Proceedings of 9th Annual American Business Research Conference
10 - 11 July 2017, Sheraton LaGuardia East Hotel, New York, USA
ISBN: 978-1-925488-41-8

A New Approach for Forecasting Stock Prices using Principal Components

Mahsa Ghorbani* and Edwin K. P. Chong**

Stock price prediction is one of the most challenging problems in the finance world. In this paper we focus on forecasting stock prices over a short period of time using historical price data. Given past data, the multivariate conditional mean is used as a point estimator to minimize the mean square error of prediction. However, the calculation of the condition mean and covariance involves the numerical inverse of an ill-conditioned matrix, leading to numerical issues. To overcome this problem, we propose a filtering operation based on principle component analysis. Projecting the noisy observation onto a principle subspace leads to a significantly better numerical condition. We use daily historical price data for General Electric Company from 1996 to 2015 to illustrate our method, which shows promising results in terms of the estimation performance.

Field/Track: Finance (Stock Market, Financial Modelling, Quantitative Finance)

1. Introduction

Prediction of stock prices is one of the most widely studied and challenging problems and is receiving considerable attention from researchers. The successful forecasting of potential stock prices can provide significant profit. We focus on estimating future price values based on past data. More specifically, assume we are given the price values for the past 90 days, and based on that, we want to estimate the price for the next 10 days. Given historical data, multivariate conditional mean is a suitable point estimator because it minimizes the mean square error of the estimation. However the numerical results cannot always be trusted because estimating future values using this method is often not a well-conditioned problem. Our main goal is to propose a method with similar forecasting power that improves the reliability of the numerical results. We propose a filtering operation on the noisy observed data onto a principle subspace using principle component analysis (PCA), and we investigate the results based on mean square error of prediction, as a measure of performance. We use historical daily stock prices of General Electric from 1996 to 2015 to illustrate our method.

2. Literature Review

Technical traders base their analysis on the promise that the patterns in market prices are assumed to recur in the future, and thus these patterns can be used for predictive purposes(Gencay 1998). Studies show fundamental variables such as earnings yield, cash flow yield, size and book to market ratio (Chen, Leung and Daouk 2003; Fama and French

*Mrs. Mahsa Ghorbani, Department of Systems Engineering, Colorado State University, United States. Email: mahsa.ghorbani@colostate.edu

**Dr. Edwin K. P. Chong, Department of Electrical and Computer Engineering, Colorado State University, United States. Email: edwin.chong@colostate.edu

ISBN: 978-1-925488-41-8

1992), and macroeconomic factors such as interest rates, expected inflation and dividend have some power to predict stock returns (Fama and French 1986; Fama and French 1988).The literature proposes strong evidence that price can also be predicted from past price/return data as well as other fundamental or macroeconomic variables. Some studies found significant auto-correlation for daily and weekly returns (Lo and MacKinlay 1988; French and Roll 1986), others demonstrate correlation over the horizon of several months or years (Cutler, Poterba and Summers 1991; Fama and French 1986).

Principal component analysis (PCA) is a well-established mathematical procedure for dimensionality reduction and has wide applications across various fields such as time-series prediction (Hotelling 1933), pattern recognition, feature extraction, data compression, and visualization (Jolliffe 2002). In the field of quantitative finance, PCA has relevance in exploring financial time series (Incea and Trafalisb 2007), dynamic trading strategies (Fung and Hsieh 1997), financial risk computations (Alexander 2009;Fung and Hsieh 1997;Jain, Bakshi and Kalele 2015), and statistical arbitrage (Shukla and Trzcinka 1990). In this work, we employ PCA in forecasting stock prices.

Subspace filtering methods are based on the orthogonal decomposition of the noisy data space onto a signal subspace and a noise subspace. This decomposition is possible under the assumption of a low-rank model for the data, and on the availability of an estimate of the noise correlation matrix (Hermus, Wambacq and van Hamme 2007). This task can be done based on a modified singular value decomposition (SVD) of data matrices (Tufts, Kumaresan and Kirsteins 1982). The orthogonal decomposition into frame-dependent signal and noise subspace can be performed by an SVD of the noisy signal observation matrix or equivalently by an eigenvalue decomposition of the noisy signal correlation matrix (Hermus, Wambacq and van Hamme 2007).

