A State Space Framework for Models with Uncertainty

Willie Mohlongo (willie@aims.ac.za)
African Institute for Mathematical Sciences (AIMS)

Supervised by Dr. Dennis Ikpe
Co-Supervised by Prof. Ronnie Becker

24 October 2019

Submitted in partial fulfillment of a structured masters degree at AIMS South Africa

Abstract

This essay is aimed at formulating a filtering model for a bond market with particular attention on long term maturities and uncertainty in models parameters. In developing economies countries such as South Africa, it is challenging to calibrate bond market models correctly because prices of very long maturity bonds are not entirely available. Our filtering framework exploits linear state-space representation and Kalman filter to estimate the current term structure of the underlying price process for the long-term bond market. The measurement system is viewed as partially observed price to maturity process, in the form of a state variable with a white noise proces. On the other hand, the transition process is the Ornstein-Uhlenbeck with uncertain parameters.

Declaration

I, the undersigned, hereby declare that the work contained in this research project is my original work, and that any work done by others or by myself previously has been acknowledged and referenced accordingly.

Willie Mohlongo, 24 October 2019

Contents

1. Introduction

1.1 Background Study

The use of state-space representation and Kalman filtering involving time series and cross-section data have gained popularity in the formulation of the bond market based on the work of Duan and Simonato (1999), Lund (1997), Geyer and Frühwirth&Schnatter (1996), De Jong, 1999, and Babbs and Nowman (1999). The major setback with the classic Kalman filter theory in the state-space bond market is that it assumes the previous knowledge regarding the underlying interest rate model Kalman (1960). It also believes that long-term bond yields are entirely observable. In reality, complexity within the systems makes it typically impractical to calibrate bond market models correctly since the prices of long-term bonds are not entirely observable. Thus, there is enormous ambiguity about the system model, and the error-free model assumption of classical Kalman filtering is no longer realistic in applications Bulut and Bayat (2012).

In reality, we wish to construct a framework that captures uncertainty in bond market models. This approach is incredibly helpful, all the time when the underlying state variables are not observable. Thus it is practical to optimally estimate the parameter set when the underlying factors are unobservable, and this relies on the relationship within underlying state variables and the bond price. The classical Kalman filtering approach can be changed to accommodate uncertainty. The formulation of state-space representation combined with Kalman filter can optimally estimate the existing long-term bond term structure model. We begin with the relationship between bond prices and state variables whereby the observed system is the measurement equation. The unobserved system is the transition equation, which describes the dynamics of state variables, which may be seen as Ornstein-Ulhenbeck process Bolder (2001). The measurement equation together with the transition equation, represents a state-space form model.

The Kalman filter applies the state-space model to recursively produces conclusions relating to the unobserved values of the state variables acquisition on the observed market of bond prices. The recursive inferences are used to construct the maximum likelihood function to estimate parameter set optimally. Kalman filter can be seen as state-space subjected to ambiguity within the observation process. The investigation of the Ornstein Uhlenbeck model in Kalman filter application is due to the most recent work by Mastro (2013) whereby unobserved state variables are optimally estimated using Kalman filter equations.

1.2 Outline of the essay

This essay is broken down into four chapters, chapter 1 introduce the study, by giving the background of the study, problem statement. Chapter 2 provides a brief discussion on the interest rates product, definitions, state-space formulation framework with a demonstration example and Kalman filter equations. with the simulation example of Chapter 3 presents the Ornstein-Ulhenbeck process applies the state-space representation of the model, derives the Kalman filter equations of the model and thereafter demonstrate the simulation of the model. Chapter 4 gives the conclusion, recommendations, and further work.

2. Preliminary Background

A state-space form is a system given by two equations, measurement equation and transition equation. Kalman filter can only be put to use once our model form is in the state space form. The sections below explain the formulation of state-space representation, whereas the latter parts expand on the Kalman filter.

2.1 Interest Rate Products

In the section, we describe the specific underlying interest rate products for this specify essay.

Debt Instruments is a mechanism entity used to utilize escalation in the capital. It is a mandatory agreement that gives funds to an entity in return for a promise from the entity to pay back the lender following the contract terms. The contracts of the instrument incorporate a full detailed requirement of the deal protection involved, interest rate, the time frame to the due date if applicable and the duration of interest payments.

Bonds is known as fixed-payment debt instrument. It is debt security under the agreement between the creditor and lender that the creditor is forced to pay interest coupons (depending on terms of the bond) and the principal amount at bond expiry date called maturity date. The bond interest is repaid at firm intervals half-yearly, yearly, or monthly in some cases. One of the advantages of bonds is the way they payoff. The income is received through the interest payments, and at maturity, the principal amount is paid back. This makes bonds to be safer.

Coupon Bond is a financial obligation type of bond with attached payments (coupons) to it, and pays periodic (semi-annual) interest coupons. The bond comes up with a fixed interest rate called coupon rate, which is also referred to as bond yield on the date of issuance. The bond holder receives these payments during the duration between the delivery and the expiry date of the bond.

Zero Coupon Bonds is a discount bond debt instrument that does not pay interest to the lender. The bond sells at a deep discount less than its par value, rendering profit to the lender when the bond regains its nominal value at the expiration date. One of the advantages of such a bond is that it inherent pure profit, and the holder of the bond is sure to get beyond what they have invested at maturity furthermore to the flat rate of interest.

Term Structure of Interest Rates is the association between the yield on bonds of the similar credit trait Nevertheless, the dissimilar expiration date is usually referred to as the term structure of interest rates. The plotting depiction is generally referred to as the yield curve (Focardi and Fabozzi, 2004). The curve usually demonstrates the association between time to the due date of bonds and yield to the expiration date of bonds, and this relationship has a significant role in the financial situation. The term structure often takes into consideration of market players about the forthcoming comportment of the interest rate. In general, usually, longer expiration date bonds tend to have a higher yield than short end date bonds, but in some instances, the yield curve is reversed vice-versa. The yield curve changes depending on the bid and offer for short-term, medium-term, and long-term bonds varies, to some extent, independently (Ibhagui, 2010).

2.2 State Space Representation

The general state-space form appears in a multi-variable setting. It comprises of output variable and the state variable. Let $\mathbf{y}_t$ denote the output vector with N components and is expressed as $\mathbf{y}_t = (\mathbf{y}_{1t}, \mathbf{y}_{2t}, \ldots, \mathbf{y}_{Nt})$. The time series model can be expressed through the state-space model comprising of two basic equations, which are measurement and transition equations, respectively. The observable variables of $\mathbf{y}_t$ are related to the state variable α_t through measurement equation. In this essay, we will only consider the case where our model is linear, which implies the observation variables in the two system equations are the linear function of α_t, which is governed by the Markov process.

2.2.1 Measurement Equation. Specifies the observed variables and defines the relationship between the time series observed, $\mathbf{y}_t$ and the unobserved state variables α_t. It is given by

$$\mathbf{y}_t = Z_t\alpha_t + d_t + \epsilon_t \quad t = 1, \ldots, T \tag{2.2.1}$$

where $\mathbf{y}_t$ is an $N \times 1$ actual measurement vector of α_t at time t containing N components called observation of the system, the variable α_t is an $m \times 1$ vector called the state vector, Z_t is an $N \times m$ noiseless matrix connection between state variable and the measurement vector, d_t is an $N \times 1$ vector and lastly ϵ_t is an $N \times 1$ vector associated with measurement error called measurement noise. The measurement noise is additive and normally distributed with mean zero and known covariance H_t, given by $H_t = E[\epsilon_t\epsilon_t']$.

2.2.2 Transition Equation. Is the first order stochastic process describing the evolution of state variables from previous state α_{t-1} to the current state α_t, it is given by

$$\alpha_t = T_t\alpha_{t-1} + c_t + R_t\eta_t \quad t = 1, \ldots, T \tag{2.2.2}$$

where α_t is an $m \times 1$ state vector process at time t, T_t is $m \times m$ state transition matrix from the state at $t-1$ to the state at t, c_t is an $m \times 1$ control input vector, R_t is an $m \times g$ matrix and lastly η_t is a $g \times 1$ vector called process noise. The noise of the process is additive and normally distributed with mean zero and covariance Q_t , given by $Q_t = E[\eta_t\eta_t']$. If α_t is a scaler then Q_t is the variance for the process noise.

