January 22, 2006

Greene Problems

3.1

$y = a + bx + \varepsilon \qquad \min_{a,b} \Sigma\left[y - a - bx\right]^2$

a)FOC: $\quad -2\Sigma\left[y-a-bx\right] = 0 \qquad -2\Sigma\left[y-a-bx\right]x = 0$

Hence $\Sigma\left[y-a-bx\right] = \Sigma e = 0$, $\Sigma\left[y-a-bx\right]x = \Sigma ex = 0$

b)$0 = \Sigma\left[y-a-bx\right] = \Sigma\left[y\right] - a\Sigma\left[1\right] - b\Sigma\left[x\right] = n(\overline{y} - a - b\overline{x})$

Therefore, $a = \overline{y} - b\overline{x}$

c) $0 = \Sigma\left[y-a-bx\right]x = \Sigma\left[y - \overline{y} + b\overline{x} - bx\right]x = \Sigma\left[(y-\overline{y}) - b\left(x - \overline{x}\right)\right]x$

$0 = \Sigma(y-\overline{y}) = \Sigma(y-\overline{y})\overline{x} \qquad 0 = b\Sigma\left(x-\overline{x}\right) = b\Sigma\left(x-\overline{x}\right)\overline{x}$

Hence, $\Sigma\left(y-\overline{y}\right)\left(x-\overline{x}\right) = b\Sigma\left(x-\overline{x}\right)\left(x-\overline{x}\right)$ and $b = \frac{\Sigma(y-\overline{y})(x-\overline{x})}{\Sigma(x-\overline{x})(x-\overline{x})}$

d) SOC: $\quad \left|\frac{\partial^2\left(\Sigma[y-a-bx]^2\right)}{\partial(a,b)^2}\right| = \begin{vmatrix} 2n & 2n\overline{x} \\ 2n\overline{x} & 2n\overline{x^2} \end{vmatrix} = 4n^2\left(\overline{x^2} - \overline{x}^2\right) = 4n\Sigma\left(x-\overline{x}\right)^2 > 0$

3.3

$Z = XP, \qquad e_x = M_x y = \left[I - X\left(X'X\right)^{-1}X'\right]y$

$e_z = M_z y = \left[I - XP\left(P'X'XP\right)^{-1}P'X'\right]y = \left[I - XPP^{-1}\left(X'X\right)^{-1}\left(P'\right)^{-1}P'X'\right]y = M_x y = e_x$

The fit of the regression does not change.

3.6

$X_{ns} = \begin{bmatrix} X_n \\ X_s \end{bmatrix} \qquad y_{ns} = \begin{bmatrix} y_n \\ y_s \end{bmatrix} \qquad b_n = \left(X_n'X_n\right)^{-1}\left(X_n'y_n\right)$

$b_{ns} = \left(X_{ns}'X_{ns}\right)^{-1}X_{ns}'y_{ns} = \left(X_n'X_n + X_s'X_s\right)^{-1}\left(X_n'y_n + X_s'y_s\right)$

Compute $\left(X_n'X_n + X_s'X_s\right)\left(b_n + \left(X_s\left(X_n'X_n\right)^{-1}X_s' + 1\right)^{-1}\left(X_n'X_n\right)^{-1}X_s'\left(y_s - X_s b_n\right)\right)$

$b_n + \left(X_s\left(X_n'X_n\right)^{-1}X_s' + 1\right)^{-1}\left(X_n'X_n\right)^{-1}X_s'\left(y_s - X_s b_n\right) =$

$= \left(X_n'X_n\right)^{-1}X_n'y_n + \left(X_s\left(X_n'X_n\right)^{-1}X_s' + I_s\right)^{-1}\left(X_n'X_n\right)^{-1}X_s'\left(y_s - X_s\left(X_n'X_n\right)^{-1}X_n'y_n\right)$

$\left(X_n'X_n + X_s'X_s\right)\left(X_n'X_n\right)^{-1}X_n'y_n = \left(I_k + X_s'X_s\left(X_n'X_n\right)^{-1}\right)X_n'y_n$

$\left(X_n'X_n + X_s'X_s\right)\left(X_n'X_n\right)^{-1}X_s'\left(y_s - X_s\left(X_n'X_n\right)^{-1}X_n'y_n\right) =$

$= \left(I_k + X_s'X_s\left(X_n'X_n\right)^{-1}\right)\left(X_s'y_s - X_s'X_s\left(X_n'X_n\right)^{-1}X_n'y_n\right)$

$\left(I_k + X_s'X_s\left(X_n'X_n\right)^{-1}\right)X_n'y_n + \left(X_s\left(X_n'X_n\right)^{-1}X_s' + 1\right)^{-1} *$

$* \left(I_k + X_s'X_s\left(X_n'X_n\right)^{-1}\right)\left(X_s'y_s - X_s'X_s\left(X_n'X_n\right)^{-1}X_n'y_n\right) =$

$\left(I_k + X_s'X_s\left(X_n'X_n\right)^{-1}\right)X_n'y_n + \left(X_s'y_s - X_s'X_s\left(X_n'X_n\right)^{-1}X_n'y_n\right) = X_n'y_n + X_s'y_s$

Hence, they are equal, and new data changes the estimate only if it is not exactly on the predicted line.

