

Matrix Algebra Concepts

Concepts from matrix algebra that the course participants should be familiar with:

- matrix
- vector (row and column)
- some special matrices: square, symmetric, diagonal, identity
- scalar
- matrix addition/subtraction/multiplication
- transpose of a matrix/vector
- inverse of a matrix
- trace of a matrix

We will not actually carry out any computation with matrix algebra by hand, but familiarity with these concepts will be useful. Writing out certain equations without matrix algebra is extremely cumbersome. With matrix algebra, we can write the same equations in a much shorter way.

Simple Regression with Matrix Algebra

Once you have reviewed the items on the list above, it is useful to see an example how these concepts can be put into practice. For this, here is an illustration of using matrix algebra to carry out the computations for simple linear regression. Suppose you have collected the following data on the midterm and final exam scores of 5 students:

Person	Midterm	Final
1	5	4
2	7	5
3	7	6
4	7	8
5	9	10

Now suppose you want to conduct a regression analysis with the final exam score as the dependent variable and the midterm exam score as the independent variable. Enter the data given above into a software package of your choice and conduct the regression analysis. For example, in SPSS, the output looks as follows:

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	-3.900	3.311		-1.178	.324
midterm	1.500	.465	.881	3.223	.048

a Dependent Variable: final

Therefore, we see that the estimate of the intercept is equal to -3.9 and the estimate of the slope is equal to 1.5.

Therefore:

predicted (average) final exam score = $-3.9 + 1.5$ midterm exam score.

The predicted (average) final exam score for students who score a 7 on the midterm exam is therefore:

$$-3.9 + 1.5(7) = 6.6.$$

Now we will obtain the same results with matrix algebra. First, define the following (column) vector and matrix:

$$\mathbf{y} = \begin{bmatrix} 4 \\ 5 \\ 6 \\ 8 \\ 10 \end{bmatrix} \quad \text{and} \quad \mathbf{X} = \begin{bmatrix} 1 & 5 \\ 1 & 7 \\ 1 & 7 \\ 1 & 7 \\ 1 & 9 \end{bmatrix}.$$

So \mathbf{y} is a column vector with the values of the dependent variable and \mathbf{X} is a matrix with only 1's in the first column and the values of the independent variable in the second column.

The (matrix) equation to obtain the parameter estimates is given by

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}$$

We will now go through this equation step by step. First, we take the transpose of \mathbf{X} (denoted by \mathbf{X}'), which is equal to

$$\mathbf{X}' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 5 & 7 & 7 & 7 & 9 \end{bmatrix}.$$

Next, we multiply \mathbf{X}' with \mathbf{X} :

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 5 & 7 & 7 & 7 & 9 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 1 & 7 \\ 1 & 7 \\ 1 & 7 \\ 1 & 9 \end{bmatrix} = \begin{bmatrix} 5 & 35 \\ 35 & 253 \end{bmatrix}.$$

Next, we need to take the inverse of $\mathbf{X}'\mathbf{X}$. Note that $\mathbf{X}'\mathbf{X}$ is a square matrix with two rows and two columns. In general, the inverse of a 2x2 matrix \mathbf{M} with elements

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

is equal to

$$\mathbf{M}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -b & a \end{bmatrix}.$$

Applying this to our example, we find:

$$\begin{aligned} (\mathbf{X}'\mathbf{X})^{-1} &= \frac{1}{5(253) - 35(35)} \begin{bmatrix} 253 & -35 \\ -35 & 5 \end{bmatrix} \\ &= \frac{1}{40} \begin{bmatrix} 253 & -35 \\ -35 & 5 \end{bmatrix} \\ &= \begin{bmatrix} 6.325 & -0.875 \\ -0.875 & 0.125 \end{bmatrix}. \end{aligned}$$

Next, we compute $\mathbf{X}'\mathbf{y}$:

$$\mathbf{X}' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 5 & 7 & 7 & 7 & 9 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \\ 8 \\ 10 \end{bmatrix} = \begin{bmatrix} 33 \\ 243 \end{bmatrix}.$$

So finally, we have all of the elements to compute

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}$$

namely

$$\mathbf{b} = \begin{bmatrix} 6.325 & -0.875 \\ -0.875 & 0.125 \end{bmatrix} \begin{bmatrix} 33 \\ 243 \end{bmatrix} = \begin{bmatrix} -3.9 \\ 1.5 \end{bmatrix}.$$

The first element in \mathbf{b} is the intercept estimate and the second element is the slope estimate.

Using matrix algebra, we can get the predicted value for students who score a 7 on the midterm exam by first defining the row vector:

$$\mathbf{x}_i = [1 \quad 6]$$

and then calculating:

$$\mathbf{x}_i \mathbf{b} = [1 \quad 6] \begin{bmatrix} -3.9 \\ 1.5 \end{bmatrix} = 6.6.$$