Non-Linear Filtering in the Estimation of a Term Structure Model of Interest Rates

SHU WU
Department of Management Science and System Engineering
Nanjing University of Aeronautics and Astronautics
29 Yu Dao Street, Nan Jing, Jiang Su Province
P.R.CHINA 210007
iamwushu@hotmail.com

Abstract: - The methods of the class of Kalman filters have recently been used in the estimation of the term structure of interest rates. These methods can employ both time-series and cross-sectional aspects of term structure models. This paper compares the performance of two kinds of non-linear Kalman filter algorithms - Extended Kalman Filter (EKF) and Square-Root Unscented Kalman Filter (SRUKF) in estimating one popular exponential-affine term structure model. Simulation results show that SRUKF is of higher approximation accuracy and stronger numerical stability than EKF is.

Key-Words: - Non-linear filtering, Square-Root Unscented Kalman Filter, Extended Kalman Filter, Term structure of interest rates, Exponential-affine term structure model, One-factor Vasicek model

1 Introduction

The term structure of interest rates describes the relationship between bond rates of different terms and it is of great interest to people due to its position of a leading indicator for economic activities. The modeling and estimation of the dynamics of term structure is a main issue in term structure analysis.

Given the high correlation among bond yields of different maturities, many term structure models have been established to explain the joint movements. In view of the stochastic volatility characteristics of short-term interest rates, a class of equilibrium models has been proposed. Merton's paper in1971 is the first study that uses a dynamic continuous-time model to describe the term structure of interest rates [1]. Vasicek notices the equilibrium characterization of the term structure and derives the expected rate of return on any bond in excess of the spot rate [2]. Cox, Ingersoll and Ross use a general equilibrium asset pricing model to study the term structure [3]. More recently, Duffie and Kan have developed a unified framework of exponential-affine representation for many term structure models, which contains models of Vasicek, Cox, Ingersoll and Ross, Longstaff and Schawartz [4] and so on [5].

According to the number of factors, term structure models are classified into one-factor and multi-factor models. One-factor models have often been analyzed in the literature for their simplicity. What's more, there is evidence showing that almost 90 percent of the variation in the changes of the yield curve of bond rates is attributable to the variation in the first factor based on principle component analysis. For example, for most one-factor models, the factor is generally taken to be the instantaneous short rate [6].

Most modeling approaches are to divide the stochastic movement of the instantaneous short rate into two parts using a stochastic differential equation. The first part is the drift of the process, which is deterministic. The second part is the volatility component of the process, which is the random part. Examples are one-factor Vasicek model [2], multi-factor extensions of Vasicek model [7], two-factor Brennan-Schwartz model [8], one-factor Cox-Ingersoll-Ross model [3] and two-factor extensions of Cox-Ingersoll-Ross model [4], among many others.

Although the modeling progress keeps going, the estimation techniques are relatively immature because of the complexity in the models. For many models, the unknown probability distribution of bond rates or yields and the unobserved state variables pose the challenge. In recent years, it has been popular to express term structure models in the state-space forms and then adopt methods of the class of Kalman filters to do the estimation [9,10,11]. These filtering estimation methods have provided a new vision for the study of term structure models. Filtering is a natural approach when the underlying state is unobserved. Kalman Filter (KF) is an optimal filter for recursive estimation for unobserved state variables [12]. KF applies where the process and measurement noises follow Gaussian distributions and the system is linear. Moreover, a wide variety of financial models are nonlinear, so it is necessary to

use nonlinear filtering algorithms in which Extended Kalman Filter (EKF) is one of the most popular. But when a model presents highly nonlinear nature, EKF-based estimator may diverge mainly because EKF handles nonlinear systems based on an approximate first-order Taylor series expansion around the mean values. To compensate this, many options are proposed. Examples are Particle Filter (PF) [13], Unscented Kalman Filter (UKF) [14] and Square-Root Unscented Kalman Filter (SRUKF) [15]. Due to huge computational expense of PF, this paper considers Unscented Filter. Furthermore, to prevent the covariance matrix from becoming non-positive semi-definite, the Square-Root form is adopted. SRUKF is put forward to avoid keeping calculating the square roots of state covariance matrix in UKF and thus to avoid more calculation errors [15]. The central idea of SRUKF is to use the form of Cholesky decomposition to directly propagate forward and update the square roots of state covariance matrix. During the process, two powerful linear algebra techniques are used: QR decomposition and Cholesky decomposition, which help enhancing the calculation efficiency.

2 EKF Implementation

As is well-known, Kalman Filter is to acquire optimal dynamic estimation according to minimum mean-square state error rule under the condition of linear systems and Gaussian noises. However, strictly speaking, in the real world, nearly all the systems are non-linear. Most are even of high non-linearity. During the real time estimation of dynamic systems, usually the non-linearity nature is one of the important factors affecting optimal estimation results. To solve the estimation problem of non-linear systems, the usual way is to transform the non-linear issue into a linear one by virtue of some linearization techniques. The most common way is first-order Taylor series expansion adopted in Extended Kalman Filter.

