Practical Homework 1

Gorlevich Daniil, Andreicheva Polina

December 2021

Contents

1 Data description

The VAR model was made with Blanchard-Quah (1989) results in mind. We decided to take quarterly US data on real GNP and unemployment rate for period from 1983 to 2021. The data was taken from the site of Federal Reserve Bank of St. Louis [1][2]. GNP was calculated in chained 2012 dollars. To follow the original paper we decided to use the natural logarithm of GNP.

2 Univariate Model

2.1 Preprocessing

Figure 1 depicts an unprocessed log GNP data.

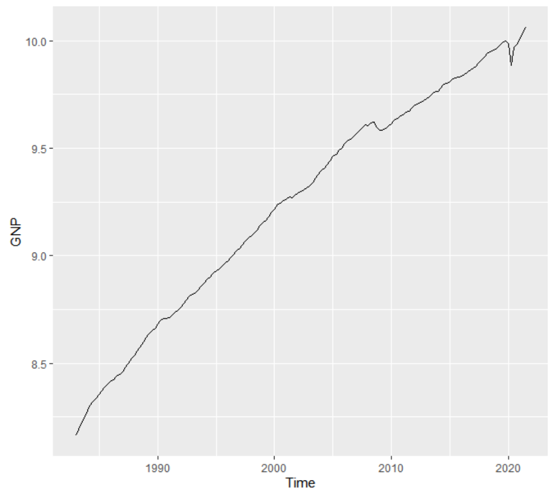

Figure 1: ln(GNP)

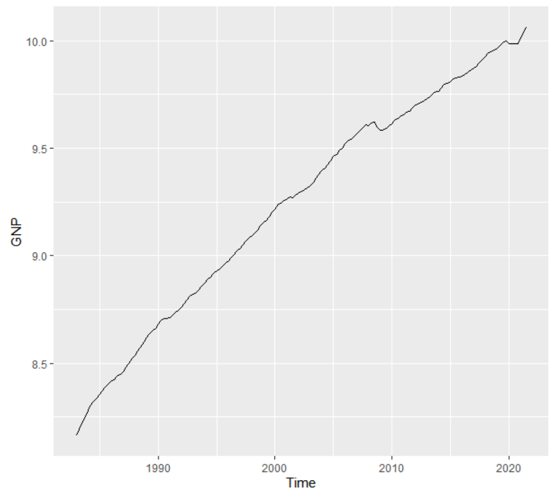

Figure 2: ln(GNP) without outliers

Firstly, the data had outliers - 2008 and coronavirus crises stood out from the data. We used *forecast::tsclean()* function to remove outliers. The procedure is to fit a smoother model on seasonality and then marking the normal data in range $(25\text{th quantile} - 3 \times IQR, 75\text{th quantile} + 3 \times IQR)$. Outliers then are substituted with linear interpolation. Figure 2 shows data after *forecast::tsclean()* procedure.

2.2 Models

After outliers were removed from the data, we decided to use automated procedure to build ARIMA model. The best model was ARIMA(1, 2, 3)(1, 0, 2)[4].
To verify the quality of the fit, we checked the residuals to verify that they are uncorrelated. Figure 3 shows residual diagnostics
We see that ACF does not have a significant lags, except one. The histogram of the residuals is more concentrated around the mathematical expectation than the normal distribution plot.
Formally testing the hypotheses on the absence of correlation between residuals and normality of distribution is more precise than the graphic analysis. Firstly, we tested the Box-Ljung test.
The null hypothesis of the test is "independence in a given series". We receive P-value of 0.3624. The null hypothesis cannot be rejected, which means that there is no residual correlations in the data.
The results of the Portmanteau test indicate that residuals look like a white noise process, which means that the automated fit is good.

2.3 Forecasting

After making sure that the data is fitted, we can forecast the future values of the GNP (Figure 4).

