Nowcasting Economic Activity

a comparison of nowcasting, econometrics and machine learning models

Romain legrand
Capital Fund Management

Friday 30$^{\text{th}}$ October, 2020

Outline

1.1 - Context: Now-casting.com

- now-casting.com trial at CFM
- Now-casting.com: real time forecasts (nowcasts) for quaterly GDP growth
- based on large monthly macro dataset
- use the Dynamic Factor Model of Giannone et al. (2008)

Conclusion

- GDP: works well for some countries (Euro area, Japan), poorly for others (US)
- other features: some well predicted, others terribly predicted

1.1 - Context: Now-casting.com

- now-casting.com trial at CFM
- Now-casting.com: real time forecasts (nowcasts) for quaterly GDP growth
- based on large monthly macro dataset
- use the Dynamic Factor Model of Giannone et al. (2008)

Conclusion

- GDP: works well for some countries (Euro area, Japan), poorly for others (US)
- other features: some well predicted, others terribly predicted

1.2 - The project

- replicate the dfm (possibly: improve the results)
- compare its performance with other models

Proposition

- nowcasting dataset for the US: quarterly GDP + monthly features
- nowcasting: dynamic factor model, MIDAS, mixed frequency Bayesian VAR
- econometrics: Bayesian VAR, time-varying Bayesian VAR
- machine learning: LSTM, boosting, random forest
- compare predictive performance

1.2 - The project

- replicate the dfm (possibly: improve the results)
- compare its performance with other models

Proposition

- nowcasting dataset for the US: quarterly GDP + monthly features
- nowcasting: dynamic factor model, MIDAS, mixed frequency Bayesian VAR
- econometrics: Bayesian VAR, time-varying Bayesian VAR
- machine learning: LSTM, boosting, random forest
- compare predictive performance

2.1 - a US macro dataset

- quarterly series of real GDP growth
- 30 monthly series of macro variables
- sample dates: 1992m1 - 2020m8 (T = 320 m /106 q)
- series made stationary by differentiation / growth rate

Limit of the exercise

- pseudo real time only!

2.1 - a US macro dataset

- quarterly series of real GDP growth
- 30 monthly series of macro variables
- sample dates: 1992m1 - 2020m8 ($T = 320$ m /106 q)
- series made stationary by differentiation / growth rate

Limit of the exercise

- pseudo real time only!

2.2 - dataset overview

- business indicators: pmi, business outlook survey, business confidence index, consumer confidence index
- production and sales: industrial production, Markit GDP index, business sales, residential sales, inventories
- labor and wage: unemployment rate, employment, weekly hours, hourly earnings, consumer credit, personal income
- macro aggregates: federal debt, exports, imports
- prices: ppi, cpi
- money and credit: monetary base, bank asset and liabilities, mortgage rate
- interest rates and finance: federal funds rate, treasury bill, treasury 10 years, effective exchange rate, spot Euro/US, NYSE index, VIX

3.1 - Dynamic factor: the model in brief

- features x_t are linear functions of a few factors f_t:
  $x_t = \mu + \Lambda f_t + \xi_t \qquad \xi_t \sim N(0, diag(\psi))$
- factors f_t follow an AR process:
  $f_t = A f_{t-1} + B u_t \qquad u_t \sim N(0, I_q)$
- gdp growth $\hat{y}_t$ is a linear function of projected factors:
  $\hat{y}_t = \alpha + \beta \hat{f}_t \qquad \hat{f}_t = E(f_t | \Omega_t)$

3.2 - estimation procedure

- raw estimates of f_t by running PCA on x_t.
- estimation of μ, Λ, A and B from f_t.
- model in state-space form: obtain smoothed estimates $\hat{f}_t$ from Kalman filtering
- estimate α and β by simple OLS

Predictions

- $\hat{y}_{t+h} = \alpha + \beta \hat{f}_{t+h}$
- $\hat{f}_{t+h}$ from Kalman filter

3.2 - estimation procedure

- raw estimates of f_t by running PCA on x_t.
- estimation of μ, Λ, A and B from f_t.
- model in state-space form: obtain smoothed estimates $\hat{f}_t$ from Kalman filtering
- estimate α and β by simple OLS