The extended Kalman filter allows simultaneous estimation of states and parameters. These parameters are considered as extra states in an augmented state vector. This augmented model is non-linear because of multiplication of states. Thus, it must be linearized along the state trajectory to give a linear perturbation model [16]. Extended Kalman Filter is applied to dynamic models that are in a state-space representation, which include state and observation equations. Consider the following general nonlinear state-space system:

$$x_k = f(x_{k-1}) + G_k w_k \tag{1}$$

$$y_k = h(x_k) + v_k \tag{2}$$

where equation (1) is state equation, equation (2) is observation equation, the two equations form a space-state model, $f(\bullet)$ and $h(\bullet)$ are both nonlinear functions, x_k is state vector, y_k is observation vector, G_k denotes the intensity of the noise w_k. w_k and v_k are zero-mean, white and Gaussian noises. Some former models neglect observation noise. In fact, many factors affect testing data. Observation data includes disturbance, which is called "observation noise". Observation noise has white noise and colored noise. To compute simply, we treat disturbance as white noise with zero mean [17].The joint covariance matrix of w_k and v_k is:

$$E\left[\begin{pmatrix} w_k \\ v_k \end{pmatrix}\begin{pmatrix} w_k^T & v_k^T \end{pmatrix}\right] = \begin{bmatrix} Q_k & 0 \\ 0 & R_k \end{bmatrix} \tag{3}$$

1) The initial condition:

$$\hat{x}(0|-1) = E\{x_0\} = \bar{x}_0 \tag{4}$$

$$P(0|-1) = E\left\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\right\} = \Sigma_0 \tag{5}$$

where $\hat{x}(\bullet)$ denotes priori estimation, $P(\bullet)$ is the corresponding error covariance matrix of the state variable.

2) Approximate linearized system:

Because of the nonlinearity of $f(\bullet)$ and $h(\bullet)$, the conventional Kalman updating cannot be applied for such cases. A solution for this is to approximate the linearity by first-order approximation, i.e.,

$$f(x) \approx f(x_0) + \nabla f\big|_{x_0}(x - x_0) \tag{6}$$

With this approximation, we can evaluate the Kalman updating in the vicinity of estimated state. This is the so called Extended Kalman Filter. Its idea is clearly shown in Fig.1 [18].

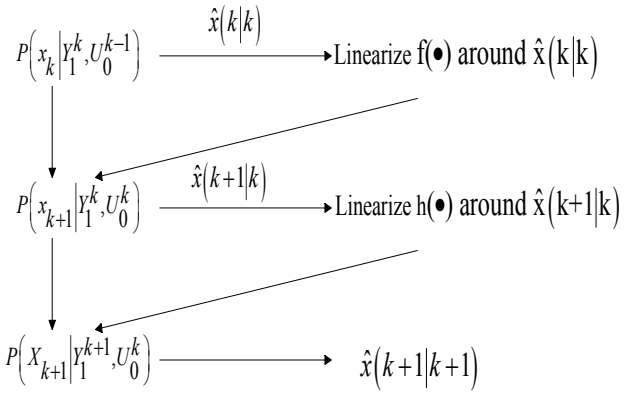

Fig.1 Dynamic Concept of Extended Kalman Filter

If we denote $F_{k,k-1} = \left.\frac{\partial f(x_{k-1})}{\partial x}\right|_{\hat{x}_{(k-1|k-1)}}$, we have

$$\begin{aligned} x_k &= f(\hat{x}_{k-1}) + F_{k,k-1}(x_{k-1} - \hat{x}_{k-1}) + G_k w_k \\ &= F_{k,k-1}x_{k-1} + [f(\hat{x}_{k-1}) - F_{k,k-1}\hat{x}_{k-1}] + G_k w_k \end{aligned} \quad (7)$$

where $[f(\hat{x}_{k-1}, k) - F_{k,k-1}\hat{x}_{k-1}]$ is known fixed value.

Similarly, if we denote $H_{k,k} = \left.\frac{\partial h(x_k)}{\partial x}\right|_{\hat{x}_{(k|k-1)}}$, we have [19]

$$\begin{aligned} y_k &= h(\hat{x}_{k|k-1}) + H_{k,k}(x_k - \hat{x}_{k|k-1}) + v_k \\ &= H_{k,k}x_k + [h(\hat{x}_{k|k-1}) - H_{k,k}\hat{x}_{k|k-1}] + v_k \end{aligned} \quad (8)$$

With the above-said linearization and simplification, the Kalman update can be applied.

3) Time updates:

The mean x_{k+1} and variance $P_{k+1|k}$ can be calculated recursively by an application of the one-step ahead prediction equations:

$$\hat{x}_k = f(\hat{x}_{k|k-1}) \quad (9)$$

$$P_{k|k-1} = F_{k,k-1}P_{k-1}F_{k,k-1}^T + G_{k-1}Q_{k-1}G_{k-1}^T \quad (10)$$

4) Measurement updates:

If we define the Kalman Gain here as

$$K_k = P_{k,k-1}H_{k,k}^T(H_{k,k}P_{k,k-1}H_{k,k}^T + R_{k-1})^{-1} \quad (11)$$

we have

$$\hat{x}_k = \hat{x}_{k|k-1} + K_k[y_k - h(\hat{x}_{k|k-1})] \quad (12)$$

$$P_k = P_{k|k-1} - K_k H_k^k P_{k|k-1} \quad (13)$$

The equation (12) indicates a very important conclusion that the filtered state estimate equals to the predicted state estimate plus the Kalman Gain multiplied by the innovation.

Although compared with KF, EKF not only mains the computationally efficient recursive update form of the KF, but also presents an improved way in solving nonlinear problems, there are several defects in EKF. First, linearized transformations are only reliable if the error propagation can be well approximated by a linear function [20]. Second, derivation is a necessary step in EKF algorithm, but it is not always the case that Jacobian matrix exists for every system. Last, sometimes it is nearly impossible to calculate the Jacobian matrix or even if it is calculated by great patience, the whole process is prone to errors.

3 SRUKF Implementation

Similar as the conventional robust Kalman filter, the robust implementation of UKF also utilizes the square root algorithm for covariance propagation. The detail of this algorithm can be found in [15].

