Application of the Kalman Filter in Fitting a Cox-Ingersoll-Ross Affine Term Structure Model

Kenrick So

1 Introduction

The term structure of interest rates relates the interest rates on debt instruments with respect to varying dates of maturity. For example, the term structure of interest rates can help explain why the interest rate of a zero-coupon bond is 3.5% for a bond maturing in one year and 5% for a bond maturing in five years. The interest rate of a zero-coupon bond is influenced by its length to maturity; typically, longer maturities attract higher yields due to the extended time horizon and the associated risks. Additionally, prevailing market conditions, such as inflation expectations and economic prospects, play a crucial role in determining these rates. In normal circumstances, investors expect to receive greater compensation for committing their funds for a longer period, which is reflected in the higher interest rate of the longer-maturity bond.

Recently, there has been interest in a subset of term structure models characterized by the affine relationship between future dynamics and state variables. These models are known as affine term structure models. Mathematically, affine refers to a linear plus constant relationship. The primary advantage of affine term structure models is that one can induce arbitrarily many state variables into the model and unique closed-form solutions to the corresponding partial differential equation continue to exist. Furthermore, the linear relationship between the instantaneous rate of interest and the underlying state variables reduces complexity in simulating paths for the term structure.

After specifying the representation of bond prices as affine functions of the underlying state variable, it is important to fit the parameters of the affine term structure model. Determining the appropriate parameter estimates for a term structure model is of utmost importance since a model with poorly specified parameters will not provide accurate forecasts. In this project, the Kalman Filter will be utilized to estimate parameters.

The Kalman Filter involves a set of observed system of equations called the measurement system and an unobserved system of equations called the transition system. Together, the Kalman filter uses these equations to make inferences about the unobserved values of the state variables from the transition system by conditioning on the observed rates from the measurement system. Lastly, the recursive inferences are used to construct and maximize a log-likelihood function to determine the optimal parameter values. This application of the Kalman filter is particularly relevant in the context of this project, as it allows the derivation of accurate parameter estimates for the affine term structure model.

This project sets out to model a single-factor affine Cox-Ingersoll-Ross (CIR) process. The CIR model is used in the valuation of interest rate derivatives since the model precludes negative interest rates if the Feller conditions are satisfied. The CIR model is a sub-class of square-root diffusion models that can be guaranteed to be non-negative given some condition and remain almost as tractable as the traditional Gaussian process.

2 Preliminaries

This section presents preliminary concepts to aid the reader in understanding this project. The setup of the problem is described, affine term structure models are introduced, and the Kalman filter procedure is discussed in this section.

2.1 Cox–Ingersoll–Ross Process

The CIR model relates the instantaneous interest rate $r(t)$ with a Feller square-root process with stochastic differential equation given by,

$$dr(t) = \alpha(\beta - r(t))dt + \sigma\sqrt{r(t)}dW_t \tag{1}$$

where W_t is a standard Wiener process, α is the mean-reversion speed parameter, β is the mean level of the process and σ is the volatility parameter. The CIR model avoids the possibility of negative interest rates if for all positive values of α and β, the following condition is satisfied:

$$2\alpha\beta \geq \sigma^2.$$

This is also known as the Feller condition. To understand how the CIR model avoids negative interest rates consider the following: suppose that the instantaneous interest rate $r(t)$ approaches zero, the standard deviation $\sigma\sqrt{r(t)}$ likewise becomes very small resulting in a dampened effect of the random shock to the rate. This result in the instantaneous interest rate being driven primarily by the drift term, which pushes to increase the instantaneous interest rate resulting in a model that precludes negative values.

