ANNALES SCIENTIFIQUES
DE L'UNIVERSITÉ DE CLERMONT-FERRAND 2
Série Probabilités et applications

TRAN HUNG THAO

Optimal state estimation of a Markov process from point process observations

Annales scientifiques de l'Université de Clermont-Ferrand 2, tome 96, série *Probabilités et applications*, nº 9 (1991), p. 1-10
<http://www.numdam.org/item?id=ASCFPA_1991__96_9_1_0>

OPTIMAL STATE ESTIMATION OF A MARKOV PROCESS FROM POINT PROCESS OBSERVATIONS

TRAN HUNG THAO
INSTITUTE OF MATHEMATICS
HANOI, VIETNAM

Abstract. The aim of this paper is to study the problem of filtering of a Feller- Markov process from point process observations. A modification of a result of Kunita will be made and a stochastic diferential equation for quasi-filtering will be investigated.

0. INTRODUCTION

In [3], H. Kunita has studied the nonlinear filtering of Markov processes from observations driven by a standard Wiener process. In this paper we consider the filtering from Poisson process observations. In Section 1 we recall some recent results of filtering of a semimartingale from point process observations. In Sections 2, 3 the method of Kunita is used to the case where the observation process is a Poisson process; the unnormalized distribution of state estimation (quasi-filtering) is considered as a stochastic process taking values in the set of all probability measures over the state space.

1. FILTERING OF A SEMIMARTINGALE FROM POINT PROCESS OBSERVATIONS

Let (Ω, F, p) be a complete probability space on which all processes are defined and adapted to a filtration (F_t). "Usual conditions" are supposed to be satisfied by (F_t).

The system process is a real semimartingale

$$X_t = X_0 + \int_0^t H_s ds + Z_t \tag{1.1}$$

where Z_t is a F_t-martingale, H_t is a bounded F_t-progressive process and $E[sup_{s\leq t}|X_s|] < \infty$.

The observation is given by an N-vector of point process F_t-semimartingales of the form

$$Y_t = \int_0^t h_s ds + M_t \tag{1.2}$$

where M_t is an N-dimensional F_t-martingale with mean 0 $h_t = h(X_t)$ is an N-vector of positive bounded F_t-progressive processes $h_t = (h_t^1, ..., h_t^N)$.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Denote by F_t^Y the natural filtration of Y which provide observation datas concerning X_t.

Suppose that the processes $u_s^i = \frac{d}{ds} < Z, M^i >_s, (s \leq t), 1 \leq i \leq n)$, are F_s-predictable, where $<,>$ stands for the quadratic variation of two processes. Denote also by $\widehat{u}_s$ the F_t^Y-predictable projection of u_s.

The filtering process $\pi(X_t)$ is defined by

$$\pi(X_t) = E(X_t \mid F_t^Y) , \tag{1.3}$$

where F_t^Y is the σ-algebra generated by all $Y_s, s \leq t : F_t^Y = \sigma(Y_s, s \leq t)$.

Let $\pi(h_t)$ be the N-dimensional filtering process corresponding to the process in (1.2). The following facts are well known :

1. The process

$$m_t = Y_t - \int_0^t \pi(h_s) ds \tag{1.4}$$

is an N-dimensional point process F_t^Y-martingale and it can be expressed by

$$m_t = M_t - \int_0^t [h_s - \pi(h_s)] ds . \tag{1.5}$$

Furthermore, for any t, the future σ-field $\sigma(m_u - m_t : u \geq t$ is independent of F_t^Y.

2. The σ-field generated by m_t is included in $F_t^Y : \sigma(m_s ; s \leq t) \subset F_t^Y$ and m_t is called the innovation of the point process Y_t.

