

On Markov Modulated Mean-Reverting Price-Difference Models

W. P. Malcom*, Lakhdar Aggoun and Mohamed Al-Lawati **

**National ICT Australia, Canberra, Australia; **Department of Mathematics and Statistics, College of Science, Sultan Qaboos University, P.O.Box 36, PC 123 Al-Khoud, Muscat, Sultanate of Oman, Email: laggoun@squ.edu.om.*

ABSTRACT: In this paper we develop a stochastic model incorporating a double-Markov modulated mean-reversion model. Unlike a price process the basis process X can take positive or negative values. This model is based on an explicit discretisation of the corresponding continuous time dynamics. The new feature in our model is that we suppose the mean reverting level in our dynamics as well as the noise coefficient can change according to the states of some finite-state Markov processes which could be the economy and some other unseen random phenomenon.

KEYWORDS: Double-Markov Modulated Mean-Reversion Model; Filtering; Smoothing.

1. Introduction

The main contribution of this article is to further extend the primary ideas presented in Elliott C. *et al.* (2005). The subject matter of our work is the dynamics that describe the difference between two prices, for example the prices of two different stocks. What we would like to do, is construct the dynamics which model price differences, and in addition to this, capture important but unseen random phenomena. To do this, we consider regime switching mean reversion. Mean-reverting models are well known in quantitative finance and were introduced by Vasicek. The extension of Vasicek's ideas to Markov-modulated mean reversion has been investigated for interest rate models (see Elliott R.J. *et al.*, 1999).

One common domain of application for price difference models is in the natural gas market. In the natural gas market the basis is the difference in the price of gas at two delivery points. The usual reference in the U.S.A. for a basis differential is NYMEX. For example, if the May Henry Hub price is \$5.25 and the May NYMEX

price is \$5.45 then the basis differential for May NYMEX is \$0.20 to Henry. The usual reference for Canada is the price at the AECO facility.

In this article we propose to model the basis as a mean reverting diffusion, $X = \{X_t, t \geq 0\}$. Unlike a price process the basis process X can take positive or negative values. The new feature in our model is that we suppose the mean reverting level in our dynamics can change according to the state of the economy. The economy is modeled as a finite state Markov chain $Z = \{Z_t, t \geq 0\}$ and the economy can perhaps be in two states ('good' and 'bad'), or possibly three states. Our continuous time model is discretized and the results of Elliott R.J. *et al* (2005), are adapted to obtain a recursive filter for the state of the economy given observations of X . In turn, this allows predictions to be made of the basis at the next time. If the observed basis is then higher or lower than the predicted value, it suggests one price is possibly higher than it should be and the other lower. Consequently, a trading strategy can be implemented based on these predictions.

2. Stochastic dynamics

All models are, initially, on the probability space (Ω, F, P) . Write $X = \{X_u, 0 \leq u \leq t\}$, for the basis (price difference) process. $X_t \in R$. Suppose L is a mean reversion level and $\alpha \in R_+$ is the rate-parameter, that is, a parameter determining how fast the level L is attained by the process X . X has dynamics:

$$X_t = X_0 + \alpha \int_0^t (L - X_u) du + \sigma W_t. \quad (1)$$

where W is a standard Wiener process, and $\sigma \in R$.

Remark 1. *The dynamics at (1) exhibit a mean reversion¹ character of the model when written in stochastic differential equation form:*

$$dX_t = \alpha(L - X_t) dt + \sigma dW_t. \quad (2)$$

Ignoring the noise σdW_t , if $X_t > L$ then $\alpha(L - X_t) < 0$, while if $X_t < L$ then $\alpha(L - X_t) > 0$, and so the right-hand side of (2) is continually trying to reach the level L .

