SQU Journal For Science, 15(2010) 87-100

Filtering and M-ary Detection of Markov Modulated Mean Reverting Model

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

**Department of Mathematics and Statistics, College of Science, Sultan Qaboos University, P.O. Box 36, Postal code 123, Muscat, Sultanate of Oman, Email: laggoun@squ.edu.om.*
***Australian National University, Canberra Australia, Email: Malcolm@maths.anu.edu.au.*

نموذج مركوفي للفرق بين الأسعار في سوق المال

الأخضر عجون، محمد اللواتي و و.ب. مالكولم

خلاصة: هذا البحث يطور نموذج يحتوي على سلسلتان من نوع مركوف تقوم بالتأثير على نموذج يقوم بدراسة حركة للفرق بين الأسعار في سوق المال.السلسلتان من نوع مركوف تقوم بتوصيف أحداث غير معروفة لكن ذات تأثير على الأسعارفي سوق المال. وتستعمل في هذا البحث طرق تغيير القياس لتقدير التوزيع الشرطي المتكرر.

ABSTRACT: In an earlier paper we developed a stochastic model incorporating a double-Markov modulated mean-reversion model. The model is based on an explicit discretisation of the corresponding continuous time dynamics. Here we discuss parameter estimation via the technique of M-ary detection.

KEYWORDS: Double-Markov modulated mean-reversion model, Filtering, M-Ary detection, Continuous-Time Dynamics.

mathematics subject classification. 60g35, 62m05, 62m20, 91b30, 91b70.

1. Introduction

The model we developed in Malcolm *et al.* (2004) is 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 discretization of the corresponding continuous time dynamics. In that model 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. In this paper we wish to discuss M-ary detection for this model. The term M-ary detection is used in Electrical Engineering to describe sequential hypothesis testing for more than two candidate model hypotheses. Here we are interested in model-parameter hypotheses. In effect our formulation is something like a discrete and finite version of the EM algorithm by Baum and Petrie (1966), Dempster *et al*. (1977) where, rather than considering an uncountable collection of model parameter sets in the space of all admissible models, we consider a finite collection in this space.

We assume that we have a list of M candidate models, from which to choose, describing the model dynamics over time. These candidate models will be denoted by H_h, $h = 1, \cdots, M$. Let β be a simple random variable denoting a specific model, with states indexed by $1 \leq h \leq M$. We assume that β is taking on values in the canonical basis $(b_1, \ldots . b_M)$ of R^M. We suppose β is an indicator random variable such that $\beta = b_h$, that is $\langle \beta, b_h \rangle = 1$ if and only if hypothesis H_h holds. Here $\langle , \rangle$ is the usual inner product. We shall be interested in computing the posterior probabilities$P(\beta = h \mid \mathcal{O}_n)$, where $\mathcal{O}_n$ denotes information contained in some observation process. It will be shown that this problem separates into a pure filtering component and a pure estimation component. In the context of M-ary detection, this is known as the Separation Theorem (Poor 1988). This paper is organized as follows. In §2 & §3 we recall the model dynamics as well as the construction of a new probability measure under which all processes are independent. In §4 M-ary Detection Filters are derived. In §5 & §6 our results are adapted to continuous time dynamics.

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:

$$X_t = X_0 + \alpha \int_0^t (L - X_u) du + \sigma W_t . \quad (2.1)$$

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

Remark 1. The dynamics at (2.1) exhibit a mean reversion[1] character of the model when written in stochastic differential equation form:

$$dX_t = \alpha (L - X_u)\, dt + \sigma dW_t . \quad (2.2)$$

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 side of is continually trying to reach the level L.

Now suppose that parameters L and σ are stochastic and can switch between different levels $L_1, L_2, \ldots, L_m$ and $\sigma_1, \ldots, \sigma_n$ respectively. We assume here that these levels are determined by the states of two Markov chains Z and $\mathfrak{Z}$ respectively.

Without loss of generality, we take the state spaces of our Markov chains to be the canonical basis $L = \{e_1, e_2, \ldots, e_m\}$ of R^m and the canonical basis $S = \{f_1, f_2, \ldots, f_n\}$ of R^n respectively.

[1] Modeling 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.

Write

$$\pi_{(j,i)} \triangleq P(Z_{k+1} = e_j \mid Z_k = e_i),$$
$$p_{(s,r)} \triangleq P(\mathfrak{Z}_{k+1} = f_s \mid \mathfrak{Z}_k = f_r), \tag{2.3}$$

$$\Pi = \left[\pi_{(j,i)}\right]_{\substack{1\le j\le m\\ 1\le i\le m}},$$
$$\mathcal{P} = \left[p_{(s,r)}\right]_{\substack{1\le s\le n\\ 1\le r\le n}}.$$

Write

$$\mathcal{Z}_t \triangleq \sigma\{Z_u, \mathfrak{Z}_u, 0 \le u \le t\}. \tag{2.4}$$

Then

$$Z_{k+1} = \Pi Z_k + M_{k+1},$$
$$\mathfrak{Z}_{k+1} = \mathcal{P}\mathfrak{Z}_k + \mathfrak{M}_{k+1}. \tag{2.5}$$

Here, M and $\mathfrak{M}$ are martingale increments.

