# Assignment #6 (mandatory)

feedback session for this assignment: 01-02-2021

**General remarks: This assignment deals with first steps in multivariate time series analysis on real economic data. It starts with the implementation of an ADF-test to assess the properties of the series at hand and then moves on to estimating the parameters of a reduced form VAR and trying to link composite and idiosyncratic innovations. Please note the following:**

- **At the time of uploading the assignment, you are already able to deal with tasks 1 and 2 of the assignment – the next two lectures will give you more theoretical background regarding task 3 (to account for that, our feedback session takes place on Feb 1)**
- **You will notice that the VAR-part of the assignment contains some tasks that "only" ask you to specifically write down some matrices in your pdf (and do not require any coding at all). This is not a mistake but on purpose: In VAR analysis, it is very important that you are at all times aware about the ordering of your variables and the dimensions of your matrices.**

## 1. Applying the Augmented Dickey-Fuller test to economic data

1. The SWISS_1976_2014.xlsx file contains Swiss seasonally-adjusted macroeconomic data: quarterly interest rate in percent (first column), quarterly Consumer Price Index (second column), quarterly Gross Domestic Product in Mio CHF (third column) and quarterly Money Stock M1 in Mio CHF (fourth column). The first observation is from the first quarter of 1976 and the last observation is from the last quarter of 2014; hence $T = 156$. Load the data into your program.
   **Hint:** Use the MATLAB function `xlsread` to load an Excel file into MATLAB.
2. Plot the four time series and make your graphs look professional. For each series, suggest an appropriate case (1, 2 or 4) to test for a unit root.
3. Use an Augmented Dickey-Fuller test to test for the presence of unit roots. For each of the series estimate the model:

$$y_t = \alpha + \delta t + \rho y_{t-1} + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + u_t, \tag{1}$$

   where $\Delta y_t = y_t - y_{t-1}$. The parameters $\alpha$ and $\delta$ are optional and only included in equation (1) if you believe that the respective series is best described using a constant or a time trend. Write a function whose input argument is the series for which you want to perform the analysis. You may want to use a second (auxiliary) input variable to determine whether or not a constant and/or trend are included in the model. Estimate the parameters by OLS via $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. Make sure that you set up your $\mathbf{X}$ matrix appropriately (use an if statement and the auxiliary input variable to include a constant and/or trend if you wish to do so). Output argument of your function should be a $(1 \times 3)$ vector consisting of $\hat{\rho}$, $s.e.(\hat{\rho})$, and $\frac{(\hat{\rho}-1)}{s.e.(\hat{\rho})}$.
   **Hint 1:** The MATLAB function `lagmatrix` (see MATLAB documentation) lags a series $n$ times. Be aware of the fact that we do not have a $y_0$ value - so you actually lose observations when taking lags.
   **Hint 2:** $\widehat{\mathrm{Var}}(\mathbf{b})$=$s^2(\mathbf{X}'\mathbf{X})^{-1}$, where $s^2 = \frac{1}{T-m}\sum(y_t - \hat{\alpha} - \hat{\delta}t - \hat{\rho}y_{t-1} - \hat{\zeta}_1\Delta y_{t-1} - \hat{\zeta}_2\Delta y_{t-2})^2$ with $m$ denoting the number of estimated parameters ($m = 5$) and $T$ equaling the number of rows in the $\mathbf{X}$ matrix (you lose some observations due to taking lags).