2.2 Model Setup

A pure discount bond is a contract that pays one unit of currency at maturity. Assume that the pure discount bond price $P(t,T)$ is a function of the interest rate. The bond price function has the following form

$$P(t,T) = P(t,r,T). \tag{2}$$

Furthermore, given any pure discount bond price for any maturity, the spot rate of interest for that date denoted by $z(t,T)$ is the continuously compounded rate of return that generates a price of unity. This is given by

$$\begin{aligned}
P(t,T)e^{(T-t)z(t,T)} &= 1,\\
\ln\left(P(t,T)e^{(T-t)z(t,T)}\right) &= 0,\\
\ln\left(e^{(T-t)z(t,T)}\right) &= -\ln P(t,T),\\
z(t,T) &= -\frac{\ln P(t,T)}{T-t}.
\end{aligned} \tag{3}$$

Moreover, assume that $P(t,r,T)$ is a $C^{1,2}$ function of its arguments (i.e. once continuously differentiable in time and twice continuously differentiable with respect to the instantaneous interest rate). Applying Ito's lemma yields the following equation:

$$\begin{aligned}
dP(t,T,r(t)) &= P_t dt + P_r dr(t) + \frac{1}{2}(\sigma\sqrt{r(t)})^2 P_{rr} dt\\
&= P_t dt + P_r\left[\alpha(\beta - r(t))dt + \sigma\sqrt{r(t)}dW_t\right] + \frac{1}{2}\sigma^2 r(t) P_{rr} dt\\
&= \left[P_t + \alpha(\beta - r(t))P_r + \frac{\sigma^2 r(t)}{2}P_{rr}\right]dt + \sigma\sqrt{r(t)}P_r dW_t.
\end{aligned} \tag{4}$$

The goal is to construct a self-financing portfolio comprised of a contingent claim and the underlying asset. To achieve this feat, one must select portfolio weights to eliminate the underlying source of uncertainty driven by the Brownian motion process. This deterministic portfolio must earn the equivalent risk-free rate to preclude the existence of arbitrage opportunities in the market. Unlike the Black-Scholes option pricing formula, the instantaneous interest rate is not an asset traded in the marketplace. To over this, introduce two discount bonds with arbitrary maturities denoted by s_1 and s_2 and use these two discount bonds to construct a self-financing portfolio denoted by V. Define the portfolio weights of each bond in the portfolio as u_1 and u_2. The scaled return on the self-financing portfolio at each time increment is given by the weighted return of each of the two bonds in the self-financing portfolio. Mathematically, this is given by

$$\frac{dV(t)}{V(t)} = u_1 \frac{dP_1(t,s_1)}{P_1(t,s_1)} + u_2 \frac{dP_2(t,s_2)}{P_2(t,s_2)}. \tag{5}$$

Plugging in values from Equation (4) into Equation (5) yields,

$$\begin{aligned}
\frac{dV(t)}{V(t)} &= u_1 \frac{\left[P_{1,t} + \alpha(\beta - r(t))P_{1,r} + \frac{\sigma^2 r(t)}{2} P_{1,rr}\right] dt + \sigma\sqrt{r(t)} P_{1,r} dW_t}{P_1} \\
&\quad + u_2 \frac{\left[P_{2,t} + \alpha(\beta - r(t))P_{2,r} + \frac{\sigma^2 r(t)}{2} P_{2,rr}\right] dt + \sigma\sqrt{r(t)} P_{2,r} dW_t}{P_2} \\
&= u_1 \left[\frac{P_{1,t} + \alpha(\beta - r(t))P_{1,r} + \frac{\sigma^2 r(t)}{2} P_{1,rr}}{P_1} dt + \frac{\sigma\sqrt{r(t)} P_{1,r}}{P_1} dW_t\right] \\
&\quad + u_2 \left[\frac{P_{2,t} + \alpha(\beta - r(t))P_{2,r} + \frac{\sigma^2 r(t)}{2} P_{2,rr}}{P_2} dt + \frac{\sigma\sqrt{r(t)} P_{2,r}}{P_2} dW_t\right] \\
&= u_1(\mu_{s,1} dt + \sigma_{s,1} dW) + u_2(\mu_{s,2} dt + \sigma_{s,2} dW) \\
&= (u_1\mu_{s,1} + u_2\mu_{s,2})dt + (u_1\sigma_{s,1} + u_2\sigma_{s,2})dW.
\end{aligned} \tag{6}$$