Now suppose that parameters L and σ are stochastic and can switch between different levels L_1, L_2, \dots, L_m and $\sigma_1, \dots, \sigma_n$, respectively. We assume here that these levels are determined by the states of two Markov chains Z and A , respectively. Without loss of generality, we take the state spaces of our Markov chains to be the canonical basis $L = \{e_1, e_2, \dots, e_m\}$ of R^m and the canonical basis $S = \{f_1, f_2, \dots, f_n\}$ of R^n , respectively. Write

$$\begin{aligned} \pi_{(j,i)} &= P(Z_{k+1} = e_j / Z_k = e_i), \\ p_{(s,r)} &= P(A_{k+1} = f_s / A_k = f_r), \end{aligned} \quad (3)$$

$$\Pi = [\pi_{(j,i)}]_{\substack{1 \leq j \leq m \\ 1 \leq i \leq m}}, P = [p_{(s,r)}]_{\substack{1 \leq s \leq n \\ 1 \leq r \leq n}}. \quad (4)$$

¹ Modelling a mean reversion process is widely used in finance, for example in interest rates models such as the Vasicek Model. This class of models assumes an (static) average value will be attained, not unlike the notion of an equilibrium state, or steady state of a dynamical system in the physical sciences

ON MARKOV MODULATED MEAN-REVERTING PRICE-DIFFERENCE MODELS

Write $\mathbb{Z}_t = \sigma\{Z_u, A_u, 0 \leq u \leq t\}$. Then

$$\begin{aligned} Z_{k+1} &= \Pi Z_k + M_{k+1}, \\ A_{k+1} &= P A_k + m_{k+1}. \end{aligned} \tag{5}$$

Here, M and m are martingale increments. The scalar-valued Markov processes taking values L_1, \dots, L_m and $\sigma_1, \dots, \sigma_n$, are obtained by

$$\begin{aligned} \langle Z_t, \mathbf{L} \rangle &= \sum_{\ell=1}^m 1_{\{\omega: Z_t(\omega) = e_\ell\}} L_\ell, \\ \langle \mathbf{S}, A_t \rangle &= \sum_{i=1}^n 1_{\{\omega: A_t(\omega) = f_i\}} \sigma_i. \end{aligned} \tag{6}$$

Here $\mathbf{L} = (L_1, L_2, \dots, L_m)'$, $\mathbf{S} = (\sigma_1, \sigma_2, \dots, \sigma_n)'$; $\langle \cdot, \cdot \rangle$ denotes an inner product and $1_{\{A\}}$ denotes an indicator function for the event A .

What also we wish to impose is that the two Markov chains Z and A are not independent, that is, information on the behavior of one conveys some knowledge on the behavior of the other. More precisely, we assume the dynamics:

$$Z_{k+1} \otimes A_{k+1} = \mathbf{P} Z_k \otimes A_k + \mathbf{M}_{k+1}. \tag{7}$$

where $\mathbf{P} = (\mathbf{p}_{js,ir})$ denotes a $mn \times mn$ matrix, or tensor, mapping $\mathbb{R}^m \times \mathbb{R}^n$ into $\mathbb{R}^m \times \mathbb{R}^n$ and

$\mathbf{p}_{js,ir} = P(Z_{k+1} = e_j, A_{k+1} = f_s / Z_k = e_i, A_k = f_r)$, $1 \leq r, s \leq n, 1 \leq i, j \leq m$. Again \mathbf{M}_{k+1} is a martingale increment. The dynamics at (1) take the form

$$X_t = X_0 + \alpha \int_0^t (\langle Z_u, \mathbf{L} \rangle - X_u) du + \langle \mathbf{S}, A_t \rangle W_t. \tag{8}$$

Remark 2. We defined Z and A as inherently discrete-time. Here, we "read" Z and A as the output of a sample and hold circuit, or CADLAG processes.

What we wish to do now, is discretise the dynamics at (8) and then compute a corresponding filter and detector. We will use an Euler-Maruyama discretisation scheme to obtain discrete-time dynamics, although many other schemes can be used.