Mean square error (MSE) is considered an appropriate metric to measure the performance of predictive tools (Gencay 1998; Fama and French 1988), which is the average squared difference between the actual and predicted price value. Given historical data, assuming normality, one efficient way to estimate future prices, is by simply using the multivariate conditional mean as the point estimator, because it minimizes the mean square error of the estimate (Scharf 1991). The numerical problem that occurs in most cases, is that the proposed estimator requires calculating the inverse of the covariance matrix of the observed data. Because this matrix is often ill conditioned, meaning its condition number is very large, calculation of its inverse can lead to significant numerical errors. We propose and implement a dimensionality reduction method to resolve this issue.

In this paper we focus on forecasting stock price from daily historical price data. Our method shows great potential for investigating the relevance of other fundamental and macroeconomic variables in price predictability as well.

3. The Methodology

3.1. General setting

Suppose that we have K samples of vector data, each of length N, where $N < K$. Call these vectors $x_1, x_2, ..., x_K$, where each $x_i \in R^N (i = 1, ..., K)$ is a vector of length N,

$$x_i = [x_{i1} \quad x_{i2} \quad \cdots \quad x_{iN}]. \qquad [1]$$

We assume that the vectors $x_1, x_2, \dots, x_K$ are drawn from the same underlying distribution. Let $M \le N$ and suppose the first M data points of vector x_i represent the end-of-day prices of a company stock over the past M consecutive trading days. The overall goal is to predict the next $M+1$ to N data points, which is company stock prices over the next $N-M$ trading days using the observed values of the past consecutive M days.

3.2. Normalizing and Centering the Data

In the case of stock-price data, the vectors $x_1, x_2, \dots, x_K$, might come from prices spanning several years or more. If so, the basic assumption that they are drawn from the same distribution may not hold because stock prices have changed over time. Even the value of a US dollar has changed over time, as a result of inflation. To overcome this issue, a scaling approach should be used to meaningfully normalize the prices. One such approach is presented here. Suppose that $t_i = [t_i(1), t_i(2), \dots, t_i(N)]'$ is a vector of stock prices centered as described above, over N consecutive trading days. Suppose that $Q \le N$ is also given. Then we apply the following normalization to obtain x_i:

$$x_i = \frac{t_i}{t_i(Q)}. \qquad [2]$$

This normalization has the interpretation that the x_i vector contains stock prices as a fraction of the value on the Qth day, and is meaningful if we believe that the pattern of such fractions over the days $1, \dots, N$ are drawn from the same distribution. Note that $x_i(Q) = 1$.

For the purpose of applying our method based on PCA, we assume that the vectors $x_1, x_2, \dots, x_K$ are drawn from the same underlying distribution and that the mean, $\bar{x}$, is equal to zero. However because x_i represents price values, in general the mean is not zero. The mean $\bar{x}$ can be estimated by averaging $x_i \in R^N (i = 1, ..., K)$,

$$\bar{x} = \frac{1}{K}\sum_{i=1}^{K} x_i, \qquad [3]$$

and then this average vector is deducted from each x_i to center the data. For convenience, in the following, we use the notation x_i for normalized and centered data.

3.3. Estimation Techniques

As mentioned before, we have K samples of vector data, each of length N, where $N < K$: $x_1, x_2, ..., x_K$, where each $x_i \in R^N (i = 1, ..., K)$ is a vector of length N. We can stack these vectors together as rows of a $K \times N$ matrix:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ x_{K1} & x_{K2} & \cdots & x_{KN} \end{bmatrix}.$$

ISBN: 978-1-925488-41-8

Let $M \le N$ and suppose that we are given a vector $y \in R^M$ representing the first M data points of a vector we believe is drawn from the same distribution as $x_1, x_2, \dots, x_K$. Again, these M data points represent the end-of-day prices of a company stock over the past M consecutive trading days. Let z be the price of the next $N-M$ days in the future. We wish to estimate z from y.

A. Gauss-Bayes or Conditional Estimation of z given y

The vector x_i is a multivariate random vector and can be partitioned in the form $x_i = [y_i \quad z_i]$, where y_i has length M and z_i has length $N-M$.
Accordingly the data matrix X can be divided into two sub-matrices Y and Z as follow:

$$[X] = [Y \quad Z]. \qquad [4]$$

We can think of Y as a data matrix consisting of samples of historical data, and Z as a data matrix consisting of the true future values of prices. The covariance matrix for the data can be written as

$$\Sigma_{xx} = \begin{bmatrix} \Sigma_{yy} & \Sigma_{yz} \\ \Sigma_{zy} & \Sigma_{zz} \end{bmatrix}. \qquad [5]$$