The scalar-valued Markov processes taking values $L_1,\ldots,L_m$ and $\sigma_1,\ldots,\sigma_n$, are obtained by

$$\langle Z_t, \mathbf{L}\rangle = \sum_{\ell=1}^{m} \mathbf{1}_{\{\omega \mid Z_t(\omega)=e_\ell\}} L_\ell,$$
$$\langle \mathbf{S}, \mathfrak{Z}_t\rangle = \sum_{i=1}^{n} \mathbf{1}_{\{\omega:\ \mathfrak{Z}_t(\omega)=f_i\}} \sigma_i. \tag{2.6}$$

Here $\mathbf{L} = (L_1, L_2, \ldots, L_m)'$, $\mathbf{S} = (\sigma_1, \sigma_2, \ldots, \sigma_n)'$, $\langle\cdot,\cdot\rangle$ 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 $\mathfrak{Z}$ be not independent, that is, information on the behavior of one conveys some knowledge of the behavior of the other. More precisely, we assume the dynamics:

$$Z_{k+1} \otimes \mathfrak{Z}_{k+1} = \mathbf{P} Z_k \otimes \mathfrak{Z}_k + \mathbf{M}_{k+1}. \tag{2.7}$$

where $\mathbf{P} = \left(\mathbf{p}_{js,ir}\right)$ 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

$$\mathbf{p}_{js,ir} = P\left(Z_{k+1} = e_j,\ \mathfrak{Z}_{k+1} = f_s \mid Z_k = e_i,\ \mathfrak{Z}_k = f_r\right), 1 \le r, s \le n,\ 1 \le i, j \le m.$$

Again $\mathbf{M}_{k+1}$ is a martingale increment.

The dynamics at (1) take the form

$$X_t = X_0 + \alpha \int_0^t \Big(\langle Z_u, L\rangle - X_u\Big) du + \langle \mathbf{S}, \mathfrak{Z}_t\rangle W_t. \tag{2.8}$$

Remark 2. We defined Z and $\mathfrak{Z}$ as inherently discrete-time. Here, we "read" Z and $\mathfrak{Z}$ 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; see, for example, *Numerical Solution of Stochastic Differential Equations* by Kloeden and Platen (1992).

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

$$\mathcal{M}^{(K)} \triangleq \{0 = t_0, t_1, \ldots, t_K = T\}. \tag{2.9}$$

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

$$\begin{aligned} X_{k+1} &= X_k + \alpha\langle Z_k, L\rangle\Delta_t - \alpha X_k\Delta_t + \langle \mathfrak{Z}_k, \mathbf{S}\rangle\big(W_{k+1} - W_k\big) \\ &= aX_k + b\langle Z_k, L\rangle + \langle \mathfrak{Z}_k, \mathbf{c}\rangle v_k. \end{aligned} \tag{2.10}$$

Here

$$\begin{aligned} a &\triangleq \big(1 - \alpha\Delta_t\big), \\ b &\triangleq \alpha\Delta_t, \\ \mathbf{c} &\triangleq \big(\sqrt{\Delta_t}\langle\sigma, e_1\rangle, \ldots, \sqrt{\Delta_t}\langle\sigma, e_m\rangle\big)'. \end{aligned}$$

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

$$P \qquad \begin{cases} Z_{k+1} = \Pi Z_k + M_k \\ \mathfrak{Z}_{k+1} = \mathcal{P}\mathfrak{Z}_k + \mathfrak{M}_{k+1} \\ Z_{k+1} \otimes \mathfrak{Z}_{k+1} = \mathbf{P} Z_k \otimes \mathfrak{Z}_k + \mathbf{M}_{k+1} \\ X_{k+1} = aX_k + b\langle Z_k, L\rangle + \langle \mathfrak{Z}_k, \mathbf{c}\rangle v_k. \end{cases} \tag{2.11}$$

Write

$$\begin{aligned} \mathcal{Z}_k &\triangleq \sigma\big\{Z_0, Z_1, \ldots, Z_k, \mathfrak{Z}_0, \mathfrak{Z}_1, \ldots, \mathfrak{Z}_k\big\}, \\ \mathcal{F}_k &\triangleq \sigma\big\{X_0, X_1, \ldots, X_k\big\}, \\ \mathcal{G}_k &\triangleq \sigma\big\{Z_0, Z_1, \ldots, Z_k, \mathfrak{Z}_0, \mathfrak{Z}_1, \ldots, \mathfrak{Z}_k, X_1, X_2, \ldots, X_k\big\}. \end{aligned}$$

3. 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, see Elliott *et al.* (1995) and more recently Aggoun and Elliott (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 $\mathfrak{Z}$ are Markov chains 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 $\mathfrak{Z}$ are again Markov chains with initial distributions p_0 and $\mathbf{p}_0$ respectively.
4. The sequence v, where

$$v_{\ell+1} = \frac{X_{\ell+1} - aX_\ell - b\langle Z_\ell, L\rangle}{\langle \mathfrak{Z}_k, \mathbf{c}\rangle}, \tag{3.1}$$

