S.-S. Lee
Nagoya Math. J.
Vol. 91 (1983), 173–184

NON-LINEAR PREDICTION PROBLEMS FOR ORNSTEIN-UHLENBECK PROCESS

SHEU-SAN LEE

§ 0. Introduction

We shall discuss in this paper some problems in non-linear prediction theory. An Ornstein-Uhlenbeck process $\{U(t)\}$ is taken to be a basic process, and we shall deal with stochastic processes $X(t)$ that are transformed by functions f satisfying certain condition. Actually, observed processes are expressed in the form $X(t) = f(U(t))$. Our main problem is to obtain the best non-linear predictor $\hat{X}(t, \tau)$ for $X(t + \tau)$, $\tau > 0$, assuming that $X(s)$, $s \leq t$, are observed. The predictor is therefore a non-linear functional of the values $X(s)$, $s \leq t$.

Non-linear prediction theory that discusses how to obtain such non-linear predictors has been considered in various situations. For instance, I. I. Gihman and A. V. Skorohod (cf. [1], § 8, Chapter IV, Vol. I) have considered optimum mean square predictor of $X(t + \tau)$, $\tau > 0$, assuming that the basic process $V(s)$, $s \leq t$, itself is observed. As is well-known the predictor $\hat{X}(t, \tau)$ is given by the conditional expectation:

$$(1) \qquad \hat{X}(t, \tau) = E\{X(t + \tau) | \mathscr{B}_t(V)\} , \qquad \mathscr{B}_t(V) = \sigma\{V(s);\ s \leq t\} .$$

While A. M. Yaglom [5] has discussed the optimum mean square predictor assuming that Markov process $V(t)$ is transformed by a function f with inverse f^{-1} and that $X(s) = f(V(s))$, $s \leq t$, are observed.

In this case, it holds evidently that

$$(2) \qquad \begin{aligned} \hat{X}(t, \tau) &= E\{X(t + \tau) | X(s);\ s \leq t\} = E\{X(t + \tau) | X(t)\} \\ &= E\{X(t + \tau) | V(t)\} = E\{X(t + \tau) | V(s);\ s \leq t\} . \end{aligned}$$

Yaglom's situation coincides with ours in the sense that the $X(s)$ are assumed to be given for $s \leq t$. In this case, too, the predictor (2) coincides with (1) actually, because f is invertible.

Received May 24, 1982.

https://doi.org/10.1017/S002776300002050X Published online by Cambridge University Press

We now clarify the best non-linear predictor of $X(t+\tau)$, $\tau > 0$, which is the main topic of this paper. Now let $\mathscr{B}_t(X)$ be the least σ-field generated by $X(s)$; $s \leq t$, and let $\hat{X}(t, \tau)$ be the $\mathscr{B}_t(X)$-measurable random variable for which

$$E|X(t+\tau) - \hat{X}(t, \tau)|^2 \tag{3}$$

attains the minimum of $E|X(t+\tau) - Y|^2$, Y being $\mathscr{B}_t(X)$-measurable random variable with finite variance. Such a random variable $\hat{X}(t, \tau)$ always exists and it is called the best non-linear predictor of $X(t+\tau)$, $\tau > 0$. Evidently, it is given by the conditional expectation:

$$\hat{X}(t, \tau) = E\{X(t+\tau) | \mathscr{B}_t(X)\} . \tag{4}$$

It depends on the properties of the basic processes and on the structure of the function f whether the explicit value of $\hat{X}(t, \tau)$ can be given or not.

We shall discuss how to obtain the best non-linear predictor $\hat{X}(t, \tau)$ of $X(t+\tau)$, $\tau > 0$, by means of the observed values $X(s)$, $s \leq t$, here the process $\{X(s)\}$ is a new process transformed from Ornstein-Uhlenbeck process $\{U(s)\}$ by the function f, namely $\{X(s)\} = \{f(U(s))\}$.

As is mentioned above, the explicit form of $\hat{X}(t, \tau)$ depends on the structure of function f. Therefore our discussion will be restricted to the following several cases: (1)–(4). In each case we are able to obtain the exact value of the predictors.