Assuming that y and z are jointly normally distributed, knowing the prior distribution of $x = [y, z]$, the Bayesian posterior distribution of z given y is given by

$$\begin{aligned} \hat{z}_{z|y} &= \Sigma_{zy}\Sigma_{yy}^{-1} y \\ \hat{\Sigma}_{z|y} &= \Sigma_{zz} - \Sigma_{zy}\Sigma_{yy}^{-1}\Sigma_{yz}, \end{aligned} \qquad [6]$$

The $\hat{\Sigma}_{z|y}$ matrix is also called the Schur complement of Σ_{yy} in Σ_{xx}. Note that the posterior variance does not depend on the specific value of y. The Gauss-Bayes estimator, the conditional mean, minimizes the mean square error (Scharf 1991). The same set of equations can be obtainedin development of Kalman filtering. Kalman's own view of this approach is a complete deterministic operation (Byrnes, Lindquist and Zhou 1994), and not computing the Gaussian posterior distribution, which means the normality assumption is not mandatory. Although the point estimator $\hat{z}_{z|y}$ is optimal in term of MSE, in practice there are numerical complications involved in this method: The matrix of Σ_{yy} is not well conditioned, so the numerical calculation of Σ_{yy}^{-1} cannot always be trusted. To overcome this problem, we propose a better conditioned estimator, which has a behavior close to Guess-Bayes.

B. Principal Components and Estimation in Lower Dimension

Principal component analysis (PCA) is a well-established mathematical procedure for dimensionality reduction of the data and has wide applications across various fields. In this work, we consider its application in forecasting stock prices. PCA involves calculating the eigenvalue decomposition of the covariance matrix of the data or equivalently singular value decomposition (SVD) of the matrix of data. Consider the SVD of X:

$$X = USV', \qquad [7]$$

where S is a diagonal matrix of the same dimension as X with non-negative diagonal elements in decreasing order, and U and V are unitary matrices ($UU^T = I_K$ and $VV^T = I_N$). The square of diagonal elements of S are called eigenvalues.Equivalently, the SVD of the covariance matrix of data can be considered. A little bit of algebra shows that

$$\Sigma_{xx} = \frac{1}{K-1} X'X \quad = \frac{1}{K-1}(USV')'(USV) = V\frac{S^2}{K-1}V'. \qquad [8]$$

Hence the matrix of eigenvectors V for the covariance matrix are the same as the singular vectors from SVD for the data matrix, and the eigenvalues generated in this case are just the squares of the singular values from the first SVD. From now on we focus on SVD of covariance matrix. In general, the first few eigenvalues account for the bulk of the sum of all singular values. The eigenvectors with the greatest eigenvalues are called the principal components.

Let $L < N$ be such that the first L singular values in S account for the bulk part (say 85% or more) of the sum of the singular values. Let V_L be the first L column of unitary matrix V in the SVD of X. Let $A = S'U'$, so the ith column of X is $x_i = Va_i$, and let vector $\alpha_i \in R^L$ be the first L components of a_i. Then each x_i is approximately equal to the linear combination of the first L columns of V:

$$x_i \approx V_L \alpha_i. \qquad [9]$$

Because L is a small number compared to N, equation (9) suggests that a less noisy subspace with a lower dimension can represent most of the information in X. Projecting onto this principle subspace can resolve the ill-conditioned problem of Σ_{yy}. The idea is that instead of using all eigenvalues, which vary greatly in magnitude, we use a subset which only includes the big ones, and therefore has a smaller range of eigenvalues. The same concept is implemented in signal subspace filtering methods, which are based on the orthogonal decomposition of noisy speech observation space onto a signal subspace and a noise subspace (Hermus, Wambacq and van Hamme 2007). Let $V_{M,L}$ be the first M rows and first L columns of V. We have

$$y = V_{M,L}\alpha + \text{Noise}. \qquad [10]$$

Mathematically resolving noisy observation vector y onto the principle subspace can be written as a filtering operation in the form of

$$w = Gy, \qquad [11]$$

where G is given by

$$G = (V'_{M,L}V_{M,L})^{-1}V'_{M,L} \qquad [12]$$

The vector w is actually calculating the coordinates of the orthogonal projection. Substituting y by w in (6) leads to a better conditioned set of equations,

$$\begin{aligned} \hat{z}_{z|w} &= \Sigma_{zw}\Sigma_{ww}^{-1}w \\ \hat{\Sigma}_{z|w} &= \Sigma_{zz} - \Sigma_{zw}\Sigma_{ww}^{-1}\Sigma_{wz}, \end{aligned} \qquad [13]$$

because the condition number of Σ_{ww} is much lower than that of Σ_{yy} as we will see later.