For all time discretisations we will consider a partition, on the interval $[0, T]$ and write

$$M^{(K)} \triangleq \{0 = t_0, t_1, \dots, t_K = T\}. \tag{9}$$

This partition is strict, $t_0 < t_1 < \dots$, regular and the $\Delta_t = t_k - t_{k-1}$ are identical for indices k . Applying the Euler-Maruyama scheme to (8), we get,

$$\begin{aligned} X_{k+1} &= X_k + \alpha \langle Z_k, \mathbf{L} \rangle \Delta_t - \alpha X_k \Delta_t + \langle A_k, \mathbf{S} \rangle (W_{k+1} - W_k) \\ &= a X_k + b \langle Z_k, \mathbf{L} \rangle + c \langle A_k, \mathbf{S} \rangle v_k. \end{aligned} \tag{10}$$

Here

$$a \triangleq (1 - \alpha \Delta_t), \tag{11}$$

$$b \triangleq \alpha \Delta_t, \tag{12}$$

$$c \triangleq \sqrt{\Delta_t}. \quad (13)$$

The Gaussian process v is an independently and identically distributed $N(0,1)$. Our stochastic system now, under the measure P , has the form:

$$P \begin{cases} Z_{k+1} = \Pi Z_k + M_k \\ A_{k+1} = PA_k + M_{k+1} \\ Z_{k+1} \otimes A_{k+1} = \mathbf{P}Z_k \otimes A_k + \mathbf{M}_{k+1} \\ X_{k+1} = aX_k + b\langle Z_k, L \rangle + c\langle A_k, \mathbf{S} \rangle v_k. \end{cases} \quad (14)$$

Write

$$\begin{aligned} Z_k &= \sigma\{Z_0, Z_1, \dots, Z_k, A_0, A_1, \dots, A_k\} \\ F_k &= \sigma\{X_0, X_1, \dots, X_k\} \\ G_k &= \sigma\{Z_0, Z_1, \dots, Z_k, X_1, A_0, A_1, \dots, A_k, X_2, \dots, X_k\}. \end{aligned}$$

1. State estimation filters

The approach we take to compute our filters is the so-called reference probability method. This technique is widely used in Electrical Engineering, (Elliott *et al.*, 1995 and more recently Aggoun *et al.*, 2004). We define a probability measure P^\dagger on the measurable space (Ω, F) , such that, under P^\dagger , the following two conditions hold :

1. The state processes Z and A are Markov processes with transition matrices Π and P and initial distributions p_0 and \mathbf{p}_0 , respectively.
2. The observation process X is independently and identically distributed and is Gaussian with zero mean and unit variance.

With P^\dagger defined, we construct P , such that under P the following hold:

3. The state processes Z and A are Markov processes with transition matrices Π and P and initial distributions p_0 and \mathbf{p}_0 , respectively.
4. The sequence $\{v_1, v_2, \dots\}$, where

$$v_{\ell+1} = \frac{X_{\ell+1} - aX_\ell - b\langle Z_\ell, L \rangle}{c\langle A_\ell, \mathbf{S} \rangle}, \quad (15)$$

is a sequence of independently and identically distributed Gaussian $N(0,1)$ random variables.

Write

$$\phi(\xi) \triangleq \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \xi \xi'\right).$$

Definition 1. For $\ell = 1, 2, \dots$,

$$\lambda_\ell = \frac{\phi\left(\frac{X_{\ell+1} - aX_\ell - b\langle Z_\ell, L \rangle}{c\langle A_\ell, \mathbf{S} \rangle}\right)}{\langle A_\ell, \mathbf{S} \rangle \phi(X_{\ell+1})} \quad (16)$$

$$\Lambda_k = \prod_{\ell=0}^k \lambda_\ell, \quad \lambda_0 = 1. \quad (17)$$

The "real world" probability P is now defined in terms of the probability measure P^\dagger by setting

$$\frac{dP}{dP^\dagger} \Big|_{G_t} = \Lambda_k.$$

Lemma 1. Under P , the sequence $v = \{v_1, v_2, \dots\}$ is a sequence of independently and identically distributed $N(0,1)$ random variables, where

$$v_{k+1} = \frac{X_{k+1} - aX_k - b\langle Z_k, L \rangle}{c\langle A_k, \mathbf{S} \rangle}.$$

That is, under P ,

$$X_{k+1} = aX_k + b\langle Z_k, L \rangle + c\langle A_k, \mathbf{S} \rangle v_{k+1} \quad (18)$$

Lemma 2. Under the measure P , the process Z remains a Markov process, with transition matrix Π and initial distribution p_0 .