In UKF, the most computationally expensive part lies in calculating new sigma points during time update every time and this requires the computation of a matrix square-root of the state covariance matrix [21]. However, recently, Merwe R Vander and Wan have introduced the SRUKF [15], a reimplementation of the general UKF which delivers exactly the same results (to within machine accuracy), but which cleverly avoids the decomposition by directly propagating the Cholesky factor rather than the covariance [22]. In SRUKF, the matrix square-root will be propagated directly, avoiding the re-factorizing at each re-sampling point. Consequently, numerical stability is well kept. In the algorithm of SRUKF, the QR decomposition (noted as $qr(\bullet)$), Cholesky factor updating and efficient least square techniques are employed. The key idea is to propagate the covariance matrix in its cholesky factor instead of the full covariance matrix.

From the previous section, we know that for the estimation problems of non-linear systems, the traditional linearization way is to approximate the nonlinear mapping itself and then apply the standard KF method. In fact, it is easier to approximate a probability distribution than it is to approximate an arbitrary nonlinear function or transformation [20]. Thus, in [14], UT transformation is put forward. SRUKF is just based on the idea of UT transformation. The UT approach is illustrated in Fig.2. A set of points (sigma points) are chosen so

that their mean and covariance are $\bar{x}$ and S. The nonlinear function is applied to each point, in turn, to yield a cloud of transformed points. The statistics of the transformed points can then be calculated to form an estimate of the nonlinearly transformed mean and covariance [20].

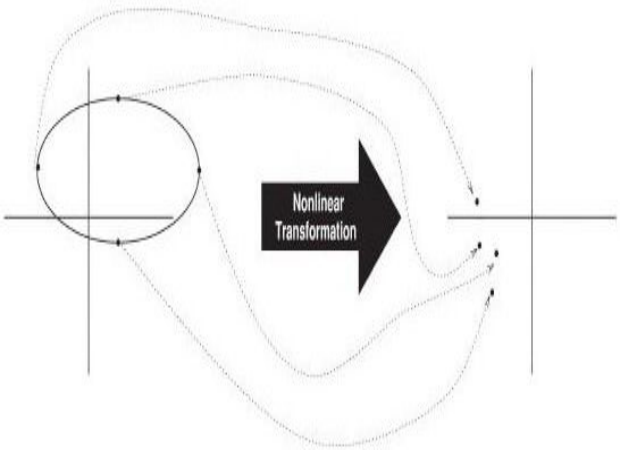

Fig.2 the Principle of the UT [20]

In terms of simplicity conceptually and convenience in application, the UT algorithm is of the same appeal as linearization for the EKF. Yet the UT method exhibits much more sufficient accuracy in the case of some highly nonlinear systems.

Since the time when the UT transformation was proposed, the UT method has found a number of applications in various nonlinear systems. More recently, van der Merwe and Wan have developed a square root formulation of the UT method. For the algorithm of SRUKF, the system model is assumed to be of the same form of equation (1) and equation (2). The definition of each variable is also the same as that in the corresponding part in the previous section. The algorithm of SRUKF is described as following:

1) The initial condition:

$$\hat{x}_0 = E(x_0) \tag{14}$$

$$S_0 = chol\left\{E\left[\left(x_0 - \hat{x}_0\right)\left(x_0 - \hat{x}_0\right)^T\right]\right\} \tag{15}$$

where *chol* means Cholesky factorizing [15]. All symmetric nonnegative definite matrices have Cholesky factors. If $A = XX^T$, where A is a symmetric nonnegative definite matrix, X is a triangular matrix, then X is a triangular Cholesky factor, indicated as $X = chol(A)$.

2) Computation of the set of sigma points:

$$\begin{aligned} X_{k-1} &= \left[\left(X_{k-1}\right)_0 \quad \left(X_{k-1}\right)_i \quad \left(X_{k-1}\right)_j\right] \\ &= \left[\hat{x}_{k-1} \quad \hat{x}_{k-1} + \sqrt{L+\lambda}S_k \quad \hat{x}_{k-1} - \sqrt{L+\lambda}S_k\right] \end{aligned} \tag{16}$$

where $i = 1,2,...,L$, $j = 1,2,...,L$, $\lambda = \alpha^2\left(L+\kappa\right) - L$ is a scaling parameter which can be adjusted to enhance the approximation of the distribution of the state vector, α is a positive scaling parameter which can be made arbitrarily small to minimize the higher order effects, κ is a secondary scaling parameter. There are $2L+1$ sigma points to be required and L is the dimension of the state vector.

3) Time updates:

The transformed set propagates through the nonlinear state equation:

$$X_{k|k-1} = f\left(X_{k-1}\right) \tag{17}$$

The predicted mean is calculated as:

$$\hat{x}_k^- = \sum_{i=0}^{2L} W_i^{(m)} X_{i,k|k-1} \tag{18}$$

where $W_0^{(m)} = \dfrac{\lambda}{L+\lambda}$, $W_i^{(m)} = \dfrac{1}{2L+2\lambda}$, $i = 1,2,...,2L$.

The Cholesky form of the covariance is predicted using:

$$S_k^- = qr\left[\sqrt{W_1^{(c)}}\left(X_{1:2L,k|k-1} - \hat{x}_k^-\right) \quad \sqrt{Q_k}\right] \tag{19}$$

$$S_k^- = cholupdate\left\{S_k^-, X_{0,k} - \hat{x}_k^-, W_0^{(c)}\right\} \tag{20}$$

where *qr* represents the QR decomposition of the matrix and cholupdate represents the Cholesky factor update. The QR decomposition of a matrix $A \in \Re^{M\times N}$ is given by $A^T = QR$, where $Q \in \Re^{N\times N}$ is orthogonal, $R \in \Re^{N\times M}$ is upper triangular and $N \geq M$ [23].

