A Hidden Markov Multi-Assets Price Model

Robert J. Elliott
Haskayne School of Business
University of Calgary
E-mail: relliott@ucalgary.ca

Tao Lin
Norwegian School of Economics
and Business Administration
tao.lin@nhh.no

Hong Miao
Haskayne School of Business
University of Calgary
E-mail: hmiao@ucalgary.ca

August 28, 2007

Abstract

Prices for some real world tradable assets, for instance natural gas and oil, are correlated, and the price dynamics for those assets are different in different economic environments. In this paper we extend the mean reverting model to multi-assets and model correlation between prices. Our model also allows the means and the mean reverting factors to switch between different regimes by including a Hidden Markov chain which models the different economic environments, or "states of the world". We then obtain approximate estimates for the parameters by applying filters and the EM algorithm. Approximate derivative prices are also given.

Keywords: Mean Reverting, Commodity Markets, Regime Switching, Hidden Markov Chain, EM Algorithm

1 Introduction

After the publication of the Black and Scholes (1973)[1] paper on option pricing, Geometric Brownian Motion has become a standard tool for modelling financial assets. However, over the last two decades, much attention has been paid to the modeling of commodities, and researchers have found that commodity markets are different from financial markets because of physical constraints. The price dynamics for commodities clearly exhibit strong mean reversion and seasonality characteristics. Therefore, mean reverting models are often used to describe the behavior of asset prices, particularly commodity prices.

Some work has been done by combining regime switching and mean reversion into one model. Elliott et al. (1999)[4] developed a mean reverting and regime switching model by allowing the long term means and the mean reverting factors to switch between different states. Wu and Elliott (2005)[6] proposed a regime switching model with mean reversion and jumps and estimate all the parameters using the Expectation Maximization (EM) algorithm.

It is well known that the prices of some commodities are highly correlated, for instance natural gas and crude oil. However, previous models consider only the spot prices for a particular commodity, instead of modeling the spot prices of a group of correlated commodities together. To integrate the correlations between the prices of multi-assets, we generalize one factor mean reverting models to a multi-asset situation by including correlations between the prices of assets. Moreover, it is known that price dynamics differ under different economic environments. We introduce a hidden Markov chain to represent the "states of the world", or of the economy. We derive approximate expressions for all the parameters by applying the filtering techniques and the EM algorithm. We also obtain approximate formulas for derivative prices.

The paper is organized as follows. In section two, we give the spot price dynamics. We derive approximate expressions for forward and option prices in section three. All the parameters are estimated by applying the EM algorithm in section four. Section five concludes the paper and proposes further research topics.

The authors would like to thank the anonymous referee for his/her helpful comments. Robert Elliott is graterful for the financial support from the Social Sciences and Humanities Research Council.

2 Spot Price Dynamics

Prices of tradable assets are often highly correlated. Therefore, we shall model the prices of a number of assets together.

Suppose we work on a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$, where $\mathcal{P}$ is a real-world probability measure. We wish to price Q correlated assets, for instance natural gas, futures on natural gas, crude oil, and futures on crude oil. It is clear there are correlations between all these prices. Let the price process for the asset i be $S_i = S_i(t)$, for $0 \leq t \leq T$, $1 \leq i \leq Q$.

We suppose the logarithm of the price of the *i-th* asset, $R_i(t) = \log S_i(t)$ follows a mean-reverting diffusion process,

$$dR_i(t) = \alpha_i(t)(\beta_i(t) - R_i(t))dt + \sum_{j=1}^{Q} \sigma_{i,j} dB_j(t). \tag{1}$$

Here $\alpha_i(t)$ is a mean reverting factor, $\beta_i(t)$ is the long-term equilibrium mean for the i-th asset, and $B_j, (1 \leq j \leq Q)$, are Q independent Brownian motions. Therefore, the price of the i-th individual asset is described by a mean reverting process and also depends on its correlations with the other assets. Since we observe mean reverting behavior in energy markets, and also the mean reverting behavior may differ in different economics, we introduce a Markov Chain $\{X_t, 0 \leq t \leq T\}$ with a finite state space $\mathbf{B}$ to represent the different "states of the world". Without loss of generality we can identify $\mathbf{B}$ with the set of unit vectors:

$$S = \{e_1, e_2, ..., e_N\}, \qquad \text{where } e_i = (0, ..., 1..., 0)' \in \mathbb{R}^N.$$

Suppose X has a transition rate matrix $\tilde{A} = p_{ji} \in \mathbb{R}^{N \times N}$. Let $\mathcal{F}_T^X$ be the σ-algebra generated by the hidden Markov Chain process up to time T, that is, $\mathcal{F}_T^X = \sigma\{X_s, 0 \leq s \leq T\}$. As shown in Elliott et al. (1993)[3], we can write the dynamics of X_t as

$$X_t = X_0 + \int_0^t \tilde{A} X_u du + M_t,$$

where $M_t \in R^N$ and $E[M_t | \mathcal{F}_s^X] = M_s \in \mathbb{R}^N$, so M is a martingale. Of course, X defines the martingale M and in turn M determines the dynamics of X. We assume the $\alpha_i(t)$ and $\beta_i(t)$ for all $1 \leq i \leq Q$ take different values in different "states of the world". Therefore, we let the $(\alpha_i(t), \beta_i(t))$ take values in a finite set $B = \{(\alpha_i^n, \beta_i^n) : 1 \leq i \leq Q, 1 \leq n \leq N\}$, that is,

$$\alpha_i = (\alpha_i^1, \alpha_i^2, \ldots, \alpha_i^N)',$$
$$\beta_i = (\beta_i^1, \beta_i^2, \ldots, \beta_i^N)'.$$

Then we suppose,

$$\alpha_i(t) = \langle \alpha_i, X_t \rangle,$$
$$\beta_i(t) = \langle \beta_i, X_t \rangle.$$

The solution to (1) is:

$$R_i(t) = e^{-a_i(t)} \left[R_i(0) + \int_0^t e^{a_i(s)} \alpha_i(s) \beta_i(s) ds + \sum_{j=1}^{Q} \sigma_{i,j} \int_0^t e^{a_i(s)} dB_j(s) \right], \tag{2}$$

where we write $a_i(t) = \int_0^t \alpha_i(s) ds$, $0 \leq t \leq T$.

If $\mathcal{F}_T^X$ and the initial value $R_i(0)$ are given, $R_i(T)$ is a Gaussian random variable with conditional mean

$$\begin{aligned} \mu_i(T) &= E[R_i(T) | \mathcal{F}_T^X \vee R_i(0)] \\ &= e^{-a_i(T)} \left[R_i(0) + \int_0^T e^{a_i(s)} \alpha_i(s) \beta_i(s) ds \right], \end{aligned}$$

and variance

$$\begin{aligned} \rho_i(T) &= Var[R_i(T) | \mathcal{F}_T^X \vee R_i(0)] \\ &= E\left[(R_i(T) - E[R_i(T) | \mathcal{F}_T^X \vee R_i(0)])^2 \right] \\ &= E\left[\left[e^{-a_i(T)} \sum_{j=1}^{Q} \sigma_{i,j} \int_0^T e^{a_i(s)} dB_j(s) \right]^2 | \mathcal{F}_T^X \vee R_i(0) \right] \\ &= e^{-2a_i(T)} \sum_{j=1}^{Q} \sigma_{i,j}^2 \int_0^T e^{2a_i(s)} ds. \end{aligned}$$

Although we have the expressions for both the conditional mean and the variance, they involve the behavior of X over the time interval $[0, T]$.

We have to estimate the terms $e^{-2a_i(T)}$, $\int_0^T e^{a_i(s)}\alpha_i(s)\beta_i(s)ds$, and $\int_0^T e^{2a_i(s)}ds$ to obtain estimates for the values of the conditional mean and variance. Then we can estimate the statistical behavior of the spot price dynamics and price forwards and options. We first introduce the following *Lemma* which simplifies the expressions. The lemma connects the mean reverting factors with the occupation time, the amount of time that the Markov chain spends in a state up to a point of time t.

Lemma 1 *Write the occupation time for X_t in state j up to time t as $\mathcal{O}_t^j = \int_0^t \langle X_s, e_j\rangle ds$. Then $a_i(t) = \int_0^t \langle \alpha_i, X_s\rangle ds = \langle \alpha_i, \mathcal{O}_t\rangle$, where $\mathcal{O}_t = (\mathcal{O}_t^1, \mathcal{O}_t^2, \ldots, \mathcal{O}_t^N)'$.*

Proof. By definition

$$X_s = \sum_{j=1}^{N} \langle X_s, e_j\rangle e_j.$$

Then

$$a_i(t) = \int_0^t \langle \alpha_i, X_s\rangle ds = \sum_{j=1}^{N} \int_0^t \langle \alpha_i, e_j\rangle\langle X_s, e_j\rangle ds = \langle \alpha_i, \mathcal{O}_t\rangle.$$

$\square$

From *Lemma 1*,

$$\mu_i(T) = e^{\langle -\alpha_i, \mathcal{O}_T\rangle}\left[R_i(0) + \int_0^t e^{\langle \alpha_i, \mathcal{O}_s\rangle}\langle \alpha_i \odot \beta_i, X_s\rangle ds\right], \tag{3}$$

and

$$\rho_i(T) = e^{\langle -2\alpha_i, \mathcal{O}_T\rangle}\sum_{j=1}^{Q}\sigma_{i,j}^2 \int_0^T e^{\langle 2\alpha_i, \mathcal{O}_s\rangle}ds. \tag{4}$$

Here $\alpha_i \odot \beta_i = (\alpha_i^1\beta_i^1, \alpha_i^2\beta_i^2, \ldots, \alpha_i^N\beta_i^N)'$. The only remaining random terms are terms related to $e^{\langle \alpha_i, \mathcal{O}_t\rangle}$. The next result gives a general expression for calculating these.

