MATRICES (MATERIAL NOT EXAMINABLE)

An $r \times c$ **matrix** A is a rectangular arrangement of numbers with r rows and c columns;

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1c} \\ a_{21} & a_{22} & \cdots & a_{2c} \\ \vdots & \vdots & \vdots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rc} \end{bmatrix}$$

Some rules for manipulating matrices are given below:

- **Transpose:** the transpose operator $^{\mathsf{T}}$ means "flipping" a $r \times c$ matrix into a $c \times r$ matrix. That is

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1c} \\ a_{21} & a_{22} & \cdots & a_{2c} \\ \vdots & \vdots & \vdots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rc} \end{bmatrix} \qquad \Longleftrightarrow \qquad A^{\mathsf{T}} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{r1} \\ a_{12} & a_{22} & \cdots & a_{r2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1c} & a_{2c} & \cdots & a_{rc} \end{bmatrix}$$

  For example, if $r = 2$ and $c = 4$

$$A = \begin{bmatrix} 5 & -4 & 0 & 1 \\ 3 & 5 & -2 & 0 \end{bmatrix} \qquad \Longleftrightarrow \qquad A^{\mathsf{T}} = \begin{bmatrix} 5 & 3 \\ -4 & 5 \\ 0 & -2 \\ 1 & 0 \end{bmatrix}$$

  A square matrix A is termed **symmetric** if $A = A^{\mathsf{T}}$.

- **Matrix Multiplication:** If A and B are two matrices, where A is a $r_1 \times c$ matrix and B is a $c \times r_2$ matrix, then the product $A.B$ (also written AB) is an $r_1 \times r_2$ matrix, with (i, j)th element

$$\sum_{k=1}^{c} a_{ik} b_{kj} \qquad\qquad i = 1, \ldots, r_1, \; j = 1, \ldots, r_2.$$

  For example,

$$\begin{bmatrix} 5 & -4 & 0 & 1 \\ 3 & 5 & -2 & 0 \end{bmatrix} \begin{bmatrix} 3 & 3 & -3 \\ -1 & 2 & -2 \\ 0 & -2 & 0 \\ -5 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 5 & -6 \\ 4 & 23 & -19 \end{bmatrix}$$

  That is, for the first entry in the result matrix, we multiply the **first row** of the first matrix by the **first column** of the second matrix:

$$(5 \times 3) + (-4 \times -1) + (0 \times 0) + (1 \times -5) = 15 + 4 - 5 = 14$$

  Note that for matrix multiplication to work, we need the first matrix to have the same number of columns as the number of rows in the second matrix. If this holds, the matrices are termed **conformable**. In general, for rectangular matrices

$$A.B \neq B.A \qquad \text{and} \qquad A.B.C = A.(B.C) = (A.B).C$$

- **Matrix Identity :** A square $k \times k$ with ones along the main diagonal, and zeros elsewhere, is termed the **identity** matrix, and denoted I_k

$$I_k = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \qquad \text{so that} \qquad I_k.A = A \qquad \text{for any } k \times k \text{ matrix } A$$

- **Matrix Inversion :** A square $k \times k$ matrix A has an **inverse**, denoted A^{-1} if

$$A.A^{-1} = A^{-1}.A = I_k$$

Matrices in Linear Regression

- $n \times 1$ vector $\underset{\sim}{y} = [y_1, \ldots, y_n]^\intercal$
- $n \times 2$ matrix $\boldsymbol{X}$ given by

$$\boldsymbol{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix}^\intercal$$

- 2×1 Parameter estimate vector $\underset{\sim}{\widehat{\beta}} = \left[\widehat{\beta}_0, \widehat{\beta}_1\right]^\intercal$

It can be shown that

$$\underset{\sim}{\widehat{\beta}} = (\boldsymbol{X}^\intercal \boldsymbol{X})^{-1} \boldsymbol{X}^\intercal \underset{\sim}{y}$$

The other quantities of interest in statistical inference for the simple linear regression are also available in matrix form.

- SSE:

$$\text{SSE} = S(\, \underset{\sim}{\widehat{\beta}}\,) = (\underset{\sim}{y} - \boldsymbol{X}\underset{\sim}{\widehat{\beta}})^\intercal (\underset{\sim}{y} - \boldsymbol{X}\underset{\sim}{\widehat{\beta}})$$

- Residual error variance estimate, $\widehat{\sigma}^2$:

$$\widehat{\sigma}^2 = \frac{S(\, \underset{\sim}{\widehat{\beta}}\,)}{n-2} = \frac{1}{n-2} (\underset{\sim}{y} - \boldsymbol{X}\underset{\sim}{\widehat{\beta}})^\intercal (\underset{\sim}{y} - \boldsymbol{X}\underset{\sim}{\widehat{\beta}})$$

- Variance/Standard Errors of the Parameter estimates:

$$Var[\, \underset{\sim}{\widehat{\beta}}\,] = \widehat{\sigma}^2 (\boldsymbol{X}^\intercal \boldsymbol{X})^{-1}$$

  This is a 2×2 matrix, with diagonal entries equal to the squared estimated standard errors for $\widehat{\beta}_0$ and $\widehat{\beta}_1$, $s^2_{\widehat{\beta}_0}$ and $s^2_{\widehat{\beta}_1}$ respectively.
- Fitted-values :

$$\underset{\sim}{\widehat{y}} = \boldsymbol{X}\underset{\sim}{\widehat{\beta}} = \boldsymbol{X}(\boldsymbol{X}^\intercal \boldsymbol{X})^{-1} \boldsymbol{X}^\intercal \underset{\sim}{y} = \boldsymbol{H}\underset{\sim}{y}$$

  say, where $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}^\intercal \boldsymbol{X})^{-1} \boldsymbol{X}^\intercal$
- Residuals:

$$\widehat{e} = \underset{\sim}{y} - \underset{\sim}{\widehat{y}} = \underset{\sim}{y} - \boldsymbol{H}\underset{\sim}{y} = (I_n - \boldsymbol{H})\underset{\sim}{y}$$

- Prediction: if $\boldsymbol{x}_p = [1, x_p]^\intercal$, then the prediction is at the value x_p is

$$y_p = \boldsymbol{x}_p^\intercal \, \underset{\sim}{\widehat{\beta}}$$

  and the prediction error variances are

$$\begin{aligned} \text{Expected Value} &\;:\; \widehat{\sigma}^2 \boldsymbol{x}_p^\intercal (\boldsymbol{X}^\intercal \boldsymbol{X})^{-1} \boldsymbol{x}_p \\ \text{Individual Value} &\;:\; \widehat{\sigma}^2 (1 + \boldsymbol{x}_p^\intercal (\boldsymbol{X}^\intercal \boldsymbol{X})^{-1} \boldsymbol{x}_p) \end{aligned}$$