LINEAR FILTERING OF SYSTEMS WITH MEMORY

A. INOUE, Y. NAKANO AND V. ANH

ABSTRACT. We study the linear filtering problem for systems driven by continuous Gaussian processes V_1 and V_2 with memory described by two parameters. The processes V_j have the virtue that they possess stationary increments and simple semimartingale representations simultaneously. It allows for straightforward parameter estimations. After giving the semimartingale representations of V_j by innovation theory, we derive Kalman-Bucy-type filtering equations for the systems. We apply the result to the optimal portfolio problem for an investor with partial observations. We illustrate the tractability of the filtering algorithm by numerical implementations.

1. INTRODUCTION

In this paper, we use the following Gaussian process $(V(t))_{t \in \mathbf{R}}$ with stationary increments as the driving noise process:

$$(1.1) \quad V(t) = W(t) - \int_0^t \left(\int_{-\infty}^s p e^{-(q+p)(s-u)} dW(u) \right) ds \qquad (t \in \mathbf{R}),$$

where p and q are real constants such that

$$(1.2) \qquad\qquad 0 < q < \infty, \quad -q < p < \infty,$$

and $(W(t))_{t \in \mathbf{R}}$ is a one-dimensional Brownian motion satisfying $W(0) = 0$. The parameters p and q describe the memory of $V(\cdot)$. In the simplest case $p = 0$, $V(\cdot)$ is reduced to the Brownian motion, i.e., $V(\cdot) = W(\cdot)$.

It should be noticed that (1.1) is not a semimartingale representation of $(V(t))_{0 \le t \le T}$ with respect to the natural filtration $\mathcal{F}^V(\cdot)$ of $(V(t))_{0 \le t \le T}$ since $(W(t))_{0 \le t \le T}$ is not $\mathcal{F}^V(\cdot)$-adapted. Using the innovation theory as described in Liptser and Shiryayev [21] and a result in Anh et al. [2], we show (Theorem 2.1) that there exists a one-dimensional Brownian motion $(B(t))_{0 \le t \le T}$, called the *innovation process*, satisfying

$$\sigma(B(s) : 0 \le s \le t) = \sigma(V(s) : 0 \le s \le t) \qquad (0 \le t \le T),$$

and

$$(1.3) \qquad V(t) = B(t) - \int_0^t \left(\int_0^s l(s,u) dB(u) \right) ds \qquad (0 \le t \le T),$$

where $l(t,s)$ is a Volterra kernel given by

$$(1.4) \quad l(t,s) = pe^{-(p+q)(t-s)} \left\{ 1 - \frac{2pq}{(2q+p)^2 e^{2qs} - p^2} \right\} \qquad (0 \le s \le t \le T).$$

With respect to the natural filtration $\mathcal{F}^B(\cdot)$ of $B(\cdot)$, which is equal to $\mathcal{F}^V(\cdot)$, (1.3) gives the semimartingale representation of $V(\cdot)$. Thus the process $V(\cdot)$ has the virtue that it possesses the property of a stationary increment process with memory and a simple semimartingale representation simultaneously. We know no other process with this kind of properties. The two properties of $V(\cdot)$ become a great advantage, for example, in its parameter estimation.

In [1] and Anh et al. [2, 3], the process $V(\cdot)$ is used as the driving process for a financial market model with memory. The empirical study for S&P 500 data in Anh et al. [4] shows that the model captures very well the memory effect when the market is stable. The work in these references suggests that the process $V(\cdot)$ can serve as an alternative to Brownian motion when the random disturbance exhibits dependence between different observations.

In this paper, we are concerned with the filtering problem of the two-dimensional partially observable process $(X(t), Y(t))_{0 \le t \le T}$ governed by the following linear system of equations:

$$(1.5) \qquad \begin{cases} dX(t) = \theta X(t) dt + \sigma dV_1(t), & X(0) = X_0, \\ dY(t) = \mu X(t) dt + dV_2(t), & Y(0) = 0. \end{cases}$$

Here $X(\cdot)$ and $Y(\cdot)$ represent the state and the observation respectively. For $j = 1, 2$, the noise process $V_j(\cdot)$ is described by (1.1) with $(p,q) = (p_j, q_j)$ and $W(\cdot) = W_j(\cdot)$. We assume that the Brownian motions $W_1(\cdot)$ and $W_2(\cdot)$, whence $V_1(\cdot)$ and $V_2(\cdot)$, are independent. The coefficients $\theta, \sigma, \mu \in \mathbf{R}$ with $\mu \neq 0$ are known constants, and X_0 is a centered Gaussian random variable independent of (V_1, V_2).

The basic filtering problem for the linear model (1.5) with memory is to calculate the conditional expectation $\mathbb{E}[X(t)|\mathcal{F}^Y(t)]$, called the (optimal) *filter*, where $\mathcal{F}^Y(\cdot)$ is the natural filtration of the observation process $Y(\cdot)$. In the classical Kalman-Bucy theory (see Kalman [12], Kalman and Bucy [13], Bucy and Joseph [5], Davis [6] and [21]), Brownian motion is used as the driving noise. Attempts have been made to resolve the filtering problem of dynamical systems with memory by replacing Brownian motion by other processes. In Kleptsyna et al. [16, 17, 18] and others, fractional Brownian motion was used. Notice that fractional Brownian

motion is not a semimartingale. In the discrete-time setting, autoregressive processes are used as driving noise (see, e.g., [12], [5] and Jazwinski [11]). Our model may be regarded as a continuous-time analogue of the latter since it is shown in [1] that $V(\cdot)$ is governed by a continuous-time AR(∞)-type equation.

The Kalman-Bucy filter is a computationally tractable scheme for the optimal filter of a Markovian system. We aim to derive a similar effective filtering algorithm for the system (1.5) which has memory. However, rather than considering (1.5) itself, we start with a general continuous Gaussian process $(X(t))_{0 \leq t \leq T}$ as the state process and Y defined by

$$Y(t) = \int_0^t \mu(s)X(s)ds + V(t)$$

as the observation process, where $\mu(\cdot)$ is a deterministic function and $V(\cdot)$ is a process which is independent of $X(\cdot)$ and given by (1.1). Using (1.3) and (1.4), we derive explicit Volterra integral equations for the optimal filter (Theorem 3.1). In the special case (1.5), the integral equations are reduced to differential equations, which gives an extension to Kalman-Bucy filtering equations (Theorem 3.4). Due to the non-Markovianness of the formulation (1.5), it is expected that the resulting filtering equations would appear in the integral equation form (cf. Kleptsyna et al. [16]). The fact that good Kalman-Bucy-type differential equations can be obtained here is due the special properties of (1.5). This interesting result does not seem to hold for any other formulation where memory is inherent.

We apply the results to an optimal portfolio problem in a partially observable financial market model. More precisely, we consider a stock price model that is driven by the solution $V(\cdot)$ to (1.1). Assuming that the investor can observe the stock price but not the drift process, we discuss the portfolio optimization problem of maximizing the expected logarithmic utility from terminal wealth. To solve this problem, we make use of our results on filtering to reduce the problem to the case where the drift process is adapted to the observation process. We then use the martingale methods (cf. Karatzas and Shreve [14]) to work out the explicit formula for the optimal portfolio (Theorem 4.1).

This paper is organized as follows. In Section 2, we prove the semimartingale representation (1.3) with (1.4) for $V(\cdot)$. Section 3 is the main body of this paper. We derive closed form equations for the optimal filter. In Section 4, we apply the results to finance. In Section 5, we illustrate the advantage of $V(\cdot)$ in parameter estimation. Some numerical results on Monte Carlo simulation are presented. Finally, in Section 6, we numerically compare the performance of our filter with the Kalman-Bucy filter in the presence of memory effect.

2. Driving noise process with memory

Let $T \in (0, \infty)$, and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. For a process $(A(t))_{0 \le t \le T}$, we denote by $\mathcal{F}^A(t)$ the $\mathbb{P}$-augmentation of the filtration $\sigma(A(s) : 0 \le s \le t)$, $0 \le t \le T$.

Let $V(\cdot)$ be the process given by (1.1). The process $(V(t))_{t \in \mathbf{R}}$ is a continuous Gaussian process with stationary increments. The aim of this section is to prove (1.3) with (1.4).

By [21, Theorem 7.16], there exists a one-dimensional Brownian motion $(B(t))_{0 \le t \le T}$, called the innovation process, such that $\mathcal{F}^B(t) = \mathcal{F}^V(t)$, $0 \le t \le T$ and that

$$(2.1) \qquad V(t) = B(t) - \int_0^t \alpha(s)ds, \qquad (0 \le t \le T),$$

$$\alpha(t) = \mathbb{E}\left[\left. \int_{-\infty}^t pe^{-(q+p)(t-u)}dW(u) \right| \mathcal{F}^V(t)\right] \qquad (0 \le t \le T).$$

Thus, $(V(t))_{0 \le t \le T}$ is a $\mathcal{F}^V$-semimartingale with representation (2.1).

