ISSN 1011-8888

INSTITUTE OF ECONOMIC STUDIES

WORKING PAPER SERIES

W00:18 December 2000

Estimation of Correlations in Financial Markets when Trading is Infrequent

Helgi Tomasson

Address: Department of Economics
University of Iceland
Oddi at Sturlugata
101 Reykjavik
Iceland

Email: helgito@hi.is

Estimation of correlations in financial markets when trading is infrequent

Helgi Tómasson

University of Iceland

Faculty of Economics and Business Administration

and Institute of Economic Studies

Oddi v/Sturlugötu

IS-101 Reykjavík

ICELAND

helgito@rhi.hi.is

Institute of Economic Studies at the University of Iceland

Working paper: W00:18, 2000

Abstract

Observed transaction prices on a financial market are assumed to be noisy discrete measurements on a multivariate Wiener-process. The likelihood function is calculated by means of Kalman filtering. A three-stage procedure of estimating the covariance matrix is proposed. The method is applied to Icelandic stock market data by numerically maximizing the likelihood. The application shows that the method is feasible and that the information about many relations on the Iceland Stock Exchange is sparse.

Keywords: Continuous time, discrete observations, financial market, state-space models.

1 Introduction

In a financial market where many assets are traded, a measure of the relation between movements of assets is of interest. Some risk management systems, e.g., CAPM(Sharpe, 1964; Lintner, 1965; Mossin, 1966), are based on correlations between movements of assets. When stocks trade non-synchronously, estimation from observed data poses a problem. In this paper a direct procedure for estimation of the correlation coefficient from nonsynchronous data is given.

When a new stock market opens it is typically characterized by infrequent trading and low volumes. Traditional ways of reporting closing day prices may give misleading signals about the market, undermining agents' confidence. Reporting closing day prices of various assets is actually quoting information of varying quality, as some assets are traded more frequently than others. When reporting a closing day price of an asset that was not traded that day, or perhaps has not been traded at all for weeks, the information about that asset, is obviously of inferior quality, compared with information on one which was traded just before closing. Using time series methods on multivariate data requires simultaneous observations of all coordinates of the data vector at each sampling point. Relaxing that requirement by guessing the missing observations demands special precautions. Harvey (1993) discusses methods for dealing with missing observations. In the traditional time series literature, missing observations are typically treated like some kind of exception. In the environment of financial markets, missing observations for some assets, is perhaps rather the rule rather than the exception. Analyzing the movement of a partially observed multivariate vector, Bassett et al. (1991) approach the problem by treating the dimension of the observed vector as variable. The approach in the present paper is different. The setup is a very simple, continuous-time stochastic model. The simplicity of the model is justified by the intended application to stock market data, where the efficient market hypothesis implies that it should not be possible to predict future values. It is assumed that the observed data are discretely observed points of a multivariate diffusion process. The diffusion process describes the value of each asset at each point in time. At the time of a transaction, a part of the vector, usually only one coordinate, is observed.

Economic theory typically abstracts from the mechanics of trading, whereas the market-microstructure literature deals with explaining what impact the trading mechanism has on price formation. The purpose of market-microstructure theory is also to explain why prices exhibit particular time-series properties (O'Hara, 1995). It is well known that time-series analysis of real data from financial markets tends to show significant serial correlation in returns (Campbell et al., 1997; Cuthbertson, 1996). The existence of such serial correlation for stock-market data has troubled many analysts as it suggests that the returns of some assets are predictable, and thereby contradicting the efficient-market-hypothesis. Some authors consider these findings as a possible violation of the efficient market hypothesis. Campbell et al. (1997), page 129, discuss empirical findings in discrete time on the impact of nonsynchronous trading on serial correlations; individual security returns with non-zero means tend to induce, among other things, negative serial correlation, well-diversified portfolios tend to have positive serial correlations, etc. They also discuss the impact of daily non-trading probabilities on autocorrelations of observed returns.

The intended application for this work is stock market data, and the approach is to assume the efficient-market-hypothesis in continuous time and formalize it as a continuous martingale in log-prices. When data consist of non-synchronous observations of a multivariate Wiener-process, it can be shown that information on one asset can help predict the price of another, provided that they are correlated. The example described in Section 2 illustrates this in a simple manner. The transactions on a financial market take place in a bargaining environment. The nature of bid-ask data is quite complex; thus, to allow for the impact of bargaining and other microstructure phenomena, a noise component is added to the model. The interpretation of observed prices is that they are the sum of the true value and microstucture noise. The decomposition of transaction prices into signal and noise for univariate financial data is described in Tómasson (2000). That model is now extended to deal with the relations between many stocks. To be more specific, the variable under consideration here consists of a k-dimensional vector, k, being the number of assets traded on the market. The vector of values is defined at each point in time, whereas the transactions consist of discrete noisy observations of one or more coordinates of this vector. It is assumed that the value vector has a continuous path in each coordinate, therefore a multivariate diffusion-process is a natural candidate. This obviously means that when real discontinuities occur, e.g., when a dividend is paid, or a stock split takes place, some precautions are necessary. The quantitative measure of relation between the various assets is the covariance matrix of the innovations of the process. The model is written down explicitly in Section 3. The model is a multivariate one, whereas the observations are univariate in nature, i.e., at most time points only one coordinate of the vector is observed. Due to bargaining and the existence of many bids/asks at some time points, it is possible to observe two or more transactions of the same asset simultaneously, not necessarily at the same transaction price. It is assumed, though, that at each point in time, the value of each asset is unique. In Section 4, a statistical model for dealing with discrete observations is derived from the theoretical model in Section 3. Section 2 shows that when two assets, driven by a correlated bivariate Wiener-process, are traded non-synchronously, the expected value of an asset, A, conditional on past values of both assets, will be a function of the latest observation of asset A, and the two latest observations of asset B, if there are transactions with asset B later than the latest transaction of asset A. This conditional, expected value will also be a function of the four time points of transaction times for assets A and B, respectively. When extending these results to more than two assets, it is clear that the conditional expected value of an asset will become a function of every transaction price that has taken place since that asset was last traded, as well as the time points of the transactions. It is evident that the extension of the formulas of Section 2 to high dimensions will be quite messy, not to mention the impact of introducing the noise component. The model can easily be written in state-space form, as the interpretation is that the state equation represents the dynamics of the latent value, and the measurement equation is the observed transaction price. Using the state-space representation, the Kalman-filter recursions are a straightforward way of calculating the conditional expectations. Assuming normality, the Kalman-filter algorithm also gives a method of calculating the likelihood function. The parameters of the model can then be estimated by numerically maximizing the likelihood. The dimension of the system is a problem, so a stepwise procedure is proposed in Section 5. The steps consist of several univariate or low-dimensional optimization problems, constraining the estimates to form a valid covariance matrix, which is used as a starting value for an EM (expectation maximum) algorithm. A brief description of an application of the method to Icelandic stock market data is given in Section 6.

