4.0 Matrix Notation and Literacy

4.1 Glossary

A matrix is a table of numbers arranged like a spread sheet into r rows and c columns

e.g.
$$\begin{bmatrix} 1 & 4 & 3 \\ 2 & 6 & 8 \end{bmatrix}$$

We denote a matrix with a "$\sim$" underscore:

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 4 & 3 \\ 2 & 6 & 8 \end{bmatrix}$$

The ij^{th} element of $\underset{\sim}{X}$ is the contents of the cell in the i^{th} row and j^{th} column

The dimension of a matrix is the number of rows and columns expressed as

$r \times c$

Often we use a short-hand notation

$$\underset{\sim}{X} = [x_{ij}] \qquad \begin{array}{l} i = 1, 2, \ldots, r \\ j = 1, 2, \ldots, c \end{array}$$

A square matrix has the same number of rows as columns

A vector is a matrix with one column.

e.g. $\underset{\sim}{y} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$

The transpose of a matrix is a version where the rows and column's have been interchanged. It is denoted by "T"

e.g.

$$\underset{\sim}{X}^T = \begin{bmatrix} 1 & 2 \\ 4 & 6 \\ 3 & 8 \end{bmatrix} \qquad \underset{\sim}{y}^T = \begin{bmatrix} 4 & 6 \end{bmatrix}$$

A matrix is symmetric if $x_{ij} = x_{ji} \quad \forall i \,\&\, j$

A symetric matrix is diagonal if $x_{ij} = 0 \quad \forall i \neq j$

The identity matrix of size n is the $n \times n$ diagonal matrix with $x_{ii} = 1 \quad i = 1, 2, \ldots, n$

We also define two special vectors

$$\underset{\sim}{1}_n = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \qquad \underset{\sim}{0}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Two matrices are equal if, and only if, they have the same dimension and all their elements are equal

Thinking about matrix - vector multiplication, etc

Let $\underset{\sim}{x} = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$ $\underset{\sim}{y} = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$

scalar - multiplication and scalar - addition

$3\underset{\sim}{x}$

$\underset{\sim}{x} + 4$

$2\underset{\sim}{x} + 2$

vector - addition : The elements of $\underset{\sim}{x} + \underset{\sim}{y}$ are $x_i + y_i$

$\underset{\sim}{x} + \underset{\sim}{y} =$

Some results

$$\underset{\sim}{x} + \underset{\sim}{y} = \underset{\sim}{y} + \underset{\sim}{x}$$

$$a\underset{\sim}{x} + b\underset{\sim}{x} = (a+b)\underset{\sim}{x}$$

$$b(a\underset{\sim}{x}) = (ab)\underset{\sim}{x}$$

$$(a\underset{\sim}{x})^T = a\underset{\sim}{x}^T$$

matrix - vector multiplication

$$\underset{\sim}{x}^T\underset{\sim}{y} = (\underset{\sim}{x}^T)\underset{\sim}{y} = \sum_{i=1}^{n} x_i y_i$$

$$\underset{\sim}{A} = \begin{pmatrix} 3 & -1 & 2 \\ 4 & 0 & 5 \end{pmatrix} \qquad \underset{\sim}{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

4.2 Matrix and vector multiplication

A scalar is a 1×1 matrix. That is an element of a matrix

Let a be a scalar

Consider multiplying each element of a matrix by a

e.g.

$$\begin{bmatrix} a1 & a4 & a3 \\ a2 & a6 & a8 \end{bmatrix}$$

we denote this process by $a\underset{\sim}{X}$

How can we define the multiplication of two matrices?

Let $\underset{\sim}{X}$ be a $r \times c$ matrix

and $\underset{\sim}{Y}$ be a $c \times s$ matrix

we define $\underset{\sim}{X}\,\underset{\sim}{Y}$ to be the matrix

$$\left[\sum_{k=1}^{c} x_{ik} y_{kj} \right] \qquad \begin{matrix} i = 1, 2, \ldots, r \\ j = 1, 2, \ldots, s \end{matrix}$$

For example

$$\underset{\sim}{X} = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} \qquad \underset{\sim}{Y} = \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix}$$

$$\underset{\sim}{Z} = \underset{\sim}{X}\,\underset{\sim}{Y} = \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix} =$$

$z_{11} =$

$z_{12} =$

$z_{21} =$

$z_{22} =$

Also $\quad \underset{\sim}{x} = \begin{bmatrix} 2 \\ 5 \end{bmatrix} \qquad \underset{\sim}{y} = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$