(1) The function f has a single valley (peak) and is not symmetric (cf. Theorem 1).

(2) The function f has a single valley (peak) and is symmetric on a bounded interval (cf. Theorem 2).

(3) The function f is symmetric on $(-\infty, \infty)$ (cf. Theorem 3).

(4) The function f has several valleys (peaks) (cf. Theorem 4).

We have hopes that similar results would be obtained in more general cases where the basic processes are taken to be a multiple Markov process. Such an approach will be discussed in separate paper.

ACKNOWLEDGEMENT. I am very grateful to Professor I. Kubo for having given me kind help for writing this paper.

§1. Background

We will take a canonical Ornstein-Uhlenbeck process $\{U(t)\}$ as the basic process, and will assume regular condition on f. By a canonical

https://doi.org/10.1017/S002776300002050X Published online by Cambridge University Press

Ornstein-Uhlenbeck process we mean a Gaussian process with continuous paths, with expectation zero and with covariance $E\{U(t+\tau)U(t)\} = e^{-|\tau|}$.

The Ornstein-Uhlenbeck process $\{U(t)\}$ is not only a Gaussian stationary process but also a strong Markov process together with $\{U(-t)\}$. These properties will play important roles in the later discussion.

The semi-group $\{T_\tau;\ \tau \geq 0\}$ of $U(t)$ is given as follows

$$(5)\qquad (T_\tau f)(x) = \int_{-\infty}^{\infty} f(y)[2\pi(1-e^{-2\tau})]^{-1/2} \exp\left\{-\frac{(y-e^{-\tau}x)^2}{2(1-e^{-2\tau})}\right\}dy$$

provided the integral exists.

LEMMA 1. (i) $E\{X(t+\tau)\,|\,\mathscr{B}_t(U)\} = (T_\tau f)(U(t))$,

$$\mathscr{B}_t(U) = \sigma\{U(s);\ s \leq t\}\,,$$

(ii) $\hat{X}(t,\tau) = E\{(T_\tau f)(U(t))\,|\,\mathscr{B}_t(X)\}$.

Proof. (i) is clear from the Markov property of $U(t)$. (ii) is also obvious since we have (i) and

$$\hat{X}(t,\tau) = E\{X(t+\tau)\,|\,\mathscr{B}_t(X)\} = E\{E[X(t+\tau)\,|\,\mathscr{B}_t(U)]\,|\,\mathscr{B}_t(X)\}\,.$$

Our approach will be illustrated by the following three examples.

EXAMPLE 1. If f is a strictly monotone function, then we see that $\mathscr{B}_t(U) = \mathscr{B}_t(X)$, since the σ-fields generated by a single random variable $U(t)$ and by $X(t)$, respectively, coincide with each other for each t. Therefore we easily verify

$$(6)\qquad \hat{X}(t,\tau) = (T_\tau f)(U(t)) = (T_\tau f)(f^{-1}(X(t)))\,.$$

Such cases have been discussed by Yaglom [5], Zabotina [6] and others.

EXAMPLE 2. Let $H_n(x)$ be the Hermite's polynomial of degree n defined by

$$\exp\left\{sx - \frac{1}{2}s^2\right\} = \sum_{n=0}^{\infty} \frac{s^n}{n!} H_n(x)\,,$$

and put $X(t) = H_n(U(t))$. If n is odd, then although the σ-fields generated by $U(t)$ and $X(t)$ respectively, do not coincide with each other for each t, we can still show $\mathscr{B}_t(U) = \mathscr{B}_t(X)$, and hence the best non-linear predictor of $X(t+\tau)$ and the mean square error are given as follows

$$(7)\qquad \hat{X}(t,\tau) = e^{-n\tau}X(t)\,, \qquad \sigma^2(\tau) = (1-e^{-2n\tau})n!\,.$$

https://doi.org/10.1017/S002776300002050X Published online by Cambridge University Press

Generally, if the equality $\mathscr{B}_t(U) = \mathscr{B}_t(X)$ holds, then the best non-linear predictor $\hat{X}(t, \tau)$ is represented by using an explicit function $(T_\tau f)(U(t))$, namely

$$\hat{X}(t, \tau) = (T_\tau f)(U(t)) .$$

However the value of $U(t)$ is not always determined by means of observed values $X(s) = f(U(s))$ for $s \leq t$. We are therefore interested in the cases where the value of $U(t)$ is determined from observed values under suitable conditions for f, so as the predictor is obtained explicitly.

