

matrix by matrix

let $\mathbf{A} = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix}$ and let $\mathbf{A} = \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix}$

then

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix} = \begin{bmatrix} 2 \cdot 4 + 5 \cdot 5 & 2 \cdot 6 + 5 \cdot 8 \\ 4 \cdot 4 + 1 \cdot 5 & 4 \cdot 6 + 1 \cdot 8 \end{bmatrix} = \begin{bmatrix} 33 & 52 \\ 21 & 32 \end{bmatrix}$$

Matrix Multiplication?

Is $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$?

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} \cdot \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix} = \begin{bmatrix} 2 \cdot 4 + 5 \cdot 5 & 2 \cdot 6 + 5 \cdot 8 \\ 4 \cdot 4 + 1 \cdot 5 & 4 \cdot 6 + 1 \cdot 8 \end{bmatrix} = \begin{bmatrix} 33 & 52 \\ 21 & 32 \end{bmatrix}$$

and

$$\mathbf{B} \cdot \mathbf{A} = \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix} \cdot \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 4 \cdot 2 + 6 \cdot 4 & 4 \cdot 5 + 6 \cdot 1 \\ 5 \cdot 2 + 8 \cdot 4 & 5 \cdot 5 + 8 \cdot 1 \end{bmatrix} = \begin{bmatrix} 32 & 26 \\ 42 & 33 \end{bmatrix}$$

Matrix Multiplication

In order to multiply two matrices, the inner dimensions most agree

let

$$\mathbf{A}_{m \times n} \cdot \mathbf{B}_{n \times p}$$

then

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{C}_{m \times p}$$

Special Types of Matrices

Symmetric

If $\mathbf{A}^T = \mathbf{A}$ then $\mathbf{A}$ is symmetric

Example

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

Special Types of Matrices

Diagonal

A matrix that only have values in the diagonal

$$\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}$$

Special Types of Matrices

Identity

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

for any square matrix $\mathbf{A}$ we have

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

Special Types of Matrices

Scalar matrix

$$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

Linear Dependence

Example

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 5 & 1 \\ 2 & 2 & 10 & 6 \\ 3 & 4 & 15 & 1 \end{bmatrix}$$

Think of columns here as single vectors.

We observe that

$$\begin{bmatrix} 5 \\ 10 \\ 15 \end{bmatrix} = 5 \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Linear Dependence

since

$$\begin{bmatrix} 5 \\ 10 \\ 15 \end{bmatrix} = 5 \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

we say that the columns of $\mathbf{A}$ are linearly dependent

In other words, they contain redundant information

Vectors are linear dependent if one vector can be expressed as a linear combination of the others

Rank of a matrix

Is defined to be the maximum number of linearly independent columns in the matrix

For $\mathbf{A}$ on the last slide we have $\text{rank}(\mathbf{A}) = 3$

For $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ then $\text{rank}(\mathbf{C}) \leq \min(\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B}))$

Inverse of a matrix

For a number 6, the inverse is $\frac{1}{6}$ such that $\frac{1}{6} \cdot 6 = 1$

For square (invertible) matrices we have that

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

The invertible matrix theorem

The invertible matrix theorem [edit]

Let **A** be a square *n* by *n* matrix over a field *K* (e.g., the field **R** of real numbers). The following statements are equivalent (i.e., they are either all true or all false for any given matrix):[4]

A is invertible, that is, **A** has an inverse, is nonsingular, or is nondegenerate.

A is row-equivalent to the *n*-by-*n* identity matrix $\mathbf{I}_n$.

A is column-equivalent to the *n*-by-*n* identity matrix $\mathbf{I}_n$.

A has *n* pivot positions.

det $\mathbf{A} \neq 0$. In general, a square matrix over a commutative ring is invertible if and only if its determinant is a unit in that ring.

A has full rank; that is, rank $\mathbf{A} = n$.

The equation $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$.

The kernel of **A** is trivial, that is, it contains only the null vector as an element, $\ker(\mathbf{A}) = \{\mathbf{0}\}$.

The equation $\mathbf{Ax} = \mathbf{b}$ has exactly one solution for each **b** in K^n.

The columns of **A** are linearly independent.

The columns of **A** span K^n.

Col $\mathbf{A} = K^n$.

The columns of **A** form a basis of K^n.

The linear transformation mapping **x** to **Ax** is a bijection from K^n to K^n.

There is an *n*-by-*n* matrix **B** such that $\mathbf{AB} = \mathbf{I}_n = \mathbf{BA}$.

The transpose $\mathbf{A}^\mathrm{T}$ is an invertible matrix (hence rows of **A** are linearly independent, span K^n, and form a basis of K^n).

The number 0 is not an eigenvalue of **A**.

The matrix **A** can be expressed as a finite product of elementary matrices.

The matrix **A** has a left inverse (that is, there exists a **B** such that $\mathbf{BA} = \mathbf{I}$) *or* a right inverse (that is, there exists a **C** such that $\mathbf{AC} = \mathbf{I}$), in which case both left and right inverses exist and $\mathbf{B} = \mathbf{C} = \mathbf{A}^{-1}$.