Introduction to matrix notation

Alain F. Zuur and Elena N. Ieno
Highland Statistics Ltd.
www.highstat.com

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

1.1 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 $\boldsymbol{\beta}$ is 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 number. For example, $2 \times \boldsymbol{\beta}$ is defined as

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

The new vector is also of dimension 3 by 1.

1.2 Matrix

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

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

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 a vector we can multiply a matrix with a number. For example $2 \times \boldsymbol{X}$ is given by

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

1.3 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

$$\boldsymbol{X} \times \boldsymbol{\beta} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \times \begin{pmatrix} 2 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 1\times2+2\times5+3\times7 \\ 4\times2+5\times5+6\times7 \end{pmatrix} = \begin{pmatrix} 33 \\ 75 \end{pmatrix}$$

The resulting matrix is of dimension 2 by 1.

1.4 Why all the fuss?

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 abundance of birds at site *i*, and suppose we sampled 100 sites. This means that the index *i* runs from 1 to 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

$$\begin{aligned} Y_1 &= \beta_1 + \beta_2 \times X_1 + \varepsilon_1 \\ Y_2 &= \beta_1 + \beta_2 \times X_2 + \varepsilon_2 \\ Y_3 &= \beta_1 + \beta_2 \times X_3 + \varepsilon_3 \\ Y_4 &= \beta_1 + \beta_2 \times X_4 + \varepsilon_4 \\ &\dots\dots \\ Y_{100} &= \beta_1 + \beta_2 \times X_{100} + \varepsilon_{100} \end{aligned}$$

This takes quite some space. Matrix notation can be used to simplify the notation. Define the vector **Y**, the matrix **X**, the vector $\boldsymbol{\beta}$ and the vector $\boldsymbol{\varepsilon}$ as

$$\boldsymbol{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{100} \end{pmatrix} \qquad \boldsymbol{X} = \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_{100} \end{pmatrix} \qquad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \qquad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{100} \end{pmatrix}$$

Using matrix notation we can write the linear regression model as

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{100} \end{pmatrix} = \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_{100} \end{pmatrix} \times \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{100} \end{pmatrix}$$

And this can be written as

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

If we use notation we can easily derive expressions for estimated parameters, standard errors, fitted values, and 95% confidence intervals for the fitted values in regression models, GLMs, GAMs, mixed models, etc.

In R we can obtain **X** with the `model.matrix` function and **β** with the `coef` function (or `fixef` for mixed models).