EXAMPLE 3. We shall then discuss a case where $\mathscr{B}_t(U) \neq \mathscr{B}_t(X)$ but

$$\hat{X}(t, \tau) = (T_\tau f)(U(t)) \tag{8}$$

does hold. For instance, let n be an even number in Example 2. Then we can show that $\mathscr{B}_t(U) \neq \mathscr{B}_t(X)$ but (8) does hold, actually (7) does hold (cf. Theorem 3 and Example 5). The mean square error of the best non-linear predictor given by (8) is

$$\sigma^2(\tau) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)^2 \exp\left(-\frac{x^2}{2}\right) dx - \frac{1}{2\pi\alpha} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) f(y) \times \exp\left\{-\frac{(x^2+y^2) - 2e^{-2\tau}xy}{2\alpha^2}\right\} dxdy ,$$

with $\alpha = (1 - e^{-4\tau})^{1/2}$.

The canonical Ornstein-Uhlenbeck process can be canonically represented (with respect to the past and to future, respectively) by some canonical Brownian motion $B(t)$ and $\tilde{B}(t)$ in such a way that

$$U(t) = \sqrt{2}\, e^{-t} \int_{-\infty}^{t} e^{\lambda} dB(\lambda) = \sqrt{2}\, e^{t} \int_{t}^{\infty} e^{-\lambda} d\tilde{B}(\lambda) . \tag{9}$$

Here it is noted that

$$\begin{aligned} U(t) - U(s) &= \sqrt{2}\,[B(t) - B(s)] - \int_s^t U(r) dr \\ &= \sqrt{2}\,[\tilde{B}(s) - \tilde{B}(t)] + \int_s^t U(r) dr . \end{aligned} \tag{10}$$

PROPOSITION 1. *If the function f is differentiable and $T = T(\omega)$ $(<\infty)$ is a stopping time of $U(t)$, then*

$$O(T, \omega) \equiv \limsup_{r \downarrow T} \frac{f(U(r)) - f(U(T))}{2\sqrt{|r - T| \log\log 1/|r - T|}} = |f'(U(T))| \tag{11}$$

https://doi.org/10.1017/S002776300002050X Published online by Cambridge University Press

holds. Moreover, if T is a stopping time of the time reversed Ornstein-Uhlenbeck process $\{U(-t)\}$, then also

$$(11)' \qquad O(T, \omega) \equiv \limsup_{r \uparrow T} \frac{f(U(r)) - f(U(T))}{2\sqrt{|r-T| \log\log 1/|r-T|}} = |f'(U(T))| .$$

Proof. By using the law of the iterated logarithm for Brownian motion:

$$(12) \qquad \limsup_{r \to s} \frac{B(r, \omega) - B(s, \omega)}{\sqrt{2|r-s| \log\log 1/|r-s|}} = 1 ,$$

the formula (10) and the strong Morkov property imply the formula (11).

§ 2. The best non-linear prediction problem for functions with single valley

Now, we assume that the function f is continuous on $(-\infty, \infty)$, strictly monotone decreasing (resp. increasing) and differentiable on $(-\infty, 0)$ (resp. on $(0, \infty)$) and that f is normalized as $f(0) = 0$.

