School of Electronics and Computer Science
University of Southampton

COMP6041 Computational Finance 2015/16, Assignment 4 (of 4) 25%

Issue	03 May 2016
Due	18 May 2016

---

1. In a recent paper, Mahler [1] shows how the Kalman filter for sequential estimation of a state space model can be combined with the *Lasso* approach [2] for l_1 penalized regression. The result is to select lags (past values) of explanatory variables that are relevant for the prediction of future values of a time series. The work shows how residual signal arising from a time-series model might be explained by exogenous information in the environment.

   Implement this model and critically evaluate the claims in the paper:

   - First implement a simple second / third order autoregressive model on the index data and calculate the residual. The variance of this residual can be an estimate you can use as observation noise variance for the Kalman filter.
   - Implement a Kalman filter to estimate AR model parameters recursively. The process noise variance has to be tuned; observe if there is anything systematic about the role of this hyper-parameter. Implementing a Kalman filter is tricky, so you should first construct a synthetic time series (in which you know the correct answer to the AR model coefficients), gain some experience in working with this before applying the model to real data.
   - Model the error signal as a linear function of the variables considered in Mahler's paper ([1]) with a sparsity inducing regulariser (lasso) and explore their relative relevance.

   Write up your findings as a report of no more than six pages.

   - Use the convex programming package `cvx` from `http://cvxr.com/` to implement sparse regression. You are familiar with this from the index tracking exercise. Note you have used this tool in designing sparse index tracking portfolios.
   - `S&P500` time series can be obtained at `http://finance.yahoo.com`
   - Other data used in [1] can be obtained from `http://research.stlouisfed.org/fred2/`

Marking Scheme

Implementation fo the time series model	10
LagLasso to explain the residual	15

References

[1] N. Mahler, "Modeling the S & P 500 index using the Kalman filter and the LagLasso," in *Machine Learning for Signal Processing, 2009. MLSP 2009. IEEE International Workshop on*, Sept 2009, pp. 1–6.

[2] R. Tibshirani, "Regression srinkage and selection via the lasso," *Journal of the Royal Statistical Society, Series B*, vol. 58, no. 1, pp. 267–288, 1996.

Mahesan Niranjan
May 2016