Here,

$$\mu_{s,i} = \frac{P_{i,t} + \alpha(\beta - r(t))P_{i,r} + \frac{\sigma^2 r(t)}{2} P_{i,rr}}{P_i}, \quad i = 1,2 \tag{7}$$

$$\sigma_{s,i} = \frac{\sigma\sqrt{r(t)} P_{i,r}}{P_i} \quad i = 1,2. \tag{8}$$

The objective is to select weights u_1 and u_2 such that the Brownian motion term is removed hence the uncertainty is eliminated. This results in finding values for u_1 and u_2 such that the following system holds:

$$\begin{aligned}
u_1 + u_2 &= 1, \\
u_1\sigma_{s,1} + u_2\sigma_{s,2} &= 0.
\end{aligned}$$

A solution to this system is given by

$$u_1 = \frac{-\sigma_{s,2}}{\sigma_{s,1} - \sigma_{s,2}} \tag{9}$$

$$u_2 = \frac{\sigma_{s,1}}{\sigma_{s,1} - \sigma_{s,2}}. \tag{10}$$

Plugging in the solution obtained in Equation (9) into the equation for the dynamics of the self-financing portfolio Equation (6) results in:

$$\begin{aligned}
\frac{dV}{V} &= \left(\frac{-\sigma_{s,2}}{\sigma_{s,1} - \sigma_{s,2}} \mu_{s,1} + \frac{\sigma_{s,1}}{\sigma_{s,1} - \sigma_{s,2}} \mu_{s,2}\right) dt + \left(\frac{-\sigma_{s,2}}{\sigma_{s,1} - \sigma_{s,2}} \sigma_{s,1} + \frac{\sigma_{s,1}}{\sigma_{s,1} - \sigma_{s,2}} \sigma_{s,2}\right) \\
&= \left(\frac{-\sigma_{s,2}}{\sigma_{s,1} - \sigma_{s,2}} \mu_{s,1} + \frac{\sigma_{s,1}}{\sigma_{s,1} - \sigma_{s,2}} \mu_{s,2}\right) dt.
\end{aligned} \tag{11}$$

The self-financing portfolio V is by specification, riskless over the interval dt because we have eliminated the source of uncertainty driven by the Brownian motion. In order to avoid arbitrage opportunities, the self-financing portfolio V must earn the risk-free rate $r(t)$ over each incremental period of time dt. Otherwise, consider the situation where the portfolio earns more than the risk-free rate, then an investor could take a short position in the risk-free asset and use the profit to simultaneously take a long position in the self-financing portfolio hence an arbitrage opportunity exists. Thus to ensure no arbitrage opportunities exist, the rate of return of the portfolio $\frac{dV}{V}$ must equate to $r(t)dt$:

$$\begin{aligned}\frac{dV}{V} &= r(t)dt \\ \left(\frac{-\sigma_{s,2}}{\sigma_{s,1}-\sigma_{s,2}}\mu_{s,1} + \frac{\sigma_{s,1}}{\sigma_{s,1}-\sigma_{s,2}}\mu_{s,2}\right)dt &= r(t)dt \\ \frac{\sigma_{s,1}\mu_{s,2}-\sigma_{s,2}\mu_{s,1}}{\sigma_{s,1}-\sigma_{s,2}} &= r(t) \\ \sigma_{s,1}\mu_{s,2}-\sigma_{s,2}\mu_{s,1} &= r(t)\sigma_{s,1} - r(t)\sigma_{s,2} \\ \sigma_{s,1}(\mu_{s,2}-r(t)) &= \sigma_{s,2}(\mu_{s,1}-r(t)) \\ \frac{\mu_{s,2}-r(t)}{\sigma_{s,2}} &= \frac{\mu_{s,1}-r(t)}{\sigma_{s,1}}. \end{aligned} \tag{12}$$

