PhD Econometrics 1: Study Questions Class 6
Imperial College London

Hormoz Ramian

Question 1 Suppose that the sample is n, the log-likelihood function is denoted by $\ell(\boldsymbol{\beta})$ where $\boldsymbol{\beta} \in \boldsymbol{B}$ is a $k \times 1$ vector of parameters.

(1.1) State the maximum likelihood (ML) method's assumptions.

(1.2) Use mean value theorem to show the ML estimator is asymptotically normally distributed $\sqrt{n}(\widehat{\boldsymbol{\beta}}_{ML} - \boldsymbol{\beta}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathcal{I}^{-1})$.

(1.3) Now consider the normal linear model below:

$$\boldsymbol{y} \;=\; \boldsymbol{X\beta} + \boldsymbol{\epsilon}, \qquad \epsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{I}_n) \tag{1}$$

where n is the sample size and $n > k$ with k showing the column rank of nonrandom matrix $\boldsymbol{X}$, $\sigma^2 > 0$ and that $\boldsymbol{y}$ and $\boldsymbol{\epsilon}$ are $n \times 1$ vectors. Given a positive definite matrix $\boldsymbol{X}'\boldsymbol{X}$, Apply the general definitions of W, LM and LR to testing $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0$ against $H_1 : \boldsymbol{\beta} \neq \boldsymbol{\beta}_0$.

(1.4) Express LM and LR in terms of W, and then show that $W \geq LR \geq LM$.

(1.5) (Optional) Obtain asymptotic distributions of W and LR.

Question 2 Let $\boldsymbol{y}$ be a $n \times 1$ vector, and $\boldsymbol{X}$ and $\boldsymbol{Z}$ be a $n \times k$ and a $n \times q$ matrices, respectively. Define $\boldsymbol{u}(\boldsymbol{\theta}) := \boldsymbol{y} - \boldsymbol{X\theta}$ then the GMM estimator for this model $\widehat{\boldsymbol{\theta}}_n = \arg\min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} Q_n(\boldsymbol{\theta})$ where:

$$Q_n(\boldsymbol{\theta}) = \left\{ \frac{1}{n} \boldsymbol{u}(\boldsymbol{\theta})' \boldsymbol{Z} \right\} \boldsymbol{W}_n \left\{ \frac{1}{n} \boldsymbol{Z}' \boldsymbol{u}(\boldsymbol{\theta}) \right\} \tag{2}$$

is the GMM minimand.

(2.1) Derive the first-order condition for the minimization above and provide an expression for the estimator $\widehat{\boldsymbol{\theta}}_{GMM}$.

(2.2) Derive the asymptotic distribution of random variable $\widehat{\boldsymbol{\theta}}_{GMM}$.

(2.3) (Optional) Propose an approximate large sample $100(1 - \alpha)\%$ confidence interval for $\theta_{0,j}$, where j is the j^{th} element in vector $\boldsymbol{\theta}$.

Question 3 (optional) Consider the following factor model approach to represent a large set of yields $\boldsymbol{y}_t$ (with various maturities at each point in time e.g. 3-months to 10 years) as a function of a small set of unobserved factors. The Nelson and Siegel (1987) methodology expresses the yields with the following normal linear system[1]:

$$\boldsymbol{y}_t \;=\; \Lambda \boldsymbol{f}_t + \boldsymbol{\epsilon}_t \tag{3}$$

$$\boldsymbol{f}_t - \boldsymbol{\mu} \;=\; \boldsymbol{A}(\boldsymbol{f}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{\eta}_t \tag{4}$$

$$(\boldsymbol{\eta}_t, \boldsymbol{\epsilon}_t)' \;\sim\; \mathcal{N}\left[\mathbf{0}, \operatorname{diag}(\boldsymbol{Q}, \boldsymbol{H})\right] \tag{5}$$

where $\boldsymbol{f}_t' = \{\beta_{1,t}, \beta_{2,t}, \beta_{3,t}\} = \{L_t, S_t, C_t\}$, $\boldsymbol{\Lambda} = \left[1,\ (1 - e^{-\lambda\tau})/(\lambda\tau),\ [1 - e^{-\lambda\tau}/(\lambda\tau)] - e^{\lambda\tau}\right]$ is the factor loadings, λ is a known parameter and $\beta_{j,t}$ $\forall_j$ are factors or time-varying parameters particularly determining yield curve's level, slope and curvature, respectively, and that factors themselves are assumed to follow a first order de-meaned model shown in equation (4) with $\boldsymbol{\mu}' = (\mu_L, \mu_S, \mu_C)$.

[1] https://www.jstor.org/stable/2352957?socuuid=fa6258f4-1197-41cb-a737-70e8e297f4a3 and https://www.sciencedirect.com/science/article/pii/S030440760500014X

(3.1) The term $\boldsymbol{f}_t$ contains characteristics of the yield curve, given the data $\boldsymbol{y}_t = (y_{1,t}, \ldots, y_{6,t})'$ for the US treasuries[2] with 3, 6, 12, 24, 60, 120 months, which are unobservable to us. Estimate $\boldsymbol{f}_t$ using the maximum likelihood principle.

(3.2) Suppose that the characteristics of the yield curve remains unchanged within a fixed rolling window of an arbitrary size. Use the OLS method to estimate $\boldsymbol{f}_t$ within rolling windows as an approximation to the results obtain in the previous part and compare the estimated time-varying factors with results in the previous part.

[2] US Treasury Department database for yield curve rates:
https://www.treasury.gov/resource-center/data-chart-center/Pages/index.aspx