Lemma 2 *For $u = (u_1, u_2, \ldots, u_N) \in \mathbb{R}^N$, write $D_u = diag(u_1, u_2, \ldots, u_N)$, and define $Z(t) = e^{\langle u, \mathcal{O}_t\rangle}X_t$. Then*

$$E[Z(t)|\mathcal{F}_s^X] = e^{(\tilde{A}+D_u)(t-s)}Z(s).$$

Proof. Consider

$$Z(t) = e^{\langle u, \mathcal{O}_t \rangle} X_t, \tag{5}$$

where, as before,

$$\begin{aligned}
\mathcal{O}_t =& (\mathcal{O}_t^1, \mathcal{O}_t^2, \dots, \mathcal{O}_t^N)' \\
=& (\int_0^t \langle X_s, e_1 \rangle ds, \int_0^t \langle X_s, e_2 \rangle ds, \dots, \int_0^t \langle X_s, e_N \rangle ds)'.
\end{aligned}$$

Differentiating equation (5), we obtain

$$\begin{aligned}
dZ(t) =& e^{\langle u, \mathcal{O}_t \rangle} dX_t + X_t de^{\langle u, \mathcal{O}_t \rangle} \\
=& (\tilde{A} X_t dt + dM(t)) e^{\langle u, \mathcal{O}_t \rangle} \\
& + X_t e^{\langle u, \mathcal{O}_t \rangle} \left(u_1 \langle X_t, e_1 \rangle dt + \dots + u_N \langle X_t, e_N \rangle \right) dt.
\end{aligned}$$

That is,

$$\begin{aligned}
Z(t) =& Z(s) + \int_s^t \tilde{A} X_r e^{\langle u, \mathcal{O}_r \rangle} dr + \int_s^t e^{\langle u, \mathcal{O}_r \rangle} dM(r) \\
& + \int_s^t X_r e^{\langle u, \mathcal{O}_r \rangle} (u_1 \langle X_r, e_1 \rangle + \dots + u_N \langle X_r, e_N \rangle) dr.
\end{aligned}$$

Because the integral with respect to M is a martingale,

$$\begin{aligned}
& E[Z(t) | \mathcal{F}_s^X] \\
& = Z(s) + \int_s^t \tilde{A} E[Z(r) | \mathcal{F}_s^X] dr + \int_s^t diag(u_1, \dots, u_N) E[Z(r) | \mathcal{F}_s^X] dr.
\end{aligned}$$

Writing

$$E[Z(t) | \mathcal{F}_s^X] = \gamma_t \in \mathbb{R}^N,$$

the previous equation is

$$\gamma_t = Z(s) + \int_s^t (\tilde{A} + D_u) \gamma_r dr, \tag{6}$$

where $D_u = diag(u_1, u_2, \dots, u_N)$.

Therefore,

$$\gamma_t = e^{(\tilde{A} + D_u)(t-s)} Z(s). \in \mathbb{R}^N,$$

and,

$$E[Z(t)|\mathcal{F}_s^X] = e^{(\tilde{A}+D_u)(t-s)}Z(s).$$

$\square$

By applying this *Lemma*, we can estimate the means and variances for individual assets after we obtain the estimates for the parameters in section 4. In the next section we estimate prices of forwards and options on those assets. Since forwards and options on commodities are frequently traded, these expressions are useful.

3 Derivative Prices

In energy markets, forward contracts, futures and various types of options are traded daily. The introduction of future contracts and the mechanism of marking to market have been considered two of the most important breakthroughs in modern finance. Although there are future contracts whose safety is guaranteed by the clearing houses, forward contracts are still heavily traded in over-the-counter markets. Below, we shall estimate prices of forwards and options whose underling is an asset whose spot price follows the dynamics given in the previous section.

3.1 Forward Prices

Forwards are particularly important in energy markets as prices are volatile. An oil refiner may enter into a forward contract to secure crude oil for its future operations and so to avoid both volatility in spot oil prices and the need to store oil for extended periods. Theoretically speaking, the price for a forward contract is just the expectation of the spot price at a fixed future date. However, it is not easy to price a forward contract when the correlations between the related assets are taken into consideration as the computation becomes very complicated and usually there are not closed form expressions.

Suppose we want to price a forward contract whose underlying is the i-th asset. The forward price, for $0 \leq t \leq T$, is then

$$F_i(T,t) = E[S_i(T)|R_i(0)].$$

That is, the value of the forward is the expectation of the spot price at time T conditional on the initial price, which is usually observable at the time when the forward contract is agreed. As, given $\mathcal{F}_T^X$ and the initial value $R_i(0)$, $R_i(T)$ is conditionally Gaussian:

$$\begin{aligned}F_i(T,t) =&E[e^{R_i(T)}|R_i(0)]\\ =&E[E[e^{R_i(T)}|\mathcal{F}_T^X \vee R_i(0)]|R_i(0)]\\ =&E[exp(\mu_i(T) - \frac{1}{2}\rho_i(T))|R_i(0)].\end{aligned}$$

Writing $K = \mu_i(T) - \frac{1}{2}\rho_i(T)$, and $\bar{K} = E[\mu_i(T) - \frac{1}{2}\rho_i(T)|R_i(0) \vee X_0]$, this becomes $F_i(T,t) = E[e^K|R_i\,(0) \vee X_0]$. Although we can not evaluate this exactly, we can obtain an approximation using a Taylor expansion around $\bar{K}$.

$$\begin{aligned}E[e^K|R_i\,(0) \vee X_0] &= E[e^{\bar{K}}e^{K-\bar{K}}|R_i\,(0) \vee X_0]\\ &= e^{\bar{K}}E\left[1 + (K-\bar{K}) + \frac{(K-\bar{K})^2}{2!} + o(K-\bar{K})^2|R_i\,(0) \vee X_0\right]\\ &\approx e^{\bar{K}}\left(1 + \frac{E\left[(K-\bar{K})^2|R_i\,(0) \vee X_0\right]}{2!}\right). \qquad (7)\end{aligned}$$

Now we have to compute $e^{\bar{K}}$ and $E\left[(K-\bar{K})^2|R_i\,(0) \vee X_0\right]$. Since $E[(K-\bar{K})^2|R_i\,(0) \vee X_0] = E[K^2|R_i\,(0) \vee X_0] - \bar{K}^2$ we shall calculate the following terms: $\bar{K} = E\,[K|R_i\,(0) \vee X_0]$, $E[K^2|R_i\,(0) \vee X_0]$, and $\bar{K}^2$. The following lemmas give the values of these terms.

Lemma 3 *The value of* $\bar{K} = E\,[K|R_i\,(0) \vee X_0] = E[\mu_i(T) - \frac{1}{2}\rho_i(T)|R_i\,(0) \vee X_0]$ *is given by:*

$$\begin{aligned}\bar{K} =& \langle e^{(\tilde{A}-D_{\alpha_i})T}X_0, \underline{1}\rangle R_i(0)\\ &+ \int_0^T \langle (e^{(\tilde{A}'-D_{\alpha_i})(T-s)}\underline{1}) \odot \alpha_i \odot \beta_i, e^{\tilde{A}s}X_0\rangle ds\\ &- \frac{1}{2}\sum_{j=1}^{Q}\sigma_{i,j}^2 \int_o^T \langle e^{(\tilde{A}'-2D_{\alpha_i})(T-s)}e^{\tilde{A}'s}\underline{1}, X_0\rangle ds.\end{aligned}$$

Proof. We first notice that

$$
\begin{aligned}
K &= \mu_i(T) - \frac{1}{2}\rho_i(T) \\
&= e^{-a_i(T)}R_i(0) + \int_0^T e^{a_i(s)-a_i(T)}\alpha_i(s)\beta_i(s)ds - \frac{1}{2}\sum_{j=1}^{Q}\sigma_{i,j}^2\int_0^T e^{2(a_i(s)-a_i(T))}ds.
\end{aligned}
$$

To simplify the expression we write

$$I_1 = e^{-\langle \alpha_i, \mathcal{O}_T\rangle}R_i(0), \tag{8}$$

$$I_2 = \int_0^T e^{\langle \alpha_i, \mathcal{O}_s - \mathcal{O}_T\rangle}\langle \alpha_i \odot \beta_i, X_s\rangle ds, \tag{9}$$

and

$$I_3 = -\frac{1}{2}\sum_{j=1}^{Q}\sigma_{i,j}^2\int_o^T e^{2\langle \alpha_i, \mathcal{O}_s - \mathcal{O}_T\rangle}ds. \tag{10}$$