LEMMA 2. *Putting $\theta = \inf\{u > 0; f(u) \neq f(-u)\}$, we have the following:*

(i) *If $\theta = 0$, then for any $\varepsilon > 0$ there exist u and $\bar{u}$ in the neighbourhood $D = \{u; |u| < \varepsilon\}$ such that*

$$f(u) = f(\bar{u}) , \quad |f'(u)| \neq |f'(\bar{u})| , \quad u > 0, \quad \bar{u} < 0 .$$

(ii) *If $0 < \theta < \infty$, then for any $\varepsilon > 0$ there exist u and $\bar{u}$ in the neighbourhood $D = \{u; 0 < f(u) - f(\theta) < \varepsilon\}$ such that*

$$f(u) = f(\bar{u}) , \quad |f'(u)| \neq |f'(\bar{u})| , \quad u > 0 , \quad \bar{u} < 0 .$$

(iii) *If $\theta = \infty$, then for any u, $f(u) = f(-u)$ holds, namely the function f is symmetric.*

We are going to discuss our prediction problems dividing them into three cases, $\theta = 0$, $0 < \theta < \infty$, $\theta = \infty$, by virtue of the lemma above.

For $h > 0$ the inverse image of the function f consists of two points with different sign; denote by $f_{+}^{-1}(h)$ the positive one and by $f_{-}^{-1}(h)$ the negative one. Moreover, we define a stopping time $T(h, t, \omega)$ of the time reversed Ornstein-Uhlenbeck process $\{U(-t)\}$ for $h \geq 0$ by

$$(13) \qquad T(h) \equiv T(h, t, \omega) \equiv \sup\{q;\ X(q) = h,\ q < t\} .$$

If $\theta = 0$, then we can choose a positive monotonical sequence $h_n \downarrow 0$ such that

https://doi.org/10.1017/S002776300002050X Published online by Cambridge University Press

$$|f'(f_+^{-1}(h_n))| \neq |f'(f_-^{-1}(h_n))| \tag{14}$$

holds for each n, by Lemma 2 (i).

THEOREM 1. *If $\theta = 0$, then $U(t)$ is $\mathscr{B}_t(X)$-measurable and the best non-linear predictor of $X(t+\tau)$, $\tau > 0$, is given by*

$$\hat{X}(t, \tau) = (T_\tau f)(U(t)) .$$

Actually, if $X(t) = 0$, then the value of $U(t)$ is equal to zero. If $X(t) > 0$, then taking a sequence $\{h_n\}$ as above and choosing n with $h_n < X(t)$, we are given the value of $U(t)$ by

$$U(t) = \begin{cases} f_+^{-1}(X(t)) & \text{if } O(T(h_n, t, \omega), \omega) = |f'(f_+^{-1}(h_n))| , \\ f_-^{-1}(X(t)) & \text{if } O(T(h_n, t, \omega), \omega) = |f'(f_-^{-1}(h_n))| . \end{cases}$$

Proof of Theorem 1. We assume that for a fixed t the values $\{X(r); r \leq t\}$ are observed. By the conditions of this theorem and Lemma 2 (i) there exists an h_n which satisfies (14). Let $T(h_n, t, \omega)$ be the stopping time defined by (13). Then by using the Proposition 1 we have

$$O(T(h_n, t, \omega), \omega) = |f'(U(T(h_n)))| . \tag{15}$$

Thus from (14) and (15) the values of $U(T(h_n))$ are determined as follows

$$U(T(h_n)) = \begin{cases} f_+^{-1}(h_n) & \text{if } O(T(h_n, t, \omega), \omega) = |f'(f_+^{-1}(h_n))| , \\ f_-^{-1}(h_n) & \text{if } O(T(h_n, t, \omega), \omega) = |f'(f_-^{-1}(h_n))| . \end{cases}$$

Then the question is how to determine the value of $U(t)$ by means of the value $U(T(h_n))$. If $X(t, \omega) > 0$ and h_n is chosen so as to hold $h_n < X(t, \omega)$, then by the definition of the $T(h_n, t, \omega)$, $X(r)$ does not pass through the point zero in the time interval $(T(h_n), t)$. Hence if $U(T(h_n)) > 0$ (<0), then $U(t) > 0$ (<0). Namely

$$U(t) = \begin{cases} f_+^{-1}(X(t)) & \text{if } U(T(h_n)) = f_+^{-1}(h_n) , \\ f_-^{-1}(X(t)) & \text{if } U(T(h_n)) = f_-^{-1}(h_n) . \end{cases}$$

Therefore, the value of $U(t)$ is uniquely determined by the observed values of $X(r)$ for $r \leq t$. Hence $U(t)$ is $\mathscr{B}_t(X)$-measurable. Thus

$$\hat{X}(t, \tau) = E\{(T_\tau f)(U(t)) \,|\, \mathscr{B}_t(X)\} = (T_\tau f)(U(t)) .$$

The above results are also valid when the graph of the equation $y = f(x)$ involves parallel counter parts.

