BOSTON COLLEGE
Department of Economics
EC771: Econometrics
Spring 2004
Prof. Baum, Ms. Uysal

SOLUTION KEY FOR PROBLEM SET 2

1. **The Two Variable Regression.** For the regression model $y = \alpha + \beta x + \epsilon$,
(a) Show that the least squares normal equations imply $\sum_i e_i = 0$ and $\sum_i x_i e_i = 0$.
(b) Show that the solution for the constant term is $a = \bar{y} - b\bar{x}$.
(c) Show that the solution for b is $b = \frac{\left[\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\right]}{\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]}$.
(d) Prove that these two values uniquely minimize the sum of squares by showing that the diagonal elements of the second derivatives matrix of the sum of squares with respect to the parameters are both positive and that the determinant is

$$4n\left[(\sum_{i=1}^{n} x_i^2) - n\bar{x}^2\right] = 4n\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right],$$

which is positive unless all values of x are the same.

(a) $\min L = e_i^2 = \sum_i (y_i - \alpha - \beta x_i)^2$

$$\frac{\partial L}{\partial \alpha} = -2\sum_i (y_i - \alpha - \beta x_i) = 0$$

$$\frac{\partial L}{\partial \beta} = -2\sum_i (y_i - \alpha - \beta x_i)x_i = 0$$

This implies that $\sum_i e_i = 0$ and $\sum_i x_i e_i = 0$.

(b) Use $\sum_i e_i = 0$ to conclude from the first normal equation that $a = \bar{y} - b\bar{x}$.

$$\begin{aligned}
\sum_i e_i &= \sum_i (y_i - a - bx_i) = 0 \\
&= \sum_i y_i - (n)a - b\sum_i x_i = 0 \\
& \frac{1}{n}\sum_i y_i + \frac{1}{n} b \sum_i x_i = a \\
& \bar{y} - b\bar{x} = a.
\end{aligned}$$

(c) We know that $\sum_i e_i = 0$ and $\sum_i x_i e_i = 0$. It follows then that $\sum_i (x_i - \bar{x})e_i = 0$. Further, the latter implies $\sum_i (x_i - \bar{x})(y_i - a - bx_i) = 0$ or $\sum_i (x_i -$

$\bar{x})(y_i - \bar{y} - b(x_i - \bar{x})) = 0$ from which the result follows.

(d) Second derivatives are as follows,

$$\frac{\partial^2 L}{\partial \alpha^2} = 2n$$
$$\frac{\partial^2 L}{\partial \beta^2} = 2\sum_i x_i^2$$
$$\frac{\partial^2 L}{\partial \alpha \partial \beta} = -2n\bar{x}$$

The diagonal elements of the second derivatives matrix of the sum of squares with respect to the parameters are both positive and that the determinant is $4n\left[\left(\sum_{i=1}^n x_i^2\right) - n\bar{x}^2\right] = 4n\left[\sum_{i=1}^n (x_i - \bar{x})^2\right]$, which is positive unless all values of x are the same.

2. Suppose $\mathbf{b}$ is the least squares coefficient vector in the regression of $\mathbf{y}$ on $\mathbf{X}$ and $\mathbf{c}$ is another Kx1 vector. Prove that the difference in the two sums of squared residuals is

$$(\mathbf{y} - \mathbf{Xc})'(\mathbf{y} - \mathbf{Xc}) - (\mathbf{y} - \mathbf{Xb})'(\mathbf{y} - \mathbf{Xb}) = (\mathbf{c} - \mathbf{b})'\mathbf{X}'\mathbf{X}(\mathbf{c} - \mathbf{b}).$$

Prove that this difference is positive.

Write $\mathbf{c}$ as **b+(c-b)**. Then, the sum of the squared residuals based on $\mathbf{c}$ is

$$\begin{aligned}
(\mathbf{y} - \mathbf{Xc})'(\mathbf{y} - \mathbf{Xc}) &= [\mathbf{y} - \mathbf{X}(\mathbf{b} + (\mathbf{c} - \mathbf{b}))]'[\mathbf{y} - \mathbf{X}(\mathbf{b} + (\mathbf{c} - \mathbf{b}))] \\
&= [(\mathbf{y} - \mathbf{Xb}) + \mathbf{X}(\mathbf{c} - \mathbf{b})]'[(\mathbf{y} - \mathbf{Xb}) + \mathbf{X}(\mathbf{c} - \mathbf{b})] \\
&= (\mathbf{y} - \mathbf{Xb})'(\mathbf{y} - \mathbf{Xb}) + (\mathbf{c} - \mathbf{b})'\mathbf{X}'\mathbf{X}(\mathbf{c} - \mathbf{b}) + \mathbf{2}(\mathbf{c} - \mathbf{b})'\mathbf{X}'(\mathbf{y} - \mathbf{Xb}).
\end{aligned}$$

But, the third term is zero, as $\mathbf{2}(\mathbf{c} - \mathbf{b})'\mathbf{X}'(\mathbf{y} - \mathbf{Xb}) = \mathbf{2}(\mathbf{c} - \mathbf{b})'\mathbf{X}'\mathbf{e} = \mathbf{0}$. Therefore,

$$\begin{aligned}
(\mathbf{y} - \mathbf{Xc})'(\mathbf{y} - \mathbf{Xc}) &= \mathbf{e}'\mathbf{e} + (\mathbf{c} - \mathbf{b})'\mathbf{X}'\mathbf{X}(\mathbf{c} - \mathbf{b}) \\
(\mathbf{y} - \mathbf{Xc})'(\mathbf{y} - \mathbf{Xc}) - \mathbf{e}'\mathbf{e} &= (\mathbf{c} - \mathbf{b})'\mathbf{X}'\mathbf{X}(\mathbf{c} - \mathbf{b}).
\end{aligned}$$

The right hand side can be written as **d'd** where **d=X(c-b)**, so it is necessaril y positive. This confirms what we knew at the outset, least squares is least squares.