Define the market price of risk λ as the standardized excess return over the risk-free rate. Mathematically, this is given by

$$\lambda(t) = \frac{\mu_t - r(t)}{\sigma_t}. \tag{13}$$

Returning to Equation (12), the equation states that the market price of risk must be equal for arbitrary maturities s_1 and s_2. As a result, the market price of risk must be constant across all maturities. In fact, one can derive a partial differential equation that describes the dynamics of any interest-rate contingent claim. A Feynman-Kac representation solution is given by:

$$\begin{aligned}\lambda(t) &= \frac{\mu_t - r(t)}{\sigma_t} \\ \frac{P_t + \alpha(\beta - r(t))P_r + \frac{\sigma^2 r(t)}{2}P_{rr}}{P} - r(t) &= \lambda(t)\frac{\sigma\sqrt{r(t)}P_r}{P} \\ P_t + \alpha(\beta - r(t))P_r + \frac{\sigma^2 r(t)}{2}P_{rr} - r(t)P &= \lambda(t)\sigma\sqrt{r(t)}P_r \\ P_t + \left[\alpha(\beta - r(t)) - \lambda(t)\sigma\sqrt{r(t)}\right]P_r + \frac{\sigma^2 r(t)}{2}P_{rr} - r(t)P &= 0 \end{aligned} \tag{14}$$

2.3 Affine Term Structure Model

Consider the case where the term structure $\{P(t,T), t \in [0,T]\}$ has the following form

$$P(t,T) = e^{A(t,T)-B(t,T)r(t)}. \tag{15}$$

where $A(t)$ and $B(t)$ are deterministic functions. Introduce the change of variables $\tau = T - t$ is the remaining time to maturity and define $A(\tau) \equiv A(t,T)$ and $B(\tau) \equiv B(t,T)$. Hence Equation (15) becomes,

$$P(\tau) = e^{A(\tau)-B(\tau)r(t)}. \tag{16}$$

Consider the first and second order partial derivatives of $P(\tau)$ given by

$$\begin{aligned} P_t &= [-A'(\tau) + B'(\tau)r]\, e^{A(\tau)-B(\tau)r} = [-A'(\tau) + B'(\tau)r]\, P(\tau) \\ P_r &= -B(\tau)e^{A(\tau)-B(\tau)r} = -B(\tau)P(\tau) \\ P_{rr} &= B^2(\tau)e^{A(\tau)-B(\tau)r} = B^2(\tau)P(\tau). \end{aligned} \tag{17}$$

Here,

$$\begin{aligned} A'(\tau) &= \frac{\partial A(\tau)}{\partial \tau}\underbrace{\frac{\partial \tau}{\partial t}}_{=-1} = -\frac{\partial A(\tau)}{\partial t}, \\ B'(\tau) &= \frac{\partial B(\tau)}{\partial \tau}\frac{\partial \tau}{\partial t} = -\frac{\partial B(\tau)}{\partial t}. \end{aligned} \tag{18}$$

Plugging the values from Equation (17) into Equation (14) results in

$$
\begin{aligned}
\left[-A'(\tau)+B'(\tau)r(t)\right]P(\tau)+\left[\alpha(\beta-r(t))-\lambda(t)\sigma\sqrt{r(t)}\right]\left[-B(\tau)P(\tau)\right]+\frac{\sigma^2 r(t)}{2}B^2(\tau)P(\tau)-r(t)P(\tau)&=0\\
-A'(\tau)P(\tau)+B'(\tau)r(t)P(\tau)+[\alpha\beta-\alpha r(t)-\lambda(t)\sigma\sqrt{r(t)}][-B(\tau)P(\tau)]+\frac{\sigma^2 r(t)}{2}B^2(\tau)P(\tau)-r(t)P(\tau)&=0\\
-A'(\tau)+B'(\tau)r(t)-\alpha\beta B(\tau)+\alpha r(t)B(\tau)+\lambda(t)\sigma\sqrt{r(t)}B(\tau)+\frac{\sigma^2 r(t)}{2}B^2(\tau)-r(t)&=0\\
-A'(\tau)+(B'(\tau)-1)r(t)+\alpha B(\tau)r(t)-\alpha\beta B(\tau)+\lambda(t)\sigma\sqrt{r(t)}B(\tau)+\frac{\sigma^2 r(t)}{2}B^2(\tau)&=0\\
-A'(\tau)+(B'(\tau)-1)r(t)+(\alpha r(t)-\alpha\beta+\lambda(t)\sigma\sqrt{r(t)})B(\tau)+\frac{\sigma^2 r(t)}{2}B^2(\tau)&=0.
\end{aligned}
\tag{19}
$$

