MFE230E Problem Set 7

Due May 12 10:00am via bCourses

Note: Some questions are intentionally somewhat "vague." Use you own judgment to come up with the "best" choices and answers. Imagine these questions were an assignment in your internship.

1. Download the data set `PS7_LP_Portfolios.csv` from bCourses. The file includes excess returns of decile portfolios based on sorts of 37 characteristics. The data is described in more detail in Lettau and Pelger "Factors that Fit the Time Series and Cross-Section of Stock Returns" (see link on https://sites.google.com/view/martinlettau/).

   (a) Summarize the most important aspects of the data. (Use your own judgment about what is "most important".)

   (b) Compute the pricing errors from the 3- and 5-factor Fama-French models. Summarize the results.

   (c) Use the `numpy` library to compute the eigenvalues and eigenvectors of the covariance and correlation matrices of portfolio return sof the 370 portfolio returns. Construct principal components and compute their variances. Show scree plots for decompositions of the covariance and correlation matrices. Compare the results. What do you conclude?

   (d) Compare your results with those obtained by the `statsmodels PCA` package.

   (e) How many PCs are "significant"? Describe the PCs. Do they have any interesting economic interpretation?

   (f) Consider linear factor models with PCs as factors and evaluate the pricing errors. You may want to consider models with different sets of PCs.

   (g) How do the factor models based on PCs compare to the Fama-French models?

2. (a) Use the `statsmodels` library to estimate Kalman filter CAPM models with time-varying α_t and β_t for the four "corner" of the 25ME/BM-sorted FF portfolios. Summarize the results.

   (b) Repeat (a) for the 3-factor FF model.

   (c) Next, estimate the model for monthly Microsoft and IBM returns.

   (d) Compare the results for individual stocks to the results for portfolios.