TU Dortmund University
Department of Statistics
Summer 2022

Prof. Dr. Matei Demetrescu
M.Sc. Alexander Gerharz
M.Sc. Karsten Reichold

Case Studies – Project 1
One-Quarter-Ahead Forecasts of US GDP Growth
A Vector Autoregression Approach

Deadline: 5 May 2022, at 2:10 pm
Submission: Submit your report (together with corresponding `R` code in a separate file) electronically via Moodle **and** as a printed version (without `R` code) in the course.

This project is concerned with forecasting US real gross domestic product (GDP) growth. With y_t denoting US real GDP at time $t = 1, \ldots, T$, quarterly GDP growth (expressed in percent) is defined as

$$g_{y,t} := \frac{y_t - y_{t-1}}{y_{t-1}} \times 100, \tag{1}$$

$t = 2, \ldots, T$.[1] A simple one-quarter-ahead forecast of GDP growth at time $t+1$ given the information up to time t is give by $\hat{g}_{y,t+1|t} := \hat{a}_{0|t} + \hat{a}_{1|t} g_{y,t}$, where $\hat{a}_{0|t}$ and $\hat{a}_{1|t}$ are the ordinary least squares (OLS) estimates of $a_0, a_1 \in \mathbb{R}$ in the autoregression model of order one (AR(1) model)

$$g_{y,s} = a_0 + a_1 g_{y,s-1} + \epsilon_s,$$

$3 \leq s \leq t$. Instead of forecasting GDP growth using only its own history, it might be advantageous to exploit information contained in several additional economic variables denoted by $x_{1,t}, \ldots, x_{n,t}$. To this end, let $X_t := [g_{y,t}, x_{1,t}, \ldots, x_{n,t}]'$ and consider the vector autoregressive model of order $p \geq 1$ (VAR(p) model)

$$X_s = A_0 + A_1 X_{s-1} + \ldots, A_p X_{s-p} + \varepsilon_s, \tag{2}$$

$2 + p \leq s \leq t$, where $A_j \in \mathbb{R}^{(n+1)\times(n+1)}$, $j = 0, 1, \ldots, p$. The corresponding one-quarter-ahead forecast of GDP growth at time $t+1$ given the information up to time t is given by the first element of

$$\hat{X}_{t+1|t} := \hat{A}_{0|t} + \hat{A}_{1|t} X_t + \ldots + \hat{A}_{p|t} X_{t-p+1},$$

with $\hat{A}_{0|t}, \hat{A}_{1|t}, \ldots, \hat{A}_{p|t}$ denoting the OLS estimates of $A_0, A_1, \ldots, A_p$ in (2), see, e.g., Kilian and Lütkepohl (2017, Chapter 2.3).

Go to the FRED-QD database and download the file `2022-02.csv`.[2] Create a separate Excel file with the following time series from 1959Q1 to 2021Q4:

`GDPC1`: real GDP (billions of chained 2012 dollars)

`CUMFNS`: manufacturing industry capacity utilization (percent of capacity)

---

[1] This definition uses the previous quarter (i.e., y_{t-1}) as a reference point. An alternative definition uses the same quarter in the previous year (i.e., y_{t-4}) as a reference point. You could discuss advantages and disadvantages of either definition in your report.

[2] Make sure you download the file containing quarterly data.

`UNRATESTx`: unemployment rate less than 27 weeks (percent)

`CPIAUCSL`: consumer price index for all urban consumers (all items, index 1982-1984=100)

`FEDFUNDS`: effective Federal Funds rate (percent)

`M1REAL`: real M1 money stock (billions of 1982-84 dollars), deflated by CPI

`S&P 500`: S&P's common stock price index

The Appendix in McCracken and Ng (2020) provides a detailed description of the variables.

Before you start, calculate the quarterly growth rates (expressed in percent, see (1)) of `GDPC1`, `CPIAUCSL`, `M1REAL` and `S&P 500` (and potentially add them to your Excel file). Use the growth rates rather than the levels of these variables in the following.[3]

Tasks:

a) Generate a plot that shows the seven variables over the period 1959Q2–2021Q4.

b) Create one-quarter-ahead forecasts of GDP growth for 1959Q3 to 2021Q4 using an AR(1) model. Illustrate the forecasts together with the actual growth rates in one plot. Calculate the root mean squared forecasting error over the whole period.[4]

c) Create one-quarter-ahead forecasts of GDP growth for 1959Q3 to 2021Q4 using a VAR(1) model containing the six additional variables. Illustrate the forecasts together with the actual growth rates and the forecasts based on the AR(1) model in one plot, potentially extending the plot in (b). Calculate the root mean squared forecasting error over the whole period.

d) Explain the concept of Granger causality (compare the discussion in Kilian and Lütkepohl, 2017, Chapter 2.5 and Chapter 7) and identify the variables in the VAR(1) model that Granger cause GDP growth using all information up to 2021Q4.

e) Create one-quarter-ahead forecasts of GDP growth for 1959Q2+p to 2021Q4 using a VAR(p) model containing the six additional variables. To determine p use Akaike's (1974) information criterion as described in Kilian and Lütkepohl (2017, p. 55).[5] Illustrate the forecasts together with the actual growth rates and the forecasts based on the AR(1) and VAR(1) models in one plot, potentially extending the plot in (c). Calculate the root mean squared forecasting error over the whole period.

f) Compare the performance of the different forecasting approaches. Does it suffice to merely fit an AR(1) model? In which periods is it particularly advantageous to exploit information in other variables? Are there any periods in which the forecasts fit poorly to the actual observations? Are these periods potentially related to "economic events"? Is there any structure in the data not captured by the methods used? How could you account for these features?

[3] Note that the growth rate of `CPIAUCSL` is known as the inflation rate.

[4] Using generic notation, let $\hat{z}_{t+1|t}$ be a forecast of the variable z at time $t+1$ given the information up to time t and let z_{t+1} denote the actual value of z at time $t+1$. The mean squared forecasting error over the whole period is then defined as $\text{MSFE}_z := \sqrt{T^{-1}\sum_{t=1}^{T-1}(z_{t+1} - \hat{z}_{t+1|t})^2}$.

[5] Restrict the maximum number of lags to four.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**, 716 – 723.

Kilian, L., Lütkepohl, H. (2017). Structural Vector Autoregressive Analysis. Cambridge University Press, Cambridge.

McCracken, M.W., Ng S. (2020). FRED-QD: A Quarterly Database for Macroeconomic Research. Federal Reserve Bank of St. Louis Working Paper 2020-005.
URL https://doi.org/10.20955/wp.2020.005