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

Write

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

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

$$\lambda_\ell \triangleq \frac{\phi\left(\frac{X_{\ell+1} - aX_\ell - b\langle Z_\ell, L\rangle}{\langle \mathfrak{Z}_k, \mathbf{c}\rangle}\right)}{\langle \mathfrak{Z}_k, \mathbf{c}\rangle\, \phi(X_{\ell+1})} \tag{3.2}$$

$$\Lambda_k = \prod_{\ell=0}^{k} \lambda_\ell, \quad \lambda_0 = 1. \tag{3.3}$$

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

$$v_{k+1} \triangleq \frac{X_{k+1} - aX_\ell - b\langle Z_\ell, L\rangle}{c\langle \mathfrak{Z}_k, \mathbf{S}\rangle}.$$

That is, under P,

$$X_{k+1} = aX_k + b\langle Z_k, L\rangle + \langle \mathfrak{Z}_k, \mathbf{c}\rangle v_{k+1}. \tag{3.4}$$

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:

$$P^\dagger \qquad \begin{cases} Z_{k+1} = \Pi Z_k + M_k \\ \mathfrak{Z}_{k+1} = \mathcal{P}\mathfrak{Z}_k + \mathfrak{M}_{k+1} \\ Z_{k+1} \otimes \mathfrak{Z}_{k+1} = \mathbf{P} Z_k \otimes \mathfrak{Z}_k + \mathbf{M}_{k+1} \\ X_{k+1} = v_{k+1} \end{cases} \tag{3.5}$$

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

$$E\big[Z_k \otimes \mathfrak{Z}_k \mid \mathcal{F}_{k+1}\big] = \frac{E^\dagger\big[\Lambda_{k+1} Z_k \otimes \mathfrak{Z}_k \mid \mathcal{F}_{k+1}\big]}{E^\dagger\big[\Lambda_{k+1} \mid \mathcal{F}_{k+1}\big]} = \frac{q_k(Z_k \otimes \mathfrak{Z}_k)}{q_k(1)}. \tag{3.6}$$

Note that

$$\sum_{\ell=1}^{m}\sum_{r=1}^{n}\Big\langle E^\dagger\big[\Lambda_{k+1} Z_k \otimes \mathfrak{Z}_k \mid \mathcal{F}_{k+1}\big], e_\ell \otimes f_r\Big\rangle$$
$$= E^\dagger\big[\Lambda_{k+1}\sum_{\ell=1}^{m}\sum_{r=1}^{n}\langle Z_k \otimes \mathfrak{Z}_k, e_\ell \otimes f_r\rangle \mid \mathcal{F}_{k+1}\big] = E^\dagger\big[\Lambda_{k+1} \mid \mathcal{F}_{k+1}\big].$$

The following result is proven in Malcolm *et al.* (2008).

Theorem 1. *Information State Recursion*. Suppose the Markov chains Z and $\mathfrak{Z}$ are observed though the unit-delay discrete-time dynamics at (2.10). The information state for the corresponding filtering problem is computed by the recursion:

$$q_k(Z_k \otimes \mathfrak{Z}_k) \triangleq E^\dagger\big[\Lambda_{k+1} Z_k \otimes \mathfrak{Z}_k \mid \mathcal{F}_{k+1}\big] = \Gamma_{k+1}\mathbf{P} q_k(Z_{k-1} \otimes \mathfrak{Z}_{k-1}). \tag{3.7}$$

Here

$$\Gamma_{k+1} \triangleq \operatorname{diag}\{\gamma_{k+1}^{1,1}, \gamma_{k+1}^{1,2}, \ldots, \gamma_{k+1}^{m,n}\}, \tag{3.8}$$

and

$$\gamma_{k+1}^{\ell,r} \triangleq \frac{\phi\Big(\frac{X_{k+1} - aX_k - bL_\ell}{\langle f_r, \mathbf{c}\rangle}\Big)}{\langle f_r, \mathbf{c}\rangle\, \phi(X_{\ell+1})} \tag{3.9}$$

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

$$\begin{aligned}\widehat{Z_k \otimes \mathfrak{Z}_k} &\triangleq \frac{1}{\langle q_k(Z_k \otimes \mathfrak{Z}_k), (1, \ldots, 1)\rangle} \\ &\times \Big\{\langle q_k(Z_k \otimes \mathfrak{Z}_k), e_1 \otimes f_1\rangle e_1 \otimes f_1, \ldots, \langle q_k(Z_k \otimes \mathfrak{Z}_k), e_m \otimes f_n\rangle e_m \otimes f_n\Big\}\end{aligned} \tag{3.10}$$

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

$$\begin{aligned}\widehat{Z_k \otimes \mathfrak{Z}_k} &\triangleq \frac{1}{\langle q_k(Z_k \otimes \mathfrak{Z}_k), (1, \ldots, 1)\rangle} \\ &\times \operatorname{argmax}\Big\{\langle q_k(Z_k \otimes \mathfrak{Z}_k), e_1 \otimes f_1\rangle e_1 \otimes f_1, \ldots, \langle q_k(Z_k \otimes \mathfrak{Z}_k), e_m \otimes f_n\rangle e_m \otimes f_n\Big\}\end{aligned} \tag{3.11}$$

Marginal distributions for the Markov chains are obtained by multiplying $q_k(Z_k \otimes \mathfrak{Z}_k)$ on the right with the n-dimensional row vector $(1,\ldots,1)$ or on the left with the m-dimensional column vector $(1,\ldots,1)$ respectively.

