Filtering the Microstructure Noise with a Robust Kalman Filter

Yuh-Dauh Lyuu* Yu-Cheng Tien† Yun-Cheng Tsai‡

Abstract

Normally, the state approach is used to estimate the true price process from high-frequency data with microstructure noise with excellent results. When the log-price process is normally distributed, the microstructure noise is both independent and identically distributed Gaussian with zero mean and independent of true prices. In this case, the observed and the hidden vectors' covariance matrices are known, so the Kalman filter can be used to produce optimal estimates, but those covariance matrices are never known exactly and may be hard to estimate in practice. We introduce a recursive method that replaces those covariance matrices with matrices that are obtained by purely computational means in each step when applying the standard Kalman filter, proving its convergence. Simulation results show that our method gives more robust estimates than the Kalman filter ones, allowing us to obtain efficient estimators for the variance of the true price process, calculate the variance of the microstructure noise, and generate the filtered series of the log-price process.

Keywords: high-frequency, microstructure noise, Kalman filter, Kalman gain.

---

*Department of Finance and Department of Computer Science & Information Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan. E-mail: lyuu@csie.ntu.edu.tw.
The author was supported in part by the National Science Council of Taiwan, under grant 100-2221-E-002-111-MY3.

†Futures Proprietary Trading, Chinatrust Commercial Bank, No.3, Songshou Rd., Taipei 11051, Taiwan. E-mail: yu.cheng.tien@chinatrust.com.tw.

‡Corresponding author. Department of Computer Science & Information Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan. E-mail: d98922012@csie.ntu.edu.tw.
The author was supported in part by the National Science Council of Taiwan, under grant 100-2221-E-002-111-MY3.

1 Introduction

Volatility of asset prices is one of the most important variables in finance, but its estimation is non-trivial mainly because it is not directly observable. Quadratic variation, also called realized variance, is a standard tool to estimate volatility, whose error diminishes as the sampling frequency increases (see Shreve (2004) and Barndorff-Nielsen and Shephard (2002)). But as prices are sampled at finer intervals, the microstructure noise issue becomes pronounced (see Hansen and Lunde (2006)). The idea is that microstructure noise captures a variety of frictions inherent in the price process of high-frequency data such as bid-ask spread, discreteness of price changes, differences in trade sizes or informational content of price changes, gradual response of prices to a block trade, strategic component of the order flow, inventory control effects, and so on (see Hansen and Lunde (2006)). Therefore, using noisy high-frequency data may lead to entirely different volatilities from the true data, where the highest frequency is of course tick-by-tick.

Estimating volatility in the presence of microstructure noise has been widely studied (see Andersen and Bollerslev (1998), Andersen et al. (2001), Maheu and McCurdy (2002), and Areal and Taylor (2002)). Easley and O'Hara (1992) assume that the observed price equals the true price plus an error term that captures the microstructure noise. Zhou (1996) assumes the microstructure noise is independent and identically distributed Gaussian with zero mean and independent of true prices. Based on this model, Zhang et al. (2005) look into the optimal frequency to sample a continuous-time diffusion process and propose the two-scale realized variance (TSRV) estimator that uses high-frequency data to estimate the volatility in the presence of the microstructure noise.

The existence of microstructure noise evidence that the true prices are hidden variables because it is added to them in order to produce the observed data. Filtering out the noise is a key issue. Towards that goal, Cartea and Karyampas (2011) propose a batch Kalman filter-based approach, which is a general method to estimate the hidden variables, but their batch Kalman filter needs the covariance matrix of the microstructure noise and that of the true price process, both of which are difficult to identify precisely. Cartea and Karyampas (2011) use the expectation-maximization (EM) algorithm of Dempster et al. (1977) to estimate the above-mentioned covariance matrices. The initial guesses are the TSRV estimation of the true price covariance matrix and an estimated microstructure noise's covariance matrix by calculating the observed prices quadratic variation. Once the de-noised prices are available, standard estimators can be used to produce the volatility of the de-noised price, which is an efficient estimator of the true volatility, and the covariances between any two assets as well as the variance of the microstructure noise. For example, this approach estimates the variance-covariance matrix of the true prices by that of the de-noised prices.

