ECON675 – Assignment 3

Anirudh Yadav

October 23, 2018

Contents

1 Non-linear least squares

1.1 Identifiability

This is a standard M-estimation problem. The parameter vector $\boldsymbol{\beta}_0$ is assumed to solve the population problem

$$\boldsymbol{\beta}_0 = \arg\min_{\beta \in \mathbb{R}^d} \mathbb{E}[(y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}))^2].$$

For $\boldsymbol{\beta}_0$ to be identified, it must be the *unique* solution to the above population problem (i.e. the unique minimizer). In math, this means for all $\epsilon > 0$ and for some $\delta > 0$:

$$\sup_{||\beta - \beta_0|| > \epsilon} M(\boldsymbol{\beta}) \geq M(\boldsymbol{\beta}_0) + \delta$$

where $M(\boldsymbol{\beta}) = \mathbb{E}[(y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}))^2]$. Of course $\boldsymbol{\beta}_0$ can be written in closed form if $\mu(\cdot)$ is linear. In this case, we know that

$$\boldsymbol{\beta}_0 = \mathbb{E}[\boldsymbol{x}_i\boldsymbol{x}_i']^{-1}\mathbb{E}[\boldsymbol{x}_i y_i].$$

1.2 Asymptotic normality

The M-estimator is asymptotically normal if:

1. $\hat{\boldsymbol{\beta}} \to_p \boldsymbol{\beta}_0$
2. $\boldsymbol{\beta}_0 \in int(B)$ and $m(\boldsymbol{x}_i, \boldsymbol{\beta}) \equiv (y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}))^2$ is 3 times continuously differentiable.

3. $\Sigma_0 = \mathbb{V}[\frac{\partial}{\partial\beta} m(\boldsymbol{x}_i; \beta_0)] < \infty$ and $H_0 = \mathbb{E}[\frac{\partial^2}{\partial\beta\partial\beta'} m(\boldsymbol{x}_i; \beta_0)]$ is full rank (and therefore invertible).

Now, the FOC for the M-estimation problem is

$$0 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}))\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}))\boldsymbol{x}_i$$

where $\dot{\mu} = \frac{\partial}{\partial\beta}\mu(\boldsymbol{x}_i'\boldsymbol{\beta})$. So, we've converted the M-estimation problem into a Z-estimation problem. Then we can use the standard asymptotic normality result to arrive at a precise form of the asymptotic variance:

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \rightarrow_d \mathcal{N}(0, H_0^{-1}\Sigma_0 H_0^{-1}).$$

Now, taking the second derivative gives the Hessian

$$\begin{aligned}
H_0 &= \mathbb{E}[\frac{\partial^2}{\partial\beta\partial\beta'} m(\boldsymbol{x}_i; \beta_0)] \\
&= \mathbb{E}\left[-\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}_0))\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}_0))\boldsymbol{x}_i\boldsymbol{x}_i' + (y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0))\ddot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}_0))\boldsymbol{x}_i\boldsymbol{x}_i'\right] \\
&= -\mathbb{E}[\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}_0))^2\boldsymbol{x}_i\boldsymbol{x}_i']
\end{aligned}$$

by LIE. And, the variance of the score is

$$\begin{aligned}
\Sigma_0 &= \mathbb{V}[\frac{\partial}{\partial\beta} m(\boldsymbol{x}_i; \beta_0)] \\
&= \mathbb{E}\left[(y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0))^2\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}_0))^2\boldsymbol{x}_i\boldsymbol{x}_i'\right] \\
&= \mathbb{E}[\sigma^2(\boldsymbol{x}_i)\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}_0))^2\boldsymbol{x}_i\boldsymbol{x}_i']
\end{aligned}$$