In (13) we have:

$$\Sigma_{zw} = E\left[zw'\right] = \Sigma_{zy}G', \qquad [14]$$

and

$$\Sigma_{ww} = E\left[ww'\right] = G\Sigma_{yy}G' \qquad [15]$$

ISBN: 978-1-925488-41-8

If the posterior distribution of z estimated based on (13) has a similar behavior to the distribution estimated by (6), it can be considered a good substitute for the Gauss-Bayes method. Our simulation demonstrates promising results, which will be presented in the following sections.

C. Unconditional Estimation of Normalized Data

Recall that $z_i \in R^{N-M}$ represents the last $N-M$ points of the x_i vector, which are the historical sample stock price values. The sample mean and covariance matrix for sub-matrix z are unbiased estimators of the mean and covariance of the random vector z_i, and therefore can be used as an estimate for future price values, albeit one that ignores y. We have:

$$\hat{z}_{\text{Uncon}} = \frac{1}{K}\sum_{i=1}^{K} z_i$$
$$\hat{\Sigma}_{\text{Uncon}} = \frac{1}{K-1}(Z^{'}Z), \qquad [16]$$

where $\hat{z}_{\text{Uncon}}$ is the sample mean and $\hat{\Sigma}_{\text{Uncon}}$ is the sample covariance of z .It is important to notice that these statistics are different from sample mean and covariance of the original price data because we are applying the formulas above to the normalized and centered data.

We run the simulation for the dimension-reduction method to estimate the stock prices over the next $N-M$ days and the results are compared to Gauss-Bayes method. The unconditional estimation results are also included for comparison.

3.4. Performance Metric

To compare the performance of these methods we evaluate the expected value of squared error between the real and estimated values. The mean squared error over all different observations is formulated as

$$MSE = E\left[\left\|z-\hat{z}\right\|^2\right] = E\left[\left\|z^2\right\|\right] + E\left[\left\|\hat{z}^2\right\|\right] - 2E\left[\left\|z'\hat{z}\right\|\right] \qquad [17]$$

In each case, the MSE formula is updated by substituting $\hat{z}$ by the corresponding point estimator. The Gauss-Bayes estimator is unbiased and in fact the covariance matrix in this method provides the lower limit on the MSE for all unbiased estimators.

4. Experiments

The daily historical price data from 1996 to 2015 for General Electric Company was downloaded from finance.yahoo.com. This data set is transformed into a Hankel matrix and then centered and normalized to construct the data matrices, as described earlier. In this paper we focus on short-term prediction, meaning just a few days. We compare the estimation values from each technique in terms of MSE.

ISBN: 978-1-925488-41-8

4.1. Constructing Data Matrix

Our data is transformed into a Hankel matrix with K rows, samples of vector data, each of length N. We get that by stocking K rows (K samples), each one time shifted from the previous one, all in one big matrix, called the Hankel matrix.
More precisely, the Hankel matrix H_x for this problem is constructed from the vector, x_i, formed in the following format:

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_K \end{bmatrix} = \begin{bmatrix} P(1) & P(2) & \cdots & P(N) \\ P(2) & P(3) & \cdots & P(N+1) \\ \cdots & \cdots & \cdots & \cdots \\ P(K) & P(K+1) & \cdots & P(K+N-1) \end{bmatrix}.$$

where $P(N)$ represents the price for day N. This is our matrix of data, X. Then the sample covariance matrix is calculated as $\Sigma_{xx} = H_x^T H_x / (N-1)$.

