Econometrics 2

OLS – A Complete Numerical Example

1 Setup and Data

Suppose we observe $n = 4$ individuals. For each person i we record:

- Y_i: wage (hundreds of GEL)
- X_{i1}: years of education
- We include a constant: $X_{i0} = 1$ for all i

Individual i	**Education** X_{i1}	**Wage** Y_i
1	6	4
2	8	5
3	10	8
4	12	9

Table 1: Observed data ($n = 4$)

The model is $Y_i = \beta_0 + \beta_1 X_{i1} + e_i$, or in matrix form: $\mathbf{Y} = \mathbf{X}\beta + \mathbf{e}$.

2 The Matrices Y and X

Stack all observations into:

$$\mathbf{Y} = \begin{bmatrix} 4 \\ 5 \\ 8 \\ 9 \end{bmatrix}_{4\times 1}, \qquad \mathbf{X} = \begin{bmatrix} 1 & 6 \\ 1 & 8 \\ 1 & 10 \\ 1 & 12 \end{bmatrix}_{4\times 2}, \qquad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}_{2\times 1}$$

We have $k = 2$ parameters (intercept plus one slope coefficient).

3 Deriving the OLS Estimator

The OLS estimator is $\widehat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$. We compute each piece in turn.

3.1 Computing X′X

$\mathbf{X}'\mathbf{X}$ is a 2×2 symmetric matrix:

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 6 & 8 & 10 & 12 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 1 & 8 \\ 1 & 10 \\ 1 & 12 \end{bmatrix} = \begin{bmatrix} \sum 1 & \sum X_{i1} \\ \sum X_{i1} & \sum X_{i1}^2 \end{bmatrix}$$

Evaluating each entry:

- $(1,1)$: $\sum 1 = 4$
- $(1,2)$: $\sum X_{i1} = 6 + 8 + 10 + 12 = 36$
- $(2,2)$: $\sum X_{i1}^2 = 36 + 64 + 100 + 144 = 344$

Therefore:

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 4 & 36 \\ 36 & 344 \end{bmatrix} \tag{1}$$

The determinant is $\det(\mathbf{X}'\mathbf{X}) = 4 \times 344 - 36 \times 36 = 1376 - 1296 = 80$, and the inverse is:

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{80} \begin{bmatrix} 344 & -36 \\ -36 & 4 \end{bmatrix} \tag{2}$$

3.2 Computing $\mathbf{X}'\mathbf{Y}$

$\mathbf{X}'\mathbf{Y}$ is a 2×1 vector:

$$\mathbf{X}'\mathbf{Y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 6 & 8 & 10 & 12 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_{i1} Y_i \end{bmatrix}$$

Evaluating each entry:

- Entry 1: $\sum Y_i = 4 + 5 + 8 + 9 = 26$
- Entry 2: $\sum X_{i1} Y_i = 6(4) + 8(5) + 10(8) + 12(9) = 24 + 40 + 80 + 108 = 252$

Therefore:

$$\mathbf{X}'\mathbf{Y} = \begin{bmatrix} 26 \\ 252 \end{bmatrix} \tag{3}$$

3.3 Computing $\hat{\beta}$

Multiplying (2) and (3):

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \frac{1}{80} \begin{bmatrix} 344 & -36 \\ -36 & 4 \end{bmatrix} \begin{bmatrix} 26 \\ 252 \end{bmatrix}$$

$$= \frac{1}{80} \begin{bmatrix} 344(26) - 36(252) \\ -36(26) + 4(252) \end{bmatrix} = \frac{1}{80} \begin{bmatrix} 8{,}944 - 9{,}072 \\ -936 + 1{,}008 \end{bmatrix} = \frac{1}{80} \begin{bmatrix} -128 \\ 72 \end{bmatrix}$$

$$\hat{\beta} = \begin{bmatrix} -1.6 \\ 0.9 \end{bmatrix} \tag{4}$$

Economic interpretation: each additional year of education is associated with a 90 GEL increase in monthly wages, holding everything else fixed. The intercept $\hat{\beta}_0 = -1.6$ is a mechanical anchor – zero years of education is far outside the sample support and carries no direct interpretation.

4 Predicted Values

The vector of fitted values is $\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta}$:

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta} = \begin{bmatrix} 1 & 6 \\ 1 & 8 \\ 1 & 10 \\ 1 & 12 \end{bmatrix} \begin{bmatrix} -1.6 \\ 0.9 \end{bmatrix} = \begin{bmatrix} -1.6 + 0.9(6) \\ -1.6 + 0.9(8) \\ -1.6 + 0.9(10) \\ -1.6 + 0.9(12) \end{bmatrix} = \begin{bmatrix} 3.8 \\ 5.6 \\ 7.4 \\ 9.2 \end{bmatrix}$$

Each $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{i1}$ is the point on the fitted regression line at X_{i1}.

