Quantitative Methods in Finance

Tutorial, Part 2:
Basic matrix algebra recapitulated.
Derivation of the least squares estimator and its properties.

Example 1: Derive the least squares estimator formula for estimation of regression coefficients and apply it in case of the following regression function:

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i,$$

where we know the following quantities for 10 observations:

$$\overline{y} = 2.4; \quad \overline{x}_2 = 0.4; \quad \overline{x}_3 = 0.8; \quad \sum y_i^2 = 184; \quad \sum x_{2i} x_{3i} = -2;$$
$$\sum x_{2i}^2 = 10; \quad \sum x_{3i}^2 = 12; \quad \sum y_i x_{2i} = 40; \quad \sum y_i x_{3i} = -2.$$

Derivation of the least squares estimator

The linear population regression model with two numerical explanatory variables, x_2 and x_3, is given by the following expression:

$$\begin{aligned} y_i &= \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i, \\ E(y_i \mid x_{2i}, x_{3i}) &= \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i}. \end{aligned} \tag{1}$$

As the parameter values from the population model above are not known, one should establish the procedure of their estimation from the sample data. Analogous to (1), one can form the following linear sample regression model:

$$\begin{aligned} y_i &= b_1 + b_2 x_{2i} + b_3 x_{3i} + e_i, \\ \hat{y}_i &= b_1 + b_2 x_{2i} + b_3 x_{3i}. \end{aligned} \tag{2}$$

Among the different possible criteria, we choose the method of ordinary least squares (OLS) in order to estimate the regression coefficients of the population regression function, such that the sum of squared residuals of the model (2) is minimal. Given that a residual is defined as the difference between the observed and the estimated value of dependent variable, the sum of its squares is a linear function of the estimators of regression coefficients:

$$\begin{aligned} e_i &= y_i - \hat{y}_i = y_i - b_1 - b_2 x_{2i} - b_3 x_{3i}, \\ \sum_{i=1}^{n} e_i^2 &= \sum_{i=1}^{n} (y_i - b_1 - b_2 x_{2i} - b_3 x_{3i})^2 = S(b_1, b_2, b_3). \end{aligned} \tag{3}$$

This function can be further broken down as:

$$
\begin{aligned}
S(b_1,b_2,b_3) = \sum_{i=1}^{n}(y_i - b_1 - b_2x_{2i} - b_3x_{3i})^2 &= \sum_{i=1}^{n} y_i^2 + nb_1^2 + b_2^2\sum_{i=1}^{n} x_{2i}^2 + \\
&+ b_3^2\sum_{i=1}^{n} x_{3i}^2 - 2b_1\sum_{i=1}^{n} y_i - 2b_2\sum_{i=1}^{n} x_{2i}y_i - 2b_3\sum_{i=1}^{n} x_{3i}y_i + \\
&+ 2b_1b_2\sum_{i=1}^{n} x_{2i} + 2b_1b_3\sum_{i=1}^{n} x_{3i} + 2b_2b_3\sum_{i=1}^{n} x_{2i}x_{3i},
\end{aligned} \tag{4}
$$

and then partially differentiated with respect to b_j:

$$
\begin{aligned}
\frac{\partial S}{\partial b_1} &= 2nb_1 + 2b_2\sum_{i=1}^{n} x_{2i} + 2b_3\sum_{i=1}^{n} x_{3i} - 2\sum_{i=1}^{n} y_i \ , \\
\frac{\partial S}{\partial b_2} &= 2b_1\sum_{i=1}^{n} x_{2i} + 2b_2\sum_{i=1}^{n} x_{2i}^2 + 2b_3\sum_{i=1}^{n} x_{2i}x_{3i} - 2\sum_{i=1}^{n} x_{2i}y_i \ , \\
\frac{\partial S}{\partial b_3} &= 2b_1\sum_{i=1}^{n} x_{3i} + 2b_2\sum_{i=1}^{n} x_{2i}x_{3i} + 2b_3\sum_{i=1}^{n} x_{3i}^2 - 2\sum_{i=1}^{n} x_{3i}y_i \ .
\end{aligned} \tag{5}
$$