2.2.3 The State Space Model Assumptions . The additional assumptions complete the specification of the model.

a) we assume that α_0 is the initial state vector of the process with expectation $\mathbf{a}_0$ and covariance $\mathbf{P}_0$, denoted by

$$E(\alpha_0) = \mathbf{a}_0 \quad \text{and} \quad Var(\alpha_0) = \mathbf{P}_0 \tag{2.2.3}$$

b) the process noise η_t, the initial state and the measurement noise ϵ_t are uncorrelated in all periods of time, therefore

$$E(\epsilon_t\eta_v') = 0 \quad \forall t, v = 1, \ldots, T \tag{2.2.4}$$

and

$$E(\epsilon_t\alpha_t') = 0, \quad E(\eta_t\alpha_t') = 0 \quad t = 1, \ldots, T \tag{2.2.5}$$

The matrices $Z_t, d_t, H_t, T_t, c_t, R_t$ and Q_t in equation (2.2.1) and equation (2.2.2) are called the system matrices. They change with time in a predetermined way. If the system matrices don't vary over time, then the model is said to be invariant with time, which can be view as a special case model, see (Harvey, 1990).

2.2.4 State Space Representation Example . AR(2) is an example of state space representation given by,

$$\begin{aligned} \mathbf{y}_t &= \mu + c_t \\ c_t &= \phi_1 c_{t-1} + \phi_2 c_{t-2} + \eta_t \end{aligned}$$

we want to transform our AR(2) model to state space representation in the form of equation (2.2.1) and equation (2.2.2). We do this by defining a state variable α_t, which is given by

$$\alpha_t = (c_t, c_{t-1})^T \tag{2.2.6}$$

therefore, our transition equation is given by

$$\begin{pmatrix} c_t \\ c_{t-1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c_{t-1} \\ c_{t-2} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \eta_t$$

where

$$\begin{aligned} \alpha_t &= (c_t, c_{t-1})^T \\ \alpha_{t-1} &= (c_{t-1}, c_{t-2})^T \end{aligned}$$

now, we match our AR(2) model measurement equation to equation (2.2.1), therefore we can write $\mathbf{y}_t$ in terms of α_t

$$\begin{aligned} Z_t &= (1, 0) \\ d_t &= \mu \\ \epsilon_t &= 0 \\ H_t &= 0 \end{aligned}$$

where

$$\mu = \frac{\alpha_t}{1 - (\phi_1 + \phi_2)}$$

thus, we see that μ appears in the state space representation through α_t.

The initial state variables has mean zero see (Zivot, 2006). The unknown parameters, ϕ_1 and ϕ_2, are enclosed in a system matrix. Our goal is to optimally estimate these parameter, and the Kalman filter is a good technique to use when estimating unknown parameter set.

2.3 The Kalman Filter

In this section, after the brief given in the introduction, we give a detailed demonstration of the Kalman filter. Now that our model is in state-space representation discussed in detail in the preceeding section, we can apply the Kalman filter algorithm in our model. The method was developed by Rudolph Kalman 1960. It is a series of mathematical equations that use state-space representation inferences recursively to optimally estimating the unobservable. The algorithm starts by giving the initial values for the state variables with the logical guess, and it thus measures the ambiguity of its initial guesses, which is the mean and variance in our case. The measurement system predicts the observed value, which are then used to update our observed inferences via the transition equation. The updated values are exploited to estimate the next state variables value. If we continue in this sense in our next iteration and so forth, we eventually construct a time series of unobserved state variables. This approach provides the optimal estimator by reducing the covariance of the estimation error of the state vector, given that the noise process and initial state variables are normally distributed. When we drop the assumption of normality, there is no assurance that the Kalman filter provides the conditional expectation of the state vector. Nevertheless, this provides the optimal estimator of the state variables and reduces the covariance estimation error (Harvey, 1990).

The two parameters, state estimate and error covariance matrix represent the state of the filter as a measure of the expected state estimate accuracy.The Kalman filter can be written as a single formula, but it is conceived most often as two distinct phases: Predict and Update. The predict process uses the previous time-state estimate to generate an estimation of the current time-state. In the update process, to refine the state estimate, the current forecast is combined with current observation information.

2.3.1 Derivation of Kalman Filter Equations. From the preceeding section of state space formulation consider equation (2.2.1) and equation (2.2.2). Our aim is to estimate the state α_t on the basis of our knowledge of system dynamics and the availability of noisy measurements. The amount of information available to us for our estimation of the state varies depending on the specific problem we are trying to solve. If we have all the information available in our calculation up to $\mathbf{y}_{t-1}$, but not including $\mathbf{y}_t$ available to use at time step t in our estimate of α_t, we call our estimate *prior estimate*. If all the information is available up to and including $\mathbf{y}_t$ at time step t in our estimate of α_t, we have call the estimate *posteriori estimate* see (Simon, 2006).

We start the estimation process by letting $\hat{\mathbf{a}}_{t-1|t-1}$ denote the state estimate of α_{t-1}. The notation $\hat{\mathbf{a}}_{t-1|t-1}$ represents the estimate of α_{t-1} at time index $t-1$, based on all information available up to and including time index $t-1$, thus we can form our *posteriori estimate*. This estimate is given by

$$\hat{\mathbf{a}}_{t-1|t-1} = E[\alpha_{t-1} \mid \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}] \tag{2.3.1}$$

Based on our knowledge of process dynamics and the availability of observed measurements, our goal is to estimate the state α_t. Now given $\hat{\mathbf{a}}_{t-1|t-1}$, the *priori estimate* of α_t for time t step is computed by taking the expected value of α_t conditioned on all information available up to $\mathbf{y}_{t-1}$ at time t but not including $\mathbf{y}_t$, therefore we obtain

$$\begin{aligned}
\hat{\mathbf{a}}_{t|t-1} &= E[\alpha_t \mid \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}] \\
\hat{\mathbf{a}}_{t|t-1} &= E[T_t\alpha_{t-1} + c_t + R_t\eta_t \mid \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}] \\
\hat{\mathbf{a}}_{t|t-1} &= T_t E[\alpha_{t-1} \mid \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}] + E[c_t \mid \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}] + R_t E[\eta_t \mid \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}]
\end{aligned}$$

the process noise is normally distributed with mean zero,therefore

$$E[\eta_t] = 0 \tag{2.3.2}$$

therefore, using the definition of (2.3.1) and (2.3.2), we get

$$\hat{\mathbf{a}}_{t|t-1} = T_t\hat{\mathbf{a}}_{t-1|t-1} + c_t \tag{2.3.3}$$

Intuitively, this makes sense. We don't have any measurement available to help us update our calculation between *posteriori estimate* at time step $(t-1)$ to *priori estimate* at time t, so we should only update the estimate based on our knowledge of the dynamics of the system see (Simon, 2006).

The predicted measurement at time step t given all the information available up to $\mathbf{y}_{t-1}$ but not including $\mathbf{y}_t$ is given by simply substituting state estimate in equation (2.2.1) see (Mastro, 2013), therefore we have

$$\begin{aligned}
\hat{\mathbf{y}}_{t|t-1} &= E[\mathbf{y}_t \mid \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}] \\
\hat{\mathbf{y}}_{t|t-1} &= E[Z_t\alpha_t + d_t + \epsilon_t \mid \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}] \\
\hat{\mathbf{y}}_{t|t-1} &= Z_tE[\alpha_t \mid \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}] + E[d_t \mid \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}] + E[\epsilon_t \mid \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1}]
\end{aligned}$$

the measrement noise is normally distributed with mean,

$$E[\epsilon_t] = 0 \tag{2.3.4}$$

thus, using definition of (2.3.3) and (2.3.4), we have

$$\hat{\mathbf{y}}_{t|t-1} = Z_t\hat{\mathbf{a}}_{t|t-1} + d_t \tag{2.3.5}$$

Let $\mathbf{P}_{t-1|t-1}$ denotes the error covariance matrix of α_{t-1} given the observations up to and including $\mathbf{y}_{t-1}$. The covariance error matrix $\mathbf{P}_{t-1|t-1}$ is equivalent to the mean squared error, and it is given by the covariance difference between the actual state α_{t-1} at time step $t-1$ and the estimated state $\hat{\mathbf{a}}_{t-1|t-1}$. This error is denoted by,

$$\mathbf{P}_{t-1|t-1} = E\Big[(\alpha_{t-1} - \hat{\mathbf{a}}_{t-1|t-1})(\alpha_{t-1} - \hat{\mathbf{a}}_{t-1|t-1})'\Big] \tag{2.3.6}$$