Then $K = I_1 + I_2 + I_3$, and

$$\bar{K} = E\left[K|R_i\left(0\right) \vee X_0\right] = E\left[I_1|R_i\left(0\right) \vee X_0\right] + E\left[I_2|R_i\left(0\right) \vee X_0\right] + E\left[I_3|R_i\left(0\right) \vee X_0\right].$$

From *Lemma 2* we know $E[Z(t)|R_i\left(0\right) \vee X_0] = e^{(\tilde{A}+D_u)t}X_0$, for a general process $Z(t) = e^{\langle u, \mathcal{O}_t\rangle}X_t$, and $s = 0$, therefore,

$$
\begin{aligned}
&E[\langle Z(t), \underline{1}\rangle|R_i\left(0\right) \vee X_0] \\
&= \langle E[Z(t)|R_i\left(0\right) \vee X_0], \underline{1}\rangle \\
&= \langle e^{(\tilde{A}+D_u)t}X_0, \underline{1}\rangle \\
&= E[e^{\langle u, \mathcal{O}_t\rangle}\langle X_t, \underline{1}\rangle|R_i\left(0\right) \vee X_0].
\end{aligned}
$$

Here, $\underline{1} = (1, 1, 1, \cdots, 1)'$. That is,

$$E[e^{\langle u, \mathcal{O}_t\rangle}|R_i\left(0\right) \vee X_0] = \langle e^{(\tilde{A}+D_u)t}X_0, \underline{1}\rangle.$$

Thus, we have

$$E[I_1|R_i\left(0\right) \vee X_0] = \langle e^{(\tilde{A}-D_{\alpha_i})T}X_0, \underline{1}\rangle R_i(0),$$

where $D_{\alpha_i} = diag(\alpha_i)$.

Now we turn to the computation of $E[I_2|R_i(0) \vee X_0]$:

$$E[I_2|R_i(0) \vee X_0] = \int_0^T E\left[E\left[e^{\langle \alpha_i, \mathcal{O}_s - \mathcal{O}_T \rangle} \langle \alpha_i \odot \beta_i, X_s \rangle | \mathcal{F}_s^X\right] | R_i(0) \vee X_0\right] ds$$
$$= \int_0^T E\left[E\left[e^{-\langle \alpha_i, \mathcal{O}_T \rangle} | \mathcal{F}_s^X\right] e^{\langle \alpha_i, \mathcal{O}_s \rangle} \langle \alpha_i \odot \beta_i, X_s \rangle | R_i(0) \vee X_0\right] ds.$$

Notice that

$$E[e^{-\langle \alpha, \mathcal{O}_T \rangle} X_T | \mathcal{F}_s^X] = e^{(\tilde{A} + D_{\alpha_i})(T-s)} e^{-\langle \alpha, \mathcal{O}_s \rangle} X_s,$$

so,

$$E[e^{-\langle \alpha, \mathcal{O}_T \rangle} | \mathcal{F}_s^X] = e^{-\langle \alpha, \mathcal{O}_s \rangle} \langle e^{(\tilde{A} + D_{\alpha_i})(T-s)} X_s, \underline{1} \rangle$$
$$= e^{-\langle \alpha, \mathcal{O}_s \rangle} \langle e^{(\tilde{A}' + D_{\alpha_i})(T-s)} \underline{1}, X_s \rangle.$$

Thus, we obtain

$$E[I_2|R_i(0) \vee X_0] = \int_0^T E\left[e^{-\langle \alpha_i, \mathcal{O}_s \rangle} \langle e^{(\tilde{A}' + D_{\alpha_i})(T-s)} \underline{1}, X_s \rangle e^{\langle \alpha_i, \mathcal{O}_s \rangle} \langle \alpha_i \odot \beta_i, X_s \rangle | R_i(0) \vee X_0\right] ds$$
$$= \int_0^T \langle (e^{(\tilde{A}' - D_{\alpha_i})(T-s)} \underline{1}) \odot \alpha_i \odot \beta_i, e^{\tilde{A}s} X_0 \rangle ds.$$

Similarly,

$$E[I_3|R_i(0) \vee X_0] = E\left[-\frac{1}{2} \sum_{j=1}^{Q} \sigma_{i,j}^2 \int_o^T e^{2\langle \alpha_i, \mathcal{O}_s - \mathcal{O}_T \rangle} ds | R_i(0) \vee X_0\right]$$
$$= -\frac{1}{2} \sum_{j=1}^{Q} \sigma_{i,j}^2 \int_o^T \langle e^{(\tilde{A}' - 2D_{\alpha_i})(T-s)} \underline{1}, e^{\tilde{A}s} X_0 \rangle ds.$$

Then finally,

$$\bar{K} = \langle e^{(\tilde{A} - D_{\alpha_i})T} X_0, \underline{1} \rangle R_i(0)$$
$$+ \int_0^T \langle (e^{(\tilde{A}' - D_{\alpha_i})(T-s)} \underline{1}) \odot \alpha_i \odot \beta_i, e^{\tilde{A}s} X_0 \rangle ds$$

$$-\frac{1}{2}\sum_{j=1}^{Q}\sigma_{i,j}^2\int_o^T\langle e^{(\tilde{A}'-2D_{\alpha_i})(T-s)}e^{\tilde{A}'s}\underline{1},X_0\rangle ds.$$

□

Lemma 3 gives us the expression of $\bar{K}$, so the computation of $e^{\bar{K}}$ and $\bar{K}^2$ is just straightforward. The only remaining term is $E[K^2|R_i(0)\vee X_0]$.

Lemma 4 *The value of* $E[K^2|R_i(0)\vee X_0]$, *conditional on the initial value of the Markov Chain, is given by*

$$\begin{aligned}
&E[K^2|R_i(0)\vee X_0]\\
&=\langle e^{(\tilde{A}-2D_{\alpha_i})T}X_0,\underline{1}\rangle R(0)^2\\
&\quad+2\int_0^T\int_0^s\langle e^{(\tilde{A}'-D_{\alpha_i})(s-r)}[e^{(\tilde{A}'-2D_{\alpha_i})(T-s)}\underline{1}\odot\alpha_i\odot\beta_i]\odot\alpha_i\odot\beta_i,e^{\tilde{A}r}X_0\rangle drds\\
&\quad+\frac{1}{2}(\sum_{j=1}^{Q}\sigma_{i,j}^2)^2\int_0^T\int_0^s\langle e^{(\tilde{A}'-4D_{\alpha_i})(T-s)}\underline{1},e^{(\tilde{A}-2D_{\alpha_i})(s-r)}e^{\tilde{A}r}X_0\rangle drds\\
&\quad+2R(0)\int_0^T\langle(e^{(\tilde{A}'-2D_{\alpha_i})(T-s)}\underline{1})\odot\alpha_i\odot\beta_i,e^{(\tilde{A}-D_{\alpha_i})s}X_0\rangle ds\\
&\quad-\left(\sum_{j+1}^{Q}\sigma_{i,j}^2\right)R(0)\int_0^T\langle e^{(\tilde{A}'-3D_{\alpha_i})(T-s)}e^{(\tilde{A}-2D_{\alpha_i})s}X_0,\underline{1}\rangle ds\\
&\quad-\left(\sum_{j+1}^{Q}\sigma_{i,j}^2\right)\int_0^T\int_0^s\langle e^{(\tilde{A}'-D_{\alpha_i})(s-r)}[e^{(\tilde{A}'-3D_{\alpha_i})(T-s)}\underline{1}]\odot\alpha_i\odot\beta_i,e^{\tilde{A}r}X_0\rangle drds\\
&\quad-\left(\sum_{j+1}^{Q}\sigma_{i,j}^2\right)\int_0^T\int_0^s\langle e^{(\tilde{A}'-2D_{\alpha_i})(s-r)}[e^{(\tilde{A}'-3D_{\alpha_i})(T-s)}\underline{1}]\odot\alpha_i\odot\beta_i,e^{\tilde{A}r}X_0\rangle drds.
\end{aligned}$$

Proof. We have

$$K=I_1+I_2+I_3,$$

where I_1, I_2, and I_3 are defined in equation $(8-10)$, respectively. Thus

$$K^2=(I_1+I_2+I_3)^2=I_1^2+I_2^2+I_3^2+2I_1I_2+2I_1I_3+2I_2I_3.$$

We write the six terms in the formula as M_i, $(1\leq i\leq 6)$, therefore,

$$K^2=M_1+M_2+M_3+M_4+M_5+M_6.$$

Now

$$E[K^2|R_i\,(0)\vee X_0] = E[M_1|R_i\,(0)\vee X_0] + E[M_2|R_i\,(0)\vee X_0] + E[M_3|R_i\,(0)\vee X_0] \\ + E[M_4|R_i\,(0)\vee X_0] + E[M_5|R_i\,(0)\vee X_0] + E[M_6|R_i\,(0)\vee X_0].$$