To find a good representation of $\alpha(\cdot)$, we recall the following result from [2, Example 5.3]:

$$\alpha(t) = \int_0^t k(t, s)dV(s) \qquad (0 \le t \le T)$$

with

$$k(t, s) = p(2q + p)\frac{(2q+p)e^{qs} - pe^{-qs}}{(2q+p)^2 e^{qt} - p^2 e^{-qt}} \qquad (0 \le s \le t).$$

From the theory of Volterra integral equations, there exists a function $l(t, s) \in L^2[0, T]^2$, called the resolvent of $k(t, s)$, such that, for almost every $0 \le s \le t \le T$,

$$(2.2) \qquad \begin{aligned} l(t, s) - k(t, s) + \int_s^t l(t, u)k(u, s)du &= 0, \\ l(t, s) - k(t, s) + \int_s^t k(t, u)l(u, s)du &= 0 \end{aligned}$$

(see [6, Chapter 4, Section 3] and [10, Chapter 9]). Using $l(t, s)$, we have the following representation of α in terms of the innovation process B:

$$(2.3) \qquad \alpha(t) = \int_0^t l(t, s)dB(s) \qquad (0 \le t \le T).$$

We shall solve (2.2) explicitly to obtain $l(t, s)$.

Theorem 2.1. *The expression* (1.4) *holds.*

Proof. We have $k(t,s) = a(t)b(s)$ for $0 \le s \le t$, where, for $t \in [0,T]$,

$$a(t) = \frac{p(2q+p)}{(2q+p)^2 e^{qt} - p^2 e^{-qt}}, \quad b(t) = (2q+p)e^{qt} - pe^{-qt}.$$

Fix $s \in [0,T]$ and define $x(t) = x_s(t)$ and $\lambda = \lambda_s$ by

$$x(t) = \int_s^t b(u)l(u,s)du \quad (s \le t \le T), \quad \lambda = b(s).$$

Then, from (2.2) we obtain

$$\frac{dx}{dt}(t) + a(t)b(t)x(t) = \lambda a(t)b(t), \quad x(s) = 0.$$

The solution x is given by

$$x(t) = \lambda - \lambda e^{-\int_s^t a(u)b(u)du},$$

whence $l(t,s)$ is obtained as

$$l(t,s) = a(t)b(s)e^{-\int_s^t a(u)b(u)du} = k(t,s)e^{-\int_s^t k(u,u)du} \quad (0 \le s \le t).$$

We have

$$k(u,u) = p - \frac{2p^2 q}{(2q+p)^2 e^{2qu} - p^2}.$$

By the change of variable $x(u) = (2q+p)^2 e^{2qu} - p^2$, we obtain

$$2p^2 q \int_s^t \frac{du}{(2q+p)^2 e^{2qu} - p^2} = 2p^2 q \int_{x(s)}^{x(t)} \frac{1}{2qx(x+p^2)}dx$$

$$= 2p^2 q \int_{x(s)}^{x(t)} \left\{ \frac{1}{x} - \frac{1}{x+p^2} \right\} dx = \log \frac{x(t)(x(s)+p^2)}{x(s)(x(t)+p^2)}$$

$$= \log \left\{ e^{-2q(t-s)} \frac{(2q+p)^2 e^{2qt} - p^2}{(2q+p)^2 e^{2qs} - p^2} \right\},$$

so that

$$e^{-\int_s^t k(u,u)du} = e^{-p(t-s)} e^{-2q(t-s)} \frac{(2q+p)^2 e^{2qt} - p^2}{(2q+p)^2 e^{2qs} - p^2}.$$

Thus

$$l(t,s) = (2q+p)pe^{-(p+q)(t-s)} \frac{(2q+p)e^{2qs} - p}{(2q+p)^2 e^{2qs} - p^2}$$

or (1.4), as desired. $\quad\square$

3. Filtering equations

3.1. **General result.** Let $(X(t), U(t))_{0 \le t \le T}$ be a two-dimensional centered continuous Gaussian process. The process $X(\cdot)$ represents the state process, while $U(\cdot)$ is another process which is related to the dynamics of $X(\cdot)$.

Let $(B(t))_{0 \le t \le T}$ be a one-dimensional Brownian motion that is independent of (X, U). We define the processes $V(\cdot)$ and $\alpha(\cdot)$ by (2.1) and (2.3), respectively. In this subsection, we assume that $l(t, s)$ in (2.3) is an arbitrary Volterra-type bounded measurable function (i.e., $l(t, s) = 0$ for $s > t$) rather than the special form (1.4). Of course, the function $l(t, s)$ in (1.4) satisfies this assumption. We consider the observation process $Y(\cdot)$ satisfying

$$Y(t) = \int_0^t \mu(s) X(s) ds + V(t) \qquad (0 \le t \le T),$$

where $\mu(\cdot)$ is a bounded measurable deterministic function on $[0, T]$ such that $\mu(t) \ne 0$ for $0 \le t \le T$.

As in Section 2, let $\mathcal{F}^Y$ be the augmented filtration generated by $Y(\cdot)$. For d-dimensional column vector processes $(A(t))_{0 \le t \le T}$ and $(C(t))_{0 \le t \le T}$, we write

$$\widehat{A}(t) = \mathbb{E}[A(t) | \mathcal{F}^Y(t)] \qquad (0 \le t \le T),$$
$$\Gamma_{AC}(t, s) = \mathbb{E}[A(t) C^*(s)] \qquad (0 \le s \le t \le T),$$

where $*$ denotes the transposition. Notice that $\Gamma_{AC}(t, s) \in \mathbf{R}^{d \times d}$.

We put

$$Z(t) = (X(t), U(t), \alpha(t))^* \qquad (0 \le t \le T),$$

and define the error matrix $P(t, s) \in \mathbf{R}^{3 \times 3}$ by

$$P(t, s) = \mathbb{E}[Z(t)(Z(s) - \widehat{Z}(s))^*] \qquad (0 \le s \le t \le T).$$

The next theorem gives an answer to the filtering problem for the partially observable process $(X(t), Y(t))_{0 \le t \le T}$. It turns out that this will be useful in the filtering problem for (1.5) for example.

Theorem 3.1. *The filter $\widehat{Z}(\cdot)$ satisfies the stochastic integral equation*

$$(3.1) \qquad \widehat{Z}(t) = \int_0^t \{P(t, s) + D(t, s)\} a(s) \{dY(s) - a^*(s)\widehat{Z}(s) ds\},$$

6

and the error matrix $P(t, s)$ is the unique solution to the following matrix Riccati-type integral equation such that $P(t, t)$ is symmetric and nonnegative definite for $0 \leq t \leq T$:

$$(3.2) \quad P(t, s) = -\int_0^s \{P(t, r) + D(t, r)\}a(r)a^*(r)\{P(s, r) + D(s, r)\}^* dr$$
$$+ \Gamma_{ZZ}(t, s) \qquad (0 \leq s \leq t \leq T),$$

where

$$D(t, s) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ l(t, s)/\mu(s) & 0 & 0 \end{pmatrix}, \quad a(s) = \begin{pmatrix} \mu(s) \\ 0 \\ -1 \end{pmatrix}.$$

Proof. Since (X, U) is independent of B, (X, U, α, Y) forms a Gaussian system. We have

$$Y(t) = \int_0^t \{\mu(s)X(s) - \alpha(s)\}ds + B(t).$$

Thus we can define the innovation process $I(\cdot)$ by

$$I(t) = Y(t) - \int_0^t (\mu(s)\widehat{X}(s)ds - \widehat{\alpha}(s)) \qquad (0 \leq t \leq T),$$

which is a Brownian motion satisfying $\mathcal{F}^Y = \mathcal{F}^I$ (cf. [21, Theorem 7.16]). Notice that $I(\cdot)$ can be written as

$$(3.3) \qquad I(t) = \int_0^t (Z(s) - \widehat{Z}(s))^* a(s)ds + B(t).$$

By [21, Corollary of Theorem 7.16], there exists an $\mathbf{R}^3$-valued Volterra-type function $F(t, s) = (F_1(t, s), F_2(t, s), F_3(t, s))^*$ on $[0, T]^2$ such that

$$\int_0^t |F(t, s)|^2 ds < +\infty \quad (0 \leq t \leq T),$$
$$(3.4) \qquad \widehat{Z}(t) = \int_0^t F(t, s)dI(s) \quad (0 \leq t \leq T),$$

where $|\cdot|$ denotes the Euclidean norm

Now let $g(t) = (g_1(t), g_2(t), g_3(t))$ be an arbitrary bounded measurable row-vector function on $[0, T]$. Then, for $t \in [0, T]$,

$$\mathbb{E}\left[\int_0^t g(s)dI(s) \cdot (Z(t) - \widehat{Z}(t))\right] = 0.$$

From this, (3.3), (3.4) and the fact that (X, U) and B are independent, we have

$$\int_0^t g(s)F(t, s)ds = \mathbb{E}\left[\int_0^t g(s)dI(s) \cdot Z(t)\right]$$

$$= \mathbb{E}\left[\int_0^t g(s)\{(Z(s) - \widehat{Z}(s))^*a(s)ds + dB(s)\} \cdot Z(t)\right]$$

$$= \int_0^t g(s)\mathbb{E}[Z(t)(Z(s) - \widehat{Z}(s))^*]a(s)ds + \int_0^t g_3(s)l(t, s)ds$$

$$= \int_0^t g(s)P(t, s)a(s)ds + \int_0^t g(s)D(t, s)a(s)ds.$$

Since $g(\cdot)$ is arbitrary, we deduce that $F(t, s) = (P(t, s) + D(t, s))a(s)$ or