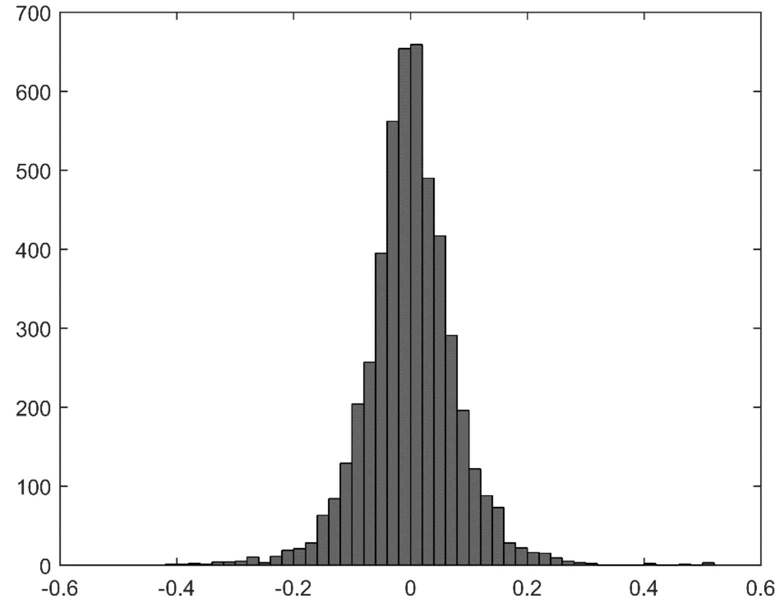
End-of-day stock prices for General Electric Company for about 5000 consecutive days are converted into Hankel matrices with different lengths. We vary M from 20 to 740, with a 30 day interval, to investigate the effect of length of observation vector on the results, which means 25 sets of data are evaluated in this study.

As explained earlier, we first normalize each row by $y(Q)$ and then subtract the average vector $\overline{x}$ from each row. After running the simulation, to make use of the predicted values, we should add back the average vector $\overline{x}_{N-M}$ (last $N-M$ components of $\overline{x}$) from days $M+1$ through N and also multiply the result by $y(Q)$ to get back to actual stock prices. Different values for Q have been tested in terms of MSE. For the purpose of this study, $Q = M$ has been chosen because we believe it shows the best results in this setting. Note that $x_i(Q) = 1$. This column is removed from the data matrix because it does not provide any information. From now on matrix X represents normalized price data. In each case, the average of the data set is deducted from each observation to center the data.

The histogram of normalized data is graphed as a representation of the distribution of data. Figure 1 represent the first predictor (first column) in matrix X, and the plot resembles a bell shape.

ISBN: 978-1-925488-41-8

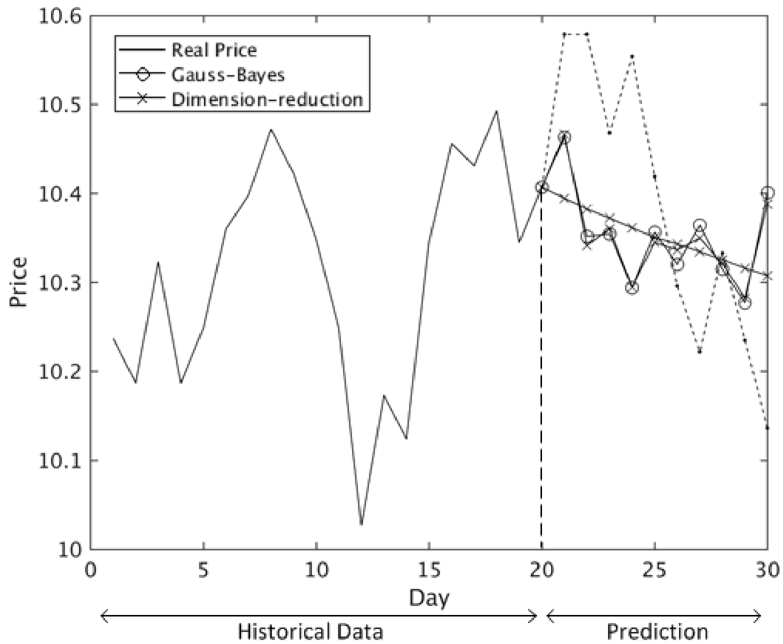
Figure 1: Histogram of historical data in normalized domain



4.2. Simulation

For each of the data sets constructed above, we implement three different techniques of estimation for the next 10 days (day $M+1$ to N). Figure 2 shows an example of our stock prediction. Assume we are given the price values for the past 20 days, and we want to use those values to predict the future prices over the next 10 business days, from day $M+1$ to day N. In our dimension-reduction technique, we can get a very smooth plot for a relatively small L, to a plot almost the same as Gauss-Bayes, for bigger values of L.

Figure 2: Predicting price for the next 10 days using historical data, M=20, N=30, L=2 andL=11.



The general goal, as mentioned above, is an estimation technique that has a similar behavior as Gauss-Bayes but does not have the associated calculation difficulties resulting from ill-conditioning. As mentioned before, mean squared error (MSE) is a common and appropriate measure of performance. We implement our dimension-reduction technique for different M s, and for different numbers of principal eigenvalues, L.