2 A two-dimensional example

Suppose that we have a pair of correlated processes:

$$\begin{bmatrix} dX_1(t) \\ dX_2(t) \end{bmatrix} = \begin{bmatrix} \sigma_1 dW_1(t) \\ \sigma_2 dW_2(t) \end{bmatrix} \tag{1}$$

$$E(dW_1(t)) = E(dW_2(t)) = 0 \quad , \quad E(dW_1^2(t)) = E(dW_2^2(t)) = dt \quad \text{and } E(dW_1(t)dW_2(t)) = \rho dt$$

Both $X_1(t)$ and $X_2(t)$ are standard univariate Wiener-processes, i.e., the best forecast value is the last observed one:

$$E(X_1(t)|X_1(s)) = X_1(s) \quad \text{and} \quad E(X_2(t)|X_2(s)) = X_2(s) \quad \text{for } t > s \tag{2}$$

However, information about a new value of, say, $X_1(t)$, will affect the conditional expected value of $X_2(t)$:

$$E(X_2(t)|X_1(t), X_1(s), X_2(s)) = X_2(s) + \rho\sigma_2/\sigma_1(X_1(t) - X_1(s)) \quad \text{for } t > s \tag{3}$$

From equation (3) it is clear that when the processes are correlated, the arrival of new information in one of the processes will affect the conditional expectation. This fact is illustrated in Figures 1 to 3. In Figure 1 two correlated random-walk processes are plotted, and the optimal forecast $\hat{X}_i(t) = X_i(30)$ for $t > 30$ for $i = 1$ and $i = 2$. In Figure 2 a new observation, $A = X_2(35)$, on one of the processes is obtained, and in Figure 3 it is shown how this new information will affect the conditional expectation, $B = E(X_1(35)|X_2(35), X_1(30), X_2(30))$. It is clear that the best forecast for $X_1(t)$ is affected by this new information.

A four-dimensional vector consisting of two correlated Wiener-processes, X_1 measured at time-points t_1 and t_2, and X_2 measured at time-points s_1 and s_2, is normally distributed with

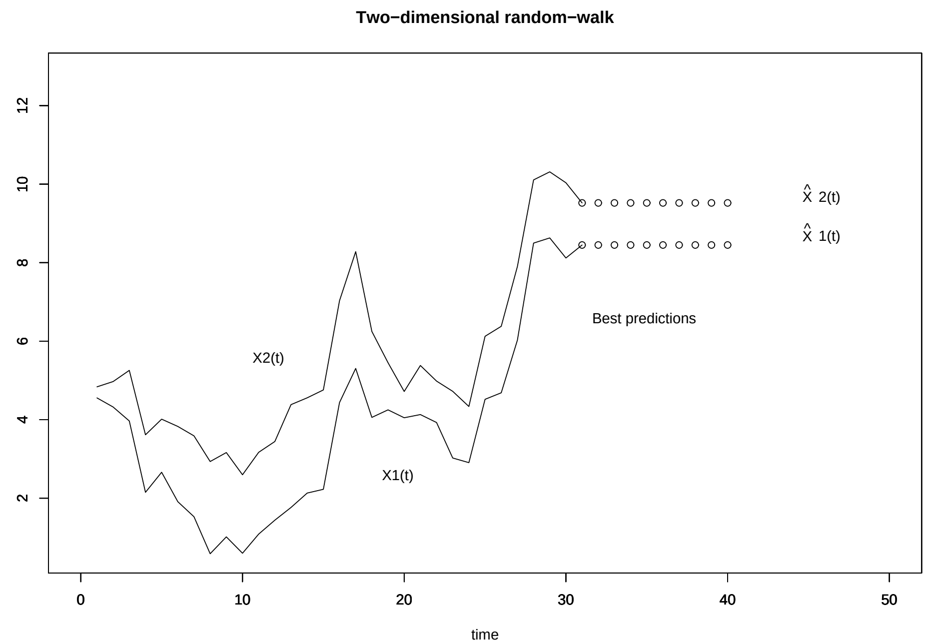

Figure 1: Two correlated random walks.

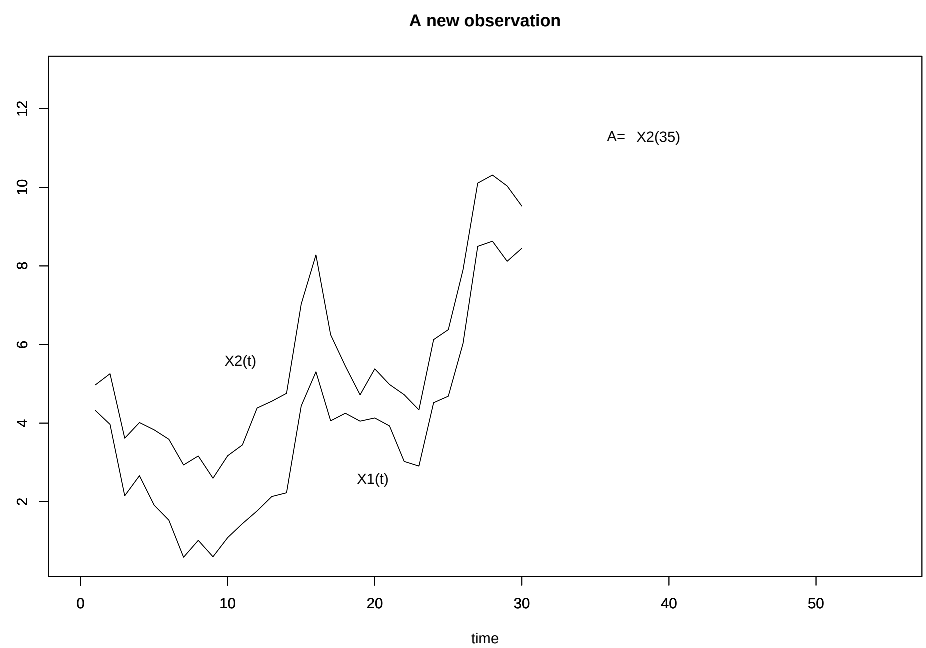

Figure 2: Two correlated random walks with a new observation added.

zero mean and a covariance matrix given by:

$$\begin{bmatrix} \sigma_1^2 t_1 & \sigma_1^2 min(t_1, t_2) & \rho\sigma_1\sigma_2 min(t_1, s_1) & \rho\sigma_1\sigma_2 min(t_1, s_2) \\ \sigma_1^2 min(t_1, t_2) & \sigma_1^2 t_2 & \rho\sigma_1\sigma_2 min(t_2, s_1) & \rho\sigma_1\sigma_2 min(t_2, s_2) \\ \rho\sigma_1\sigma_2 min(t_1, s_1) & \rho\sigma_1\sigma_2 min(t_2, s_1) & \sigma_2^2 s_1 & \sigma_2^2 min(s_1, s_2) \\ \rho\sigma_1\sigma_2 min(t_1, s_2) & \rho\sigma_1\sigma_2 min(t_2, s_2) & \sigma_2^2 min(s_1, s_2) & \sigma_2^2 s_2 \end{bmatrix} =$$

$$\begin{bmatrix} \sigma_1^2 t_1 & \Sigma_{AB} \\ \Sigma'_{AB} & \Sigma_{BB} \end{bmatrix} \tag{4}$$

By using equation (4), equation (3) can be generalized by use of standard normal theory, which gives the expected value of $X_1(t_1)$, conditional on $(X_1(t_2), X_2(s_1), X_2(s_2))$ as:

$$E(X_1(t_1)|X_1(t_2), X_2(s_1), X_2(s_2)) = \Sigma_{AB}\Sigma_{BB}^{-1}(X_1(t_2), X_2(s_1), X_2(s_2))' \tag{5}$$

If $t_1 > t_2 > s_1 > s_2$ the observation intervals are disjoint, then $E(X_1(t_1)|X_1(t_2), X_2(s_1), X_2(s_2)) = X_1(t_2)$, i.e., disjoint observation intervals do not yield any information about the relation between the two processes. Equation (3) can easily be generalized to higher dimensions, but it