(3.5) $\qquad \widehat{Z}(t) = \int_0^t (P(t, s) + D(t, s))a(s)dI(s) \quad (0 \le t \le T).$

The SDE (3.1) follows from (3.5) and the representation

$$I(t) = Y(t) - \int_0^t a^*(s)\widehat{Z}(s)ds \quad (0 \le t \le T).$$

The equation (3.2) follows from (3.5) and the equality

$$P(t, s) = \mathbb{E}[Z(t)Z^*(s)] - \mathbb{E}[\widehat{Z}(t)\widehat{Z}^*(s)].$$

The matrix $P(t, t)$ is clearly symmetric and nonnegative definite. Finally, the uniqueness of the solution to (3.2) follows from Proposition 3.2 below. $\square$

Proposition 3.2. *The solution $P(t, s)$ to the matrix integral equation (3.2) such that $P(t, t)$ is symmetric and nonnegative definite for $0 \le t \le T$ is unique.*

Proof. By coninuity, there exists a positive constant $C(T)$ such that

$$\|\Gamma_{ZZ}(t, s)\| \le C(T) \qquad (0 \le s \le t \le T),$$

where $\|A\| := (\text{trace}(A^*A))^{1/2}$ for $A \in \mathbf{R}^{3 \times 3}$. Let P be a solution to (3.2) such that $P(t, t)$ is symmetric and nonnegative definite for $t \in [0, T]$. We put $Q(t, s) = P(t, s) + D(t, s)$. Then (3.2) with $s = t$ is

$$\Gamma_{ZZ}(t, t) - P(t, t) = \int_0^t Q(t, r)a(r)a^*(r)Q^*(t, r)dr.$$

8

From this, we have

$$\int_0^t |Q(t,r)a(r)|^2 dr = \int_0^t \text{trace}\left\{Q(t,r)a(r)a^*(r)Q^*(t,r)\right\} dr$$
$$= \text{trace}(\Gamma_{ZZ}(t,t) - P(t,t)) \le \text{trace}(\Gamma_{ZZ}(t,t))$$
$$\le \sqrt{3}\|\Gamma_{ZZ}(t,t)\| \le \sqrt{3}C(T).$$

Therefore, $\|P(t,s)\|$ is at most

$$\|\Gamma_{ZZ}(t,s)\| + \int_0^s \|Q(t,r)a(r)a^*(r)Q^*(s,r)\| dr$$

$$\le \|\Gamma_{ZZ}(t,s)\| + \left(\int_0^t |Q(t,r)a(r)|^2 dr\right)^{1/2} \cdot \left(\int_0^s |Q(s,r)a(r)|^2 dr\right)^{1/2}$$

$$\le (1 + \sqrt{3})C(T).$$

Let P_1 and P_2 be two solutions of (3.2). We define $Q_i(t,s) = P_i(t,s) + D(t,s)$ for $i = 1, 2$. We put $P_i(t,s) = 0$ for $s > t$ and $i = 1, 2$. Since μ and l are bounded, it follows from the above estimate that there exists a positive constant $K(T)$ satisfying

$$\|a(s)a^*(s)Q_i(t,s)\| \le K(T) \qquad (0 \le s \le t \le T, \ i = 1, 2).$$

It follows that

$$\|Q_1(t,r)a(r)a^*(r)Q_1(s,r)^* - Q_2(t,r)a(r)a^*(r)Q_2(s,r)^*\|$$
$$\le \|Q_1(t,r)a(r)a^*(r)(Q_1(s,r)^* - Q_2(s,r)^*)\|$$
$$+ \|(Q_1(t,r) - Q_2(t,r))a(r)a^*(r)Q_2(s,r)^*\|$$
$$\le 2K(T)\|Q_1(s,r) - Q_2(t,r)\| = 2K(T)\|P_1(s,r) - P_2(t,r)\|.$$

From this and (3.2), we obtain

$$\sup_{0 \le t \le T} \|P_1(t,s) - P_2(t,s)\| \le 2K(T)\int_0^s \sup_{0 \le t \le T} \|P_1(t,r) - P_2(t,r)\| dr.$$

Therefore, Gronwall's lemma gives

$$\sup_{0 \le t \le T} \|P_1(t,s) - P_2(t,s)\| = 0 \qquad (0 \le s \le T).$$

Thus the uniqueness follows. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 1. We consider the case in which $\alpha = 0$ and the state process $X(\cdot)$ is the Ornstein-Uhlenbeck process satisfying

$$dX(t) = \theta X(t)dt + \sigma dW(t), \quad X(0) = 0,$$

where $\theta, \sigma \neq 0$ and $W(\cdot)$ is a one-dimensional Brownian motion that is independent of $B(\cdot)$. We also assume that $\mu(\cdot) = \mu$, a constant. Then $X(t) = \sigma \int_0^t e^{\theta(t-u)} dW(u)$ and

$$\Gamma_{XX}(t, s) = \frac{\sigma^2}{2\theta}(e^{\theta(t+s)} - e^{\theta(t-s)}) \quad (0 \leq s \leq t \leq T).$$

By Theorem 3.1, we have

$$(3.6) \qquad \widehat{X}(t) = \int_0^t \mu P_{XX}(t, s)(dY(s) - \mu \widehat{X}(s)ds),$$

$$(3.7) \qquad P_{XX}(t, s) = \Gamma_{XX}(t, s) - \int_0^s \mu^2 P_{XX}(t, r) P_{XX}(s, r) dr,$$

where $P_{XX}(t, s) = \mathbb{E}[X(t)(X(s) - \widehat{X}(s))]$ for $0 \leq s \leq t \leq T$. Let $\mathcal{F}(t)$, $0 \leq t \leq T$, be the $\mathbb{P}$-augmentation of the filtration generated by $(W(t), B(t))_{0 \leq t \leq T}$. Then $P_{XX}(t, s)$ is

$$\mathbb{E}[X(t)(X(s) - \widehat{X}(s))] = \mathbb{E}[\mathbb{E}[X(t)|\mathcal{F}(s)](X(s) - \widehat{X}(s))]$$
$$= \mathbb{E}[e^{\theta(t-s)}X(s)(X(s) - \widehat{X}(s))] = e^{\theta(t-s)}\gamma(s)$$

with $\gamma(s) = \mathbb{E}[X(s)(X(s) - \widehat{X}(s))]$. Thus (3.6) is reduced to

$$(3.8) \qquad d\widehat{X}(t) = (\theta - \mu^2\gamma(t))\widehat{X}(t)dt + \mu\gamma(t)dY(t).$$

Differentiating (3.7) in $s = t$, we get

$$(3.9) \qquad \frac{d\gamma(t)}{dt} = \sigma^2 + 2\theta\gamma(t) - \mu^2\gamma(t)^2.$$

The equations (3.8) and (3.9) are the well-known Kalman-Bucy filtering equations (see [13], [5], [6], [11] and [21]).

3.2. **Linear systems with memory.** Let $(X(t), Y(t))_{0 \leq t \leq T}$ be the two-dimensional partially observable process satisfying (1.5). For $j = 1, 2$, the noise process $V_j(\cdot)$ there is described by (1.1) with $W(\cdot) = W_j(\cdot)$ and $(p, q) = (p_j, q_j)$ satisfying (1.2). The Brownian motions $W_1(\cdot)$ and $W_2(\cdot)$, whence $V_1(\cdot)$ and $V_2(\cdot)$, are independent. In (1.5), the coefficients $\theta, \sigma, \mu \in \mathbf{R}$ with $\mu \neq 0$ are known constants and the initial value X_0 is a Gaussian random variable that is independent of (V_1, V_2). The processes $X(\cdot)$ and $Y(\cdot)$ represent the state and the observation, respectively.

By Theorem 2.1, we have $V_j(t) = B_j(t) - \int_0^t \alpha_j(s)ds$ with $\alpha_j(t) = \int_0^t l_j(t, s)dB_j(s)$ for $j = 1, 2$ and $0 \leq t \leq T$, where $B_1(\cdot)$ and $B_2(\cdot)$ are two independent Brownian motions, and, for $0 \leq s \leq t \leq T$,

$$l_j(t, s) = p_j e^{-(p_j+q_j)(t-s)} \left\{ 1 - \frac{2p_j q_j}{(2q_j + p_j)^2 e^{2q_j s} - p_j^2} \right\}.$$

10

We put $l_j(t) = l_j(t,t)$ for $j = 1, 2$, that is,

$$(3.10) \quad l_j(t) = p_j \left\{ 1 - \frac{2p_j q_j}{(2q_j + p_j)^2 e^{2q_j s} - p_j^2} \right\} \quad (j = 1, 2, \ 0 \le t \le T).$$

It holds that $l_j(t,s) = e^{-r_j(t-s)} l_j(s)$, where $r_j = p_j + q_j$.

