Section 9: "Optimal" estimators

ARE 210

November 2, 2017

1) Prove (CRI Lemma).

Taking the first order condition of the minimization problem, $\mathbf{E}[(Y - Z'\beta(P))Z] = 0$. Simplifying under our assumptions yields $\beta(P) = \mathbf{E}[ZZ']^{-1}\mathbf{E}[ZY]$. Note that this just says that OLS recovers the minimum variance ("best") linear estimator of Y (or its conditional expectation with respect to Z).

Second, let $\mu_L(Z) = Z'\beta(P)$. We want to show $V[Y] \geq V[\mu_L(Z)]$. First, note, $V[Y] = V[\mu_L(Z) + (Y - \mu_L(Z))] = V[\mu_L(Z)] + V[Y - \mu_L(Z)] + 2\text{Cov}[\mu_L(Z), Y - \mu_L(Z)]$. Note that $V[Y - \mu_L(Z)] \geq 0$, so it suffices to show $\text{Cov}[\mu_L(Z), Y - \mu_L(Z)] = 0$.

To show this, first note that if $\mathbf{E}[Y - \mu_L(Z)] = 0$, then the first order condition to the minimization problem implies $\text{Cov}[\mu_L(Z), Y - \mu_L(Z)] = 0$. Next, note that if Z contains a constant, and if $\mathbf{E}[Y - \mu_L(Z)] \neq 0$, we can improve the minimization by adjusting the term on the constant, a contradiction, so $\mathbf{E}[Y - \mu_L(Z)] = 0$.

Together, these give (1) $\beta(P) = \mathbf{E}[ZZ']^{-1}\mathbf{E}[ZY]$, and (2) $V[Y] \geq V[\mu_L(Z)] = \beta(P)'V[Z]\beta(P)$.

2) (**Identification in Exponential Families)** Consider the (canonical) exponential family $\left\{p\left(y,\theta\right) : \theta \in \Theta \subset \mathbb{R}^d\right\}$ where

$$p\left(y,\theta\right) = h\left(y\right)\exp\left\{\sum_{k=1}^{d}\theta_k T_k\left(y\right) - A\left(\theta\right)\right\}$$

and suppose that $A\left(\theta\right)$ is continuously differentiable for all $\theta \in \Theta$. Show that θ is identified if the $d \times d$ matrix

$$I\left(\theta^*\right) = \mathbb{E}_{\theta^*}\left(\frac{d\log p\left(y,\theta^*\right)}{d\theta}\frac{d\log p\left(y,\theta^*\right)'}{d\theta}\right)$$

is non-singular for every $\theta^* \in \Theta$ (where $\frac{d\log p(y,\theta^*)}{d\theta} = \left.\frac{d\log p(y,\theta)}{d\theta}\right|_{\theta=\theta^*}$)

We assume that $I(\theta^*) = \mathbf{E}_{\theta^*}\left[\frac{d\log p(y,\theta^*)}{d\theta}\frac{d\log p(y,\theta^*)'}{d\theta}\right]$ is non-singular $\forall\ \theta^* \in \Theta$. Suppose θ is not identified. Then, $\exists\ \theta_1 \neq \theta_2$ such that $p(y,\theta_1) = p(y,\theta_2)$ for all y. Assuming

$h(y) \neq 0$, this implies $A(\theta_1) - A(\theta_2) = \sum(\theta_{2k} - \theta_{1k})T_k(y)$. Applying the mean value theorem, there exists $\theta^* = \alpha\theta_1 + (1-\alpha)\theta_2$ such that $A(\theta_1) - A(\theta_2) = (\theta_1 - \theta_2)'\frac{dA(\theta^*)}{d\theta}$. Substituting into our expression for the derivative of $\log p(y, \theta)$ yields

$$(\theta_1 - \theta_2)'\frac{d \log p(y, \theta^*)}{d\theta} = 0$$

for all y. This implies $V_{\theta^*}\left[(\theta_1 - \theta_2)'\frac{d \log p(y,\theta^*)}{d\theta}\right] = 0$.

Next, note that $\mathbf{E}_{\theta^*}\left[\frac{d \log p(y,\theta^*)}{d\theta}\right] = 0$.[1] Therefore, $V_{\theta^*}\left[\frac{d \log p(y,\theta^*)}{d\theta}\right] = I(\theta^*)$. Since this matrix is non-singular, it must be the case that $V_{\theta^*}\left[(\theta_1 - \theta_2)'\frac{d \log p(y,\theta^*)}{d\theta}\right] \neq 0$, a contradiction. Therefore θ is identified.

3) Prove the Hausman principle.

($\Rightarrow$, in class notes) First, assume W is UMVUE of $\tau(\theta)$. Consider an unbiased estimator U of 0. Let $W_a = W + aU$, which is trivially unbiased for $\tau(\theta)$. Note that $V_\theta[W_a] = V_\theta[W] + a^2 V_\theta[U] + 2a\text{Cov}_\theta[W, U]$. Fix some θ such that $\text{Cov}_\theta[W, U] \neq 0$, then we can pick a such that $V_\theta[W_a] < V_\theta[W]$.

($\Leftarrow$) Assume W is unbiased for $\tau(\theta)$, and that W is uncorrelated with all unbiased estimators of 0. Suppose $\exists\ W_a$ which is unbiased for $\tau(\theta)$ and has lower variance than W for some θ_0. Then, define $U = W - W_a$, which is clearly an unbiased estimator of $\tau(\theta)$. Note that $\text{Cov}_{\theta_0}[U, W] = V_{\theta_0}[W] - \text{Cov}_{\theta_0}[W, W_a] > 0$, since $V_{\theta_0}[W] > V_{\theta_0}[W_a]$. This is a contradiction, so it must be the case that W is UMVUE.

Note that one way to interpret this result is that Lehmann-Scheffe is basically an if and only if - in other words, given ϕ UMVUE for $g(\theta)$, ϕ is complete and sufficient for $g(\theta)$. The Hausman principle is basically showing that any ancillary statistic is uncorrelated with ϕ (implying it's complete for $g(\theta)$), and by assuming it's unbiased for $g(\theta)$ we already know that it's sufficient, and that if ϕ is unbiased for $g(\theta)$, ϕ is UMVUE if and only if ϕ is complete and sufficient for $g(\theta)$, again loosely speaking.

[1] This is true because θ^* maximizes the likelihood function when it is the true parameter.