This paper proposes a robust Kalman filter that is recursive and does not need to pre-compute the two above-mentioned covariance matrices to serve as inputs to the filter. Furthermore, when the hidden variables satisfy certain conditions (see Section 3.1), our method is provably convergent, extending Nilsson (2006) to higher dimensions. The filter is then used to estimate the true prices for two models with microstructure noise: geometric Brownian motion and Heston's stochastic-volatility model (see Heston (1993)). The root-mean-square deviation (RMSD) of the robust Kalman filter de-noised prices is found to be bounded. Furthermore, the magnitude of the RMSD depends on the two above-mentioned covariance matrices. When the covariances of both the microstructure noise and the true prices are all of the same order, our filter gives essentially identical numerical estimates as the traditional Kalman filter, holding even when the last one has access to the correct covariance matrices. Surprisingly, when the covariances of the microstructure noise are much bigger than those of the true prices, our filter performs better than the traditional Kalman filter. This finding holds again even if the traditional Kalman filter has access to the correct covariance matrices. Roughly speaking, our filter performs better because it takes advantage of the extra structures on the data and model. Because Cartea and Karyampas's approach has to estimate the covariance matrices before the Kalman filter part can proceed, it cannot perform better than it, which is supplied with the correct covariance matrices to begin with (see Lainiotis and Sims (1970) and Madjarov and Mihaylova (1993)). Hence our filter should perform better than theirs by transitivity. Like Cartea and Karyampas's approach, our filter generates de-noised prices which allows us to obtain efficient estimator of the true volatility, and the covariances between any two assets as well as the variance of the microstructure noise, all without first estimating the covariance matrices that the Kalman filter requires as inputs.

The rest of the paper is organized as follows. Section 2 gives basic definitions and a summary of the Kalman filter. Section 3 describes our method. Section 4 compares the traditional Kalman filter and our robust Kalman filter using geometric Brownian motion and Heston's stochastic-volatility model. Section 5 concludes.

2 Preliminaries

2.1 Price Process with Microstructure Noise

To fix ideas, let S_t denote the price process. The log-price is defined as $X_t = \log S_t$. Following Zhang et al. (2005), we assume that the log-price follows this process:

$$dX_t = \mu_t \, dt + \sigma_t \, dB_t, \tag{1}$$

where μ_t is the mean rate of return, σ_t is the volatility, which can be constant or stochastic, and $B_t, t \geq 0$ is a Brownain motion. When the price is sampled at higher frequencies, microstructure noise will matter. Hasbrouck and Sofianos (1993) incorporate the microstructure noise in the log-price process by postulating the observed log-price process to follow

$$Y_t = X_t + \epsilon_t, \tag{2}$$

where ϵ_t are independent noise added to the true, hidden X_t.

2.2 Kalman Filter

Kalman (1960) proposes a filter that is a fast estimate algorithm for filtering out the noise of the observations. Given a time series of data containing noise, it produces an estimate of the hidden variable. The Kalman filter assumes the true but hidden state X_t at time t evolves from its state at $t-1$ according to the following precess:

$$X_t = F_t\, X_{t-1} + M_t\, \mu_t + w_t, \tag{3}$$

where $X_t \in \mathbb{R}^{n\times 1}$, $F_t \in \mathbb{R}^{n\times n}$, $M_t \in \mathbb{R}^{n\times n}$, and $\mu_t \in \mathbb{R}^{n\times 1}$. In particular, $w_t \in \mathbb{R}^{n\times 1}$ is a random vector drawn from a zero-mean multivariate normal distribution with covariance matrix $Q_t \in \mathbb{R}^{n\times n}$.

At time t, an observation $Y_t \in \mathbb{R}^{m\times 1}$ of the hidden state X_t is made according to:

$$Y_t = H_t\, X_t + v_t,$$

where $H_t \in \mathbb{R}^{m\times n}$ maps the hidden space into the observed space and $v_t \in \mathbb{R}^{m\times 1}$ is the noise which is assumed to be a zero-mean Gaussian white noise with covariance matrix $R_t \in \mathbb{R}^{m\times m}$.

The Kalman filter is recursive. It means only the estimated state from the previous time step and the current observation are needed to compute the estimate for the current state. The prediction state of the filter is represented by $(\widehat{X}_{t|t-1},\ P_{t|t-1})$. Here, $\widehat{X}_{t|t-1}$ is the estimate of X_t given the observations up to, and including time $t-1$, and $P_{t|t-1}$ is the estimate of the a priori error covariance matrix. The updated state of the filter is represented by $(\widehat{X}_{t|t},\ P_{t|t})$. Here, $\widehat{X}_{t|t}$ is the estimate of X_t given the observations up to, and including time t, and $P_{t|t}$ is the estimate of the a posteriori error covariance matrix. Two phases alternate in the execution of the filter, with the prediction phase advancing the state until the next observation, and the update phase incorporating the observation. Specifically, the prediction steps are:

$$\begin{aligned} &\text{Predicted (a priori) state estimate: } \widehat{X}_{t|t-1} = F_t\, \widehat{X}_{t-1|t-1} + M_t\, \mu_t, \\ &\text{Predicted (a priori) error covariance: } P_{t|t-1} = F_t\, P_{t-1|t-1}\, {F_t}^T + Q_t. \end{aligned} \tag{4}$$

The updating steps are:

$$\text{Innovation: } \widetilde{Y}_t = Y_t - H_t\, \widehat{X}_{t|t-1},$$

$$\text{Optimal Kalman gain: } K_t = P_{t|t-1}\, {H_t}^T \left(H_t\, P_{t|t-1}\, {H_t}^T + R_t\right)^{-1}, \tag{5}$$

$$\text{Updated (a posteriori) state estimate: } \widehat{X}_{t|t} = \widehat{X}_{t|t-1} + K_t\, \widetilde{Y}_t, \tag{6}$$

$$\text{Updated (a posteriori) error covariance: } P_{t|t} = P_{t|t-1} - K_t\, H_t\, P_{t|t-1}, \tag{7}$$

where $H_t\, P_{t|t-1}\, {H_t}^T + R_t$ is invertible.