The *priori* error covariance matrix at time step t denoted by $\mathbf{P}_{t|t-1}$. The notation $\mathbf{P}_{t|t-1}$ represents the error covariance estimate of α_t at time index t, based on all information available up to and including time index $t-1$. The error covariance is equivalent to the mean squared error, and it is given by the covariance difference between the actual state α_t at time step t and the estimated state $\hat{\mathbf{a}}_{t|t-1}$. This error is given by,

$$\begin{aligned}
e_{t|t-1} &= \alpha_t - \hat{\mathbf{a}}_{t|t-1} \\
e_{t|t-1} &= T_t\alpha_{t-1} + c_t + R_t\eta_t - [T_t\hat{\mathbf{a}}_{t-1|t-1} + c_t]
\end{aligned}$$

collecting simalar terms gives,

$$e_{t|t-1} = T_t(\alpha_{t-1} - \hat{\mathbf{a}}_{t-1|t-1}) + R_t\eta_t \tag{2.3.7}$$

therefore, the error covariance is given by

$$\begin{aligned}
\mathbf{P}_{t|t-1} &= cov(\alpha_t - \hat{\mathbf{a}}_{t|t-1}) \\
\mathbf{P}_{t|t-1} &= E[e_{t|t-1}e'_{t|t-1}] \\
\mathbf{P}_{t|t-1} &= E\Big[(T_t(\alpha_{t-1} - \hat{\mathbf{a}}_{t-1|t-1}) + R_t\eta_t)((\alpha_{t-1} - \hat{\mathbf{a}}_{t-1|t-1})T_t + R_t\eta_t)'\Big]
\end{aligned}$$

At this point it is clear that $(\alpha_{t-1} - \hat{\mathbf{a}}_{t-1|t-1})$ is uncorrelated with process noise, the expectation can be re-written as

$$\mathbf{P}_{t|t-1} = T_t E\Big[(\alpha_{t-1} - \hat{\mathbf{a}}_{t-1|t-1})(\alpha_{t-1} - \hat{\mathbf{a}}_{t-1|t-1})'\Big]T'_t + R_t E\Big[\eta_t\eta'_t\Big]R'_t$$

where

$$E[\eta_t\eta'_t] = Q_t \tag{2.3.8}$$

using the definition of (2.3.6) and (2.3.8), we get

$$\mathbf{P}_{t|t-1} = T_t\mathbf{P}_{t-1|t-1}T'_t + R_tQ_tR'_t \tag{2.3.9}$$

thus, the state estimate $\hat{\mathbf{a}}_{t|t-1}$ and covariance error matrix $\mathbf{P}_{t|t-1}$ are called predict equations.

Now evolving in time observations such as $\mathbf{y}_t$ becomes available. We can update the estimator of α_t and $\hat{\mathbf{a}}_{t|t-1}$ respectively. Assuming that *priori estimate* of $\hat{\mathbf{a}}_{t|t}$ is given by $\hat{\mathbf{a}}_{t|t-1}$ and it was obtained by having the knowledge of the system. The update equation of the new estimate which is the *posteriori estimate* is given by combining the previous estimate which is the *priori estimate* with the measurement information using the correction term weighted by the Kalman gain, see (Mastro, 2013). Thus we have,

$$\begin{aligned}
\hat{\mathbf{a}}_{t|t} &= \hat{\mathbf{a}}_{t|t-1} + \mathbf{K}_t(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}) \\
\hat{\mathbf{a}}_{t|t} &= \hat{\mathbf{a}}_{t|t-1} + \mathbf{K}_t\mathbf{y}_t - \mathbf{K}_t\hat{\mathbf{y}}_{t|t-1} \\
\hat{\mathbf{a}}_{t|t} &= \hat{\mathbf{a}}_{t|t-1} + \mathbf{K}_t(Z_t\alpha_t + d_t + \epsilon_t) - \mathbf{K}_t(Z_t\hat{\mathbf{a}}_{t|t-1} + d_t) \\
\hat{\mathbf{a}}_{t|t} &= \hat{\mathbf{a}}_{t|t-1} - \mathbf{K}_tZ_t\hat{\mathbf{a}}_{t|t-1} + \mathbf{K}_tZ_t\alpha_t + \mathbf{K}_t\epsilon_t + \mathbf{K}_td_t - \mathbf{K}_td_t
\end{aligned}$$

by collecting like terms, we get

$$\hat{\mathbf{a}}_{t|t} = \hat{\mathbf{a}}_{t|t-1}(I - \mathbf{K}_tZ_t) + \mathbf{K}_tZ_t\alpha_t + \mathbf{K}_t\epsilon_t \tag{2.3.10}$$

where $\mathbf{K}_t$ is the Kalman gain, the term $\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}$ is known to be the correction term or measurement residual denoted by i_t,

$$i_t = \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1} \tag{2.3.11}$$

Kalman filter's key feature is the weighting through the Kalman gain of the residual calculation. Typically the Kalman gain is set to minimize the variance of errors $e_{t|t} = \alpha_t - \hat{\mathbf{a}}_{t|t}$.

The *posteriori* covariance error matrix $\mathbf{P}_{t|t}$ at time t is equivalent with the mean squared error, and is given by,

$$\mathbf{P}_{t|t} = cov(\alpha_t - \hat{\mathbf{a}}_{t|t}) = E\Big[(\alpha_t - \hat{\mathbf{a}}_{t|t})(\alpha_t - \hat{\mathbf{a}}_{t|t})'\Big]$$

$$\mathbf{P}_{t|t} = E\Big[\Big(\alpha_t - \hat{\mathbf{a}}_{t|t-1}(I - \mathbf{K}_t Z_t) - \mathbf{K}_t Z_t \alpha_t - \mathbf{K}_t \epsilon_t\Big)\Big(\alpha_t - \hat{\mathbf{a}}_{t|t-1}(I - \mathbf{K}_t Z_t) - \mathbf{K}_t Z_t \alpha_t - \mathbf{K}_t \epsilon_t\Big)'\Big]$$

$$\mathbf{P}_{t|t} = E\Big[\Big(\alpha_t(I - \mathbf{K}_t Z_t) - \hat{\mathbf{a}}_{t|t-1}(I - \mathbf{K}_t Z_t) - \mathbf{K}_t \epsilon_t\Big)\Big(\alpha_t(I - \mathbf{K}_t Z_t) - \hat{\mathbf{a}}_{t|t-1}(I - \mathbf{K}_t Z_t) - \mathbf{K}_t \epsilon_t\Big)'\Big]$$

$$\mathbf{P}_{t|t} = E\Big[\Big((I - \mathbf{K}_t Z_t)(\alpha_t - \hat{\mathbf{a}}_{t|t-1}) - \mathbf{K}_t \epsilon_t\Big)\Big((I - \mathbf{K}_t Z_t)(\alpha_t - \hat{\mathbf{a}}_{t|t-1}) - \mathbf{K}_t \epsilon_t\Big)'\Big]$$

At this point it is clear that $(\alpha_t - \hat{\mathbf{a}}_{t|t-1})$ s uncorrelated with the measurement noise and, therefore we re-write our expectation as

$$\mathbf{P}_{t|t} = \Big(I - \mathbf{K}_t Z_t\Big) E\Big[(\alpha_t - \hat{\mathbf{a}}_{t|t-1})(\alpha_t - \hat{\mathbf{a}}_{t|t-1})'\Big]\Big(I - \mathbf{K}_t Z_t\Big)' + \mathbf{K}_t E\Big[\epsilon_t \epsilon_t'\Big]\mathbf{K}_t'$$

using our *priori* error covariance $\mathbf{P}_{t|t-1}$ and the definition of $E(\epsilon_t \epsilon_t')$, we get

$$\mathbf{P}_{t|t} = (I - \mathbf{K}_t Z_t)\mathbf{P}_{t|t-1}(I - \mathbf{K}_t Z_t)' + \mathbf{K}_t H_t \mathbf{K}_t' \tag{2.3.12}$$

This equation (2.3.12) is valid for any $\mathbf{K}_t$ value (Sometimes referred to as the covariance update equation Joseph form). It turns out that it can be further simplified if $\mathbf{K}_t$ is the optimal Kalman gain that will be shown shortly.

