ECON 221

Assignment 4

Due: October 25, 2012

1. {Goldberger 14.4} Show that every symmetric idempotent matrix is nonnegative definite.

2. What happens when you include an intercept? Consider the following models:

$$\begin{aligned} y_i &= \beta_1 + \beta_2 x_i + \varepsilon_i, && i = 1, \ldots, n \\ y_i &= \beta_1 + \beta_2 (x_i - \overline{x}) + \varepsilon_i, && i = 1, \ldots, n \\ y_i &= \beta_1 (x_i - \overline{x}) + \varepsilon_i, && i = 1, \ldots, n \end{aligned}$$

   (a) Write each model in matrix form.

   (b) Show that $|\mathbf{X'X}|$ is the same for the first two models.

   (c) Show that $\mathbf{X}\hat{\beta}$ is the same for the first two models, and explain how the estimates of $\hat{\beta}$ for those models are different from the estimate in the final model specification.

   (d) Interpret this result.

3. {Goldberger 17.3} Suppose that the $n \times k$ matrix $\mathbf{X} = (\mathbf{X_1}, \mathbf{X_2})$ has full column rank. Let $\mathbf{X_2^*} = \mathbf{M_1 X_2}$ be the $n \times k_2$ matrix of residuals from the auxiliary regression of $\mathbf{X_2}$ on $\mathbf{X_1}$. Show that $\text{rank}(\mathbf{X_2^*}) = k_2$. Hint: Use proof by contradiction.

4. {Goldberger 15.4} The Classical Regression model applies along with the usual notation. For each of the following statements, indicate whether it is true or false, and justify your answer.

   (a) The random variable $t = \mathbf{b'b}$ is an unbiased estimator of the parameter $\theta = \beta'\beta$.

   (b) Since $\hat{\mathbf{y}} = \mathbf{Ny}$ where $\mathbf{N} = \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}$, it follows that $\mathbf{y} = \mathbf{N}^{-1}\hat{\mathbf{y}}$.

   (c) Since $E(\hat{\mathbf{y}}) = E(\mathbf{y})$, it follows that the sum of the residuals is zero.

   (d) If b_1 and b_2 are the first two elements of $\mathbf{b}$, $t_1 = b_1 + b_2$, and $t_2 = b_1 - b_2$, then $V(t_1) \geq V(t_2)$.

5. Suppose you are addressing the question of geographical effects on unemployment and have unemployment data on 50 different cities in each region of the US: Northeast, Midwest, South, and West.[1] In order to study your question, you set your explanatory variables as $x_1 = 1$ and the set $\{x_i\}$ = a binary variable that is equal to 1 if the city is in region i and 0 otherwise for $i \in \{Northeast, Midwest, South, West\}$. Will the $\mathbf{X}$ matrix have full column rank? Why?

[1] Regions according to the U.S. Census Beuro: `http://www.census.gov/geo/www/us_regdiv.pdf`