4. Extend your function from task 1.3 to determine the optimal number of lagged differences in
$$y_t = \alpha + \delta t + \rho y_{t-1} + \zeta_1 \Delta y_{t-1} + \ldots + \zeta_k \Delta y_{t-k} + u_t, \tag{2}$$
for each of the series using the (residual-based) Bayes-Schwarz information criterion (SBC) and the (residual-based) Akaike information criterion (AIC). In addition to the input arguments mentioned in task 1.3, a further argument is now the number of lagged differences that should be included in your model. Use a loop statement and the `lagmatrix` to construct your **X** matrix. Estimate the model parameters and compute $s.e.(\hat{\rho})$ and $\frac{(\hat{\rho}-1)}{s.e.(\hat{\rho})}$ as above. Additionally, compute the SBC and the AIC. Output argument of your procedure should be a $(1 \times 6)$ vector consisting of the number of lagged differences currently included in your model, $\hat{\rho}$, $s.e.(\hat{\rho})$, $\frac{(\hat{\rho}-1)}{s.e.(\hat{\rho})}$, the SBC and the AIC values.
**Hint:** $SBC = \ln(s^2) + \frac{m}{T}\ln(T)$, and $AIC = \ln(s^2) + \frac{2m}{T}$, and $s^2 = \frac{1}{T-m}\sum(y_t - \hat{\alpha} - \hat{\delta}t - \hat{\rho}y_{t-1} - \hat{\zeta}_1\Delta y_{t-1} - \ldots - \hat{\zeta}_k\Delta y_{t-k})^2$, where $m = k+3$.
5. For each of the series, use a loop statement to call your function from task 1.4 for the number of included lagged differences ranging from 0 ($y_t = \alpha + \delta t + \rho y_{t-1} + u_t$) to 16 ($y_t = \alpha + \delta t + \rho y_{t-1} + \zeta_1 \Delta y_{t-1} + \ldots + \zeta_{16} \Delta y_{t-16} + u_t$). Store your results in a matrix and print it to the command window.
6. Use your results from task 1.5 and argue for each of the series how many lagged differences you would include in your model based on the SBC and the AIC, respectively. Explain briefly why the two criteria may not always prefer the same number of lags. For task 1.7, only consider the SBC-preferred model specification.
7. Decide – based on the value of the Dickey-Fuller $t$-statistic – whether you can reject or not reject the null hypothesis of a unit root at the 5% significance level. Obtain the respective critical value from the table in Hamilton (p. 762) and use the entries for infinitely long series, as it is usually done in applied work.
**Hint:** If you could not solve tasks 1.4–1.6, use your results from task 1.1.
8. Transform the series by computing (log) differences:
For the interest rate series, use $\tilde{y}_t = \Delta y_t = y_t - y_{t-1}$; for all other series compute $\tilde{y}_t = \Delta \ln(y_t) = \ln(y_t) - \ln(y_{t-1})$.
9. Plot the series $\tilde{y}_t$ and describe the plot. For each series, suggest an appropriate case (1, 2 or 4) to test for a unit root.
10. Repeat tasks 1.5–1.7 for $\tilde{y}_t$ (for all four time series) and discuss your results.
11. Compute and interpret the 95% confidence bounds of $\rho$ using the transformed GDP and M1 series and the model specification preferred by the SBC.
**Hint:** Use the MATLAB function `norminv` to obtain quantiles of a normal distribution (see MATLAB documentation).

**2. Estimation of a structural VAR**:

From now on, we are going to use the following notation: $\tilde{m}$ refers to tranformed money stock series that you created in task 1.8, $\tilde{r}$ refers to the transformed interest rate series, $\tilde{p}$ refers to the transformed CPI series, and $\tilde{g}$ refers to the transformed GDP series.

1. We want to model the effects of Swiss monetary policy and consider the following structural VAR in its primitive form:

$$\begin{aligned}
\tilde{m}_t + b_{12}^{(0)}\tilde{r}_t + b_{13}^{(0)}\tilde{p}_t + b_{14}^{(0)}\tilde{g}_t &= k_1 + b_{11}^{(1)}\tilde{m}_{t-1} + b_{12}^{(1)}\tilde{r}_{t-1} + b_{13}^{(1)}\tilde{p}_{t-1} + b_{14}^{(1)}\tilde{g}_{t-1} + u_{1t} \\
b_{21}^{(0)}\tilde{m}_t + \tilde{r}_t + b_{23}^{(0)}\tilde{p}_t + b_{24}^{(0)}\tilde{g}_t &= k_2 + b_{21}^{(1)}\tilde{m}_{t-1} + b_{22}^{(1)}\tilde{r}_{t-1} + b_{23}^{(1)}\tilde{p}_{t-1} + b_{24}^{(1)}\tilde{g}_{t-1} + u_{2t} \\
b_{31}^{(0)}\tilde{m}_t + b_{32}^{(0)}\tilde{r}_t + \tilde{p}_t + b_{34}^{(0)}\tilde{g}_t &= k_3 + b_{31}^{(1)}\tilde{m}_{t-1} + b_{32}^{(1)}\tilde{r}_{t-1} + b_{33}^{(1)}\tilde{p}_{t-1} + b_{34}^{(1)}\tilde{g}_{t-1} + u_{3t} \\
b_{41}^{(0)}\tilde{m}_t + b_{42}^{(0)}\tilde{r}_t + b_{43}^{(0)}\tilde{p}_t + \tilde{g}_t &= k_4 + b_{41}^{(1)}\tilde{m}_{t-1} + b_{42}^{(1)}\tilde{r}_{t-1} + b_{43}^{(1)}\tilde{p}_{t-1} + b_{44}^{(1)}\tilde{g}_{t-1} + u_{4t},
\end{aligned} \tag{3}$$