Figure 3 shows the values of MSE over all days of estimation versus value of Lin the normalized domain, for 25 different lengths of observation vector$M$, from 20 to 740. It turns out that MSE value is not that sensitive to the value of L for sufficiently large$L$. As we can see, initially, the MSE values fall quickly for small values of$L$, but then remain relatively constant, so if we have a particular constraint on condition number, we do not lose that much in terms of MSE by choosing a lower dimension subspace, which leads to a better conditioned problem.

We are interested in the sum of MSE values over all days of estimation as shown in Figure 4. The figure illustrates the sum of MSE over all days, subject to an upper limit on condition number ofΣ_{ww}. There is a trade-off between$M$, length of observation vector, and value of MSE. In general by increasing$M$, more information is available in each observation, resulting in better performance of the estimation. For each length of$M$, the values for MSE are investigated based on condition number ofΣ_{ww}. The top plot in Figure 4 corresponds to the MSE values corresponding to the Unconditional method. The last plot on the bottom corresponds to Gauess-Bayes MSE values, which indicates the optimal performance. The other 4 plots correspond to our dimension-reduction method, subject to 4 different upper limits on Σ_{ww} condition number, from 10^3 to10^6.

ISBN: 978-1-925488-41-8

Figure 3: MSE versus L in the normalized domain for different lengths of observation vector

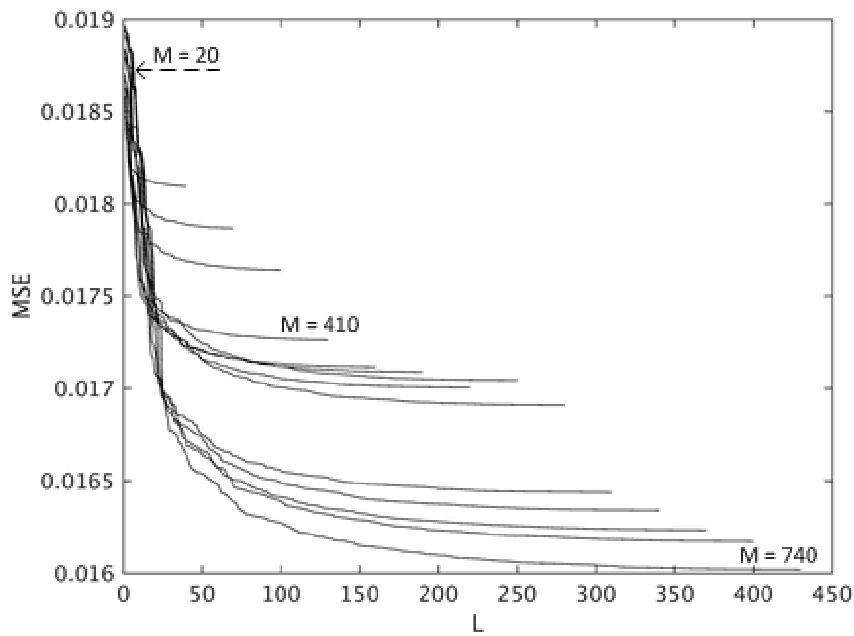
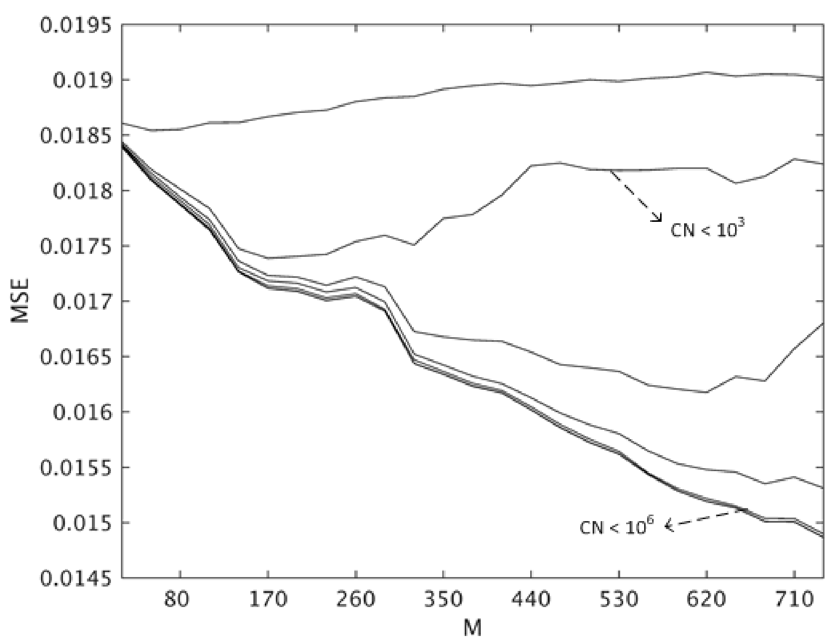