3. If m_t is the innovation of Y_t and R_t is an N-dimensional F_t^Y-martingale then

$$R_t = R_0 + \int_0^t (K_s, dm_s) \tag{1.6}$$

where K_t is an N-dimensional F_t^Y-predictable process such that

$$\sum_{i=1}^N \int_0^t K_s^i \pi(h_s^i) ds$$

and the notation (,) is understood as an inner product in R^N :

$$\int_0^t (K_s, dm_s) = \sum_{n=1}^N \int_0^t K_s^i dm_s^i . \tag{1.7}$$

(See, for instance, [1]).

4. THEOREM 1.1 (cf. [4]). The filtering process $\pi(X_t)$ is given by

(1.8) $$\pi(X_t) = \pi(X_0) + \int_0^t (K_s, dm_s) \ ,$$

where

$$K_s = (K_s^1, ..., K_s^N)$$

and

$$K_s^i = [\pi(h_s^i)]^{-i}[\pi(X_{s-}h_s^i) - \pi(X_{s-})\pi(h_s^i) + u_s^i]dm_s \ , 1 \leq i \leq N \ .$$

The equation is up to an indistinguishability.

5. Filtering by the method of Probability of Reference

The filtering equation (1.8) obtained by the innovation method, have some disvantages in applications because of the appearance of the factors $[\pi(h_s^i)]_{-1}$ and the quadratic terms $\pi(X_{s-})\pi(h_s^i)$. These disvantages will be disappeared when we consider the Zakai equation for unnormalized distributions derived from the method of probability of reference.

Suppose that under the reference probability Q the observations are given by an N vector of standard Poisson processes. Then

$$\mu_t^i \stackrel{def}{=} Y_y^i - t \ , \ i = 1, ..., N$$

are $Q-$, F_t)-martingales, where $F_t = F_\infty^X \vee F_t^Y$, F_∞^Y is the σ-algebra $\sigma(X_s \ , \ sgeq0)$.

Suppose now that the probability P is obtained from Q by an absolutely continuous change of measure $Q \longrightarrow P$ such that

(1.9) $$m_t^i = Y_t^i - \int_0^t \pi(h_s^i)ds \ , \ i = 1, ..., N$$

are P, F_t-martingales. Let us denote

(1.10) $$E[\frac{dP}{dQ} \mid F_t] = L_t \ ,$$

A Bayes formula give us (see, for instance, [1]) :

(1.11) $$E_P[X_t \mid F_t^Y] = \frac{E_Q[X_t L_t \mid F_t^Y]}{E_Q[L_t]}$$

Denote by $\sigma(X_t)$ the filtering under Q of $X_t L_t : \sigma(X_t) = E_Q[X_t L_t \mid F_t^Y]$ we have

(1.12) $$\pi(X_t) = \frac{\sigma(X_t)}{\sigma(1)_t}$$

DEFINITION : $\sigma(X_t)$ is called the quasi-filtering process of X_t from point process observations Y_t.

It is known that an equation for quasi-filtering (Zakai equation) is given by :

THEOREM 1.2 (R. Elliott, cf. [2])

The system process and the observation process are supposed as in the begining of §1. Then we have :

$$\sigma(X_t) = \sigma(X_0) + \int_0^t \sigma_s(H_s)ds + \int_0^t (K_s, d\mu_s) \ , \tag{1.13}$$

where

$$K_t = (K_t^1, ..., K_t^N) \ ,$$
$$K_t^i = \sigma(X_{t-}h_t) - \sigma(X_{t-}) \ ,$$
$$\mu_t = (\mu_t^1, ..., \mu_t^N) \ ,$$
$$\mu_t^i = Y_t^i - t \ .$$

This equation is linear in σ.

2. FILTERING OF A MARKOV PROCESS

In this section, a modification of a theorem of Kunita [3] for the case of point process observations will be made.

The system process X_t is supposed now to be a homogeneous Feller Markov process taking values in a compact separable Hausdorff space S. The semigroup P_t , $t \geq 0$ associated with the transition probabilities $P_t(x, E)$ is a Feller semigroup, that is

$$P_t f(x) = \int P_t(x, dy) f(y) \tag{2.1}$$

maps C(S) into itself for all $t \geq 0$ and satisfies

$$lim_{t \to 0} P_t f(x) = f(x)$$

uniformly in S for all $f \in C(S)$ where C(S) is the space of all real continuous functions over S.

