Introduction to matrix notation

Highland Statistics Ltd
www.highstat.com

In this document elementary steps from matrix algebra and matrix notation are presented.

Vector

Define the vector $\boldsymbol{\beta}$ with elements 2, 5 and 7 as

$$\boldsymbol{\beta} = \begin{pmatrix} 2 \\ 5 \\ 7 \end{pmatrix}$$

This vector has 3 rows and 1 column and therefore the dimension of $\boldsymbol{x}$ is defined as 3 by 1. We call $\boldsymbol{\beta}$ a vector if it has multiple rows and only one column. If $\boldsymbol{\beta}$ has multiple rows *and* multiple columns we call it a matrix. We can multiply the elements of $\boldsymbol{\beta}$ with a scalar value (that is a number). For example, $2 \times \boldsymbol{\beta}$ is defined as

$$2 \times \boldsymbol{\beta} = 2 \times \begin{pmatrix} 2 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 2 \times 2 \\ 2 \times 5 \\ 2 \times 7 \end{pmatrix} = \begin{pmatrix} 4 \\ 10 \\ 14 \end{pmatrix}$$

Matrix

Define a matrix $\boldsymbol{X}$ with 2 rows and 3 columns, with the values 1, 2, 3 for the first row, and the second row is 4, 5 and 6:

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

The dimension of $\boldsymbol{X}$ is 2 by 3 (we have 2 rows and three columns). It is custom to put vectors in bold lower case and matrices in bold capital case, hence the $\boldsymbol{\beta}$ and the $\boldsymbol{X}$. Just as for the vector we can multiply a matrix with a scalar. For example $2 \times \boldsymbol{X}$ is given by

$$2 \times \boldsymbol{X} = 2 \times \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{pmatrix}$$

Multiplying a matrix with a vector

It is also possible to multiply a matrix with a vector, provided the dimensions match. With this we mean that a matrix of dimension a by b can be multiplied with a vector of dimension b by 1. For example $\boldsymbol{X} \times \boldsymbol{\beta}$ is defined as

$$\boldsymbol{X} \times \boldsymbol{\beta} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \times \begin{pmatrix} 2 \\ 5 \\ 7 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \times 2 + 2 \times 5 + 3 \times 7 \\ 4 \times 2 + 5 \times 5 + 6 \times 7 \end{pmatrix}$$

$$= \begin{pmatrix} 2 + 10 + 11 \\ 8 + 25 + 42 \end{pmatrix}$$

Using matrix algebra for multiple linear regression

Now that we have defined vectors and matrices we show how matrix notation can be used for multiple linear regression models. Let y_i be the number of birds at site i, and suppose we sampled at 100 sites; hence i = 1, …, 100. We model the number of birds as a function of temperature. Let x_i be the temperature at site i. A bivariate linear regression model for the bird abundance is given by:

$$y_i = \beta_1 + \beta_2 \times x_i + \varepsilon_i$$

The parameters β_1 and β_2 are the intercept and slope, respectively. We can write this equation for each site, resulting in:

$$\begin{aligned} y_1 &= \beta_1 + \beta_2 \times x_1 + \varepsilon_1 \\ y_2 &= \beta_1 + \beta_2 \times x_2 + \varepsilon_2 \\ y_3 &= \beta_1 + \beta_2 \times x_3 + \varepsilon_3 \\ &\vdots \\ y_{100} &= \beta_1 + \beta_2 \times x_{100} + \varepsilon_{100} \end{aligned}$$

This takes quite some space. Matrix notation can be used to simplify the notation. Define the vectors ***y, b*** and ***ε*** as:

$$\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{100} \end{pmatrix} \qquad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \qquad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{100} \end{pmatrix}$$

And define the matrix ***X*** as:

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_{100} \end{pmatrix}$$

Using matrix notation we can write the linear regression model as

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{100} \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_{100} \end{pmatrix} \times \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{100} \end{pmatrix}$$

$$\boldsymbol{y} = X \times \boldsymbol{\beta} + \varepsilon$$

In R we can obtain $\boldsymbol{X}$ with the `model.matrix` function and $\boldsymbol{\beta}$ with the `coef` function (or `fixef` for mixed models).