It has been shown in Chen and Scott (2003); Duan and Simonato (1999); Geyer and Pichler (1999) that the following structure for $A(\tau)$ and $B(\tau)$ is a solution to Equation (19),

$$
\begin{aligned}
B(\tau)&=\frac{2\left(e^{\gamma\tau}-1\right)}{(\gamma+\alpha+\lambda)\left(e^{\gamma\tau}-1\right)+2\gamma}\\
A(\tau)&=\frac{2\alpha\beta}{\sigma^2}\ln\left(\frac{2\gamma e^{\frac{(\gamma+\alpha+\lambda)\tau}{2}}}{(\gamma+\alpha+\lambda)\left(e^{\gamma\tau}-1\right)+2\gamma}\right),
\end{aligned}
$$

where

$$\gamma=\sqrt{(\alpha+\lambda)^2+2\sigma^2}. \tag{20}$$

From Equation (3), the following relationship between the zero-coupon yield and the price of a zero-coupon bond is obtained

$$z(t,T)=-\frac{\ln P(t,T)}{T-t}=\frac{-A(\tau)+B(\tau)r(t)}{\tau}. \tag{21}$$

2.4 Kalman Filter

The Kalman filter is a technique in control engineering to solve a filtering problem. A filtering problem involves an observable stream of data that is often cluttered with noise, and the objective is to effectively differentiate between the valuable signal and the irrelevant noise. For example, consider a scenario where you're trying to track the position of a moving object, such as a car, using GPS data that may be affected by interference from buildings or weather conditions. The goal is to extract the accurate location of the car from the disrupted data. The Kalman filter is a recursive technique that begins this process with an initial inference, typically the unconditional mean and variance of the state variables—quantities that describe the internal state of the system being observed and evolve over time. From this baseline, the filter updates and refines its predictions based on incoming observed values, continuously evaluating the conditional mean and variance against the initial inference. This iterative updating allows for the isolation of the signal from the noise, enabling a clearer understanding of the system's behavior over time. The Kalman filter proceeds to infer values of the measurement equation from the initial inference of the state variables. This process involves evaluating the conditional mean and conditional variance given the initial inference for the state variables. Next, as a natural progression from the initial inference, the prediction from the measurement system leads to the acquisition of an observed value. This observed value is crucial, as it feeds back into the process, updating our inference of the current value of the transition system. By doing so, we maintain the continuity inherent in this recursive process, which is consistently repeated over the period of interest.

To implement the Kalman filter, a discrete-time state space representation must be made. The time interval $[0,T]$ is discretized into N subintervals where $t_i=i\frac{T}{N}$ for $i=1,2,\ldots,n$. Denote each time step as $\Delta t = t_i - t_{i-1}$. The measurement system can be represented in state-space form as,