Assume that the observation Y_t is an N-vector of Poisson processes of intensities $h_t^i = h^i(X_t) \in C(S)$, $(i = 1, ..., N)$.

The filtering process is defined now by the conditional distributions

$$\pi(f(X_t)) = E[f(X_t) \mid F_t^Y] \ , \ f \in C(S) \tag{2.2}$$

Denote by m_t the N-dimensional innovation :

$$m_t = (m_t^1, ..., m_t^N) \tag{2.3}$$

where $m_t^i = Y_t^i - \int_0^t \pi(h_s^i)ds$ $(i = 1, \dots, N)$ are F_t^Y - martingales. Note that the σ-field F_t^Y and $\sigma(m_v - m_u; t \le u \le v)$ are independent for all $t \ge 0$.

1. Filtering by Innovation Method.

THEOREM 2.1. If A is the infinitesimal generator of semigroup P_t of the system process X_t then the filtering process $\pi(f(X_t))$ satisfies the two following equations

$$\pi(f(X_t)) = \pi(f(X_0)) + \int_0^t \pi(Af(X_s))ds + \int_0^t (K_s, dm_s) \tag{2.4}$$

where $f \in \mathcal{D}(A)$,

$$K_s = (K_s^1, \dots, K_s^N) \ , \ 0 \le s \le t$$
$$K_s^i = (\pi(h_s))^{-1}[\pi(f(X_s)h^i(X_s) - \pi(f(X_{s-})\pi(h^i(X_s))] \ , \ i = 1, \dots, N \ .$$

$\pi(f(X_t)) = \pi_0(P_t f) + \int_0^t (G_s, dm_s)$, where $f \in \mathcal{D}(A)$ (2.5)

$$G_s = (G_s, \dots, G_s^N) \ , \ 0 \le s \le t \ ,$$
(2.5) $$G_s^i = (\pi(h_s^i))^{-1}[\pi(h^i(X_s)P_{t-s}f(X_{s-})) - \pi(P_{t-s}f(X_{s-}))] \ , i = 1, \dots, N \ .$$

Proof. a) Note that the process

$$C_t^f \overset{def}{=} f(X) - f(X_0) - \int_0^t Af(X_s)ds$$

is an F_t-martingale, so a direct application of the formula (1.8) for the semimartingale

$$f(X_t) = f(X_0) + \int_0^t Af(X_s)ds + C_t^f \tag{2.6}$$

yields (2.4) in noticing that the corresponding process u is 0 hence $\hat{u} = 0$, because of the independence of C_t^f from $M = (m)_t$.

b) It is known also that if $f \in C(S)$ and $t \ge 0$, the process

$$Q_t \equiv M_t^{f,t} \overset{def}{=} \begin{cases} f(X_t) & if \quad s \ge t \\ P_{t-s}f(X_s) & if \quad s \le t \end{cases} \tag{2.7}$$

is an F_t-martingale.

In writing the equation (1.8) for the signal Q_t at a fixed instant t then using an argument on a monotone class, we get (2.5)

2. Zakai equation

We have now

$$\pi(f(X_t)) = \frac{\sigma(f(X_t))}{\sigma(1)_t} \tag{2.8}$$

where $\sigma(f(X_t)) = E[L_t.f(X_t) \mid F_t^Y]$ which is caled now the quasi - filtering of X_t by f .