We denote by $\mathcal{F}(t)$, $0 \le t \le T$, the $\mathbb{P}$-augmentation of the filtration $\sigma(X_0, (V_1(s), V_2(s))_{0 \le s \le t})$, $0 \le t \le T$.

Lemma 3.3. *For $0 \le s \le t \le T$, we have*

$$\mathbb{E}[X(t)|\mathcal{F}(s)] = e^{\theta(t-s)} X(s) - \sigma b(t-s)\alpha_1(s),$$

where

$$b(t) = \begin{cases} (e^{\theta t} - e^{-r_1 t})/(\theta + r_1) & (\theta + r_1 \ne 0), \\ t e^{\theta t} & (\theta + r_1 = 0). \end{cases}$$

Proof. For $t \in [0, T]$, $\int_0^t e^{\theta(t-s)} \alpha_1(s) ds$ is

$$\int_0^t l_1(u) \left\{ \int_u^t e^{\theta(t-s)} e^{-r_1(s-u)} ds \right\} dB_1(u) = \int_0^t b(t-u) l_1(u) dB_1(u).$$

Since $X(t) = e^{\theta t} X_0 + \sigma \int_0^t e^{\theta(t-u)} dV_1(u)$ or

$$X(t) = e^{\theta t} X_0 + \sigma \int_0^t e^{\theta(t-u)} dB_1(u) - \sigma \int_0^t e^{\theta(t-u)} \alpha_1(u) du,$$

$\mathbb{E}[X(t)|\mathcal{F}(s)]$ with $s \le t$ is equal to

$$e^{\theta t} X_0 + \sigma e^{\theta(t-s)} \int_0^s e^{\theta(s-u)} dB_1(u) - \sigma \int_0^s b(t-u) l_1(u) dB_1(u)$$

$$= e^{\theta(t-s)} X(s) - \sigma \int_0^s \left\{ b(t-u) - e^{\theta(t-s)} b(s-u) \right\} l_1(u) dB_1(u).$$

However, by elementary calculation, we have

$$b(t-u) - e^{\theta(t-s)} b(s-u) = b(t-s) e^{-r_1(s-u)} \quad (0 \le u \le s \le t).$$

Thus the lemma follows. $\qquad\square$

We put, for $0 \le t \le T$,

$$F = \begin{pmatrix} -\theta & \sigma & 0 \\ 0 & r_1 & 0 \\ 0 & 0 & r_2 \end{pmatrix}, \quad D(t) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \mu^{-1} l_2(t) & 0 & 0 \end{pmatrix}, \quad a = \begin{pmatrix} \mu \\ 0 \\ -1 \end{pmatrix},$$

$$G(t) = \begin{pmatrix} \sigma^2 & \sigma l_1(t) & 0 \\ \sigma l_1(t) & l_1^2(t) & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad H(t) = \begin{pmatrix} -\theta & \sigma & 0 \\ 0 & r_1 & 0 \\ \mu l_2(t) & 0 & r_2 - l_2(t) \end{pmatrix}.$$

11

We also put
$$Z(t) = (X(t), \alpha_1(t), \alpha_2(t))^* \qquad (0 \le t \le T).$$

Recall that $\Gamma_{ZZ}(0) = \mathbb{E}[Z(0)Z^*(0)]$ and that
$$\widehat{Z}(t) = \mathbb{E}[Z(t)|\mathcal{F}^Y(t)] \qquad (0 \le t \le T).$$

We define the error matrix $P(t) \in \mathbf{R}^{3 \times 3}$ by
$$P(t) = \mathbb{E}[Z(t)(Z(t) - \widehat{Z}(t))^*] \qquad (0 \le t \le T).$$

Here is the solution to the optimal filtering problem for (1.5).

Theorem 3.4. *The filter* $\widehat{Z}(\cdot)$ *satisfies the stochastic differential equation*

(3.11)
$$\begin{aligned} d\widehat{Z}(t) = &-\{F + (P(t) + D(t))aa^*\}\widehat{Z}(t)dt \\ &+ (P(t) + D(t))adY(t) \quad (0 \le t \le T), \end{aligned}$$

with $\widehat{Z}(0) = (\mathbb{E}[X_0], 0, 0)^*$, *and* $P(\cdot)$ *follows the matrix Riccati equation*

(3.12)
$$\begin{aligned} \frac{dP(t)}{dt} = &G(t) - H(t)P(t) - P(t)H(t)^* - P(t)aa^*P(t) \\ &(0 \le t \le T) \end{aligned}$$

with $P_{ij}(0) = \delta_{i1}\delta_{j1}\mathbb{E}[(X_0)^2]$ *for* $i, j = 1, 2, 3$.

Proof. For $0 \le s \le t \le T$, we put
$$P(t, s) = \mathbb{E}[Z(t)(Z(s) - \widehat{Z}(s))^*].$$

Then we have $P(t) = P(t, t)$. We also put, for $0 \le s \le t \le T$,
$$D(t, s) = e^{-r_2(t-s)}D(s),$$
$$Q(t, s) = P(t, s) + D(t, s), \quad Q(s) = Q(s, s) = P(s) + D(s).$$

By
$$P(t, s) = \mathbb{E}[\mathbb{E}[Z(t)|\mathcal{F}(s)](Z(s) - \widehat{Z}(s))^*]$$

and Lemma 3.3, we have $P(t, s) = M(t - s)P(s)$, where
$$M(t) = \begin{pmatrix} e^{\theta t} & -\sigma b(t) & 0 \\ 0 & e^{-r_1 t} & 0 \\ 0 & 0 & e^{-r_2 t} \end{pmatrix}$$

with $b(\cdot)$ as in Lemma 3.3. We also see that $D(t, s) = M(t-s)D(s)$. Combining, $Q(t, s) = M(t - s)Q(s)$. However, $M(t) = e^{-tF}$ since $dM(t)/dt = -FM(t)$ and $M(0)$ is the unit matrix. Thus we obtain

(3.13)
$$Q(t, s) = e^{-(t-s)F}Q(s).$$

12

From (3.13) and Theorem 3.1 with $U = \alpha_1$ and $\alpha = \alpha_2$, it follows that

$$(3.14) \qquad \widehat{Z}(t) = \int_0^t e^{-(t-s)F} Q(s) a\{dY(s) - a^*\widehat{Z}(s)ds\},$$

$$(3.15) \qquad P(t) = \Gamma_{ZZ}(t) - \int_0^t e^{-(t-u)F} Q(u) aa^* Q^*(u) e^{-(t-u)F^*} du.$$

The SDE (3.11) follows easily from (3.14).

Differentiating both sides of (3.15) yields

$$(3.16) \qquad \begin{aligned} \dot{P} &= \dot{\Gamma}_{ZZ} + F(\Gamma_{ZZ} - P) + (\Gamma_{ZZ} - P)F^* - Qaa^*Q^* \\ &= \dot{\Gamma}_{ZZ} + F\Gamma_{ZZ} + \Gamma_{ZZ}F^* - Daa^*D - HP - PH^* - Paa^*P. \end{aligned}$$

Since

$$dZ(t) = -FZ(t)dt + dR(t)$$

with

$$R(t) = \left(\sigma B_1(t), \int_0^t l_1(s)dB_1(s), \int_0^t l_2(s)dB_2(s) \right)^*,$$

we see by integration by parts that $Z(t)Z(t)^* - Z(0)Z(0)^*$ is equal to

$$\int_0^t Z(s)dZ(s)^* + \int_0^t dZ(s)Z(s)^* + \mathbb{E}[R(t)R^*(t)].$$

It follows that $\Gamma_{ZZ}(t) - \Gamma_{ZZ}(0)$ is equal to

$$\mathbb{E}\left[\int_0^t Z(s)dZ(s)^* \right] + \mathbb{E}\left[\int_0^t dZ(s)Z(s)^* \right] + \mathbb{E}[R(t)R^*(t)]$$

$$= -\int_0^t \Gamma_{ZZ}(s)F^*ds - \int_0^t F\Gamma_{ZZ}(s)ds + \mathbb{E}[R(t)R^*(t)].$$

Thus

$$\Gamma_{ZZ}(t) + \int_0^t F\Gamma_{ZZ}(s)ds + \int_0^t \Gamma_{ZZ}(s)F^*ds - \Gamma_{ZZ}(0)$$

$$= \mathbb{E}[R(t)R^*(t)] = \int_0^t \left(G(s) + D(s)aa^*D^*(s) \right) ds.$$

This and (3.16) yield (3.12). $\qquad \square$

Remark 2. We can easily extend Theorem 3.4 to a more general setting where the functions $l_j(t, s)$ take the form $l_j(t, s) = e^{c(t-s)}g(s)$.