COROLLARY. *Suppose that the function f is continuously differentiable on $(-\infty, \infty)$ except at a point v_0, strictly monotone in both sides of the point v_0 respectively and not symmetric (with respect to v_0) in any neighbourhood of the point v_0. Then $\mathscr{B}_t(X) = \mathscr{B}_t(U)$ holds and the best non-linear predictor of $X(t+\tau) = f(U(t+\tau))$, $\tau > 0$, is given by*

$$\hat{X}(t, \tau) = (T_\tau f)(U(t)) .$$

THEOREM 2. *If $0 < \theta < \infty$, then the best non-linear predictor of $X(t+\tau)$, $\tau > 0$, is given by*

$$\hat{X}(t, \tau) = \begin{cases} (T_\tau f)(U(t)) & \text{if } \omega \in W , \\ \frac{1}{2}[(T_\tau f)(f_+^{-1}(X(t))) + (T_\tau f)(f_-^{-1}(X(t)))] & \text{if } \omega \notin W , \end{cases}$$

where $W \equiv \{\omega;\ X(s, \omega) > f(\theta),\ T(0, t, \omega) < {}^{\exists}s \leq t\}$ and *$T(0, t, \omega)$ is given by (13). The value of $U(t, \omega)$ is given by (18) for $\omega \in W$.*

Proof. Since $0 < \theta < \infty$, by Lemma 2 (ii) we can choose a monotone decreasing sequence $\{h_n\}$ such that $h_n \downarrow f(\theta)$ as $n \to \infty$ and

$$f'(f_+^{-1}(h_n)) \neq -f'(f_-^{-1}(h_n)) \tag{16}$$

holds.

Since for any $\omega \in W$ there exists s satisfying $X(s, \omega) > f(\theta)$, $T(0, t, \omega) < s \leq t$, by the continuity of the path, there exists n such that

$$T(0, t, \omega) < T(h_n, t, \omega) \leq t . \tag{17}$$

Moreover by the Proposition 1 we have

$$O(T(h_n, t, \omega), \omega) = |f'(U(T(h_n)))| \qquad \text{for every } n, \text{ a.e. } \omega$$

and hence the value of $U(T(h_n))$ is determined by

$$U(T(h_n)) = \begin{cases} f_+^{-1}(h_n) & \text{if } O(T(h_n), \omega) = |f'(f_+^{-1}(h_n))| , \\ f_-^{-1}(h_n) & \text{if } O(T(h_n), \omega) = |f'(f_-^{-1}(h_n))| . \end{cases}$$

Furthermore by using the arguments similar to the proof of Theorem 1, we obtain

$$U(t) = \begin{cases} f_+^{-1}(X(t)) & \text{if } O(T(h_n), \omega) = |f'(f_+^{-1}(h_n))| , \quad \omega \in W , \\ f_-^{-1}(X(t)) & \text{if } O(T(h_n), \omega) = |f'(f_-^{-1}(h_n))| , \quad \omega \in W . \end{cases} \tag{18}$$

Therefore the value of $U(t)$ is uniquely determined in terms of the observed values $X(r)$ for $r \leq t$ under the condition W that there exists

s such that $X(s) > f(\theta)$ and $T(0, t, \omega) < s \leq t$. Namely under this condition $U(t)$ is $\mathscr{B}_t(X)$-measurable. However, since

$$\begin{aligned}\hat{X}(t, \tau) &= E\{(T_\tau f)(U(t)) \mid \mathscr{B}_t(X)\} \\ &= (T_\tau f)(U(t))\chi_W + \chi_{W^c} E\{(T_\tau f)(U(t)) \mid \mathscr{B}_t(X)\} ,\end{aligned}$$