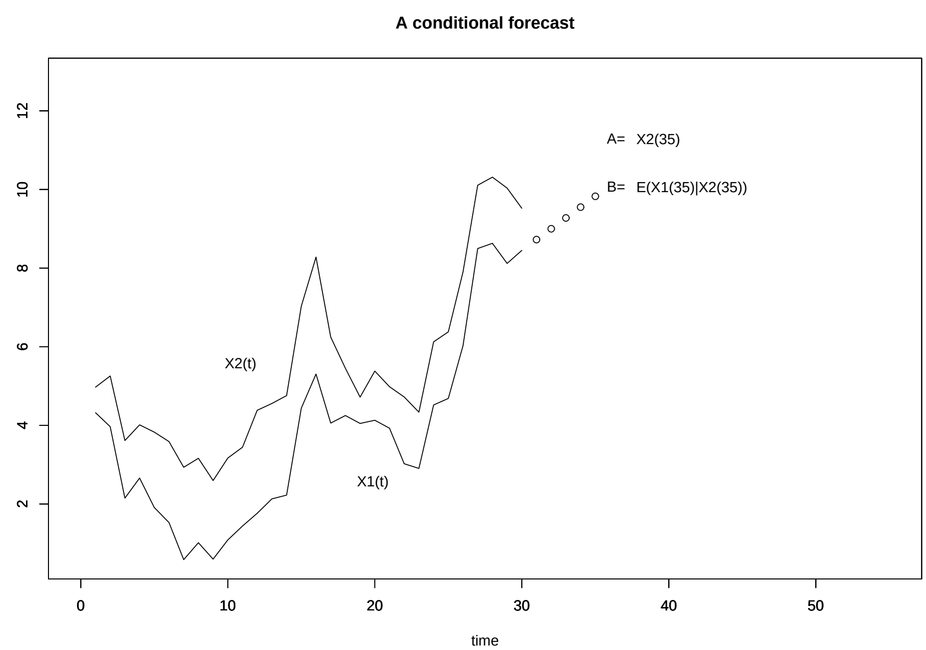

Figure 3: Two correlated random walks and a conditional forecast.

is evident that the conditional distribution of $X_1(t_1)$ will be a function of everything that has happened since the last observation of X_1. Therefore, the corresponding formula will be quite messy.

3 A simple continuous-time model of dependent assets

The example in Section 2 shows that in the environment of non-synchronous trading, it is very natural, even under the efficient market hypothesis, to observe a significant serial-correlation structure of asset returns. In this paper a slight modification of the random-walk model presented in Section 2 will be introduced. The approach is similar to Tómasson (2000), i.e., that the true prices are assumed to follow a continuous time diffusion process with added noise. The observed transaction price is the sum of the true price and a noise component. The purpose of the noise component is to capture microstructure phenomena. The true price of asset j, at time t, is denoted with $\alpha_j(t)$, and the k-dimensional vector $\boldsymbol{\alpha}(t)' = (\alpha_1(t), \dots, \alpha_k(t))$, is defined accordingly. The dynamics of $\boldsymbol{\alpha}(t)$ are assumed to follow a correlated Wiener-process:

$$d\boldsymbol{\alpha}(t) = d\tilde{\boldsymbol{W}}(t) \quad \text{with} \quad E(d\tilde{\boldsymbol{W}}(t)d\tilde{\boldsymbol{W}}(t)') = \boldsymbol{Q} \tag{6}$$

The observation vector $\boldsymbol{y}(t)$ consists of an observation of $\boldsymbol{\alpha}(t) + \boldsymbol{\varepsilon}(t)$. The vector $\boldsymbol{\varepsilon}(t) = (\varepsilon_1(t), \dots, \varepsilon_k(t))$ consists of the microstructure noise for each asset. It is conceivable to observe simultaneous observations, i.e. many realizations of $\boldsymbol{\varepsilon}(t)$. In a previous paper by Tómasson (2000) several structures of drift term are tested on empirical stock market data. The results indicated that the volatility term was dominating, and that the choice of drift structure affected the volatility estimates only marginally, which perhaps is a natural result of the efficient market hypothesis. Therefore, for simplicity in this paper no drift term is put into equation (6). It is also conceivable to allow for some structure of the noise term, $\boldsymbol{\varepsilon}(t)$, but for simplicity, the noise is assume to be i.i.d. within each asset. The noise components are assumed to be independent, zero-mean processes for all assets, but each asset is allowed to have its own asset-specific variance of the noise. The vector of standard-deviations of $\boldsymbol{\varepsilon}(t)$ is denoted with $\boldsymbol{\sigma_\varepsilon} = (\sigma_{\varepsilon,1}, \dots, \sigma_{\varepsilon,k})'$. The matrix $\boldsymbol{Q}$ consists of elements, σ_{ij} describing the simultaneous comovements of assets. The square root of the diagonal, σ_{jj}, is the daily volatility of stock j.

4 Application to discrete data

In real markets transactions take place in a bargaining environment, which results in observed transaction prices being different from the true price. Transactions take place at time points, $t_1, t_2, \ldots$, the possibility of simultaneous transactions of the same asset at different transaction prices is allowed, as well as non-synchronous trading of assets in general. In fact, only a part of the k-dimensional vector, $\boldsymbol{y}(t)$, of log-transaction prices, is observed at each time point, and, as it is possible to obtain many different observations of a coordinate of $\boldsymbol{y}(t)$, the vector $\boldsymbol{y}(t)$ should be interpreted as a sample. The observations are therefore in practice univariate, or at most bivariate. At time, t_i, we observe the price, $y_j(t_i)$, and the number, $j = j(t_i)$, of the asset which is traded at time t_i. To be more specific, the observed price at time t_i is:

$$y_j(t_i) = \alpha_j(t_i) + \varepsilon_j(t_i) \tag{7}$$

where $y_j(t_i)$ denotes the observed transaction price of asset j, $\alpha_j(t_i)$ denotes its true price, and $\varepsilon_j(t_i)$ denotes a market microstructure noise component due to phenomena like bargaining. The vector of the true asset prices at time, t, is denoted by $\boldsymbol{\alpha}(t)$ of equation (6). The true price is unobservable, and there an evident time dependency is described by the stochastic differential equation (6). Solving the stochastic differential equation at time point t_i conditional on the value at time point t_{i-1} gives:

$$\boldsymbol{\alpha}(t_i) = \boldsymbol{\alpha}(t_{i-1}) + \boldsymbol{\xi}(t_i) \tag{8}$$

with

$$\boldsymbol{\xi}(t_i) = \tilde{\boldsymbol{W}}(t_i) - \tilde{\boldsymbol{W}}(t_{i-1})$$

It is therefore natural to interpret equation (7) and write a measurement equation as (9) into a state space model:

$$y(t_i) = \boldsymbol{z}(t_i)\boldsymbol{\alpha}(t_i) + \varepsilon_j(t_i) \tag{9}$$

Here $y(t_i) = y_j(t_i)$ is a transaction in asset $j = j(t_i)$ taking place at time point t_i; the vector $\boldsymbol{z}(t_i)$ is a k-dimensional vector of zeroes, except $z_j(t_i) = 1$. As information, i.e., data, $y(t_1), y(t_2), \ldots$, arrives, the Kalman-filter recursive algorithm gives an objective estimate of the true price for all assets at each point in time. If $\boldsymbol{\alpha}(t|t_i)$ denotes the expected value of $\boldsymbol{\alpha}(t)$, conditional on information available at time t_i, and $\boldsymbol{P}(t|t_i)$ as the covariance-matrix for this estimate, then

$$\boldsymbol{\alpha}(t|t_i) = \boldsymbol{\alpha}(t_i|t_i) \tag{10}$$

$$\boldsymbol{P}(t|t_i) = \boldsymbol{P}(t_i|t_i) + (t - t_i)\boldsymbol{Q} \tag{11}$$