3.7

Assume we have y (n x 1) and X (n x 1). $\qquad \widehat{b} = (X'X)^{-1}X'y$

If follow the 'strategy' we get

$$\begin{bmatrix} y_1 \\ ... \\ y_n \\ 0 \end{bmatrix} = \begin{bmatrix} x_{11} & ... & x_{1k} & 0 \\ ... & ... & ... & 0 \\ x_{n1} & ... & x_{nn} & 0 \\ x_{n+1,1} & ... & x_{n+1,n} & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ ... \\ b_k \\ b_{k+1} \end{bmatrix} + \begin{bmatrix} e_1 \\ ... \\ e_n \\ 0 \end{bmatrix}$$

The first n rows are equivalent to the initial problem. The last row uniquely determines the value of b_{k+1} as here the last error can be made exactly zero. Given that the old estimate is still optimal since the new observation hasn't increased the objective function. Hence, the fit does not change.

$$\begin{bmatrix} y_1 \\ \dots \\ y_n \\ y' \end{bmatrix} = \begin{bmatrix} 1 & x_{1k} & 0 \\ \dots & \dots & 0 \\ 1 & x_{nn} & 0 \\ 1 & x_{n+1,n} & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} + \begin{bmatrix} e_1 \\ \dots \\ e_n \\ 0 \end{bmatrix}$$

Similarly, here any value of $y\prime$ could be plugged in since it will only change c, not the fit.

4.1

$E\left[c_1\theta_1 + c_2\theta_2\right] = c_1 E\left[\theta_1\right] + c_2 E\left[\theta_2\right] = \left(c_1 + c_2\right) E\left[\theta\right] = E\left[\theta\right] \qquad \Rightarrow \qquad c_1 + c_2 = 1$

$\min_c \left[c^2\nu_1 + (1-c)^2 \nu_2\right] \qquad$ FOC: $\qquad c\nu_1 = (1-c)\,\nu_2 \qquad \Rightarrow \qquad c = \frac{\nu_2}{\nu_1+\nu_2}$

$\widehat{\theta} = c_1\widehat{\theta}_1 + c_2\widehat{\theta}_2 = \frac{\nu_2}{\nu_1+\nu_2}\widehat{\theta}_1 + \frac{\nu_1}{\nu_1+\nu_2}\widehat{\theta}_2 = \frac{\frac{1}{\nu_1}}{\frac{1}{\nu_1}+\frac{1}{\nu_2}}\widehat{\theta}_1 + \frac{\frac{1}{\nu_2}}{\frac{1}{\nu_1}+\frac{1}{\nu_2}}\widehat{\theta}_2$

$V\left[\widehat{\theta}\right] = \left(\frac{\nu_2}{\nu_1+\nu_2}\right)^2 \nu_1 + \left(\frac{\nu_1}{\nu_1+\nu_2}\right)^2 \nu_2 = \frac{\nu_1\nu_2}{\nu_1+\nu_2} = \frac{1}{\frac{1}{\nu_1}+\frac{1}{\nu_2}}$

So the intuition is in harmonization: the higher the variance the lower the weight.

4.2

$\widehat{\beta} = c'y = c'x + c'\varepsilon \qquad E\left[\widehat{\beta}\right] = \beta c'x \qquad V\left[\widehat{\beta}\right] = \left(c'c\right)\sigma^2$

$\min_c MSE\left[\widehat{\beta}\right] = \min_c \left\{\left[\beta\left(c'x-1\right)\right]^2 + \left(c'c\right)\sigma^2\right\}$

FOC: $\frac{\partial}{\partial c}\frac{1}{2}\left\{\beta^2\left(c'x-1\right)^2 + \left(c'c\right)\sigma^2\right\} = \beta^2\left(c'x-1\right)x + \sigma^2 c = 0$

$\beta^2\left(c'x-1\right)x'x = -\sigma^2 x'c = -\sigma^2 c'x \qquad \Rightarrow \qquad c'x = \frac{\beta^2 x'x}{\beta^2 x'x+\sigma^2} \qquad \Rightarrow$