4) Measurement updates:

The sigma points propagate through the nonlinear measurement equation:

$$Y_{k|k-1} = h\left(X_{k-1}\right) \tag{21}$$

The predicted mean observation is calculated as:

$$\hat{y}_k^- = \sum_{i=0}^{2L} W_i^{(m)} Y_{i,k|k-1} \tag{22}$$

The innovation Cholesky covariance is given by:

$$S_{\tilde{y}_k} = qr\left[\sqrt{W_1^{(c)}}\left(Y_{1:2L,k} - \hat{y}_k^-\right) \quad \sqrt{R_k}\right] \tag{23}$$

$$S_{\tilde{y}_k} = cholupdate\left\{S_{\tilde{y}_k}, Y_{0,k} - \hat{y}_k^-, W_0^{(c)}\right\} \tag{24}$$

The cross-covariance matrix of x and y is determined by:

$$P_{x_k y_k} = \sum_{i=0}^{2L} W_i^{(c)} \left(X_{i,k|k-1} - \hat{x}_k^- \right) \left(Y_{i,k|k-1} - \hat{y}_k^- \right)^T \quad (25)$$

where $W_0^{(c)} = \frac{\lambda}{L+\lambda} + \left(1 - \alpha^2 + \beta\right)$, $W_i^{(c)} = \frac{1}{2L+2\lambda}$, $i = 1,2,...,2L$, β is an extra degree of freedom scalar parameter used to incorporate extra prior knowledge of the distribution of the state variable.

The Kalman gain is:

$$K_k = \left(P_{x_k y_k} / S_{\tilde{y}_k}^T \right) / S_{\tilde{y}_k} \quad (26)$$

The update mean is:

$$\hat{x}_k = \hat{x}_k^- + K_k \left(y_k - \hat{y}_k^- \right) \quad (27)$$

The update Cholesky factor is:

$$S_k = cholupdate\left\{ S_k^-, K_k S_{\tilde{y}_k}, -1 \right\} \quad (28)$$

The computational complexity associated with the square-root implementation of the unscented Kalman filter is similar to that of the original unscented Kalman filter, but the square-root implementation is more stable numerically [24].

In the previous part, the general formulation for state estimate is outlined using the unscented Kalman filter in square root form. As mentioned in [22], the SRUKF is in general $O(N^3)$ for state estimation, but is $O(N^2)$ for parameter estimation. To estimate parameters using the unscented Kalman filter, it is necessary to introduce additional state variables $x^p(t)$ representing the unknown parameters, i.e.

$$x(t) := \{x^s(t) \quad x^p(t)\}$$

Where the dynamics of the additional parameters follow a white-noise process

$$x_k^p = w_k^p$$

and w_k^p is a zero mean Gaussian process noise with covariance Q_k^p [25]. Although the parameters are usually assumed to be constant in the model of the system, the process noise is used to aid in convergence of the estimate particularly when there may be a lot of uncertainty in the initial estimate. Note that the initial covariance of the unknown parameters also affects the rate of convergence of the estimate. It can also be important for online estimation of parameters to quickly detect changes to the system parameters [25].

To further understand the superiority of UT method intuitively, two examples are shown in Fig.3 and Fig. 4.

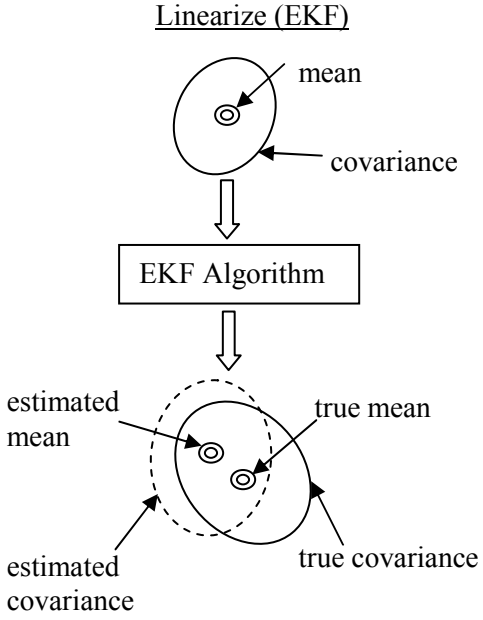

Fig.3 Example of EKF for Mean and Covariance Propagation

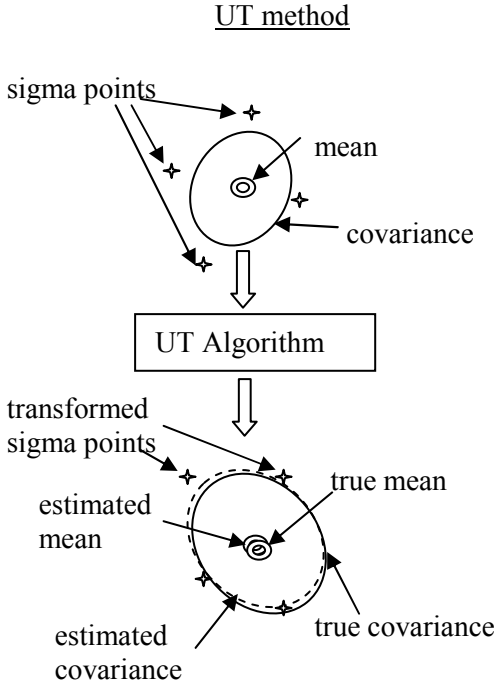

Fig.4 Example of UT for Mean and Covariance Propagation

The two figures are revised and plotted based on the examples often seen in the literature of E. Wan, R. Merwe and A. Nelson [26]. Fig.3 shows the results using a linearization approach as would be done in the EKF. Fig.4 shows the performance of the UT. The superior performance of the UT is clear.