$$\begin{bmatrix} z(t_i, t_{z_1}) \\ z(t_i, t_{z_2}) \\ \vdots \\ z(t_i, t_{z_n}) \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{-A(t_i,t_{z_1})}{t_{z_1}-t_i} \\ \frac{-A(t_i,t_{z_2})}{t_{z_2}-t_i} \\ \vdots \\ \frac{-A(t_i,t_{z_n})}{t_{z_n}-t_i} \end{bmatrix}}_{A} + \underbrace{\begin{bmatrix} \frac{B(t_i,t_{z_1})}{t_{z_1}-t_i} \\ \frac{B(t_i,t_{z_2})}{t_{z_2}-t_i} \\ \vdots \\ \frac{B(t_i,t_{z_n})}{t_{z_n}-t_i} \end{bmatrix}}_{H} \begin{bmatrix} r(t_i) \end{bmatrix} + \begin{bmatrix} \nu_{1,t_i} \\ \nu_{2,t_i} \\ \vdots \\ \nu_{n,t_i} \end{bmatrix}. \tag{22}$$

Alternatively,

$$z_{t_i} = A + H r_{t_i} + \nu_{t_i}, \tag{23}$$

where,

$$\nu_t \sim \mathcal{N}(0, R),$$
$$R = \begin{bmatrix} [r_1]^2 & 0 & \cdots & 0 \\ 0 & [r_2]^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & [r_n]^2 \end{bmatrix}.$$

Likewise, the transition system derived from solving the CIR stochastic differential equation analytically can be written as follows:

$$r_{t_i} = \underbrace{\beta(1 - e^{-\alpha\Delta t})}_{C} + \underbrace{e^{-\alpha\Delta t}}_{F} r_{t_{i-1}} + \epsilon_{t_i} \tag{24}$$

$$= C + F r_{t_{i-1}} + \epsilon_{t_i}. \tag{25}$$

Here,

$$\varepsilon_{t_i} \mid \mathcal{F}_{t_{i-1}} \sim \mathcal{N}(0, Q_{t_i}),$$
$$Q_{t_i} = \begin{bmatrix} \xi_1 & 0 & 0 \\ 0 & \xi_2 & 0 \\ 0 & 0 & \xi_3 \end{bmatrix},$$
$$\xi = \frac{\beta\sigma^2}{2\alpha}\left(1 - e^{-\alpha\Delta t}\right)^2 + \frac{\sigma^2}{\alpha}\left(e^{-\alpha\Delta t} - e^{-2\alpha\Delta t}\right) r_{t_{i-1}}.$$

An outline of the Kalman filter process is detailed below. Define the filtration generated by the measurement system $\mathcal{F}_{t_i}$ as,

$$\mathcal{F}_{t_i} = \sigma\{z_0, z_1, \ldots, z_i\}, \quad t_i = i\frac{T}{N} \text{ over } [0, T]$$

First, the initial values for the recursion must be initialized. The unconditional mean and unconditional variance of the transition system are used. For the CIR model, the unconditional mean is given by

$$\mathbb{E}[r_1] = \mathbb{E}[r_1 \mid \mathcal{F}_0] = \beta.$$

Likewise, the unconditional variance is given by

$$\text{var}[r_1] = \text{var}[r_1 \mid \mathcal{F}_0] = \frac{\sigma^2\beta}{2\alpha}.$$

Second, the measurement equation is forecasted. The conditional forecast of the measurement equation is given by

$$\mathbb{E}\left[z_{t_i} \mid \mathcal{F}_{t_{i-1}}\right] = A + H\mathbb{E}\left[r_{t_i} \mid \mathcal{F}_{t_{i-1}}\right].$$

The unconditional variance is given by

$$\text{var}\left[z_{t_i} \mid \mathcal{F}_{t_{i-1}}\right] = H \,\text{var}\left[r_{t_i} \mid \mathcal{F}_{t_{i-1}}\right] H^T + R$$