we have

$$\hat{X}(t, \tau) = (T_\tau f)(U(t)) , \qquad \omega \in W . \tag{19}$$

If $\omega \notin W$, i.e. $X(s, \omega) \leq f(\theta)$ for any s in the time interval $(T(0, t, \omega), t]$, then during $(T(0, t, \omega), t]$, $U(s)$ stays within the interval $[-\theta, \theta]$, on which the function f is symmetric with respect to the axis of the ordinate. Although the value of $U(s)$ is not uniquely determined by the value of $X(s)$, the best non-linear predictor of $X(t + \tau)$, $\tau > 0$, under the condition W^c is given by

$$\begin{aligned}\hat{X}(t, \tau) &= E\{(T_\tau f)(U(t)) \mid \mathscr{B}_t(X)\} \\ &= E\{(T_\tau f)(U(t))\chi_{W^c} \mid \mathscr{B}_t(X)\} \qquad \text{for } \omega \notin W .\end{aligned} \tag{20}$$

To complete the proof of the theorem it is sufficient to show that

$$E\{(T_\tau f)(U(t))\chi_{W^c} \mid \mathscr{B}_t(X)\} = \tfrac{1}{2}\{(T_\tau f)(|U(t)|) + (T_\tau f)(-|U(t)|)\}\chi_{W^c} . \tag{21}$$

Since $W \in \mathscr{B}_t(X)$, the $\chi_{W^c}(\omega)$ is $\mathscr{B}_t(X)$-measurable so is the right-hand side of (21). Therefore for any $G \in \mathscr{B}_t(X)$, we must show

$$\begin{aligned}&\int_G (T_\tau f)(U(t))\chi_{W^c} dP(\omega) \\ &\quad = \int_G \frac{1}{2}\{(T_\tau f)(|U(t)|) + (T_\tau f)(-|U(t)|)\}\chi_{W^c} dP(\omega) .\end{aligned} \tag{22}$$

It can be shown by noting the strong Markov property of $U(t)$, the symmetry of the probability measure of Ornstein-Uhlenbeck process starting from the origin.

THEOREM 3. *If $\theta = \infty$, then the best non-linear predictor of $X(t + \tau)$,* $\tau > 0$, is given by

$$\hat{X}(t, \tau) = (T_\tau f)(U(t)) = (T_\tau f)(f_+^{-1}(X(t))) = (T_\tau f)(f_-^{-1}(X(t))) .$$

Proof. Since $f(u) = f(-u)$ holds for any u, it is easy to verify

$$(T_\tau f)(u) = (T_\tau f)(-u) = (T_\tau f)(|u|) .$$

Hence we have

https://doi.org/10.1017/S002776300002050X Published online by Cambridge University Press

$$\begin{aligned}\hat{X}(t,\tau) &= E\{(T_\tau f)(U(t))\,|\,\mathscr{B}_t(X)\} = E\{(T_\tau f)(U(t))\,|\,f(U(s));\, s\le t\}\\ &= E\{(T_\tau f)(|U(t)|)\,|\,|U(s)|;\, s\le t\} = (T_\tau f)(|U(t)|)\\ &= (T_\tau f)(-|U(t)|) = (T_\tau f)(f_+^{-1}(X(t))) = (T_\tau f)(f_-^{-1}(X(t)))\,.\end{aligned}$$

EXAMPLE 4. Set

$$f(u) = \begin{cases} -au & \text{if } u<0\,,\\ bu & \text{if } u\ge 0\end{cases}$$

with $a \neq b$, $ab > 0$, and put $X(t) = f(U(t))$. Then we can easily have $\theta = 0$, and hence we obtain

$$\begin{aligned}\hat{X}(t,\tau) = (T_\tau f)(U(t)) &= \frac{(a+b)\delta}{\sqrt{2\pi}}\exp\left\{-\frac{e^{-2\tau}}{2\delta^2}U(t)^2\right\}\\ &\quad + (b-a)e^{-\tau}U(t)\left\{\Phi\left(\frac{e^{-\tau}}{\delta}U(t)\right) - \Phi\left(-\frac{e^{-\tau}}{\delta}U(t)\right)\right\},\end{aligned}$$

where

$$\Phi(r) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{r}\exp\left(-\frac{y^2}{2}\right)dy \quad\text{and}\quad \delta = \sqrt{1-e^{-2\tau}}\,.$$