All these six terms can be computed explicitly as follows:

$$E[M_1|R_i\,(0)\vee X_0] = \langle e^{(\tilde{A}-2D_{\alpha_i})T}X_0, \underline{1}\rangle R(0)^2.$$

$$E[M_2|R_i\,(0)\vee X_0] = 2\int_0^T\int_0^s \langle e^{(\tilde{A}'-D_{\alpha_i})(s-r)}[e^{(\tilde{A}'-2D_{\alpha_i})(T-s)}\underline{1}\odot\alpha_i\odot\beta_i]\odot\alpha_i\odot\beta_i, e^{\tilde{A}r}X_0\rangle drds.$$

$$E[M_3|R_i\,(0)\vee X_0] = \frac{1}{2}(\sum_{j=1}^{Q}\sigma_{i,j}^2)^2\int_0^T\int_0^s \langle e^{(\tilde{A}'-4D_{\alpha_i})(T-s)}\underline{1}, e^{(\tilde{A}-2D_{\alpha_i})(s-r)}e^{\tilde{A}r}X_0\rangle drds.$$

$$E[M_4|R_i\,(0)\vee X_0] = 2R(0)\int_0^T \langle (e^{(\tilde{A}'-2D_{\alpha_i})(T-s)}\underline{1})\odot\alpha_i\odot\beta_i, e^{(\tilde{A}-D_{\alpha_i})s}X_0\rangle ds.$$

$$E[M_5|R_i\,(0)\vee X_0] = -\left(\sum_{j+1}^{Q}\sigma_{i,j}^2\right)R(0)\int_0^T \langle e^{(\tilde{A}'-3D_{\alpha_i})(T-s)}e^{(\tilde{A}-2D_{\alpha_i})s}X_0, \underline{1}\rangle ds.$$

and,

$$E\,[M_6|R_i\,(0)\vee X_0] \\ = -\left(\sum_{j+1}^{Q}\sigma_{i,j}^2\right)\int_0^T\int_0^s \langle e^{(\tilde{A}'-D_{\alpha_i})(s-r)}[e^{(\tilde{A}'-3D_{\alpha_i})(T-s)}\underline{1}]\odot\alpha_i\odot\beta_i, e^{\tilde{A}r}X_0\rangle drds \\ -\left(\sum_{j+1}^{Q}\sigma_{i,j}^2\right)\int_0^T\int_0^s \langle e^{(\tilde{A}'-2D_{\alpha_i})(s-r)}[e^{(\tilde{A}'-3D_{\alpha_i})(T-s)}\underline{1}]\odot\alpha_i\odot\beta_i, e^{\tilde{A}r}X_0\rangle drds.$$

See the Appendix A for details of the computation of $E[M_2|R_i\,(0)\vee X_0]$, and the computations of the others are similar. □

Lemmas 3 and *4* give expressions for all the terms we need to obtain an approximate forward price as:

$$F_i(T,t) \approx e^{\bar{K}}\left(1 + \frac{E\left[(K-\bar{K})^2\right]}{2!}\right). \tag{11}$$

3.2 Option Prices

The price $C(S_i(T))$ of a European call on the i-th asset is given by

$$\begin{aligned} C(S_i, t, T) &= E[(S_i(T) - K_i)^+ | X_0] \\ &= E[E[(S_i(T) - K_i)^+ | \mathcal{F}_T^X] | X_0], \end{aligned}$$

where K_i is the exercise price. Knowing the history $\mathcal{F}_T^X$ of X, the inner expectation is a European call written on a mean-reverting asset S_i which is log-normally distributed with the mean (3) and the variance (4). Therefore, the Black-Scholes type option pricing formula can be applied.

Clewlow and Strickland (1999)[2] proposed the following formula for the price of a European option:

$$E[(S_i(T) - K_i)^+ | \mathcal{F}_T^X] = e^{-r(T-t)}[F_i\,(t,T)\,N(h_i) - KN(h_i - \sqrt{\rho_i})],$$

Here $F_i(t,T)$ is the forward price on asset $i, 1 \leq i \leq Q$ at time t; and r is the risk-free interest rate,

$$h_i = \frac{ln(F_i(t,T)/K_i) + \frac{1}{2}\rho_i}{\sqrt{\rho_i}}, \tag{12}$$

and

$$\rho_i = e^{-2a_i(T)} \sum_{j=1}^{Q} \sigma_{i,j}^2 \int_0^T e^{2a_i(s)} ds.$$

Therefore, using our estimates in our switching model the price of a European option is approximately:

$$\begin{aligned} C(S_i, t, T) &= E_Q[E_Q[C_{CH}(S_i, t, T) | X_0]] \\ &= e^{-r(T-t)} E_Q[F_i(t,T) N(h_i) - K_i N(h_i - \sqrt{\rho_i}) | X_0] \\ &= e^{-r(T-t)} \left\{ F_i(t,T) E[N(h_i) | X_0] - K_i E_Q[N(h_i - \sqrt{\rho_i}) | X_0] \right\}. \end{aligned} \tag{13}$$

Therefore, we must estimate $E_Q\left[N\left(h_i\right)|X_0\right]$, and $E_Q[N(h_i - \sqrt{\rho_i})|X_0]$. To do this we use the approximations $N\left(E\left[h_i|X_0\right] - \sqrt{E[\rho_i|X_0]}\right)$ for $E_Q[N(h_i - \sqrt{\rho_i})|X_0]$, and $N\left(E_Q\left[h_i|X_0\right]\right)$ for $E_Q\left[N\left(h_i\right)|X_0\right]$. From equation (12), we have

$$E_Q\left[h_i|X_0\right] = E_Q\left[\frac{ln(F_i(t,T)/K_i) + \frac{1}{2}\rho_i}{\sqrt{\rho_i}}|X_0\right]$$
$$\simeq \frac{ln(F_i(t,T)/K_i)}{\sqrt{E[\rho_i|X_0]}} + \frac{1}{2}\sqrt{E_Q[\rho_i|X_0]}.$$

We notice that $E_Q[\rho_i|X_0]$ can be computed as following,

$$E_Q[\rho_i|X_0] = E_Q\left[e^{-2\langle\alpha_i,\mathcal{O}_T\rangle}\sum_{j=1}^{Q}\sigma_{i,j}^2\int_0^T e^{2\langle\alpha_i,\mathcal{O}_s\rangle}ds|X_0\right]$$
$$= \sum_{j=1}^{Q}\sigma_{i,j}^2\int_0^T E_Q[E_Q[e^{-2\langle\alpha_i,\mathcal{O}_T\rangle}|\mathcal{F}_s^X]e^{2\langle\alpha_i,\mathcal{O}_s\rangle}|X_0]ds.$$

By applying *Lemma 2*, this becomes

$$= \sum_{j=1}^{Q}\sigma_{i,j}^2\int_0^T E_Q[e^{-2\langle\alpha_i,\mathcal{O}_s\rangle}\langle e^{(\tilde{A}-2D_{\alpha_i})(T-s)}X_s,\underline{1}\rangle e^{2\langle\alpha_i,\mathcal{O}_s\rangle}|X_0]ds$$
$$= \sum_{j=1}^{Q}\sigma_{i,j}^2\int_0^T \langle e^{(\tilde{A}-2D_{\alpha_i})(T-s)}e^{\tilde{A}s}X_0,\underline{1}\rangle]ds.$$

Inserting these approximations in the equation (13), we can estimate the price of an European option.

4 Filtering and Estimation

In the previous sections, we have derived closed form expressions for forward prices and approximate expressions for option prices. In this section, we shall use filters and the EM algorithm to estimate the parameters in the spot price dynamics from historical data. We only observe the dynamics of spot prices, but not the hidden Markov Chain, X_t. We suppose the logarithm of returns of the Q assets follow the dynamics (1) .

Observing the discrete time spot prices, we use the following discrete time forms for the dynamics.

We suppose that X has discrete time dynamics

$$X_{k+1} = AX_k + M_{k+1}, \quad \text{with } A = a_{ji} \in \mathbb{R}^{N \times N}.$$

and R follows:

$$R(k+1) = \Gamma(k+1) + \Upsilon(k+1)R(k) + \Omega W(k+1).$$

Here

$$\begin{aligned}
A &= I + \tilde{A}, \\
R(k) &= (R_1(k), R_2(k), \ldots, R_Q(k))' \in \mathbb{R}^Q, \\
\Gamma(k+1) &= (\gamma_1, \gamma_2, \ldots, \gamma_Q)' \\
&= (\langle (\alpha\beta)_1, X_k \rangle, \ldots, \langle (\alpha\beta)_Q, X_k \rangle)' \in \mathbb{R}^Q, \\
(\alpha\beta)_i &= (\alpha_i^1 \beta_i^1, \alpha_i^2 \beta_i^2, \ldots, \alpha_i^N \beta_i^N)' \in \mathbb{R}^N, \\
\Upsilon(k+1) &= diag(\upsilon_1, \upsilon_2, \ldots, \upsilon_Q), \\
&= diag(1 - \langle \alpha_1, X_k \rangle, \ldots, 1 - \langle \alpha_Q, X_k \rangle) \in \mathbb{R}^{Q \times Q}, \\
W(k+1) &= (\omega_1(k+1), \ldots, \omega_Q(k+1)) \in \mathbb{R}^Q, \\
\omega_i(k) &\sim N(0,1), i.i.d..
\end{aligned}$$

Ω is the variance-covariance matrix $(\sigma_{i,j}) \in \mathbb{R}^{Q \times Q}$, which we assume to be non-singular.