We derive the optimal Kalman gain by minimizing the expected value of $E[(\alpha_t - \hat{\mathbf{a}}_{t|t})^2]$. This is equivalent to reducing the *posteriori* error covariance matrix trace. By expanding out the terms in equation (2.3.12), we get

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{K}_t Z_t - \mathbf{P}_{t|t-1}\mathbf{K}_t' Z_t' + \mathbf{K}_t Z_t \mathbf{P}_{t|t-1}\mathbf{K}_t' Z_t' + \mathbf{K}_t H_t \mathbf{K}_t'$$

simplifying, we get

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{K}_t Z_t - \mathbf{P}_{t|t-1}\mathbf{K}_t' Z_t' + \mathbf{K}_t(Z_t \mathbf{P}_{t|t-1} Z_t' + H_t)\mathbf{K}_t' \tag{2.3.13}$$

$\mathbf{P}_{t|t}$ is the error covariance matrix updation equation, which conatins the mean squared error in the diagonal as shown below see (Lacey, 2019)

$$\mathbf{P}_{t|t} = \begin{bmatrix} E[e_{t-1}e_{t-1}'] & E[e_t e_{t-1}'] & E[e_{t+1}e_{t-1}'] \\ E[e_{t-1}e_t'] & E[e_t e_t'] & E[e_{t+1}e_t'] \\ E[e_{t-1}e_{t+1}'] & E[e_t e_{t+1}'] & E[e_{t+t}e_{t+1}'] \end{bmatrix} \tag{2.3.14}$$

the trace of the covariance error matrix is the sum of the mean squared errors, thus we minimize the mean squared error by minimizing the trace of $\mathbf{P}_{t|t}$, which is

$$trace[\mathbf{P}_{t|t}] = trace[\mathbf{P}_{t|t-1}] - trace[\mathbf{P}_{t|t-1}\mathbf{K}_t Z_t] - trace[\mathbf{P}_{t|t-1}\mathbf{K}_t' Z_t'] + trace[\mathbf{K}_t(Z_t \mathbf{P}_{t|t-1} Z_t' + H_t)\mathbf{K}_t']$$

Note that the trace of a matrix is the same as the trace of the same matrix being transposed, so we have

$$trace[\mathbf{P}_{t|t}] = trace[\mathbf{P}_{t|t-1}] - 2trace[\mathbf{P}_{t|t-1}\mathbf{K}_t Z_t] + trace[\mathbf{K}_t(Z_t \mathbf{P}_{t|t-1} Z_t' + H_t)\mathbf{K}_t']$$

where $trace[\mathbf{P}_{t|t}]$ is the trace of the covariance error matrix $\mathbf{P}_{t|t}$. The trace is minimized when its derivative matrix is zero with respect to the matrix of Kalman gain.

$$\frac{d[trace[\mathbf{P}_{t|t}]]}{d\mathbf{K}_t} = 0 \tag{2.3.15}$$

Using the gradient matrix rules and the matrix symmetry, we get that

$$\frac{d[trace[\mathbf{P}_{t|t}]]}{d\mathbf{K}_t} = -2\mathbf{P}_{t|t-1}Z_t' + 2\mathbf{K}_t(Z_t\mathbf{P}_{t|t-1}Z_t' + H_t) \tag{2.3.16}$$

setting equation (2.3.16) to zero and re-arrange gives,

$$2\mathbf{P}_{t|t-1}Z_t' = 2\mathbf{K}_t(Z_t\mathbf{P}_{t|t-1}Z_t' + H_t)$$

solving for $\mathbf{K}_t$ gives the Kalman gain expression with minimum mean squared error, therefore we have

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}Z_t'\Big(Z_t\mathbf{P}_{t|t-1}Z_t' + H_t\Big)^{-1} \tag{2.3.17}$$

we let F_t be the measurement residual associated with measurement prediction and is defined by

$$F_t = Z_t\mathbf{P}_{t|t-1}Z_t' + H_t$$

substituting F_t in (2.3.17), we have

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}Z_t'F_t^{-1} \tag{2.3.18}$$

This gain, is referred to as the optimal Kalman gain, is the one that gives estimates of the MMSE (Minimum Mean Squared Error) when used..

Now, substituting equation (2.3.18) into equation (2.3.13) below, we have

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{K}_tZ_t - \mathbf{P}_{t|t-1}\mathbf{K}_t'Z_t' + \mathbf{K}_t(Z_t\mathbf{P}_{t|t-1}Z_t' + H_t)\mathbf{K}_t$$
$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}(\mathbf{P}_{t|t-1}Z_t'F_t^{-1})Z_t - \mathbf{P}_{t|t-1}(\mathbf{P}_{t|t-1}Z_t'F_t^{-1})'Z_t' + (\mathbf{P}_{t|t-1}Z_t'F_t^{-1})F_t(\mathbf{P}_{t|t-1}Z_t'F_t^{-1})'$$

we see that last two terms cancel out, giving

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}(\mathbf{P}_{t|t-1}Z_t'F_t^{-1})Z_t$$
$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1}(I - \mathbf{P}_{t|t-1}Z_t'F_t^{-1}Z_t)$$

therefore, using (2.3.18) we have

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1}(I - \mathbf{K}_tZ_t) \tag{2.3.19}$$

So we have the update equation of error covariance matrix with optimal gain with minimum mean squared error. This equation is computationally cheaper and is therefore practically always used, but is only accurate for the optimal benefit. If arithmetic accuracy is unusually low causing numerical stability problems or if a deliberate use is made of a non-optimal Kalman gain, this simplification can not be applied; the equation of a posteriori error covariance (Joseph form) must be used(Mastro, 2013).

2.3.2 The Kalman Filter Equations. The prediction equations together with the updation equations makes up the Kalman filter equations. The mathematical Kalman filter equations are as follows

1. State prediction equation

$$\hat{\mathbf{a}}_{t|t-1} = T_t \hat{\mathbf{a}}_{t-1|t-1} + c_t$$

2. Covariance prediction equation

$$\mathbf{P}_{t|t-1} = T_t \mathbf{P}_{t-1|t-1} T_t' + R_t Q_t R_t'$$

3. Kalman Gain equation

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} Z_t' F_t^{-1}$$

4.State updation equation

$$\hat{\mathbf{a}}_{t|t} = \hat{\mathbf{a}}_{t|t-1}(I - \mathbf{K}_t Z_t) + \mathbf{K}_t Z_t \alpha_t + \mathbf{K}_t \epsilon_t$$

5.Covarience updation equation

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1}(I - \mathbf{K}_t Z_t)$$

3. Application of Kalman Filter in Ornstein-Uhlenbeck Model

The Ornstein-Uhlenbeck method is the most common stochastic process defining the characteristic of the process of drifting towards the mean. The Ornstein-Uhlenbeck process and its applications are being discussed in the preceeding sections. This approach relies on the state space representation and the Kalman filter. The formulation of state-space representation combined with Kalman filtering can optimally estimate the parameter set of the observation price. We begin with the relationship between the observed price and state variables whereby the observed system is the measurement equation. The unobserved system is the transition equation, which describes the dynamics of state variables, and is given as Ornstein-Uhlenbeck Model. The measurement equation together with the transition equation, represents a state-space bond market model.

3.1 Ornstein-Uhlenbeck Model

Ornstein-Uhlenbeck (1930) gave the first description of an ordinary mean process of reversion. The Ornstein-Uhlenbeck process is the only non-trivial process that meets these three conditions, up to allowing linear space and time variables to be transformed. Over time, the process continues to return to its long-term mean: it is called a mean-reverting process. The commodity price behaviour, economic principle, interest rates and foreign exchange rates are well described by mean reverting process. The Ornstein-Uhlenbeck process, however, has emerged in finance in recent years as a measure of the instability of the underlying price system.

3.1.1 Derivation of Ornstein-Uhlenbeck Process Solution. The Ornstein–Uhlenbeck process X_t is defined by the following stochastic differential equation:

$$dX_t = -\lambda dt + \sigma dW_t \tag{3.1.1}$$

where $\lambda > 0$ and $\sigma > 0$ are the model parameters and W_t denotes a the Wiener process see (Mastro, 2013). Sometimes an additional term of drift is added, then giving Ornstein–Uhlenbeck process to be,

$$dX_t = \lambda(\mu - X_t)dt + \sigma dW_t \tag{3.1.2}$$

where λ is the speed of reversion, μ is the long-term mean rate and σ is the volatility. This model in (3.1.2) can also being refer to as the Vasicek model in financial mathematics see (Focardi and Fabozzi, 2004).