[1] GNP: https://fred.stlouisfed.org/series/GNPC96
[2] Unemployment: https://fred.stlouisfed.org/series/LRUN64TTUSQ156S

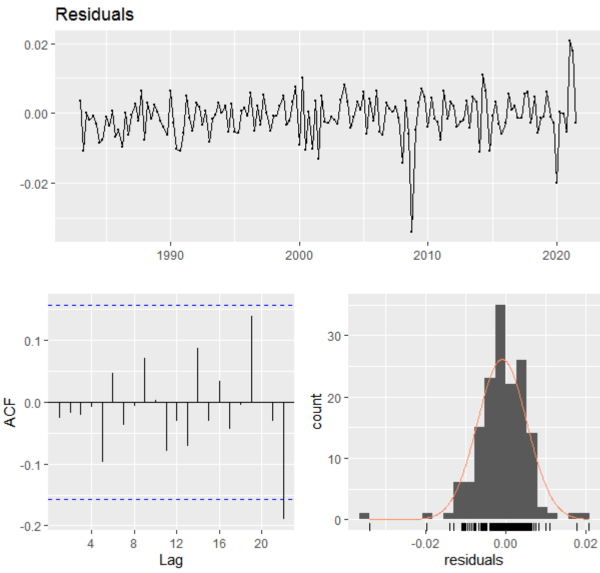

Figure 3: Residual diagnostics

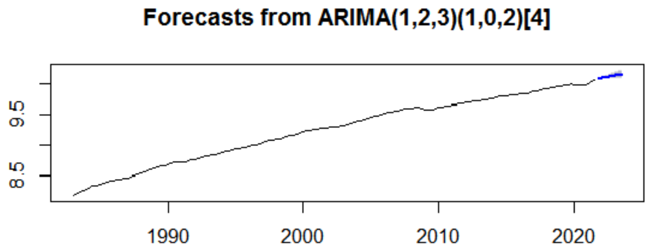

Figure 4: ln(GNP) forecast

2.4 Quality

On horizon of 8 quarters RMSE of ARIMA was 0.006245809, against RMSE of 0.004097395 of Random Walk Forecast and 0.004097287 of Random Walk with Drift forecast. Cross-validation RMSE for ARIMA was 0.003510494, which means that the model is better than the benchmarks.

3 Multivariate Model

3.1 Preprocessing

VARs have a more strict requirements for stationarity than the automatic ARIMA, so we need additional preprocessing in this part.
Firstly, Figure 5 shows unprocessed data. We can see that the unemployment possibly follows a polynomial trend if any or a seasonal component that might reflect business cycles.
We use the same procedure as in 2.1 to remove the outliers.
Both ADF and KPSS tests show that GNP is non-stationary. Unemployment is shown to be stationary. However, unemployment seems to have a seasonal component. To remove trend and seasonality we use *forecast::decompose* for both variables. Random part of the variables is shown in Figure 7
After the decomposition both ADF and KPSS tests show stationarity for both variables.

3.2 VAR

We select VAR order via automatic procedure. The selection criteria suggest the following ordering:

AIC(n)	HQ(n)	SC(n)	FPE(n)
5	2	1	5

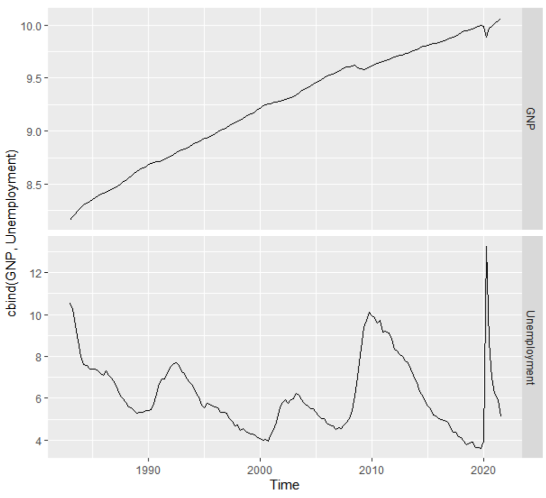

Figure 5: ln(GNP) and Unemployment

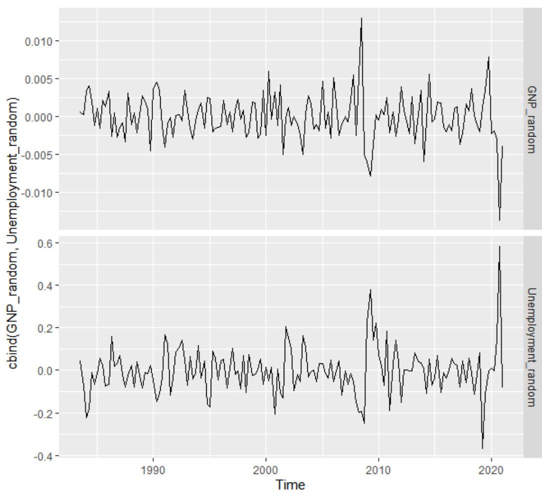

Figure 6: ln(GNP) and Unemployment, detrended

After automatic procedure we select lags based on Portmanteau Asymptotic Test and Breusch-Godfrey test. Both of them suggest using VAR(5)

Order	PT	BG
1	0.02719	0.0004783
2	0.03933	0.004089
3	0.01156	0.02678
4	0.02442	0.04888
5	0.05396	0.05274

Multivariate Jarque-Bera test on VAR(5) shows P-value equal to 5.551e-16, so the residuals of the model are not normally distributed.

3.3 Granger Causality

We test Granger causality for both variables. GNP does Granger-cause Unemployment. P-value on the Granger test is 0.01643. Unemployment does Granger-cause GNP. P-value on the Granger test is 0.01036.

3.4 Impulse response function

We use VMA representation of VAR to build impulse response function and forecast error decompositions.
Figure 7 shows the impulse response functions. Upper panels represents the reactions of variables to the shock of GNP, lower panels show the reactions of variables to the shock of unemployment.
The graphs indicate that all shocks affect both variables permanently, which contradicts the Blanchard, Quah (1989). GNP shock positively affects GNP and negatively affects Unemployment. Unemployment shocks negatively affects GNP and positively affects unemployment.
The variables are volatile for the first 12 periods after the shock. After the 12 periods they seem to reach a new steady state and the volatility reduces.
VMA impulse functions contradicts Blanchard and Quah results.

3.5 Forecast error variance decomposition

Figure 9 shows forecast error variance decomposition. Evidently, more than 90% of variance of GNP is explained by shocks in GNP. Variance of unemployment is more affected by shocks in GNP, they explain approximately 20% of variance of this variable.

3.6 Forecast

Figure 8 shows forecast plots of VAR(5).