4. *M-ary* Detection Filters

To denote a specific model hypothesis for the discrete-time dynamics given at (2.5), (2.7) and (2.10) we write,

$$H_j \triangleq \{\mathbf{P}^{H_j}, a^{H_j}, b^{H_j}, L^{H_j}, \mathbf{c}^{H_j}\}. \tag{4.1}$$

Here $j \in \{1, 2, \ldots, M\}$. Using the simple random variable α, as before, we are interested to compute the detector expectation

$$q_k = E^{\dagger}\big[\Lambda_k \langle \alpha, \mathfrak{f}_j \rangle \mid \mathcal{Y}_k\big]. \tag{4.2}$$

Here the sigma algebra $\mathcal{Y}_k$ is taken as generated by a model with parameter set H_j, and similarly the Radon-Nikodym derivative Λ_k, is constructed according to H_j. Further, to make a clear distinction between the filter information state defined for specific model H_j, and the corresponding un-normalised detector probability for model H_j, we write, respectively

$$q_k^{H_j} = E^{\dagger}\big[\Lambda_{k+1} Z_k \otimes \mathfrak{z}_k \mid \mathcal{Y}_{k+1}\big], \quad \langle q_k^{\mathrm{Det}}, \mathfrak{f}_j \rangle = \langle E^{\dagger}\big[\Lambda_k \langle \alpha, \mathfrak{f}_j \rangle \mid \mathcal{Y}_k\big], \mathfrak{f}_j \rangle.$$

Theorem 2 (*M-ary Detection Filter*)
The M-ary detection filter for the model hypothesis H_j is computed by the recursion

$$\langle q_{k+1}^{\mathrm{Det}}, \mathfrak{f}_j \rangle = \sum_{r=1}^{n} \sum_{\ell=1}^{m} \frac{\phi\Big(\dfrac{X_{k+1} - a^{H_j} X_k - b^{H_j} \langle e_\ell, L^{H_j} \rangle}{\langle f_r, \mathbf{c}^{H_j} \rangle}\Big)}{\langle f_r, \mathbf{c}^{H_j} \rangle \, \phi(X_{k+1})} \frac{\langle \mathbf{P}^{H_j} q_{k-1}^{H_j}, e_\ell \otimes f_r \rangle}{\langle \mathbf{P}^{H_j} q_{k-1}^{H_j}, (1, 1, \ldots, 1) \rangle} \times \langle q_k^{\mathrm{Det}}, \mathfrak{f}_j \rangle.$$

Proof:

$$\begin{aligned}
\langle q_{k+1}^{\mathrm{Det}} \mathfrak{f}_j \rangle &= E^{\dagger}\big[\langle \alpha, \mathfrak{f}_j \rangle \Lambda_{k+1} \mid \mathcal{Y}_{k+1}\big] \\
&= E^{\dagger}\big[\langle \alpha, \mathfrak{f}_j \rangle \Lambda_k \lambda_{k+1} \mid \mathcal{Y}_{k+1}\big]
\end{aligned}$$

$$= E^{\dagger}\left[\langle \alpha, \mathfrak{f}_j \rangle \Lambda_k \left\{ \frac{\phi\Big(\dfrac{X_{k+1} - a^{H_j} X_k - b^{H_j} \langle Z_k, L^{H_j} \rangle}{\langle \mathfrak{z}_k, \mathbf{c}^{H_j} \rangle}\Big)}{\langle \mathfrak{z}_k, \mathbf{c}^{H_j} \rangle \, \phi(X_{k+1})} \right\} \mid \mathcal{Y}_{k+1}\right]$$

$$= E\left[\langle \alpha, \mathfrak{f}_j \rangle \left\{ \frac{\phi\Big(\dfrac{X_{k+1} - a^{H_j} X_k - b^{H_j} \langle Z_k, L^{H_j} \rangle}{\langle \mathfrak{z}_k, \mathbf{c}^{H_j} \rangle}\Big)}{\langle \mathfrak{z}_k, \mathbf{c}^{H_j} \rangle \, \phi(X_{k+1})} \right\} \mid \mathcal{Y}_k\right] E^{\dagger}\big[\Lambda_k \mid \mathcal{Y}_k\big]$$