3. Consider the least squares regression of $\mathbf{y}$ on K variables (with a constant), $\mathbf{X}$. Consider an alternative set of regressors, $\mathbf{Z} = \mathbf{XP}$, where $\mathbf{P}$ is a nonsingular matrix. Thus, each column of $\mathbf{Z}$ is a mixture of some of the columns of $\mathbf{X}$. Prove that the residual vectors in the regressions of $\mathbf{y}$ on$\mathbf{X}$ and $\mathbf{y}$ on $\mathbf{Z}$ are identical. What relevance does this have to the question of changing the fit of a regression by changing the units of measurement of independent variables?

The residual vector in the regression of $\mathbf{y}$ on $\mathbf{X}$ is $\mathbf{M_x y} = [\mathbf{I} - \mathbf{X}(\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'}]\mathbf{y}$. The residual vector in the regression of $\mathbf{y}$ on $\mathbf{Z}$ is,

$$\begin{aligned}
\mathbf{M_z y} &= [\mathbf{I} - \mathbf{Z}(\mathbf{Z'Z})^{-\mathbf{1}}\mathbf{Z'}]\mathbf{y} \\
&= [\mathbf{I} - \mathbf{XP}((\mathbf{XP})'(\mathbf{XP}))^{-\mathbf{1}}(\mathbf{XP})']\mathbf{y} \\
&= [\mathbf{I} - \mathbf{XPP^{-1}}(\mathbf{X'X})^{-\mathbf{1}}(\mathbf{P'})^{-\mathbf{1}}\mathbf{P'X'}]\mathbf{y} \\
&= [\mathbf{I} - \mathbf{X}(\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'}]\mathbf{y} \\
&= \mathbf{M_x y}
\end{aligned}$$

Since the residual vectors are identical, the fits must be as well. Changing the units of measurement of the regressors is equivalent to postmultipliying by a diagonal $\mathbf{P}$ matrix whose kth diagonal element is the scale factor to be applied to the kth variable (1 if it is to be unchanged). It follows from the result above that this will not change the fit of the regression.

4. Let Y denote total expenditure on consumer durables, nondurables, and services, and E_d, E_n, and E_s are the expenditures on the three categories. As defined, $Y = E_d + E_n + E_s$. Now, consider the expenditure system

$$\begin{aligned}
E_d &= \alpha_d + \beta_d Y + \gamma_{dd} P_d + \gamma_{dn} P_n + \gamma_{ds} P_s + \epsilon_d \\
E_n &= \alpha_n + \beta_n Y + \gamma_{nd} P_d + \gamma_{nn} P_n + \gamma_{ns} P_s + \epsilon_n \\
E_s &= \alpha_s + \beta_s Y + \gamma_{sd} P_d + \gamma_{sn} P_n + \gamma_{ss} P_s + \epsilon_s.
\end{aligned}$$

Prove that if all equations are estimated by ordinary least squares, then the sum of the income coefficients will be 1 and the four other column sums in the preceding model will be zero.

For convenience, reorder the variables so that $\mathbf{X} = [\mathbf{i}, \mathbf{P_d}, \mathbf{P_n}, \mathbf{P_s}, \mathbf{Y}]$. The three dependent variables are E_d, E_n, and E_s, and $\mathbf{Y} = \mathbf{E_d} + \mathbf{E_n} + \mathbf{E_s}$. The coefficient vectors are

$$\mathbf{b_d} = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'E_d}, \mathbf{b_n} = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'E_n}, and \mathbf{b_s} = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'E_s}.$$

The sum of the three vectors is

$$\mathbf{b} = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'}[\mathbf{E_d} + \mathbf{E_n} + \mathbf{E_s}] = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'Y}.$$

Now, $\mathbf{Y}$ is the last column of $\mathbf{X}$, so the preceding sum is the vector of least squares coefficients in the regression of the last column of $\mathbf{X}$ on all of the columns of $\mathbf{X}$, including the last. Of course, we get a perfect fit. In addition, $\mathbf{X'}[\mathbf{E_d} + \mathbf{E_n} + \mathbf{E_s}]$ is the last column of $(\mathbf{X'X})$, so the matrix product is equal to the last column of an identity matrix. Thus, the sum of the coefficients on all variables except income is 0, while that on income is 1.

5. Suppose you have two independent unbiased estimators of the sample parameter, θ, say $\hat{\theta_1}$ and $\hat{\theta_2}$, with different variances, ν_1 and ν_2. What linear combination, $\hat{\theta} = c_1\hat{\theta_1} + c_2\hat{\theta_2}$ is the minimum variance unbiased estimator of θ?

Consider the optimization problem of minimizing the variance of the weighted estimator. If the estimate is to be unbiased, it must be of the form $c_1\hat{\theta_1} + c_2\hat{\theta_2}$ where c_1 and c_2 sum to 1. Thus, $c_2 = 1 - c_1$. The function to minimize is $Min_{c_1} L_* = c_1^2\nu_1 + (1 - c_1)^2\nu_2$. The necessary condition is $\frac{\partial L_*}{\partial c_1} = 2c_1\nu_1 - 2(1 - c_1)\nu_2 = 0$ which implies $c_1 = \frac{\nu_2}{\nu_1+\nu_2}$. A more intuitively appealing form is obtained by dividing numerator and denominator by $\nu_1\nu_2$ to obtain $c_1 = \frac{(1/\nu_1)}{[(1/\nu_1)+(1/\nu_2)]}$. Thus, the weight is proportional to the inverse of the variance. The estimator with the smaller variance gets the larger weight.