4 Model Description

We utilize a result of Duffie and Kan which establishes the sufficient and necessary conditions for the obtainment of an exponential-affine term structure model. The exponential-affine term structure model is a class of models in which the yields to maturity are affine functions in some abstract state variable vector as X_t indicated in equation (29). It is constructed by assuming that bond yields are a linear function of the underlying state variables that provide uncertainty in the model. The Duffie-Kan model is as follows [5]:

$$dX_t = U(X_t, \mathbb{R})dt + V(X_t, \mathbb{R})dz_t \tag{29}$$

where X_t is a state variable; $\mathbb{R}$ is a set containing model parameters; dz_t denotes an independent Wiener process and $dz_t = \varepsilon\sqrt{dt}$, $\varepsilon \sim N(0,1)$. In this model, the state variable driving the dynamics of the term structure is specified in a stochastic process. The change of the value of the state variable (dX_t) is divided into two parts, the first part is deterministic ($U(X_t, \mathbb{R})$), called the drift of the process; and the second is a random part ($V(X_t, \mathbb{R})$), which is the variance of the process. This involves the assumption, that the interest rate process is generated by a standard Brownian motion, also known as a Wiener process [11].

The corresponding bond pricing formula for this class of model can be generically expressed as:

$$P_t(X_t, \mathbb{R}, \tau) = A(\mathbb{R}, \tau)e^{-B(\mathbb{R},\tau)X_t} \tag{30}$$

where $\tau(t,T) = T - t$ denotes term with t representing the time the bond starts and T representing the time the bond matures; $P_t(X_t, \mathbb{R}, \tau)$ is the bond price. Duffie and Kan have shown that $P(\bullet)$ is generically exponential-affine if and only if $U(\bullet)$ and $V(\bullet)$ are affine in X_t .

Among the existing exponential-affine term structure models, one-factor Vasicek model has gained prominence in the literature of derivative contract pricing. The Vasicek (1977) model is a one-factor partial equilibrium model and starts out with the specification of a time series process for the instantaneous spot interest rate which is treated as the only factor of uncertainty [2]. The no-arbitrage restriction then permits the derivation of a bond pricing formula whereby the bond price is a function of the unobserved instantaneous spot rate and the model's parameters [11].

In the Duffie-Kan model, if we set some limits to the variables, we can get the representation of one-factor Vasicek model. In Vasicek model, the unobserved state variable is the instantaneous interest rate r_t .

In the following part of this section, we introduce the state-space formulation of the interest rate stochastic volatility model - one-factor Vasicek model and use this example to illustrate the effectiveness of SRUKF.

As a popular one-factor stochastic volatility model, Vasicek model is widely used in term structure literature. The one-factor Vasicek model is characterized by one factor, the instantaneous interest rate r_t that evolves in continuous time. The model is given by the following first-order stochastic differential equation [2]:

$$dr_t = k(\theta - r_t)dt + \beta_1 dz_t \tag{31}$$

In the movement of the instantaneous interest rate, the interest rate appears to be pulled back to some long-term average level over time and this phenomenon is known as mean reversion. In equation (31), k , θ , β_1 are all positive constants, k is the mean-reverting intensity, θ is the long-run average of the instantaneous interest rate r_t which is the state variable . When the instantaneous interest rate deviates from its long-term mean, θ , it will revert back to this mean at a speed governed by the parameter k . For the second part of this stochastic differential equation, β_1 is the volatility parameter of the process, dz_t denotes an independent Wiener process and $dz_t = \varepsilon\sqrt{dt}$, $\varepsilon \sim N(0,1)$. This process is hampered in its ability to revert back to its mean level just by the second random part, known as the diffusion term, which essentially shocks the process at each step in time.

The process specified by the aforementioned stochastic differential equation is defined in continuous time, while the observed data are sampled at discrete time intervals. Before applying directly the algorithm of EKF and SRUKF, we try to put equation (31) into discrete form. If we use daily data

to do the estimation, we can set $dt = 1/365$. Then we get:

$$\begin{aligned} r_{t+1} &= r_t + \left[k(\theta - r_t)\right]/365 + \left(\beta_1 \varepsilon\right)/\sqrt{365} \\ &= \left(1 - k/365\right) r_t + \left(k\theta\right)/365 + \left[\beta_1/\sqrt{365}\right] * \varepsilon \\ &= \left(1 - k/365\right) r_t + \left(k\theta\right)/365 + w_t \end{aligned} \quad (32)$$

where $w_t = \left[\beta_1/\sqrt{365}\right] * \varepsilon$, ε is zero-mean, white Gaussian noise with unit variance.

In Vasicek model, the instantaneous interest rate r_t is the unobserved state variable, while the corresponding bond price data can be observed. The corresponding bond pricing formula is:

$$\begin{aligned} P_t(r_t, \mathbb{R}, \tau) &= A_j(\mathbb{R}, \tau) e^{-B_j(\mathbb{R}, \tau) r_t} + \beta_2 \varepsilon' \\ &= A_j(\mathbb{R}, \tau) e^{-B_j(\mathbb{R}, \tau) r_t} + v_t \end{aligned} \quad (33)$$

where $\mathbb{R}$ is a set containing model parameters, $\tau(t,T) = T - t$ denotes term with t representing the time the bond starts and T representing the time the bond matures, ε' is zero-mean, white Gaussian noise with unit variance, $P_t(r_t, \mathbb{R}, \tau)$ is the bond price. Here the measurement noise v_t is added to equation (33) because the equation involves the problem of estimation. The noise $v_t = \beta_2 \varepsilon'$ is zero-mean, white Gaussian noise with variance β_2^2. In equation (33),

$$A(\mathbb{R}, \tau) = \exp\left\{\gamma\left[B\left(\mathbb{R}, \tau\right) - \tau\right] - \frac{\beta_1^2 B^2\left(\mathbb{R}, \tau\right)}{4k}\right\} \quad (34)$$

$$B(\mathbb{R}, \tau) = \frac{1}{k}\left(1 - e^{-k\tau}\right) \quad (35)$$

$$\gamma = \theta + \frac{\beta_1 \varphi}{k} - \frac{\beta_1^2}{2k^2} \quad (36)$$

where φ denotes the market price of risk.

