For this assignment, you may work in groups of two or three students. Part 1 is theoretical and part 2 is an application of your results from part 1. For each sub-question, you can earn three points. It is crucial to motivate your answers: No points will be awarded for answers without motivation. Furthermore, reports that are not typeset as a single PDF file in LaTeX will not be considered. Please also make sure to submit your R (and/or Python) codes for question 2, separately, together with a CSV file of the data. For question 2, your code should reproduce all reported results without returning errors. Results that are not returned by your code cannot be graded.

The deadline for this assignment is Monday, March 11, 15:00.

Part 1. For the next couple of subquestions we assume a m-dimensional VAR(p) process for $\mathbf{y}_t = \begin{pmatrix} y_{1t} & y_{2t} & \cdots y_{mt} \end{pmatrix}'$, generated by

$$\mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + ... + \boldsymbol{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t,$$

conditional on the assumed to be available presample values $\mathbf{y}_{-p+1}, \ldots, \mathbf{y}_0$. Here $\boldsymbol{\alpha}$ is the $(m \times 1)$ vector of intercepts, $\boldsymbol{\Phi}_i$ are $(m \times m)$ coefficient matrices and $\boldsymbol{\varepsilon}_t$ is the $(m \times 1)$ vector of disturbance terms, where we assume that $\boldsymbol{\varepsilon}_t \sim$ i.i.d. $\mathcal{N}(0, \boldsymbol{\Omega}_\varepsilon)$. assume that

You will now deduce the Least Squares estimator for the coefficients in a couple of steps. We will begin by defining some convenient notation:

$\mathbf{Y} = \begin{pmatrix} \mathbf{y}_1 & \cdots & \mathbf{y}_n \end{pmatrix}$, dimensions $(m \times n)$
$\mathbf{B} = \begin{pmatrix} \boldsymbol{\alpha} & \boldsymbol{\Phi}_1 & \cdots & \boldsymbol{\Phi}_p \end{pmatrix}$, dimensions $(m \times (mp+1))$

$\mathbf{z}_t = \begin{pmatrix} 1 & \mathbf{y}_t' & \ldots & \mathbf{y}_{t-p+1}' \end{pmatrix}'$, dimensions $((mp+1) \times 1)$

$\mathbf{Z} = \begin{pmatrix} \mathbf{z}_0 & \mathbf{z}_1 & \cdots & \mathbf{z}_{n-1} \end{pmatrix}$, dimensions $((mp+1) \times n)$
$\mathbf{E} = \begin{pmatrix} \boldsymbol{\varepsilon}_1 & \cdots & \boldsymbol{\varepsilon}_n \end{pmatrix}$, dimensions $(m \times n)$
$\mathbf{y} = \mathrm{vec}(\mathbf{Y})$, dimensions $(mn \times 1)$
$\boldsymbol{\beta} = \mathrm{vec}(\mathbf{B})$, dimensions $((m^2p + m) \times 1)$
$\boldsymbol{\varepsilon} = \mathrm{vec}(\mathbf{E})$, dimension $(mn \times 1)$

Using this notation, the full VAR(p) model for $t = 1, ..., n$ can be written as $\mathbf{Y} = \mathbf{BZ} + \mathbf{E}$.

(a) Use the information above and the properties $\mathrm{vec}(\mathbf{A} + \mathbf{C}) = \mathrm{vec}(\mathbf{A}) + \mathrm{vec}(\mathbf{C})$ and

$\text{vec}(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\text{vec}(\mathbf{B})$, to show that

$$\mathbf{y} = (\mathbf{Z}' \otimes \mathbf{I}_m)\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Note that the covariance matrix of $\boldsymbol{\varepsilon}$ is given by $\boldsymbol{\Omega} = \mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon$. This implies that with GLS estimation, we should minimize

$$S(\boldsymbol{\beta}) = \boldsymbol{\varepsilon}'(\mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon)^{-1}\boldsymbol{\varepsilon}$$

(b) Show that this corresponds to minimizing

$$S(\boldsymbol{\beta}) = \mathbf{y}'(\mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon^{-1})\mathbf{y} + \boldsymbol{\beta}'(\mathbf{ZZ}' \otimes \boldsymbol{\Omega}_\varepsilon^{-1})\boldsymbol{\beta} - 2\boldsymbol{\beta}'(\mathbf{Z} \otimes \boldsymbol{\Omega}_\varepsilon^{-1})\mathbf{y}$$

To reach this result you should use the result from (a) and the properties $(\mathbf{A} \otimes \mathbf{C})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{C}^{-1}$, $(\mathbf{A} \otimes \mathbf{C})' = \mathbf{A}' \otimes \mathbf{C}'$ and $(\mathbf{A} \otimes \mathbf{C})(\mathbf{F} \otimes \mathbf{G}) = (\mathbf{AF}) \otimes (\mathbf{CG})$.