   which is in short

$$\mathbf{B}_0\mathbf{y}_t = \mathbf{k} + \mathbf{B}_1\mathbf{y}_{t-1} + \mathbf{u}_t, \tag{4}$$

   where

$$\mathbb{E}_t[\mathbf{u}_t\mathbf{u}_\tau'] = \begin{cases} \mathbf{D} & \text{for} \quad t = \tau \\ \mathbf{0} & \text{otherwise} \end{cases} \tag{5}$$

   In your pdf, write down the elements in Equation (4) ($\mathbf{B}_0$, $\mathbf{y}_t$, $\mathbf{k}$, $\mathbf{B}_1$, $\mathbf{y}_{t-1}$ and $\mathbf{u}_t$) explicitly.

2. The estimation of the VAR is performed in its standard form, which is

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{B}_0^{-1}\mathbf{k} + \mathbf{B}_0^{-1}\mathbf{B}_1\mathbf{y}_{t-1} + \mathbf{B}_0^{-1}\mathbf{u}_t \\
\mathbf{y}_t &= \mathbf{c} + \mathbf{\Phi}_1\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t.
\end{aligned} \tag{6}$$

   It can be written extensively as

$$\begin{aligned}
\tilde{m}_t &= c_1 + \phi_{11}\tilde{m}_{t-1} + \phi_{12}\tilde{r}_{t-1} + \phi_{13}\tilde{p}_{t-1} + \phi_{14}\tilde{g}_{t-1} + \varepsilon_{1t} \\
\tilde{r}_t &= c_2 + \phi_{21}\tilde{m}_{t-1} + \phi_{22}\tilde{r}_{t-1} + \phi_{23}\tilde{p}_{t-1} + \phi_{24}\tilde{g}_{t-1} + \varepsilon_{2t} \\
\tilde{p}_t &= c_3 + \phi_{31}\tilde{m}_{t-1} + \phi_{32}\tilde{r}_{t-1} + \phi_{33}\tilde{p}_{t-1} + \phi_{34}\tilde{g}_{t-1} + \varepsilon_{3t} \\
\tilde{g}_t &= c_4 + \phi_{41}\tilde{m}_{t-1} + \phi_{42}\tilde{r}_{t-1} + \phi_{43}\tilde{p}_{t-1} + \phi_{44}\tilde{g}_{t-1} + \varepsilon_{4t}.
\end{aligned} \tag{7}$$

   Why do we consider the standard form of the VAR (instead of the structural form) for estimation purposes?

3. Explain why the composite innovations of this VAR in standard form are contemporaneously correlated by construction.

4. Include the VAR estimation toolbox into your working directory and study the code of the function file `PHI.m`.
   **Hint:** The toolbox can be dowloaded from Ilias. It is a folder called `VAR2_toolbox` which contains five MATLAB function files. The toolbox can be included into your program by writing

```
addpath('VAR2_toolbox');
```

   somewhere at the beginning of your program.

5. Estimate the VAR from Equation (7) using the function file `PHI.m` given in the toolbox:

```
[constant,Phi_1,Phi_2,Omega]=PHI(x_t,p).
```

Read the documentation of the function code carefully. `x_t` refers to the data set and `p` denotes the number of lags. The ordering of the macroeconomic variables should be as implied by Equation (3). Hence, you have to re-order the columns of the data matrix that you loaded into the program. The function allows for only 1 or 2 lags. Make sure to understand what the function returns to you.

6. Using the results of your estimation, write $\hat{\mathbf{c}}$ and $\hat{\mathbf{\Phi}}_1$ explicitly in your command window and in your pdf.

