Addendum on Likelihood Theory: The Cramer-Rao Bound

We are building on the argumentation that resulted in the derivation of the information matrix equality to derive an asymptotic version of the Cramer-Rao bound:

Suppose that τ_n is an estimator of θ_0 such that $\lim_{n \to +\infty} \mathbb{E}_\theta\left(n^{1/2}(\tau_n - \theta)\right) = 0_{k \times 1}$, $\forall \theta \in B_{\theta_0}$ (it is locally at θ_0 asymptotically unbiased) and such that $V^* := \lim_{n \to +\infty} \mathbb{E}_\theta\left[n(\tau_n - \theta)(\tau_n - \theta)'\right]$ exists $\forall \theta \in B_{\theta_0}$. Suppose also that the likelihood function $\ell_n(\theta)$ is such that

$$\sup_{\theta \in B_{\theta_0}} \mathbb{E}_\theta \left[\sup_{\theta \in B_{\theta_0}} \left\| \frac{\partial^2 l_n(\theta)}{\partial \theta \partial \theta'} \right\| \right] = O_p(1).$$ Then it

can be proven (among others via an argument that uses the concept of dominated convergence and the definition of the derivative)

that: $\forall \theta \in \operatorname{Int} B_\theta$,

$$\frac{\partial}{\partial \theta'} \lim_{n \to +\infty} \mathbb{E}_\theta \left(n^{1/2} (\tilde{\tau}_n - \theta) \right) = \frac{\partial O_{k \times 1}}{\partial \theta'} \Rightarrow$$

$$\lim_{n \to +\infty} \frac{\partial}{\partial \theta'} \int n^{1/2} (\tilde{\tau}_n - \theta) f(y_n, \theta) \, dy_n = O_{k \times k} \Rightarrow$$

$$\lim_{n \to +\infty} \left[\int - n^{1/2} \frac{\partial \theta}{\partial \theta'} f(y_n, \theta) \, dy_n + \int n^{1/2} (\tilde{\tau}_n - \theta) \frac{\partial f(y_n, \theta)}{\partial \theta'} \, dy_n \right]$$

$$= O_{k \times k} \Rightarrow$$

$$\lim_{n\to+\infty}\Big[-n^{1/2}\underbrace{I_{k\times k}}_{\text{the identity } k\times k \text{ matrix}}\underbrace{\int f(y_n,\theta)dy_n}_{=1} +$$

$$+\int n^{1/2}(\tilde{\theta}_n-\theta)\frac{\partial \ln f(y_n,\theta)}{\partial\theta'} f(y_n,\theta)dy_n\Big] = 0_{k\times k}$$

$$\Leftrightarrow \lim_{n\to+\infty}\Big[-n^{1/2}I_{k\times k}+\mathbb{E}_\theta\Big[n^{1/2}(\tilde{\theta}_n-\theta)\,n\frac{\partial l_n(\theta)}{\partial\theta'}\Big]\Big]=0_{k\times k}$$

CMT

$$\Rightarrow \lim_{n\to+\infty}\frac{1}{n^{1/2}}\Big[-n^{1/2}I_{k\times k}+\mathbb{E}\Big[n^{1/2}(\tilde{\theta}_n-\theta)\,n^{1/2}\frac{\partial l_n(\theta)}{\partial\theta'}\Big]\Big]=0_{k\times k}$$

$$\Rightarrow \lim_{n\to+\infty}\Big[-\underbrace{I_{k\times k}}_{\text{independent of } n}+\mathbb{E}\Big[n^{1/2}(\tilde{\theta}_n-\theta)\,n^{1/2}\frac{\partial l_n(\theta)}{\partial\theta'}\Big]\Big]=0_{k\times k}$$

$$\Rightarrow \boxed{\lim_{n\to+\infty}\mathbb{E}_\theta\Big[n^{1/2}(\tilde{\theta}_n-\theta)\,n^{1/2}\frac{\partial l_n(\theta)}{\partial\theta'}\Big] = I_{k\times k},\ \forall\theta\in\mathcal{B}_\theta}$$

which says that the limiting P_θ-covariance matrix between $n^{1/2}[\tilde{\theta}_n-\theta]$ and the $n^{1/2}$-scaled score, exists and equals the identity.

[Remember that in our framework we have that $\mathrm{Var}_{\theta_0}\left[n^{1/2}\frac{\partial l_n(\theta_0)}{\partial\theta}\right] = -\mathbb{E}\left[\frac{\partial^2 l_n(\theta_0)}{\partial\theta\partial\theta'}\right]$

and the r.h.s. is independent of n]