$$= E\left[\left\{ \frac{\phi\Big(\dfrac{X_{k+1} - a^{H_j} X_k - b^{H_j} \langle Z_k, L^{H_j} \rangle}{\langle \mathfrak{z}_k, \mathbf{c}^{H_j} \rangle}\Big)}{\langle \mathfrak{z}_k, \mathbf{c}^{H_j} \rangle \, \phi(X_{k+1})} \right\} \mid \alpha = \mathfrak{f}_j \,\&\, \mathcal{Y}_k\right] \times$$
$$E\big[\langle \alpha, \mathfrak{f}_j \rangle \mid \mathcal{Y}_k\big] E^{\dagger}\big[\Lambda_k \mid \mathcal{Y}_k\big]$$

$$= E\left[\left\{ \frac{\phi\Big(\dfrac{X_{k+1} - a^{H_j} X_k - b^{H_j} \langle Z_k, L^{H_j} \rangle}{\langle \mathfrak{z}_k, \mathbf{c}^{H_j} \rangle}\Big)}{\langle \mathfrak{z}_k, \mathbf{c}^{H_j} \rangle \, \phi(X_{k+1})} \right\} \mid \alpha = \mathfrak{f}_j \,\&\, \mathcal{Y}_k\right] \langle q_k^{\mathrm{Det}}, \mathfrak{f}_j \rangle.$$

The expectation in the last line of the calculation is

$$E\left[\left\{\frac{\phi\Big(\dfrac{X_{k+1}-a^{H_j}X_k-b^{H_j}\langle Z_k,L^{H_j}\rangle}{\langle \mathfrak{Z}_k,\mathbf{c}^{H_j}\rangle}\Big)}{\langle \mathfrak{Z}_k,\mathbf{c}^{H_j}\rangle\,\phi(X_{k+1})}\right\}\mid \alpha=\mathfrak{f}_j\,\&\,\mathcal{Y}_k\right]$$
$$=\sum_{r=1}^{n}\sum_{\ell=1}^{m}\frac{\phi\Big(\dfrac{X_{k+1}-a^{H_j}X_k-b^{H_j}\langle e_\ell,L^{H_j}\rangle}{\langle f_r,\mathbf{c}^{H_j}\rangle}\Big)}{\langle f_r,\mathbf{c}^{H_j}\rangle\,\phi(X_{k+1})}P(Z_k\otimes\mathfrak{Z}_k=e_\ell\otimes f_r\mid\alpha=\mathfrak{f}_j\,\&\,\mathcal{Y}_k).$$

The normalized probabilities $P(Z_k\otimes\mathfrak{Z}_k=e_\ell\otimes f_r\mid\alpha=\mathfrak{f}_j\,\&\,\mathcal{Y}_k)$ are computed by the normalized one step predictor information state, that is, for the model hypothesis H_j and the event $Z_k\otimes\mathfrak{Z}_k=e_\ell\otimes f_r$, we compute

$$\begin{aligned}
&P\big(Z_k\otimes\mathfrak{Z}_k=e_\ell\otimes f_r\mid\alpha=\mathfrak{f}_j\,\mathcal{Y}_k\big)\\
&=\langle E\big[Z_k\otimes\mathfrak{Z}_k\mid\alpha=\mathfrak{f}_j\,\&\,\mathcal{Y}_k\big],e_\ell\otimes f_r\rangle\\
&=\frac{\langle E^\dagger\big[\Lambda_k Z_k\otimes\mathfrak{Z}_k\mid\alpha=\mathfrak{f}_j\,\&\,\mathcal{Y}_k\big],e_\ell\otimes f_r\rangle}{E^\dagger\big[\Lambda_k\mid\mathcal{Y}_k\big]}\\
&=\frac{\langle E^\dagger\big[\Lambda_k\big(\mathbf{P}^{H_j}Z_{k-1}\otimes\mathfrak{Z}_{k-1}+\mathbf{M}_k\big)\mid\mathcal{Y}_k\big],e_\ell\otimes f_r\rangle}{E^\dagger\big[\Lambda_k\mid\mathcal{Y}_k\big]}\\
&=\frac{\langle\mathbf{P}^{H_j}E^\dagger\big[\Lambda_k Z_{k-1}\otimes\mathfrak{Z}_{k-1}\mid\mathcal{Y}_k\big],e_\ell\otimes f_r\rangle}{E^\dagger\big[\Lambda_k\mid\mathcal{Y}_k\big]}\\
&=\frac{\langle\mathbf{P}^{H_j}q_{k-1}^{H_j},e_\ell\otimes f_r\rangle}{\langle\mathbf{P}^{H_j}q_{k-1}^{H_j},(1,1,\ldots,1)\rangle}
\end{aligned}$$

Here q_k is the information state for the filter computed earlier. Since we need the normalised form of the expectation at (4.2), the M-ary detector has the form:

$$\langle q_{k+1}^{\mathrm{Det}},\mathfrak{f}_j\rangle=\sum_{r=1}^{n}\sum_{\ell=1}^{m}\frac{\phi\Big(\dfrac{X_{k+1}-a^{H_j}X_k-b^{H_j}\langle e_\ell,L^{H_j}\rangle}{\langle f_r,\mathbf{c}^{H_j}\rangle}\Big)}{\langle f_r,\mathbf{c}^{H_j}\rangle\,\phi(X_{k+1})}\frac{\langle\mathbf{P}^{H_j}q_{k-1}^{H_j},e_\ell\otimes f_r\rangle}{\langle\mathbf{P}^{H_j}q_{k-1}^{H_j},(1,1,\ldots,1)\rangle}\times\langle q_k^{\mathrm{Det}},\mathfrak{f}_j\rangle.$$