Given that transaction with asset j takes place at time t_i, the predicted transaction price is:

$$\hat{y}(t_i) = \boldsymbol{z(t_i)}\boldsymbol{\alpha}(t_i|t_{i-1}) \tag{12}$$

and the corresponding forecast variance is given by:

$$f(t_i) = \boldsymbol{z(t_i)}\boldsymbol{P}(t_i|t_{i-1})\boldsymbol{z(t_i)'} + \sigma^2_{\varepsilon,j} \tag{13}$$

The updating equations of the Kalman filter are

$$\boldsymbol{\alpha}(t_i|t_i) = \boldsymbol{\alpha}(t_i|t_{i-1}) + \boldsymbol{P}(t_i|t_{i-1})\boldsymbol{z(t_i)'}(y(t_i) - \hat{y}(t_i))/f(t_i) \tag{14}$$

$$\boldsymbol{P}(t_i|t_i) = \boldsymbol{P}(t_i|t_{i-1}) - \boldsymbol{P}(t_i|t_{i-1})\boldsymbol{z(t_i)})'\boldsymbol{z(t_i)}\boldsymbol{P}(t_i|t_{i-1})/f(t_i) \tag{15}$$

If all the parameters of the matrix $\boldsymbol{Q}$, are known, these recursions give an optimal estimate of the value and the corresponding variance at each point in time. The problem of daily non-trading is thus eliminated, and an eventual inter-asset correlation is incorporated in a proper way, and there is no need to interpolate on non-trading days.

5 Estimation of $\boldsymbol{Q}$ from data

Assuming normality the Kalman-filter recursions of the previous section provide a direct way of calculating the likelihood function.

$$log(L(y(t_1), \ldots, y(t_n)|\boldsymbol{Q}, \boldsymbol{\sigma}_\varepsilon)) = -\frac{n}{2}log(2\pi) - \frac{1}{2}\sum_{i=1}^{n} log(f(t_i)) - \frac{1}{2}\sum_{i=1}^{n}(y(t_i) - \hat{y}(t_i))^2/f(t_i) \tag{16}$$

The parameter $\boldsymbol{\sigma}_\varepsilon$ consists of k company-specific standard deviations measuring the bargaining scope for each asset. The matrix $\boldsymbol{Q}$ consists of $k(k+1)/2$ parameters describing the movements and comovements in time of the k assets, i.e., the total number of parameters to be estimated is $k+k(k+1)/2$. Even in a moderately sized market, this is a substantial number for numerical methods; therefore, the following approximative three-stage method is suggested.

The first step consists of decomposing each individual series into signal and noise, as described in Tómasson (2000), i.e., for each individual stock, j, estimate the model:

$$y_j(t) = \alpha_j(t) + \varepsilon_j(t) \quad \text{where} \quad V(\varepsilon_j(t)) = \sigma^2_{\varepsilon,j} \tag{17}$$

$$d\alpha_j(t) = \sigma_j dW_j(t) \tag{18}$$

where $\sigma_{\varepsilon,j}$ denotes the microstructure noise, and σ_j in equation (18) describes the daily dynamics. The Kalman filter is used to calculate the likelihood function and maximized with numerical routines. This step consists of k two-dimensional optimization problems. Having got the signal-noise parameters for each of the k stocks, one proceeds to the second step, estimating the pairwise correlations. The matrix, $\boldsymbol{Q}$, consists of $q_{lj} = \sigma_l \sigma_j \rho_{lj}$. Having obtained the estimates of σ_i and σ_j in the first step, the Kalman-filter algorithm is used to calculate the likelihood, and then it is maximized numerically. This step thus consists of $k \times (k-1)/2$ one-dimensional optimization problems. The setup is as follows:

$$y(t_i) = \boldsymbol{z}(t_i)\boldsymbol{\alpha}(t_i) + \varepsilon(t_i) \tag{19}$$

$$\boldsymbol{\alpha}(t_i) = \boldsymbol{\alpha}(t_{i-1}) + \xi(t_{i-1}, t_i) \tag{20}$$

where $\boldsymbol{z}$ is a two-dimensional vector, $(1,0)$ or $(0,1)$, depending on whether the transaction at time t_i is on stock l or j. Similarly, the variance of $\varepsilon(t_i)$ is taken as either $\hat{\sigma}^2_l$ or $\hat{\sigma}^2_j$, from the first step, depending on which stock is traded. The likelihood is calculated using equation (16) and maximized numerically with respect to $q_{lj} = \sigma_i \sigma_j \rho_{lj}$.

Repeating this for all pairs would yield an estimate, $\hat{\boldsymbol{Q}}$, for the matrix $\boldsymbol{Q}$. However, this is generally not a valid correlation matrix, i.e., due to sampling error, the matrix $\hat{\boldsymbol{Q}}$, would not necessarily be non-negative definite. To find a valid way of estimating the correlation matrix, it is necessary to enforce the matrix to be non-negative definite.

A way out is to estimate the $k \times (k-1)/2$ correlations recursively, using Choleski decomposition. The Choleski method assumes that we write the correlation matrix $\boldsymbol{\rho} = LL'$ where L is a lower triangular matrix. The recursions are as follows:

$$\rho_{ij} = \sum_{r=1}^{i-1} l_{ir} l_{jr} + l_{ij} l_{jj} \quad \text{for i=1, \dots , j-1} \tag{21}$$

$$\text{i.e. } l_{ij} = (\rho_{ij} - \sum_{r=1}^{i-1} l_{ir} l_{jr})/l_{jj} \tag{22}$$

$$1 = \rho_{ii} = \sum_{r=1}^{i-1} l_{ir}^2 + l_{ii}^2 \quad \text{for i=1, \dots k} \tag{23}$$

$$\text{i.e. } l_{ii} = \sqrt{1 - \sum_{r=1}^{i-1} l_{ir}^2} \tag{24}$$

First, the likelihood equation (16) is maximized for, say, the two most frequently traded assets, giving an estimate, $\hat{l}_{21}$, of $l_{21} = \rho_{21}/1$, that gives an estimate, $\hat{l}_{22}$, of $l_{22} = \sqrt{1 - l_{21}^2}$. Next, one maximizes the likelihood with respect to $l_{31} = r_{31}/l_{11}$, then $l_{32} = (r_{32} - l_{31} l_{21})/l_{22}$, giving $l_{33} = \sqrt{1 - l_{31}^2 - l_{32}^2}$, then $l_{41} = r_{41}/l_{11}$, etc. It is clear that the first element of L that is estimated is not restricted at all, except being smaller than 1 in absolute value. Due to equation (24), subsequent estimates of l_{ij} are subject to the condition, $(1 - \sum_{r=1}^{i-1} l_{ir}^2) > 0$. A large value of $|\hat{l}_{ij}|$ will therefore restrict later $\hat{l}_{im}$'s, for $m > j$, more than a small one would do. For correlation where there is a lot of information in the data, this restriction is not likely to be binding, but for pairs where data is sparse, it is likely that the maximum of the likelihood function will be on the boundary of the parameter space. It is therefore a good idea to estimate the correlations, l_{ij}, of the most frequently traded stocks first to prevent low-information data-pairs from enforcing restrictions on high-information data-pairs. Having an estimate of l_{ij} hitting the boundary results in a zero element of the diagonal, l_{ii}, as well as the remaining elements of the i-the line of L being zero. This is infeasible and would result in breakdown of the method. A natural remedy is to shrink the estimate of l_{ij} away from the boundary towards zero. There are a number of various principles for shrinkage, empirical Bayes, pre-test, James-Stein estimators (Judge and Bock, 1978). Here, a very simple shrinkage estimator for l_{ij} is proposed:

$$\tilde{l}_{ij} = \hat{l}_{ij}\frac{\hat{l}_{ij}^2}{(\hat{l}_{ij}^2 + V(\hat{l}_{ij}))} \tag{25}$$

