Econ 203B: Single Equation Models

Solutions for Winter 2005 Problem Set 2

Michael Powell
Department of Economics, UCLA
February 12th, 2006

1. The CNLR model applies to

$$E\left(Y|\, X_1, X_2\right) = X_1\beta_1 + X_2\beta_2$$

A sample of size $n = 102$ gives

$$\begin{pmatrix} \hat\beta_1 \\ \hat\beta_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \end{pmatrix};\ (X'X) = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix};\ \hat\varepsilon'\hat\varepsilon = 80$$

Let $\theta = \beta_1 - \beta_2$. Test at the 5% significance level the null hypothesis that $\theta = 1$.

Solution Since this model satisfies the CNLR assumptions, we can use the test statistic:

$$t_0 = \frac{\Gamma\hat\beta - \gamma_0}{\left(\frac{\hat\varepsilon'\hat\varepsilon}{n-k}\Gamma\left(X'X\right)^{-1}\Gamma'\right)^{\frac{1}{2}}} \sim t\left(n-k\right)$$

Here, we want to test the hypothesis $H_0 : \Gamma\beta = \gamma_0$ where

$$\Gamma = \begin{bmatrix} 1 & -1 \end{bmatrix},\ \gamma_0 = 1$$

It remains just to fill in the missing values for the t statistic.

$$\left(X'X\right)^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}^{-1} = \frac{1}{6-1}\begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{pmatrix}$$

$$\Gamma\left(X'X\right)^{-1}\Gamma' = \begin{bmatrix} 1 & -1 \end{bmatrix}\begin{bmatrix} \frac{3}{5} & -\frac{1}{5} \\ -\frac{1}{5} & \frac{2}{5} \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{7}{5}$$

Also, $n = 102$, $k = 2$, and finally,

$$\frac{\hat\varepsilon'\hat\varepsilon}{n-k} = \frac{80}{100} = \frac{4}{5}$$

Therefore,

$$t_0 = \frac{5-2-1}{\sqrt{\frac{4}{5}\frac{7.}{5}}} = 1.8898$$

The critical value for this test is $c^*_{0.025,t(100)} = 1.9839$. Since $t_0 \leq c^*_{0.025,t(100)}$, we fail to reject the null hypothesis that $\theta = 1$.

2. The CNLR model applies to

$$E\left(Y|\, X_1, X_2\right) = X_1\beta_1 + X_2\beta_2$$

with $\sigma^2 = V\left(Y_i|\, X_1, X_2\right) = 2$ and

$$X'X = \begin{pmatrix} 5 & 2 \\ 2 & 4 \end{pmatrix}$$

The sample produces

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

a. Construct a 95% confidence interval for $\theta = \beta_1 + \beta_2$.

Solution Since the CNLR assumptions hold and we know σ^2, we know that

$$z = \frac{\Gamma\hat{\beta} - \Gamma\beta}{\sqrt{\sigma^2 \Gamma (X'X)^{-1} \Gamma'}} \sim N(0,1)$$

It follows that

$$\begin{aligned} 0.95 &= \Pr\left[\left|\frac{\Gamma\hat{\beta} - \Gamma\beta}{\sqrt{\sigma^2 \Gamma (X'X)^{-1} \Gamma'}}\right| \leq 1.96\right] \\ &= \Pr\left[-1.96\sqrt{\sigma^2 \Gamma (X'X)^{-1} \Gamma'} \leq \Gamma\hat{\beta} - \Gamma\beta \leq 1.96\sqrt{\sigma^2 \Gamma (X'X)^{-1} \Gamma'}\right] \\ &= \Pr\left[\Gamma\hat{\beta} - 1.96\sqrt{\sigma^2 \Gamma (X'X)^{-1} \Gamma'} \leq \Gamma\beta \leq \Gamma\hat{\beta} + 1.96\sqrt{\sigma^2 \Gamma (X'X)^{-1} \Gamma'}\right] \end{aligned}$$

