Introductory Econometrics
Tutorial 4 Solutions

PART A:

1. Consider the matrix:
$$\mathbf{X} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$
Which one of the following matrices is $\mathbf{X}'\mathbf{X}$?

   (a)
$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
This option can be easily eliminated because $\mathbf{X}'\mathbf{X}$ must be a 3×3 matrix.

   (b)
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
Only this option and the last option are 3×3 matrices. Computing a couple of elements of $\mathbf{X}'\mathbf{X}$ makes it clear that this option is not correct.

   (c)
$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 1 \\ 1 & 2 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$
This option can be easily eliminated because $\mathbf{X}'\mathbf{X}$ must be a 3×3 matrix.

   (d)
$$\begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$$
This option can be easily eliminated because $\mathbf{X}'\mathbf{X}$ must be a 3×3 matrix.

   (e)
$$\begin{pmatrix} 4 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$
This is the right answer because
$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

2. Let

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Which one of the following matrices is $(\mathbf{X}'\mathbf{X})^{-1}$? (Hint: You do not need to know how to calculate an inverse of a matrix, you only need to check which of the following satisfies the definition of an inverse.)

(a)

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

This can be easily eliminated because the inverse of a 3×3 matrix must be 3×3.

(b)

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

This can be eliminated because this is an identity matrix. Identity multiplied by any matrix equals that matrix again. That is not the definition of an inverse.

(c)

$$\begin{pmatrix} 0.25 & 0.5 & 1 \\ 0.5 & 0.5 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

The test is if the product of this matrix an ($\mathbf{X}'\mathbf{X}$ is an identity matrix:

$$\begin{pmatrix} 0.25 & 0.5 & 1 \\ 0.5 & 0.5 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3.0 & 1.5 & 1.25 \\ 3.0 & 2.0 & 0.5 \\ 5 & 2 & 2 \end{pmatrix}$$

which is not an identity matrix.

(d)

$$\begin{pmatrix} 4 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Applying the same test:

$$\begin{pmatrix} 4 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 21 & 12 & 5 \\ 12 & 8 & 2 \\ 5 & 2 & 2 \end{pmatrix}$$

which is not an identity matrix.

(e)

$$\begin{pmatrix} 1.0 & -1.0 & -1.0 \\ -1.0 & 1.5 & 1.0 \\ -1.0 & 1.0 & 2.0 \end{pmatrix}$$

Applying the same test:

$$\begin{pmatrix} 1.0 & -1.0 & -1.0 \\ -1.0 & 1.5 & 1.0 \\ -1.0 & 1.0 & 2.0 \end{pmatrix} \begin{pmatrix} 4 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

which means that this is the correct answer.

3. We have estimated the model $wage = \beta_0 + \beta_1$ experience $+u$ using OLS based on a sample of 4 observations. We know that the matrix of explanatory variables is

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix}$$

and we are told that the OLS residuals are:

$$\hat{\mathbf{u}} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

We can immediately say that the residual vector:

(a) is reported correctly because it sums to zero: Summing to zero only implies that $\hat{\mathbf{u}}$ is orthogonal to the first column of $\mathbf{X}$,

$$\begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} = 0$$

but the OLS residual vector has to be orthogonal to all columns of $\mathbf{X}$

(b) is reported incorrectly because it should be a 2×1 vector: This is rubbish. Every observation will have a residual, so the dimension of $\hat{\mathbf{u}}$ is the same as the dimension of $\mathbf{y}$

(c) is reported correctly but shows a poor fit because $\mathbf{X}'\hat{\mathbf{u}} \neq \mathbf{0}$: This cannot be true because if $\hat{\mathbf{u}}$ was correct $\mathbf{X}'\hat{\mathbf{u}}$ would be equal to zero

(d) is reported correctly because it is linearly independent of columns of $\mathbf{X}$: while it is true that $\hat{\mathbf{u}}$ is linearly independent of columns of $\mathbf{X}$, but any vector that is not a linear combination of columns of $\mathbf{X}$ would be linearly independent of columns of $\mathbf{X}$. But not all such vectors are orthogonal to columns of $\mathbf{X}$.

(e) is reported incorrectly because it is not orthogonal to the second column of $\mathbf{X}$: yes, this is the correct answer because:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -4 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

4. In the bivariate linear regression model

$$
\begin{aligned}
\widehat{\beta}_1 &= \frac{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2} \\
&= \frac{\widehat{\sigma}_{xy}}{\widehat{\sigma}_x^2} \\
&= \frac{100}{5} \\
&= 20.
\end{aligned}
$$

5. By definition

$$
\begin{aligned}
R^2 &= \frac{SSE}{SST} \\
&\Rightarrow 0.5 = \frac{10}{SST} \\
&\Rightarrow SST = 20
\end{aligned}
$$

$$
\begin{aligned}
SSR &= SST - SSE \\
&= 20 - 10 \\
&= 10.
\end{aligned}
$$

PART B: You do not need to hand this part in. It will be covered in the tutorial. It is still a good idea to attempt these questions before the tutorial.