Denote again $\mu_t = (\mu_t^1, ..., \mu_t^N)$ with

$$\mu_t^i = Y_t^i - t \qquad (i = 1, ..., N) .$$

A direct application of Theorem 1.2 yields :

THEOREM 2.2. The quasi - filtering of X_t by f satisfies the following two types of stochastic differential equation :

$$\sigma(f(X_t)) = \sigma(f(X_0)) + \int_0^t \sigma(Af(X_s))ds$$
$$+ \int_0^t (\sigma(h_s.f(X_{s-}) - \sigma(f(X_{s-}), d\mu_s) \ , f \in C(S) \tag{2.9}$$

$$\sigma(f(X_t)) = \sigma(P_t f(X_0)) + \int_0^t (\sigma_s(h_s P_{t-s} f(X_{s-}))-$$
$$- \ \sigma(P_{t-s} f(X_{s-})), d\mu_s) \ , f \in C(S) \ . \tag{2.10}$$

Remarks. If X_t is of continuous sample paths, $X_{s-} = X_s \quad (0 \le \sigma \le t)$ hence the two above equations can be briefly rewritten as follows

$$\sigma_t(f) = \sigma_0(f) + \int_0^t \sigma_s(Af)ds$$
$$+ \int_0^t (\sigma_s(hf) - \sigma_s(f), d\mu_s) \ , f \in C(S) \ . \tag{2.9*}$$

$$\sigma_t(f) = \sigma_0(P_t f) + \int_0^t (\sigma_s(hP_{t-s}f)-$$
$$- \ \sigma_s(P_{t-s}f), d\mu_s) \ , f \in C(S) \ . \tag{2.10*}$$

3. A STOCHASTIC DIFFERENTIAL EQUATION

The process X_t is supposed as in Section 2, in particular it takes values in a compact Hausdorff space. Denote by M(S) the set of all probability measures over S. Then M(S) is also a compact Hausdorff space with the induced topology. Assume that $Y_t \ , t \ge 0$ is an N - vector ~~standard Poisson~~ point process and set $\mu_t = (Y_t^1 - t, ..., Y_t^N - t)$, $M_t^i =$

$Y_t^i - t$, $i = 1, ..., N$. Let σ_0 be an M(S) - valued random variable independent of (μ_t) defined on a probability space (Ω, F, P).

An M(S) - valued stochastic process σ_t is called a solution of the following equation

$$\sigma_t(f) = \sigma_0(P_t f) + \int_0^t (\sigma_t(hP_{t-s}f) - \sigma_t(P_{t-s}f), d\mu_s) \ , \tag{3.1}$$

where $\sigma_t(f) = \int f(X_t) d\sigma_t$ for $f \in C(S)$, if σ_s is independent of the σ - field $\sigma(\mu_v - \mu_u \ ; s \le u \le v)$ for all $s \ge 0$ and satisfies this equation. The quasi - filtering σ_t defined in §2 can be considered as a solution of (3.1), where Y_t is the observation process.

THEOREM 3.1. There exists a unique solution σ_t of (3.1) for arbitrary initial condition σ_0. Furthermore, this solution σ_t is $\sigma(\mu_s - \mu_0 \ ; 0 \le s \le t) \vee F(\sigma_0)$ measurable, where $F(\sigma_0)$ is the σ - field generated by the M(S) - valued random variable σ.

Proof. The method of Kunita will be repeated step by step to prove the above stated theorem. For simplicity we shall prove the theorem in case N = 1.

a/ We prove first the uniqueness of the solution of (1.3). Let σ_t and σ_t' be solutions of (1.3) with the same initial condition σ_0. Set

$$\varrho_t(f) = E(|\sigma_t(f) - \sigma_t'(f)|^2) \ . \tag{3.2}$$

$$\varrho_t(f) \le 2E(|\sigma_t(f)|^2 + |\sigma_t'(f)|^2) \le 4||f||^2 \ . \tag{3.3}$$

where $||f|| = sup_{x \in S} f(x)$.