Therefore, equation (32) and equation (33) are the discrete time state-space specifications of Vasicek model. This is a non-linear system, and in addition, the measured input data are also corrupted by noises. All these factors require the use of an observer that can provide a good filtering and unknown parameter estimation. The extended Kalman filter (EKF), which is a full-order stochastic observer, can be the solution to these issues [27]. Moreover, as an improved non-linear filtering method, SRUKF is also applied here to do the comparison of state estimation and parameter estimation with that of EKF.

5 Empirical Results

Various kinds of methods have been put forward in the finance literature for the estimation of the term structure of interest rates, including that of the one-factor Vasciek model. Usually, these methods can be classified as the cross section method and the time series method. For the cross section method, only information on the yields of bonds with different maturities at a point in time is used in the estimation process. This generates a different set of parameters for each time period [11]. For the time series method, on the contrary, only the dynamic implications of the model is studied and the cross-sectional information is ignored. Obviously, both methods are ex parte. However, as two kinds of the class of Kalman filter methods, SRUKF and EKF can make full use of both the cross-sectional information and the time series data. Traditional estimation methods use either time series data or cross-sectional data while Kalman filters class methods can use both at the same time. What's more, in traditional research, short-term interest rate, the unobserved state variable, is often dealt with as a proxy variable, while under Kalman filters methods, short-term interest rate is taken as unobserved to do further estimation to enhance the reliability. The main advantage of the class of Kalman filter methods just stems from the fact that they allow the state variable to be unobserved magnitude.

The performances of the SRUKF and EKF are now examined for yielding state estimate and parameter estimate to the term structure of interest rates. For the sake of comparison, SRUKF and EKF are respectively applied to the system in 100 simulations.

During the simulation, we set $k = 0.1$, $\theta = 0.2$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\varphi = 0.16$ based on historical experience. In addition, the process noise covariance matrix is $Q = 0.001\mathrm{I}$ and the measurement noise covariance matrix is $R = 0.2\mathrm{I}$. The three parameters of SRUKF estimator are assumed to be $\alpha = 1e-1$ (usually $1e-2 \le \alpha \le 1$), $\beta = 2$ (for Gaussian distribution, $\beta = 2$ is optimal), $\kappa = 0$ (usually set to either 0 or 3-L).

The abilities of SRUKF-based estimator and EKF-based estimator to track the term structure are presented in Fig. 5 and Fig. 6 separately. The two figures show examples of the state estimates produced by SRUKF and EKF respectively compared to the truth.

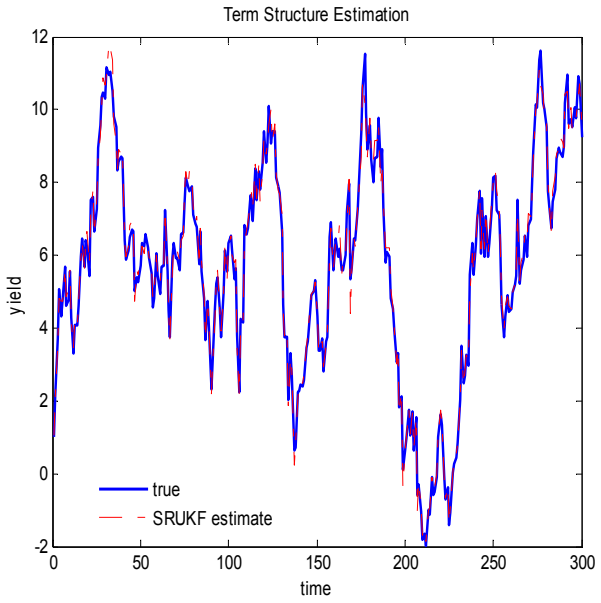

Fig.5 Estimation of Term Structure Using SRUKF

Fig. 5 shows an example of the state estimates produced by SRUKF compared to the truth.

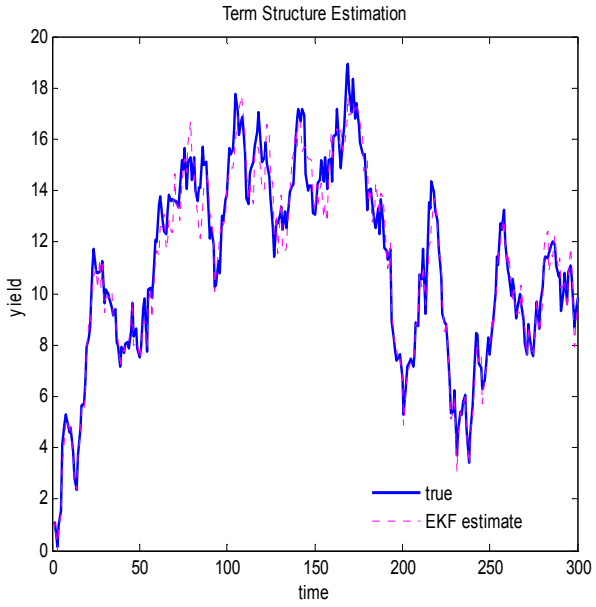

Fig.6 Estimation of Term Structure Using EKF

Fig. 6 shows an example of the state estimates produced by EKF compared to the truth. EKF produces state estimates that are a little far from the truth. SRUKF is the correct filter and captures the time history of the term structure very well. The displayed results indicate that EKF estimator demonstrates larger errors than SRUKF one does.

