**Escola de Economia de São Paulo - Fundação Getulio Vargas**

**Course:** Econometria 2

**Instructor:** Vitor Possebom

**Problem Set 3 - Total = 430 points**

**Question 1 (General Time Series Regression - 90 points)**

*In this question, you will use the dataset* `data_brazil.csv`*. It contains four variables:*

- `date`*: Year (1901-2021)*
- `real_gdp_growth_pct`*: Annual Real GDP Growth measured in %.*
- `exchange_rate_real_dolar_annual_average`*: Annual Average Exchange Rate between Real and Dollar measured in* $^{R\$}/_{US\$}$*.*
- `ipc_fipe_pct`*: Annual Inflation measured by IPC-FIPE in %.*

*Your goal is to forecast GDP growth in 2020. To do so, we will use data from 1942 to 2019.*

1. *Model 1: Run an ADL(2,1) using GDP growth as your dependent variable and Exchange Rate as your predictor.*
   (a) *Report the estimated coefficient and their standard errors. (10 points)*
   (b) *Predict GDP growth in 2020 using model 1. (10 points)*
2. *Model 2: Run an ADL(2,2) using GDP growth as your dependent variable and inflation as your predictor.*
   (a) *Report the estimated coefficient and their standard errors. (10 points)*
   (b) *Predict GDP growth in 2020 using model 2. (10 points)*
3. *Model 3: Run an general time series regression model using GDP growth as your dependent variable and two lags of GDP growth, Exchange Rate and Inflation as your predictors.*

(a) *Report the estimated coefficient and their standard errors. (10 points)*

(b) *Predict GDP growth in 2020 using model 3. (10 points)*

4. *Model 4: Run an ARMA(2,0) using GDP growth as your dependent variable.*

   (a) *Report the estimated coefficient and their standard errors. (10 points)*

   (b) *Predict GDP growth in 2020 using model 4. (10 points)*

5. *Which model generate the prediction that is closest to the realized value? (10 points)*

**Question 2 (Expected Value of a $VAR(p)$ Process - 20 points points)**

*Assume that the vector process $\{Y_t\}$ is covariance-stationary and follows a $VAR(p)$ model. Find its mean $\mu =: \mathbb{E}[Y_t]$.*

**Question 3 (The $j-$th Autocovariance Matrix - 20 points)**

*Let $\{Y_t\}$ be a covariance-stationary $n$-dimensional vector process with mean $\mu$. Its $j$-th autocovariance is defined to be the following $(n \times n)$ matrix:*

$$\Gamma_j = \mathbb{E}\left[(Y_t - \mu)(Y_{t-j} - \mu)'\right].$$

*Prove that*

$$\Gamma_j' = \Gamma_{-j}.$$

**Question 4 (Vector $MA(q)$ Process - 40 points)**

*A vector moving average process of order $q$ ($VAR(p)$) takes the form*

$$Y_t = \mu + \epsilon_t + \Theta_1 \epsilon_{t-1} + \Theta_2 \epsilon_{t-2} + \ldots + \Theta_q \epsilon_{t-q},$$

*where $\{\epsilon_t\}$ is a vector white noise process and $\Theta_j$ denotes a $(n \times n)$ matrix of $MA$ coefficients for $j \in \{1, 2, \ldots, q\}$.*

*Prove that any vector $MA(q)$ process is covariance-stationary.*

**Question 5 (A Type of LLN - 40 points)**

*Prove Proposition 1 in the Lecture Notes. (I wrote this proposition below for your convenience.)*

**Proposition 1** *Let $\{Y_t\}$ be a covariance-stationary process with moments given by*

$$\mathbb{E}[Y_t] = \mu,$$
$$\mathbb{E}\left[(Y_t - \mu)(Y_{t-j} - \mu)'\right] = \Gamma_j$$

*and with absolutely summable autocovariances (i.e., $\left(\sum_{v=\infty}^{+\infty} \Gamma_v\right) \in \mathbb{R}$). Assume we have a sample of size $T$ drawn from $\{Y_t\}$. Then, the sample mean $\bar{Y}_T := \frac{\sum_{t=1}^{T} Y_t}{T}$ satisfies*

1. $\bar{Y}_T \xrightarrow{p} \mu$
2. $\lim_{T\to+\infty} \left\{T \cdot \mathbb{E}\left[\left(\bar{Y}_T - \mu\right)\left(\bar{Y}_T - \mu\right)'\right]\right\} = \sum_{v=\infty}^{+\infty} \Gamma_v$.

**Question 6 ($VAR(p)$ Model - 220 points)**

*In this question, you will, once more, use the dataset `data_brazil.csv`. It contains four variables:*

- `date`: *Year (1901-2021)*
- `real_gdp_growth_pct`: *Annual Real GDP Growth measured in %.*
- `exchange_rate_real_dolar_annual_average`: *Annual Average Exchange Rate between Real and Dollar measured in R\$/US\$.*
- `ipc_fipe_pct`: *Annual Inflation measured by IPC-FIPE in %.*

*Your goal is to forecast GDP growth in 2020. To do so, we will use data from 1942 to 2019.*

1. *First, we will use a $VAR(p)$ to forecast GDP growth in 2020.*
   (a) *Model A: Set $p = 1$ and estimate a reduced-form $VAR(1)$ model.*
      i. *Report the estimated coefficient and their standard errors. (10 points)*

ii. *Predict GDP growth in 2020 using model A. (10 points)*

(b) *Model B: Set* $p = 2$ *and estimate a reduced-form* $VAR(2)$ *model.*

i. *Report the estimated coefficient and their standard errors. (10 points)*

ii. *Predict GDP growth in 2020 using model B. (10 points)*

(c) *Model C: Set* $p = 3$ *and estimate a reduced-form* $VAR(3)$ *model.*

i. *Report the estimated coefficient and their standard errors. (10 points)*

ii. *Predict GDP growth in 2020 using model C. (10 points)*

(d) *Which model generate the prediction that is closest to the realized value? (10 points)*

2. *Now, we will focus on a* $VAR(2)$ *model and focus on structural IRFs.*

(a) *Choose the order of your variables and justify your exclusion restrictions. (40 points)*

(b) *Estimate and plot all nine structural impulse response functions and their 90%-confidence intervals based on 1,000 bootstrap repetitions. (90 points)*

(c) *Do you believe your results are credible? Justify your answer. (20 points)*