Write $\mathcal{F}_k^X = \sigma\{X_0, \ldots, X_k\}$ for the σ-field generated by $X_0, \ldots, X_k$. Similarly,

$$\mathcal{R}_k = \sigma\{R(1), R(2), \ldots, R(k)\},$$

and

$$\mathcal{G}_k = \sigma\{X_0, X_1, \ldots, X_k, R(1), R(2), \ldots, R(k)\}.$$

Then the filtrations $\{\mathcal{F}_k^X\}, \{\mathcal{R}_k\}$ and $\{\mathcal{G}_k\}$ model histories of the state process X, the observation return processes R and $\{X, R\}$, respectively.

4.1 Change of measure

A basic technique for filtering is the change of measure technique. We work under a reference measure for which the processes have nice properties, and change back to the real world space to obtain the estimates for the parameters.

The density function of a multivariate normal distribution $Y \sim N(0, I)$ is

$$\phi(y_1, y_2, \dots, y_Q) = \frac{1}{(2\pi)^{Q/2}} e^{-\frac{1}{2}Y'Y}.$$

Suppose under a reference probability measure $\bar{P}$, $R = \{R(k), k = 0, 1, 2, \dots\}$ is a sequence of independent normally distributed random variables each with $R(k) \sim N(0, I)$, and X is a Markov chain with transition matrix A.

Define

$$\lambda_0 = 1, \quad \lambda_k = |\Omega^{-1}| \frac{\phi(W(k))}{\phi(R(k))} \text{ For } k > 1,$$
$$\Lambda_t = \prod_{k=1}^{t} \lambda_k,$$

where

$$W(k) = \Omega^{-1}[R(k) - \Gamma(k) - \Upsilon(k)R(k-1)].$$

We then define the measure P by setting $dP/d\overline{P}|\mathcal{F}_k^X = \Lambda_k$. The following lemma tells us that P is the "real world" probability.

Lemma 5 *Under P, the $\{W(k)\}$ is a sequence of independent $N(0, I)$ random variables and X remains a Markov chain with transition matrix A.*

Proof. See Appendix B. □

In the "real world", we only observe the processes of spot prices R. We estimate X using the conditional expectation, $E[X_k|\mathcal{R}_k]$. By Bayes' theorem this can be expressed in terms of $\bar{P}$:

$$E[X_k|\mathcal{R}_k] = \frac{\overline{E}\left[\Lambda_k X_k|\mathcal{R}_k\right]}{\overline{E}\left[\Lambda_k|\mathcal{R}_k\right]}.$$

Write $\overline{E}[\Lambda_k X_k|\mathcal{R}_k] = q_k \in \mathbb{R}^n$. We shall derive a recursive estimate for $q_k(\cdot)$, and consequently obtain a estimator for X.

Lemma 6 *q satisfies the recurrence $q_k = D_\mu A q_{k-1}$,where*

$$D_\mu = diag(\mu^j), 1 \le j \le n,$$
$$\mu^j = \frac{|\Omega^{-1}|\phi(\Omega^{-1}(R(k) - \Gamma^j - \Upsilon^j R(k-1))}{\phi(R(k))}. \tag{14}$$

Here Γ^j and Υ^j are respectively the values of Γ and Υ under the state j, that is, when $X_k = e_j$.

Proof.

$$\begin{aligned} q_k &= \overline{E}\left[\Lambda_k X_k | \mathcal{R}_k\right] \\ &= \sum_{j=1}^{N} \overline{E}\left[\Lambda_{k-1}\mu^j \langle X_k, e_j\rangle e_j | \mathcal{R}_k\right] \\ &= \sum_{j=1}^{N} \overline{E}\left[\Lambda_{k-1}\mu^j \langle AX_{k-1} + M_k, e_j\rangle e_j | \mathcal{R}_k\right] \\ &= \sum_{j=1}^{N} \mu^j \langle Aq_{k-1}, e_j\rangle e_j = D_\mu A q_{k-1}, \end{aligned}$$

where $D_\mu = diag(\mu^j), 1 \le j \le n$. □

The following corollary gives the optimal estimate for the state variable.

Corollary 1 *The optimal estimate for the state variable is given by:*

$$E[X_k|\mathcal{R}_k] = \frac{q_k}{\langle q_k, \underline{1}\rangle}. \tag{15}$$

For estimating the parameters, we need the following quantities:

1. The number of jumps from e_r to e_s up to time k,

$$\mathcal{J}_k^{rs} = \sum_{t=1}^{k} \langle X_{t-1}, e_r\rangle \langle X_t, e_s\rangle. \tag{16}$$

2. The occupation time at the state e_r, up to time k,

$$\begin{aligned} \mathcal{O}_k^r &= \sum_{i=1}^{N}\sum_{t=1}^{k} \langle X_{t-1}, e_r\rangle \langle X_t, e_i\rangle \\ &= \sum_{t=1}^{k} \langle X_{t-1}, e_r\rangle. \end{aligned}$$

Before calculating recursive estimates of the these two dynamics, we consider the more general process:

$$\mathcal{Z}_k^{rs} = \sum_{t=1}^{k} \langle X_{t-1}, e_r \rangle \langle X_t, e_s \rangle F(R_j(t)) G(R_j(t-1)). \tag{17}$$

Here $F(R_j(t))$ and $G(R_j(t-1))$ are functions of $R_j(t)$ and $R_j(t-1)$ respectively, for $1 \le j \le Q$. The following Lemma gives a recursive estimate for this general dynamic.

Lemma 7 *Define* $\overline{E}[\Lambda(k)\mathcal{Z}_k^{rs} X_k | \mathcal{R}_k] = \Xi_k^{rs}$.*Then the following formula updates* Ξ_k^{rs} *:*

$$\Xi_k^{rs} = D_\mu A \Xi_{k-1}^{r,s} + F(R_j(k)) G(R_j(k-1)) \langle \lambda_k q_{k-1}, e_r \rangle a_{sr} e_r.$$

Proof.

$$\begin{aligned}
&\overline{E}\left[\Lambda_k \mathcal{Z}_k^{rs} X_k | \mathcal{R}_k\right] \\
&= \overline{E}\left[\lambda_k \Lambda_{k-1} \mathcal{Z}_k^{rs} X_k | \mathcal{R}_k\right] \\
&= \overline{E}\left[\lambda_k \Lambda_{k-1} [\mathcal{Z}_{k-1}^{rs} + F(R(k)) G(R(k-1)) \langle X_{k-1}, e_r \rangle \langle X_k, e_s \rangle X_k | \mathcal{R}_k\right] \\
&= \overline{E}\left[\lambda_k \Lambda_{k-1} \mathcal{Z}_{k-1}^{rs} (A X_{k-1} + M_k) | \mathcal{R}_k\right] \\
&\quad + \overline{E}\left[F(R_j(k)) G(R_j(k-1)) \lambda_k \Lambda_{k-1} \langle X_{k-1}, e_r \rangle \langle X_k, e_s \rangle X_k | \mathcal{R}_k\right] \\
&= \sum_{i=1}^{N} \overline{E}\left[\mu^i \Lambda_{k-1} \mathcal{Z}_{k-1}^{rs} \langle X_{k-1}, e_i \rangle | \mathcal{R}_k\right] A e_i \\
&\quad + \overline{E}\left[F(R_j(k)) G(R_j(k-1)) \langle \lambda_k \Lambda_{k-1} X_{k-1}, e_r \rangle a_{sr} e_r | \mathcal{R}_k\right] \\
&= \sum_{i=1}^{N} \langle \overline{E}\left[\mu^i \Lambda_{k-1} \mathcal{Z}_{k-1}^{rs} X_{k-1} | \mathcal{R}_k\right], e_i \rangle A e_i \\
&\quad + F(R_j(k)) G(R_j(k-1)) \langle \overline{E}\left[\lambda_k \Lambda_{k-1} X_{k-1} | \mathcal{R}_k\right], e_r \rangle a_{sr} e_r.
\end{aligned}$$

Here μ^i is as defined as equation (14) in *Lemma 5.* Thus

$$\begin{aligned}
\Xi_k^{r,s} &= \sum_{i=1}^{N} \langle \mu^i \Xi_{k-1}^{r,s}, e_i \rangle A e_i + \langle F(R_j(k)) G(R_j(k-1)) \lambda_k q_{k-1}, e_r \rangle a_{sr} e_r \\
&= D_\mu A \Xi_{k-1}^{r,s} + F(R_j(k)) G(R_j(k-1)) \langle \lambda_k q_{k-1}, e_r \rangle a_{sr} e_r.
\end{aligned}$$

This is a recurrence for $\Xi_k^{r,s}$. $\square$

Write $F(R_j(k)) = G(R_j(k-1)) = 1$ in $\mathcal{Z}_k^{rs}$. We then obtain the following Corollaries.

Corollary 2 *Write* $\overline{E}[\Lambda(k)\mathcal{J}_k^{rs}X_k|\mathcal{R}_k] = \Phi_k^{r,s}$. *Then the following formula updates* $\Phi_k^{r,s}$:

$$\Phi_k^{r,s} = D_\mu A \Phi_{k-1}^{rs} + \langle \lambda_k q_{k-1}, e_r \rangle a_{sr} e_r.$$

Corollary 3 *The optimal estimate of* $\mathcal{J}_k^{rs}$ *is:* $\widehat{\mathcal{J}}_k^{rs} = \dfrac{\langle \Phi_k^{r,s}, \underline{1} \rangle}{\langle q_k, \underline{1} \rangle}$.