The idea is that if l_{ij} is poorly estimated, then the maximum-likelihood estimate will be shrunken more than those that are well estimated. The new estimates will never be on the boundary, and all $\tilde{l}_{ii}$ will be strictly positive. This will result in an estimate of L, and the corresponding estimate of $\boldsymbol{Q}$ is given by

$$\hat{\boldsymbol{Q}} = \hat{\boldsymbol{\sigma}}\hat{L}\hat{L}'\boldsymbol{\sigma}' \tag{26}$$

The estimate of $\boldsymbol{Q}$ given by equation (26) is not a maximum-likelihood estimate. A third step towards a global maximum-likelihood estimator is therefore proposed. Maximizing the likelihood directly with gradient methods is not feasible due to the dimensionality. However, with reasonably good starting values, e.g., given by equation (26), one can approximate the maximum-likelihood estimator using a version the EM algorithm. The following version is based on a discrete time version shown in Harvey (1989). Considering the states as given, the log-likelihood for the measurements and states is given by equation (27)

$$log(L(\boldsymbol{y},\boldsymbol{A}|\boldsymbol{\sigma_\varepsilon},\boldsymbol{Q})) = -\frac{n}{2}log(2\pi) - \frac{1}{2}\sum_{i=1}^{n} log(\sigma_{\varepsilon,j(i)}) - \frac{1}{2}\sum_{i=1}^{n}(y(t_i) - \boldsymbol{z}(t_i)\boldsymbol{\alpha}(t_i))^2/\sigma_{\varepsilon,j(i)}^2 \tag{27}$$

$$-\frac{nk}{2}log(2\pi) - \frac{1}{2}\sum_{\delta_i>0} log(\delta_i) - \frac{n^*}{2}log(|\boldsymbol{Q}|) - \frac{1}{2}\sum_{\delta_i>0}(\boldsymbol{\alpha}(t_i) - \boldsymbol{\alpha}(t_{i-1}))'\delta_i^{-1}\boldsymbol{Q}^{-1}(\boldsymbol{\alpha}(t_i) - \boldsymbol{\alpha}(t_{i-1})) \tag{28}$$

where $\boldsymbol{y} = (y(t_1), y(t_2), \ldots, y(t_n))$ is a vector consisting of the measurements, and $\boldsymbol{A} = \boldsymbol{\alpha}(t_1), \ldots, \boldsymbol{\alpha}(t_n)$ is a matrix consisting of the states, $\boldsymbol{\alpha}(t)$ at each time point of measurement. The parameter, $\boldsymbol{\sigma_\varepsilon} = (\sigma_{\varepsilon,1}, \ldots, \sigma_{\varepsilon,k})$ is a vector of the standard deviation of the measurement noise for the k assets, and $\boldsymbol{Q}$, is the covariance matrix of the k-dimensional innovations for the diffusion process of states. The time between measurements is denoted by $\delta_i = t_i - t_{i-1}$, the function $j(i)$ gives the numerical code of the asset traded at time point t_i, and $\boldsymbol{z}(t_i) = (0, \ldots, 1, 0, \ldots)$ is a vector of zeroes except at coordinate $j(i)$, i.e., $\boldsymbol{z}(t_i)$ is designed to pick the relevant coordinate of $\boldsymbol{\alpha}(t_i)$. The value of n^* denotes the number non-zero δ_i's. Conditional on $\boldsymbol{y}$ and $\boldsymbol{A}$, solving the equations:

$$E(\frac{\partial log(L)}{\partial \boldsymbol{Q}}) = 0 \quad \text{and} \quad E(\frac{\partial log(L)}{\partial \sigma_{\varepsilon,j}}) = 0 \tag{29}$$

for $\boldsymbol{Q}$ and $\sigma_{\varepsilon,j}$ gives:

$$\hat{\boldsymbol{Q}} = E_n(\frac{1}{n^*}\sum_{\delta_i>0}(\boldsymbol{\alpha}(t_i)-\boldsymbol{\alpha}(t_{i-1}))(\boldsymbol{\alpha}(t_i)-\boldsymbol{\alpha}(t_{i-1}))'\delta_i^{-1} \tag{30}$$

$$\hat{\sigma}^2_{\varepsilon,j} = \frac{1}{n_j}\sum_{j(t_i)=j}(y(t_i)-\hat{y}(t_i))^2/f(t_i) \tag{31}$$

where n_j denotes the number of transactions with stock j and E_n denotes the expectation, conditional on information available at time t_n. The innovation from t_{i-1} to t_i can be written as:

$$\boldsymbol{\xi}(t_{i-1},t_i) = \boldsymbol{\alpha}(t_i)-\boldsymbol{\alpha}(t_i|t_n)-\boldsymbol{\alpha}(t_{i-1}|t_n)+\boldsymbol{\alpha}(t_{i-1})+\boldsymbol{\xi}(t_{i-1},t_i|t_n) \tag{32}$$

which gives

$$E_n(\boldsymbol{\xi}(t_{i-1},t_i)\boldsymbol{\xi}(t_{i-1},t_i)') =$$

$$P(t_i|t_n)+P(t_{i-1}|t_n)-P(t_i,t_{i-1}|t_n)-P(t_i,t_{i-1}|t_n)'+\boldsymbol{\xi}(t_{i-1},t_i|t_n)\boldsymbol{\xi}(t_{i-1},t_i|t_n)' \tag{33}$$

where

$$\xi(t_i,t_{i-1}|t_n) = \boldsymbol{\alpha}(t_i|t_n)-\boldsymbol{\alpha}(t_{i-1}|t_n) \tag{34}$$

and

$$P(t_i,t_{i-1}|t_n) = E(\boldsymbol{\alpha}(t_i)-\boldsymbol{\alpha}(t_i|t_n))(\boldsymbol{\alpha}(t_{i-1})-\boldsymbol{\alpha}(t_{i-1}|t_n))' \tag{35}$$

The expression in equations (34) and (35) can be calculated using the smoothing rules of the Kalman filter. Having calculated the smoothed values, $\boldsymbol{\alpha}(t_i|t_n)$ and $P(t_i,t_{i-1}|t_n)$, substituting equations (34) and (35) into equation (30) gives an new estimate of $\boldsymbol{Q}$. That estimate can then be used to get a new estimate, and so on, until convergence is achieved. Calculating the entire smoothed history is a tedious process, so an approximation of $\boldsymbol{\alpha}(t_i|t_n)$ is possible, i.e., only smoothing a fixed lag of the history. This can be achieved on-line by augmentation of the state-vector. This is illustrated as follows; to get an on-line estimate of $\boldsymbol{\alpha}(t_{i-1}|t_i)$ one can write:

$$\boldsymbol{\alpha}^*(t_i) = \begin{bmatrix} \boldsymbol{\alpha}(t_i) \\ \boldsymbol{\alpha}(t_{i-1}) \end{bmatrix} \tag{36}$$

The state equation of the augmented system is given by:

$$\boldsymbol{\alpha}^*(t_i) = \boldsymbol{\alpha}^*(t_{i-1}) + \begin{bmatrix} I_k \\ 0_k \end{bmatrix} \boldsymbol{\xi}(t_i) \tag{37}$$

Running the Kalman filter on the augmented state equation with the same measurement equation gives the smoothed estimate of $\boldsymbol{\xi}(t_{i-1}, t_i|t_i)$ as well as the matrix given in equation (35). If it is assumed that $\boldsymbol{\alpha}(t_{i-1}|t_i) \approx \boldsymbol{\alpha}(t_i|n)$ and $\boldsymbol{\xi}(t_{i-1}, t_i|t_i) \approx \boldsymbol{\xi}(t_{i-1}, t_i|t_n)$, these approximations can substituted into to equation (30) to get a new estimate of $\boldsymbol{Q}$. This type of augmenting of the state space is a version of the fixed interval smoother (Anderson and Moore, 1979). If smoothing further back is deemed necessary, further augmentation or direct smoothing algorithms will be needed. In both cases the computation effort will be substantial if k is, say, of the order 100.

