Spring 1998 John Rust
Economics 551b 37 Hillhouse, Rm. 27

MIDTERM EXAM (first installment)
(Due: Monday, February 2, 1998)

QUESTION 1 Suppose the random variables $(\tilde{y}, \tilde{X})$ are multivariate normal, where the dimension of $\tilde{y}$ is (1×1) and the dimension of $\tilde{X}$ is $(1 \times K)$. Show that

$$E\{\tilde{y}|\tilde{X} = X\} = X\beta^*$$

where β^* is the $(K \times 1)$ vector of least squares coefficients:

$$\beta^* = \left[E\{\tilde{X}'\tilde{X}\}\right]^{-1} E\{\tilde{X}'\tilde{y}\}$$

In other words, you have shown that when the random variables are normally distributed the best nonlinear predictor $E\{\tilde{y}|\tilde{X}\}$ and the best linear predictor $\tilde{X}\beta^*$ coincide.

Hint: Show this result in several steps, following the path below:

A. **(Step 1)** Show that if $(\tilde{y}, \tilde{X})$ are any random variables where $E\{\tilde{X}'\tilde{X}\}$ is finite and nonsingular, and $E\{\tilde{X}'\tilde{y}\}$ is finite, then we can write

$$\tilde{y} = \tilde{X}\beta^* + \epsilon$$

where ϵ is a random variable satisfying $E\{\tilde{X}'\epsilon\} = \mathbf{0}$ where $\mathbf{0}$ is a $(K \times 1)$ vector of 0s.

B. **(Step 2)** Use the result in part A to show that β^* can also be written as

$$\beta^* = \left[\mathrm{cov}(\tilde{X}, \tilde{X})\right]^{-1} \left[\mathrm{cov}(\tilde{X}', \tilde{y}) + E\{\epsilon\}E\{\tilde{X}'\}\right]$$

(hint) note that if $\tilde{x}_j$ is the j^{th} random variable in the vector $\tilde{X}$ we can write

$$\begin{aligned} \mathrm{cov}(\tilde{y}, \tilde{x}_j) &= \mathrm{cov}(\tilde{x}_1\beta_1^* + \cdots \tilde{x}_K\beta_K^* + \epsilon, \tilde{x}_j) \\ &= \mathrm{cov}(\tilde{x}_1, \tilde{x}_j)\beta_1^* + \cdots \mathrm{cov}(\tilde{x}_K, \tilde{x}_j)\beta_K^* \end{aligned}$$

C. Show that the above result implies that when $E\{\epsilon\} = 0$ or $E\{\tilde{X}'\} = \mathbf{0}$, we have two equivalent expressions for the OLS coefficients β^*:

$$\begin{aligned} \beta^* &= \left[E\{\tilde{X}'\tilde{X}\}\right]^{-1} E\{\tilde{X}'\tilde{y}\} \\ &= \left[\mathrm{cov}(\tilde{X}, \tilde{X})\right]^{-1} \mathrm{cov}(\tilde{X}', \tilde{y}). \end{aligned}$$

D. Now, fill in the details of Greene's exposition of the marginal and conditional distributions of the multivariate normal in section 3.10.1 of his book, and show that if $(\tilde{y}, \tilde{X}')$ has a joint multivariate normal distribution, then the conditional density $f(y|X)$ (i.e. the conditional density of $\tilde{y}$ given that $\tilde{X} = X$) is normally distributed, $N(\mu_X, \Sigma_X)$, where

$$\mu_X = E\{\tilde{y}|\tilde{X} = X\} = X\beta^*$$

and

$$\Sigma_X = \mathrm{var}(\tilde{y}|\tilde{X} = X) = \mathrm{var}(\tilde{y}) - \beta^{*\prime}\left[\mathrm{cov}(\tilde{X}, \tilde{X})\right]\beta^*.$$