Assume the noise is white and the observed Y's and the hidden X's covariance matrices R and Q, respectively, are known without errors. The Kalman filter produces an unbiased a posteriori state estimate $\widehat{X}_{t|t}$ with a minimum a posteriori error covariance $P_{t|t}$. It is well-known that the Kalman filter is a minimum-square-error estimator among all linear estimators (see Harvey (1991) and Talluri and Ryzin (2005)). The use of covariances matrices different from R and Q may not guarantee unbiased and optimality (see Lainiotis and Sims (1970) and Madjarov and Mihaylova (1993)).

2.3 Properties of Matrix Norm

We now give a few fundamental results in matrix theory for use later (see Golub and Van Loan (1996)). A vector p-norm $||x||_p$ of $x \in \mathbb{R}^{n\times 1}$ is defined as

$$||x||_p \equiv \left(|x_1|^p + \cdots + |x_n|^p\right)^{\frac{1}{p}}, p \geq 1.$$

The matrix p-norm $||A||_p$ of a square real matrix A is a nonnegative number:

$$||A||_p = \max_{||x||_p=1} ||Ax||_p.$$

Fact 2.1 (Golub and Van Loan (1996)) If $A = \operatorname{diag}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{n\times n}$ is a diagonal matrix with $\sigma_1, \sigma_2, \ldots, \sigma_n > 0$, then $||A||_2 = \max_i \sigma_i$ and $||A^{-1}||_2 = 1/\min_i \sigma_i$.

Fact 2.2 (Golub and Van Loan (1996)) If $A \in \mathbb{R}^{n\times n}$ and $||A||_p < 1$, then $A + I_n$ is nonsingular and

$$(A + I_n)^{-1} = \sum_{k=0}^{\infty} (-A)^k,$$

with

$$||\,(A + I_n)^{-1}\,||_p \leq \frac{1}{1 - ||A||_p}.$$

From now on, we assume the Euclidean norm $||\cdot||_2$.

3 Methodology

3.1 State Formulations of the Multivariate Process

This section describes a state approach for the multivariate process like (1) and the corresponding Kalman filter. The approach treats this multivariate process as a hidden state variable.

We assume $m = n$, X_t is a vector random variable and the discrete time interval is Δt. The hidden vector $X_t \in \mathbb{R}^{n\times 1}$ can be a log-price process, for example. From equation (1), the true state of the Kalman filter at time t is the following processes:

$$X_t = X_{t-1} + \mu_t + w_t, \tag{8}$$

where $\mu_t \in \mathbb{R}^{n\times 1}$ is the mean rate of return, and $w_t = \sqrt{Q_t \Delta t}\, \xi_t$ is the $n \times 1$ random disturbance with $\xi_t \sim N(0, I_n)$, where $I_n \in \mathbb{R}^{n\times n}$ is the identity matrix. Let Q_t be the hidden X_t's covariance matrix, which is by definition symmetric. Our robust filter assumes equation (8). Comparing equation (8) with equation (3), we obtain $F_t = M_t = I_n$. The observed vector $Y_t \in \mathbb{R}^{n\times 1}$ will have the microstructure noise added:

$$Y_t \;=\; H_t X_t + v_t, \tag{9}$$

where $H_t \in \mathbb{R}^{n\times n}$ can be interpreted as portfolio weights, and the vector $v_t \in \mathbb{R}^{n\times 1}$ is the microstructure noise. We require that H_t be invertible, square, have full column-rank and $||H_t||_2 > 1$. We assume $v_t \sim N(0, R_t)$ is uncorrelated so $R_t \in \mathbb{R}^{n\times n}$ is a positive diagonal matrix (so R_t^{-1} exists). Equations (8) and (9) constitute the model for the data.

For applications in finance, typically neither Q_t nor R_t is known and they must be estimated. Estimating them accurately is key to the success of the traditional Kalman filter. But it is not needed by our robust estimator; instead, they will be computed in each step via an optimization procedure.