Corollary 3.5. *We assume that $p_2 = 0$, i.e., $V_2(\cdot) = W_2(\cdot)$. Let $Z(t) = (X(t), \alpha_1(t))^*$ and $P(t) = \mathbb{E}[Z(t)(Z(t) - \widehat{Z}(t))^*] \in \mathbf{R}^{2\times 2}$. Then the filter*

13

$\widehat{Z}(\cdot)$ and the error matrix $P(\cdot)$ satisfy, respectively,

$$d\widehat{Z}(t) = -\{F + P(t)aa^*\}\widehat{Z}(t)dt + P(t)adY(t), \quad \widehat{Z}(0) = (\mathbb{E}[X_0], 0)^*,$$

$$\frac{dP(t)}{dt} = G(t) - FP(t) - P(t)F^* - P(t)aa^*P(t), \quad P_{ij}(0) = \delta_{i1}\delta_{j1}\mathbb{E}[(X_0)^2],$$

where

$$F = \begin{pmatrix} -\theta & \sigma \\ 0 & r_1 \end{pmatrix}, \quad a = \begin{pmatrix} \mu \\ 0 \end{pmatrix}, \quad G(t) = \begin{pmatrix} \sigma^2 & \sigma l_1(t) \\ \sigma l_1(t) & l_1(t)^2 \end{pmatrix}.$$

Corollary 3.6. *We assume that $p_1 = 0$, i.e., $V_1(\cdot) = W_1(\cdot)$. Let $Z(t) = (X(t), \alpha_2(t))^*$ and $P(t) = \mathbb{E}[Z(t)(Z(t) - \widehat{Z}(t))^*] \in \mathbf{R}^{2 \times 2}$. Then the filter $\widehat{Z}(\cdot)$ and the error matrix $P(\cdot)$ satisfy, respectively,*

$$d\widehat{Z}(t) = -\{F + (P(t) + D(t))aa^*\}\widehat{Z}(t)dt + (P(t) + D(t))adY(t),$$

$$\frac{dP(t)}{dt} = G - H(t)P(t) - P(t)H(t)^* - P(t)aa^*P(t), \quad ,$$

with $\widehat{Z}(0) = (\mathbb{E}[X_0], 0)^*$ and $P_{ij}(0) = \delta_{i1}\delta_{j1}\mathbb{E}[(X_0)^2]$, where

$$F = \begin{pmatrix} -\theta & 0 \\ 0 & r_2 \end{pmatrix}, \quad D(t) = \begin{pmatrix} 0 & 0 \\ \mu^{-1}l_2(t) & 0 \end{pmatrix}, \quad a = \begin{pmatrix} \mu \\ -1 \end{pmatrix},$$

$$G = \begin{pmatrix} \sigma^2 & 0 \\ 0 & 0 \end{pmatrix}, \quad H(t) = \begin{pmatrix} -\theta & 0 \\ \mu l_2(t) & r_2 - l_2(t) \end{pmatrix}.$$

Example 3.7. We consider the estimation problem of the value of a signal ρ from the observation process $(Y(t))_{0 \le t \le T}$ given by

$$dY(t) = \rho dt + dV(t), \quad Y(0) = 0,$$

where $V(\cdot)$ and $\alpha(\cdot)$ are as in Section 2. We assume that ρ is a random variable having the normal distribution with mean zero and variance v^2. This is the special case $\theta = \sigma = 0$ of the setting of Corollary 3.6. Let $r = p + q$ and $l(\cdot)$ be as above. Let $H(t)$ and a be as in Corollary 3.6 with $\mu = 1$ and $\theta = 0$. We define $P(t) = (P_{ij}(t))_{1 \le i,j \le 2}$ by $P(t) = \mathbb{E}[Z^*(t)(Z(t) - \widehat{Z}(t))]$ with $Z(t) = (\rho, \alpha(t))^*$. Then, by Corollary 3.6, the filter $(\widehat{\rho}(t), \widehat{\alpha}(t))$ satisfies

$$d\widehat{\rho}(t) = -\{P_{11}(t) - P_{12}(t)\}\{\widehat{\rho}(t) - \widehat{\alpha}(t)\}dt + \{P_{11}(t) - P_{12}(t)\}dY(t),$$

$$d\widehat{\alpha}(t) = [\{P_{21}(t) - P_{22}(t) + l(t) - r\}\widehat{\alpha}(t) - \{P_{21}(t) - P_{22}(t) + l(t)\}\widehat{\rho}(t)]dt$$
$$+ \{P_{21}(t) - P_{22}(t) + l(t)\}dY(t)$$

with $(\widehat{\rho}(0), \widehat{\alpha}(0)) = (\mathbb{E}[X_0], 0)$, and the error matrix $P(\cdot)$ follows

$$\frac{dP(t)}{dt} = -H(t)P(t) - P(t)H(t)^* - P(t)aa^*P(t), \quad P_{ij}(0) = \delta_{i1}\delta_{j1}\mathbb{E}[\rho^2].$$

By the linearization method as described in [5, Chapter 5], we can solve the equation for $P(\cdot)$ to obtain

$$P(t) = \frac{v^2}{1 + v^2\eta(t) + v^2\xi(t)\phi(t)} \begin{pmatrix} 1 & -\phi(t)/\psi(t) \\ -\phi(t)/\psi(t) & v^2\phi(t)^2/\psi(t)^2 \end{pmatrix},$$

where

$$\psi(t) = \exp\left\{\int_0^t (r - l(s))ds\right\}, \quad \phi(t) = \int_0^t l(s)\psi(s)ds,$$

$$\xi(t) = \int_0^t \frac{\psi(s) + \phi(s)}{\psi(s)^2}ds,$$

$$\eta(t) = \int_0^t \left\{1 - \psi(s)\xi(s) + \frac{\phi(s)}{\psi(s)}\right\}ds.$$

The analytical forms of ψ, ϕ, ξ and η can be derived. We omit the details.

4. APPLICATION TO FINANCE

In this section, we apply the results in the previous section to the problem of expected utility maximization for an investor with partial observations.

Let $(V_j(t))_{0 \leq t \leq T}$, $(\alpha_j(t))_{0 \leq t \leq T}$, $j = 1, 2$, be as in Section 3. In particular, $V_1(\cdot)$ and $V_2(\cdot)$ are independent. We consider the financial market consisting of a share of the money market with price $S_0(t)$ at time $t \in [0, T]$ and a stock with price $S(t)$ at time $t \in [0, T]$. The stock price $S(\cdot)$ is governed by the stochastic differential equation

$$(4.1) \qquad dS(t) = S(t)\{U(t)dt + \eta dV_2(t)\}, \quad S(0) = s_0,$$

where s_0 and η are positive constants and $U(\cdot)$ is a Gaussian process following

$$(4.2) \qquad dU(t) = \{\delta + \theta U(t)\}dt + \sigma dV_1(t), \quad U(0) = \rho.$$

The parameters θ, δ and σ are constants, and ρ is a Gaussian random variable that is independent of (V_1, V_2). For simplicity, we assume that

$$S_0(\cdot) = 1, \quad \eta = 1, \quad \delta = 0.$$

Let $\mathcal{F}(t)$, $0 \leq t \leq T$, be the $\mathbb{P}$-augmentation of the filtration generated by $(V_1(s), V_2(s))_{0 \leq s \leq t}$ and ρ. Then $U(\cdot)$ is $\mathcal{F}$-adapted but not $\mathcal{F}^S$-adapted; recall from Section 2 that $\mathcal{F}^S$ is the augmented filtration generated by $S(\cdot)$. Suppose that we can observe neither the disturbance process $V_2(\cdot)$ nor the drift process $U(\cdot)$ but only the price process $S(\cdot)$. Thus only $\mathcal{F}^S$-adapted processes are observable.

In many references such as [7], [8], and [9], the partially observable model described by (4.1) and (4.2) with V_j's replaced by Brownian motions, i.e., $V_j = B_j$, is studied.

15

We consider the following expected logarithmic utility maximization from terminal wealth: for given initial capital $x \in (0, \infty)$,

(4.3)   maximize $\mathbb{E}[\log(X^{x,\pi}(T))]$ over all $\pi \in \mathcal{A}(x)$,

where

$$\mathcal{A}(x) = \left\{ (\pi(t))_{0 \le t \le T} : \begin{array}{l} \pi(\cdot) \text{ is } \mathbf{R}\text{-valued, } \mathcal{F}^S\text{-progressively measurable,} \\ \int_0^T \pi^2(t)dt < \infty, \ X^{x,\pi}(t) \ge 0 \ (0 \le t \le T) \text{ a.s.} \end{array} \right\},$$

and

(4.4) $$X^{x,\pi}(t) = x + \int_0^t \frac{\pi(u)}{S(u)} dS(u).$$

The value $\pi(t)$ is the dollar amount invested in the stock at time t, whence $\pi(t)/S(t)$ is the number of units of stock held at time t. The process $X^{x,\pi}(\cdot)$ is the wealth process associated with the self-financing portfolio determined uniquely by $\pi(\cdot)$.

An analogue of the problem (4.3) for full observations is solved in [2]. For related work, see [15], [19], [20] and the references therein. We solve the problem (4.3) by combining the results above on filtering and the martingale method as described in [14].