Figure 4: Best MSE subject to different upper limits on condition number of Σ_{ww}



The performance of the lower-dimension estimation method is close to Gauss-Bayes in terms of MSE up to a certain point for all different limits of condition number. After $M = 200$, or in other words after roughly 7 months, the values from lower rank methods

ISBN: 978-1-925488-41-8

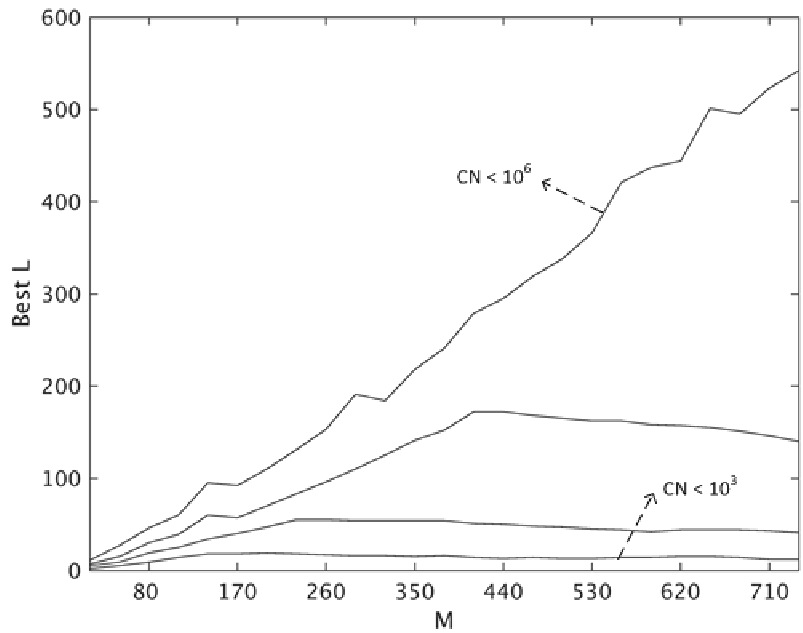
start deviating from Gauss-Bayes in some cases. The second line on top is the best performance of lower dimension method subject to condition number of Σ_{ww} being less than 10^3 which is about 1000 times better than the condition number of Σ_{yy}. However, in this it case seems like that after $M = 360$, the performance of lower rank estimator starts deteriorating. At this point, theerror in the proposed algorithm starts diminishing the predictor power. When we fix some constraint on condition number, we are actually limiting the value of L, and by increasing M, after a certain point, we mostly increase the noise, and the MSE value gets worse. As a result if we really do have a constraint on condition number, we need to pick a value of M that is appropriate.

Moving on to the next plot, which is corresponding to the 10^4 limit, the MSE value is almost decreasing for all different values of M. In this case, where the condition number is about 100 times better than Gauess-Bayes, the MSE values are very close in both techniques. As we can see, in trading off between MSE for condition number, there is little detriment in MSE values.

Investigating the dimension of the target subspace provides a better understanding of how the method works. Remember L represents the number of eigenvalues required from diagonal matrix S to represent the bulk part of the information. Value of L corresponding to best MSE for different M s, subject to different limits on condition number, is plotted in Figure 5.As the upper limit on condition number increases, the value of MSE improves as M increases, and we need a bigger subspace, bigger L, to extract the information. However, as you can see in the first three plots in Figure 5, the value for best L is almost constant after a certain point, which is consistent with Figure 3.Since L is in fact the dimension of our principal subspace, you can see how using the lower dimension algorithms decreases the dimensions of calculations, which leads to a better conditioned problem.