Similarly, we define

$$\mathcal{N}_k^r = \sum_{t=1}^{k} \langle X_{t-1}, e_r \rangle F(R_j(t)) G(R_j(t-1)), \tag{18}$$

where $F(R_j(t))$ and $G(R_j(t-1))$ are two functions of $R_j(t)$ and $R_j(t-1)$ respectively. We obtain the following result.

Lemma 8 *Define* $\overline{E}[\Lambda(k)\mathcal{N}_k^r X_k|\mathcal{R}_k] = \Delta_k^r$. *Then the following formula updates* Φ_k :

$$\Delta_k^r = D_\mu A \Delta_{k-1}^r + F(R_j(k)) G(R_j(k-1)) \langle A \lambda_k q_{k-1}, e_r \rangle e_r.$$

Corollary 4 *Define* $\overline{E}[\Lambda(k)\mathcal{O}_k^r X_k|\mathcal{R}_k] = \Psi_k^r$. *Then the following formula updates* Ψ_k :

$$\Psi_k^r = D_\mu A \Psi_{k-1}^r + \langle \lambda_k q_k, e_r \rangle e_r.$$

Corollary 5 *The optimal estimate of* $\mathcal{O}_k^r$ *is:* $\widehat{\mathcal{O}}_k^r = \dfrac{\langle \Psi_k^r, e_r \rangle}{\langle q_k, \underline{1} \rangle}$.

4.2 Parameter Estimates

We shall use the EM algorithm to estimate the parameters. Write the parameters of our model as

$$\theta := \{a_{ji}, \Gamma, \Upsilon, \Omega,\ 1 \le i, j \le N\},$$

where $a_{ji} \ge 0$, $\sum_{j=1}^{N} a_{ji} = 1$. The basic idea of the EM algorithm is:

- We start with appropriate initial values $\hat{\theta}_0$ for

$$\theta := \{a_{ji}, \Gamma, \Upsilon, \Omega,\ 1 \le i, j \le N\},$$

  which satisfy constraints for the parameters.

- After some observations of R, we compute new estimates.

- Using these values, we re-estimate the parameters iteratively until some stopping criterion is satisfied.

- After more observations we repeat the process again.

Since the EM algorithm improves the estimates monotonically, the expected log-likelihood increases with each re-estimation. In this section, we apply this algorithm to obtain the estimates for all the parameters recursively.

4.2.1 Estimate of a_{ji}

We first estimate the entries of the transition matrix for the Markov Chain by applying the change of measure technique.

Lemma 9 *Define* $\Lambda_0 = 1$ *and* $\Lambda_k = \prod_{l=1}^{k} \left(\sum_{j,i=1}^{N} (\frac{\hat{a}_{ji}}{a_{ji}}) \langle X_l, e_j \rangle \langle X_{l-1}, e_i \rangle \right)$. *(If* $a_{ji} = 0$, *take* $\hat{a}_{ji} = 0$ *and* $\frac{\hat{a}_{ji}}{a_{ji}} = 1$. *) Define* $P_{\hat{\theta}}$ *by setting* $\frac{dP_{\hat{\theta}}}{dP_{\theta}} | \mathcal{R}_k = \Lambda_k.$ *Then under the new measure* $P_{\hat{\theta}}$, X *is a Markov Chain with transition* $A = (\hat{a}_{ji})$.

Lemma 10 *Given* $\mathcal{R}_k$ *and parameter set* $\theta := \{a_{ji}, \Gamma, \Upsilon, \Omega,\ 1 \le i, j \le N\}$, *the EM estimates of* a_{ji} *are given by*

$$\hat{a}_{ji} = \frac{\widehat{\mathcal{J}}_k^{ji}}{\widehat{\mathcal{O}}_k^{i}} = \frac{\langle \Phi_k^{ji}, \underline{1} \rangle}{\langle \Psi_k^{i}, \underline{1} \rangle}.$$

Proof. As above we define $P_{\hat{\theta}}$ by:

$$\frac{dP_{\hat{\theta}}}{dP_{\theta}} | \mathcal{R}_k = \Lambda_k = \prod_{l=1}^{k} \left(\sum_{i,j=1}^{N} (\frac{\hat{a}_{ji}}{a_{ji}}) \langle X_l, e_j \rangle \langle X_{l-1}, e_i \rangle \right).$$

Then

$$\log \frac{dP_{\hat{\theta}}}{dP_\theta}|\mathcal{R}_k = \sum_{l=1}^{k}\sum_{i,j=1}^{N} \langle X_l, e_j\rangle \langle X_{l-1}, e_i\rangle (\log \hat{a}_{ji} - \log a_{ji})$$

$$= \sum_{i,j=1}^{N} \mathcal{J}_k^{ij} \log \hat{a}_{ji} + i(a),$$

where $i(a)$ does not depend on $\hat{a}_{ji}$. Then

$$L(\hat{\theta}) = E\left[\log \frac{dP_{\hat{\theta}}}{dP_\theta}|\mathcal{R}_k\right] = \sum_{i,j=1}^{N} \hat{\mathcal{J}}_k^{ij} \log \hat{a}_{ji} + i(a). \tag{19}$$

Recall that $\sum_{j=1}^{N} \hat{a}_{ji} = 1$, and $\sum_{j=1}^{N} \mathcal{J}_k^{ij} = \mathcal{O}_k^i$. Then, the optimal estimate of $\hat{a}_{ji}$ is the value that maximizes the right side of (19), and subject to $\sum_{j=1}^{N} \hat{a}_{ji} = 1$. Let c be the Lagrange multiplier and put

$$L(\hat{P}, \lambda) = \sum_{i,j=1}^{N} \hat{\mathcal{J}}_k^{ij} \log \hat{a}_{ji} + i(a) + c(\sum_{j=1}^{N} \hat{a}_{ji} - 1).$$

Differentiating in $\hat{a}_{ji}$ and c and equating the derivatives to 0, we obtain two equations. By solving the equations, we have:

$$c = -\hat{\mathcal{O}}_k^i,$$

and

$$\hat{a}_{ji} = \frac{\hat{\mathcal{J}}_k^{ij}}{\hat{\mathcal{O}}_k^i}.$$

This provides estimates for the elements of the transition matrix. □

4.2.2 Estimate of Γ, Υ and Ω

Lemma 11 *The estimate for the parameter Γ, at time k, $\hat{\Gamma}$ is given by*

$$\hat{\Gamma}^r = \frac{\Delta_k^{r,1} - \Upsilon^r \Delta_k^{r,2}}{\hat{\mathcal{O}}_k^r},$$

where

$$
\begin{aligned}
\Gamma^r &= \gamma_r, \Upsilon^r = \upsilon_r, \\
\Delta_k^{r,1} &= (\Delta_{k,1}^{r,1}, \ldots, \Delta_{k,Q}^{r,1}),\ \ \Delta_k^{r,2} = (\Delta_{k,1}^{r,2}, \ldots, \Delta_{k,Q}^{r,2}), \\
\Delta_{k,j}^{r,1} &= D_\mu A \lambda_k \Delta_{k-1,j}^{r,1} + R_j(k) \langle A\lambda_k q_k, e_r \rangle e_r, \\
\Delta_{k,j}^{r,2} &= D_\mu A \lambda_k \Delta_{k-1,j}^{r,2} + R_j(k-1) \langle A\lambda_k q_k, e_r \rangle e_r.
\end{aligned}
$$

Proof. The density which changes Γ to $\hat{\Gamma}$ is given by

$$
\Lambda_k = \frac{dP_{\hat{\Gamma}}}{dP_\Gamma} = \prod_{t=1}^{k} \frac{\phi(\hat{W}(t))}{\phi(W(t))},
$$

where

$$
\begin{aligned}
W(t) &= \Omega^{-1}[R(t) - \Gamma(t) - \Upsilon(t) R(t-1)], \\
\hat{W}(t) &= \Omega^{-1}[R(t) - \hat{\Gamma}(t) - \Upsilon(t) R(t-1)], \\
\hat{\Gamma}(t) &:= (\hat{\gamma}_1, \hat{\gamma}_2, \ldots, \hat{\gamma}_Q)' \\
&\quad = (\langle(\hat{\alpha}\hat{\beta})_1, X_{t-1}\rangle, \ldots, \langle(\hat{\alpha}\hat{\beta})_Q, X_{t-1})' \in \mathbb{R}^Q, \\
(\hat{\alpha}\hat{\beta})_i &= (\hat{\alpha}_i^1 \hat{\beta}_i^1, \hat{\alpha}_i^2 \hat{\beta}_i^2, \ldots, \hat{\alpha}_i^N \hat{\beta}_i^N)' \in \mathbb{R}^N \\
&\quad 1 \leq i \leq Q, \\
\phi(\hat{W}(t)) &= \frac{1}{2\pi^{Q/2}} e^{-\frac{1}{2}\hat{W}(t)'\hat{W}(t)}.
\end{aligned}
$$