On the other hand we have by definition (3.2) of $\varrho_t(f)$:

$$\varrho_t(f) \le E[\int_0^t (\sigma_s(hP_{t-s}f) - \sigma_s'(hP_{t-s}f))d\mu_s]^2 + E[\int_0^t (\sigma_s(P_{t-s}f) - \sigma_s'(P_{t-s}f)d\mu_s]^2$$
$$\le \int_0^t E[\sigma_s(hP_{t-s}f) - \sigma_s'(hP_{t-s}f)]^2 ds + \int_0^t E[\sigma_s(P_{t-s}f) - \sigma_s'(P_{t-s}f)]^2 ds \ .$$

Thus

$$\varrho_t(f) \le \int_0^t \varrho_s(hP_{t-s}f)ds + \int_0^t \varrho_s(f)ds \ . \tag{3.4}$$

By virtue of (3.3(the latter relation can be rewritten as follows

$$\varrho_t(f) \le \int_0^t 4||hP_{t-s}f||ds + \int_0^t 4||f||ds$$

or in noting that $||P_{t-s}f|| \leq ||f||$:

(3.5) $$\varrho_t(f) \leq 4||f||(||h||+1)t \ .$$

Substitute (3.5) to the right hand side of (3.4) and repeat this procedure n - times. We get then

(3.6) $$\varrho_t(f) \leq 4||f||(||h||+1)^n \frac{t^n}{n!} \ .$$

Let n tend to infinity : the right hand side of (3.6) tends to zero and we have for any $f \in C(S)$

$$\varrho_t(f) = 0$$

that prove the uniqueness of the solution.

b/ To prove the measurability of the solution, consider the following equation

(3.7) $$\sigma_t(f) = \sigma_0(P_t f) + \int_0^t [\hat{\sigma}_s(hP_{t-s}f) - \hat{\sigma}(P_{t-s}f)]d\mu_s \ .$$

where $\hat{\sigma}_t(f)$ is the truncated function of $\sigma_t(f)$ by a finite positive constant K :

(3.8) $$\hat{\sigma}_t(f) = \sigma_t(f) \quad if \quad |\sigma_t(f)| \leq K$$

and $\hat{\sigma}_t(f) = K \quad if \quad |\sigma_t(f)| \geq K$

Set $\sigma_t^{(0)}(f) = \sigma_0(P_t f)$

(3.9) $$\sigma_t^{(1)}(f) = \sigma_0(P_t f) + \int_0^t [\hat{\sigma}_s^{(0)}(hP_{t-s}f) - \hat{\sigma}_s^0(P_{t-s}f)d\mu_s$$

$$\sigma_t^{(n)}(f) = \sigma_0(P_t f) + \int_0^t [\hat{\sigma}_s^{(n-1)}(hP_{t-s}f) - \hat{\sigma}_s^{(n-1)}(P_{t-s}f)]d\mu_s$$

(3.10) $$\varrho_t^{(n)}(f) = E(|\sigma_t^{(n+1)}(f) - \sigma_t^{(n)}(f)|^2)$$

A computational analogous to that of for (3.4) yields

(3.11) $$\varrho_t^{(n)}(f) \leq \int_0^t [\varrho_t^{(n-1)}(hP_{t-s}f) + \varrho_t^{(n-1)}(P_{t-s}f)]ds \ , n \geq 1$$

On the other hand, in choosing $K > ||h||$ we have

$$
\begin{aligned}
(3.12)\qquad \varrho_t^{(0)}(f) &= E(|\sigma_t^{(1)}(f) - \sigma_t^{(0)}(f)|^2) \\
&\le \int_0^t E(|\sigma_s^{(0)}(hP_{t-s}f) - \sigma_s^{(0)}(P_{t-s}f)|^2)ds \\
&\le (K^2||f||^2 + ||f||^2)t = (K^2+1)||f||^2 t\ .
\end{aligned}
$$