The proofs of Lemma 1 and 2 are routine.

Remark 1. The objective in estimation via reference probability is to choose a measure P^\dagger which facilitates and or simplifies calculations. In Filtering and Prediction, we wish to evaluate conditional expectations.

Under the measure P^\dagger , our dynamics have the form:

$$\begin{cases} Z_{k+1} = \Pi Z_k + M_k \\ A_{k+1} = P A_k + M_{k+1} \\ Z_{k+1} \otimes A_{k+1} = \mathbf{P} Z_k \otimes A_k + \mathbf{M}_{k+1} \\ X_{k+1} = v_{k+1} \end{cases}$$

In what follows we shall use the following version of Bayes' rule.

$$E[Z_k \otimes A_k / F_{k+1}] = \frac{E^\dagger[\Lambda_{k+1} Z_k \otimes A_k / F_{k+1}]}{E^\dagger[\Lambda_{k+1} / F_{k+1}]} = \frac{\sigma(Z_k \otimes A_k)}{\sigma(1)}.$$

Note that

$$\begin{aligned} \sum_{\ell=1}^m \sum_{r=1}^n \langle E^\dagger[\Lambda_{k+1} Z_k \otimes A_k / F_{k+1}], e_\ell \otimes f_r \rangle &= E^\dagger[\Lambda_{k+1} \sum_{\ell=1}^m \sum_{r=1}^n \langle Z_k \otimes A_k, e_\ell \otimes f_r \rangle / F_{k+1}] \\ &= E^\dagger[\Lambda_{k+1} / F_{k+1}]. \end{aligned}$$

Theorem 1. Information State Recursion. Suppose the Markov chain Z and A are observed through the unit-delay discrete-time dynamics at (10). The information state for the corresponding filtering problem is computed by the recursion:

$$\sigma(Z_k \otimes A_k) \triangleq E^\dagger[\Lambda_{k+1} Z_k \otimes A_k / F_{k+1}] = \Gamma_{k+1} \mathbf{P} \sigma(Z_{k-1} \otimes A_{k-1}). \quad (19)$$

Here

$$\Gamma_{k+1} = \text{diag}\{\gamma_{k+1}^{1,1}, \gamma_{k+1}^{1,2}, \dots, \gamma_{k+1}^{m,n}\}, \quad (20)$$

and

$$\gamma_{k+1}^{\ell,r} = \frac{\phi\left(\frac{X_{k+1} - aX_k - bL_\ell}{c\sigma_r}\right)}{\sigma_r \phi(X_{\ell+1})} \quad (21)$$

Proof of Theorem 1

$$\begin{aligned}
 \sigma(Z_k \otimes Z_k) &= E^\dagger[\Lambda_{k+1} Z_k \otimes A_k / F_{k+1}] \\
 &= E^\dagger[\Lambda_{k+1} \sum_{\ell=1}^m \sum_{r=1}^n \langle Z_k \otimes A_k, e_\ell \otimes f_r \rangle e_\ell \otimes f_r / F_{k+1}] \\
 &= \sum_{\ell=1}^m \sum_{r=1}^n \langle E^\dagger[\Lambda_{k+1} Z_k \otimes A_k / F_{k+1}], e_\ell \otimes f_r \rangle e_\ell \otimes f_r \\
 &= \sum_{\ell=1}^m \sum_{r=1}^n \langle E^\dagger[\Lambda_k \lambda_{k+1} Z_k \otimes A_k / F_{k+1}], e_\ell \otimes f_r \rangle e_\ell \otimes f_r \\
 &= \sum_{\ell=1}^m \sum_{r=1}^n \langle E^\dagger[\Lambda_k \frac{\phi(\frac{X_{k+1} - aX_k - bL_\ell}{c\sigma_r})}{\sigma_r \phi(X_{k+1})} \\
 &\quad \times Z_k \otimes A_k / F_{k+1}], e_\ell \otimes f_r \rangle e_\ell \otimes f_r
 \end{aligned} \tag{22}$$