We want to derive the solution the equation (3.1.2). Thus expanding out equation (3.1.2), we get

$$\begin{aligned} dX_t &= \lambda\mu dt - \lambda X_t dt + \sigma dW_t \\ dX_t + \lambda X_t dt &= \lambda\mu dt + \sigma dW_t \end{aligned}$$

the equation above is the stochastic differential equation form see (Mastro, 2013). We solve this equation by multiplying with the integration factor $e^{\lambda t}$, thus we have,

$$e^{\lambda t}dX_t + e^{\lambda t}\lambda X_t dt = \lambda\mu e^{\lambda t}dt + \sigma e^{\lambda t}dW_t \tag{3.1.3}$$

the left-hand side of (3.1.3) becomes the exact differential using product rule,

$$d(e^{\lambda t}X_t) = \lambda\mu e^{\lambda t}dt + \sigma e^{\lambda t}dW_t \tag{3.1.4}$$

thus, integrating (3.1.4) from s to T, where $0 \leq s < T$, we get

$$\int_s^T d(e^{\lambda t}X_t) = \int_s^T \lambda\mu e^{\lambda t}dt + \int_s^T \sigma e^{\lambda t}dW_t \tag{3.1.5}$$

evaluating the integrals of (3.1.5), we have

$$e^{\lambda T}X_T - e^s X_s = \lambda\mu\frac{e^{\lambda T} - e^s}{\lambda} + \sigma\int_s^T e^{\lambda t}dW_t \tag{3.1.6}$$

multiplying (3.1.5) by $e^{-\lambda T}$, we get

$$X_T - X_s e^{-\lambda(T-s)} = \mu(1 - e^{-\lambda(T-s)}) + \sigma e^{-\lambda T}\int_s^T e^{\lambda t}dW_t \tag{3.1.7}$$

re-arranging the equation we get the solution of the Ornstein-Uhlenbeck.

$$X_T = X_s e^{-\lambda(T-s)} + \mu(1 - e^{-\lambda(T-s)}) + \sigma\int_s^T e^{\lambda(T-t)}dW_t \tag{3.1.8}$$

It is easy to see that X is normally distributed, one just needs to recall that the integral of the deterministic function with respect to brownian is simply Gaussian see (Mastro, 2013). Thus we can calculate the mean and the variance.

From the first moment, the mean is given by, taking the expected value both side in (3.1.8)

$$E[X_T] = E\Big[X_s e^{-\lambda(T-s)} + \mu(1 - e^{-\lambda(T-s)}) + \sigma\int_s^T e^{\lambda(T-t)}dW_t] \tag{3.1.9}$$

Note that the expected value of the deterministic function with respect to brownian is zero, thus we have

$$E[X_T] = X_s e^{-\lambda(T-s)} + \mu(1 - e^{-\lambda(T-s)}) \tag{3.1.10}$$

The variance is found from the integral of the stochastic process, and is given by

$$Var[X_T] = [(X_T - E[X_t])^2] \tag{3.1.11}$$

sustituting (3.1.10) in (3.1.11), we get the variance

$$Var[X_T] = E\Big[\Big(\sigma\int_s^T e^{-\lambda(T-t)}dW_t\Big)^2\Big] \tag{3.1.12}$$

thus, we have deterministic integral that we can integrate

$$Var[X_T] = \sigma^2\int_s^T e^{-2\lambda(T-t)}dt = \sigma^2 e^{-2\lambda T}\Big[\frac{e^{2\lambda T} - e^{-2\lambda s}}{2\lambda}\Big]$$

thus, giving our desired variance to be

$$Var[X_T] = \sigma^2\frac{(1 - e^{-2\lambda(T-s)})}{2\lambda} \tag{3.1.13}$$

therefore, we have the solution of the Ornstein-Uhlenbeck process

3.2 State Space Formulation and Kalman Filter

We now see how the model's unobservable variables and parameter set can be estimated using the Kalman filter equations. Firstly, the Ornstein-Uhlenbeck process solution is written in a state space representation that allows measurement errors and then the Kalman filter equations are obtained using results from the preceeding chapter.

The Ornstein-Uhlenbeck process is the simplest mean-reverting model given by:

$$dX_t = \lambda(\mu - X_t)dt + \sigma dW_t$$

in the case where X_t=ln(S_t) is the log-price in (3.1.2). The stochastic differential equation can be discretized following the derivation of the Ornstein-Uhlenbeck process solution in the previous section and is given by

$$\ln S_t = \ln(S_{t-1})e^{-\lambda\delta_t} + \mu(1 - e^{-\lambda\delta_t}) + \sigma\sqrt{\frac{(1 - e^{-2\lambda\delta_t})}{2\lambda}} \tag{3.2.1}$$

where lnS_t is the observation log-price , lnS_{t-1} is the previous observation log-price and δ_t is the change in time.

Our aim is to represent equation (3.2.1) in state space representation. More insight is gained through the application of the Kalman filter to a real process.

We want to match Ornstein-Uhlenbeck process's solution to the state transition equation in (2.2.2), thus we can relate (3.2.1) with our process equation,

$$\ln S_t = \mathsf{M}\ln(S_{t-1}) + \mathsf{C} + \sigma\sqrt{\mathsf{Q}_t} \tag{3.2.2}$$

where,

$$\mathsf{M} = e^{-\lambda\delta_t} \quad , \mathsf{C} = \mu(1 - e^{-\lambda\delta_t}) \quad , \sqrt{\mathsf{Q}_t} = \sqrt{\frac{(1 - e^{-2\lambda\delta_t})}{2\lambda}}$$

The term Q_t is the variance of the process noise. One of the applications of the Kalman filter would be to predict in real time the future price. This includes a previous estimate of the parameter set $\{\mu, \delta, \lambda\}$.

Thus transition equation is given by,

$$\ln S_t = \mathsf{M}\ln(S_{t-1}) + \mathsf{C} + \eta_t \tag{3.2.3}$$

and the measurement equation for the oberseved price is given by,

$$\mathbf{y}_t = \mathsf{Z}\ln S_t + \epsilon_t \tag{3.2.4}$$

where, observation model $\mathsf{Z} = 1$.

We have represented our model in state space form, which is specified by (3.2.4) and (3.2.3). We use two techniques to estimate the parameters of an Ornstein-Uhlenbeck process. We use the Kalman filter to estimate the specified parameter set $\{\mu, \delta, \lambda\}$. Nonetheless this approach requires an auxiliary program to use the Kalman filter as an objective function to search for the optimal parameter set. The reliability of the parameter set of the model is evaluated by the value of the maximum multivariate likelihood based on the measurement residual given as follows,

$$V_t = \mathbf{y}_t - \mathsf{Z}\hat{\mathsf{x}}_{t|t-1}$$

and the covariance measurement residual,

$$F_t = \mathsf{Z}\mathbf{P}_{t|t-1}\mathsf{Z}' + \mathsf{H}_t$$

The n sequential data points conditional density function x_i with a normal distribution for any mean μ and the typical σ variance is

$$f(x_1, \ldots, x_n \mid \mu, \hat{\sigma}) = \prod \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}} = \frac{2\pi^{(-n/2)}}{\sigma^n} e^{-\frac{\sum(x_i-\mu)^2}{2\sigma^2}}$$

The logarithm of likelihood is more convenient to measure as it is expressed in terms of summation rather than multiplication.

