Escola de Economia de São Paulo - Fundação Getulio Vargas

Course: Econometria 2

Instructor: Vitor Possebom

Problem Set 5 - Total = 280 points

Question 1 (GMM with known gradient - 40 points)

We want to estimate the parameter λ of an Exponential distribution using the generalized method of moments. The probability density function of the Exponential distribution is given by

$$f(x) = \lambda_0 \cdot \exp(-\lambda_0 \cdot x).$$

In our GMM, we will use the first and second moments of the exponential distribution:

$$\mathbb{E}[g(\lambda_0, X)] = \mathbb{E}\begin{bmatrix} \frac{1}{\lambda_0} - X \\ \\ \frac{1}{\lambda_0^2} - \left(X - \frac{1}{\lambda_0}\right)^2 \end{bmatrix} = 0.$$

1. *Derive the gradient $G(\lambda) = \dfrac{\partial \mathbb{E}[g(\lambda, X)]}{\partial \lambda}$. (20 points)*
2. *Setting your seed to 220122, randomly draw 100,000 observations from an exponential distribution with $\lambda_0 = 5$. Using this vector as your data, estimate λ_0 using a GMM estimator where you provide the correctly specified gradient function.*[1] *Report your estimated coefficient and its estimated standard error. (20 points)*

Question 2 (Estimating an ARMA(1,2) using GMM - 80 points)

ARMA(p,q) models are usually estimated by Maximum Likelihood. However, estimating an ARMA(1,2) by GMM serves two purposes:

- *It illustrates how to write an IV model as a GMM model.*

[1] To numerically solve the optimization problem in this exercise, you will need to use Brent's method through the option `method`. To do so, you also need to specify the initial search interval through the options `lower` and `upper`.

- *It uses techniques that will be useful when we are analyzing dynamic panel data models.*

Let $\{Y_t\}$ *be a stationary stochastic process such that*

$$Y_t = \phi_1 \cdot Y_{t-1} + U_t \tag{1}$$

and

$$U_t = \epsilon_t + \theta_1 \cdot \epsilon_{t-1} + \theta_2 \cdot \epsilon_{t-2}, \tag{2}$$

where $\phi_1, \theta_1, \theta_2 > 0$ *and* $\epsilon_t \sim i.i.d.N(0,1)$.

1. *Show that you cannot estimate Equation 1 using OLS. (10 points)*
2. *Show that* Y_{t-2} *is not a valid instrument for* Y_{t-1} *in Equation 1. (10 points)*
3. *Show that* Y_{t-3} *is a valid instrument for* Y_{t-1} *in Equation 1. (20 points)*
4. *Setting your seed to 220122, randomly draw 100,000 observations from ARMA(1,2) process where* $\phi_1 = 0.2$, $\theta_1 = 0.1$, $\theta_2 = 0.1$. *Using this vector as your data, estimate* ϕ_1 *and the intercept using a GMM estimator where* Y_{t-3} *is used as an instrument for* Y_{t-1}. *Report the estimated coefficients (intercept and* ϕ_1*) and their estimated standard errors. (40 points)*

Question 3 (Testing the CAPM model - 160 points)

In Lecture 6B, we saw that the CAPM model implies a set of parametric moment conditions and restrictions on their parameters. In particular, we saw that the CAPM model implies that:

$$\mathbb{E}\left[\{(R_{i,t} - R_{f,t}) - \alpha_i - \beta_i \cdot (R_{m,t} - R_{f,t})\} \cdot (R_{m,t} - R_{f,t})\right] = 0, \tag{3}$$

where $\alpha_i = 0$ *for every stock* i *if the CAPM model is valid,* $R_{i,t}$ *is the return of stock* i *in day* t, $R_{f,t}$ *is the risk-free rate in day* t *and* $R_{m,t}$ *is the market portfolio's return in day* t.

Our goal is to check the validity of the CAPM model by estimating Equation (3) with Brazilian data[2] *and testing whether* $\alpha_i = 0$ *for every stock* i.

[2]The data for this exercise is messy on purpose. It is important to learn how to clean data in an appropriate way.

1. *Files* `BVSP.csv`, `BBDC3.csv`, `ABEV3.csv` *and* `ITUB3` *come from Yahoo Finance and contain information about the IBOVESPA index and the ordinary shares of Bradesco, AMBEV and Itaú. In those files, you can find the opening price, the closing price, the daily highest price, the daily lowest price, the adjusted closing price and the trading volume.*

   Create one data set with daily returns (based on closing prices) during 2021. (10 points)

2. *File* `selic.csv` *comes from the Brazilian Central Bank and contain information about the Selic Rate. In this file, you can find information about the annual Selic rate, the daily Selic factor, the daily trade volume, the number of daily operations, the daily mean, the daily median, the daily mode, the daily standard deviation and the daily kurtosis.*

   Compute the daily Selic rate and merge it with the dataset created in the previous item. You should keep only the data from 2021. (10 points)

3. *Create the variable* $(R_{m,t} - R_{f,t})$ *and the variables* $(R_{i,t} - R_{f,t})$ *for each share* i *in your dataset. (10 points)*

4. *Estimate the unrestricted CAPM model using GMM and the dataset built in the last item. Report all 6 coefficients and their standard errors. (120 points)*

5. *Test the validity of the CAPM model by testing the null hypothesis* $H_0 : \alpha_1 = 0, \alpha_2 = 0, \alpha_3 = 0$. *Report the p-value of this asymptotic test. (10 points)*