3.2 An Optimization Procedure for $\widehat{Q}_t$ and $\widehat{R}_t$

Our robust filter replaces Q_t with a diagonal matrix $\widehat{Q}_t$ in equation (4) and we use $P_{t|t-1} = P_{t-1|t-1} + \widehat{Q}_t$ for $t = 0, 1, \ldots$ to be the predicted a priori error covariance because $F_n = I_n$ for our model. Our filter also replaces R_t with an $\widehat{R}_t$. Both $\widehat{Q}_t$ nad $\widehat{R}_t$ are found by an optimization procedure with an idea similar to Nilsson (2006) but in higher dimensions. Without loss of generality, we assume $P_{0|0} = 0$.

Let $\widehat{Q}_t = \mathrm{diag}(q, \ldots, q)$, $q > 0$. So $P_{t|t-1} = P_{t-1|t-1} + q\, I_n$. Let $\widehat{R}_t = \mathrm{diag}(\sigma_t, \ldots, \sigma_t)$, where

$$\sigma_t \equiv ||H_t||_2^2\, ||P_{t|t-1}||_2\, \zeta_t \tag{10}$$

for some $\zeta_t > 1$. Both q and σ_t will be determined later. We go over some of the needed properties first. Note that

$$\sigma_t > ||H_t||_2^2\, ||P_{t|t-1}||_2. \tag{11}$$

Let σ_t further satisfy

$$\left|\left|\left(\frac{1}{\sigma_t} P_{t|t-1}\, {H_t}^T H_t + I_n\right)^{-1}\right|\right|_2 < 1. \tag{12}$$

By Fact 2.1,

$$\sigma_t = 1/||\widehat{R}_t^{-1}||_2. \tag{13}$$

Define

$$C_t \equiv H_t\, P_{t|t-1}\, {H_t}^T \widehat{R}_t^{-1}. \tag{14}$$

Lemma 3.1 If $||H_t||_2^2\, ||P_{t|t-1}||_2\, ||\widehat{R}_t^{-1}||_2 < 1$, then $C_t + I_n$ is invertible.

Proof: By the sub-multiplicative property of 2-norm,

$$\begin{aligned} ||C_t||_2 &= ||H_t\, P_{t|t-1}\, {H_t}^T \widehat{R}_t^{-1}||_2 \\ &\leq ||H_t||_2^2\, ||P_{t|t-1}||_2\, ||\widehat{R}_t^{-1}||_2 < 1. \end{aligned}$$

Hence $C_t + I_n$ is invertible by Fact 2.2. ■

Lemma 3.2 If $||H_t||_2^2\, ||P_{t|t-1}||_2\, ||\widehat{R}_t^{-1}||_2 < 1$, then $H_t\, P_{t|t-1}\, {H_t}^T + \widehat{R}_t$ is invertible.

Proof: As $\widehat{R}_t$ is invertible,

$$H_t\, P_{t|t-1}\, {H_t}^T + \widehat{R}_t = (H_t\, P_{t|t-1}\, {H_t}^T \widehat{R}_t^{-1} + I_n)\widehat{R}_t.$$

By the sub-multiplicative property of 2-norm,

$$||H_t\, P_{t|t-1}\, {H_t}^T \widehat{R}_t^{-1}||_2 \leq ||H_t||_2^2\, ||P_{t|t-1}||_2\, ||\widehat{R}_t^{-1}||_2 < 1.$$

Thus $H_t\, P_{t|t-1}\, {H_t}^T \widehat{R}_t^{-1} + I_n$ is invertible by Fact 2.2. Finally, $H_t\, P_{t|t-1}\, {H_t}^T + \widehat{R}_t$ is invertible because $\widehat{R}_t$ is. ■

By Lemmas 3.1, inequality (11) and equation (13), $C_t + I_n$ is invertible. Furthermore, $H_t\, P_{t|t-1}\, {H_t}^T + \widehat{R}_t$ is invertible by Lemmas 3.2, inequality (11) and equation (13). By equation (4) with $\widehat{Q}_t$ replacing Q_t, equation (5) with $\widehat{R}_t$ replacing R_t our Kalman gain equation is now

$$\begin{aligned} K_t &= P_{t|t-1}\, {H_t}^T \left(H_t\, P_{t|t-1}\, {H_t}^T + \widehat{R}_t\right)^{-1} \\ &= {H_t}^{-1}\, C_t\, \widehat{R}_t \left(C_t\, \widehat{R}_t + \widehat{R}_t\right)^{-1} \\ &= {H_t}^{-1}\, C_t\, \widehat{R}_t \widehat{R}_t^{-1}\, (C_t + I_n)^{-1} \\ &= {H_t}^{-1}\, C_t\, (C_t + I_n)^{-1}. \end{aligned}$$