Thus the log-likelihood function for equation (3.2.3) is given by,

$$\begin{aligned}
L &= \log(f(\ln(S_t) \mid \ln(S_{t-1}))), \mu, \hat{\sigma}, \lambda) \\
L &= \sum_{t=1}^{l_n} \log\Big(f(\ln(S_t) \mid \ln(S_{t-1}))), \mu, \hat{\sigma}, \lambda) \Big) \\
L &= -\sum_{t=1}^{l_n} \left[\frac{n}{2}\ln(2\pi) + \frac{1}{2} \mid F_t \mid + \frac{1}{2} V_t (F_t)^{-1} V_t' \right]
\end{aligned}$$

where , V_t is the correction term called measurement residual, and is given by

$$V_t = \mathbf{y}_t - \mathsf{Z}\hat{\mathsf{x}}_{t|t-1}$$

and the covariance of the measurement residual is given by

$$F_t = E\left[\Big(\mathbf{y}_t - \mathsf{Z}\hat{\mathsf{x}}_{t|t-1}) \Big) \Big(\mathbf{y}_t - \mathsf{Z}\hat{\mathsf{x}}_{t|t-1}) \Big)' \right]$$

Thus log-likelihood function of Ornstein-Uhlenbeck method solution in (3.2.1) after substituting the covariance of the measurement residual and measurement residual V_t is given by,

$$L = -\frac{n}{2}\ln(2\pi) + \sum_{t=1}^{l_n} \left[-\frac{1}{2}\log \mid F_t \mid -\frac{1}{2} \Big(\mathbf{y}_t - \mathsf{Z}\hat{\mathsf{x}}_{t|t-1} \Big)' (F_t)^{-1} \Big(\mathbf{y}_t - \mathsf{Z}\hat{\mathsf{x}}_{t|t-1} \Big) \right]$$

The logarithm price is used to suit the Kalman filter's linear structure. The unobservable parameters of the state evolve through the transition equation

$$\hat{x}_{t|t-1} = \mathsf{M}\hat{x}_{t-1|t-1} + \mathsf{C} \tag{3.2.5}$$

sustituting M and C, we get

$$\hat{x}_{t|t-1} = e^{-\lambda\delta_t}\hat{x}_{t-1|t-1} + \mu(1 - e^{-\lambda\delta_t}) + \eta_t \tag{3.2.6}$$

The process noise η_t is normally distributed with has a zero mean and covariance of,

$$\mathsf{Q}_t = \hat{\sigma}^2 = \sigma^2 \sqrt{\frac{(1 - e^{-2\lambda\delta_2})}{2\lambda}}$$

This is due to the underlying randomness of a stochastic system dampened by mean reversion. As a direct measurement of the price, $\mathsf{Z} = 1$ is unity. The measurement equation is given by,

$$\hat{\mathbf{y}}_{t|t-1} = \mathsf{Z}\hat{x}_{t-1|t-1} + \epsilon_t \tag{3.2.7}$$

where the measurement noise results from bid-ask distribution, errors in data feed and related errors. Now that we have our measurement equation (3.2.4) and our transition equation (3.2.3).

The equation of log-price is represented in state space form. We can now apply Kalman Filter algorithm. We'll only derive the two predictions equation, the updation equations follows as it was shown in the preceeding chapter.

Let $\hat{x}_{t-1|t-1}$ be the estimate of $\ln(S_{t-1})$ given all the measurement information up to and including time index $t-1$ be given by

$$\hat{x}_{t-1|t-1} = E[\ln(S_{t-1}) \mid \mathbf{y}_1, \ldots, \mathbf{y}_{i-1}]$$

From chapter 2 we know that the *priori* estimate of $\ln(S_t)$ is given by,

$$\begin{aligned}
\hat{x}_{t|t-1} &= E[\ln(S_t) \mid \mathbf{y}_1, \ldots, \mathbf{y}_{t-1}] \\
&= E[\mathsf{M}\ln(S_{t-1}) + \mathsf{C} + \eta_t \mid \mathbf{y}_1, \ldots, \mathbf{y}_{t-1}] \\
&= \mathsf{M}E[\ln(S_{t-1}) \mid \mathbf{y}_1, \ldots, \mathbf{y}_{t-1}] + E[\mathsf{C} \mid \mathbf{y}_1, \ldots, \mathbf{y}_{t-1}] + E[\eta_t \mid \mathbf{y}_1, \ldots, \mathbf{y}_{t-1}]
\end{aligned}$$

by definition of $E[\eta_t] = 0$ we have

$$\hat{x}_{t|t-1} = \mathsf{M}\hat{x}_{t-1|t-1} + \mathsf{C} \tag{3.2.8}$$

thus, giving us the *prior estimate* of $\ln(S_t)$.

The error covariance matrix is equivalent to the mean squared error, and is given by the covariance between the true state and the estimated state. Let $\mathbf{P}_{t-1|t-1}$ denote the error covariance of $\ln(S_{t-1})$ given all the measurement information up to and including time index $t-1$, and is given by

$$\mathbf{P}_{t-1|t-1} = E[(\ln(S_{t-1}) - \hat{x}_{t-1|t-1})(\ln(S_{t-1}) - \hat{x}_{t-1|t-1})']$$

Thus the *priori* error covariance matrix of $\ln(S_t)$ is given by

$$\begin{aligned}
\mathbf{P}_{t|t-1} &= E[(\ln(S_t) - \hat{x}_{t|t-1})(\ln(S_t) - \hat{x}_{t|t-1})'] \\
&= E[(\mathsf{M}\ln(S_{t-1}) + \mathsf{C} + \eta_t - \mathsf{M}\hat{x}_{t-1|t-1} - \mathsf{C})(\mathsf{M}\ln(S_{t-1}) + \mathsf{C} + \eta_t - \mathsf{M}\hat{x}_{t-1|t-1} - \mathsf{C})'] \\
&= E\left[\Big(\mathsf{M}(\ln(S_{t-1}) - \hat{x}_{t-1|t-1}) + \eta_t\Big)\Big(\mathsf{M}(\ln(S_{t-1}) - \hat{x}_{t-1|t-1}) + \eta_t\Big)'\right]
\end{aligned}$$

At this point it is clear that $(\ln(S_{t-1}) - \hat{x}_{t-1|t-1})$ is uncorrelated with process noise. We can re-write the expectation as

$$\mathbf{P}_{t|t-1} = \mathsf{M}E\left[\Big(\ln(S_{t-1}) - \hat{x}_{t-1|t-1}\Big)\Big(\ln(S_{t-1}) - \hat{x}_{t-1|t-1}\Big)'\right]\mathsf{M}' + E\left[\eta_t\eta_t'\right]$$

using the definition of $\mathbf{P}_{t-1|t-1}$ and $E[\eta_t\eta_t']$, we get

$$\mathbf{P}_{t|t-1} = \mathsf{M}\mathbf{P}_{t-1|t-1}\mathsf{M}' + \mathsf{Q}_t$$

thus giving the *prior* error covariance estimate. The updation equations follows from the state prediction equations. The update equations proofs follows as demonstrated in chapter two. Thus the Kalman Filter equation of the Ornstein-Uhlenbeck process for price are given as follows,

1. State Prediction

$$\hat{x}_{t|t-1} = \mathsf{M}\hat{x}_{t-1|t-1} + \mathsf{C}$$

2. Error Covariance Prediction

$$\mathbf{P}_{t|t-1} = \mathsf{M}\mathbf{P}_{t-1|t-1}\mathsf{M}' + \mathsf{Q}_t$$

3. Kalman Gain

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathsf{Z}F_t^{-1}$$

4. State Updation

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \mathbf{K}_t(\mathbf{y}_t - \mathsf{Z}\hat{\mathsf{x}}_{t|t-1})$$

5. Error Covariance Updation

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{K}_t\mathsf{Z}$$

3.2.1 Simulation of Ornstein Ulhenbeck Process . In this simulation of our model process we generate the simulated data of real price, and we then use Kalman filter to estimate the parameter set $\{\mu, \sigma, \lambda\}$ and then compares with the true parameters. In our python code the function SimulateOUkalman creates a true mean reversion parameters for the data set of real price with system noise as a mean reversion process. The observed price is calculated by adding measurement noise to the actual price system. We use the function called OUloglikelihoodfn that uses the specific set of model parameters and the observed price to predict the price through the Kalman filter at each time point iteratively. The scipy optimizer function called fmin helps us to estimates the parameters of the system by maximising the log likelihood which is the output of the function OUloglikelihoodfn. See the table below comparing our true paramters set and our estimated parameter set.

Parameters	μ	λ	σ
True	1	3	0.5
Estimated	1.03	2.82	0.45

Measurement Noise Covariance $= \mathsf{H} = 0.0076$

we see that the Kalman filter's maximum likelihood offers a good approximation of the true parameters from a relatively small sampling space:using the generated observed prices.