In order to further quantify the performance of the two estimators, the mean and variance of Root Mean Square Error (RMSE) have been generated. 100 samples (observation times) were simulated. 300 different realizations of the proposed EKF and SRUKF were generated. The comparison is in terms of RMSE, computed as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Forecasted\ yield\ -\ Realized\ yield\right)^2}$$

On one hand, we obtain from the simulation 300 forecasted interest rate term structures. On the other hand, we observe 300 realized interest rate term structures. We are thus able to measure the performance of the simulation by computing the RMSE.

The average results of RMSE are shown in Table 1:

algorithms	SRUKF	EKF
RMSE(mean)	0.63532	0.97559
RMSE(var)	0.38323	1.7298

Table 1 RMSE of 100 Times Simulation

As expected, the SRUKF-based term structure estimator presents a good performance. Thus the simulations demonstrate the superiority of the SRUKF-based term structure approaches over the EKF-based one.

Fig.7 presents the parameter estimation of theta using EKF and SRUKF respectively. As the figure indicates, SRUKF estimator is able to more accurately track the long-run average of the instantaneous interest rate. However EKF estimator exhibits poorer performance.

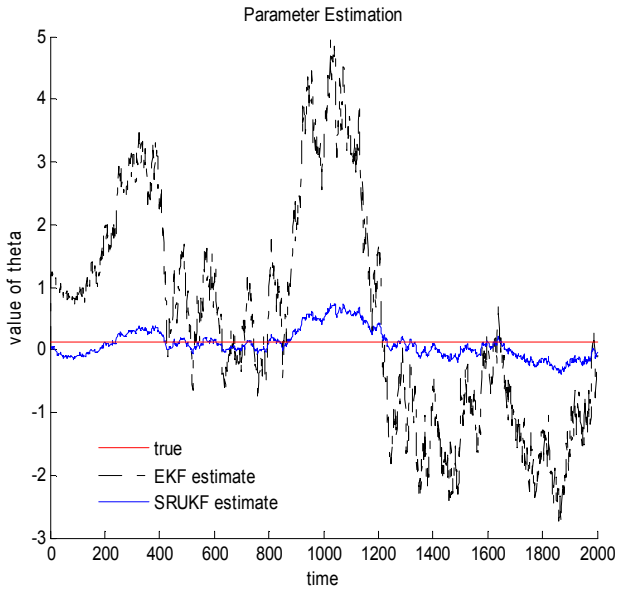

Fig.7 Parameter Estimation of Term Structure

The simulation results show that SRUKF estimator seems to have stronger noise power immunity and higher precision than EKF estimator does. Such better numerical performance is mainly because that SRUKF estimator can achieve third-order accuracy and EKF estimator approximates the linearity by first-order approximation through Taylor expansion series.

6 Conclusion

The Kalman filter is a well-known recursive algorithm that takes the stochastic state space model of the system together with measured outputs to achieve the optimal estimation of states in multi-input, multi-output systems. The filter takes system and measurement noises into account in the form of white noise. The optimality of the state estimation is achieved with the minimization of the mean estimation error [28]. In this study, EKF and SRUKF estimators are studied, which are two extended forms of Kalman filter that could be used for nonlinear systems.

SRUKF and EKF are two filters aiming at solving state and/or parameter estimation problems of nonlinear system. Compared with EKF estimator, SRUKF estimator is derivative-free and need not calculate Jacobian matrix associated with EKF algorithm and sometimes difficult to calculate in case of large and complicated systems. For SRUKF, it only needs relatively simple algebra calculation. This fact makes it much easier for SRUKF-based estimation. To linearize system equations is the most difficult part during the use of EKF and such approximation can only reach first-order accuracy in terms of Taylor series expansion. However, SRUKF does filtering directly according to the nonlinear system instead of linearization and so it can avoid the errors brought about from that linearization.

Simulation results tell us that SRUKF-based estimator provides higher precision because of the higher-order approximation than that of EKF. What's more, the square-root form embedded in SRUKF helps effectively reduce the covariance of important weights and guarantees positive definiteness of the underlying state covariance. SRUKF estimator exhibits improved accuracy and numerical performance relative to the EKF one and consequently can be a better alternative in case of nonlinear system estimation than the EKF estimator.

References:

[1] Robert C. Merton, Optimum Consumption and Portfolio Rules in a Continuous-Time Model, *Journal of Economic Theory,* November 1971: Vol. 3: 373-413.

[2] Oldrich Vasicek, An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics,* Vol. 5, Issue 2, November 1977, pp.177-188.

[3] John C. Cox, Jonathan E. Ingersoll, Jr. and Stephen A. Ross, An Intertemporal General Equilibrium Model of Asset Prices, *Econometrica,* Vol. 53, No. 2, March 1985, pp. 363-384.

[4] Francis A. Longstaff and Eduardo S. Schwartz, Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model, *The Journal of Finance,* Vol. 47, No. 4, September 1992, pp.1259-1282.

[5] D Duffie and R Kan, A Yield-factor Model of Interest Rates, *Mathematical Finance,* Oct. 1996: Vol. 6, No. 4: 379-406.