Hence

$$\begin{aligned}
 \sigma(Z_k \otimes Z_k) &= \sum_{\ell=1}^m \sum_{r=1}^n \gamma_{k+1}^{\ell,r} \langle E^\dagger[\Lambda_k Z_k \otimes A_k / F_{k+1}], e_\ell \otimes f_r \rangle e_\ell \otimes f_r \\
 &= \sum_{\ell=1}^m \sum_{r=1}^n \gamma_{k+1}^{\ell,r} \langle E^\dagger[\Lambda_k (\mathbf{P}Z_{k-1} \otimes A_{k-1} + \mathbf{M}_k) / F_k], e_\ell \otimes f_r \rangle e_\ell \otimes f_r \\
 &= \sum_{\ell=1}^m \sum_{r=1}^n \gamma_{k+1}^{\ell,r} \mathbf{P} \sigma(Z_{k-1} \otimes A_{k-1}), e_\ell \otimes f_r \rangle e_\ell \otimes f_r \\
 &= \Gamma_{k+1} \mathbf{P} \sigma(Z_{k-1} \otimes A_{k-1})
 \end{aligned} \tag{23}$$

The recursion given in Theorem 1 provides a scheme to estimate the conditional probabilities for events of the form $\{\omega: Z_k \otimes A_k(\omega) = e_\ell \otimes f_r\}$, given the information F_{k+1} . In practice, one would use the vector-valued information state $\sigma(Z_k \otimes A_k)$, to compute an estimate for the state $Z_k \otimes A_k$. In general two approaches are adopted; one computes either a conditional mean, that is

$$\overline{Z_k \otimes A_k} = \frac{1}{\langle \sigma(Z_k \otimes A_k), (1, \dots, 1) \rangle} \tag{24}$$

$$\times \{ \langle \sigma(Z_k \otimes A_k), e_1 \otimes f_1 \rangle e_1 \otimes f_1, \dots, \langle \sigma(Z_k \otimes A_k), e_m \otimes f_n \rangle e_m \otimes f_n \}$$

or the so-called Maximum-a-Posteriori (MAP) estimate, that is

$$\overline{Z_k \otimes A_k} = \frac{1}{\langle \sigma(Z_k \otimes A_k), (1, \dots, 1) \rangle} \tag{25}$$

$$\times \operatorname{argmax} \{ \langle \sigma(Z_k \otimes A_k), e_1 \otimes f_1 \rangle e_1 \otimes f_1, \dots, \langle \sigma(Z_k \otimes A_k), e_m \otimes f_n \rangle e_m \otimes f_n \}$$

Marginal distributions for Z_k and Z_k are obtained by multiplying $\sigma(Z_k \otimes Z_k)$ on the right-hand side with the

n -dimensional row vector $(1, \dots, 1)$ or on the left-hand side with the m -dimensional column vector $(1, \dots, 1)$, respectively.

2. Prediction/Forecasting

What we would like to do is to predict the difference X in the *next* time period, and with this information, develop a trading strategy. Let us first compute the n -step predictor, where $n \in N \setminus \{0\}$.

Lemma 3. (n -Step Predictor)

$$\begin{aligned}
 E[Z_{k+n} / F_{k+1}] &= \mathbf{P}^n E[Z_k \otimes A_k / F_{k+1}] \\
 &= \mathbf{P}^n \frac{\sigma(Z_k \otimes A_k)}{\langle \sigma(Z_k \otimes A_k), 1 \rangle} \\
 &= \mathbf{P}^n \frac{\Gamma_{k+1} \mathbf{P} \sigma(Z_{k-1})}{\langle \Gamma_{k+1} \mathbf{P} \sigma(Z_{k-1}), 1 \rangle}
 \end{aligned} \tag{26}$$