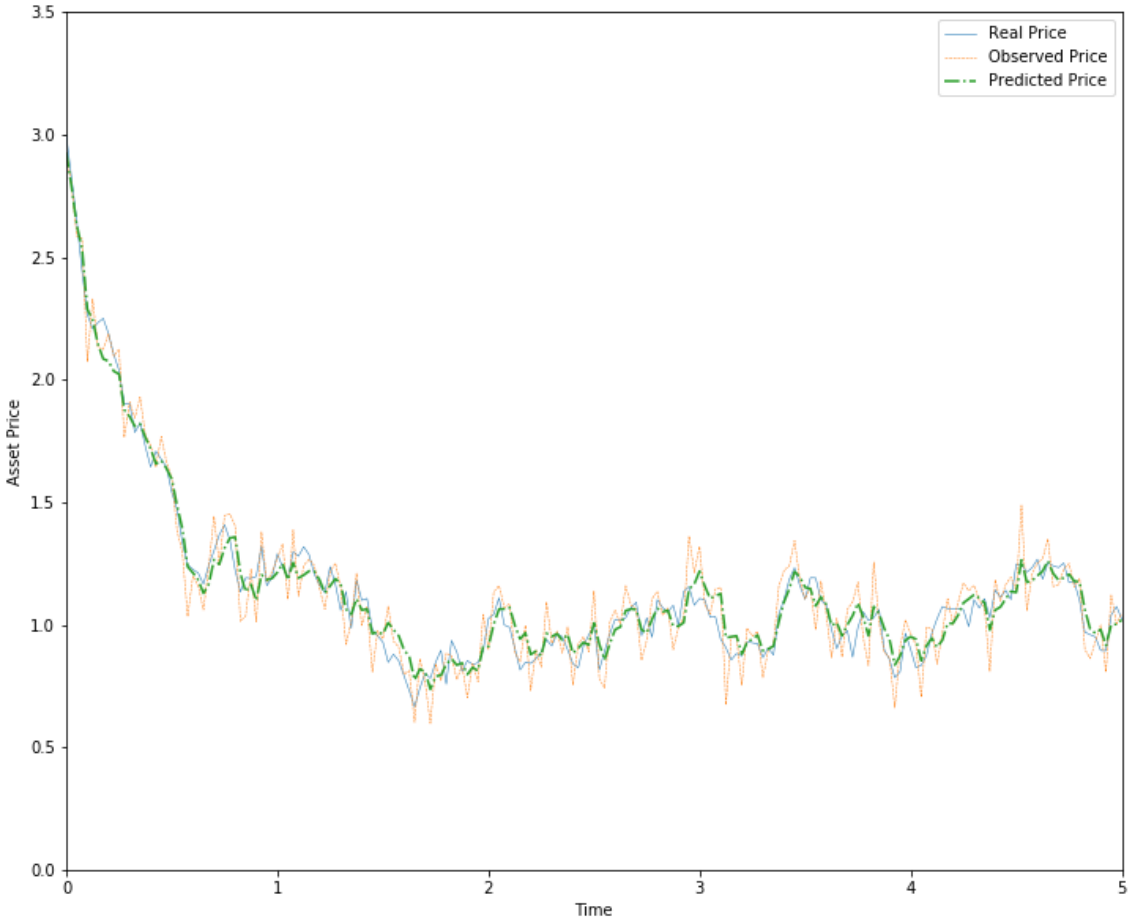

Figure 3.1: Simulation of Ornstein Uhlenbeck Process

3.3 Model Simulation Results

This section presents the simulation results of the actual bond market data. The raw data used in our empirical work consisted of U.S. Treasury Bonds.The data was downloaded from the website of Yahoo finance. The raw data includes the Date, Open, High, Low, Close, Adj Close and Volume. The data sample daily quotes ranging from the 31^{st} of December 1992 to the 8^{th} of November 2017. This raw data uses 10 years maturity, see the partial listing of U.S. Treasury Bond below

	Date	Open	High	Low	Close	Adj Close	Volume
0	1992-12-31	6.70	6.70	6.70	6.70	6.70	0.0
1	1993-01-04	6.60	6.60	6.60	6.60	6.60	0.0
2	1993-01-05	6.61	6.61	6.61	6.61	6.61	0.0
3	1993-01-06	6.63	6.63	6.63	6.63	6.63	0.0
4	1993-01-07	6.76	6.76	6.76	6.76	6.76	0.0
5	1993-01-08	6.75	6.75	6.75	6.75	6.75	0.0
6	1993-01-11	6.71	6.71	6.71	6.71	6.71	0.0
7	1993-01-12	6.72	6.72	6.72	6.72	6.72	0.0
8	1993-01-13	6.71	6.71	6.71	6.71	6.71	0.0
9	1993-01-14	6.65	6.65	6.65	6.65	6.65	0.0

Figure 3.2: Partial listing of U.S. Treasury Bond maturity

We use Adj Close price as it adjusts the closing price of a stock to correctly reflect the value of that stock after accounting for any corporate shares. The estimated parameter set are given in the table below.

Parameters	μ	λ	σ
True	3	1	2
Estimated	3.26	0.86	2.04

Measurement Noise Covariance $=$ H $=$ 0.0104

The Kalman filter's maximum likelihood offers a good approximation of the true parameter set as shown in the table above. The simulated price with the Ornstein-Uhlenbeck process are shown below

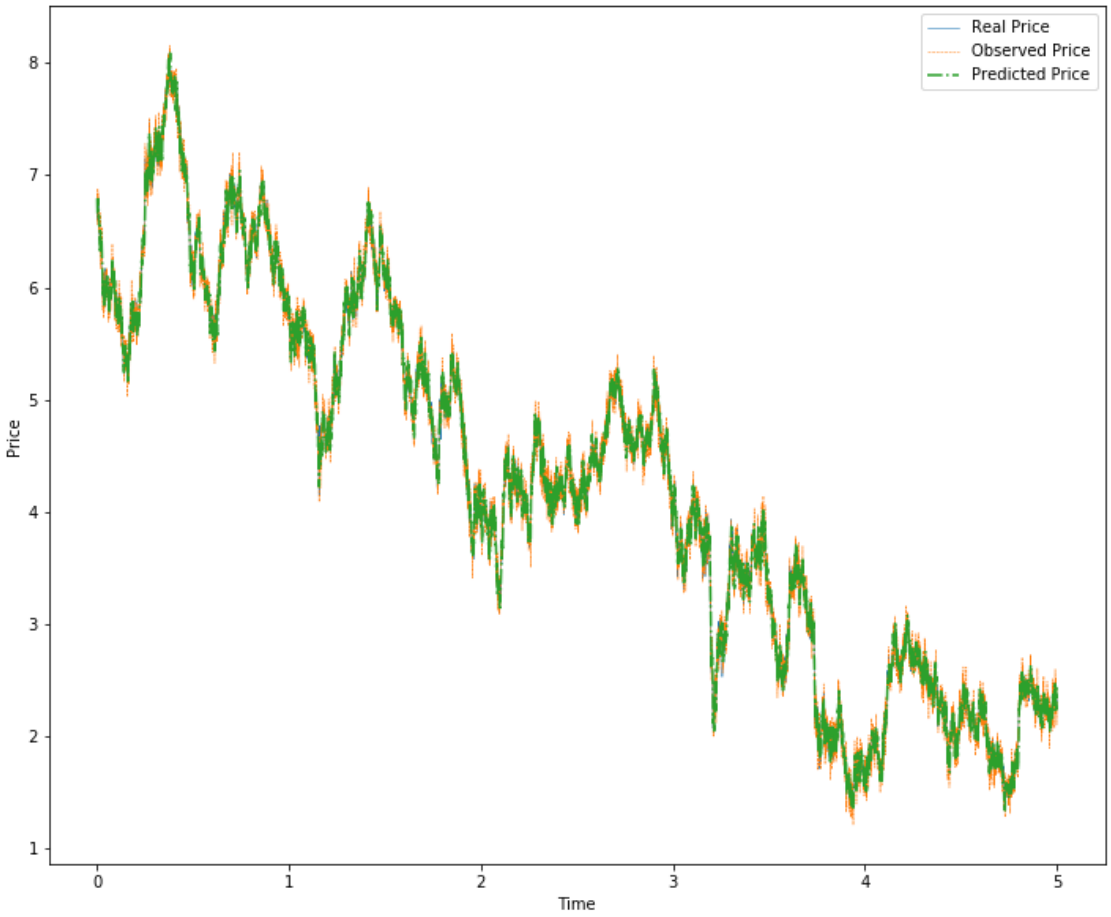

Figure 3.3: Simulation of Ornstein Uhlenbeck Process

Figure 3.1 and **Figure 3.3** above, shows that a mean reversion cycle matches the real price (solid line). Increasing a measurement noise that is usually distributed to the real price gives the observed value system (thin line). The predicted price is calculated using the Kalman filter (dot-dashed line) which is real close to the observed price. Our parameter μ is actually 7% off towards the true μ, λ is 16% off towards the true mean speed of reversion and lastly σ is actually 7% off towards the true σ. The Kalman filter with Gaussian distributions is quite robust and adaptable to univariating and multivariating economic system variables. The application of a Kalman filter in this instance to match price information basically replicates a linear fitting routine.