Third, the error in the conditional prediction is given by

$$\zeta_{t_i} = z_{t_i} - \mathbb{E}[z_{t_i|\mathcal{F}_{t_{i-1}}}].$$

This prediction error is then used to update the inferences about the unobserved transition system. Mathematically,

$$\mathbb{E}\left[r_{t_i} \mid \mathcal{F}_{t_i}\right] = \mathbb{E}\left[r_{t_i} \mid \mathcal{F}_{t_{i-1}}\right] + K_{t_i}\zeta_{t_i},$$

where

$$K_{t_i} = \operatorname{var}\left[r_{t_i} \mid \mathcal{F}_{t_{i-1}}\right] H^T \operatorname{var}\left[z_{t_i} \mid \mathcal{F}_{t_{i-1}}\right]^{-1},$$

is the Kalman gain matrix, which determines the weight given to new observations in the updated forecast. The conditional expectation and conditional variance is also updated,

$$\operatorname{var}\left[r_{t_i} \mid \mathcal{F}_{t_i}\right] = \left(I - K_{t_i}H\right) \operatorname{var}\left[r_{t_i} \mid \mathcal{F}_{t_{i-1}}\right].$$

The fourth step is to forecast the unknown values of the state system in the next time period conditioning on the updated values in the previous period. The conditional expectation and condition variance are given by

$$\begin{aligned}
\mathbb{E}\left[r_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] &= C + F\mathbb{E}\left[r_{t_i} \mid \mathcal{F}_{t_i}\right] \\
\operatorname{var}\left[r_{t_{i+1}} \mid \mathcal{F}_{t_i}\right] &= \operatorname{var}\left[r_{t_i} \mid \mathcal{F}_{t_{i-1}}\right] - F \operatorname{var}\left[r_{t_i} \mid \mathcal{F}_{t_i}\right] F^T + Q.
\end{aligned}$$

Lastly, after the values in the previous steps are obtained, the log-likelihood function is constructed with the following form

$$\begin{aligned}
\ell(\theta) &= \sum_{i=1}^{N} \ln\left[(2\pi)^{-\frac{n}{2}} \det\left(\operatorname{var}\left[r_{t_i} \mid \mathcal{F}_{t_{i-1}}\right]\right)^{-\frac{1}{2}} e^{-\frac{1}{2}\zeta_{t_i}^T \operatorname{var}\left[r_{t_i} \mid \mathcal{F}_{t_{i-1}}\right]^{-1}\zeta_{t_i}}\right], \\
&= -\frac{nN\ln(2\pi)}{2} - \frac{1}{2}\sum_{i=1}^{N}\left[\ln\left(\det\left(\operatorname{var}\left[r_{t_i} \mid \mathcal{F}_{t_{i-1}}\right]\right)\right) + \zeta_{t_i}^T \operatorname{var}\left[r_{t_i} \mid \mathcal{F}_{t_{i-1}}\right]^{-1}\zeta_{t_i}\right].
\end{aligned}$$

3 Numerical Experiment

This section implements the Kalman filter technique to estimate parameter values for a CIR model.

To demonstrate the methodology, start by simulating the interest rate process for various maturities. Define an initial set of parameter values then use Equation (21) to generate values of for the spot rate of interest at each time step.

We consider zero-coupon bonds with four different time-to-maturities: $\tau = 3/12$, $6/12$, 2, and 5, representing 3 months, 6 months, 2 years, and 5 years, respectively. Figure (1) presents a plot of simulated paths over a period of 120 months (10 years) for the CIR model across the given time-to-maturities.

Next, the variables from the Kalman filter section are defined, and a function that calculates the log-likelihood using the Kalman filter algorithm is obtained. For each time-to-maturity, the simulation process is repeated 300 times, applying the Kalman filter for parameter estimation. A Newton-type gradient method is employed to minimize the log-likelihood function, which is crucial for parameter estimation. By minimizing this function, we enhance the model's performance, ensuring that the estimated parameters best represent the observed data. This process leads to more reliable predictions and improved accuracy in capturing the underlying dynamics of the interest rate process.