Filling in the missing values for this particular sample, we have: $\Gamma = \begin{bmatrix} 1 & 1 \end{bmatrix}$,

$$(X'X)^{-1} = \begin{pmatrix} 5 & 2 \\ 2 & 4 \end{pmatrix}^{-1} = \frac{1}{20-4}\begin{pmatrix} 4 & -2 \\ -2 & 5 \end{pmatrix} = \begin{pmatrix} \frac{4}{16} & -\frac{2}{16} \\ -\frac{2}{16} & \frac{5}{16} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & -\frac{1}{8} \\ -\frac{1}{8} & \frac{5}{16} \end{pmatrix}$$

And

$$\Gamma (X'X)^{-1} \Gamma' = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & -\frac{1}{8} \\ -\frac{1}{8} & \frac{5}{16} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{5}{16}$$

This gives us the 95% confidence interval for $\beta_1 + \beta_2$:

$$\begin{aligned} &\left(\Gamma\hat{\beta} - 1.96\sqrt{\sigma^2 \Gamma (X'X)^{-1} \Gamma'}, \Gamma\hat{\beta} + 1.96\sqrt{\sigma^2 \Gamma (X'X)^{-1} \Gamma'}\right) \\ =& \left(5 - 1.96\sqrt{2 \cdot \frac{5}{16}}, 5 + 1.96\sqrt{2 \cdot \frac{5}{16}}\right) \\ =& (3.4505, 6.5495) \end{aligned}$$

b. Construct a 90% confidence region for the pair (β_1, β_2).

Under the CNLR assumptions, we know that

$$W = \left(\Gamma\hat{\beta} - \Gamma\beta\right)' \left(\sigma^2 \Gamma (X'X)^{-1} \Gamma'\right)^{-1} \left(\Gamma\hat{\beta} - \Gamma\beta\right) \sim \chi^2(2)$$

The 10% critical value for this distribution is $c^*_{0.10,\chi^2(2)} = 4.61$. The 90% confidence region is therefore defined by:

$$0.90 = \Pr\left(\left(\Gamma\hat{\beta} - \Gamma\beta\right)' \left(\sigma^2 \Gamma (X'X)^{-1} \Gamma'\right)^{-1} \left(\Gamma\hat{\beta} - \Gamma\beta\right) \leq 4.61\right)$$

Here, we have that $\Gamma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, and

$$\sigma^2 \Gamma (X'X)^{-1} \Gamma' = 2I_2 \begin{bmatrix} \frac{1}{4} & -\frac{1}{8} \\ -\frac{1}{8} & \frac{5}{16} \end{bmatrix} I_2' = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{5}{16} \end{bmatrix}$$

Taking the inverse:

$$\left(\sigma^2\Gamma\left(X'X\right)^{-1}\Gamma'\right)^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{5}{16} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{10}{3} & \frac{8}{3} \\ \frac{8}{3} & \frac{16}{3} \end{bmatrix}$$

This gives us:

$$\begin{aligned} & \left(\Gamma\hat{\beta} - \Gamma\beta\right)' \left(\sigma^2\Gamma\left(X'X\right)^{-1}\Gamma'\right)^{-1} \left(\Gamma\hat{\beta} - \Gamma\beta\right) \\ = & \left(\begin{bmatrix} 3 \\ 2 \end{bmatrix} - \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}\right)' \begin{bmatrix} \frac{10}{3} & \frac{8}{3} \\ \frac{8}{3} & \frac{16}{3} \end{bmatrix} \left(\begin{bmatrix} 3 \\ 2 \end{bmatrix} - \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}\right) \\ = & \frac{1}{3} \begin{bmatrix} 3-\beta_1 & 2-\beta_2 \end{bmatrix} \begin{bmatrix} 10 & 8 \\ 8 & 16 \end{bmatrix} \begin{bmatrix} 3-\beta_1 \\ 2-\beta_2 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} = & \frac{1}{3} \begin{bmatrix} 30 - 10\beta_1 + 16 - 8\beta_2 & 24 - 8\beta_1 + 32 - 16\beta_2 \end{bmatrix} \begin{bmatrix} 3-\beta_1 \\ 2-\beta_2 \end{bmatrix} \\ = & \frac{1}{3} \begin{bmatrix} 46 - 10\beta_1 - 8\beta_2 & 56 - 8\beta_1 - 16\beta_2 \end{bmatrix} \begin{bmatrix} 3-\beta_1 \\ 2-\beta_2 \end{bmatrix} \\ = & \frac{1}{3}\left(138 - 30\beta_1 - 24\beta_2 - 46\beta_1 + 10\beta_1^2 + 8\beta_1\beta_2\right) \\ & + \frac{1}{3}\left(112 - 16\beta_1 - 32\beta_2 - 56\beta_2 + 8\beta_1\beta_2 + 16\beta_2^2\right) \\ = & \frac{1}{3}\left(250 - 92\beta_1 - 112\beta_2 + 10\beta_1^2 + 16\beta_2^2 + 16\beta_1\beta_2\right) \end{aligned}$$