5. Continuous-Time Dynamics

We consider here a continuous time Markov chain Z. Again we use the canonical representation of an arbitrary Markov chain. That is, without loss of generality we take the state space for Z to be the set $\mathcal{L}=\{e_1,e_2,\ldots,e_n\}$, whose elements e_i are column vectors with unity in the i^{th} position and zero elsewhere. The key benefit of this representation is that it admits the dynamics:

$$Z_t = Z_0 + \int_0^t A Z_u du + V_t.$$

Here V_t is a $(P, \sigma\{Z_u, 0 \leq u \leq t\})$-martingale and $A \in \mathbb{R}^{n\times n}$ is a time invariant rate matrix, whose elements are the infinitesimal intensities of X. To denote an element of the matrix A at row i and column j, we write $\langle Ae_i, e_j\rangle$. Here $\langle\cdot\,,\,\cdot\rangle$ denotes an inner product.
Now we consider the continuous-time dynamics

$$X_t = X_0 + \alpha \int_0^t \Big(\langle Z_u, L\rangle - X_u\Big) du + \sigma W_t \tag{5.1}$$

Under $P^\dagger$ the state and observation process dynamics have the form:

$$P^\dagger \quad \begin{cases} dZ_t = A\, Z_t\, dt + dV_t, \\ dX_t = \sigma dW_t. \end{cases}$$

Let

$$\Lambda_t = 1 + \int_0^t \Lambda_u \Big(\langle Z_u, L\rangle - X_u\Big)\, dX_u$$

where X is given by equation (5.1).
Then the 'real world' probability P is defined via

$$\left.\frac{dP}{dP^\dagger}\right|_{\mathcal{G}_t} = \Lambda_t.$$

Under P the dynamics have the form:

$$P \quad \begin{cases} dZ_t = A\, Z_t dt + dV_t, \\ dX_t = \Big(\langle Z_t, L\rangle - X_t\Big)\, dt + \sigma dW_t \end{cases}$$

Notation: Suppose $H = \{H_u, 0 \leq u\}$ is any $\mathcal{G}$-adapted process and we wish to estimate $E\big[H_t \mid \mathcal{Y}_t\big]$. Using Bayes' rule (Elliott *et al.* 1995)

$$E\big[H_t \mid \mathcal{Y}_t\big] = \frac{E^\dagger\big[\Lambda_t H_t \mid \mathcal{Y}_t\big]}{E^\dagger\big[\Lambda_t \mid \mathcal{Y}_t\big]} = \frac{\sigma(H_t)}{\sigma(1)}.$$

6. Continuous-Time Detection Schemes

State Estimation Filters

With $\quad q_t \triangleq E^\dagger\big[\Lambda_t Z_t \mid \mathcal{Y}_t\big] \in \mathbb{R}^n$

$$q_t = q_0 + \int_0^t A\, q_u\, du + \int_0^t \mathrm{diag}\{L - X_u\} q_u\, dX_u \ \in \mathbb{R}^n.$$

Then
$$P(Z_t = e_i \mid \mathcal{Y}_t) = \frac{\langle q_t, e_i \rangle}{\sum_{\ell=1}^{n} \langle q_t, e_\ell \rangle}.$$

M-ary Detection Filters

define a matrix-valued process $\mathbf{Z} = \{\mathbf{Z}_u, 0 \le u \le t\}$, where

$$\mathbf{Z}_t \triangleq \alpha Z_t' = \begin{bmatrix} \langle \alpha, f_1 \rangle \langle Z_t, e_1 \rangle & \dots & \langle \alpha, f_1 \rangle \langle Z_t, e_n \rangle \\ \langle \alpha, f_2 \rangle \langle Z_t, e_1 \rangle & \dots & \langle \alpha, f_2 \rangle \langle Z_t, e_n \rangle \\ \vdots & \vdots \ \ddots & \vdots \\ \langle \alpha, f_M \rangle \langle Z_t, e_1 \rangle & \dots & \langle \alpha, f_M \rangle \langle Z_t, e_n \rangle \end{bmatrix} \subseteq \mathbb{R}^{M \times n}.$$

Here α is the simple random variable defined above. The state space for the process $\mathbf{Z}$ is a canonical basis of matrix-valued indicator functions $F_{(j,i)} = f_j e_i'$

$Z \in \mathcal{M} = \{F_{(j,i)}\}$

$$= \left\{ \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \dots & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & \dots & \dots & 0 \\ \vdots & \dots & \dots & \vdots \\ 0 & & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{bmatrix}, \right.$$
$$\vdots$$
$$\left. \begin{bmatrix} 0 & \dots & 0 & 1 \\ \vdots & \dots & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & \dots & \dots & 0 \\ \vdots & \dots & \dots & \vdots \\ \vdots & & \dots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix} \right\}.$$

The process Z takes values on a canonical basis of matrix-valued indicator functions, each of which jointly indicates a particular model hypothesis, and a particular value taken by the state process.