(c) Now solve the first order conditions w.r.t $\boldsymbol{\beta}$ to show that the Least Squares estimator is given by

$$\hat{\boldsymbol{\beta}} = ((\mathbf{ZZ}')^{-1}\mathbf{Z} \otimes \mathbf{I}_m)\mathbf{y}$$

and check that the Hessian matrix of $S(\boldsymbol{\beta})$ is positive definite, which indeed implies minimization. All properties needed to derive this result, have already been mentioned.

Interestingly, the GLS estimator $\hat{\boldsymbol{\beta}}$ is equal to the (multivariate) OLS estimator, obtained from minimising $\bar{S}(\boldsymbol{\beta}) = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}$. This somewhat surprising equivalence result follows from the fact that we apply LS estimation to a multiple equation model where the regressors are the same in all equations.

(d) Follow a series of three steps to derive the OLS estimator for $\mathbf{B}$:

(i) Show that we can write $\hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + ((\mathbf{ZZ}')^{-1}\mathbf{Z} \otimes \mathbf{I}_m)\boldsymbol{\varepsilon}$

(ii) Show that $\text{vec}(\hat{\mathbf{B}}) = \text{vec}(\mathbf{YZ}'(\mathbf{ZZ}')^{-1})$

(iii) Show that $\hat{\mathbf{B}} = \mathbf{B} + \mathbf{EZ}'(\mathbf{ZZ}')^{-1}$

(e) Assuming that $\boldsymbol{\Omega}_\varepsilon = \mathbb{E}[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t']$ and using that $\sum_{t=1}^n \hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t' = \hat{\mathbf{E}}\hat{\mathbf{E}}'$, derive that a reasonable estimator for $\boldsymbol{\Omega}_\varepsilon$ is given by

$$\tilde{\boldsymbol{\Omega}}_\varepsilon = \frac{1}{n}\mathbf{Y}(\mathbf{I}_n - \mathbf{Z}'(\mathbf{ZZ}')^{-1}\mathbf{Z})\mathbf{Y}'.$$

Note: This estimator is often corrected for the degrees of freedom, leading to the final estimator $\hat{\boldsymbol{\Omega}}_\varepsilon = \frac{n}{n-mp-1}\tilde{\boldsymbol{\Omega}}_\varepsilon$.

(f) Derive the log-likelihood of such a VAR(p) process, using the normality assumption made at the start of the exercise. Explain why GLS and ML estimators of the coefficients coincide.

Part 2. In each of the sub-questions below, make sure to program the required log-likelihoods yourself. Besides base R/Python, your code from previous assignments, packages for plotting and optimization, you are not allowed to use any other packages for your final results. Here, 'base Python' refers to any basic operations captured by Numpy and other packages providing operations that translate to R.

Download daily data from the website of Yahoo Finance for the 13-week Treasury Bill (ticker: ˆ IRX) and for the 5-year Treasury Yield (ticker: ˆ FVX) from January 2, 2002, up to and including December 29, 2023. Call these time series x_t and y_t respectively, from now on. Save as a CSV file and enable your script to read the data. Use the adjusted closing prices as our time series. Note that the data may contain days without a value (non-trading days). Make sure you remove these empty entries and the two series have the same length.

(a) Estimate the parameters for the two-dimensional VAR(1) model:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

Which is in matrix form:

$$\mathbf{Y}_t = \boldsymbol{\alpha} + \boldsymbol{\Phi}\mathbf{Y}_{t-1} + \boldsymbol{\varepsilon}_t$$

Here we assume that $\boldsymbol{\varepsilon}_t \sim \text{i.i.d.}\mathcal{N}(0, \boldsymbol{\Omega})$, with $\boldsymbol{\Omega} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$.

(b) With the resulting coefficients, what can you say about Granger causality in this process? Plot the impulse response functions (4 plots in total) and give an interpretation.

(c) Now, estimate the covariance matrix $\boldsymbol{\Omega}$ for the VAR(1) model and calculate the log-likelihood of the estimated model.

(d) Also perform the estimations for VAR models with two and three lags, VAR(2) and VAR(3), and form a table with the AIC, the log-likelihood and the total number of parameters for the VAR(1), VAR(2) and VAR(3) models. Choose the optimal model based on the results.

(e) Write down the implied ARMA processes for x_t and y_t, for the model you have picked in (d).

(f) Perform the Johansen's trace test for cointegration on the optimal estimated model and show the results. Draw your conclusion on the cointegration relation between x_t and y_t in this model.