Therefore, our confidence region is:

$$R = \left\{(\beta_1, \beta_2) : \frac{1}{3}\left(10\beta_1^2 + 16\beta_2^2 + 16\beta_1\beta_2 - 92\beta_1 - 112\beta_2 + 250\right) \leq 4.61\right\}$$

3. A multiple regression of y on a constant, x_1 and x_2 produces the following results:

$$\hat{y} = 4 + 0.4x_1 + 0.9x_2,\ R^2 = \frac{8}{60},\ \hat{\varepsilon}'\hat{\varepsilon} = 520,\ n = 29,\ X'X = \begin{bmatrix} 29 & 0 & 0 \\ 0 & 50 & 10 \\ 0 & 10 & 80 \end{bmatrix}$$

a. Test the hypothesis that the two slopes sum to 1 at the 5% significance level under normality of y.

Solution Here, $H_0 : \Gamma\beta = \gamma_0$ where

$$\Gamma = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix},\ \gamma_0 = 1$$

We will test this hypothesis using the t statistic:

$$t_0 = \frac{\Gamma\hat{\beta} - \gamma_0}{\sqrt{\hat{\sigma}^2\Gamma\left(X'X\right)^{-1}\Gamma'}}$$

which is distributed $t\left(n-k\right) = t\left(26\right)$ under the standard CNLR assumptions. Filling in the missing data, we have:

$$\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k} = \frac{590}{26} = 22.692$$

Recall the result that for a diagonal block matrix $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{bmatrix}$. Here, if we let $X'X = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$,

$$B^{-1} = \begin{bmatrix} 50 & 10 \\ 10 & 80 \end{bmatrix}^{-1} = \frac{1}{3900} \begin{bmatrix} 80 & -10 \\ -10 & 50 \end{bmatrix} = \begin{bmatrix} \frac{8}{390} & -\frac{1}{390} \\ -\frac{1}{390} & \frac{5}{390} \end{bmatrix}$$

And therefore,

$$\begin{bmatrix} 29 & 0 & 0 \\ 0 & 50 & 10 \\ 0 & 10 & 80 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{29} & 0 & 0 \\ 0 & \frac{8}{390} & -\frac{1}{390} \\ 0 & -\frac{1}{390} & \frac{8}{390} \end{bmatrix}$$

Which gives us:

$$\Gamma \left(X'X\right)^{-1} \Gamma' = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{29} & 0 & 0 \\ 0 & \frac{8}{390} & -\frac{1}{390} \\ 0 & -\frac{1}{390} & \frac{8}{390} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \frac{14}{390}$$

The resulting test statistic is:

$$t_0 = \frac{0.4 + 0.9 - 1}{\sqrt{\frac{590.}{26} \cdot \frac{14}{390}}} = 0.3324$$

And we fail to reject the hypothesis that the coefficients sum to one.

b. Test the hypothesis at the 5% significance level that the slope on x_1 is zero by running the restricted regression and comparing the two sums of squared deviations. As in part (a), carry the test assuming normality of y.

Solution In order to proceed with what little information we are given, it is first necessary to derive an expression which relates the residuals from the unrestricted model to the residuals from the restricted model. I will first do so with some degree of generality, and then I will apply it to this particular problem afterwards as an application. Consider the following unrestricted model. Let $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ and $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$

$$Y = X\beta + \varepsilon_{UR}$$

If we estimate this model using OLS, we can write it as:

$$\begin{aligned} Y &= X\hat{\beta} + \hat{\varepsilon}_{UR} \\ &= X_1\hat{\beta}_1 + X_2\hat{\beta}_2 + \hat{\varepsilon}_{UR} \qquad (1) \end{aligned}$$

