Econ 482

Giorgio Primiceri

October 6, 2016

# PROBLEM SET 1

SUBMIT BY EMAIL BEFORE TUESDAY, OCTOBER 25

## 1. Exercise 1

(1) Consider the quarterly dataset "dataVARmedium.mat," which starts in 1959:Q1 and ends in 2008:Q4, and includes (i) log-real GDP; (ii) log-GDP deflator; (iii) the federal funds rate; (iv) log-real consumption; (v) log-real investment; (vi) log-hours worked; (vii) log-real wages. Estimate recursively a 5-lag VAR using flat priors and conditioning on the initial 5 observations. Start with the estimation sample that ranges from 1959:Q1 to 1974:Q4, and then iterate the same procedure extending the estimation sample, one quarter at a time, until the end of the sample, i.e. 2008:Q4. For each estimation sample, set the VAR coefficients at their posterior mode and generate the 1–quarter and 4–quarter–ahead forecasts of log-real GDP and log-GDP deflator (ignoring both estimation uncertainty and the uncertainty about the realization of the shocks). Denote these out-of-sample forecasts by $\hat{y}_{i,t+h|t}$, where $h = 1$ or 4, and $i = 1$ or 2, consistent with the position of log-real GDP and the log-GDP deflator in the vector of endogenous variables $y$. Using $\hat{y}_{i,t+h|t}$, compute the forecast of the $h-$period ahead average growth rates of GDP and the GDP deflator as

$$\hat{z}_{i,t+h|t} = \frac{1}{h}\left[\hat{y}_{i,t+h|t} - y_{i,t}\right].$$

Compute the mean squared forecast error of these forecasts of $z$ (MSFE, i.e. the average squared deviation of the model-based forecasts ($\hat{z}_{i,t+h|t}$) from the actual realized values ($z_{i,t+h|t}$)).

(2) Repeat the exercise in part (1) by estimating the VAR with the following prior

$$\begin{aligned}
\beta|\Sigma &\sim N\left(b, \Sigma \otimes \Omega\right) \\
\Sigma &\sim IW\left(\Psi, d\right),
\end{aligned}$$

where I am using the same notation we have used in class. Set $d = n + 2$, and $\Psi$ equal to a diagonal matrix with $\left[\hat{\sigma}_1^2, ..., \hat{\sigma}_n^2\right]$ on the main diagonal, where $\hat{\sigma}_j^2$ is the variance of the residuals of an AR(1) process (with a constant) estimated using variable $j$ and the entire sample available at the moment of the forecast (therefore, $\hat{\sigma}_j^2$ changes each time you extend the sample and re-estimate the model). Set $b$ and $\Omega$ in order to reproduce the mean and the variance of the Minnesota prior, as we have seen in class. In other words, set $b$ and $\Omega$ so that

$$\begin{aligned}
(B_s)_{i,j} | \Sigma &\sim N\left(\gamma(s,i,j), \frac{\Sigma_{ii}}{\hat{\sigma}_j^2}\left(\frac{\lambda}{s}\right)^2\right), \\
\gamma(s,i,j) &= \begin{cases} 1 \text{ if } s = 1 \text{ and } i = j \\ 0 \qquad\quad \text{otherwise} \end{cases} \\
cov\left((B_s)_{i,j}, (B_r)_{h,m} | \Sigma\right) &= \begin{cases} \frac{\Sigma_{ih}}{\hat{\sigma}_j^2}\left(\frac{\lambda}{s}\right)^2 & \text{if } m = j \text{ and } r = s \\ \qquad 0 & \quad\text{otherwise} \end{cases} .
\end{aligned}$$

Set $\lambda = 0.2$ and compare the MSFE results to those obtained in part (1).

(3) Repeat the exercise in part (2), but this time choose $\lambda$ optimally at each point in time, by maximizing its posterior under flat hyperprior (refer to the formulas we have seen in class, also reported in Giannone, Lenza, and Primiceri, 2015). Compare the MSFE of this forecasting exercise to those of part (1) and part (2). Also, plot the optimal choice of $\lambda$ over time.

## 2. Exercise 2

(1) Using the same dataset, estimate a VAR with 5 lags on the entire set of data and a flat prior on the model parameters. Starting from the initial 5 observations (those you have been conditioning on for the estimation), and setting the VAR parameters to their posterior mode, compute the $(T - p)$–period ahead forecast ($T$ denotes the sample size and $p$ is the number of lags). In producing these forecasts, ignore the uncertainty about the parameters and the realization of the shocks. This is what we have defined in class as the deterministic component implied by the estimated model. Plot projections and actual realizations for the following variables: (i) log-real GDP; (ii) quarterly price inflation; (iii) the federal funds rate; (iv) quarterly nominal wage inflation; (v) log-labor share (defined as $log\left(\frac{\text{nominal wages * hours worked}}{\text{prices * real GDP}}\right)$); (vi)

log-consumption ratio (defined as $log\left(\frac{\text{consumption}}{GDP}\right)$). How much of the variation of each time series is explained by the deterministic component? Do you find this plausible?

(2) Repeat the same exercise using the Minnesota prior (exactly the same prior of exercise 1, with $\lambda = 0.2$). Do you notice any sizable difference?

(3) Repeat the same exercise by using the sum-of-coefficients prior that we have seen in class. Experiment with the hyperparameter $\mu = 1$ and 5. Comment on the results.