$\underset{\sim}{x}^T =$

$\underset{\sim}{x}^T \underset{\sim}{y} =$

4.2.1 Sums as vector notation

Consider

$$\sum_{k=1}^{n} x_k = x_1 + x_2 + \cdots + x_k$$

Let $\underset{\sim}{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix}$ and $\underset{\sim}{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$

then

$$\sum_{k=1}^{n} x_k = \underset{\sim}{1}^T \underset{\sim}{x}$$

$$\sum_{k=1}^{n} x_k^2 = \underset{\sim}{x}^T \underset{\sim}{x}$$

4.3 Matrix and vector random variables

Let $\underset{\sim}{y}$ be a vector of random variables

i.e. y_i $i = 1, 2, \ldots, n$ are random variables

For example $\underset{\sim}{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ where $y_1 \sim N(a, \sigma^2)$
$y_2 \sim N(b, \sigma^2)$
and y_1 is independent of y_2

By $\mathbb{E}(\underset{\sim}{y})$ we mean the vector of expectations

$$\begin{bmatrix} \mathbb{E}(y_1) \\ \mathbb{E}(y_2) \\ \vdots \\ \mathbb{E}(y_n) \end{bmatrix}$$

Note that $\mathbb{E}(a\underset{\sim}{y} + b) = a\mathbb{E}(\underset{\sim}{y}) + b$

Let σ_{ij}^2 be the covariance between y_i and y_j

i.e. $\sigma_{ij}^2 = \mathbb{E}\left[(y_i - \mathbb{E}(y_i))(y_j - \mathbb{E}(y_j)) \right]$

The matrix with elements σ_{ij}^2 is called the covariance matrix of $\underset{\sim}{y}$ and denoted by

For example, in the above case

$$\sigma_{11} = \text{variance of } y_1 = \sigma^2$$

$$\sigma_{12} = \text{covariance of } y_1 \text{ and } y_2 = 0$$

$$\sigma_{22} = \text{variance of } y_2 = \sigma^2$$

so $$\mathbb{V}(\underset{\sim}{y}) = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$$

Some useful formulas

$$\mathbb{V}(a\underset{\sim}{y}) = a^2\,\mathbb{V}(\underset{\sim}{y})$$

If $\mathbb{V}(y_i) = \sigma_i^2$ and the $y_1, \ldots, y_n$ are independent then

$$\mathbb{V}(\underset{\sim}{y}) = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sigma_n^2 \end{bmatrix}$$

If $\underset{\sim}{a}$ is a vector

$$\mathbb{E}(\underset{\sim}{a}^T\underset{\sim}{y}) = \underset{\sim}{a}^T(\mathbb{E}(\underset{\sim}{y}))$$

$$\mathbb{V}(\underset{\sim}{a}^T\underset{\sim}{y}) = \underset{\sim}{a}^T(\mathbb{V}(\underset{\sim}{y}))\underset{\sim}{a}$$

In general if

$$W = AY$$

$$\mathbb{E}(A) =$$

$$\mathbb{E}(W) =$$

$$\mathbb{V}[W] = \mathbb{V}[AY] = A\,\mathbb{V}[Y]A^T$$

4.4 Simple linear regression in matrix terms

Our basic data is

case		
1	y_1	x_1
2	y_2	x_2
$\vdots$	$\vdots$	$\vdots$
n	y_n	x_n

We denote

$$\underset{\sim}{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \qquad \underset{\sim}{X} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$$

Our model is

$$
\begin{aligned}
y_1 &= \beta_0 + \beta_1 x_1 + \epsilon_1 \\
y_2 &= \beta_0 + \beta_1 x_2 + \epsilon_2 \\
&\vdots \\
y_n &= \beta_0 + \beta_1 x_n + \epsilon_n
\end{aligned}
$$

Let $\underset{\sim}{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ $\quad \underset{\sim}{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$

Then the above system of equations is

$$\underset{\sim}{y} = \underset{\sim}{X} \underset{\sim}{\beta} + \underset{\sim}{\epsilon}$$

Also

$$\mathbb{E}(y_1) = \beta_0 + \beta_1 x_1$$
$$\vdots$$
$$\mathbb{E}(y_n) = \beta_0 + \beta_1 x_n$$

so

$$\mathbb{E}(\underset{\sim}{y}) = \underset{\sim}{X}\,\underset{\sim}{\beta}$$

$$\mathbb{E}(\underset{\sim}{\epsilon}) = \underset{\sim}{0}_n$$

$$\mathbb{V}(\underset{\sim}{\epsilon}) = \begin{bmatrix} \sigma^2 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \sigma^2 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \sigma^2 & 0 & & \vdots \\ \vdots & & 0 & \sigma^2 & & \vdots \\ \vdots & & & & \ddots & 0 \\ 0 & 0 & \cdots & & \cdots & \sigma^2 \end{bmatrix}$$

$$= \sigma^2\, \underset{\sim}{I}_n$$

The inverse of a matrix

Let A be a square matrix

Then B is an inverse of A

if

$$AB = \underset{\sim}{I}_n$$

and $$BA = \underset{\sim}{I}_n$$

Ex

If $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ then $B = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$

Not every square matrix has an inverse

If an inverse exists, A is invertible or non-singular

otherwise call A a singular matrix