Write

$$\langle q_t^{\text{Det}}, f_j \rangle \triangleq E^{\dagger}\big[\Lambda_t \langle \alpha, f_j \rangle \mid \mathcal{Y}_t\big].$$

The unnormalised probability process $\langle q_t^{\text{Det}}, f_j \rangle$, satisfies the stochastic integral equation

$$\langle q_t^{\text{Det}}, f_j \rangle = \langle q_0^{\text{Det}}, f_j \rangle + \int_0^t \left(\sum_{i=1}^n L_i^{H_j} \widehat{\langle Z_u, e_i \rangle} - X_u \right) \langle q_u^{\text{Det}}, f_j \rangle dX_u,$$

where $\widehat{\langle Z_u, e_i \rangle} = E\big[\langle Z_u, e_i \rangle \mid \mathcal{Y}_u, \alpha = f_j\big]$ is evaluated under the probability measure P, given that the hypothesis H_j holds.

The corresponding normalized detection probabilities are computed, for example, by

$$P\big(\alpha = f_j \mid \mathcal{Y}_t\big) = \frac{\langle q_t^{\text{Det}}, f_j \rangle}{\langle q_t^{\text{Det}}, 1 \rangle}$$

Write

$$\mathbf{q}_t \triangleq E^{\dagger}\big[\Lambda_t \mathbf{Z}_t \mid \mathcal{Y}_t\big] = E^{\dagger}\big[\Lambda_t\, \alpha\, Z_t' \mid \mathcal{Y}_t\big] \in \mathbb{R}^{M \times n}, \tag{6.1}$$

define

$$G = \begin{bmatrix} \langle L^{H_1}, e_1 \rangle & \langle L^{H_1}, e_2 \rangle & \dots & \langle L^{H_1}, e_n \rangle \\ \langle L^{H_2}, e_1 \rangle & \langle L^{H_2}, e_2 \rangle & \dots & \langle L^{H_2}, e_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle L^{H_M}, e_1 \rangle & \langle L^{H_M}, e_2 \rangle & \dots & \langle L^{H_M}, e_n \rangle \end{bmatrix} \in \mathbb{R}^{M \times n}.$$

The process $\mathbf{q}$, defined by equation (6.1) satisfies the dynamics

$$\mathbf{q}_t = \mathbf{q}_0 + \sum_{j=1}^m \sum_{i=1}^n \int_0^t \big(f_j' \mathbf{q}_u e_i\big) F_{(j,i)} A_{H_j}' du + \int_0^t G \odot \mathbf{q}_u \, dX_u - \int_0^t \mathbf{q}_u \, X_u dX_u.$$

The symbol $\odot$ in the previous equation denotes a point-wise matrix product, where for two matrices of the same dimensions, the point-wise product is

$$A \odot B = [a_{(i,j)} b_{(i,j)}]_{\substack{i=1,\dots,n \\ j=1,\dots,m}}.$$

Write

$$\mathbf{1}_M = (1,1,\ldots,1)' \in \mathbb{R}^M,$$
$$\mathbf{1}_n = (1,1,\ldots,1)' \in \mathbb{R}^n.$$

Recalling the numerator in Bayes' rule, we note that

$$\begin{aligned}\mathbf{1}_M' E^{\dagger}\big[\Lambda_t \alpha \mathbf{Z}_t' \mid \mathcal{Y}_t\big]\mathbf{1}_n &= E^{\dagger}\big[\Lambda_t \mathbf{1}_M' \alpha \mathbf{Z}_t' \mathbf{1}_n \mid \mathcal{Y}_t\big] \\ &= E^{\dagger}\big[\Lambda_t \mid \mathcal{Y}_t\big].\end{aligned}$$

So, by computing the numerator in Bayes' rule, we can readily compute the normalizing denominator $E^{\dagger}\big[\Lambda_t \mid \mathcal{Y}_t\big]$. The matrix quantity$\mathbf{q}$, defined at (6.1), is an un-normalized conditional expectation, so, the corresponding normalized conditional expectation is computed by

$$E\big[\mathbf{Z}_t \mid \mathcal{Y}_t\big] = \frac{\mathbf{q}_t}{\mathbf{1}_M' \mathbf{q}_t \mathbf{1}_n}.$$

To recover the normalized M-ary detection probabilities from the quantity $\mathbf{q}_t$, one computes

$$\left\{\frac{\mathbf{q}_t}{\mathbf{1}_M' \mathbf{q}_t \mathbf{1}_n}\right\}\mathbf{1}_n = \begin{bmatrix} P(\alpha = f_1 \mid \mathcal{Y}_t) \\ P(\alpha = f_2 \mid \mathcal{Y}_t) \\ \vdots \\ P(\alpha = f_M \mid \mathcal{Y}_t) \end{bmatrix}.$$