The application of Kalman filter to estimated the parameter set using maximum likehood to optimize our parameter set proves to be almost realistic. Althought there are some errors in our estimate parameters, our predicted price looks more likely like the real prices which shows that Kalman filter is a good method to estimate unobservable parameter set. To observe the action of mean reversion, we simulated the

stochastic behaviour of bond price with a mean reversion speed equal to 1 ($\lambda = 1$) to observe the weak mean reversion.

The simulated Ornstein-Uhlenbeck model shows that prices tend to revert towards the mean. This is in agreement with the well known fact that the mean and variance of the Ornstein Uhlenbeck model tend to deterministic values as T goes to infinity for fixed s. This property makes the Ornstein-Uhlenbeck a weakly stationary process, hence a tractable model for bond market. So, by using the Ornstein-Uhlenbeck process method, we can predict the prices in the future market prices. We find that estimates of parameters are close to actual parameters. Thus Ornstein-Uhlenbeck converges as T goes to infinity.

The python code for this simulation can be found in https://github.com/Willie-Mohlongo/willie-KalmanFilter.git

4. Conclusion

The main objective of this essay was to formulate bond market models that capture the following long-term characteristics, which are market volatility, ambiguity, and partial observation in the emerging economy markets like South Africa. Our contribution was to construct a filtering model for a bond market with particular attention on long term maturities and uncertainty in interest rates models.

This filtering model was achieved by constructing the linear state-space representation, which enables both cross-sectional and time sequence constraints into consideration in the term structure. The Kalman filter algorithm was introduced to optimally predict the current term structure of the underlying asset price process for the long-term bond market.

Furthermore, the Ornstein-Uhlenbeck models was reviewed. We derive the dynamics of the the solution process and then introduced the implementation of state-space formulation combined with the Kalman filter algorithm. This implementation ensured a suitable representation of the underlying unobservable state variables. This model is helpful in optimally predicting unobservable assert with added noise using Maximum likelihood to check the quality of the model.

Future studies include the use of the Kalman filter algorithm in situations where Ornstein-Uhlenbeck model assumes the second factor of uncertatity such as inflation rate. We can develop Ornstein-Uhlenbeck process using these additional factors.

Acknowledgements

Firstly, I want to acknowledge AIMS and it's funders for supporting this work, as well as my supervisor Dr. Dennis Ikpe and co-supervisor Prof. Ronnie Becker for their encouragement and guidance until I finish this essay. I have learnt a lot from both of them during this time even though it was difficult at some point.,

I have received a great deal of support and encouragement from my family. In particular, my brothers Bongani George Mohlongo, Smanga Joseph Mohlongo, Siphiwe Mohlongo, my one and only sister Cindy Mohlongo, my niece Mathabo Mohlongo, my nephew Bonginkosi Donald and lastly my mother, Miss Lucy Mohlongo, for their Prayers and daily support during the difficult times.

I would like to thank all my tutors who have been very helpful to me, especially my current tutor Rock Stephane Koffi who have been encouraging and motivating me to work hard and be strong during the tough times. I will forever be thankful to him for making it possible for me to finish even though at times I wanted to quit. I would also like to thank Bernard Lebeko Poulo for his advices and his positive energy during difficult times and my former tutors Dr. Kenneth Dadedzi, Jordan Ngucho Yvan Mbeutchou, Elbasher Mohamed, and Dr. Taboka Prince for experience they gave me.

I would like to thank the, all the ladies and gentlemen at AIMS, my former roommate Katlego Lentswe, AIMS soccer team, January Intake soccer squad, Finhai,Tee(game), team Long Street Cafe, my twin Sebenele Tshwala, my AIMS friends Blessing Khumalo, Chabo Bobe, Irene Claudia Noharinaivo, Veronica Mtonga, Banan Mozher Mohammed, Samah Mohammed, Ahlam Abdullah Adam, Ramatsimela Lebogang Mphahlele, Zizipho Maqokolo,Tshenolo Thato Daumas, Fezile Mhlabane, Tamara Tembo, Sahada Larabu, Bouchra Er-Rabbany, Marwa Musa Babiker Adam, Aulan Lucrèce Zahoundo, Nedjla Tazdait and Refiloe Shabe.

I would also like to thank my friends back home Inge Mazibuko, Peter Motse, Bongani Ngcobo, Pule Mathews, Thabiso Toka Moshoaliba, Samson Mofokeng, Boteng Mofokeng, Lefa Thapelo Moeketsi, Clive Dlamini, Sandile Nhlapo, Peter Mhlongo and my friends from Wits University Nokuzola Mkhatshwa, Thandile Tonisi, Thato Molelekeng, Tshepang More, Katlego Motsitsi, Damon Moodley, Lungile Mhlongo, Godfrey Tshehla and Martin Masilo.

Finally, I would like to thank the AIMS family, especially Mr. Jan Groenewald for his charisma 'John Wick', our current AIMS academic director Dr. Simukai Utete, AIMS director Prof. Barry Green, former AIMS academic director Prof. Jeff Sanders, IT staff, kitchen staff, cleaning staff,maintenance staff and Igsaan Kamile for all the amazing work they put on daily basis.

References

Babbs, S. H. and Nowman, K. B. Kalman filtering of generalized vasicek term structure models. *Journal of Financial and Quantitative Analysis*, 34(1):115–130, 1999.

Benninga, S. and Wiener, Z. Term structure of interest rates. *Mathematica in Education and Research*, 7:13–21, 1998.

Bergstrom, A. R. Continuous time stochastic models and issues of aggregation over time. *Handbook of econometrics*, 2:1145–1212, 1984.

Bolder, D. J. Affine term-structure models: Theory and implementation. *Available at SSRN 1082826*, 2001.

Bulut, Y. and Bayat, O. Kalman filtering with model uncertainties. volume 5, 03 2012. doi: 10.1007/978-1-4614-2425-3_42.

De Jong, F. Time-series and cross-section information in affine term structure models. *Available at SSRN 154681*, 1999.

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

Focardi, S. M. and Fabozzi, F. J. *The mathematics of financial modeling and investment management*, volume 138. John Wiley & Sons, 2004.

Geyer, A. and Frühwirth&Schnatter, S. Bayesian estimation of econometric multi factor cox ingersoll ross models of the term structure of interest rates via mcmc methods. *Unpublished Working Paper*, 1996.

Harvey, A. C. *Forecasting, structural time series models and the Kalman filter*. Cambridge university press, 1990.

Hull, J. and White, A. One-factor interest-rate models and the valuation of interest-rate derivative securities. *Journal of financial and quantitative analysis*, 28(2):235–254, 1993.

Ibhagui, O. Application of teh kalman filter to interest rate modelling. 2010. http://archive.aims.ac.za/postgraduate-diploma-essays/2009-10/wallace.pdf.

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

Kheng, L. W. Matrix differentiation. *Matrix*, 1:36. https://atmos.washington.edu/~dennis/MatrixCalculus.pdf.

Lacey, T. Tutorial: The kalman filter. 2019. http://web.mit.edu/kirtley/kirtley/binlustuff/literature/control/Kalman%20filter.pdf.

Lund, J. *Econometric analysis of continuous-time arbitrage-free models of the term structure of interest rates*. Department of Finance, Faculty of Business Administration, Aarhus School of . . . , 1997.

Mastro, M. *Financial Derivative and Energi Market Valuation*. Wiley Online Library, 2013.

Simon, D. *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

Zivot, E. State space models and the kalman filter. *Apr*, 9:1–8, 2006. https://faculty.washington.edu/ezivot/econ584/notes/statespacemodels.pdf.