Then

$$
\begin{aligned}
E\left[\log \Lambda_k | \mathcal{F}_k^X\right] &= E\left[\sum_{t=1}^{k} -\frac{1}{2}\hat{W}(t)'\hat{W}(t) | \mathcal{F}_k^X\right] + L(\Gamma, \Upsilon, \Omega) \\
&= \sum_{r=1}^{N} E\left[\sum_{t=1}^{k} -\frac{1}{2}\langle X_{t-1}, e_r\rangle \hat{W}^r(t)'(\hat{W}^r(t)) | \mathcal{F}_k^X\right] + L(\Gamma, \Upsilon, \Omega),
\end{aligned}
$$

where $\hat{W}^r(t)$ denotes the value of $\hat{W}(t)$ under state r and

$$
\hat{W}^r(t) = \Omega^{-1}(R(t) - \hat{\Gamma}^r - \Upsilon^r R(t-1)).
$$

To maximize $E\left[\log \Lambda_k | \mathcal{F}_k^X\right]$, we differentiate it against $\hat{\Gamma}^r, 1 \le r \le N$, and let the first derivative equal zero. That is,

$$E\left[\sum_{t=1}^{k} \langle X_{t-1}, e_r\rangle (R(t) - \hat{\Gamma}^r - \Upsilon^r R(t-1)) | \mathcal{F}_k^X\right] = 0.$$

Thus,

$$\hat{\Gamma}^r = \frac{E\left[\sum_{t=1}^{k} \langle X_{t-1}, e_r\rangle (R(t) - \Upsilon^r R(t-1)) | \mathcal{F}_k^X\right]}{E\left[\sum_{t=1}^{k} \langle X_{t-1}, e_r\rangle | \mathcal{F}_k^X\right]} = \frac{\Delta_k^{r,1} - \Upsilon^r \Delta_k^{r,2}}{\hat{\mathcal{O}}_k^r},$$

Here

$$\Delta_k^{r,1} = (\Delta_{k,1}^{r,1}, \ldots, \Delta_{k,Q}^{r,1}) = E\left[\sum_{t=1}^{k} \langle X_{t-1}, e_r\rangle R_j(t) | \mathcal{F}_k^X\right],$$

$$\Delta_k^{r,2} = (\Delta_{k,1}^{r,2}, \ldots, \Delta_{k,Q}^{r,2}) = E\left[\sum_{t=1}^{k} \langle X_{t-1}, e_r\rangle R_j(t-1) | \mathcal{F}_k^X\right].$$

By *Lemma 8*, we have

$$\begin{aligned}
\Delta_{k,j}^{r,1} &= D_\mu A \lambda_k \Delta_{k-1,j}^{r,1} + R_j(k) \langle A \lambda_k q_k, e_r\rangle e_r, \\
\Delta_{k,j}^{r,2} &= D_\mu A \lambda_k \Delta_{k-1,j}^{r,2} + R_j(k-1) \langle A \lambda_k q_k, e_r\rangle e_r.
\end{aligned}$$

Therefore,

$$\hat{\Gamma}^r = \frac{\Delta_k^{r,1} - \Upsilon^r \Delta_k^{r,2}}{\hat{\mathcal{O}}_k^r}.$$

$\square$

Similarly we have the following two *Lemmas*.

Lemma 12 *The estimate for the parameter Υ, at time k, $\hat{\Upsilon}$ is given by*

$$\hat{\Upsilon}^r = \left(\Delta_k^{r,1} - \hat{\mathcal{O}}_k^r \hat{\Gamma}^r\right) \left(\Delta_k^{r,2}\right)^{-1}.$$

Lemma 13 *The estimate for the parameter* Ω *at time* k, $\hat{\Omega}$ *is given by*

$$\hat{\Omega}^2 = \frac{1}{k} \sum_{r=1}^{N} E\left[\sum_{t=1}^{k} \langle X_{t-1}, e_r \rangle \hat{Z}^r(t) (\hat{Z}^r(t))' | \mathcal{F}_k^X \right].$$

Proof. The density which changes Ω to $\hat{\Omega}$ is given by

$$\Lambda_k = \frac{dP_{\hat{\Omega}}}{dP_{\Omega}} = \prod_{t=1}^{k} \left(\frac{|\Omega|}{\left|\hat{\Omega}\right|} \frac{\phi(\hat{W}(t))}{\phi(W(t))} \right),$$

where

$$\begin{aligned} W(t) &= \Omega^{-1}[R(t) - \Gamma(t) - \Upsilon(t) R(t-1)], \\ \hat{W}(t) &= \hat{\Omega}^{-1}[R(t) - \Gamma(t) - \Upsilon(t) R(t-1)], \\ \phi(\hat{W}(t)) &= \frac{1}{2\pi^{Q/2}} e^{-\frac{1}{2}\hat{W}(t)\hat{W}(t)'}. \end{aligned}$$

Then

$$\log \Lambda_k = -k \log \left|\hat{\Omega}\right| - \frac{1}{2} \sum_{t=1}^{k} \hat{W}(t) \hat{W}(t)' + L(\Gamma, \Upsilon, \Omega)$$

$$\begin{aligned} E\left[\log \Lambda_k | \mathcal{F}_k^X\right] &= E\left[-k \log \left|\hat{\Omega}\right| - \frac{1}{2} \sum_{t=1}^{k} \hat{W}(t)' \hat{W}(t) | \mathcal{F}_k^X \right] + L(\Gamma, \Upsilon, \Omega) \\ &= -k \log \left|\hat{\Omega}\right| - \sum_{r=1}^{N} E\left[\sum_{t=1}^{k} \frac{1}{2} \langle X_{t-1}, e_r \rangle \hat{W}^r(t)' (\hat{W}^r(t)) | \mathcal{F}_k^X \right] + L(\Gamma, \Upsilon, \Omega) \\ &= -k \log \left|\hat{\Omega}\right| - \frac{1}{2} \sum_{r=1}^{N} E\left[\sum_{t=1}^{k} \langle X_{t-1}, e_r \rangle \hat{Z}^r(t)' \hat{\Omega}^{-2} (\hat{Z}^r(t)) | \mathcal{F}_k^X \right] + L(\Gamma, \Upsilon, \Omega), \end{aligned}$$

where we write

$$\hat{Z}(t) = [R(t) - \Gamma(t) - \Upsilon(t) R(t-1)].$$

To maximize $E\left[\log \Lambda_k | \mathcal{F}_k^X\right]$, we differentiate it against $\hat{\Omega}, 1 \leq r \leq N$, and let the first derivative equal zero. That is,

$$-k\left(\hat{\Omega}^{-1}\right) + \hat{\Omega}^{-3} \sum_{r=1}^{N} E\left[\sum_{t=1}^{k} \langle X_{t-1}, e_r \rangle \hat{Z}^r(t)(\hat{Z}^r(t))' | \mathcal{F}_k^X\right] = 0.$$

Therefore

$$\hat{\Omega}^2 = \frac{1}{k} \sum_{r=1}^{N} E\left[\sum_{t=1}^{k} \langle X_{t-1}, e_r \rangle \hat{Z}^r(t)(\hat{Z}^r(t))' | \mathcal{F}_k^X\right].$$

□

Lemmas 10 to 13 give recursively updated estimates for all the parameters. We can apply these *Lemmas* to real world data to estimate the parameters. For large Q and N, the computation might be expensive, but for small numbers, it is not too hard or time-consuming.

5 Conclusions

In this paper, we have presented a mean reverting model for several assets by including correlations between the assets and allowing the parameters of the spot price dynamics to switch between finite regimes. We have derived estimates for forwards and European Options on an individual asset whose spot prices follow the specified dynamics. The parameters in the model can be estimated recursively by applying the *Lemmas* of section 4. A possible extension of the model is to add jumps into the model. A convenience yield could also be included in the mean reverting terms. To our knowledge, this is the first model which considers both the correlations between the prices of several assets, and the regime switching features of real world dynamics.

References

[1] Black, F., and M., Scholes, The pricing and corporate liabilities, *Journal of Political Economy* **81**: 637-659 (1973).

[2] Clewlow, L., and C. Strickland, Valuing energy options in a one factor model fitted to forward prices, *Working paper.*

[3] R. J. Elliott, New finite-dimensional filters and smoothers for noisily observed Markov chains, *IEEE Transactions on Information Theory* **39:** no. 1, 265–271(1993).

[4] Elliott, R.J., P. Fischer and E. Platen, Filtering and parameter estimation for a mean reverting interest rate model, *Canadian Applied Math. Quarterly* **7**: 381-400 (1999).

[5] Elliott, R.J., W.P. Malcolm, and A.H. Tsoi, Robust parameter estimation for asset price models with Markov modulated volatilities, *Journal of Economic, Dynamics and Control* **27**: 1391-1409 (2003).

[6] Wu, P., and R. J. Elliott, Parameter Estimates for a Regime Switching Mean-Reverting Model with Jumps, *International Journal of Theoretical and Applied Finance* **8(6)**: 791-806 (2005).

