Stock price estimation using ensemble Kalman Filter square root method

D F Karya¹, P Katias¹, T Herlambang²
¹Management Department, University of Nahdatul Ulama Surabaya, Indonesia
²Information System Department, University of Nahdatul Ulama Surabaya, Indonesia

E-mail: denisfk@unusa.ac.id

Abstract. Shares are securities as the possession or equity evidence of an individual or corporation over an enterprise, especially public companies whose activity is stock trading. Investment in stocks trading is most likely to be the option of investors as stocks trading offers attractive profits. In determining a choice of safe investment in the stocks, the investors require a way of assessing the stock prices to buy so as to help optimize their profits. An effective method of analysis which will reduce the risk the investors may bear is by predicting or estimating the stock price. Estimation is carried out as a problem sometimes can be solved by using previous information or data related or relevant to the problem. The contribution of this paper is that the estimates of stock prices in high, low, and close category can be utilized as investors’ consideration for decision making in investment. In this paper, stock price estimation was made by using the Ensemble Kalman Filter Square Root method (EnKF-SR) and Ensemble Kalman Filter method (EnKF). The simulation results showed that the resulted estimation by applying EnKF method was more accurate than that by the EnKF-SR, with an estimation error of about 0.2 % by EnKF and an estimation error of 2.6 % by EnKF-SR.

1. Introduction

Shares are units of value or bookkeeping in various financial instruments referring to the ownership shares of a company [1]. The issuance of shares may enable companies that require long-term funding to sell their interests in business in return for cash. The fluctuation of stock value can be affected by various factors such as interest rate, investor psychology, political factor, and natural disaster [2]. The formation of stock prices occurs due to the demand and supply of the shares. The stock price index is an indicator that shows the value of the stock price. The value of the stock may fluctuate either in the form of increase or decrease. So investors should be able to predict whether stock prices are increasing or decreasing [3]. One method is the method of estimating the increase and decrease in stock prices. Estimation are made as a problem can very often be solved by using previous information or data related or relevant to the problem [3]. Kalman filter is a method of estimating state variables of a discrete linear dynamic system that minimizes the estimation error covariance [4].

Kalman filter was first introduced by Rudolph E. Kalman in 1960. It was about problem solving on linear discrete data filtering. Kalman filter methods are often used in robotics such as for estimation of missile trajectory [5], mobile robot trajectory [6] and Maglev ball position estimation [7]. A solution relies on the problem of linear discrete data filtering. However, in actual circumstances, there is often a continuous nonlinear dynamic system that requires another approach which is an extension of the Kalman filter called Ensemble Kalman Filter (EnKF). In the EnKF method, the algorithm is executed by generating a certain number of ensembles to calculate the mean value and the error covariance of the state variable [3]. The development of the EnKF method through modification of algorithm by adding a square root scheme at EnKF correction stage which can result in Ensemble Kalman Filter Square Root (EnKF-SR) method. This paper studied the implementation of the Square Root Scheme on the Ensemble Kalman Filter (EnKF) method on the Stock Function obtained from Software Mathematica which is then applied to estimate the Stock Function simulated with Matlab software so as to generate error ratio of EnKF and EnKF-SR methods for close price, high price, and
low price of stocks. The contribution of this paper is that the estimates of stock prices in high, low, and close category can be utilized as investors’ consideration for decision making in investment.

2. Ensemble Kalman Filter Square Root (EnKF-SR)

This section presents EnKF-SR algorithm to estimated nonlinear or linear dynamic system and measurement model, the algorithm Ensemble Kalman Filter Square Root (EnKF-SR) can be seen [5]:

Model system and measurement model

\[ x_{k+1} = f(x_k, u_k) + w_k \]  \hspace{1cm} (1)
\[ z_k = Hx_k + v_k \]  \hspace{1cm} (2)
\[ w_k \sim N(0, Q_k) \quad v_k \sim N(0, R_k) \]  \hspace{1cm} (3)