Afterward, the mean and standard deviation of the estimated parameters are returned. A summary of the mean and standard deviation after applying the Kalman filter technique on 300 instances of simulated CIR model paths is provided in Table (1). The Kalman filter estimates for the mean reversion level parameter β and the volatility parameter σ are reasonably close to the actual values, with minimal standard error. In contrast, the estimates for the mean reversion speed parameter α and the market price of risk parameter λ are less accurate, with larger associated standard errors. This discrepancy is consistent with the findings of Bolder (2001), who also reported large standard errors for estimating the parameters of a CIR model using the Kalman filter. Bolder (2001) attribute the poor estimation to the joint presence of the mean reversion speed and market price of risk parameters in both $A(\tau)$ and $B(\tau)$. This situation creates identifiability issues, meaning that the model struggles to distinguish between these two parameters because they influence the outcomes in similar ways, complicating the process of minimizing the log-likelihood. As a result, the estimation procedure becomes less reliable, leading to broader confidence intervals for the estimated parameters.

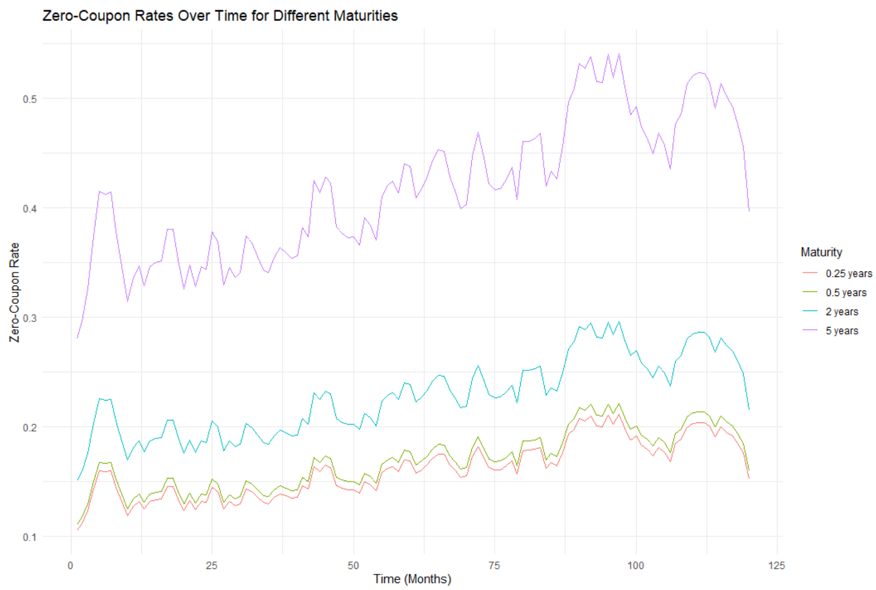

Figure 1: Plot of Sample Zero-Coupon Rates over Time for 3-months, 6-months, 2-years, and 5-years time-to-maturity respectively from bottom to top

Parameters	Actual Value	Mean Estimate	Standard Deviation Estimate	95% Confidence Interval
α	0.05	0.7530	1.5999	0.05 $\in$ (-2.3828, 3.8888)
β	0.10	0.1282	0.5002	0.10 $\in$ (-0.8522, 1.1086)
σ	0.075	0.0448	0.1138	0.075 $\in$ (-0.1783, 0.2678)
λ	-0.40	-0.1065	1.4477	-0.40 $\in$ (-2.9440, 2.7310)

Table 1: Table of Mean and Standard Deviation Estimate for Numerical Experiment along with the 95% Confidence Interval

References

Bolder, D. J. (2001). Affine term-structure models: Theory and implementation. *SSRN Electronic Journal*.

Chen, R.-R. and Scott, L. (2003). Multi-factor cox-ingersoll-ross models of the term structure: Estimates and tests from a kalman filter model. *The Journal of Real Estate Finance and Economics*, 27(2):143–172.

Duan, J.-C. and Simonato, J.-G. (1999). Estimating and testing exponential-affine term structure models by kalman filter. *Review of Quantitative Finance and Accounting*, 13(2):111–135.

Geyer, A. L. J. and Pichler, S. (1999). A state-space approach to estimate and test multifactor cox-ingersoll-ross models of the term structure. *Journal of Financial Research*, 22(1):107–130.