If we equate the partial derivatives with zero and divide them with two, we get the following system of normal equations:

$$
\begin{aligned}
b_1n + b_2\sum_{i=1}^{n} x_{2i} + b_3\sum_{i=1}^{n} x_{3i} &= \sum_{i=1}^{n} y_i, \\
b_1\sum_{i=1}^{n} x_{2i} + b_2\sum_{i=1}^{n} x_{2i}^2 + b_3\sum_{i=1}^{n} x_{2i}x_{3i} &= \sum_{i=1}^{n} x_{2i}y_i, \\
b_1\sum_{i=1}^{n} x_{3i} + b_2\sum_{i=1}^{n} x_{2i}x_{3i} + b_3\sum_{i=1}^{n} x_{3i}^2 &= \sum_{i=1}^{n} x_{3i}y_i,
\end{aligned} \tag{6}
$$

which can be rewritten as:

$$
\begin{bmatrix}
n & \sum_{i=1}^{n} x_{2i} & \sum_{i=1}^{n} x_{3i} \\
\sum_{i=1}^{n} x_{2i} & \sum_{i=1}^{n} x_{2i}^2 & \sum_{i=1}^{n} x_{2i}x_{3i} \\
\sum_{i=1}^{n} x_{3i} & \sum_{i=1}^{n} x_{2i}x_{3i} & \sum_{i=1}^{n} x_{3i}^2
\end{bmatrix}
\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}
=
\begin{bmatrix}
\sum_{i=1}^{n} y_i \\
\sum_{i=1}^{n} x_{2i}y_i \\
\sum_{i=1}^{n} x_{3i}y_i
\end{bmatrix}. \tag{7}
$$

After some definitions in matrix algebra:

$$\mathbf{X}=\begin{bmatrix}1 & x_{21} & x_{31}\\ 1 & x_{22} & x_{32}\\ \vdots & \vdots & \vdots\\ 1 & x_{2n} & x_{3n}\end{bmatrix},\quad \mathbf{X}^{\mathrm{T}}=\begin{bmatrix}1 & 1 & \dots & 1\\ x_{21} & x_{22} & \dots & x_{2n}\\ x_{31} & x_{32} & \dots & x_{3n}\end{bmatrix},\quad \mathbf{b}=\begin{bmatrix}b_1\\ b_2\\ b_3\end{bmatrix},\quad \mathbf{y}=\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix}, \tag{8}$$

we can establish that the following holds:

$$\mathbf{X}^{\mathrm{T}}\mathbf{X}=\begin{bmatrix}n & \sum_{i=1}^{n}x_{2i} & \sum_{i=1}^{n}x_{3i}\\ \sum_{i=1}^{n}x_{2i} & \sum_{i=1}^{n}x_{2i}^2 & \sum_{i=1}^{n}x_{2i}x_{3i}\\ \sum_{i=1}^{n}x_{3i} & \sum_{i=1}^{n}x_{2i}x_{3i} & \sum_{i=1}^{n}x_{3i}^2\end{bmatrix} \quad \text{in} \quad \mathbf{X}^{\mathrm{T}}\mathbf{y}=\begin{bmatrix}\sum_{i=1}^{n}y_i\\ \sum_{i=1}^{n}x_{2i}y_i\\ \sum_{i=1}^{n}x_{3i}y_i\end{bmatrix}. \tag{9}$$

This means that the system of normal equations (6) or alternatively (7) can be rewritten in the following matrix form:

$$\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{b}=\mathbf{X}^{\mathrm{T}}\mathbf{y}. \tag{10}$$

The solution of the system of linear equations is obtained by multiplying both sides of the above matrix equation by the inverse matrix $(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}$:

$$(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}(\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{b}=(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y} \quad \Rightarrow \quad \mathbf{b}=(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}. \tag{11}$$

■

Example 2: Based on sample data for 10 observations we estimated a linear population regression function of the form $y_i=\beta_1+\beta_2 x_{2i}+\beta_3 x_{3i}+u_i$, and obtained the following fitted values and the following residuals:

i	1	2	3	4	5	6	7	8	9	10
$\hat{y}_i$	5.5	5	8	5.5	6.5	a	6	4	5.5	4.25
e_i	0.7	–2	0.5	1	–1	1	–1.5	0.5	b	1.3

Calculate the missing values, denoted by *a* and *b*. Do this by applying the properties of the least squares estimator.

Finally, prove the validity of the four basic properties of the least squares estimator.

■