1. Initialization

Generate \( N \) ensemble as the first guess \( \bar{x}_0 \)

\[ x_{0,i} = [x_{0,1}, x_{0,2}, \ldots, x_{0,N}] \]  \hspace{1cm} (4)

The first Mean Ensemble: \( \bar{x}_{0,i} = x_{0,i} \quad 1_N \)  \hspace{1cm} (5)

The first Ensemble error:

\[ \bar{x}_{0,i} = x_{0,i} - \bar{x}_{0,i} = x_{k,i}(I - 1_N) \]  \hspace{1cm} (6)

2. Time Update

\[ \tilde{x}_{k,i} = f(\tilde{x}_{k-1,i}, u_{k-1,i}) + w_{k,i} \]  \hspace{1cm} (7)

where \( w_{k,i} \sim N(0, Q_k) \)

Mean Ensemble: \( \bar{x}^{-}_{k,i} = \tilde{x}_{k,i} \quad 1_N \)  \hspace{1cm} (8)

Error Ensemble:

\[ \bar{x}_{k,i} = \tilde{x}_{k,i} - \bar{x}_{k,i} = \tilde{x}_{k,i}(I - 1_N) \]  \hspace{1cm} (9)

3. Measurement Update

\[ z_{k,i} = Hx_{k,i} + v_{k,i} \]  \hspace{1cm} (10)

where \( v_{k,i} \sim N(0, R_k) \)

\[ S_k = H \bar{x}_{k,i} \quad E_k = (v_1, v_2, \ldots, V_N) \quad \text{and} \]
\[ C_k = S_k S_k^T + E_k E_k^T \]  \hspace{1cm} (11)

Mean Ensemble:

\[ \bar{x}_{k,i} = \bar{x}^{-}_{k,i} + \bar{x}^{-}_{k,i} \cdot S_k C_k^{-1} \left( \bar{x}_{k,i} - H \bar{x}_{k,i} \right) \]  \hspace{1cm} (12)

Square Root Scheme:

- eigenvalue decomposition from \( C_k = U_k \Lambda_k U_k^T \)  \hspace{1cm} (13)
- determine matrix \( M_k = \Lambda_k^{\frac{1}{2}} \cdot U_k^T \cdot S_k^{-} \)  \hspace{1cm} (14)
- determine SVD from \( M_k = Y_k L_k V_k^T \)  \hspace{1cm} (15)

Ensemble Error:

\[ \bar{x}_{k,i} = \bar{x}_{k,i}^{-} V_k \left( I - L_k L_k^T \right)^{\frac{1}{2}} \]  \hspace{1cm} (16)

Ensemble Estimation:

\[ \tilde{x}_{k,i} = \bar{x}_{k,i} + \bar{x}_{k,i}^{-} \]  \hspace{1cm} (17)

To evaluate of estimation result accuracy from EnKF and EnKF-SR algorithm, can be show with calculate Root Mean Square Error (RMSE) [8].

\[ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_{\text{obs},i}(k) - x_{\text{model},i}(k))^2} \]  \hspace{1cm} (18)

With

\[ x_{\text{obs},i}(k) \quad \text{= observation data} \quad n \quad \text{= iteration} \]
\[ x_{\text{model},i}(k) \quad \text{= model data} \]
3. Computational Result
This simulation made applying the EnKF and EnKF-SR algorithms to the stock functions obtained from Mathematical software simulation showed the stock data in Table 1. The simulation results were evaluated and compared to the established stock functions, and the stock functions for high, low and close prices in equation (19) - (21) are as follows:

\[ f_{\text{high}}(x) = 11,856x^2 - 503,089x + 8010,62 \]

\[ f'_{\text{high}}(x) = 23,71x - 503,089 \]  
\[ (19) \]

\[ f_{\text{low}}(x) = 53,51x^2 - 877,946x + 3875,59 \]

\[ f'_{\text{low}}(x) = 107,02x - 877,946 \]  
\[ (20) \]

\[ f_{\text{close}}(x) = 44,2843x^2 - 627,876x + 7868,65 \]

\[ f'_{\text{close}}(x) = 88,5686x - 627,876 \]  
\[ (21) \]