1. Suppose that

$$\underset{n\times 3}{\mathbf{X}} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \vdots & & \vdots \\ x_{n1} & x_{n2} & x_{n3} \end{bmatrix}$$

and

$$\underset{3\times 1}{\widehat{\boldsymbol{\beta}}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}$$

Show that $\widehat{\mathbf{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}}$ is an $n \times 1$ vector which is a linear combination (a weighted sum) of the columns of $\mathbf{X}$ with weights given by the elements of $\widehat{\boldsymbol{\beta}}$. That is:

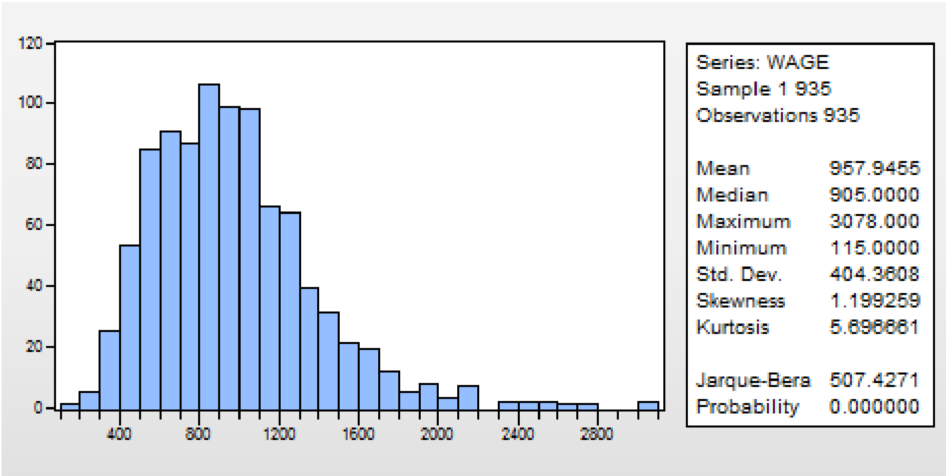
$$\widehat{\mathbf{y}} = \text{first column of } \mathbf{X} \times \hat{\beta}_1 + \text{second column of } \mathbf{X} \times \hat{\beta}_2 + \text{third column of } \mathbf{X} \times \hat{\beta}_3$$

In fact this is not specific to $\mathbf{X}$ having 3 columns. It is true for any $n \times k$ matrix $\mathbf{X}$ and $k \times 1$ vector $\widehat{\boldsymbol{\beta}}$.

- $\widehat{\mathbf{y}} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \vdots & & \vdots \\ x_{n1} & x_{n2} & x_{n3} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} x_{11}\hat{\beta}_1 + x_{12}\hat{\beta}_2 + x_{13}\hat{\beta}_3 \\ x_{21}\hat{\beta}_1 + x_{22}\hat{\beta}_2 + x_{23}\hat{\beta}_3 \\ \vdots \\ x_{n1}\hat{\beta}_1 + x_{n2}\hat{\beta}_2 + x_{n3}\hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix} \times \hat{\beta}_1 + \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix} \times \hat{\beta}_2 + \begin{bmatrix} 13 \\ x_{23} \\ \vdots \end{bmatrix} \times \hat{\beta}_3$

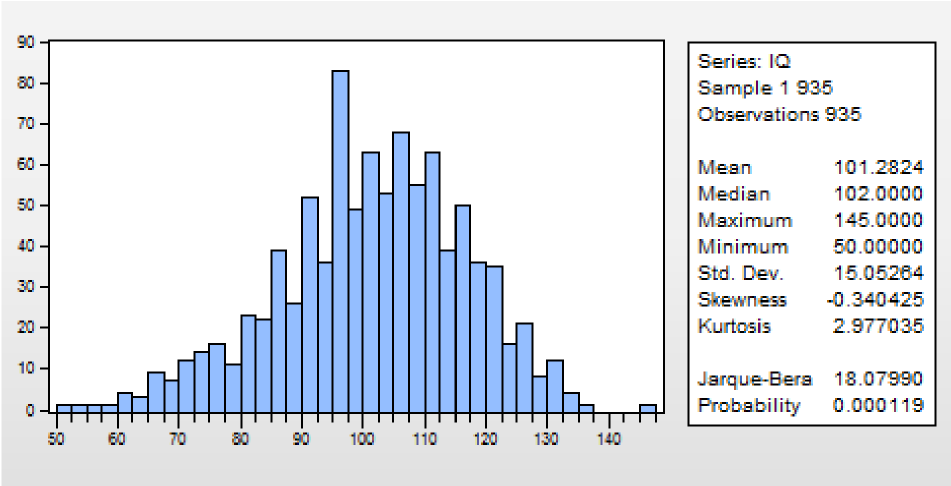
  x_{n3}

2. This question is based on question C4 in Chapter 2 of the textbook. The dependent variable is *wage* and the independent variable is IQ.
   Preliminary analysis of the data (not asked in the question, but a step that we usually take before estimating a regression equation: "Look" at these variables (meaning that examine their histograms, summary statistics, scatter plot of wage against IQ, sample correlation coefficient between wage and IQ, and summarise your insights from just looking at data through these views).