With $\Theta_t \equiv {H_t}^{-1} C_t \, (C_t + I_n)^{-1} \, H_t$, we have $K_t = \Theta_t \, {H_t}^{-1}$. Now,

$$\begin{aligned}
\widehat{X}_{t|t} &= \widehat{X}_{t|t-1} + K_t \left(Y_t - H_t \, \widehat{X}_{t|t-1} \right) \\
&= \widehat{X}_{t|t-1} + \Theta_t \, {H_t}^{-1} \left(Y_t - H_t \, \widehat{X}_{t|t-1} \right) \\
&= \widehat{X}_{t|t-1} + \Theta_t \, {H_t}^{-1} Y_t - \Theta_t \, {H_t}^{-1} H_t \, \widehat{X}_{t|t-1} \\
&= \widehat{X}_{t|t-1} + \Theta_t \, {H_t}^{-1} Y_t - \Theta_t \, \widehat{X}_{t|t-1} \\
&= (I_n - \Theta_t) \, \widehat{X}_{t|t-1} + \Theta_t \, {H_t}^{-1} Y_t.
\end{aligned}$$

Therefore,

$$\begin{aligned}
X_t - \widehat{X}_{t|t} &= X_t - (I_n - \Theta_t) \, \widehat{X}_{t|t-1} - \Theta_t \, {H_t}^{-1} Y_t \\
&= X_t - (I_n - \Theta_t) \, \widehat{X}_{t|t-1} - \Theta_t \, {H_t}^{-1} (H_t \, X_t + v_t) \\
&= X_t - (I_n - \Theta_t) \, \widehat{X}_{t|t-1} - \Theta_t \, {H_t}^{-1} H_t \, X_t - \Theta_t \, {H_t}^{-1} v_t \\
&= X_t - (I_n - \Theta_t) \, \widehat{X}_{t|t-1} - \Theta_t \, X_t - \Theta_t \, {H_t}^{-1} v_t \\
&= (I_n - \Theta_t) \, X_t - (I_n - \Theta_t) \left(\widehat{X}_{t-1|t-1} + \mu_t \right) - \Theta_t \, {H_t}^{-1} v_t \\
&= (I_n - \Theta_t) \, (X_{t-1} + \mu_t + w_t) - (I_n - \Theta_t) \left(\widehat{X}_{t-1|t-1} + \mu_t \right) - \Theta_t \, {H_t}^{-1} v_t \\
&= (I_n - \Theta_t) \left(X_{t-1} - \widehat{X}_{t-1|t-1} \right) + (I_n - \Theta_t) \, w_t - \Theta_t \, {H_t}^{-1} v_t. \qquad (15)
\end{aligned}$$

Note that $\sigma_t \equiv ||H_t||_2^2 \, ||P_{t-1|t-1} + q \, I_n||_2 \, \zeta_t$. Our robust Kalman filter simply picks any pair $q > 0$ and $\zeta_t > 1$ that let σ_t satisfy inequalities (11)–(12). Hence, both q and σ_t are determined.

Theorem 3.3 If σ_t satisfies inequalities (11)–(12), then $|| \, (I_n - \Theta_t) \, || < 1$.

Proof: Note that

$$\begin{aligned}
I_n - \Theta_t &= I_n - {H_t}^{-1} C_t \, (C_t + I_n)^{-1} \, H_t \\
&= {H_t}^{-1} \left(I_n - C_t \, (C_t + I_n)^{-1} \right) H_t \\
&= {H_t}^{-1} \left((C_t + I_n) - C_t \right) (C_t + I_n)^{-1} H_t \\
&= {H_t}^{-1} (C_t + I_n)^{-1} H_t \\
&= \left((C_t + I_n) \, H_t \right)^{-1} H_t \\
&= \left(H_t \, P_{t|t-1} \, {H_t}^T \widehat{R}_t^{-1} H_t + H_t \right)^{-1} H_t \\
&= \left(H_t \left(P_{t|t-1} \, {H_t}^T \widehat{R}_t^{-1} H_t + I_n \right) \right)^{-1} H_t \\
&= \left(P_{t|t-1} \, {H_t}^T \widehat{R}_t^{-1} H_t + I_n \right)^{-1} {H_t}^{-1} H_t \\
&= \left(P_{t|t-1} \, {H_t}^T \widehat{R}_t^{-1} H_t + I_n \right)^{-1},
\end{aligned}$$

where we recall both $C_t + I_n$ and $P_{t|t-1}\, {H_t}^T \widehat{R}_t^{-1} H_t + I_n$ are invertible. Finally, by equation (12) and inequality (13),

$$\left|\left|\left(P_{t|t-1}\, {H_t}^T \widehat{R}_t^{-1} H_t + I_n\right)^{-1}\right|\right|_2 = \left|\left|\left(\frac{1}{\sigma_t} P_{t|t-1}\, {H_t}^T H_t + I_n\right)^{-1}\right|\right|_2 < 1.$$

■

The following shows that our scheme is convergent.