Proof of Lemma 1

$$\begin{aligned}
 E[Z_{k+n} \otimes A_{k+n} / F_{k+1}] &= E[\mathbf{P}Z_{k+n-1} \otimes A_{k+n-1} + \mathbf{M}_{k+n} / F_{k+1}] \\
 &= \mathbf{P}E[\mathbf{P}Z_{k+n-1} \otimes A_{k+n-1} / F_{k+1}] \\
 &= \mathbf{P}E[\mathbf{P}Z_{k+1-2} / F_{k+1}] + E[\mathbf{M}_{k+n-1} / F_{k+1}] \\
 &= \mathbf{P}^2 E[\mathbf{P}Z_{k+n-2} \otimes A_{k+n-2} / F_{k+1}] \\
 &\quad \vdots \\
 &= \mathbf{P}^n E[Z_k \otimes A_k / F_{k+1}] \\
 &= \mathbf{P}^n \frac{\sigma(Z_k \otimes A_k)}{\langle \sigma(Z_k \otimes A_k), 1 \rangle}
 \end{aligned} \tag{27}$$

Recall now our discrete-time dynamics modeling a price difference process, *viz*

$$X_{k+1} = aX_k + b\langle Z_k, L \rangle + c\langle A_k, \mathbf{S} \rangle v_{k+1}.$$

The one-step prediction of the price difference is computed as follows:

$$\begin{aligned}
 \hat{X}_{k+1|k} &= E[X_{k+1} / F_k] \\
 &= E[aX_k + b\langle Z_k, L \rangle + c\langle A_k, \mathbf{S} \rangle v_{k+1} / F_k] \\
 &= aX_k + bE[\langle Z_k, L \rangle / F_k] + cE[\langle A_k, \mathbf{S} \rangle v_{k+1} / F_k] \\
 &= aX_k + b\langle E[Z_k / F_k], L \rangle \quad (\text{Predictor } n = 1) \\
 &= aX_k + b\langle \Pi E[Z_{k-1} / F_k], L \rangle \\
 &= aX_k + b\langle \Pi \frac{\sigma(Z_{k-1} \otimes A_{k-1}) \mathbf{1}_{n \times 1}}{\langle \sigma(Z_{k-1} \otimes A_{k-1}) \mathbf{1}_{n \times 1}, \mathbf{1}_{m \times 1} \rangle}, L \rangle.
 \end{aligned} \tag{28}$$

Remark 2. Here the usual issue of MAP/conditional-mean-estimate is irrelevant, as the price difference is continuously-valued.

3. Trading strategies

- The estimate $\hat{X}_{k+1|k}$ is the predicted value of the basis at the next time $k+1$, given observations X_0, X_1, \dots, X_k .
- If $\hat{X}_{k+1|k}$ is greater than the observed difference X_{k+1} , it suggests the higher price is too high and/or the lower price too low.
- Conversely, if $\hat{X}_{k+1|k}$ is smaller than the observed difference X_{k+1} , the higher price is possibly too low and/or the lower price too high.
- These observations might suggest trading strategies.

4. Acknowledgement

AGGOUN, L. and MALCOLM, W.P. would like to gratefully acknowledge to have benefited from technical discussions with Professor Robert J Elliott on the areas of modern quantitative finance, and MALCOLM W.P. would like to thank Sultan Qaboos University for their hospitality.

5. References

- AGGOUN, L., and ELLIOTT, R.J. 2004. *Measure Theory and Filtering: Introduction with Applications*. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge.
- ELLIOTT, C. ELLIOTT, R.J. and MALCOLM, W.P. 2005. Commodity prices and regime switching bases, Conference on Stochastic Modelling of Complex Systems, Daydream Island, Queensland Australia, 10-16, July 2005.
- ELLIOTT, R.J., AGGOUN, L. and MOORE, J.B. 1995. *Hidden Markov Models: Estimation and Control*. Springer-Verlag, New-York. Applications of Mathematics No. 29.
- ELLIOTT, R.J. 1982. *Stochastic Calculus and Applications*. Springer-Verlag, New-York. Applications of Mathematics No. 18.
- ELLIOTT, R.J., FISCHER, P. and PLATEN, E. 1999. Filtering and parameter estimation for a mean-reverting interest rate model, *Canadian Applied Mathematics Quarterly*, **7**: 381-400.
- ELLIOTT, R.J. and KOPP, P.E. 2005. *Mathematics of Financial Markets*, 2nd Edition, Springer-Verlag.
-

Received 25 March 2006

Accepted 22 September 2008