Machine Learning for Central Banking

Daily Takeaways from BSP Lectures

Paul D. McNelis, S.J.

October 2023

Outline

Introduction

- No free lunch in regression analysis:
- We have to filter our data.
- We have to know the questions we are asking.
- We have to check the assumptions of the regression model.

Regression

- Single equation is not the way to go.
- We have to take seriously the assumptions of Gauss-Markov.
- Are the regressors exogenous? Is the disturbance term IID?
- Moving to first-differenced weekly data does not get rid of serial dependence.
- Ergodicity: we can move to monthly first differences or quarterly.
- No free lunch: we lose observations.

Macroeconomics and Reality

- Chris Sims developed the Vector Autoregressive (VAR) model in the 1980s.
- It is one of the principal workhorses of policy analysis.
- As Sargent notes, it is a state-space model that brings together good dynamic econometrics with good dynamic economics.
- All variables depend on other *lagged* endogenous variables.
- Adding more lags removes higher-order serial correlation, as shown by the Ljung-Box Q statistic.

Interpreting the VAR

- We can use Granger causality to see if one variable is a cause or significant predictor of another variable.
- We can use impulse response functions to see how one-time changes in one variable affect the dynamic response of other variables.
- We can use Forecast Error Variance Decomposition (FEVD) to see the relative importance of one variable for the overall variance of other variables.
- We can use the FEVD matrix to see if one variable has more outward or inward connectedness to other variables in the system.
- The relative strength of bivariate connectedness can be visualized with Directional Graphics.

Questions about VAR Models

- Are the results of VAR regressions robust to the choice of the number of lags?
- As we increase the dimensions of the VAR, or lags, or both, we rapidly increase the number of parameters.
- For a VAR system of 10 variables with a lag structure of 5, we have 510 parameters, if we also include constant terms.
- So a VAR rapidly consumes degrees of freedom.
- There is also the ever-present danger of *overfitting*.
- Another way of putting things: we encounter the bias-variance trade-off.

Selection Criteria

- Need for *regularization* criteria
- After getting rid of serial correlation, one can add more lags and get a better fit
- So we need to handicap our models: adjust the Likelihood L by the number of parameters k for a given number of observations n:
- Akaike: $AIC = -2\ln(L) + 2k$
- Schwartz: $BIC = -2\ln(L) + k\ln(n)$
- Hannan-Quinn: $HQIC = -2\ln(L) + 2k\ln(\ln(n))$