Because the system requires discretization, so the stock functions model in equation (19) – (21) must be discretized using the finite difference method. Equation (19) and (21). If \( f_{\text{high}} \) functional of high price stock and \( f_{\text{low}} \) functional of high price stock and \( f_{\text{close}} \) functional of close price stock

\[ f_{\text{high}} = f_{\text{high}}, f_{\text{low}} = f_{\text{low}}, f_{\text{close}} = f_{\text{close}} \]  
\[ (22) \]

The change of state variables respect to the time are approximated by forward scheme of finite difference. Thus we will get

\[ f_{\text{high}} = \frac{df_{\text{high}}}{dt} \approx \frac{f_{\text{high},k+1} - f_{\text{high},k}}{\Delta t} \]  
\[ (23) \]

\[ f_{\text{low}} = \frac{df_{\text{low}}}{dt} \approx \frac{f_{\text{low},k+1} - f_{\text{low},k}}{\Delta t} \]  
\[ (24) \]

\[ f_{\text{close}} = \frac{df_{\text{close}}}{dt} \approx \frac{f_{\text{close},k+1} - f_{\text{close},k}}{\Delta t} \]  
\[ (25) \]

from equation (23) and (25) will be gotten the modified the stock functions model in (19)-(21) below

\[
\begin{bmatrix}
  f_{\text{high},k+1} \\
  f_{\text{low},k+1} \\
  f_{\text{close},k+1}
\end{bmatrix} =
\begin{bmatrix}
  (23,71x_k^2 - 503,089)\Delta t \\
  (107,02x_k^2 - 877,946)\Delta t \\
  (88,5686x_k^2 - 627,876)\Delta t
\end{bmatrix}
\]  
\[ (26) \]

Stock data for high, low and close price in Table 1.

| Month | High  | Low  | Close |
|-------|-------|------|-------|
| 1     | 7225  | 6875 | 6925  |
| 2     | 6950  | 6575 | 6625  |
| 3     | 6700  | 6150 | 6300  |
| 4     | 6425  | 5800 | 5850  |
| 5     | 6200  | 5700 | 5725  |
| 6     | 5700  | 4500 | 4900  |
| 7     | 4900  | 4495 | 4530  |
This simulation consists of three simulations of 100, 200 and 300 ensembles applied for estimation of the water level and the steam temperature. The value of $\Delta t$ used was $\Delta t = 0.1$ and the iteration was 150. Figure 1, 2 and 3 show the results of high, low and close stock price estimates by generating 300 ensembles. Figures 1, 2 and 3 show very accurate estimation results with an error of about 0.2 and accuracy of 98%. Figure 1 shows that the estimation result by the EnKF method was more accurate than that by the EnKF-SR method with a smaller error. Figure 2 and 3 also show that the EnKF method was more accurate than the EnKF-SR method.

| Month | High  | Low  | Close |
|-------|-------|------|-------|
| 8     | 4600  | 3870 | 3950  |
| 9     | 4245  | 3760 | 3760  |
| 10    | 4000  | 3725 | 3800  |
| 11    | 3900  | 3425 | 3425  |
| 12    | 3545  | 3400 | 3470  |
| 13    | 3500  | 3000 | 3125  |
| 14    | 3200  | 2885 | 3000  |
| 15    | 3400  | 2975 | 3235  |
| 16    | 4000  | 3150 | 3930  |
| 17    | 5075  | 3950 | 4290  |
| 18    | 4410  | 4030 | 4400  |
| 19    | 5500  | 4250 | 5275  |
| 20    | 6700  | 5250 | 6100  |
| 21    | 6700  | 5950 | 6150  |
| 22    | 6675  | 6025 | 6525  |
| 23    | 6550  | 5850 | 5950  |
| 24    | 6750  | 5850 | 6750  |
| 25    | 7100  | 6200 | 7000  |