Theorem 3.4 If σ_t satisfies inequalities (11)–(12), then

$$\mathbf{E}\left[X_t - \widehat{X}_{t|t}\right] = \left(I_n - \Theta_t\right) \mathbf{E}\left[\left(X_{t-1} - \widehat{X}_{t-1|t-1}\right)\right],$$

and

$$\left|\left|\mathbf{E}\left[X_t - \widehat{X}_{t|t}\right]\right|\right|_2 \leq \left(\prod_{i=0}^{t} ||(I_n - \Theta_i)||_2\right) \left|\left|\mathbf{E}\left[X_0 - \widehat{X}_{0|0}\right]\right|\right|_2.$$

Proof: First, by equation (15),

$$\begin{aligned}
\mathbf{E}\left[X_t - \widehat{X}_{t|t}\right] &= \mathbf{E}\left[\left(I_n - \Theta_t\right)\left(X_{t-1} - \widehat{X}_{t-1|t-1}\right) + \left(I_n - \Theta_t\right) w_t - \Theta_t\, {H_t}^{-1} v_t\right] \\
&= \left(I_n - \Theta_t\right) \mathbf{E}\left[\left(X_{t-1} - \widehat{X}_{t-1|t-1}\right)\right] + \left(I_n - \Theta_t\right) \mathbf{E}\left[w_t\right] - \Theta_t\, {H_t}^{-1} \mathbf{E}\left[v_t\right] \\
&= \left(I_n - \Theta_t\right) \mathbf{E}\left[\left(X_{t-1} - \widehat{X}_{t-1|t-1}\right)\right].
\end{aligned}$$

Note that $\mathbf{E}[w_t] = 0$ and $\mathbf{E}[v_t] = 0$. Therefore,

$$\begin{aligned}
\left|\left|\mathbf{E}\left[X_t - \widehat{X}_{t|t}\right]\right|\right|_2 &= \left|\left|\left(I_n - \Theta_t\right) \mathbf{E}\left[\left(X_{t-1} - \widehat{X}_{t-1|t-1}\right)\right]\right|\right|_2 \\
&\leq ||(I_n - \Theta_t)||_2 \left|\left|\mathbf{E}\left[\left(X_{t-1} - \widehat{X}_{t-1|t-1}\right)\right]\right|\right|_2 \\
&\leq \left(\prod_{i=0}^{t} ||(I_n - \Theta_i)||_2\right) \left|\left|\mathbf{E}\left[X_0 - \widehat{X}_{0|0}\right]\right|\right|_2.
\end{aligned}$$

■

4 Experimental Results

This section compares the performance of our robust Kalman filter with that of the traditional Kalman filter on two processes: geometric Brownian motion and Heston's stochastic-volatility model. The following experiments numbers are the same as Cartea and Karyampas (2011). The experiments assume (1) $\Delta t = 1$ second, (2) the secondly microstructure noise $v_t \sim N(0, R_t)$ is Gaussian with variance $R_t = 10^{-6} I_n$, (3) there are $N = 23,400$ observations per trading day, and (4) the total initial prices are $S_0 = 30$, so $X_0 = \log S_0 = 3.4012$. The experiments generate 1,000 paths in each simulation. We then compare each estimator's RMSD of $X_t - \widehat{X}_{t|t}$, the difference of the true price and the de-noised price.

4.1 Geometric Brownian Motion

Following Cartea and Karyampas (2011), the univariate geometric Brownian motion (GBM) plus microstructure noise is assumed to follow

$$\begin{aligned} dX_t &= \sigma\, dB_t, \\ Y_t &= X_t + v_t, \end{aligned}$$

where $X_t \in \mathbb{R}^{1\times 1}$ is the unobserved true log-price, $Y_t \in \mathbb{R}^{1\times 1}$ is the observed noisy price, $\sigma^2 = 0.09$ is the annual constant variance for the process and B_t is the Brownian motion, independent of the Gaussian v_t. Comparing the above with our formulations in equations (8)–(9), we have $\mu_t = 0$, $H_t = 1$ and $Q_t = \sigma^2$. Hence, the data-generating process under our formulation is

$$\begin{aligned} X_t &= X_{t-1} + w_t, \\ Y_t &= X_t + v_t. \end{aligned}$$

We observe the prices Y_t at second t for $t = 0, 1, 2, \ldots, N$, and X_t is the unobserved true log-price at the t-th second.

The univariate case can be extended to 2 assets with the same notations as above except that $X_t, Y_t \in \mathbb{R}^{2\times 1}$ and the hidden state X_t is now $X_t = (x_{1t}, x_{2t})$. Following Cartea and Karyampas (2011),

$$Q_t = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} 0.029255 & 0.014126 \\ 0.014126 & 0.045429 \end{bmatrix}.$$

The traditional Kalman filter must know the true Q_t, R_t and X_0. Cartea and Karyampas (2011) assume $X_0 = 0$ and use the expectation-maximization (EM) algorithm of Dempster et al. (1977) to estimate Q_t and R_t. Our robust Kalman filter generates the de-noised prices without estimating any of Q_t, R_t and X_0 explicitly.