**3. Obtain idiosyncratic shocks ($\mathbf{u}_t$) from composite shocks ($\varepsilon_t$):**

1. To obtain the idiosyncratic shocks from the composite shocks, we need the matrix $\mathbf{B}_0$, because:

$$\boldsymbol{\varepsilon}_t = \mathbf{B}_0^{-1}\mathbf{u}_t. \tag{8}$$

We can link the variance-covariance matrix of the composite shocks $\boldsymbol{\varepsilon}_t$ to the variance-covariance matrix of the idiosyncratic shocks $\mathbf{u}_t$ (labeled $\mathbf{D}$ in Equation (9)) by:

$$\begin{aligned} \mathbb{E}[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'] &= \mathbf{\Omega} \\ &= \mathbf{B}_0^{-1}\mathbb{E}[\mathbf{u}_u\mathbf{u}_t'](\mathbf{B}_0^{-1})' \\ &= \mathbf{B}_0^{-1}\mathbf{D}(\mathbf{B}_0^{-1})'. \end{aligned} \tag{9}$$

Print $\hat{\mathbf{\Omega}}$ to the command window.

2. When trying to obtain $\hat{\mathbf{B}}_0$ and $\hat{\mathbf{D}}$ from $\hat{\mathbf{\Omega}}$, you are facing an identification problem. How many unknown parameters are there and how many equations do you have to help you solve the system? How does the Cholesky decomposition help to solve this problem?

3. $\hat{\mathbf{\Omega}}$ is a real symmetric positive definite matrix and we know that for such a matrix there exist matrices $\mathbf{P}$, $\mathbf{A}$ and $\mathbf{C}$ such that

$$\hat{\mathbf{\Omega}} = \mathbf{P}\mathbf{P}' = \mathbf{A}\mathbf{C}\mathbf{C}'\mathbf{A}', \tag{10}$$

where $\mathbf{P}$ and $\mathbf{A}$ are lower triangular matrices and $\mathbf{C}$ is a diagonal matrix with positive elements. We use this mathematical property for the identification of $\hat{\mathbf{B}}_0^{-1}$ and $\hat{\mathbf{D}}$, link it to Equation (9), and assume that

$$\hat{\mathbf{B}}_0^{-1} = \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ a_{21} & 1 & 0 & \dots & 0 \\ a_{31} & a_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & 1 \end{bmatrix} \quad \text{and} \quad \hat{\mathbf{D}} = \mathbf{C}\mathbf{C}' = \begin{bmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ 0 & 0 & d_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & d_n \end{bmatrix},$$

$$\text{where} \quad \mathbf{C} = \begin{bmatrix} \sqrt{d_1} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{d_2} & 0 & \dots & 0 \\ 0 & 0 & \sqrt{d_3} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sqrt{d_n} \end{bmatrix} = \begin{bmatrix} p_{11} & 0 & 0 & \dots & 0 \\ 0 & p_{22} & 0 & \dots & 0 \\ 0 & 0 & p_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & p_{nn} \end{bmatrix} \tag{11}$$

Compute the lower triangular matrix $\mathbf{P}$ from the estimated variance-covariance matrix of the composite innovations $\hat{\mathbf{\Omega}}$ using a Cholesky decomposition and print the matrix to the

output window and in your pdf. Note that the ordering of the variables is important.
**Hint:** Search the MATLAB documentation for the function `chol`.

4. Obtain the matrices $\mathbf{A}$, $\hat{\mathbf{B}}_0$, $\mathbf{C}$, and $\hat{\mathbf{D}}$ using $\hat{\boldsymbol{\Omega}}$ and $\mathbf{P}$. Print all the matrices to the command window and in your pdf.
**Hint:** First compute $\mathbf{C}$ using the link to the elements of the $\mathbf{P}$ matrix that is provided in Equation (11). Then compute $\mathbf{A}$ from the relationship of $\mathbf{A}$ (unknown), $\mathbf{P}$ (known) and $\mathbf{C}$ (known) that is implied in Equation (10). Finally, obtain $\hat{\mathbf{B}}_0^{-1}$ and $\hat{\mathbf{D}}$ from $\mathbf{A}$ and $\mathbf{C}$.

---

This assignment can be handed in for grading until 11:55 pm on Sunday, February 7th, 2021. Upload your working MATLAB script file and all your associated function files as well as your report (pdf format) to ILIAS.