Substituting (3.12) into the right hand side of (3.11) we have :

$$
\varrho_t^{(1)} \le (K^2+1)^2 ||f||^2 \frac{t^2}{2!}\ .
$$

By induction

$$
(3.13)\qquad \varrho_t^{(n)} \le (K^2+1)^{n+1} ||f||^2 \frac{t^n}{2!}\ .
$$

By consequence the series $\sum_{n=0}^{\infty} [\varrho_t^{(n)}(f)]^{\frac{1}{2}}$ is convergent. We can see from the following relation

$$
\begin{aligned}
||\sigma_t^{(n)}(f) - \sigma_t^{(m)}||_{L^2(P)} &= E[|\sigma_t^{(n)}(f) - \sigma_t^{(m)}(f)|^2] \\
&\le \sum_{k=m}^{n-1} [\varrho_t^{(k)}(f)]^{\frac{1}{2}}
\end{aligned}
$$

that $\sigma_t^{(n)}(f)$ convergese in $L^2(P)$ to a stochastic process $\sigma_t^{(\infty)}(f)$. Because each $\sigma_t^{(n)}(f)$ is measurable with respect to $\sigma(\xi_v - \xi_u\ ; 0 \le u \le v \le t)$ so $\sigma_t^{(\infty)}(f)$ is.

Suppose now that σ_t is the M(S) - valued solution of (3.1). If the constant K is chosen such as $K > max(||f||, ||h||)$ then σ_t is also a solution of (3.7) since $\sigma_t(f) = \hat{\sigma}_t(f)\ , f \in C(S)$.

It holds

$$
(3.14)\qquad E|\sigma_t^{(\infty)}(f)|^{\frac{1}{2}} \le E[\sigma_t^{(0)}(f)]^{\frac{1}{2}} + \sum_{n=0} \varrho_t^{(n)}(f) \le c||f||
$$

where c is a positive constant depending on t,K only. Set $\bar{\varrho}_t(f) = E[|\sigma_t(f) - \sigma_t^{(\infty)}(f)|^2]$. We have by virtue of (3.14)

$$
(3.15)\qquad \bar{\varrho}_t(f) \le E|\sigma_t(f)|^2 + E|\sigma_t^{(\infty)}(f)|^2 \le c'||f||^2
$$

and

$$
\bar{\varrho}_t(f) \le \int_0^t [\bar{\varrho}_s(hP_{t-s}f) + \bar{\varrho}_s(P_{t-s}f)]ds\ .
$$

By an argument similar to the above we can see that $\bar{\varrho}_t(f) = 0$ hence $\sigma_t = \sigma_t^{(\infty)}$. The measurability of σ_t with respect to $\sigma(Y_v - Y_u \; ; 0 \leq u \leq v \leq t) \vee F(\sigma_0)$ is thus proved.

c/ As we see in §2, if μ_t is defined such that the Poisson process Y_t is the observation process then the quasi - filtering σ_t based on F_t^Y is the solution of (3.1) and it can be expressed by a functional

$$\sigma_t = \Phi(\sigma_0, \mu_s - \mu_0 \; ; 0 \leq s \leq t) \; .$$

When Y_t and μ_t are defined as in in the beginning of §3, then the same functional Φ expresses the solution of (3.1) in terms of σ_0 and $\mu_s - \mu_0$, $0 \leq s \leq t$. Thus, the existence of (3.1) is proved.

REMARK. As in [3] we can show that the M(S) - valued solution σ_t of (3.1) is a Feller - Markov process with the transition probability $\Pi_t(\nu, B)$ defined by

$$\Pi_t(\nu, B) = P(\sigma_t^\nu \in B) \; ; B = \text{Borel sets of M(S)} \; ,$$

where σ_t^ν is the solution corresponding to the initial condition $\nu \in M(S)$.

REFERENCES

[1] Brémaud, P. Point Processes and Queues, Martingale Dynamic, Springer Series in Mathematics, Springer Verlag, 1980.

[2] Elliott, Robert J. Filtering and Control for Point Process Observations, Universität Konstanz Fakultät für Wirtschaftswissenschaften und Statistik, Research Report No 09 - 86 - 1, Nr 110/s, May 1988.

[3] Kunita, H. Asymptotic Behavior of the Nonlinear Filtering Errors of Markov Process.

[4] Tran Hung Thao, Filtering from Point Process Observations, Preprint, Hanoi. 1986.