[6] Raymond Theoret, Pierre Rostan and Abdeljalil El Moussadek, Forecasting the Interest Rate Term Structure: Using the Model of Fong & Vasicek, the Extended Kalman Filter and the Bollinger Bands, working paper, August, 2008, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=671581.

[7] Langetieg Terence C., A Multivariate Model of the Term Structure. *Journal of Finance*, Vol. 35, Issue 1, March, 1980, pp.71-97.

[8] Michael J. Brennan and Eduardo S. Schwartz, A Continuous Time Approach to the Pricing of Bonds, *Journal of Banking and Finance*, Vol. 3, No. 2, 1979, pp. 133-155.

[9] Alois L.J.Geyer and Stefan Pichler, A State-Space Approach to Estimate and Test Multifactor Cox-Ingersoll-Ross Models of the Term Structure, *The Journal of Financial Research,* Spring 1999: Vol. XXII: No. 1:107-130.

[10] Ren-Raw Chen and Louis Scott, Multifactor Cox – Ingersoll - Ross Models of the Term Structure: Estimates and Tests from a Kalman Filter Model, *Journal of Real Estate Finance and Economics,* 2003: Vol. 27, No. 2: 143-172.

[11] Somnath Chatterjee, Application of the Kalman Filter for Estimating Continuous Time Term Structure Models-the Case of UK and Germany, Department of Economics, University of Glasgow, Jan. 2005.

[12] R.E. Kalman, A New Approach to Linear Filtering and Prediction Problems, *Transactions*

of the ASME, Series D, Journal of Basic Engineering, Vol. 82, 1960, pp.35-45.

[13] Gordon N, Salmond D. Novel Approach to Nonlinear and Non-Gaussian Bayesian State Estimation, *Proc of Institute Electric Engineering,* Vol. 140, No. 2, 1993, pp.107-113.

[14] Julier S J, Uhlmann J K, Durrant-Whyten H F, A New Approach for Filtering Nonlinear System [C], *Proc of the American Control Conference*, Washington: Seattle,1995.

[15] Merwe R Vander, Wan E A., The Square-Root Unscented Kalman Filter for State and Parameter estimation, *International Conference on Acoustics, Speech and Signal Processing*, Salt Lake City, UT, Vol.6, May 2001, pp.3461-3463.

[16] Krim Yazid, Rachid Ibtiouen, Omar Touhami and Maurice Fadel, Application of EKF to Parameters Estimation for Speed Sensorless and Neural Network Control of an Induction Motor, *Proceedings of the 6th WSEAS International Conference on Power Systems*, Lisbon, Portugal, September 22-24, 2006, pp.279-283.

[17] Guo Junhong, Liu Hongwei, Yang Xiaozong and Zuo Decheng, Application of Kalman Filter in Time Series Software Reliability Growth Model, *Proceedings of the 8th WSEAS International Conference on Applied Mathematics*, 2005, pp.184-188.

[18] M. I. Ribeiro, "Kalman and Extended Kalman Filters: Concept, Derivation and Properties", Institute for Systems and Robotics, Instituto Superior Tcnico, Feb. 2004: http:// citeseer.ist.psu.edu/ 650260. html.

[19] Vassilious C. Moussas, Structural Reliability Improvement Using Non-linear and Adaptive Multi-model Techniques, *Proceedings of the 10th WSEAS International Conference on Mathematical Methods, Computational Techniques and Intelligent Systems*, 2008, pp.307-311.

[20] Julier S J, Uhlmann J K, Unscented Filtering and Nonlinear Estimation, *Proceedings of the IEEE*, Vol. 92, No. 3, 2004, pp. 401-422.

[21] Jang-Sub Kim, Dong-Ryeol Shin, Erchin Serpedin, Adaptive Multi-user Receiver with Joint Channel and Time Delay Estimation of CDMA Signals Based on the Square-Root Unscented Filter, *Digital Signal Processing,* Vol. 19, Issue 3, May 2009, pp.504-520.

[22] Steven A. Holmes, Georg Klein and David W. Murray, An $O(N^2)$ Square Root Unscented Kalman Filter for Visual Simultaneous Localization and Mapping, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, No. 7, July 2009, pp.1251-1263.

[23] Sui Shu-lin, Yao Wen-long, Spacecraft Autonomous Optical Navigation Based on SR-UPF, *Journal of System Simulation*, Vol. 20, No. 18, September 2008, pp.4971- 4981.

[24] Peter Christoffrsen, Kris Jacobs, Lotfi Karoui and Karim Mimouni, Nonlinear Filtering in Affine Term Structure Models: Evidence from the Term Structure of Swap Rates. Date Posted: June 18, 2008, Last revised: August 01, 2009, http://papers.ssrn.com/sol3/papers.cfm?abstract _id=1147343.

[25] Paul Williams, Structural Damage Detection from Transient Responses Using Square-Root Unscented Filtering , *Acta Astronautica*, Volume 63, Issues 11-12, December 2008, pp.1259-1272.

[26] EA Wan, R Van Der Merwe, AT Nelson, Dual Estimation and the Unscented Transformation, *Advances in Neural Information Processing Systems*, Cambridge, MA: MIT Press, 2000, pp. 666-672.

[27] C Apostoaia, Z Szekely, D Gray, Feedback Signals Estimation of an Induction Machine Drive, *12th WSEAS International Conference on SYSTEMS*, Heraklion, Greece, July 22-24, 2008, pp.53-58.

[28] M Barut, S Bogosyan, M Gokasan, Sensorless Low/Zero Speed Control of Induction Motors with EKF Estimation, *Proceedings of the 6th WSEAS International Conference on Systems Theory & Scientific Computation*, Elounda, Greece, August 21-23, 2006, pp.127-132.