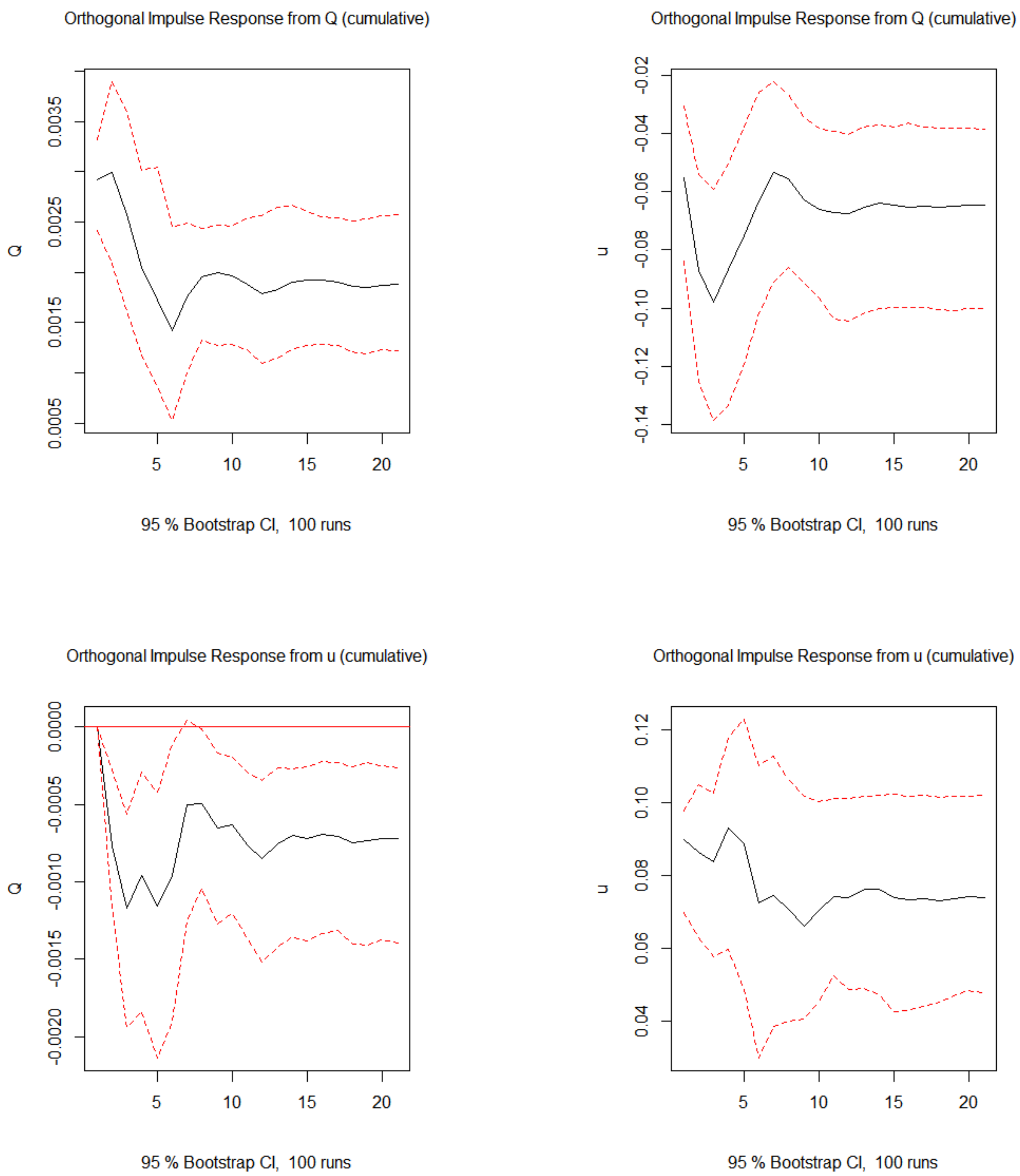

Figure 7: Impulse response function

4 Structural VAR

Blanchard, Quah assumptions does not seem to apply to the new data (1983-2021) - the log(Y) is not I(1). We tested ΔY using both ADF and KPSS test. Both tests showed that ΔY is non-stationary. However, both variables are stationary after deseasoning and detrending. Since Blanchard-Quah requirements is not satisifed, SVAR was derived through recursive identification.

4.1 Impulse response function

Figure 10 shows the impulse responses for recursively identified SVAR. We can see that in this identification shocks do not have permanent effects. The volatility after shock lasts for 20 periods. Shock of GNP seems to have a less