A Calculation of $E[M_2]$

$$
\begin{aligned}
& E[M_2|R_i\,(0)\vee X_0] \\
=& E\left[\left(\int_0^T e^{-\langle\alpha_i,\mathcal{O}_T-\mathcal{O}_s\rangle}\langle\alpha_i\odot\beta_i,X_s\rangle ds\right)^2\right] \\
=& 2E\left[\int_0^T\int_0^s\left(e^{-\langle\alpha_i,\mathcal{O}_T-\mathcal{O}_r\rangle}\langle\alpha_i\odot\beta_i,X_r\rangle dr\right)e^{-\langle\alpha_i,\mathcal{O}_T-\mathcal{O}_s\rangle}\langle\alpha_i\odot\beta_i,X_s\rangle ds|R_i\,(0)\vee X_0\right] \\
=& 2\int_0^T\int_0^s E\left[\left(e^{-\langle\alpha_i,\mathcal{O}_T-\mathcal{O}_r\rangle}\langle\alpha_i\odot\beta_i,X_r\rangle\right)e^{-\langle\alpha_i,\mathcal{O}_T-\mathcal{O}_s\rangle}\langle\alpha_i\odot\beta_i,X_s\rangle|R_i\,(0)\vee X_0\right]drds \\
& (0\le r\le s\le T) \\
=& 2\int_0^T\int_0^s E[E[E\left[e^{-2\langle\alpha_i,\mathcal{O}_T\rangle}e^{\langle\alpha_i,\mathcal{O}_s\rangle}\langle\alpha_i\odot\beta_i,X_s\rangle|\mathcal{F}_s^X\right] \\
& e^{\langle\alpha_i,\mathcal{O}_r\rangle}\langle\alpha_i\odot\beta_i,X_r\rangle|\mathcal{F}_r^x]|X_0]drds \\
=& 2\int_0^T\int_0^s E[E[e^{-2\langle\alpha_i,\mathcal{O}_s\rangle}\langle e^{(\tilde{A}-2D_{\alpha_i})(T-s)}X_s,\underline{1}\rangle e^{\langle\alpha_i,\mathcal{O}_s\rangle}\langle\alpha_i\odot\beta_i,X_s\rangle \\
& e^{\langle\alpha_i,\mathcal{O}_r\rangle}\langle\alpha_i\odot\beta_i,X_r\rangle|\mathcal{F}_r^x]|R_i\,(0)\vee X_0]drds \\
=& 2\int_0^T\int_0^s E[\langle(e^{(\tilde{A}'-2D_{\alpha_i})(T-s)}\underline{1})\odot\alpha_i\odot\beta_i,E[e^{2\langle\alpha_i,\mathcal{O}_s\rangle}X_s|\mathcal{F}_r^X]\rangle \\
& e^{\langle\alpha_i,\mathcal{O}_r\rangle}\langle\alpha_i\odot\beta_i,X_r\rangle|R_i\,(0)\vee X_0]drds \\
=& 2\int_0^T\int_0^s E[\langle(e^{(\tilde{A}'-2D_{\alpha_i})(T-s)}\underline{1})\odot\alpha_i\odot\beta_i,e^{(\tilde{A}-2D_{\alpha_1})(s-r)}e^{-\langle\alpha_i,\mathcal{O}_r\rangle}X_r\rangle \\
& e^{\langle\alpha_i,\mathcal{O}_r\rangle}\langle\alpha_i\odot\beta_i,X_r\rangle|R_i\,(0)\vee X_0]drds \\
=& 2\int_0^T\int_0^s E[\langle(e^{(\tilde{A}'-2D_{\alpha_i})(T-s)}\underline{1})\odot\alpha_i\odot\beta_i,e^{(\tilde{A}-D_{\alpha_i})(s-r)}X_r\rangle\langle\alpha_i\odot\beta_i,X_r\rangle|X_0]drds \\
=& 2\int_0^T\int_0^s\langle e^{(\tilde{A}'-D_{\alpha_i})(s-r)}[(e^{(\tilde{A}'-2D_{\alpha_i})(T-s)}\underline{1})\odot\alpha_i\odot\beta_i]\alpha_i\odot\beta_i,e^{\tilde{A}r}X_0\rangle drds.
\end{aligned}
$$

B Proof of Lemma 5

Proof. To Under P, the $\{W(k)\}$ is a sequence of independent $N(0, I)$ random variables, we must show that for any measurable function $f(Y_k)$:

$$E\left[f(W(k))|\mathcal{G}_{k-1}\right] = \int_{-\infty}^{\infty} \phi(W(k))f(W(k))dW(k).$$

Using Bayes' theorem, for a measurable function $f(W(k))$, we have

$$\begin{aligned} E[f(W(k))|\mathcal{F}_{k-1}^X] &= \frac{\overline{E}[\Lambda_k f(W(k))|\mathcal{F}_{k-1}^X]}{\overline{E}[\Lambda_k|\mathcal{F}_{k-1}]} \\ &= \frac{\overline{E}[\lambda_k \Lambda_{k-1} f(W(k))|\mathcal{F}_{k-1}^X]}{\overline{E}[\lambda_k \Lambda_{k-1}|\mathcal{F}_{k-1}^X]} \\ &= \frac{\overline{E}[\lambda_k f(W(k))|\mathcal{F}_{k-1}^X]}{\overline{E}[\lambda_k|\mathcal{F}_{k-1}^X]}. \end{aligned} \tag{20}$$

Notice that

$$\begin{aligned} &\overline{E}[\lambda_k|\mathcal{F}_{k-1}^X] \\ =&\overline{E}\left[|\Omega^{-1}|\frac{\phi(W(k))}{\phi(R(k))}|\mathcal{F}_{k-1}^X\right] \\ =&\overline{E}\left[|\Omega^{-1}|\frac{\phi(\Omega^{-1}(k)[R(k)-\Gamma(k)-\Upsilon(k)R(k-1)])}{\phi(R(k))}|\mathcal{F}_{k-1}^X\right] \\ =&\int_{-\infty}^{+\infty} |\Omega^{-1}|\frac{\phi(W(k))}{\phi(R(k))}\phi(R(k))dR(k) \\ =&\int_{-\infty}^{+\infty} |\Omega^{-1}|\phi(W(k))dR(k). \end{aligned}$$

Recall that $W(k) = \Omega^{-1}[R(k) - \Gamma(k) - \Upsilon(k)R(k-1)]$.

Using the Jacobian to change the vector variables of integration, we have $dR(k) = |\Omega| \prod_{i=1}^{Q} dW_i(k) = |\Omega| dW(k)$, and the equation becomes:

$$\int_{-\infty}^{+\infty} |\Omega^{-1}|\phi(W(k))|\Omega|dW(k) = \int_{-\infty}^{+\infty} \phi(W(k))dW(k) = 1.$$

Similarly, we consider the numerator of equation 20:

$$\begin{aligned}
&\overline{E}[\Lambda_k f(W(k))|\mathcal{F}^X_{k-1}] \\
=&\overline{E}\left[|\Omega^{-1}|\frac{\phi(W(k))}{\phi(R(k))}f(W(k))|\mathcal{F}^X_{k-1}\right] \\
=&\int_{-\infty}^{+\infty}|\Omega^{-1}|\frac{\phi(W(k))}{\phi(R(k))}\phi(R(k))f(W(k))dR(k) \\
=&\int_{-\infty}^{+\infty}\phi(W(k))f(W(k))dW(k).
\end{aligned}$$

That is, under P, the $W(k)$ is a sequence of independent $N(0,I)$ random variables, and P is the real world probability measure.

Now we show that under P, X remains a Markov chain with transition matrix A :

Using Bayes' theorem, we have

$$\begin{aligned}
E[X_k|\mathcal{R}_{k-1}] &= \frac{\overline{E}[\Lambda_k X_k|\mathcal{R}_{k-1}]}{\overline{E}[\Lambda_k|\mathcal{R}_{k-1}]} \\
&= \frac{\overline{E}[\lambda_k X_k|\mathcal{R}_{k-1}]}{\overline{E}[\lambda_k|\mathcal{R}_{k-1}]}.
\end{aligned} \tag{21}$$

We already know $\overline{E}[\lambda_k|\mathcal{R}_{k-1}]=1$. Now,

$$\begin{aligned}
&\overline{E}[\lambda_k X_k|\mathcal{R}_{k-1}] \\
&= \overline{E}\left[|\Omega^{-1}|\frac{\phi(\Omega^{-1}(k)[R(k)-\Gamma(k)-\Upsilon(k)R(k-1)])}{\phi(R(k))}X_k|\mathcal{R}_{k-1}\right] \\
&= \overline{E}\left[|\Omega^{-1}|\frac{\phi(\Omega^{-1}(k)[R(k)-\Gamma(k)-\Upsilon(k)R(k-1)])}{\phi(R(k))}AX_{k-1}+M_{k-1}|\mathcal{R}_{k-1}\right] \\
&= \overline{E}\left[|\Omega^{-1}|\frac{\phi(\Omega^{-1}(k)[R(k)-\Gamma(k)-\Upsilon(k)R(k-1)])}{\phi(R(k))}AX_{k-1}|\mathcal{R}_{k-1}\right] \\
&= \overline{E}\left[|\Omega^{-1}|\frac{\phi(\Omega^{-1}(k)[R(k)-\Gamma(k)-\Upsilon(k)R(k-1)])}{\phi(R(k))}|\mathcal{G}_{k-1}\right]AX_{k-1} \\
&= \overline{E}[\lambda_k|\mathcal{G}_{k-1}]AX_{k-1} = AX_{k-1}.
\end{aligned}$$

Thus, from equation (21) , we have $E[X_k|\mathcal{G}_{k-1}]=AX_{k-1}$. That is under P, X remains a Markov chain with transition matrix A. $\square$