Discussion of Large-Scale Dynamic Predictive Regressions (Bianchi and McAlinn)

Helmut Farbmacher

November 4, 2019

Summary

- Decouple-Recouple Synthesis/Strategy (DRS)

$$\begin{aligned} y_t &= \beta' z_{t-1} + \epsilon_t \\ &= \text{Group 1} + \text{Group 2} + \ldots + \epsilon_t \,, \qquad \epsilon_t \sim N(0, \nu_t) \end{aligned} \tag{1}$$

  with Group j $= \beta_j z_{t-1,j}$.
- Run J smaller models

$$y_t = \beta_j' z_{t-1,j} + \epsilon_{tj} \,, \qquad \epsilon_{tj} \sim N(0, \nu_{tj}) \tag{2}$$

  where $\epsilon_{tj} = \beta_{-j}' z_{t-1,(-j)} + \epsilon_t$.
- Implicit assumption for $\epsilon_{tj} \sim N(\cdot)$: all z's with $\beta_z \neq 0$ are normally distributed?

Applications and Results

- 1st application: Forecasting annual inflation in the U.S.
- 2nd application: Forecasting excess returns in the U.S. across different industries.
- In both applications the Decouple-Recouple Synthesis/Strategy (DRS) improves the out-of-sample forecasting accuracy relative to competing models (for instance, the Lasso).

Comments

- Performance may depend strongly on grouping strategy. Particularly so, if predictors are correlated.
- Authors propose to base groups on outside information (e.g. economic knowledge).
- Competitors (e.g., Lasso) may also work better with this outside information. Group Lasso includes regressors groupwise (Yuan and Lin, 2007). Procedure works like a Lasso at the group level. A group of parameters will either be all non-zero or all zero.
- If predictors are highly correlated, then Elastic Net (Zou and Hastie, 2005) may be a better competitor. Ridge component in Elastic Net effectively decorrelates the predictors.

Monte Carlo simulation

- Monte Carlo simulation setup

$$y = -2z_1 + 3z_2 + 5z_3 + \epsilon \qquad \epsilon \sim N(0, 0.01)\,,$$
$$z_1 = \frac{1}{3}z_3 + \nu_1\,, \quad \nu_1 \sim N\left(0, \frac{2}{3}\right)\,, \quad z_2 = \frac{1}{5}z_3 + \nu_2\,, \quad \nu_2 \sim N\left(0, \frac{4}{5}\right)\,,$$
$$z_3 = \nu_3\,, \quad \nu_3 \sim N\,(0, 0.01)$$

- Two key characteristics of the Monte Carlo simulation:
  1. all covariables need to be correlated
  2. existence of omitted variables

Comments

- Monte Carlo simulation setup

$$y = -2z_1 + 3z_2 + 5z_3 + \epsilon \qquad \epsilon \sim N(0, 0.01)\,,$$
$$z_1 = \frac{1}{3}z_3 + \nu_1\,, \quad \nu_1 \sim N\left(0, \frac{2}{3}\right)\,, \quad z_2 = \frac{1}{5}z_3 + \nu_2\,, \quad \nu_2 \sim N\left(0, \frac{4}{5}\right)\,,$$
$$z_3 = \nu_3\,, \quad \nu_3 \sim N\,(0, 0.01)$$

- z_1 and z_2 are very weakly correlated with the omitted predictor z_3. In the limit we expect to see $corr(z_1, z_3) \approx 0.041$ and $corr(z_2, z_3) \approx 0.022$.
- As a consequence, the omitted variable bias is tiny. For instance, with OLS we would expect rel.bias$(\beta_1) \approx 0.0025$ and rel.bias$(\beta_2) \approx 0.0092$.
- Moreover, $corr(z_1, z_2) \approx 0.001$. Makes it probably easy for the DRS to work.
- Setup is far from being a high-dimensional DGP.