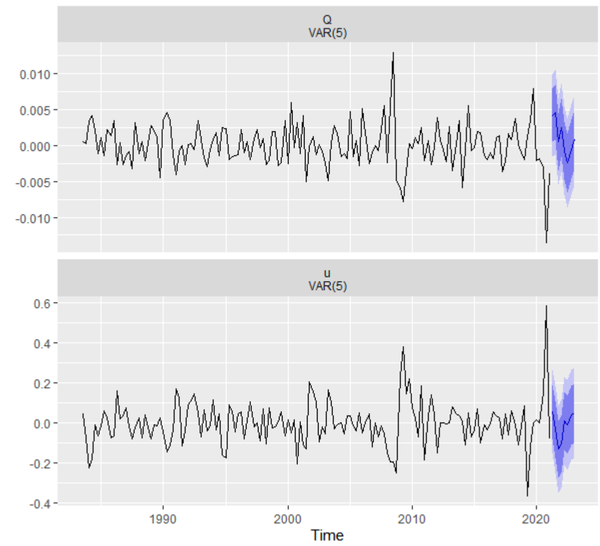

Figure 8: VAR(5) forecast

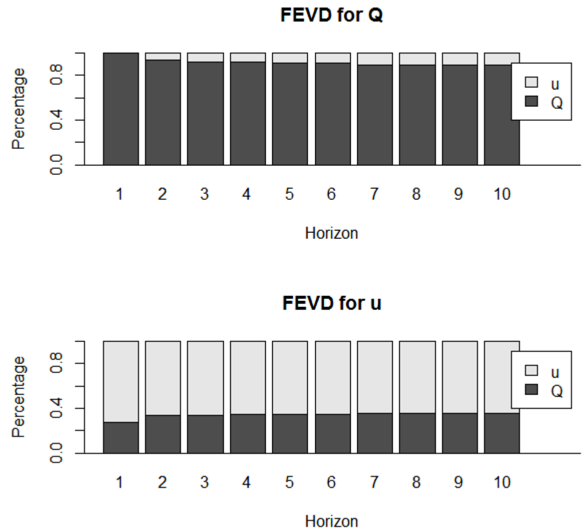

Figure 9: FEVD

powerful effect on volatility of the variables. The most unstable is the reaction of GNP to the shock of Unemployment, since the direction of the effect changes the most.
Unemployment shocks have a negative effect on output and vice versa.
The result of the identification through recursive identification contradicts the Blanchard and Quah conclusions.

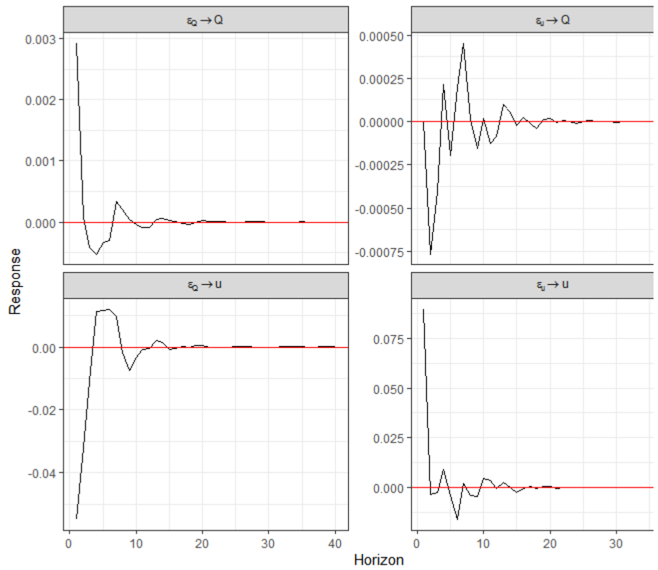

Figure 10: Cholesky identification IRF

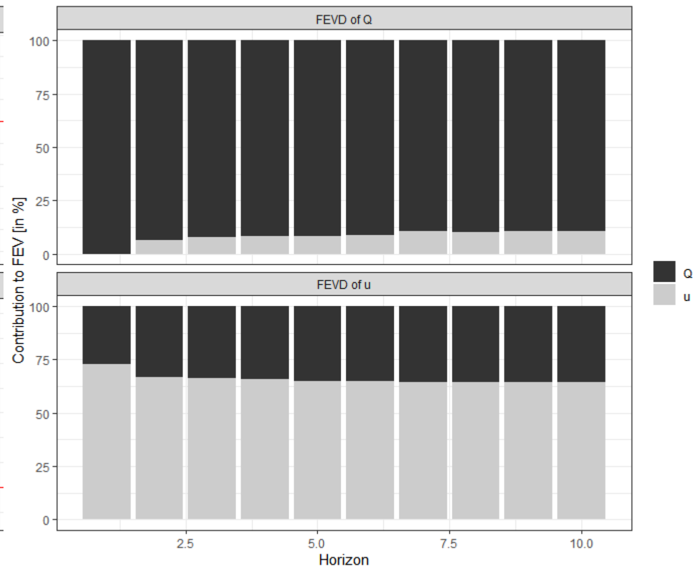

Figure 11: Cholesky identification FEVD

4.2 Forecast error variance decomposition

Figure 11 shows the forecast error variance decomposition for Cholesky identified SVAR. Cholesky's decomposition shows the same percentages of variance explained by the two variables.

4.3 Economical Interpretation

Both IRFs show that the GNP and Unemployment have a negative relation. This is consistent with the theoretical knowledge, that Unemployment is a countercyclical variable.
Forecast error variance decomposition shows that unemployment is more affected by the shocks in GNP then vise versa. These results are in line with the fact that unemployment is a lagging indicator.

Our model shows that there is no long term effects from both shocks, which contradicts Blanchard, Quah findings. Both variables reach a steady state after some amount of fluctuation.