Consider the random vector:

$$u_n := \begin{pmatrix} n^{1/2}(z_n-\theta_0) \\ n^{1/2}\frac{\partial l_n(\theta_0)}{\partial\theta}\end{pmatrix} \quad \text{and notice that}$$

$$\lim_{n\to+\infty}\mathrm{Var}_{\theta_0}(u_n) = \lim_{n\to+\infty}\mathbb{E}_{\theta_0}\left\{\begin{bmatrix} n^{1/2}(z_n-\theta_0) \\ n^{1/2}\frac{\partial l_n(\theta_0)}{\partial\theta}\end{bmatrix}\begin{bmatrix} n^{1/2}(z_n-\theta_0) \\ n^{1/2}\frac{\partial l_n(\theta_0)}{\partial\theta}\end{bmatrix}'\right\}$$

$$\underbrace{-\lim_{n\to+\infty}\mathbb{E}_{\theta_0}\begin{bmatrix} n^{1/2}(z_n-\theta_0) \\ n^{1/2}\frac{\partial l_n(\theta_0)}{\partial\theta}\end{bmatrix}\left(\mathbb{E}_{\theta_0}\begin{bmatrix} n^{1/2}(z_n-\theta_0) \\ n^{1/2}\frac{\partial l_n(\theta_0)}{\partial\theta}\end{bmatrix}\right)'}_{(*)}$$

and notice first that $(*) \stackrel{\text{cnt}}{=} A\cdot A'$ where

$$\lambda = \lim_{n\to+\infty} \mathbb{E}_{\theta_0}\begin{pmatrix} n^{1/2}(\hat{\theta}_n-\theta_0) \\ n^{1/2}\dfrac{\partial l_n(\theta_0)}{\partial\theta}\end{pmatrix} =$$

$$= \begin{pmatrix} \lim_{n\to+\infty} \mathbb{E}_{\theta_0}(\hat{\theta}_n-\theta_0) \\ \lim_{n\to+\infty} n^{1/2}\dfrac{\partial l_n(\theta_0)}{\partial\theta}\end{pmatrix} \overset{\text{why?}}{=} \begin{pmatrix} 0_{k\times 1} \\ 0_{k\times 1}\end{pmatrix} =$$

$$= 0_{2k\times 1}, \quad \text{hence } (*) = 0_{2k\times 1}\, 0_{1\times 2k} =$$

$$= 0_{2k\times 2k}.$$

→ "variance" terms

→ "covariance" terms

Thus $\lim_{n\to+\infty} \mathrm{Var}_{\theta_0} M_n =$

$$\begin{pmatrix} \lim_{n\to+\infty} \mathbb{E}_{\theta_0}\left[n(\hat{\theta}_n-\theta_0)(\hat{\theta}_n-\theta_0)'\right] & \lim_{n\to+\infty} \mathbb{E}_{\theta_0}\left(n^{1/2}(\hat{\theta}_n-\theta_0)\, n^{1/2}\dfrac{\partial l_n(\theta_0)}{\partial\theta'}\right) \\ \lim_{n\to+\infty} \mathbb{E}_{\theta_0}\left(n^{1/2}(\hat{\theta}_n-\theta_0)'\, n^{1/2}\dfrac{\partial l_n(\theta_0)}{\partial\theta}\right) & \lim_{n\to+\infty} \mathrm{Var}_{\theta_0}\left(n^{1/2}\dfrac{\partial l_n(\theta_0)}{\partial\theta}\right)\end{pmatrix}$$

But $\lim_{n\to+\infty} \mathbb{E}_{\theta_0}\left(n^{1/2}(z_n-\theta_0)' n^{1/2}\frac{\partial l_n(\theta_0)}{\partial\theta}\right) \overset{\text{CLT}}{=}$

$$\left(\lim_{n\to+\infty} \mathbb{E}_{\theta_0}\left[n^{1/2}(z_n-\theta_0)\, n^{1/2}\frac{\partial l_n(\theta_0)}{\partial\theta'}\right]\right)' = I_{k\times k}'$$

$= I_{k\times k}$, and $\lim_{n\to+\infty} \mathbb{E}_{\theta_0}\left[(z_n-\theta_0)(z_n-\theta_0)'\right] = \lim_{n\to+\infty} \mathrm{Var}_{\theta_0}\left(n^{1/2}(z_n-\theta_0)\right)$.

Hence,

$$\lim_{n\to+\infty} \mathrm{Var}_{\theta_0}(\Delta_n) = \boxed{\begin{pmatrix} \lim_{n\to+\infty} \mathrm{Var}_{\theta_0}\left(n^{1/2}(z_n-\theta_0)\right) & I_{k\times k} \\ I_{k\times k} & -E\left(\frac{\partial l_n(\theta_0)}{\partial\theta\partial\theta'}\right) \end{pmatrix}}$$

But, as a limit of psd matrices, the limiting variance has to be psd (why? use the fact that Frobenius-norm convergence implies convergence of eigenvalues), and thereby the [redacted] matrix is psd.