Consider the process $(Y(t))_{0 \le t \le T}$ defined by

$$Y(t) = \int_0^t U(s)ds + V_2(t) = B_2(t) + \int_0^t (U(s) - \alpha_2(s))ds,$$

which we regard as the observation process. Let $(I(t))_{0 \le t \le T}$ be the innovation process associated with $Y(\cdot)$ that is given by

$$I(t) = Y(t) - \int_0^t (\widehat{U}(s) - \widehat{\alpha}_2(s))ds,$$

where $\widehat{A}(t) = \mathbb{E}[A(t)|\mathcal{F}^Y(t)]$ as in the previous sections. The innovation process $I(\cdot)$ is a $\mathcal{F}^Y$-Brownian motion satisfying $\mathcal{F}^S = \mathcal{F}^Y = \mathcal{F}^I$.

Let $(L(t))_{0 \le t \le T}$ be the exponential $\mathcal{F}$-martingale given by

$$L(t) = \exp\left\{ -\int_0^t (U(s) - \alpha_2(s))dB_2(s) - \frac{1}{2}\int_0^t (U(s) - \alpha_2(s))^2 ds \right\}$$

$$= \exp\left\{ -\int_0^t (U(s) - \alpha_2(s))dY(s) + \frac{1}{2}\int_0^t (U(s) - \alpha_2(s))^2 ds \right\}.$$

By [21, Chapter 7], we find that, for $t \in [0, T]$,

$$\widehat{L}(t) = \exp\left\{ -\int_0^t (\widehat{U}(s) - \widehat{\alpha}_2(s))dY(s) + \frac{1}{2}\int_0^t (\widehat{U}(s) - \widehat{\alpha}_2(s))^2 ds \right\}$$

$$= \exp\left\{ -\int_0^t (\widehat{U}(s) - \widehat{\alpha}_2(s))dI(s) - \frac{1}{2}\int_0^t (\widehat{U}(s) - \widehat{\alpha}_2(s))^2 ds \right\}.$$

The process $(\widehat{L}(t))_{0 \le t \le T}$ is a $\mathcal{F}^Y$-martingale.

For $x \in (0, \infty)$ and $\pi \in \mathcal{A}(x)$, we see from Itô formula that the process $(\widehat{L}(t)X^{x,\pi}(t))$ is a local $\mathcal{F}^Y$-martingale, whence a $\mathcal{F}^Y$-supermartingale since $X^{x,\pi}(\cdot)$ is nonnegative. It follows that, for $x \in (0, \infty)$, $\pi \in \mathcal{A}(x)$, and $y \in (0, \infty)$,

$$(4.5) \qquad \mathbb{E}[\log(X^{x,\pi}(T))] \le \mathbb{E}[\log(X^{x,\pi}(T)) - y\widehat{L}(T)X^{x,\pi}(T)] + yx$$
$$\le \mathbb{E}[\log\{1/(y\widehat{L}(T))\} - 1] + yx,$$

where, in the second inequality, we have used

$$\log(z) - yz \le \log(1/y) - 1 \quad (y, z \in (0, \infty)).$$

The equalities in (4.5) hold if and only if

$$(4.6) \qquad\qquad\qquad X^{x,\pi}(T) = x/\widehat{L}(T) \quad \text{a.s.}$$

Thus the portfolio process $\pi(\cdot)$ satisfying (4.6) is optimal.

Put

$$\pi_0(t) = x(\widehat{U}(t) - \widehat{\alpha}_2(t))/\widehat{L}(t) \qquad (0 \le t \le T).$$

Since $(x/\widehat{L})(0) = x$ and

$$d(x/\widehat{L})(t) = \frac{x(\widehat{U}(t) - \widehat{\alpha}_2(t))}{\widehat{L}(t)}dY(t) = \frac{\pi_0(t)}{S(t)}dS(t),$$

we see from (4.4) that the process $\pi_0(\cdot)$ satisfies (4.6), whence it is the desired optimal optimal portfolio process. It also satisfies

$$\frac{\pi_0(t)}{X^{x,\pi_0}(t)} = \widehat{U}(t) - \widehat{\alpha}_2(t) \qquad (0 \le t \le T).$$

We put

$$a = (1, 0, -1)^*, \quad Z(t) = (U(t), \alpha_1(t), \alpha_2(t))^* \quad (t \in [0, T]).$$

We define the error matrix $P(t) \in \mathbf{R}^{3 \times 3}$ by $\mathbb{E}[Z(t)(Z(t) - \widehat{Z}(t))^*]$. Combining the results above with Theorem 3.4 which describes the dynamics of $\widehat{U}$ and $\widehat{\alpha}_2$, we obtain the next theorem.

Theorem 4.1. *The optimal portfolio process π_0 for the problem (4.3) is given by*

$$\pi_0(t) = xa^*\widehat{Z}(t)/\widehat{L}(t) \qquad (0 \le t \le T)$$

and satisfies

$$X^{x,\pi_0}(T) = x/\widehat{L}(T), \quad \frac{\pi_0(t)}{X^{x,\pi_0}(t)} = a^*\widehat{Z}(t) \qquad (0 \le t \le T).$$

The filter $\widehat{Z}(\cdot)$ follows

$$d\widehat{Z}(t) = -\{F + (P(t) + D(t))aa^*\}\widehat{Z}(t)dt + (P(t) + D(t))adY(t)$$

with $\widehat{Z}(0) = (\mathbb{E}[\rho], 0, 0)^*$, and the error matrix $P(\cdot)$ satisfies the matrix Riccati equation

$$\frac{dP(t)}{dt} = G(t) - H(t)P(t) - P(t)H(t)^* - P(t)aa^*P(t)$$

with $P_{ij}(0) = \delta_{i1}\delta_{j1}\mathbb{E}[\rho^2]$ $(i, j = 1, 2, 3)$, where F, $D(t)$, $G(t)$ and $H(t)$ are as in Theorem 3.4 with $\mu = 1$.

5. PARAMETER ESTIMATION

Let $V(\cdot)$ be the process given by (1.1). We can estimate the values of the parameters p and q there from given data of $V(\cdot)$ by a least-squares approach ([4]). In fact, since $V(\cdot)$ is a stationary increment process, the variance of $V(t) - V(s)$ is a function in $t - s$. To be precise,

$$\frac{1}{t-s}\text{Var}(V(t) - V(s)) = U(t - s) \qquad (0 \le s < t),$$

where the function $U(t) = U(t; p, q)$ is given by

$$(5.1) \qquad U(t) = \frac{q^2}{(p+q)^2} + \frac{p(2q+p)}{(p+q)^3} \cdot \frac{(1 - e^{-(p+q)t})}{t} \qquad (t > 0).$$

Suppose that for $t_j = j\Delta t$, $j = 1, \ldots, N$, the value of $V(t_j)$ is v_j. For simplicity, we assume that $\Delta t = 1$. The unbiased estimation u_j that corresponds to $U(t_j)$ is given by

$$u_j = \frac{1}{j(N-j-1)}\sum_{i=1}^{N-j}(v_{i+j} - v_i - m_j)^2,$$

where m_j is the mean of $v_{i+j} - v_i$'s:

$$m_j = \frac{1}{N-j}\sum_{i=1}^{N-j}(v_{i+j} - v_i).$$

Fitting $\{U(t_j; p, q)\}$ to $\{u_j\}$ by least squares, we obtain the desired estimated values of p and q.

For example, we produce sample values $v_1, v_2, \ldots, v_{1000}$ for $V(\cdot)$ with $(p, q) = (0.5, 0.3)$ by a Monte Carlo simulation. We use this data to numerically calculate the values p_0 and q_0 of p and q, respectively, that minimize

$$\sum_{j=1}^{30}(U(t_j; p, q) - u_j)^2.$$

18

It turns out that $p_0 = 0.5049$ and $q_0 = 0.2915$. In Figure 5.1, we plot $\{U(t_j; p, q)\}$ (theory), $\{u_j\}$ (sample) and $\{U(t_j; p_0, q_0)\}$ (fitted). It is seen that the fitted curve follows the theoretical curve reasonably well.
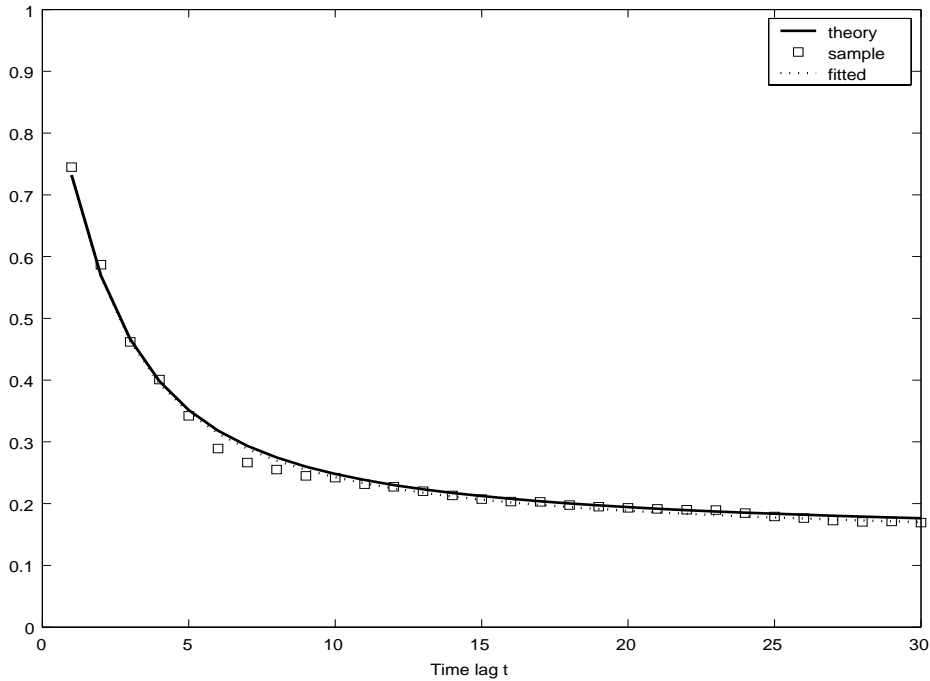

FIGURE 5.1. Plotting of the function $v(t)$, the sample data and the fitted function $v_0(t)$.

