Extremum Estimators – Analytical Exercises

Ryan Benschop

1. Kullback-Leibler information inequality

Claim: Let $\{f(y|x;\theta)\}_{\theta\in\Theta}$ *be a class of conditional density functions such that, for each* $\theta \in \Theta$, $\mathbb{E}[\ln f(y|x;\theta)]$ *exists and is finite. Then, for any* $\theta, \theta_0 \in \Theta$*:*

$$\mathbb{P}[f(y|x;\theta) \neq f(y|x;\theta_0)] > 0 \implies \mathbb{E}[\ln f(y|x;\theta)] < \mathbb{E}[\ln f(y|x;\theta_0)]$$

where the expectation is taken with respect to the density $f(y|x;\theta_0)$*.*

Proof. Suppose $\mathbb{P}[f(y|x;\theta) \neq f(y|x;\theta_0)] > 0$. Notice that the event $f(y|x;\theta) \neq f(y|x;\theta_0)$ is a subset of the event $f(y|x;\theta)/f(y|x;\theta_0) \neq 1$. Hence, the probability measure of the event $f(y|x;\theta)/f(y|x;\theta_0) \neq 1$ is greater than zero. Thus, by Jensen's inequality:

$$\mathbb{E}\left[\ln\left(\frac{f(y|x;\theta)}{f(y|x;\theta_0)}\right)\right] < \ln\left(\mathbb{E}\left[\frac{f(y|x;\theta)}{f(y|x;\theta_0)}\right]\right)$$

where these expectations are taken with respect to the density $f(y|x;\theta_0)$. By the law of iterated expectations:

$$\begin{aligned}
\mathbb{E}\left[\frac{f(y|x;\theta)}{f(y|x;\theta_0)}\right] &= \mathbb{E}_x\left[\mathbb{E}\left[\frac{f(y|x;\theta)}{f(y|x;\theta_0)}\middle| x\right]\right] \\
&= \mathbb{E}_x\left[\int \frac{f(y|x;\theta)}{f(y|x;\theta_0)} f(y|x;\theta_0)dy\right] \\
&= \mathbb{E}_x\left[\int f(y|x;\theta)dy\right] \\
&= \mathbb{E}_x[1] = 1
\end{aligned}$$

Since $\ln(1) = 0$, the above inequality becomes:

$$\mathbb{E}\left[\ln\left(\frac{f(y|x;\theta)}{f(y|x;\theta_0)}\right)\right] < 0$$

Noting that $\ln(f(y|x;\theta)/f(y|x;\theta_0)) = \ln(f(y|x;\theta)) - \ln(f(y|x;\theta_0))$, we have:

$$\mathbb{E}[\ln f(y|x;\theta)] < \mathbb{E}[\ln f(y|x;\theta_0)]$$

as desired. $\square$

2. (Information matrix equality)

Claim: Let $\{f(y|x;\theta)\}_{\theta\in\Theta}$ be a class of conditional densities. Let $\theta_0 \in \Theta$. Then:

$$-\mathbb{E}\left[\frac{\partial^2 \ln f(y|x;\theta_0)}{\partial\theta\partial\theta'}\right] = \mathbb{E}\left[\left(\frac{\partial \ln f(y|x;\theta_0)}{\partial\theta_0}\right)\left(\frac{\partial \ln f(y|x;\theta_0)}{\partial\theta_0}\right)'\right]$$

where the expectation is taken with respect to the density $f(y|x;\theta_0)$.

Proof. Notice, since $f(y|x;\theta)$ is a density:

$$\int f(y|x;\theta)dy = 1$$

Assuming conditions allowing the interchange of differentiation and integration:

$$\frac{\partial}{\partial\theta}\left[\int f(y|x;\theta)dy\right] = 0 \implies \int \frac{\partial}{\partial\theta} f(y|x;\theta)dy = 0$$

Differentiating $\ln f(y|x;\theta)$ with respect to θ, we find:

$$\frac{\partial}{\partial\theta}[\ln f(y|x;\theta)] = f(y|x;\theta)^{-1}\frac{\partial}{\partial\theta} f(y|x;\theta) \implies \frac{\partial}{\partial\theta} f(y|x;\theta) = f(y|x;\theta)\frac{\partial}{\partial\theta}[\ln f(y|x;\theta)]$$

So, the previous equality can be written as:

$$\int f(y|x;\theta)\frac{\partial}{\partial\theta}[\ln f(y|x;\theta)] = 0$$

Differentiating with respect to θ, we have:

$$\int \frac{\partial^2 \ln f(y|x;\theta)}{\partial\theta\partial\theta'} f(y|x;\theta) + \frac{\partial \ln f(y|x;\theta)}{\partial\theta}\frac{\partial f(y|x;\theta)}{\partial\theta}' dy = 0$$
$$\int \frac{\partial^2 \ln f(y|x;\theta)}{\partial\theta\partial\theta'} f(y|x;\theta) + \frac{\partial \ln f(y|x;\theta)}{\partial\theta}\frac{\partial \ln f(y|x;\theta)}{\partial\theta}' f(y|x;\theta)dy = 0$$

Rearranging the expression and evaluating it at $\theta = \theta_0$, noting that in this case the integral gives the relevant expected values with respect to the density $f(y|x;\theta_0)$, we have, as desired:

$$-\mathbb{E}\left[\frac{\partial^2 \ln f(y|x;\theta_0)}{\partial\theta\partial\theta'}\right] = \mathbb{E}\left[\left(\frac{\partial \ln f(y|x;\theta_0)}{\partial\theta}\right)\left(\frac{\partial \ln f(y|x;\theta_0)}{\partial\theta}\right)'\right]$$

$\square$