The corresponding normalized detection probabilities are computed, for example, by

$$P\big(\alpha = f_j \mid \mathcal{Y}_t\big) = \frac{\langle q_t^{\mathrm{Det}}, f_j\rangle}{\langle q_t^{\mathrm{Det}}, 1\rangle}$$

Write

$$\mathbf{q}_t \triangleq E^{\dagger}\big[\Lambda_t \mathbf{Z}_t \mid \mathcal{Y}_t\big] = E^{\dagger}\big[\Lambda_t\, \alpha\, Z_t' \mid \mathcal{Y}_t\big] \in \mathbb{R}^{M\times n}, \tag{6.1}$$

define

$$G = \begin{bmatrix} \langle L^{H_1}, e_1\rangle & \langle L^{H_1}, e_2\rangle & \ldots & \langle L^{H_1}, e_n\rangle \\ \langle L^{H_2}, e_1\rangle & \langle L^{H_2}, e_2\rangle & \ldots & \langle L^{H_2}, e_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle L^{H_M}, e_1\rangle & \langle L^{H_M}, e_2\rangle & \ldots & \langle L^{H_M}, e_n\rangle \end{bmatrix} \in \mathbb{R}^{M\times n}.$$

The process $\mathbf{q}$, defined by equation (6.1) satisfies the dynamics

$$\mathbf{q}_t = \mathbf{q}_0 + \sum_{j=1}^{m}\sum_{i=1}^{n}\int_0^t \left(f_j'\mathbf{q}_u e_i\right)F_{(j,i)}A'_{H_j}du + \int_0^t G \odot \mathbf{q}_u\, dX_u$$
$$- \int_0^t \mathbf{q}_u\, X_u dX_u.$$

The symbol $\odot$ in the previous equation denotes a point-wise matrix product, where for two matrices of the same dimensions, the point -ise product i

$$A \odot B = [a_{(i,j)}b_{(i,j)}]_{\substack{i=1,\dots,n \\ j=1,\dots,m}}.$$

Write

$$\mathbf{1}_M = (1,1,\dots,1)' \in \mathbb{R}^M,$$
$$\mathbf{1}_n = (1,1,\dots,1)' \in \mathbb{R}^n.$$

Recalling the numerator in Bayes' rule, we note that

$$\mathbf{1}'_M E^\dagger\left[\Lambda_t\alpha\mathbf{Z}'_t \mid \mathcal{Y}_t\right]\mathbf{1}_n = E^\dagger\left[\Lambda_t\mathbf{1}'_M\alpha\mathbf{Z}'_t\mathbf{1}_n \mid \mathcal{Y}_t\right]$$
$$= E^\dagger\left[\Lambda_t \mid \mathcal{Y}_t\right].$$

So, by computing the numerator in Bayes' rule, we can readily compute the normalising denominator$E^\dagger\left[\Lambda_t \mid \mathcal{Y}_t\right]$. The matrix quantity$\mathbf{q}$, defined at (6.1), is an un-normalized conditional expectation, so, the corresponding normalized conditional expectation is computed by

$$E\left[\mathbf{Z}_t \mid \mathcal{Y}_t\right] = \frac{\mathbf{q}_t}{\mathbf{1}'_M\mathbf{q}_t\mathbf{1}_n}.$$

To recover the normalized M-ary detection probabilities from the quantity $\mathbf{q}_t$, one computes

$$\left\{\frac{\mathbf{q}_t}{\mathbf{1}'_M\mathbf{q}_t\mathbf{1}_n}\right\}\mathbf{1}_n = \begin{bmatrix} P(\alpha = f_1 \mid \mathcal{Y}_t) \\ P(\alpha = f_2 \mid \mathcal{Y}_t) \\ \vdots \\ P(\alpha = f_M \mid \mathcal{Y}_t) \end{bmatrix}.$$

6. References

AGGOUN, L. and ELLIOTT, R.J. 2004 *Measure Theory and Filtering: Introduction with Applications.* Cambridge Series in Statistical and Probabilistic Mathematics.

BAUM, L.E. and PETRIE, T. 1966 Statistical Inference for Probabilistic Functions of Finite State Markov Chains. *Annals of the Institute of Statistical Mathematics*, **37:** 1554-1563.

DEMP, A.P., LAIRD, N.M. and RUBIN, D.B. 1977 Maximum Likelihood from Incomplete Data via the EM algorithm. *Jour. of the Royal Statistical Society* **B 39**: 1-38.

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*.

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.

KLOEDEN, P.E. and PLATEN, E. 1992. *Numerical Solution of Stochastic Differential Equations* Springer-Verlag, New-York. Applications of Mathematics No. 23.

MALCOLM, W.P., AGGOUN, L., and Al-LAWATIA, M. 2008. On A Markov Modulated Mean Reverting Finance Model. *Journal of Science and Technology, Sultan Qaboos University*. **13:** 55-62 .

POOR, H. 1988. *An Introduction to Signal Detection and Estimation*, Springer-Verlag, New York.

Received 28 April 2009
Accepted 8 December 2009