5 Residuals

5.1 Computing $\hat{\mathbf{e}}$

The residual vector is $\hat{\mathbf{e}} = \mathbf{Y} - \hat{\mathbf{Y}}$:

$$\hat{\mathbf{e}} = \mathbf{Y} - \hat{\mathbf{Y}} = \begin{bmatrix} 4 \\ 5 \\ 8 \\ 9 \end{bmatrix} - \begin{bmatrix} 3.8 \\ 5.6 \\ 7.4 \\ 9.2 \end{bmatrix} = \begin{bmatrix} 0.2 \\ -0.6 \\ 0.6 \\ -0.2 \end{bmatrix}$$

i	Y_i	$\hat{Y}_i$	$\hat{e}_i$	$\hat{e}_i^2$
1	4	3.8	0.2	0.04
2	5	5.6	−0.6	0.36
3	8	7.4	0.6	0.36
4	9	9.2	−0.2	0.04
Sum	26	26	0	0.80

Table 2: Predicted values and residuals

5.2 OLS Orthogonality Conditions

A fundamental property of OLS is that residuals are orthogonal to every column of $\mathbf{X}$:

$$\mathbf{X}'\hat{\mathbf{e}} = \mathbf{0}$$

This gives two conditions we can verify directly:

- $\sum \hat{e}_i = 0.2 - 0.6 + 0.6 - 0.2 = 0$
- $\sum X_{i1}\hat{e}_i = 6(0.2) + 8(-0.6) + 10(0.6) + 12(-0.2) = 1.2 - 4.8 + 6.0 - 2.4 = 0$

These are not assumptions – they are algebraic consequences of the OLS minimization, and hold in every sample.

6 Estimating the Error Variance

6.1 The Residual Sum of Squares

From Table 2:

$$\text{RSS} = \hat{\mathbf{e}}'\hat{\mathbf{e}} = \sum_{i=1}^{4} \hat{e}_i^2 = 0.80 \tag{5}$$

6.2 The Unbiased Estimator $\hat{\sigma}^2$

Estimating $k = 2$ parameters uses up 2 degrees of freedom, leaving $n - k = 4 - 2 = 2$ for estimation of σ^2:

$$\hat{\sigma}^2 = \frac{\text{RSS}}{n-k} = \frac{0.80}{2} = 0.40 \tag{6}$$

Why $n-k$ and not n? The OLS residuals $\hat{e}_i$ satisfy the k orthogonality conditions $\mathbf{X}'\hat{\mathbf{e}} = \mathbf{0}$, which imposes k linear constraints on them. Dividing by $n-k$ corrects for this, making $\hat{\sigma}^2$ an unbiased estimator of σ^2. Dividing by n instead would systematically underestimate σ^2.

7 Variance of the OLS Estimator

7.1 The Formula

Under homoskedasticity ($E[e_i^2 \mid X_i] = \sigma^2$), Theorem 2 gives:

$$\mathbf{V}_{\hat{\beta}} = \text{var}[\hat{\beta} \mid \mathbf{X}] = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

We estimate this by replacing σ^2 with $\hat{\sigma}^2$:

$$\widehat{\mathbf{V}}_{\hat{\beta}} = \hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1} \tag{7}$$

7.2 Computing $\widehat{\mathbf{V}}_{\hat{\beta}}$

Substituting $\hat{\sigma}^2 = 0.40$ and (2) into (7):

$$\widehat{\mathbf{V}}_{\hat{\beta}} = 0.40 \times \frac{1}{80}\begin{bmatrix} 344 & -36 \\ -36 & 4 \end{bmatrix} = \frac{0.40}{80}\begin{bmatrix} 344 & -36 \\ -36 & 4 \end{bmatrix} = \begin{bmatrix} 1.72 & -0.18 \\ -0.18 & 0.02 \end{bmatrix}$$

The diagonal elements are the sampling variances of each coefficient estimator.

7.3 Standard Errors

The standard errors are the square roots of the diagonal elements:

$$\widehat{\text{se}}(\hat{\beta}_0) = \sqrt{1.72} \approx 1.311, \qquad \widehat{\text{se}}(\hat{\beta}_1) = \sqrt{0.02} \approx 0.141$$

8 Hypothesis Testing

8.1 t-Statistic

Under Assumption 4 (normality of errors), the t-statistic for $H_0 : \beta_j = 0$ is:

$$t_j = \frac{\hat{\beta}_j}{\widehat{\text{se}}(\hat{\beta}_j)} \sim t_{n-k} = t_2$$

For the slope coefficient (education has no effect on wages):

$$t_1 = \frac{\hat{\beta}_1}{\widehat{\text{se}}(\hat{\beta}_1)} = \frac{0.9}{0.141} \approx 6.36$$

The critical value for a two-sided test at $\alpha = 0.05$ is $t_{2,\,0.975} \approx 4.303$. Since $|t_1| = 6.36 > 4.303$, we **reject** H_0 – education is a statistically significant predictor of wages.

8.2 Summary of the Full OLS Pipeline

Quantity	**Result**
$\mathbf{X}'\mathbf{X}$	$\begin{bmatrix} 4 & 36 \\ 36 & 344 \end{bmatrix}$, $\det = 80$
$(\mathbf{X}'\mathbf{X})^{-1}$	$\frac{1}{80}\begin{bmatrix} 344 & -36 \\ -36 & 4 \end{bmatrix}$
$\mathbf{X}'\mathbf{Y}$	$(26,\ 252)'$
$\hat{\beta}$	$(-1.6,\ 0.9)'$
$\hat{\mathbf{Y}}$	$(3.8,\ 5.6,\ 7.4,\ 9.2)'$
$\hat{\mathbf{e}}$	$(0.2,\ -0.6,\ 0.6,\ -0.2)'$
RSS	0.80
$\hat{\sigma}^2 = \text{RSS}/(n-k)$	$0.80/2 = 0.40$
$\widehat{\mathbf{V}}_{\hat{\beta}} = \hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}$	$\begin{bmatrix} 1.72 & -0.18 \\ -0.18 & 0.02 \end{bmatrix}$
t-statistic for β_1	$0.9/0.141 \approx 6.36$ – reject H_0 at 5%

Table 3: The complete OLS pipeline