LINEAR ALGEBRA

These notes have two aims:

1) Introducing linear algebra (vectors and matrices) and
2) showing how to work with these concepts in R.

Vectors

A column vector is a list of numbers stacked on top of each other, e.g.

$$a = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

A row vector is a list of numbers written one after the other, e.g.

$$b = (2, 1, 3)$$

A general n–vector has the form

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

where the a_is are numbers, and this vector shall be written $a = (a_1, \dots, a_n)$.

Transpose of vectors

Transposing a vector means turning a column (row) vector into a row (column) vector. The transpose is denoted by "$^\top$".

$$\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}^\top = [1,3,2] \quad \text{og} \quad [1,3,2]^\top = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

- Hence transposing twice takes us back to where we started:

$$a = (a^\top)^\top$$

Sum of vectors

Let a and b be n–vectors. The sum $a + b$ is the n–vector

$$a + b = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix} = b + a$$

Matrices

An $r \times c$ matrix A (reads "an r times c matrix") is a table with r rows og c columns

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

Sum of matrices

$$\begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 9 \end{bmatrix} + \begin{bmatrix} 5 & 4 \\ 8 & 2 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} 6 & 6 \\ 11 & 10 \\ 5 & 16 \end{bmatrix}$$

Multiplication of a matrix and a vector

Let A be an $r \times c$ matrix and let b be a c-dimensional column vector. The product Ab is the $r \times 1$ matrix

$$Ab = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1c} \\ a_{21} & a_{22} & \dots & a_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rc} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_c \end{bmatrix} = \begin{bmatrix} a_{11}b_1 + a_{12}b_2 + \cdots + a_{1c}b_c \\ a_{21}b_1 + a_{22}b_2 + \cdots + a_{2c}b_c \\ \vdots \\ a_{r1}b_1 + a_{r2}b_2 + \cdots + a_{rc}b_c \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 9 \end{bmatrix} \begin{bmatrix} 5 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 \cdot 5 + 2 \cdot 8 \\ 3 \cdot 5 + 8 \cdot 8 \\ 2 \cdot 5 + 9 \cdot 8 \end{bmatrix} = \begin{bmatrix} 21 \\ 79 \\ 82 \end{bmatrix}$$

Inverse of matrices

In general, the inverse of an $n \times n$ matrix A is the matrix B (which is also $n \times n$) which when multiplied with A gives the identity matrix I. That is,

$$AB = BA = I.$$

$$\text{If } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \text{ then } A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$$(A^{-1})^{-1} = A,\ (AB)^{-1} = B^{-1}.A^{-1},\ (A^{-1})^{T} = (A^{T})^{-1};\ \ (ABC\ldots)^{-1} = \ldots\ldots C^{-1}B^{-1}A^{-1}.$$

It is easy find the inverse for a 2×2 matrix. When

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

then the inverse is

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

under the assumption that $ab - bc \neq 0$. The number $ab - bc$ is called the determinant of A, sometimes written $|A|$ or $\det(A)$. A matrix A has an inverse if and only if $|A| \neq 0$.

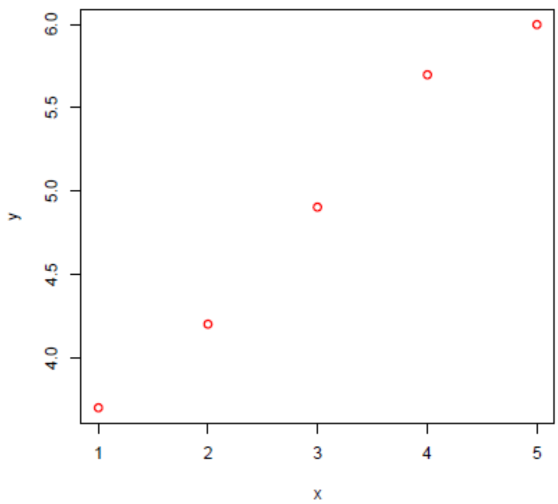
If $|A| = 0$, then A has no inverse, which happens if and only if the columns of A are linearly dependent.

Least squares

Consider the table of pairs (x_i, y_i) below.

x	1.00	2.00	3.00	4.00	5.00
y	3.70	4.20	4.90	5.70	6.00

A plot of y_i against x_i is shown in Figure 2.1.



The plot in Figure 2.1 suggests an approximately linear relationship between y and x, i.e.

$$y_i = \beta_0 + \beta_1 x_i \text{ for } i = 1, \ldots, 5$$

Writing this in matrix form gives

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_5 \end{bmatrix} \approx \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_5 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \mathbf{X}\beta$$

The first question is: Can we find a vector β such that $y = X\beta$? The answer is clearly no, because that would require the points to lie exactly on a straight line.

A more modest question is: Can we find a vector $\hat{\beta}$ such that $X\hat{\beta}$ is in a sense "as close to y as possible". The answer is yes. The task is to find $\hat{\beta}$ such that the length of the vector

$$e = y - X\beta$$

is as small as possible. The solution is

$$\hat{\beta} = (X^\top X)^{-1} X^\top y$$