ISBN: 978-1-925488-41-8

Figure 5: Best L subject to different upper limits on condition number of Σ_{ww}



5. Conclusion

In this paper, we described a general method for prediction using covariance information. We illustrated our method on daily stock price values for General Electric Company. The daily historical price data from 1996 to 2015 was transformed into Hankel matrices of 25 different lengths to investigate the impact of length of observation on estimation power. Each data matrix is then normalized for running the simulation. The multivariate conditional mean is known to minimize the mean square error and therefore is used as a suitable unbiasedestimator of future values; however, the numerical results from this method cannot be trusted because the resulting covariance matrix is not well conditioned. We proposeda filtering operation using principle component analysis to overcome this issue. Resolving theobserved data set onto a principle subspace reduces the dimensionality of the problem and in this case study, improves the condition number of the problem by orders of magnitude. The proposed method shows similar behavior to the multivariate conditional mean in terms of mean square error of the estimation (provided the imposed constraint on condition number is not excessively stringent) in a specific range for length of observation and therefore is considered a good substitute for that method. The proposed method is easily implemented and can be modified to include multiple predictors, including macroeconomic factors such as interest rates, expected inflation, or fundamental variables such as earnings yield, cash flow yield, and market capitalization. The significance of the proposed approach will be even more appreciated in estimating the future price values using multiple predictors because in that case, where observation vectors are mostly longer than the one in this study, it becomes almost impossible to rely on Gauss-Bayes due to the severe ill-conditioning of the covariance matrix.

ISBN: 978-1-925488-41-8

References

Alexander, C 2009, *Market Risk Analysis, Value at Risk Models*. Volume IV ed. Wiley, New Jersey.

Byrnes, C. I., Lindquist, A and Zhou, Y. 1994, On the nonlinear dynamics of fast filtering algorithms, *SIAM J. Control Optim.*, Vol. 32, No. 3, pp. 744-789.

Chen, A., Leung, M.T. and Daouk, H. 2003, Application of neural networks to an emerging market: forecasting and trading the Taiwan Stock Index, *Computers and Operation Research*, Vol.30, pp.901–923.

Cutler, D.M., Poterba, J.M. and Summers, L.H.1991, Speculative dynamics, *Review of Economic Studies*, Vol.58, pp.529–546.

Fama, E.F., French, K.R., 1986, Permanent and temporary components of stock prices, *Journal of Political Economy*, Vol.98, pp.246–274.

Fama, E. and French, K. 1992, The cross-section of expected stock returns, *Journal of Finance*, Vol.47, pp.427–465.

Fama, E, French, K. 1988, Dividend yields and expected stock returns, *Journal of Financial Economics*, Vol.22, pp.3–25.

French, K.R. and Roll, R. 1986, Stock return variances: The arrival of information and the reaction of traders, *Journal of Financial Economics*, Vol.17, pp.5–26.

Fung, W. and Hsieh, D. A. 1997, Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds, *Review of Financial Studies*, Vol. 10, pp.275–302.

Gencay, R 1998, The predictibility of security returns with simple technical trading rules. *Journal of Empirical Finance*, Vol.5, pp.347–359.

Hermus, K, Wambacq, K and van Hamme H. 2007, A Review of Signal Subspace Speech Enhancement and Its Application to Noise Robust Speech Recognition, *EURASIP Journal on Advances in Signal Processing*, Volume 2007, pp.1-15.

Hotelling, H 1933, Analysis of a complex of statistical variables into principal components, *Journal of Educational Psychology*, Vol. 24, pp.417-441.

Incea, H and Trafalisb, T. B. 2007, Kernel Principal Component Analysis and Support Vector Machines for Stock Price Prediction, IIE Transactions, Vol. 39, pp. 629–637.

Jain, A, Bakshi, M and Kalele 2015, A Easwar Subramanian: On Accelerating Concurrent PCA Computations for Financial Risk Applications. *Proceeding of IEEE 22nd HiPC*, pp. 175-184. Available from: IEEE Xplore digital library. [20 April 2017].

Jolliffe, I. T. 2002, *Principal Component Analysis*, 2nd ed, Springer, New York.

Lo, A.W., MacKinlay, A.C. 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies*, Vol. 1, pp.41–66.

Shukla, R and Trzcinka, C 1990, Sequential Tests of the Arbitrage Pricing Theory: A Comparison of Principal Components and Maximum Likelihood Factors, *Journal of Finance*, Vol. 45, pp. 1541–1564.

Scharf, L.L. 1991, Statistical Signal Processing, Detection, Estimation, and Time Series Analysis. Pearson, New York.

Tufts, D. W., Kumaresan, R. and Kirsteins, I. 1982, Data adaptive signal estimation by singular value decomposition of a data matrix, *Proceedings of the IEEE*, Vol. 70, No. 6, pp. 684–685.