Predictions

- $\hat{y}_{t+h} = \alpha + \beta \hat{f}_{t+h}$
- $\hat{f}_{t+h}$ from Kalman filter

4.1 - Bayesian VAR: the model in brief

- VAR model:
  $y_t = c + A_1 y_{t-1} + \cdots + A_p y_{t-p} + \varepsilon_t \qquad \varepsilon_t \sim N(0, \Sigma)$
- stack $c, A_1, \cdots, A_p$ in a single vector β.
- "estimate the model" = apply Bayes rule to get the posterior distribution of β and Σ:
  $\pi(\beta, \Sigma | y) \propto f(y|\beta, \Sigma)\pi(\beta, \Sigma)$
- likelihood function is normal:
  $f(y|\beta, \Sigma) =$
  $(2\pi)^{-nT/2}|\bar{\Sigma}|^{-1/2} \exp\left[-\frac{1}{2}(y - \bar{X}\beta)'\bar{\Sigma}^{-1}(y - \bar{X}\beta)\right]$

- prior for β is normal: $N(b_0, \Omega_0)$:
  $\pi(\beta) = (2\pi)^{-nq/2}|\Omega_0|^{-1/2} \exp\left[-\frac{1}{2}(\beta - b_0)'\Omega_0^{-1}(\beta - b_0)\right]$
- b_0 and Ω_0 set by following the "Minnesota" prior (Litterman (1986))
- b_0: most economic variables follow a random walk: 1 on first ar coefficient, 0 otherwise
- Ω_0: "Bayesian shrinkage", prior variance is tighter for further lags and cross equation coefficients

- prior for Σ is inverse Wishart: $IW(S_0, \nu_0)$:
  $\pi(\Sigma) =$
  $(2^{\nu_0 n/2}\Gamma_n(\nu_0/2))^{-1}|S_0|^{\nu_0/2}|\Sigma|^{-(\nu_0+n+1)/2} \exp\left[-\frac{1}{2}tr\{\Sigma^{-1}S_0\}\right]$
- S_0 and ν_0 chosen to match OLS estimate $\hat{\Sigma}$.

- prior for β is normal: $N(b_0, \Omega_0)$:
  $\pi(\beta) = (2\pi)^{-nq/2}|\Omega_0|^{-1/2} \exp\left[-\frac{1}{2}(\beta - b_0)'\Omega_0^{-1}(\beta - b_0)\right]$
- b_0 and Ω_0 set by following the "Minnesota" prior (Litterman (1986))
- b_0: most economic variables follow a random walk: 1 on first ar coefficient, 0 otherwise
- Ω_0: "Bayesian shrinkage", prior variance is tighter for further lags and cross equation coefficients

- prior for Σ is inverse Wishart: $IW(S_0, \nu_0)$:
  $\pi(\Sigma) =$
  $(2^{\nu_0 n/2}\Gamma_n(\nu_0/2))^{-1}|S_0|^{\nu_0/2}|\Sigma|^{-(\nu_0+n+1)/2} \exp\left[-\frac{1}{2}tr\{\Sigma^{-1}S_0\}\right]$
- S_0 and ν_0 chosen to match OLS estimate $\hat{\Sigma}$.

4.2 - estimation procedure

- impossible to marginalize the posterior analytically to get $\pi(\beta|y)$ and $\pi(\Sigma|y)$.
- Gibbs sampling algorithm: draw sequentially from the *conditional* posteriors $\pi(\beta|y,\Sigma)$ and $\pi(\Sigma|y,\beta)$.
- repeat sufficiently to converge to $\pi(\beta|y)$ and $\pi(\Sigma|y)$
- recover an *empirical* posterior distribution

Predictions

- $y_{t+h} = c + A_1 y_{t+h-1} + \cdots + A_p y_{t+h-p} + \varepsilon_t \qquad \varepsilon_t \sim N(0,\Sigma)$
- β and Σ: use simulations from Gibbs sampler

4.2 - estimation procedure

- impossible to marginalize the posterior analytically to get $\pi(\beta|y)$ and $\pi(\Sigma|y)$.
- Gibbs sampling algorithm: draw sequentially from the *conditional* posteriors $\pi(\beta|y,\Sigma)$ and $\pi(\Sigma|y,\beta)$.
- repeat sufficiently to converge to $\pi(\beta|y)$ and $\pi(\Sigma|y)$
- recover an *empirical* posterior distribution

