Econ 675 Assignment 1

Nathan Mather[*]

October 22, 2018

Contents

1 Question 1: Non-linear Least Squares

1.1 Q1 Part 1

The general non-linear least squares estimator is

$$\hat{\boldsymbol{\beta}}_n = \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^d} \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu(\boldsymbol{x_i'\beta}))^2$$

Now for $\boldsymbol{\beta}_0 = \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^d} \mathrm{E}[(y_i - \mu(\boldsymbol{x_i'\beta}))^2]$ to be identifiable we need:

$$\boldsymbol{\beta}_0 = \boldsymbol{\beta}_0^*$$

$$\iff \boldsymbol{\beta}_0^* = \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^d} \mathrm{E}[(y_i - \mu(\boldsymbol{x_i'\beta}))^2]$$

To find this start by noting that

$$\mathrm{E}[(y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}))^2] = \mathrm{E}[(y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0) + \mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0) - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}))^2]$$

$$= \mathrm{E}[(y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0))^2] + \mathrm{E}[(\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0) - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}))^2] + 2\mathrm{E}[(y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0))(\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0) - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}))]$$

$$= \mathrm{E}[(y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0))^2] + \mathrm{E}[(\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0) - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}))^2]$$

The last equality comes from the last term being zero by iterated expectations. I show this below.

[*]Shouts out to Ani, Paul, Tyler, Erin, Caitlin and others for all the help with this

$$\mathrm{E}[(y_i-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0))(\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0)-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}))]=\mathrm{E}[\mathrm{E}[(y_i-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0))(\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0)-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}))]|\boldsymbol{x}_i]$$

$$=\mathrm{E}[(\mathrm{E}[y_i|\boldsymbol{x}_i]-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0))(\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0)-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}))]=0$$

Using this fact we have that

$$\mathrm{E}[(y_i-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}))^2]=\mathrm{E}[(y_i-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0))^2]+\mathrm{E}[(\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0)-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}))^2]\geq\mathrm{E}[(y_i-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0))^2]$$

$$\forall\boldsymbol{\beta}\neq\boldsymbol{\beta}_0$$

This is strictly greater than unless $\exists\boldsymbol{\beta}\neq\boldsymbol{\beta}_0$ such that $\mathrm{E}[(\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0)-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}))^2]=0$ Thus this give us an identification condition that $\mathrm{E}[(\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0)-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}))^2]\neq 0\ \forall\boldsymbol{\beta}\neq\boldsymbol{\beta}_0$. This means that $\boldsymbol{\beta}_0$ is the unique minimizer of $\mathrm{E}[(y_i-\mu(\boldsymbol{x_i'\beta}))^2]$
Next note that if $\mu(\cdot)$ is a linear function, $\boldsymbol{\beta}_0$ is the coefficient of the best linear predictor and has the usual closed form $\boldsymbol{\beta}_0=\mathrm{E}[\boldsymbol{x}_i\boldsymbol{x}_i']^{-1}\mathrm{E}[\boldsymbol{x}_iy_i]$

1.2 Q1 Part 2

In order to set this up as a Z estimator lets take a first order condition. This gives use the following condition.

$$\mathrm{E}[(\mu(\boldsymbol{x}_i'\boldsymbol{\beta}_0)-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}))\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta})\boldsymbol{x}_i]=0$$

Now take the sample analog and let $m(\boldsymbol{z}_i,\boldsymbol{\beta})=(y_i-\mu(\boldsymbol{x}_i'\boldsymbol{\beta}))\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta})\boldsymbol{x}_i$ where $\boldsymbol{z}_i=(y_i,\boldsymbol{x}_i')'$. We can write $\hat{\boldsymbol{\beta}}_n$ as the Z-estimator that solves:

$$0=\frac{1}{n}\sum_{i=1}^{n}m(\boldsymbol{z}_i,\hat{\boldsymbol{\beta}}_n)$$

Now assuming $\hat{\boldsymbol{\beta}}_n\to\boldsymbol{\beta}_0$ and regularity conditions we get the standard M estimation result.