Now consider the the $2k \times k$ matrix,

a stacked matrix

$\begin{pmatrix} I_{k\times k} \\ -\left(\mathbb{E}\left[\frac{\partial^2 l_i(\theta_0)}{\partial\theta\,\partial\theta'}\right]\right)^{-1} \end{pmatrix}$ and notice that the p.d. of

[redacted] implies that the $k \times k$ matrix

quadratic form

$$\begin{pmatrix} I_{k\times k} \\ \left(\mathbb{E}\left[\frac{\partial^2 l_i(\theta_0)}{\partial\theta\,\partial\theta'}\right]\right)^{-1} \end{pmatrix}' \; [redacted] \; \begin{pmatrix} I_{k\times k} \\ \left(\mathbb{E}\left[\frac{\partial^2 l_i(\theta_0)}{\partial\theta\,\partial\theta'}\right]\right)^{-1} \end{pmatrix}$$ is also p.d.

(why? – use the definition along with $z_{k\times 1}$ vectors of the form $\begin{pmatrix} I_{k\times k} \\ \left(\mathbb{E}\left[\frac{\partial^2 l_i(\theta_0)}{\partial\theta\,\partial\theta'}\right]\right)^{-1} \end{pmatrix} x_{k\times 1}$ to produce this)

But the previous quadratic form, equals

($\left(\mathbb{E}\left[\frac{\partial^2 l_i(\theta_0)}{\partial\theta\,\partial\theta'}\right]\right)^{-1}$ is symmetric – why?)

$$\underbrace{\left(I_{k\times k},\ \left(\mathbb{E}\left[\frac{\partial^2 l_n(\theta_0)}{\partial\theta\partial\theta'}\right]\right)^{-1}\right)}_{k\times 2k}$$ [redacted] $$\underbrace{}_{2k\times 2k} \quad \underbrace{\begin{pmatrix} I_{k\times k} \\ \left(\mathbb{E}\left[\frac{\partial^2 l_n(\theta_0)}{\partial\theta\partial\theta'}\right]\right)^{-1}\end{pmatrix}}_{2k\times k} =$$

$$= \left(I_{k\times k},\ \left(\mathbb{E}\left[\frac{\partial^2 l_n(\theta_0)}{\partial\theta\partial\theta'}\right]\right)^{-1}\right) \begin{pmatrix} \lim\limits_{n\to+\infty} \mathrm{Var}_{\theta_0}\left(n^{1/2}(z_n-\theta_0)\right) & I_{k\times k} \\ I_{k\times k} & -\mathbb{E}_{\theta_0}\left(\frac{\partial^2 l_n(\theta_0)}{\partial\theta\partial\theta'}\right) \end{pmatrix}$$

$$\times \begin{pmatrix} I_{k\times k} \\ \left(\mathbb{E}\left[\frac{\partial^2 l_n(\theta_0)}{\partial\theta\partial\theta'}\right]\right)^{-1}\end{pmatrix} =$$

(Remember $I_{k\times k}$ is idempotent)

$$= \left(\lim_{n\to+\infty} \mathrm{Var}_{\theta_0}\left(n^{1/2}(z_n-\theta_0)\right) + \left(\mathbb{E}\left[\frac{\partial^2 l_n(\theta_0)}{\partial\theta\partial\theta'}\right]\right)^{-1},\ \underbrace{I^2_{k\times k} - I_{k\times k}}_{=0_{k\times k}}\right)$$

$$\times \begin{pmatrix} I_{k\times k} \\ \left(\mathbb{E}\left[\frac{\partial^2 l_n(\theta_0)}{\partial\theta\partial\theta'}\right]\right)^{-1}\end{pmatrix} =$$

$$\lim_{n\to+\infty} \mathrm{Var}_{\theta_0}\left(n^{1/2}(z_n-\theta_0)\right) + \left(\mathbb{E}\,\frac{\partial^2 l_n(\theta_0)}{\partial\theta\partial\theta'}\right)^{-1}$$

$$+\ 0_{k\times k}\cdot\left(\mathbb{E}\,\frac{\partial^2 l_n(\theta_0)}{\partial\theta\partial\theta'}\right)^{-1}$$

$$= \lim_{n\to+\infty} \mathrm{Var}_{\theta_0}\left(n^{1/2}(\tau_n-\theta_0)\right) + \left(\mathbb{E}_{\theta_0}\left(\frac{\partial^2 \ell_n(\theta_0)}{\partial\theta\partial\theta'}\right)\right)^{-1}$$

Hence we have established that

for any τ_n, such that $\lim_{n\to+\infty} \mathbb{E}_\theta\left(n^{1/2}(\tau_n-\theta)\right) = 0_{K\times 1}$, and $\lim_{n\to+\infty} \mathrm{Var}_\theta\left(n^{1/2}(\tau_n-\theta)\right)$ exists

$\forall\theta$ on some neighborhood of θ_0,

$$\lim_{n\to+\infty} \mathrm{Var}_{\theta_0}\left(n^{1/2}(\tau_n-\theta_0)\right) + \left(\mathbb{E}_{\theta_0}\left(\frac{\partial^2 \ell_n(\theta_0)}{\partial\theta\partial\theta'}\right)\right)^{-1}$$

is spd and thereby, the inverse of the information matrix is the minimal limiting variance achievable in this framework.