Predictions

- $y_{t+h} = c + A_1 y_{t+h-1} + \cdots + A_p y_{t+h-p} + \varepsilon_t \qquad \varepsilon_t \sim N(0,\Sigma)$
- β and Σ: use simulations from Gibbs sampler

5.1 - LSTM: model tunning

- model: features x_{t-1} and x_{t-2} as input, x_t and GDP_t as output
- MSE used as loss function
- dataset is split in train/test samples (80-20%)
- MSE on test is lowest with only one layer
- MSE on test decreases with units, and stabilizes at 1000 units
- optimizer has little impact: AdaDelta, RMSprop, Adam all yields same estimates
- 500 epochs to train the model

5.1 - LSTM: model tunning

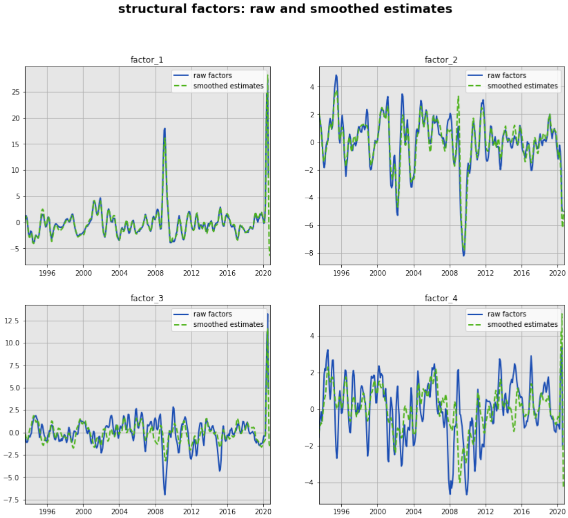
Predictions

- obtained sequentially
- predict x_{t+1} and GDP_{t+1} from x_t and x_{t-1}
- predict x_{t+2} and GDP_{t+2} from x_{t+1} and x_t ...

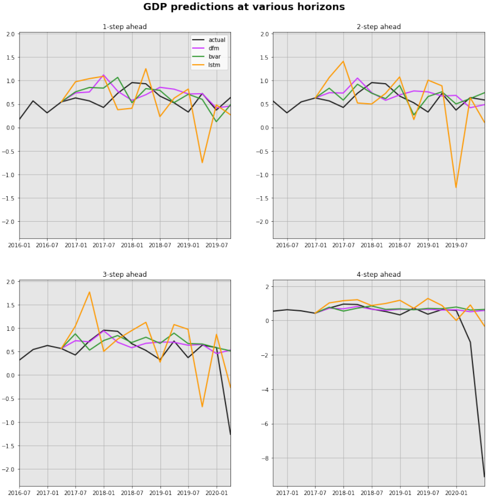
6.1 - Forecast exercise

- sequential sample: start from 1993m7 - 2016m7, then increase sequentially by one month, for 36 months (3 years)
- for each sample: produce forecasts at $t+1$, $t+2$, $t+3$ and $t+4$
- forecasts are produced for GDP and all features in x_t
- for each model, sample, horizon and features, compute $\text{RMSE} = \sqrt{(x_{t+h} - \hat{x}_{t+h})^2}$
- consider the average RMSE over the different samples

6.2 - Dynamic factors: illustration



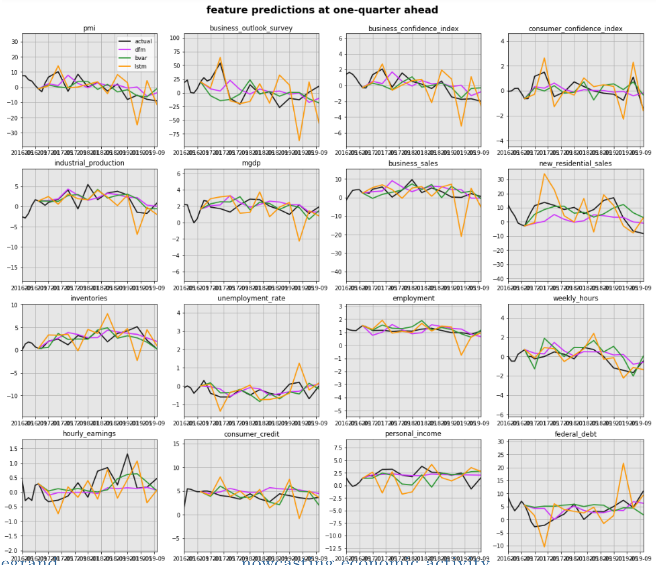
6.3 - GDP predictions: all models, all horizons



