Lecture 8:
Matrix Notation for Multiple Linear Regression
STAT 632, Spring 2020

Review of Matrices

We say that $\boldsymbol{X}$ is an $r \times c$ *matrix* if it is an array of numbers with r rows and c columns.

Ex:

$$\underset{(4\times 3)}{\boldsymbol{X}} = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 5 \\ 1 & 3 & 4 \\ 1 & 8 & 6 \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \end{pmatrix}$$

The element x_{ij} denotes the number in the i^{th} row and j^{th} column.

Review of Matrices

A *vector* is a one-column matrix.

Ex:

$$\underset{(4\times 1)}{\boldsymbol{y}} = \begin{pmatrix} 2 \\ 3 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

Notation: Bold face is used to denote matrices and vectors.

Review of Matrices

A *square matrix* has the same number of rows and columns, so $r = c$. A square matrix is *diagonal* if all elements off the main diagonal are 0.

Ex: $\boldsymbol{C}$ is a square matrix, and $\boldsymbol{D}$ is a diagonal matrix.

$$\underset{(3\times3)}{\boldsymbol{C}} = \begin{pmatrix} -5 & 1 & 3 \\ 1 & 2 & 6 \\ 3 & 6 & -4 \end{pmatrix} \qquad \underset{(3\times3)}{\boldsymbol{D}} = \begin{pmatrix} -5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -4 \end{pmatrix}$$

Review of Matrices

The *identity matrix* is a diagonal matrix with 1's on the diagonal.

Ex:

$$\boldsymbol{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Review of Matrices

The *transpose* of a matrix reverses rows and columns.

Ex:

$$\underset{(4\times3)}{\boldsymbol{X}} = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 5 \\ 1 & 3 & 4 \\ 1 & 8 & 6 \end{pmatrix} \qquad \underset{(3\times4)}{\boldsymbol{X}'} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 3 & 8 \\ 1 & 5 & 4 & 6 \end{pmatrix}$$

Review of Matrices

A square matrix $\boldsymbol{C}$ is *symmetric* if it equals its transpose, $\boldsymbol{C}' = \boldsymbol{C}$.

Ex:

$$\boldsymbol{C} = \begin{pmatrix} -5 & 1 & 3 \\ 1 & 2 & 6 \\ 3 & 6 & -4 \end{pmatrix}$$

Matrix Addition and Scalar Multiplication

$$\underset{(2\times 3)}{\boldsymbol{A}} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \qquad \underset{(2\times 3)}{\boldsymbol{B}} = \begin{pmatrix} -5 & 1 & 2 \\ 3 & 0 & -4 \end{pmatrix}$$

$$\boldsymbol{A} + \boldsymbol{B} =$$

$$\boldsymbol{A} - \boldsymbol{B} =$$

$$2\boldsymbol{A} =$$

Inner Product

The *inner product* (or dot product) of two vectors $\underset{(1\times n)}{\boldsymbol{a}'}$ and $\underset{(n\times 1)}{\boldsymbol{b}}$ is

$$\boldsymbol{a}'\boldsymbol{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i$$

Matrix Multiplication

For two matrices to be multiplied together, $\boldsymbol{AB}$, the number of columns of $\boldsymbol{A}$ must equal the number of rows of $\boldsymbol{B}$.

$$\underset{(r \times c)}{\boldsymbol{A}}\underset{(c \times q)}{\boldsymbol{B}} = \underset{(r \times q)}{\boldsymbol{AB}}$$

Ex:

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

Inverse of a Matrix

The *inverse* of a square matrix $\boldsymbol{C}$ is denoted $\boldsymbol{C}^{-1}$ and has the property

$$\boldsymbol{C}\boldsymbol{C}^{-1} = \boldsymbol{C}^{-1}\boldsymbol{C} = \boldsymbol{I}$$

Not all square matrices have an inverse. A square matrix that has an inverse is called *non-singular*; a square matrix without an inverse is called *singular*.

Ex:

$$\boldsymbol{D} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \qquad \boldsymbol{D}^{-1} = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1/4 \end{pmatrix}$$

Inverse of a Matrix

For a 2×2 matrix we can compute the inverse as

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \boldsymbol{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

If $ad - bc = 0$, then $\boldsymbol{A}$ is not invertible. Note that $ad - bc$ is called the *determinant* of $\boldsymbol{A}$, denoted $\det(\boldsymbol{A})$.