$c = \frac{\beta^2(1-c'x)x}{\sigma^2} = \frac{\beta^2\left(1-\frac{\beta^2 x'x}{\beta^2 x'x+\sigma^2}\right)x}{\sigma^2} = \frac{x}{\sigma^2/\beta^2+xx'} \qquad \Rightarrow \qquad \widehat{\beta} = c'y = \frac{x'y}{\sigma^2/\beta^2+x'x}$

$E\left[\widehat{\beta}\right] = \beta c'x = \frac{\beta x'x}{\sigma^2/\beta^2+x'x} \qquad V\left[\widehat{\beta}\right] = \left(c'c\right)\sigma^2 = \frac{x'x\sigma^2}{\left(\sigma^2/\beta^2+x'x\right)^2}$

$MSE\left[\widehat{\beta}\right] = \left[\beta\left(c'x-1\right)\right]^2 + \left(c'c\right)\sigma^2 = \frac{x'x}{\sigma^2/\beta^2+xx'} - 1 = \frac{\frac{\sigma^4}{\beta^2}+x'x\sigma^2}{\left(x'x+\frac{\sigma^2}{\beta^2}\right)^2}$

$MSE\left[\beta\right] = \frac{\sigma^2}{x'x} \qquad \Rightarrow \qquad \frac{MSE\left[\widehat{\beta}\right]}{MSE\left[\beta\right]} = \frac{x'x}{x'x+\frac{\sigma^2}{\beta^2}} = \frac{x'x\frac{\beta^2}{\sigma^2}}{x'x\frac{\beta^2}{\sigma^2}+1} = \frac{\tau^2}{1+\tau^2} \to 0 \text{ as } \tau \to 0$

This can be attributed to either $\sigma^2 \to 0$ or $x'x \to \infty$. The first case means lack of uncertainty, which leads to equivalence of the estimators. The second case is the assimptotic case. Both estimators are consistent, hence they converge to the same limit.

4.3

No constant: $\widehat{\beta}_0 = \frac{x'y}{x'x} \qquad V\left[\widehat{\beta}_0\right] = \frac{\sigma^2}{x'x}$

With constant: $\widehat{\beta}_1 = \frac{Cov[x,y]}{Var[x]} \qquad V\left[\widehat{\beta}_1\right] = \frac{\sigma^2}{Var[x]}$

$V\left[\widehat{\beta}_0\right]/V\left[\widehat{\beta}_1\right] = \frac{Var[x]}{x'x} = \frac{x'x-n\overline{x}^2}{x'x} = 1 - \frac{n\overline{x}^2}{x'x} \leq 1$

Variance of the estimate grows if we insert a redundant variable.

4.4

$y_i = (\alpha+\lambda) + \beta x_i + (\varepsilon_i - \lambda) \qquad E\left[\varepsilon_i - \lambda\right] = 0$ and $(\varepsilon_i - \lambda)$ is independent of x_i.

CLR applies and gives unbiased estimators of β and $(\alpha+\lambda)$. If $\lambda > 0$ than the constant is not equal to α.

Exercise 1

(a) Any two time trends are linearly dependent. I.e. for any trends t_1 and t_2 there exist such g and h that $t_{2i} = g + ht_{1i}$ for all i. $X_2 = g1 + hX_1$

We first state that adding a constant to X obviously can't change the decomposition. Multiplication by a constant does not change the result either:

$e_2 = M_2 y = \left(I - X_2 \left(X_2' X_2\right)^{-1} X_2'\right) y = \left(I - hX_1 \left(hX_1' hX_1\right)^{-1} \left(hX_1'\right)\right) y =$

$= \left(I - X_1 \left(X_1' X_1\right)^{-1} \left(X_1'\right)\right) y = M_1 y = e_1$

Hence the residuals cannot change and so do the fitted values.

The parameter values do change with the change of both the origin and the unit of measurement: $X_t = a + bt_2 + e_t = a + b\left(g + ht_1\right) + e_t = a + bg + ht_1 + e_t = a' + b't_1 + e_t$.

.

(b) Correlation is -0.7375. The coefficients of the regression $x = a + by + \varepsilon$ are:

a=92.175 b=-1.3426

This implies (according to the model) that if the number of deaths without beer consumption is ~ 92.2 (much greater than any number in the sample) and that it decreases on average by ~ 1.34 per 1 barrel increase in beer consumption.

.

(c) Correlation is 0.2518. This can be due to an increase in the number of accidents, a decrease wealth and in equality, severe attitude to own children and so on, while a negative sign is almost impossible to explain.

.

(d) The coefficient b is 1.0116 which means that the trends were probably not interconnected (due to some outside reasons). The short-run correlation is positive, which is economically explicable (see (c)).

.

(e) A constant in a regression of detrended variables will appear to be zero. This will increase the variance of our results (see 4.3)

.

(f) The slope of the regression $x = a + ct + by + \varepsilon$ is b=1.0116. It is the same as in part (d).