EXAMPLE 5. Set $X(t) = f(U(t)) = \beta|U(t)|$, $\beta \neq 0$. Then we obtain

$$\begin{aligned}\hat{X}(t,\tau) = (T_\tau f)(U(t)) &= \sqrt{\frac{2}{\pi}}\beta\delta\exp\left\{-\frac{1}{2}e^{-2\tau}\beta^{-2}\delta^{-2}X(t)^2\right\}\\ &\quad + \beta e^{-\tau}U(t)\left\{\Phi\left(\frac{e^{-\tau}}{\delta}U(t)\right) - \Phi\left(-\frac{e^{-\tau}}{\delta}U(t)\right)\right\}.\end{aligned}$$

EXAMPLE 6. Set $X(t) = f(U(t)) = (U(t))^n$, $n = 1, 2, \cdots$. Then we obtain

$$\hat{X}(t,\tau) = (T_\tau f)(U(t)) = \sum_{\ell=0}^{[n/2]}\binom{n}{2\ell}\exp\{-\tau(n-2\ell)\}\delta^{2\ell}(2\ell-1)!!\,X(t)^{(n-2\ell)/n}\,.$$

§ 3. The best non-linear prediction problem with several peaks

We have discussed the prediction problem for a simpler function f in Section 2. We now extend Theorem 1 for more general function f. Assume that the function f is continuously differentiable on $(-\infty, \infty)$ except at finite points $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ and that f is strictly monotone on each interval $(\alpha_\kappa, \alpha_{\kappa+1})$, $\kappa = 0, 1, 2, \cdots, n$, where $\alpha_0 = -\infty$, $\alpha_{n+1} = \infty$. Moreover assume that the function f is not symmetric in any neighbourhood of the points α_κ, namely for any $\varepsilon > 0$ there exist $u_{\kappa 1}$ and $u_{\kappa 2}$ in the neighbourhood $D_\kappa = \{u;\, |u-\alpha_\kappa| < \varepsilon\}$, $\kappa = 1, 2, \cdots, n$, such that

https://doi.org/10.1017/S002776300002050X Published online by Cambridge University Press

(23) $$f(u_{\kappa 1}) = f(u_{\kappa 2}), \quad |f'(u_{\kappa 1})| \neq |f'(u_{\kappa 2})|, \quad u_{\kappa 1} < \alpha_\kappa < u_{\kappa 2}.$$

As in Section 2, under these conditions, the predictor is given by

$$\hat{X}(t, \tau) = (T_\tau f)(U(t)),$$

and the algorithm to determine the value of $U(t)$ is seen.

For simplicity we shall consider only the case where the function f possesses one maximal value and one minimal value: $f_{\text{maximal}} = f(\alpha_1)$, $f_{\text{minimal}} = f(\alpha_2)$. Divide the region $(-\infty, \infty)$ of u into three intervals: $I_1 = (-\infty, \alpha_1]$, $I_2 = (\alpha_1, \alpha_2]$ and $I_3 = (\alpha_2, \infty)$. Denote by f_j the restriction of f to I_j and define f_j^{-1} on the interval I_j. Then of course, $-\infty = \alpha_0 < f_1^{-1}(x) \leq \alpha_1 \leq f_2^{-1}(x) \leq \alpha_2 \leq f_3^{-}(x) < \alpha_3 = \infty$.

We suppose that $X(r)$, $r \leq t$, are observed. Once we know the interval I_j to which $U(r)$ belongs at a given time r $(\leq t)$, then we can immediately determine the value of $U(r)$ by the value of $X(r)$ in such a way that

$$U(r) = f_j^{-1}(X(r)).$$

In view of this we will first show that there exists at least one random time point t_0 $(< t)$ such that the interval I_j including $U(t_0)$ is determined at the time. Secondly, we will show that we can trace the intervals which include $U(s)$ after the time t_0 by observing $X(s)$, $t_0 \leq s \leq t$.