6.4 - GDP predictions: RMSE tables

	1q_ahead	2q_ahead	3q_ahead	4q_ahead
Dynamic_factor_model	0.223051	0.212275	0.212238	0.303178
Bayesian_VAR	0.227662	0.193084	0.323950	1.135312
lstm	0.473065	0.506470	0.562262	1.249043

6.5 - Feature predictions: all models, one quarter ahead



6.6 - feature predictions: RMSE tables

	dfm 1q	bvar 1q	lstm 1q	dfm 2q	bvar 2q	lstm 2q	dfm 3q	bvar 3q	lstm 3q
pmi	3.616116	4.733442	6.344117	4.831886	6.453893	10.824125	4.893616	5.297926	7.139917
business_outlook_survey	13.270595	17.982585	28.711299	16.543634	14.509579	31.680201	15.709294	15.441356	25.866550
business_confidence_index	0.609327	0.881891	1.363845	0.989711	1.258690	2.302627	1.019017	1.087684	1.429962
consumer_confidence_index	0.477035	0.765440	0.807741	0.647770	0.747017	1.210376	0.642065	0.712899	0.983758
industrial_production	1.572153	1.711532	2.215614	1.411702	1.641410	3.622593	1.724344	2.600201	3.123910
mgdp	0.645384	0.671864	1.414081	0.621369	0.582079	1.518936	0.643031	0.983495	1.690903
business_sales	2.144843	2.621306	5.153992	2.632747	2.273208	7.230065	2.282892	3.122980	5.378684
new_residential_sales	7.530733	6.245320	9.289625	7.606331	7.522068	15.877749	7.084518	7.490126	18.944343
inventories	1.016453	1.077648	2.382430	1.145364	1.179062	3.070826	1.237835	1.308704	3.690093
unemployment_rate	0.244735	0.335247	0.441188	0.271103	0.373083	0.789802	0.320082	0.381574	0.599678
employment	0.220218	0.236808	0.385780	0.278078	0.252483	0.774189	0.362348	0.318726	0.926000
weekly_hours	0.645843	0.885315	0.660819	0.578805	0.963105	1.111286	0.590445	0.863723	1.349622
hourly_earnings	0.348683	0.402865	0.532967	0.371051	0.401727	0.522619	0.374493	0.367376	0.530256
consumer_credit	1.180668	1.280583	1.832170	1.299718	1.229607	2.783546	1.434462	1.428367	2.501960
personal_income	0.761734	1.286746	2.061013	0.694108	0.928167	1.396335	0.793089	0.838226	1.184694
federal_debt	2.079696	3.910228	4.243902	2.787228	3.922008	6.690096	2.424600	3.140373	2.531684
exports	4.566586	3.573000	4.721985	5.286181	4.204801	13.265452	5.847206	5.839289	10.312996
imports	4.787769	3.414838	8.875460	4.568961	4.325930	10.501435	5.127416	6.158419	13.424716

Conclusion

GDP

- BVAR is best at short horizon, though dfm is quite close
- dfm is much better at longer horizons
- lstm does poorly

features

- dfm >> BVAR >> lstm
- paradox: dfm does well for features (not intended), and worse than BVAR for gdp (dfm main target)
- explanation: dfm trains at monthly frequency for features, quarterly frequency for gdp: less accurate!

Nowcasting Economic Activity

a comparison of nowcasting, econometrics and machine learning models

Romain legrand

Capital Fund Management

Friday 30$^{\text{th}}$ October, 2020

References

Giannone, D., Reichlin, L., and Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, (55):665–676.

Litterman, R. (1986). Forecasting with bayesian vector autoregressions: Five years of experience. *Journal of Business And Economic Statistics*, 4(1):25–38.