Discussion of Homework 1

November 15, 2017

I delineate below some reflections based on what I saw in your homework. These are just elements to think about. The graphs and citations are from some homeworks I graded.

1 Terminology

- Centered returns vs centered moving average:
  Centered returns means that their average (over some sample) is zero, i.e. we've substracted the mean.
  Centered moving average means that the moving average is symmetric, i.e. at t it consists of an average of $y_{t-k}, .., y_t, ..., y_{t+k}$ for some $k > 0$.

- Trend vs Tendency. Tough for French speakers as they translate the same. Same thing in fact between Expectation and Expectancy. Try to have a look at a dictionnary to understand the difference.

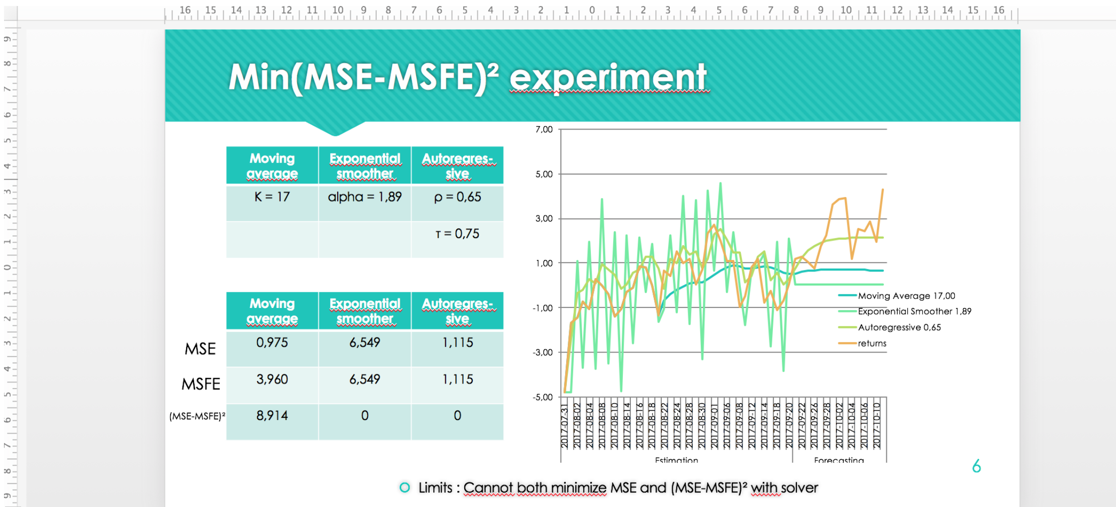
- About the first questions about the use of the proposed models.
  Do cite your sources when you answer a question based on what you read online. Also, rather than a general statement, try giving an example (using a graph or a hypothesized DGP)

2 Issues

- One-step vs multi-step forecasts
  - "*After calculating our parameters in the 'estimation' sample, we calculated the forecasts. We decided to calculate each forecast y(t+1) with its previous value y(t), instead of calculating all the forecasts with the same y(t).*
    We believed that using the same estimation y(t) for all our forecasts (y(t+1), y(t+2), y(t+3) etc) would lead to less precisions and accuracy."
    I agree but the point is whether ultimately you want to forecast the future, one period

at a time (so you do not forecast what happens in two periods) or whether you need a stream of forecasts at several horizons. It's not just a point of getting the most accurate forecast, rather it's a question of the objective (what is the forecast for?)

- MSE and MSFE
  - What model/estimated parameters would have given the best MSFE: an experiment carried out by some of you that consists in matching MSE and MSFE (see my comment starting with GC)



Considering the high difference between MSE and MSFE results, we tried to minimize (MSE-MSFE)² in order to observe the results out of curiosity.
While we were aware that this method would not allow to minimize both MSE and MSFE at the same time and would therefore present a risk of increasing MSE, we observed a significant reduction of the MSFE.

GC-> indeed it's interesting out of curiosity. In fact, it's also instructive when you look at the output
- ExpSmooth forces high oscillations, i.e. large forecast errors in sample, their amplitude matches the distance between forecasts and observations in the forecasting subsample. This is due to the fact that essentially ExpSmooth forecasts here are flat and constrained to the last smoothed value
- by contrast, the AR model forecasts converge progressively to the mean of the AR process, which is tau/(1-rho) so what you are doing here is to match this mean to the average of the observations over the forecast sample

  - an MSE close to the MSFE does not mean the forecast is accurate, it just means that you have managed to capture some stability in the dynamic modeling, so you're more likely to be entitled to extrapolate that what you observe in the sample at your disposal (the forecast accuracy) will carry on in the future.

  - Computing the MSE

    "*Moving Average Model : We had to improve the formulas about MSE, especially for the Moving Average Model. Indeed, this model leads to reduce the number of sample into the Evaluating subset, depending on its parameter. For example, if we choose to compute this model with the average of the last 30 observed values, we only have 7 values left into this subset which will have 30 previous values. Thus the MSE formula would have to be compute only depending on those 7 values.*"

A question: should you compute the MSE using the exact same observations for all models?

- Choice of sample split

1.1.a - Splitting the sample

Size of the sample : 52 entries

By default, the sample was splitted following this distribution :

- Estimation subset : 37 entries
- Forecasting subset : 15 entries
- R = Ratio estimation/forecasting : 71/29

Arbitration : Usually, we know that a good split ratio will lead to a ratio R around 80/20. However, we decided to keep this split (to start) for two reasons :

- It's important to keep a bigger estimation subset, in order to refine the model construction
- However, in order to keep a consistent MSFE (which leads to estimate the value of the model), we need to keep enough data in the forecasting. Thus we could have increased the size of the estimation subset with a bigger data sample.

if you find that some models seem to perform similarly, you may want to check wether this is a robust to a different split between training and testing subsamples.

- Forecasting and overfitting:

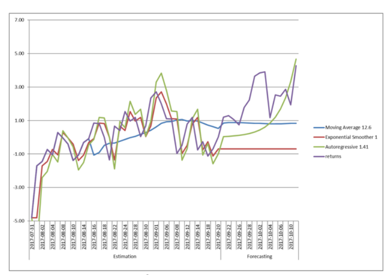
2. SOLVER

- **Optimal → average of MSFE**

Moving Average	Exponential Smoother	Autoregressive
12.6	1	1.41

MSFE Optimal

				pool
MSE	1.547	0.915	1.511	1.203
MSFE	3.514	10.443	3.292	5.750
Theil statistic.	2.27	11.41	2.18	4.78



MSE Optimal

				pool
MSE	0.967	1.131	0.720	0.822
MSFE	6.026	10.735	6.672	7.811
Theil statistic.	6.229	9.495	9.269	9.502

Interesting, here the AR model has been fitted to the forecasting subsample. Notice how it becomes explosive -¿ you see the danger over trying the get the best fit. The model is likely to be useless for forecasting as it will explode if you compute forecasts post testing sample.

3 MZ Regressions and DM test

- You need to get confidence intervals and/or perform tests in Mincer-Zarnowitz Regressions. In regression models, estimated coefficients are useless without corresponding uncertainty.
- The sign of the Diebold Mariano statistic is that of the corresponding difference in MSFEs so the whole point of the statistic is to compute a test and report the significance (or the pvalue).