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_n-\boldsymbol{\beta}_0)\to_d\mathcal{N}(0,\boldsymbol{H_0^{-1}\Sigma_0H_0^{-1}})$$

Were

$$\boldsymbol{H}_0=\mathrm{E}\left[\frac{\partial}{\partial\boldsymbol{\beta}}m(\boldsymbol{z}_i,\boldsymbol{\beta}_0)\right]=\mathrm{E}[\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}_0)^2\boldsymbol{x_ix_i'}]$$

and

$$\boldsymbol{\Sigma_0}=\mathrm{V}[m(\boldsymbol{z}_i,\boldsymbol{\beta}_0)]=\mathrm{E}[\sigma^2(\boldsymbol{x_i})\dot{\mu}(\boldsymbol{x_i'\beta_o})^2\boldsymbol{x_ix_i'}]$$

1.3 Q1 Part 3

$$\hat{\mathrm{V}}_n^{HC}=\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\boldsymbol{m}}\hat{\boldsymbol{m}}'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\boldsymbol{m}}\hat{\boldsymbol{m}}'\hat{e}_i^2\right)\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\boldsymbol{m}}\hat{\boldsymbol{m}}'\right)^{-1}$$

where $\hat{\boldsymbol{m}}=\boldsymbol{m_\beta(z_i,\hat{\beta})}$ and $hate=y_i-\boldsymbol{m(z_i,\hat{\beta})}$
The confidence interval is !!

1.4 Q1 Part 4

In this case we get $\boldsymbol{\Sigma_0} = \sigma^2 \boldsymbol{H}_0$ and the asymptotic variance reduces to

$$\boldsymbol{V}_0 = \sigma^2 \boldsymbol{H}_0^{-1} = \sigma^2 \mathrm{E}[\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}_0)^2 \boldsymbol{x_i}\boldsymbol{x_i'}]^{-1}$$

We can estimate variane using $\hat{\boldsymbol{V}} = \hat{\sigma}^2 \hat{\boldsymbol{H}}^{-1}$ where

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

and

$$\hat{\boldsymbol{H}} = \frac{a}{n}\sum_{i=1}^{n}\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}_0)^2 \boldsymbol{x_i}\boldsymbol{x_i'}$$

Which is consistent by the continuous mapping theorem. Now by the delta method and letting $g(\boldsymbol{\beta}) = \|\boldsymbol{\beta}\| = \sum_{k=1}^{d}\boldsymbol{\beta}_k^2$ we get

$$\sqrt{n}(g(\hat{\boldsymbol{\beta}_n}) - g(\boldsymbol{\beta}_0)) \rightarrow_d \mathcal{N}(0, \dot{g}(\boldsymbol{\beta}_0)\boldsymbol{V}_0\dot{g}(\boldsymbol{\beta}_0)')$$

where $\dot{g}(\boldsymbol{\beta}_0) = \frac{d}{d\boldsymbol{\beta}'}g(\boldsymbol{\beta}) = 2\boldsymbol{\beta'}$ Hence the confidence interval is given by

$$CI_{0.95} = \left[\|\hat{\boldsymbol{\beta}}_n\|^2 - 1.96\sqrt{\frac{4\hat{\boldsymbol{\beta}_n}'\hat{\boldsymbol{V}}\hat{\boldsymbol{\beta}_n}}{n}}, \|\hat{\boldsymbol{\beta}}_n\|^2 + 1.96\sqrt{\frac{4\hat{\boldsymbol{\beta}_n}'\hat{\boldsymbol{V}}\hat{\boldsymbol{\beta}_n}}{n}}\right]$$

1.5 Q1 Part 5

The conditional likelihod funciton is

$$f_{y|x}(y_i|\boldsymbol{x}_i) = \frac{1}{(2\pi)^{n/2}\sigma^2} exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}))^2\right)$$

with log likelihood

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

This gives us the following first order conditions

$$\frac{\partial}{\partial\boldsymbol{\beta}}\ell_n(\boldsymbol{\beta}, \sigma^2) = \frac{1}{\hat{\sigma}^2_{ML}}\sum_{i=1}^{n}(y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}_{ML}))\dot{\mu}(\boldsymbol{x}_i'\boldsymbol{\beta}_{ML})\boldsymbol{x}_i = 0$$

$$\frac{\partial}{\partial\sigma^2}\ell_n(\boldsymbol{\beta}, \sigma^2) = \frac{1}{2\hat{\sigma}^4_{ML}}\sum_{i=1}^{n}(y_i - \mu(\boldsymbol{x}_i'\boldsymbol{\beta}_{ML}))^2 - \frac{n}{2\hat{\sigma}^2_{ML}} = 0$$

These conditions are equivalent to those found above.