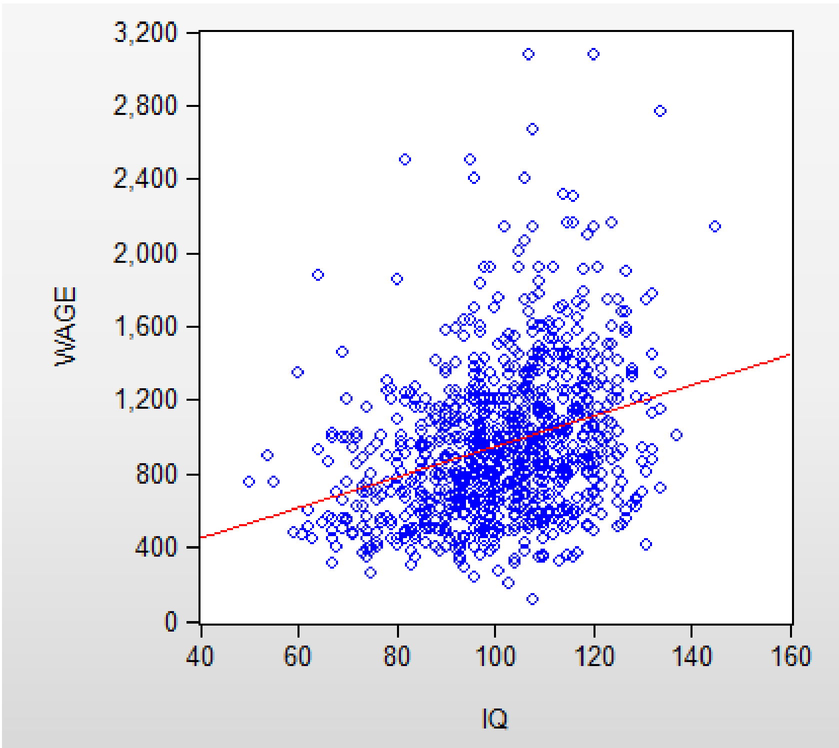


   Wage is positively skewed. The sample mean is \$958, but because of this skewness the median \$905 is a better indicator of the central tendency than the mean. Need to make sure that the one outlying observation with \$3078 wage does not affect the analysis.

IQ seems to be representative, with mean close to 100 and standard deviation close to 15. There is a very smart person with IQ of 145, so need to ensure that it does not influence the analysis too much.



The scatter plot of wage against IQ, in particular with the regression line included, shows that the relationship between wage and IQ, although seems positive (sample correlation coefficient is 0.309), is not linear. The variation of wage around the regression line seems to be higher at higher IQs.



Sample: 1 935
Included observations: 935

Correlation	IQ	WAGE
IQ	1.000000	
WAGE	0.309088	1.000000

(a) Run a regression of *wage* on a constant only. Verify that the OLS estimate of the intercept is the sample mean of wage and the standard error of the regression is the sample standard deviation of wage.

Dependent Variable: WAGE
Method: Least Squares
Sample: 1 935
Included observations: 935

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	957.9455	13.22401	72.43985	0.0000

R-squared	0.000000	Mean dependent var	957.9455
Adjusted R-squared	0.000000	S.D. dependent var	404.3608
S.E. of regression	404.3608	Akaike info criterion	14.84356

(b) *Estimation, interpretation of the slope coefficient and R^2 of the regression*: Estimate a simple regression model where a one-point increase in *IQ* changes *wage* by a constant dollar amount. Use this model to find the predicted increase in *wage* for an increase in *IQ* of 15 points. Does *IQ* explain most of the variation in *wage*? What is the relationship between the R^2 of this regression and the sample correlation coefficient between wage and IQ?

- Dependent Variable: WAGE
  Method: Least Squares
  Date: 07/28/16 Time: 12:46
  Sample: 1 935
  Included observations: 935

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	116.9916	85.64153	1.366061	0.1722
IQ	8.303064	0.836395	9.927203	0.0000

R-squared	0.095535	Mean dependent var	957.9455
Adjusted R-squared	0.094566	S.D. dependent var	404.3608
S.E. of regression	384.7667	Akaike info criterion	14.74529
Sum squared resid	1.38E+08	Schwarz criterion	14.75564
Log likelihood	-6891.422	Hannan-Quinn criter.	14.74924
F-statistic	98.54936	Durbin-Watson stat	1.802114
Prob(F-statistic)	0.000000		

If *IQ* increases by 15 points (*i.e.* one standard deviation), the predicted increase in wage is $8.303 \times 15 = 124.545$ (get students to use their calculators).
The R^2 is very low, so there is a lot of unexplained variation. Note that in a regression with only one explanatory variable, R^2 is the square of sample correlation coefficient between the dependent variable and the explanatory variable. Ask them to verify that on their calculators $0.309088^2 = 0.095535$.