again by LIE. Then we have the asymptotic variance

$$\boldsymbol{V}_0 = H_0^{-1}\Sigma_0 H_0^{-1}.$$

1.3 Variance estimator under heteroskedasticity

Under heteroskedasticity we can use the sandwich variance estimator

$$\widehat{\boldsymbol{V}}_{HC} = \hat{H}^{-1}\hat{\Sigma}\hat{H}^{-1},$$

where

$$\begin{aligned}
\hat{H} &= \frac{1}{n}\sum_{i=1}^{n}\dot{\mu}(\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}))^2\boldsymbol{x}_i\boldsymbol{x}_i' \\
\hat{\Sigma} &= \frac{1}{n}\sum_{i=1}^{n}\hat{e}_i^2\dot{\mu}(\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}))^2\boldsymbol{x}_i\boldsymbol{x}_i'
\end{aligned}$$

Now, to get an asymptotically valid CI for $||\boldsymbol{\beta}_0||^2$ we need to use the Delta Method. First, note that:

$$\begin{aligned} ||\boldsymbol{\beta}_0||^2 &= \boldsymbol{\beta}_0'\boldsymbol{\beta}_0 \\ \implies \frac{\partial}{\partial \beta}||\boldsymbol{\beta}_0||^2 &= 2\boldsymbol{\beta}_0 \end{aligned}$$

Then, using the Delta Method

$$\begin{aligned} \sqrt{n}(||\hat{\boldsymbol{\beta}}||^2 - ||\boldsymbol{\beta}_0||^2) &\to_d 2\boldsymbol{\beta}_0\mathcal{N}(0, \boldsymbol{V}_0) \\ &= \mathcal{N}(0, 4\boldsymbol{\beta}_0'\boldsymbol{V}_0\boldsymbol{\beta}_0) \end{aligned}$$

Thus, an asymptotically valid 95% CI for $||\boldsymbol{\beta}_0||^2$ is

$$CI_{95} = \left[\hat{\boldsymbol{\beta}} - 1.96\sqrt{\frac{4\hat{\boldsymbol{\beta}}'\widehat{\boldsymbol{V}}_{HC}\hat{\boldsymbol{\beta}}}{n}}, \hat{\boldsymbol{\beta}} + 1.96\sqrt{\frac{4\hat{\boldsymbol{\beta}}'\widehat{\boldsymbol{V}}_{HC}\hat{\boldsymbol{\beta}}}{n}}\right]$$

1.4 Variance estimator under homoskedasticity

Using the above results, under homoskedasticity, the asymptotic variance collapses to

$$\begin{aligned} \boldsymbol{V}_0 &= \mathbb{E}[\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}_0))^2\boldsymbol{x}_i\boldsymbol{x}_i']^{-1}\sigma^2\mathbb{E}[\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}))^2\boldsymbol{x}_i\boldsymbol{x}_i']\mathbb{E}[\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}_0))^2\boldsymbol{x}_i\boldsymbol{x}_i']^{-1} \\ &= \sigma^2\mathbb{E}[\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}_0))^2\boldsymbol{x}_i\boldsymbol{x}_i']^{-1} \end{aligned}$$

The variance estimator is now takes a simpler form

$$\widehat{\boldsymbol{V}}_{HO} = \hat{\sigma}^2\hat{H}^{-1}$$

where $\hat{H}$ is the same as above and

$$\hat{\sigma}^2 = \frac{1}{n-d}\sum_{i=1}^{n}(y_i - \mu(\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}))^2$$

Then, as above, the asymptotically valid 95% CI for $||\boldsymbol{\beta}_0||^2$ is

$$CI_{95} = \left[\hat{\boldsymbol{\beta}} - 1.96\sqrt{\frac{4\hat{\boldsymbol{\beta}}'\widehat{\boldsymbol{V}}_{HO}\hat{\boldsymbol{\beta}}}{n}}, \hat{\boldsymbol{\beta}} + 1.96\sqrt{\frac{4\hat{\boldsymbol{\beta}}'\widehat{\boldsymbol{V}}_{HO}\hat{\boldsymbol{\beta}}}{n}}\right].$$