Are RNNs Useful for Macroeconomic Forecasting?

Emre Enes Yavuz, Carl Hallmann, Federico Puglisi

September 26, 2022

Setup

Data

- Updated Stock and Watson (2012) quarterly time series for the US.
- Forecast first differences of GDP, GDP Deflator, Consumption, Investment, FedFunds rate.
- Use these variables for forecasting (mostly).

Benchmark

- Bayesian VAR with hyper-prior as in Giannone, Lenza, Primiceri (2015).

Criterion

- RMSFE out of sample since 2006q1.

Setup

Models

- Vanilla RNN.
  - One RNN layer, and one linear layer.
- Vanilla GRU.
  - One GRU layer, and one linear layer.
- GRU nesting VAR(1).
  - One GRU layer, one linear layer using original variables and GRU output.
  - Nice because regularizing shrinks it towards white noise.
  - This is also what BVAR does.

Setup

Models continued

- GRU nesting VAR(1) predicting all variables 4 steps ahead simultaneously.
  - Using same hidden states for more predictions should force them to identify more meaningful information.
- Vanilla RNN with higher dimensional auto-encoded data (AE-RNN).
  - Use 210 variables from Stock and Watson data instead of 5.
  - Train auto-encoder.
  - Predict next encoded data using encoded data with vanilla RNN.
  - Decode prediction to obtain prediction for all variables.
  - Use predictions for same 5 variables to compare performance.

Results

RMSFE	RNN	GRU	GRU-VAR	GRU-VAR-4	AE-RNN	BVAR
Mean	0.7	0.7	0.69	(0.73)	0.68	0.72
GDP	2.54	2.4	2.47	2.4	2.15	2.21
CPI	0.87	1.04	1.01	1.07	1.08	1.28
C	2.27	2.35	2.49	2.23	1.68	2.40
I	12.53	13.16	11.47	11.93	11.32	11.23
FedF	0.51	0.42	0.36	0.37	0.31	0.43
Parameters	90	210	94	(400)	(54k+1.2k)	55

Table 1

All reported RMFSE are for training data. RMSFE Mean is measured in standard deviations, equally weighted over all variables. RMSFE for GDP (output), CPI (inflation), C (consumption) and I (investment) are annual growth rates in percentage points. RMSFE for FedF (Federal Funds Rate) is measured in percentage points. Brackets indicate incompatibility because of different forecasting horizons and number of predicted variables.

Other Insights

What Models perform well?

- Any sufficiently flexible, well regularized model achieves RMSES similar to those above.
- AE-RNN only model outperforms BVAR.
- This is likely because it uses more information for prediction.
- This is despite mediocre performance of auto-encoder in test data.

Long Horizons

- RNNs have RMSFEs of about half the size of the BVAR.
- Natural that RNN outperforms, since it minimizes long horizon forecast error, while BVAR minimizes function of one step ahead forecast error.

Extensions

1 Compare forecasting performance of benchmark Bayesian FAVAR with B-FAVAR using Factors extracted using Auto-Encoder Techniques (ML)

Q1 **Does the non-linearity of the Auto-Encoder allow to extract additional relevant information/factors ?**

2 Compare B-FAVAR vs B-FAVAR with factor weights shrunk to zero using ML (Attention Model).

Q2 **Introducing a State-Dependent Ridge Non-Linearity add relevant information/factors?**

3 Use Machine Learning to predict micro-level data and then now-cast aggregate level data based on Micro-Level.

Q3 **Is there more content to be extracted from Micro Data or are Macro-Data already containing all info sufficient for Macro-Forecasting?**

Extensions

4 Using Machine learning techniques, we can construct the factors maximizing contemporaneously dual objective function:
   a Explained Variance of data (from which factor is extracted)
   b Predictive power onto Macro-Aggregates

Q4 Would this dual objective help? (Same logic as Bayesian Shrinkage maximizing fit and 1-step ahead Predictive Density?)