6 Application to Icelandic stock market data

These ideas were tested on data from the Iceland Stock Exchange (ISE). The ISE is an automated, computerized market, which opened in 1991. Active trading started in 1993 and has gradually increased, in both the number of companies registered and trading intensity. By the year 2000, over 70 companies were registered. The accessible data consist of every registered transaction since the opening of the ISE. The distribution of the transactions is tabulated in Table 1. For the years 1995-1999, the extreme quantiles seem pretty much constant, whereas there might be some upward drift in the center of the distribution. The one-percent quantile in this period is very close to the tax-deduction limit. There seems to be an upward trend in the center of the distribution.

Data for all transactions on the market were available from the opening till February 24th 2000. The variables observed were, the stock code, price per share, volume in number of shares and time of transaction. The dates of each company's annual meeting of shareholders were also available. A total of 76 stocks were traded in a little over 76000 transactions with the least traded stock only traded three times. The number of transactions has increased on average by 5% each month from 1993 to 2000, see Figure 4. The outliers in Figure 4 are due to heavy trading in December, in particular on December 31st.

Year	$q_{0.01}$	$q_{0.1}$	$q_{0.25}$	$q_{0.5}$	$q_{0.75}$	$q_{0.90}$	$q_{0.99}$	n
1993	20185	68629	100188	150415	252241	853120	6073847	1878
1994	24634	74605	117000	172000	335997	879500	5698800	2605
1995	130168	139100	200800	324000	757770	1815000	11400000	2701
1996	130020	150500	240000	400256	921000	1904000	10909000	5256
1997	130002	154724	244096	456000	971508	2190000	10500000	11821
1998	130074	176000	272337	555000	1010005	1944200	8080400	12350
1999	131078	197626	331500	773499	1320000	2860287	9900000	30153

Table 1: Quantiles of transactions in Icelandic kronur.

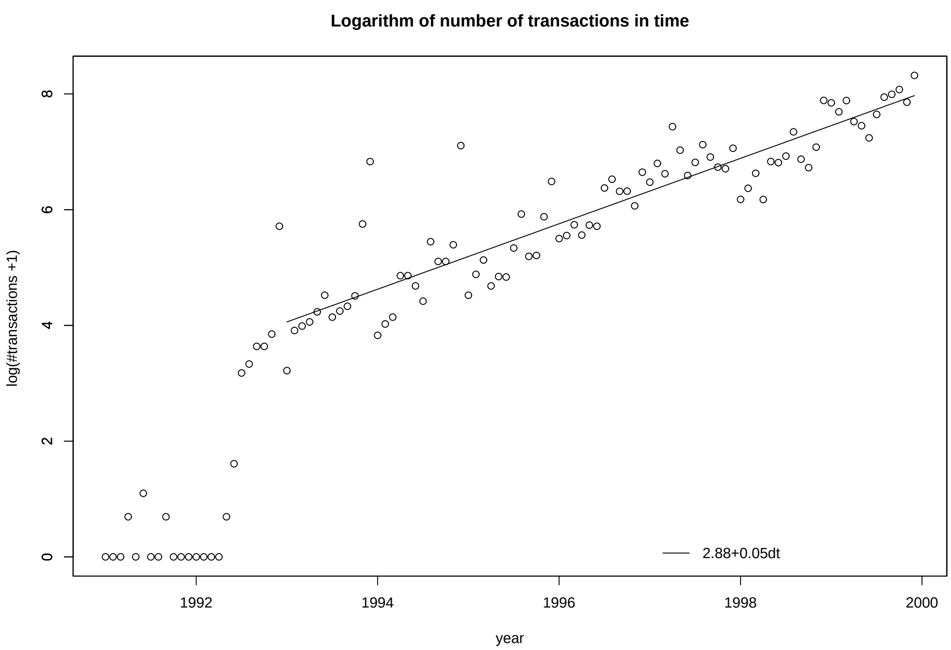

Figure 4: Logarithm of the number of monthly transactions.

The ordering of stocks affects the estimation procedure. The ordering used here is the entry date on the market, so stock 1 is the first one registered on the market. The stepwise procedure described in Section 4 was applied to the total history of transactions on the Iceland Stock Exchange in the period 1991-2000. The first step consists of getting estimates of the standard deviation of the noise and daily innovations of each stock, using the procedure described earlier (Tómasson, 2000). Conditional on these estimates all pairwise correlations are estimated one at a time. The calculation of the likelihood was performed using the author's FORTRAN subroutines, using GNU-fortran for Linux on a 200MHZ computer. For maximization of the likelihood, the nlm-function of the R-statistical language was used (Ihaka and Gentleman, 1996). The nlm-routine is for unconstrained problems but the restriction on the parameter space given by equations (22) to (24) was enforced by transforming the interval $(-\infty, \infty)$ to $(-a, a)$, using the transformation:

$$l(y) = 2a\frac{e^y}{1+e^y} - a \tag{38}$$

where a is the bound on l_{ij} set by equation (24).

For computational reasons the condition of equation (24) was sharpened. If, after shrinking the estimates of l_{ij}, the sum of the $\hat{l}_{ij}$s in a particular row, is to close to one, the corresponding diagonal element will be to close to zero or even negative. Such an event will corrupt the procedure, so the following extra condition was added. If;

$$\sum_{r=1}^{m-1} \hat{l}_{ir} > 1 - 1/i$$

all of the remaining line, except for the diagonal element, was set to zero, i.e., $\hat{l}_{ij}, j = m, \dots, i-1$ were set to zero. This condition affected 83 $\hat{l}_{ij}$'s, starting at $i = 22, j = 11$. A total of 53 pairs were deleted from the the Choleski optimizations because the stocks in question did not have overlapping trading intervals. The estimates of the correlations between those $53 + 83 = 136$ pairs of data series, are then directly decided indirectly by estimates earlier in the scheme rather than directly by their own data. Data for three stocks considered too sparse to be included at all which resulted in that, 147 correlations of of the original $75 * 74/2 = 2775$ were set to zero.