We extend the approach above to that for the estimation of the parameters p, q, θ and σ in

$$dX(t) = -\theta X(t)dt + \sigma dV(t), \quad X(0) = X_0,$$

where $\theta, \sigma \in (0, \infty)$, the process $V(\cdot)$ is given by (1.1) as above, and the initial value X_0 is independent of $(V(t))_{0 \le t \le T}$ and satisfies $\mathbb{E}[X_0^2] < \infty$. The solution $X(\cdot)$ is the following Ornstein-Uhlenbeck-type process with memory:

$$(5.2) \qquad X(t) = e^{-\theta t} X_0 + \int_0^t e^{-\theta(t-u)} dV(u) \qquad (t \in [0, T]).$$

Put

$$\varphi(t) := \int_0^t e^{(\theta - p - q)u} du \qquad (0 \le t \le T).$$

19

Proposition 5.1. *We have*

(5.3) $\quad \dfrac{1}{t-s}\mathbb{E}\left[(X(t)-e^{-\theta(t-s)}X(s))^2\right]=H(t-s)\quad(0\le s<t\le T),$

where, for $0<t\le T$, the function $H(t)=H(t;p,q,\theta,\sigma)$ is given by

$$H(t)=\sigma^2\left\{1-\frac{p(2q+p)}{(p+q)(\theta+p+q)}\right\}\frac{1-e^{-2\theta t}}{2\theta t}+\frac{\sigma^2 p(2q+p)e^{-2\theta t}\varphi(t)}{(p+q)(\theta+p+q)t}.$$

Proof. By (5.2), the left-hand side of (5.3) is equal to

$$\frac{\sigma^2 e^{-2\theta t}}{t-s}\mathbb{E}\left[\left(\int_s^t e^{\theta u}dV(u)\right)^2\right].$$

We put $c_u=pe^{-(p+q)u}I_{(0,\infty)}(u)$ for $u\in\mathbf{R}$. By Proposition 3.2 in [2], $\int_s^t e^{\theta u}dV(u)$ is given by

$$\int_s^t e^{\theta u}dW(u)-\int_{-\infty}^{\infty}\left(\int_s^t e^{\theta r}c_{r-u}dr\right)dW(u)$$

$$=\int_s^t\left(e^{\theta u}-\int_u^t e^{\theta r}c_{r-u}dr\right)dW(u)-\int_{-\infty}^s\left(\int_s^t e^{\theta r}c_{r-u}dr\right)dW(u).$$

Thus $\mathbb{E}[(\int_s^t e^{\theta u}dV(u))^2]$ is equal to

$$\int_s^t\left(e^{2\theta u}-\int_u^t e^{\theta r}c(r-u)dr\right)^2 du+\int_{-\infty}^s\left(\int_s^t e^{\theta r}c(r-u)dr\right)^2 du.$$

By integration by parts and the equalities

$$\int_u^t e^{\theta r}c_{r-u}dr=pe^{(p+q)u}\{\varphi(t)-\varphi(u)\},$$

$$e^{-(\theta-p-q)s}\{\varphi(t)-\varphi(s)\}=\varphi(t-s),$$

we obtain the desired result. $\qquad\square$

Suppose that for $t_j=j\Delta t$, $j=1,\ldots,N$, the value of $X(t_j)$ is x_j. We assume $\Delta t=1$ for simplicity. The estimation $h_j(\theta)$ that corresponds to $H(t_j;p,q,\theta,\sigma)$ is given by

$$h_j(\theta)=\frac{1}{j(N-j-1)}\sum_{i=1}^{N-j}(x_{i+j}-e^{-\theta j}x_i-m_j(\theta))^2,$$

where $m_j(\theta)$ is the mean of $x_{i+j}-e^{-\theta j}x_i$, $i=1,\ldots,N-j$:

$$m_j(\theta)=\frac{1}{N-j}\sum_{i=1}^{N-j}(x_{i+j}-e^{-\theta j}x_i).$$

Fitting $\{H(t_j; p, q, \theta, \sigma) - h_j(\theta)\}$ to $\{0\}$ by least squares, we obtain the desired estimated values of p, q, θ and σ.

For example, we produce sample values $x_1, x_2, \ldots, x_{1000}$ for $X(\cdot)$ with

$$(p, q, \theta, \sigma) = (0.2, 1.5, 0.8, 1.0)$$

by a Monte Carlo simulation. We use this data to numerically calculate the values p_0, q_0, θ_0 and σ_0 of p, q, θ and σ, repectively, that minimize

$$\sum_{j=1}^{30} (H(t_j; p, q, \theta, \sigma) - h_j(\theta))^2.$$

It turns out that

$$(p_0, q_0, \theta_0, \sigma_0) = (0.1910, 1.5382, 0.8354, 1.0184).$$

In Figure 5.2, we plot $\{H(t_j; p, q, \theta, \sigma)\}$ (theory), $\{h_j(\theta_0)\}$ (sample with estimated θ) and $\{H(t_j; p_0, q_0, \theta_0, \sigma_0)\}$ (fitted). It is seen that the fitted curve follows closely the theoretical curve.
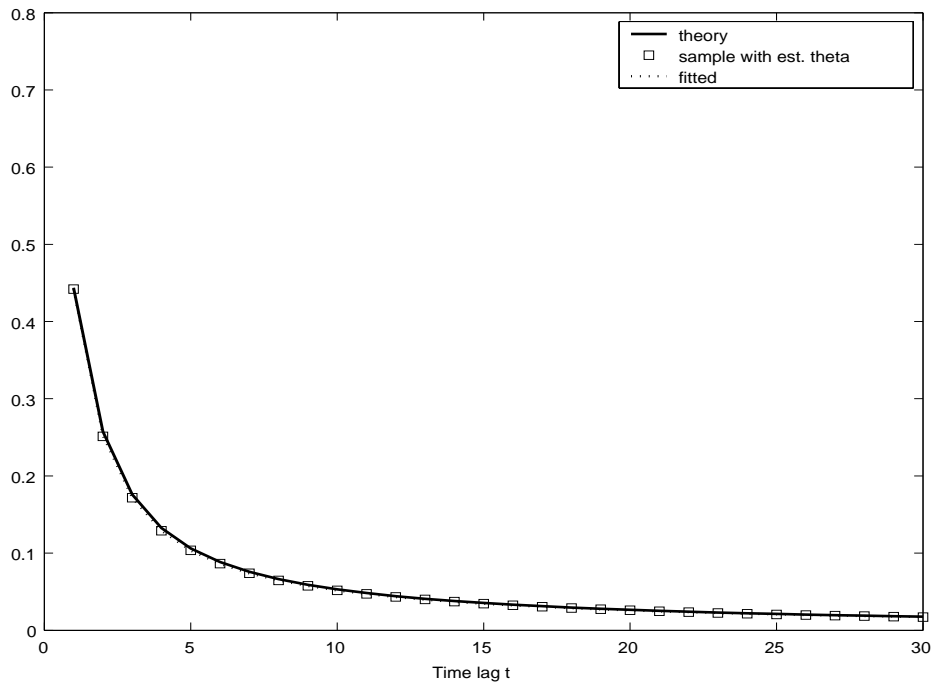

FIGURE 5.2. Plotting of the function $h(t)$, the sample data $\bar{h}(t)$ with estimated θ and the fitted function $h_0(t)$.

6. Simulation

As we have seen, the process $V(\cdot)$ described by (1.1) has both stationary increments and a simple semimartingale representation as Brownian motion does, and it reduces to Brownian motion when $p = 0$. In this sense, we may see $V(\cdot)$ as a generalized Brownian motion. Since $V(\cdot)$ is non-Markovian unless $p = 0$, we have now a wide choice for designing models driven by either white or colored noise.