**Figure 1.** Estimation of High Stock Value using 300 Ensembles
Then, comparison of estimation results with 100, 200 and 300 ensembles was done. Table 2 shows that with the 300 ensembles the most accurate results were in the estimates of high, low, and close price of the stocks by the EnKF method with errors of 0.258 for high price error, 0.251 for low price error, and 0.24 for close price error. The most accurate results by the EnKF-SR method by generating the 100 ensembles were in the estimestes of high, low, and close price of the stocks with errors of 0.273, 0.281, and 0.276 respectively. In general as seen in the table 2, the result of the three simulations were highly accurate with accuracy of 99%.
Table 2. Comparison of the values of RMSE by the EnKF and EnKF-SR method by generating 100, 200 dan 300 Ensembles

|        | 100 Ensemble |                  | 200 Ensemble |                  | 300 Ensemble |                  |
|--------|--------------|------------------|--------------|------------------|--------------|------------------|
|        | EnKF         | EnKF-SR          | EnKF         | EnKF-SR          | EnKF         | EnKF-SR          |
| High   | 0.2731       | 1.5179           | 0.26737      | 2.1947           | 0.25855      | 2.634            |
| Low    | 0.28145      | 1.5342           | 0.27019      | 2.1954           | 0.25177      | 2.7744           |
| Close  | 0.27693      | 1.5794           | 0.25515      | 2.1805           | 0.24098      | 2.6757           |

4. Conclusion
Based on the results of the simulation analysis by generating 100, 200 and 300 ensembles, it could be concluded that the EnKF and EnKF-SR methods could be applied to estimate stock functions with high accuracy. The resulting errors were less than 2-3%. Regarding the accuracy result in the implementation of both methods, it was observed that the EnKF method gave more accurate stock price estimation result than the EnKF-SR method did.

Open problem. How to implemented Unscented Kalman Filter (UKF) and Unscented Kalman filter Square Root (UKF-SR) for estimation of stock price.

Acknowledgement
We gratefully acknowledge the support from LPPM - Nahdlatat Ulama Surabaya 0f University (UNUSA).

References
[1] Sia E R and Hansun S 2014 Rancang Bangun Aplikasi Peramalan Nilai Saham Menggunakan Algoritma Kalman Filter Jurnal Ilmiah Komputer dan Informatika (Komputa) 3 2
[2] http://forex.marketiva-id.com/tag/faktor-penggerak-pasar-saham (Accessed on February 20, 2017, at 07.45 pm)
[3] Herlambang T, Djatmiko E B and Nurhadi H 2015 Navigation and Guidance Control System of AUV with Trajectory Estimation of Linear Modelling Proc. of International Conference on Advance Mechatronics, Intelligent Manufacture, and Industrial Automation (IEEE ICAMIMIA) 184-187
[4] Ermayanti E, Aprilini E, Nurhadi H and Herlambang T 2015 Estimate and Control Position Autonomous Underwater Vehicle Based on Determined Trajectory using Fuzzy Kalman Filter Method International Conference on Advance Mechatronics, Intelligent Manufacture, and Industrial Automation (ICAMIMIA)-IEEE
[5] Herlambang T 2017 Design of a Navigation and Guidance System of Missile with Trajectory Estimation Using Ensemble Kalman Filter Square Root (EnKF-SR) The 4th International conference on Computer Application and Informations Processing Technology (CAIPT 2017)
[6] Herlambang T, Rasyid R A, Hartatik S and Rahmalia D 2017 Estimasi Posisi Mobile Robot Menggunakan Metode Akar Quadrat Unsceneted Kalman Filter (AK-UKF) Technology Science and Engineering Journal 1 2
[7] Herlambang T 2017 Estimasi Posisi Magnetic Levitation Ball Menggunakan Metode Akar Kuadrat Ensemble Kalman Filter (AK-EnKF) Rekayasa Energi Manufactur Jurnal 2 1
[8] Herlambang T, Djatmiko E B and Nurhadi H 2015 Ensemble Kalman Filter with a Square Root Scheme (EnKF-SR) for Trajectory Estimation of AUV SEGOROGENI ITS International Review of Mechanical Engineering IREME Journal 9 6 553-560