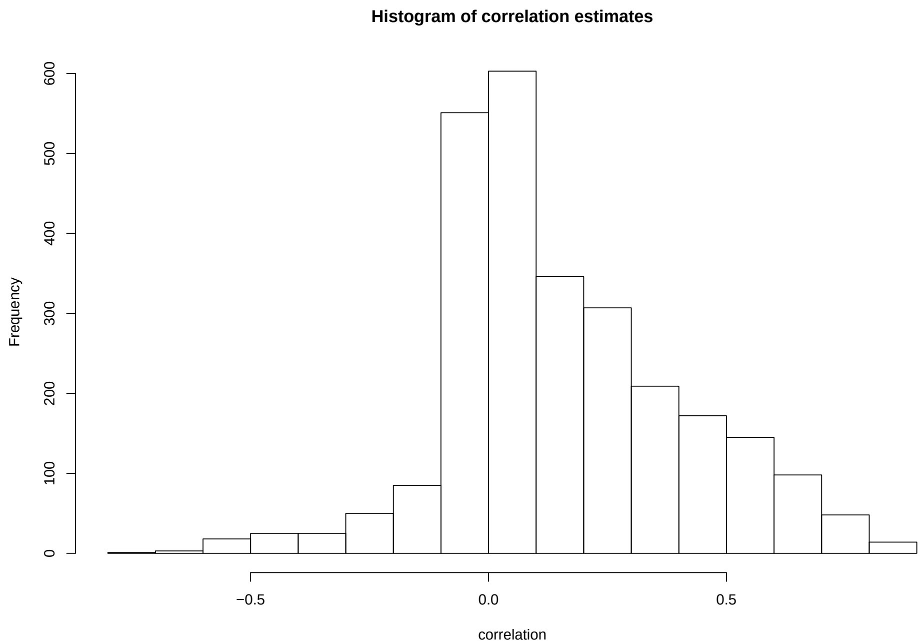

Figure 5: Histogram of correlation estimates

Figure 5 shows a histogram of the estimated correlations. Roughly 10% of the correlations are negative. The most extreme one, $\hat{\rho} = -0.76$, is based on a pair where one of the stocks had only 18 transactions. However, some estimates around -0.5 are based on a pair of more than 1000 transactions each in overlapping periods. Brief analysis suggests that most of the correlations are poorly estimated. In Table 2 the first 45 estimates of l_{ij} are shown together with the corresponding correlation. The column labeled, "st.err", represents a pseudo-standard-error, which is based on the inverse-hessian of the log-likelihood. The procedure is sequential in nature, i.e., an element in a line is restricted by earlier estimates in that line, so the Hessian is a measure of added information due to the (i, j) pair in question. It is not a direct measure of the significance of the particular correlation coefficient.

The fundamental problem of infrequent trading, i.e., lack of observations, cannot be avoided. The amount of information about the correlation of two assets depends on the number of transactions that have taken place as well as the timing of these transactions. If the trading of the two assets takes place in disjoint intervals, no information about the correlation is available. Regular transactions increase the information about the daily volatility, simultaneous observations or ones close together in time increase the information about the

i	j	$\hat{l}_{ij}$	st.err	$\hat{\rho}_{ij}$	i	j	$\hat{l}_{ij}$	st.err	$\hat{\rho}_{ij}$
2	1	0.06	0.07	0.06	8	3	0.00	0.11	0.02
3	1	0.08	0.08	0.08	8	4	0.00	0.11	0.01
3	2	0.05	0.09	0.06	8	5	0.81	0.27	0.82
4	1	0.05	0.03	0.05	8	6	0.01	0.05	0.06
4	2	0.04	0.07	0.04	8	7	0.06	0.08	0.18
4	3	0.27	0.09	0.27	9	1	0.24	0.11	0.24
5	1	0.08	0.03	0.08	9	2	0.20	0.08	0.21
5	2	0.09	0.08	0.09	9	3	0.08	0.06	0.11
5	3	0.01	0.07	0.02	9	4	0.02	0.20	0.06
5	4	0.01	0.08	0.01	9	5	0.01	0.15	0.05
6	1	0.23	0.09	0.23	9	6	0.24	0.28	0.30
6	2	0.03	0.05	0.05	9	7	0.29	0.09	0.41
6	3	0.14	0.06	0.16	9	8	-0.01	0.12	0.07
6	4	0.05	0.07	0.09	10	1	0.12	0.08	0.12
6	5	0.03	0.12	0.05	10	2	0.29	0.12	0.29
7	1	0.30	0.09	0.30	10	3	0.21	0.09	0.23
7	2	0.29	0.22	0.31	10	4	0.44	0.11	0.50
7	3	0.07	0.11	0.10	10	5	0.55	0.25	0.59
7	4	0.01	0.07	0.05	10	6	0.20	0.20	0.30
7	5	0.08	0.07	0.13	10	7	0.26	0.20	0.43
7	6	0.09	0.14	0.18	10	8	0.02	0.05	0.52
8	1	0.10	0.11	0.10	10	9	0.13	0.10	0.36
8	2	0.10	0.12	0.11					

Table 2: The first 45 estimates of l_{ij} and ρ_{ij}.

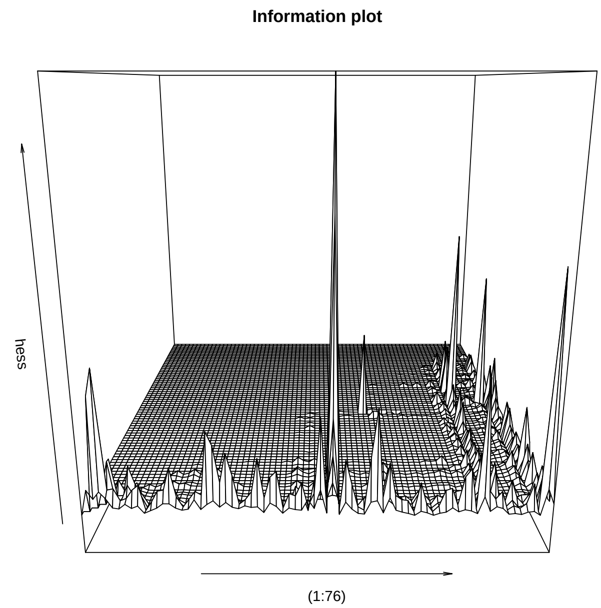

Figure 6: Information plot

standard deviation of bargaining. Similarly, two stocks traded closely in time will give more information about their correlation than those traded far apart in time. For the Icelandic stock data, a 73×73 correlation matrix, consisting of $73 \times 72/2 = 2628$ correlation coefficients was estimated by numerically maximizing the log-likelihood. The elements below the diagonal of the corresponding Hessian are plotted in Figure 6. This shows that the information is very unevenly distributed across the 2628 pairs. Actually only a few correlations can be estimated precisely with the available data. The third step in the estimation procedure, the EM iterations, was based on smoothing only one step back. This may be the explanation of why the correlations only change marginally in that step. The residual diagnostics revealed, as in analyzing the same data earlier with a different model (Tómasson, 2000), that at least the distribution of the noise component is non-normal, i.e. the kurtosis of scaled prediction errors for most stocks is much higher than 3. The interpretation of the results is that they are based on a quasi-maximum-likelihood method, the normal likelihood used for convenience. Other things might affect the normality assumption, e.g., jumps in the value which violates the continuous path assumptions. In this study the period around the annual meeting was assigned an extra-high daily innovation variance, trying to minimize the effects of jumps due to dividend payments and stock splits, but other events in the company's history might have generated jumps.

7 Discussion

This paper presents a method for estimating correlations in a multivariate Wiener-process when the sample consists of discrete observations of one or more coordinates of the vector at the time. The intended application is to stock market data, where such correlations are of interest, e.g., in risk management, VaR (value-at-risk), CAPM, etc. Despite the computational simplification steps described in Section 5, computation is still heavy. The efficient market hypothesis justifies a martingale structure for the dynamics of the value of an asset and thereby the choice of a multivariate Wiener-process. In practice one can question the efficient market hypothesis, but that would lead to a completely different, and probably more complex, setup. Shleifer (2000) gives a review of how agents might react in the environment of non-efficient markets.

The application to the Icelandic data shows that the method is feasible. It is an objective method that uses the information of each data point without adding any values, e.g., for non-trading days. It is also, in some sense, approximately optimal due to the fact that it is an approximative maximum-likelihood method. In the environment of infrequent trading, it is vital to preserve all available information.