Inverse of a Matrix

Ex: Find the inverse of $\boldsymbol{A} = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$

Properties of Matrix Multiplication:

- $\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C}$ (associativity)
- $\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{AB}+\boldsymbol{AC}$ (left distributivity)
- $(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC}+\boldsymbol{BC}$ (right distributivity)
- $\boldsymbol{I}_r\boldsymbol{A} = \underset{(r\times c)}{\boldsymbol{A}} = \boldsymbol{A}\boldsymbol{I}_c$

Matrix multiplication is not commutative. So, $\boldsymbol{AB}$ is not necessarily equal to $\boldsymbol{BA}$.

For $n \times 1$ vectors $\boldsymbol{a}$ and $\boldsymbol{b}$:

$$(\boldsymbol{a}-\boldsymbol{b})'(\boldsymbol{a}-\boldsymbol{b}) = \boldsymbol{a}'\boldsymbol{a} + \boldsymbol{b}'\boldsymbol{b} - 2\boldsymbol{a}'\boldsymbol{b}$$

Properties of the Transpose:

- $(\boldsymbol{A}')' = \boldsymbol{A}$
- $(c\boldsymbol{A})' = c(\boldsymbol{A}')$, where c is a scalar
- $(\boldsymbol{A} + \boldsymbol{B})' = \boldsymbol{A}' + \boldsymbol{B}'$
- $(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}'$

Properties of Invertible Matrices:

- $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$
- $(c\boldsymbol{A})^{-1} = \frac{1}{c}\boldsymbol{A}^{-1}$, where $c \neq 0$ is a scalar.
- If $\boldsymbol{A}$ and $\boldsymbol{B}$ are both $n \times n$ invertible matrices, then $\boldsymbol{AB}$ is invertible, and $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.
- If $\boldsymbol{A}$ is invertible, then $\boldsymbol{A'}$ is invertible and $(\boldsymbol{A'})^{-1} = (\boldsymbol{A}^{-1})'$.

Matrix Notation for Multiple Linear Regression (MLR)

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & & \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

$$\boldsymbol{Y} = \boldsymbol{X}\beta + \boldsymbol{e}$$

- $\boldsymbol{Y}$ is an $n \times 1$ response vector
- $\boldsymbol{X}$ is an $n \times (p+1)$ design matrix for the predictor variables
- $\boldsymbol{\beta}$ is a $(p+1) \times 1$ vector of regression parameters
- $\boldsymbol{e}$ is an $n \times 1$ vector of random errors, assuming $e_i \sim N(0, \sigma^2)$

Matrix Notation for MLR

Let $\boldsymbol{x}_i'$ denote the i^{th} row of $\boldsymbol{X}$. Then

$$\underset{1\times(p+1)}{\boldsymbol{x}_i'} = \begin{pmatrix} 1 & x_{i1} & x_{i2} & \cdots & x_{ip} \end{pmatrix}$$

which allows us to write

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + e_i = \boldsymbol{x}_i' \boldsymbol{\beta} + e_i$$

Matrix Notation for MLR

Let $\hat{\boldsymbol{\beta}} = \begin{pmatrix}\hat{\beta}_0 & \hat{\beta}_1 & \cdots & \hat{\beta}_p\end{pmatrix}'$ denote the least squares estimates of the unknown regression parameters $\boldsymbol{\beta} = \begin{pmatrix}\beta_0 & \beta_1 & \cdots & \beta_p\end{pmatrix}'$.

It can be shown that $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$.

Matrix Notation for MLR

The vector of fitted (or predicted) values is given by

$$\hat{\boldsymbol{Y}} = \boldsymbol{X}\hat{\beta}$$

The vector of residuals is given by

$$\hat{\boldsymbol{e}} = \boldsymbol{Y} - \hat{\boldsymbol{Y}} = \boldsymbol{Y} - \boldsymbol{X}\hat{\beta}$$

$$\hat{\boldsymbol{e}} = \begin{pmatrix} \hat{e}_1 \\ \hat{e}_2 \\ \vdots \\ \hat{e}_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} - \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{pmatrix}$$