Next, consider the restriction $\beta_2 = 0$. (Note that this is quite general, since β_2 can be a vector, but it is not completely general, since it only considers restrictions of the form $\beta_j = 0$ for some collection of j's) This restricted model can then be written as:

$$Y = X_1\beta_1 + \varepsilon_R$$

If we estimate this model using OLS, we will obtain as the residuals:

$$\hat{\varepsilon}_R = M_{X_1}Y$$

But if we substitute in (1) we get

$$\begin{aligned} \hat{\varepsilon}_R &= M_{X_1}\left(X_1\hat{\beta}_1 + X_2\hat{\beta}_2 + \hat{\varepsilon}_{UR}\right) \\ &= M_{X_1}X_1\hat{\beta}_1 + M_{X_1}X_2\hat{\beta}_2 + M_{X_1}\hat{\varepsilon}_{UR} \\ &= M_{X_1}X_2\hat{\beta}_2 + M_{X_1}\hat{\varepsilon}_{UR} \\ &= M_{X_1}X_2\hat{\beta}_2 + \hat{\varepsilon}_{UR} \end{aligned}$$

Where the last equality holds since

$$\begin{aligned} M_{X_1}\hat{\varepsilon}_{UR} &= \left(I - X_1\left(X_1'X_1\right)^{-1}X_1\right)\hat{\varepsilon}_{UR} \\ &= \hat{\varepsilon}_{UR} - X_1\left(X_1'X_1\right)^{-1}X_1\hat{\varepsilon}_{UR} \\ &= \hat{\varepsilon}_{UR} - X_1\left(X_1'X_1\right)^{-1}X_1M_XY \\ &= \hat{\varepsilon}_{UR} \end{aligned}$$

Where $X_1 M_X = (M_X X_1)' = 0' = 0$ since M_X is the annihilator matrix for all of the regressors.

Now, if we want to obtain an expression for $\hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}$, we have:

$$\begin{aligned}
\hat{\varepsilon}_R'\hat{\varepsilon}_R &= \left(M_{X_1}X_2\hat{\beta}_2 + \hat{\varepsilon}_{UR}\right)'\left(M_{X_1}X_2\hat{\beta}_2 + \hat{\varepsilon}_{UR}\right) \\
&= \hat{\beta}_2' X_2' M_{X_1} M_{X_1} X_2 \hat{\beta}_2 + \hat{\varepsilon}_{UR}' M_{X_1} X_2 \hat{\beta}_2 + \hat{\beta}_2' X_2' M_{X_1} \hat{\varepsilon}_{UR} + \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR} \\
&= \hat{\beta}_2' X_2' M_{X_1} X_2 \hat{\beta}_2 + \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}
\end{aligned}$$

Or

$$\hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR} = \hat{\beta}_2' X_2' M_{X_1} X_2 \hat{\beta}_2 \tag{2}$$

For this particular problem, we want to construct the test statistic:

$$F_0 = \frac{\left(\hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}\right)/p}{\hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}/\left(n-k\right)}$$

The restricted model is just a single restriction, so $p = 1$. $\hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}$ is given and is 520, $n - k = 29 - 3 = 26$. It remains only to calculate $\hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}$, which can be done using (2) above (modified to fit the notation of this particular problem):

$$\hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR} = \hat{\beta}_1' X_1' M_{X^*} X_1 \hat{\beta}_1$$

Where $X^* = \begin{bmatrix} 1 & X_2 \end{bmatrix}$. Recall the formula for the estimated variance of a single OLS estimator using the partitioned regression formula:

$$\begin{aligned}
\widehat{Var\left(\hat{\beta}_1 \middle| X\right)} &= Var\left(\widehat{\left(\tilde{X}_1'\tilde{X}_1\right)^{-1}} \tilde{X}_1' Y \middle| X\right) = \left(\tilde{X}_1'\tilde{X}_1\right)^{-1} \tilde{X}_1' \widehat{Var\left(Y\right| X)} \tilde{X}_1 \left(\tilde{X}_1'\tilde{X}_1\right)^{-1} \\
&= \hat{\sigma}^2 \left(\tilde{X}_1'\tilde{X}_1\right)^{-1} = \hat{\sigma}^2 \left(X_1' M_{X^*} X_1\right)^{-1}
\end{aligned}$$