The calculation of correlation in discrete time demands simultaneous observations. In general the lack of simultaneous observations distorts the analysis. When the approximation of simultaneous observations becomes poor, market-microstructure phenomena, as described in Campbell et al. (1997), are likely to arise. In the method developed in this paper, the demand for simultaneous observations is replaced by overlapping intervals of transactions. For good estimates of correlation coefficients, one should have many observations regularly distributed in time. For the young Icelandic stockmarket, many stocks are so rarely traded that there is virtually no information about an eventual correlation between stocks, i.e., for most pairs, observations in overlapping intervals are too few or too far apart in time. An objective method cannot create information about relations when there are no observations.

Acknowledgements

This research was supported by the Research Fund of Icelandic Banks and the Icelandic Science Fund. The author thanks the Iceland Stock Exchange for supplying data and assistants at the Institute of Economic Studies at the University of Iceland for their help.

References

Anderson, B. D. O. and Moore, J. B. (1979). *Optimal Filtering.* Prentice Hall.

Bassett, G. W., France, V. G., and Pliska, S. R. (1991). Kalman filter estimation for valuing non-trading securities, with applications to the mmi cash-future spread on october 19 and 20, 1987. *Review of Quantitative Finance and Accounting*, 1:135–151.

Campbell, J. Y., Lo, A. W., and MacKinley, A. C. (1997). *The Econometrics of Financial Markets.* Princeton University Press.

Cuthbertson, K. (1996). *The Quantitative Analysis of Financial Economics.* Wiley.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press.

Harvey, A. C. (1993). *Time Series Models.* Harvester Wheatsleaf.

Ihaka, R. and Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3):299–314.

Judge, G. G. and Bock, M. (1978). *The statistical implications of pre-test estimators and Stein-rule estimators in econometrics.* North-Holland.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47:13–37.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrics*, 35:768–783.

O'Hara, M. (1995). *Market Microstructure Theory.* Blackwell Publishers.

Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19:425–442.

Shleifer, A. (2000). *Inefficient Markets: An Introduction to Behavioural Finance.* Oxford University Press.

Tómasson, H. (2000). Signal-noise decomposition in financial markets: an empirical stochastic process analysis for infrequent trading. Working paper W00:11, The Institute of Economic Studies at the University of Iceland.

INSTITUTE OF ECONOMIC STUDIES WORKING PAPERS 1987-2000
Formerly Iceland Economic Papers Series

Editor Tryggvi Thor Herbertsson

A complete list of IoES Working Papers and Reprints can be accessed on our World Wide Web site at http://www.ioes.hi.is

W98:05 Thorvaldur Gylfason, Tryggvi Thor Herbertsson, and Gylfi Zoega: Ownership and Growth

W98:06 Thorarinn G. Petursson: Explaining the Term Stucture: The Expectations Hypothesis and Time Varying Term Premia

W98:07 Thorarinn G. Petursson: The Representative Houshold's Demand for Money in a Cointegrated VAR Model

W98:08 Tryggvi Thor Herbertsson and Gylfi Zoega: Trade Surpluses and Life-cycle Saving Behaviour

W98:09 Thorvaldur Gylfason: Privatization, Efficiency, and Economic Growth

W98:10 Tor Einarsson and Milton Marquis: Uncertain Labor Quality and Asset Pricing

W98:11 J. Michael Orzag: Actuarial Discretion vs. Rules in Regulation of Pension Funds

W98:12 J. Michael Orzag and Dennis Snower: From Unemployment Benefits to Unemployment Support Accounts

W98:13 Sveinn Agnarsson, Axel Hall, Tryggvi Thor Herbertsson, Sigurdur Ingolfsson, Gylfi Magnusson, and Gylfi Zoega: EMU and the Icelandic Labor Market

W98:14 Thorvaldur Gylfason: Exports, Inflation, and Growth

W99:01 Alison L. Booth, Yu-Fu Chen, and Gylfi Zoega: Hiring and Firing: A Tale of Two Thresholds

W99:02 Martin Paldam and Gert Tinggaard Svendsen: Is Social Capital an Effective Smoke Condenser?

W99:03 J. Michael Orszag and Dennis J. Snower: Anatomy of Policy Complementarities

W99:04 Thorvaldur Gylfason and Tryggvi Thor Herbertsson: Does Inflation Matter for Growth?

W99:05 Marco Biamchi, Bjorn R. Gudmundsson, and Gylfi Zoega: Iceland's Natural Experiment in Supply-side Economics

W99:06 Haukur C. Benediktsson, Tryggvi Thor Herbertsson, Gylfi Magnússon, and Marta G. Skúladóttir: Generational Accounts for Iceland

W99:07 Axel Hall and Jon Thor Sturluson: Testing a CGE Model: The Tax Free Year in Iceland as a Natural Experiment

W99:08 Edmund S. Phelps: Equilibrium and Disequilibrium in 20th Century 'Macro': With Attention to the Share Price Boom of the 1990s

W99:09 Kenneth F. Wallis: Macroeconometric Modelling

W99:10 Tryggvi Thor Herbertsson, Marta G. Skúladóttir, and Gylfi Zoega: Three Symptoms and a Cure: A Contribution to the Economics of the Dutch Disease

W99:11 Tryggvi Thor Herbertsson and J. Michael Orszag: Issues in European Pension Reforms: Supplementary Pensions

W99:12 Tryggvi Thor Herbertsson, J. Michael Orszag and Peter R. Orszag: Population Dynamics and Convergence in Fertility Rates

W99:13 Ragnar Arnason: Costs of Fisheries Management: Theoretical and Practical Implications

W99:14 Ragnar Arnason: Economic Instruments to Achieve Ecosystem Objectives in Fisheries Management

W99:15 Ragnar Arnason: Property Rights as a Means of Economic Organization

W00:01 Jerry Coakley, Ana-Maria Fuertes and Gylfi Zoega: Testing the Persistence and Structuralist Theories of Unemployment

W00:02 Thrainn Eggertsson: Norms in Economics – With Special Reference to Economic Development

W00:03 Thorvaldur Gylfason: Growing Apart

W00:04 Jon Danielsson: The Emperor has no Clothes: Limits to Risk Modelling

W00:05 Thorolfur Matthiasson: The Icelandic Debate on the Case for a Fishing Fee: A Non-Technical Introduction

W00:06 Willem H. Buiter: Is Iceland an Optimal Currency Area?

W00:07 Alison L. Booth and Gylfi Zoega: Why do Firms Invest in General Training? ‘Good’ Firms and ‘Bad’ Firms as a Source of Monopsony Power

W00:08 Eduard Hochreiter: “Exchange rate regimes and capital mobility: Issues and some lessons from central and eastern European applicant countries”

W00:09 Thorvaldur Gylfason: Fix or Flex? Alternative Exchange Rate Regimes in an Era of Global Capital Mobility

W00:10 Thorvaldur Gylfason: Natural Resources, Education and Economic Development

W00:11 Helgi Tomasson: Signal-noise Decomposition in Financial Markets: An Empirical Stochastic Process Analysis for Infrequent Trading

W00:12 Thorolfur Matthiasson: Changing Rules for Regulation of Icelandic Fisheries

W00:13 E. Tumusiime-Mutebile: Economic Reforms and their Impact in Uganda

W00:14 Sveinn Agnarsson: Productivity in Icelandic Fish Processing Industry 1985 – 1995: A Comparison of Methods

W00:15 Sveinn Agnarsson: Development of Efficiency in Icelandic Fish Processing Firms: A DEA Approach

W00:16 Jon Danielsson, Bjorn N. Jorgensen and Casper G. de Vries: Risk Management and Regulation in Incomplete Markets

W00:17 Ragnar Arnason, Gylfi Magnusson and Sveinn Agnarsson: The Norwegian Spring Spawning Herring Fishery: A Stylised Game Model

W00:18 Helgi Tomasson: Estimation of Correlations in Financial Markets when Trading is Infrequent