Discussion of Kong et al's paper: How Predictable are Components of the Aggregate Market Portfolio?

Philip H. Dybvig (大顽童)
Washington University in Saint Louis,
SMU, 西南财经大学, and 长江商学院

SMU Summer Camp 2009
July 12, 2009

Main points

Stastical model: no economics in the paper.

Estimate a large set of possible univariate models.

Combination forecasts do better

Questions

Data Snooping (in paper and across papers): how can we determine size robustly?

How to interpret combination forecasts (like Bayesian shrinkage or mixture estimator? prior on parsimony?)

What are the properties of the underlying series that would make a model with combination forecasts outperform a multivariate regression?

Can we assign any economic significance to the purely statistical data fitting? Or, are we better off just not trying to claim this is more than techical analysis?

Combination Forecasts and Shrinkage

OLS $(x'x)^{-1}x'y$ (demeaned vectors for simplicity in thinking about this)

With two covariates, if x_1 and x_2 are uncorrelated (in sample), the average estimator is *half* the bivariate estimator. In particular, if x_1 and x_2 are uncorrelated, the estimators for the betas are the same in the bivariate regression as in the individual univariate regressions. When we do prediction, we add their impact, we don't average them.

Combination Forecasts and Shrinkage: algebra

univariate OLS betas: $({x_1}'x_1)^{-1}{x_1}'y$ and $({x_2}'x_2)^{-1}{x_2}'y$ bivariate OLS betas:

$$\left(\begin{pmatrix} {x_1}' \\ {x_2}' \end{pmatrix} \begin{pmatrix} x_1 & x_2 \end{pmatrix}\right)^{-1} \begin{pmatrix} {x_1}' \\ {x_2}' \end{pmatrix} \begin{pmatrix} y_1 & y_2 \end{pmatrix}$$
$$= \begin{pmatrix} {x_1}'x_1 & 0 \\ 0 & {x_2}'x_2 \end{pmatrix}^{-1} \begin{pmatrix} {x_1}' \\ {x_2}' \end{pmatrix} \begin{pmatrix} y_1 & y_2 \end{pmatrix}$$
$$= \begin{pmatrix} ({x_1}'x_1)^{-1} & 0 \\ 0 & ({x_2}'x_2)^{-1} \end{pmatrix} \begin{pmatrix} {x_1}' \\ {x_2}' \end{pmatrix} \begin{pmatrix} y_1 & y_2 \end{pmatrix}$$
$$= \begin{pmatrix} ({x_1}'x_1)^{-1}{x_1}'y \\ ({x_2}'x_2)^{-1}{x_2}'y \end{pmatrix}$$

The corresponding predictor is *twice* the average of the univariate predictors.

Combination Forecasts and Ridge Regression

There is some flavor of a ridge regression (which adds a diagonal matrix to the covariance matrix of x, so the estimator is $(x'x+D)^{-1}x'y$) because it kills off the impact of the off-diagonal elements and shrinks the estimators. This can be given a Bayesian interpretation (as can shrinkage estimators) of the matrix D in terms of the variance of the prior for beta.

Combination Forecasts: Bayesian Interpretation

Suppose we assume that the true model is univariate and linear. We could estimate this by writing down a big multivariate model that nests all the univariate linear models. This estimation would be consistent but not very efficient. Instead, if we had an equal prior weighting on each of the models and a diffuse prior on the betas, we would get approximately (maybe exactly with the usual conjugate prior on the covariance matrix) the combination forecast model. I think this is one way of thinking about why the combination forecasts can do better than multivariate forecasts. It is in effect imposing a very strong prior on parsimony and restricting the search to a small set of models.

Recommendations

1. I would be more interested in a test of a theoretical model, but given that you are doing a pure technical exercise, do it without apologies and leave out the unconvincing economic interpretations.

2. Discuss why you think the combination forecasts do better than univariate forecasts.

3. At least mention the data snooping issue. How much does testing out-of-sample really help? Some of these variables were already known to work in this sort of a test; that is why we use them.