		Kalman filter	Cartea and Karyampas (2011)	Robust Kalman filter
case	1-asset	0.0009863	0.0010007	0.0010012
	2-asset	$\begin{bmatrix} 0.0009934 \\ 0.0010018 \end{bmatrix}$	$\begin{bmatrix} 0.0010012 \\ 0.0010025 \end{bmatrix}$	$\begin{bmatrix} 0.0009945 \\ 0.0010023 \end{bmatrix}$

Table 1: **Comparison of RMSDs of $X_t - \widehat{X}_t$ with geometric Brownian motion for three schemes when $R_t = 10^{-6} I_n$.**

Table 1 shows that our method can reach similar results to the traditional Kalman filter and Cartea and Karyampas (2011) when $R_t = 10^{-6} I_n$. Table 2 shows that our method

		Kalman filter	Cartea and Karyampas (2011)	Robust Kalman filter
case	1-asset	0.0094163	0.0121007	0.0100123
	2-asset	$\begin{bmatrix}0.0093445\\0.0094153\end{bmatrix}$	$\begin{bmatrix}0.0120123\\0.0120257\end{bmatrix}$	$\begin{bmatrix}0.0100894\\0.0100323\end{bmatrix}$

Table 2: **Comparison of RMSDs of $X_t - \widehat{X}_t$ with geometric Brownian motion for three schemes when $R_t = 10^{-3} I_n$.**

can reach similar results to the traditional Kalman filter and achieves lower RMSDs than Cartea and Karyampas (2011) when $R_t = 10^{-3} I_n$. Making a comparison between the first and the second case, we notice that R_t of last one is higher, but despite this characteristic, we have better performance.

4.2 Heston's Model

Heston (1993) proposes a stochastic-volatility model. Heston's model plus microstructure noise is assumed to follow

$$\begin{aligned} dX_t &= (r - \sigma_t^2/2)dt + \sigma_t\, dB_t, \\ d\sigma_t^2 &= \kappa(\alpha - \sigma_{t-1}^2/2)\, dt + \gamma\, \sigma_{t-1}\, dW_t, \\ Y_t &= X_t + v_t, \end{aligned}$$

where $X_t \in \mathbb{R}^{1\times 1}$ is the unobserved true log-price, $Y_t \in \mathbb{R}^{1\times 1}$ is the observed noisy price, $\sigma_0^2 = 0$ is the initial instantaneous variance for the process and B_t and W_t are Brownian motion processes with correlation ρ and independent of the Gaussian v_t. The parameters r, κ, α and γ are constants. Assume Feller's condition $2\kappa\alpha \geq \gamma^2$ to make the zero boundary unattainable by the process σ_t^2. Following Zhang et al. (2005), we set:

$$\rho = -0.5, r = 0.05, \kappa = 5, \alpha = 0.04, \gamma = 0.5.$$

This model does not fit the conditions for traditional kalman filter is designed.Below, we are interested in using the traditional Kalman and filter the rubust Kalman filter to do filtering and comparing their performances. We let $\mu_t = (r - \sigma_t^2/2)\Delta t$, $H_t = 1$ and $Q_t = \sigma_t^2$ in equations (8)–(9). Hence, the data-generating process under our formulation is

$$\begin{aligned} X_t &= X_{t-1} + \mu_t + w_t, \\ \sigma_t^2 &= \sigma_{t-1}^2 + \kappa(\alpha - \sigma_{t-1}^2/2)\, \Delta t + \gamma\, \sigma_{t-1}\, \eta_t, \\ Y_t &= X_t + v_t. \end{aligned}$$

where $\eta_t \sim N(0,1)$ is Gaussian and independents of v_t, and ρ is the correlation between ξ_t and η_t. We observe the prices Y_t at second t for $t = 0, 1, 2, \ldots, N$, and X_t is the unobserved true log-price at the t-th second.

The univariate case can be extended to 2 assets with the same notations as above except that $X_t, Y_t \in \mathbb{R}^{2\times 1}$ and the hidden state X_t is now $X_t = (x_{1t}, x_{2t})$. Following subsection 4.1, we set

$$Q_t = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} 0.029255 & 0.014126 \\ 0.014126 & 0.045429 \end{bmatrix}.$$

The traditional Kalman filter must know the true Q_t, R_t and X_0 so we assume the traditional Kalman knows real Q_t, R_t and X_0. On the other hand, our robust Kalman filter generates the de-noised prices without estimating any of Q_t, R_t and X_0 explicitly.