In this section, we compare the performance of the optimal filter with the Kalman-Bucy filter in the presence of colored noise. We consider the partially observable process $(X(t), Y(t))_{0 \le t \le T}$ governed by (1.5) with $X_0 = 0$. Suppose that an engineer uses the conventional Markovian model

$$dX'(t) = \theta X'(t)dt + \sigma dB_1(t), \quad X'(0) = 0$$

to describe the non-Markovian system process $X(\cdot)$. Then he will be led to use the Kalman-Bucy filter $\widetilde{X}(\cdot)$ governed by

$$d\widetilde{X}(t) = (\theta - \mu^2 \gamma(t))\widetilde{X}(t)dt + \mu\gamma(t)dY(t), \quad \widetilde{X}(0) = 0,$$

where $\gamma(\cdot)$ is the solution to

$$\frac{d\gamma(t)}{dt} = \sigma^2 + 2\theta\gamma(t) - \mu^2\gamma(t)^2, \quad \gamma(0) = 0$$

(see (3.8) and (3.9)), instead of the right optimal filter $\widehat{X}$ as described by Theorem 3.4.

We adopt the following parameters:

$$T = 10, \quad \Delta t = 0.01, \quad N = T/\Delta t = 1000,$$
$$t_i = i\Delta t \quad (i = 1, \dots, N),$$
$$\sigma = 1, \quad \theta = -2, \quad \mu = 5.$$

Let $n \in \{1, 2, \dots, 100\}$. For the n-th run of Monte Carlo simulation, we sample $x_n(t_1), \dots, x_n(t_N)$ for $X(\cdot)$. Let $\widetilde{x}_n(\cdot)$ and $\widehat{x}_n(\cdot)$, $n = 1, \dots, 100$, be the Kalman-Bucy filter and the optimal filter respectively. For $u_n = \widehat{x}_n$ or $\widetilde{x}_n$, we consider the *average error norm*

$$\text{AEN} := \sqrt{\frac{1}{100N} \sum_{i=1}^{N} \sum_{n=1}^{100} (x_n(t_i) - u_n(t_i))^2},$$

and the *average error*

$$\text{AE}(t_i) := \sqrt{\frac{1}{100} \sum_{n=1}^{100} (x_n(t_i) - u_n(t_i))^2} \qquad (i = 1, \dots, N).$$

In Table 6.1, we show the resulting AEN's of $\{\widehat{x}_n\}$ and $\{\widetilde{x}_n\}$ for the following five sets of $\Theta = (p_1, q_1, p_2, q_2)$:

$$\Theta_1 = (0.2, 0.3, 0.5, 0.2),$$
$$\Theta_2 = (5.2, 0.3, -0.5, 0.6),$$
$$\Theta_3 = (0.0, 1.0, 5.8, 0.7),$$
$$\Theta_4 = (5.4, 0.8, 0.0, 1.0),$$
$$\Theta_5 = (5.1, 2.3, 4.9, 1.3).$$

We see that there are clear differences between the two filters in the cases Θ_2 and Θ_4. We notice that, in these two cases, p_1 is large than the parameters p_2 and q_2. In Figure 6.1, we compare the graphs of $\mathrm{AE}(\cdot)$ for the two filters in the case $\Theta = \Theta_2$. It is seen that the error of the optimal filter is consistently smaller than that of the Kalman-Bucy filter.

Θ	Optimal filter	Kalman-Bucy filter
Θ_1	0.5663	0.5667
Θ_2	0.4620	0.5756
Θ_3	0.5136	0.5167
Θ_4	0.4487	0.5196
Θ_5	0.4294	0.4524

TABLE 6.1. Comparison of AEN's

REFERENCES

[1] V. Anh and A. Inoue, Financial markets with memory I: Dynamic models, *Stochastic Anal. Appl.*, to appear.
[2] V. Anh, A. Inoue, and Y. Kasahara, Financial markets with memory II: Innovation processes and expected utility maximization, *Stochastic Anal. Appl.*, to appear.
[3] V. Anh, A. Inoue, and Y. Nakano, Binary market models with memory, submitted.
[4] V. Anh, A. Inoue, and C. Pesee, Incorporation of memory into the Black-Scholes-Merton theory and estimation of volatility, submitted.
[5] R. S. Bucy and P. D. Joseph, *Filtering for stochastic processes with application to guidance*, Interscience, New York, 1968.
[6] M. H. A. Davis, *Linear estimation and stochastic control*, Chapman & Hall, London, 1977.
[7] J. B. Detemple, Asset pricing in a production economy with incomplete information, *J. Finance* **41** (1986), 383–391.
[8] M. U. Dothan and D. Feldman, Equilibrium interest rates and multiperiod bonds in a partially observable economy, *J. Finance* **41** (1986), 369–382.
[9] G. Genotte, Optimal portfolio choice under incomplete information, *J. Finance* **41** (1986), 733–746.
[10] G. Gripenberg, S.-O. Londen, and O. Staffans, *Volterra integral and functional equations*, Cambridge University Press, Cambridge, 1990.
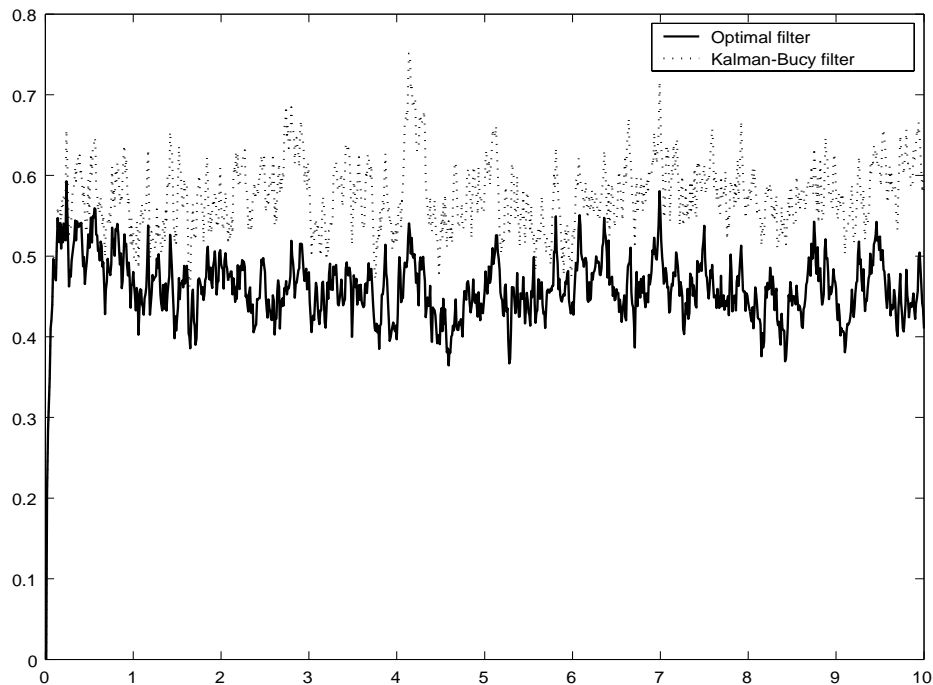
FIGURE 6.1. Plotting of AE($\cdot$) for the optimal and Kalman-Bucy filters with noise parameter $\Theta = \Theta_2$

[11] A. H. Jazwinski, *Stochastic processes and filtering theory*, Academic Press, New York, 1970.

[12] R. E. Kalman, A new approach to linear filtering and prediction problems, *J. Basic. Eng. ASME* **82** (1960), 33–45.

[13] R. E. Kalman and R. S. Bucy, New results in linear filtering and prediction theory, *J. Basic. Eng. ASME* **83** (1961), 95–108.

[14] I. Karatzas and S. E. Shreve, *Methods of mathematical finance*, Springer-Verlag, New York, 1998.

[15] I. Karatzas and X. Zhao, Bayesian adaptive portfolio optimization, Handbook of mathematical finance, 632–670, Cambridge University Press, Cambridge, 2001.

[16] M. L. Kleptsyna, P. E. Kloeden, and V. V. Anh, Linear filtering with fractional Brownian motion, *Stochastic Anal. Appl.* **16** (1998), 907–914.

[17] M. L. Kleptsyna, A. Le Breton, and M.-C. Roubaud, General approach to filtering with fractional Brownian noises – Application to linear systems, *Stoch. Stoch. Rep.* **71** (2000), 119–140.

[18] M. L. Kleptsyna and A. Le Breton, Extension of the Kalman-Bucy filter to elementary linear systems with fractional Brownian noises, *Stat. Inference Stoch. Process* **5** (2002), 249–271.

[19] P. Lakner, Utility maximization with partial information, *Stochastic Process. Appl.* **56** (1995), 247–273.

[20] P. Lakner, Optimal trading strategy for an investor: the case of partial information, *Stochastic Process. Appl.* **76** (1998), 77–97.

[21] R. S. Liptser and A. N. Shiryayev, *Statistics of random processes. I. General theory*, 2nd edition, Springer-Verlag, New York, 2001.

E-mail address: `inoue@math.sci.hokudai.ac.jp`

E-mail address: `nakano_y@math.sci.hokudai.ac.jp`

E-mail address: `v.anh@fsc.qut.edu.au`

DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE, HOKKAIDO UNIVERSITY, SAPPORO 060-0810, JAPAN

DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE, HOKKAIDO UNIVERSITY, SAPPORO 060-0810, JAPAN

SCHOOL OF MATHEMATICAL SCIENCES, QUEENSLAND UNIVERSITY OF TECHNOLOGY, GPO BOX 2434, BRISBANE, QUEENSLAND 4001, AUSTRALIA