Also, recall the alternative way of computing this variance:

$$\begin{aligned}
\widehat{Var\left(\hat{\beta}_1 \middle| X\right)} &= \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \left[\hat{\sigma}^2 \left(X'X\right)^{-1}\right] \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\
&\equiv \hat{\sigma}^2 \left[\left(X'X\right)^{-1}\right]_{2,2}
\end{aligned}$$

Which we computed in part (a): $\hat{\sigma}^2 \left[\left(X'X\right)^{-1}\right]_{2,2} = \hat{\sigma}^2 \left(\frac{8}{390}\right)$

Putting this together, we see that:

$$\left(X_1' M_{X^*} X_1\right)^{-1} = \frac{\widehat{Var\left(\hat{\beta}_1 \middle| X\right)}}{\hat{\sigma}^2} = \frac{8}{390}$$

Or

$$X_1' M_{X_2} X_1 = \frac{390}{8}$$

This gives us:

$$\begin{aligned}
\hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR} &= \hat{\beta}_1' X_1' M_{X^*} X_1 \hat{\beta}_1 \\
&= (0.4)\left(\frac{390}{8}\right)(0.4) \\
&= 7.8
\end{aligned}$$

Plugging this into the F statistic:

$$F_0 = \frac{\left(\hat{\varepsilon}_R'\hat{\varepsilon}_R - \hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}\right)/p}{\hat{\varepsilon}_{UR}'\hat{\varepsilon}_{UR}/\left(n-k\right)} = \frac{\left(7.8\right)/1}{520/26} = 0.39$$

The critical value for this test is $c^*_{0.05,F(1,26)} = 4.23$. Since $F_0 \leq c^*_{0.05,F(1,26)}$, we fail to reject the null hypothesis that the slope coefficient on x_1 is significantly different from zero.

4. Production data for 22 firms in a certain industry produce the following, where $y = \ln\left(\text{output}\right)$ and $x = \ln\left(\text{labor hours input}\right)$

$$\bar{y} = 20,\ \bar{x} = 10,\ \sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2 = 100,\ \sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 = 60,\ \sum_{i=1}^{n}\left(y_i - \bar{y}\right)\left(x_i - \bar{x}\right) = 30$$

a. Write down the model in matrix notation and state the assumptions that justify running OLS to compute the unknown coefficients. Form the $(X'X)$, $(X'X)^{-1}$, and $(X'Y)$ matrices and compute the least squares estimator of $\beta = \left(\beta_1, \beta_2\right)'$.

Solution I am assuming that we are estimating the model:

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$$

Here, then, it follows that

$$X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_{22} \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_{22} \end{bmatrix}$$

Which gives us:

$$\begin{aligned} X'X &= \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_{22} \end{bmatrix}\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_{22} \end{bmatrix} = \begin{bmatrix} 22 & \sum_{i=1}^{22} x_i \\ \sum_{i=1}^{22} x_i & \sum_{i=1}^{22} x_i^2 \end{bmatrix} \\ &= \begin{bmatrix} 22 & 22\bar{x} \\ 22\bar{x} & \sum_{i=1}^{22}\left(x_i - \bar{x}\right)^2 + 22\bar{x}^2 \end{bmatrix} = \begin{bmatrix} 22 & 220 \\ 220 & 2260 \end{bmatrix} \end{aligned}$$

Where I used the fact that

$$\begin{aligned} \sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 + n\bar{x}^2 &= \sum_{i=1}^{n}\left(x_i^2 - 2\bar{x}x_i + \bar{x}^2\right) + n\bar{x}^2 \\ &= \sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + n\bar{x}^2 + n\bar{x}^2 \\ &= \sum_{i=1}^{n} x_i^2 - 2n\bar{x}^2 + 2n\bar{x}^2 = \sum_{i=1}^{n} x_i^2 \end{aligned}$$

And,

$$\left(X'X\right)^{-1} = \frac{1}{22\left(2260\right) - 220\left(220\right)}\begin{bmatrix} 2260 & -220 \\ -220 & 22 \end{bmatrix} = \begin{bmatrix} \frac{2260}{1320} & -\frac{220}{1320} \\ -\frac{220}{1320} & \frac{22}{1320} \end{bmatrix}$$