		Kalman filter	Robust Kalman filter
case	1-asset	0.0010023	0.0010051
	2-asset	$\begin{bmatrix} 0.0010522 \\ 0.0009413 \end{bmatrix}$	$\begin{bmatrix} 0.0009843 \\ 0.0010024 \end{bmatrix}$

Table 3: **Comparison of RMSDs of $X_t - \widehat{X}_t$ with Heston's model when $R_t = 10^{-6} I_n$.**

		Kalman filter	Robust Kalman filter
case	1-asset	0.0119023	0.0091052
	2-asset	$\begin{bmatrix} 0.0121075 \\ 0.0120158 \end{bmatrix}$	$\begin{bmatrix} 0.0094149 \\ 0.0093178 \end{bmatrix}$

Table 4: **Comparison of RMSDs of $X_t - \widehat{X}_t$ with Heston's model when $R_t = 10^{-3} I_n$.**

Although the traditional Kalman filter gets better results than our method gets when $R_t = 10^{-6} I_n$ (see Table 3), our method achieves lower RMSDs than the traditional Kalman filter when $R_t = 10^{-3} I_n$ (see Table 4). Our method is nonlinear where traditional Kalman filter is linear quadratic estimation. That may be why we can get better results in Heston's Model when R_t comes higher.

5 Conclusions

Cartea and Karyampas (2011) establish a batch framework using the traditional Kalman filter to generate the de-noised prices. However, it has to estimate the covariance matrices of the true price process and the microstructure noise before the framework can be used. This paper proposes a robust Kalman filter which does not need to estimate the covariance matrices. It also outputs the de-noised series recursively. When the data generator satisfies the our convergence conditions, our gain is constrained by our algorithm for convergence, hence our filter is convergent. Our filter compares favorably with the traditional Kalman filter, based on the geometric Brownian motion and Heston's stochastic-volatility model. It is found that the RMSD of the difference between the hidden vector X_t and the estimate of X_t is bounded. The simulation results show that our robust Kalman filter gives essentially identical de-noised series as the traditional Kalman filter, which has access to true covariance matrices. When the order of the microstructure noise's covariance matrix is bigger than that of the true price process, the simulation results show that our filter gives better estimates than the traditional one in terms of estimating the true prices, all without first estimating the covariance matrices that the Kalman filter requires as inputs. Our filter allows us to generate the de-noised series, obtaining efficient estimators for the variance of the true price process, and to calculate the variance of the microstructure noise without estimating covariance matrices that the Kalman filter requires as input.

References

Andersen, T. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review 39*(4), 885–905.

Andersen, T., T. Bollerslev, F. Diebold, and H. Ebens (2001). The distribution of realized stock return volatility. *Journal of Financial Economics 61*(1), 43–76.

Areal, N. and S. Taylor (2002). The realized volatility of FTSE-100 futures prices. *Journal of Futures Markets 22*(7), 627–648.

Barndorff-Nielsen, O. and N. Shephard (2002). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics 17*(5), 457–477.

Cartea, Á. and D. Karyampas (2011). Volatility and covariation of financial assets: a high-frequency analysis. *Journal of Banking & Finance 35*(12), 3319–3334.

Dempster, A., N. Laird, and D. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1–38.

Easley, D. and M. O'Hara (1992). Time and the process of security price adjustment. *The Journal of Finance 47*(2), 577–605.

Golub, G. and C. Van Loan (1996). *Matrix Computations.* Johns Hopkins University, Press, Baltimore, MD, USA.

Hansen, P. and A. Lunde (2006). Realized variance and market microstructure noise. *Journal of Business and Economic Statistics 24*(2), 127–161.

Harvey, A. C. (1991). *Forecasting, structural time series models and the Kalman filter.* Cambridge University Press.

Hasbrouck, J. and G. Sofianos (1993). The trades of market makers: an empirical analysis of NYSE specialists. *Journal of Finance 48*(5), 1565–1593.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6*(2), 327–343.

Kalman, R. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering 82*(1), 35–45.

Lainiotis, D. and F. Sims (1970). Sensitivity analysis of discrete Kalman filters. *International Journal of Control 12*(4), 657–669.

Madjarov, N. and L. Mihaylova (1993). Sensitivity analysis of linear optimal stochastic observers. *In* Proceedings of the IEEE International Conference on Systems, Man and Cybernetics — Systems Engineering in the Service of Humans, Le Touquet, France, pp. 482–487.

Maheu, J. and T. McCurdy (2002). Nonlinear features of realized FX volatility. *Review of Economics and Statistics 84*(4), 668–681.

Nilsson, M. (2006). Kalman filtering with unknown noise covariances. *In* Proceedings of the Reglermote, Stockholm, pp. 30–31.

Shreve, S. E. (2004). *Stochastic Calculus for Finance II: Continuous-Time Models.* New York: Springer.

Talluri, K. T. and G. J. Ryzin (2005). Estimation and forecasting. *The Theory and Practice of Revenue Management*, 407–514.

Zhang, L., P. Mykland, and Y. Ait-Sahalia (2005). A tale of two time scales. *Journal of the American Statistical Association 100*(472), 1394–1411.

Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. *Journal of Business & Economic Statistics 14*, 45–52.