By the property of the function f there exist h^* and j_0 such that

$$|f'(f_j^{-1}(h^*))| \neq |f'(f_{j_0}^{-1}(h^*))|, \quad j \neq j_0.$$

For instance, if either $f(\alpha_2) > f(\alpha_0) \equiv \lim_{u \to -\infty} f(u)$ or $f(\alpha_1) < f(\alpha_3) \equiv \lim_{u \to \infty} f(u)$ holds, then we may take h^* in such way that

$$f(\alpha_2) > h^* > f(\alpha_0) \quad \text{or} \quad f(\alpha_1) < h^* < f(\alpha_3).$$

Even in the contrary case we can choose h^* which satisfies the above condition.

For the h^* define a stopping time $T(h^*, t, \omega)$ by

$$T(h^*) \equiv T(h^*, t, \omega) = \sup\{q; X(q) = h^*, q < t\},$$

then by the property of Ornstein-Uhlenbeck process we can easily see that $T(h^*, t, \omega) > -\infty$. Therefore by using the argument similar to that in the proof of the Theorem 1 we can determine the value of $U(T(h^*))$, in particular we know the interval which includes $U(T(h^*))$.

Suppose that at a given time $r\,(< t)$ the interval including $U(r)$ is known, say I_j. If the value $X(s)$ hits $f(\alpha_j)$ at a time earlier than $f(\alpha_{j-1})$ after the time r, then we can know which I_j or I_{j+1} does include $U(s)$, for s in the time interval between r and the first hitting time to the set $\{f(\alpha_{j-1}), f(\alpha_{j+1})\}$ after the above hitting time to $f(\alpha_j)$, by a similar way to the proof of Theorem 1 observing the variations of X at points h such that (23) holds with $h = f(u_{j1}) = f(u_{j2})$. Namely, the value of $U(s)$ is determined for any s in the above time interval. In the complementary case, the value of $U(s)$ is similarly determined for any s in the time interval between r and the first hitting time to $\{f(\alpha_{j-2}), f(\alpha_j)\}$ after the first hitting time to $f(\alpha_{j-1})$ after r.

We have thus known the value of $U(r)$ at the time $r = T(h^*)$. Then applying the above discussion recursively, we can determine the value of $U(s)$, $T(h^*) \leq s \leq t$, especially the value of $U(t)$, in terms of $X(s)$, $s \leq t$. Thus we see that $U(t)$ is $\mathscr{B}_t(X)$-measurable, and hence we have proved the following theorem.

THEOREM 4. *If the function f satisfies the condition explained in the b eginning of this section. Then for $X(t) = f(U(t))$ the equality $\mathscr{B}_t(X) = \mathscr{B}_t(U)$ valid is and the best non-linear predictor of $X(t + \tau)$, $\tau > 0$ is given by*

$$\hat{X}(t, \tau) = (T_\tau f)(U(t)) .$$

REFERENCES

[1] I. I. Gihman and A. V. Skorohod, The Theory of Stochastic Processes, Springer-Verlag, Berlin, 1974.

[2] Yu. A. Rozanov, Stationary random processes, Holden-Day, San Francisco, 1967.

[3] T. Hida and G. Kallianpur, The square of a Gaussian Markov process anu non-linear prediction, Multivar. Anal., **5** No. 4 (1975), 451–461.

[4] A. V. Balakrishnan, Prediction theory for Markoff processes, Pacific J. Math., **11** (1961), 1171–1182.

[5] A. M. Яглом, Примеры оптимального нелинейного зкстраполирования стационарных случайных процессов. Тр. VI Всес. совещ. по теории вероятностей и матем. статистике. Вильнюс-Паланага, 1960, Вильнюс, Гос. изд-во полит. и научн. лиа-ры ЛитССР, (1962), 275-296.

[6] Н. П. Заботина, Онтимальных нелинейные зкстраполяция, Фильтрация и интерполяция Функций от Гауссовских процессов. Проблемы передачи информаций, Том XI, Вып. **2**, (1975), 84–95.

Department of Mathematics
Faculty of Science
Nagoya University
Japan

Current Address:
Shenyang Chemical Engineering Institute
Shenyang
China