Finally,

$$\begin{aligned} X'Y &= \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_{22} \end{bmatrix}\begin{bmatrix} y_1 \\ \vdots \\ y_{22} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{22} y_i \\ \sum_{i=1}^{22} x_i y_i \end{bmatrix} \\ &= \begin{bmatrix} 22\bar{y} \\ \sum_{i=1}^{22}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) + 22\bar{x}\bar{y} \end{bmatrix} = \begin{bmatrix} 22\left(20\right) \\ 30 + 22\left(10\right)\left(20\right) \end{bmatrix} = \begin{bmatrix} 440 \\ 4430 \end{bmatrix} \end{aligned}$$

Where I used the fact that

$$\begin{aligned}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})+n\bar{x}\bar{y} &= \sum_{i=1}^{n}(x_iy_i-\bar{x}y_i-\bar{y}x_i+\bar{x}\bar{y})+n\bar{x}\bar{y}\\ &= \sum_{i=1}^{n}x_iy_i-\bar{x}\sum_{i=1}^{n}y_i-\bar{y}\sum_{i=1}^{n}x_i+n\bar{x}\bar{y}+n\bar{x}\bar{y}\\ &= \sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}-n\bar{y}\bar{x}+2n\bar{x}\bar{y}=\sum_{i=1}^{n}x_iy_i\end{aligned}$$

The estimators are therefore:

$$\hat{\beta}=(X'X)^{-1}X'Y=\begin{bmatrix}\frac{2260}{1320} & -\frac{220}{1320}\\ -\frac{220}{1320} & \frac{22}{1320}\end{bmatrix}\begin{bmatrix}440\\4430\end{bmatrix}=\begin{bmatrix}15\\ \frac{1}{2}\end{bmatrix}$$

b. Test the statistical significance of your estimates at the 5% significance level assuming that the $y_i's$ are jointly normally distributed with variance-covariance matrix the identity matrix.

Solution Here, by assumption, $\sigma^2=1$. This gives us the following standard errors:

$$\begin{aligned}\left[se\left(\hat{\beta}_1\right)\right]^2 &= \sigma^2\Gamma(X'X)^{-1}\Gamma'=\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}\frac{2260}{1320} & -\frac{220}{1320}\\ -\frac{220}{1320} & \frac{22}{1320}\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}=\frac{2260}{1320}\\ \left[se\left(\hat{\beta}_2\right)\right]^2 &= \sigma^2\Gamma(X'X)^{-1}\Gamma'=\begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix}\frac{2260}{1320} & -\frac{220}{1320}\\ -\frac{220}{1320} & \frac{22}{1320}\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix}=\frac{22}{1320}\end{aligned}$$

Or

$$\begin{aligned}se\left(\hat{\beta}_1\right) &= \sqrt{\frac{2260}{1320}}=1.3085\\ se\left(\hat{\beta}_2\right) &= \sqrt{\frac{22}{1320}}=0.1291\end{aligned}$$

And our test statistics for statistical significance are:

$$\begin{aligned}t_0^{\hat{\beta}_1} &= \frac{\hat{\beta}_1}{se\left(\hat{\beta}_1\right)}=\frac{15}{1.3085}=11.464\\ t_0^{\hat{\beta}_2} &= \frac{\hat{\beta}_2}{se\left(\hat{\beta}_2\right)}=\frac{0.5}{0.1291}=3.8730\end{aligned}$$

Using the critical value $c^*_{0.025,t(20)}=2.07$, we have that both coefficients are statistically significant. (Since $\left|t_0^{\hat{\beta}_1}\right|>c^*_{0.025,t(20)}$ and $\left|t_0^{\hat{\beta}_2}\right|>c^*_{0.025,t(20)}$)

c. Test the hypothesis that there exist constant returns to labor at the 5% significance level under the same assumption as in part (b).

Here, we are testing $H_0:\beta_2=1$. The relevant test statistic is therefore:

$$t_0^{\hat{\beta}_2-1}=\frac{\hat{\beta}_2-1}{se\left(\hat{\beta}_2\right)}=-\frac{0.5}{0.1291}=-3.8730$$

And we reject the hypothesis of constant returns to labor since $\left|t_0^{\hat{\beta}_2-1}\right|>c^*_{0.025,t(20)}$.