Acta Applicandae Mathematicae **63:** 433–439, 2000.

Entropy in Subordination and Filtering

Dedicated to Professor Takeyuki Hida on the occasion of his 70th birthday

SI SI[1] and WIN WIN HTAY[2]
[1]*Faculty of Science and Technology, Aichi Prefectural University, Aichi Prefecture, Japan*
[2]*Department of Computational Mathematics, University of Computer Studies, Yangon, Myanmar (Burma)*

(Received: 7 July 1999)

Abstract. We propose a stochastic model of transmitting random information at random time. In this model, the signal is observed as a random sampling according to an increasing stable stochastic process. Thus we are given a subordinate stochastic process which is a typical irreversible process. As the characteristic of this phenomea we observe the loss of entropy.

Mathematics Subject Classification (2000): 60H40.

Key words: white noise, subordination, filtering.

1. Introduction

We are interested in various methods of modulation of stochastic processes which are used as models in communication theory. Among them, frequency modulation, amplitude modulation are well known.

Having been inspired by astrophysical data, we propose a stochastic model of transmitting random information. Thus, we are going to introduce another modulation using Bochner's subordination in this note.

Observing typical X-ray light curves which are flickering from a black hole, we see that the curves have self-similarity (cf. [7]). We may therefore guess, in fact conclude, that those curves are sample functions of a stable stochastic process. Now, one may ask, what should be the input signal under the obtained information?

Note that if the information source is random with a limited power, it is taken to be Gaussian and to be optimal in the sense of information theory. Thus, in our model we may admit to white noise $\dot{B}(t)$ being an information source.

Since the accumulation sum of white noise according to time propagation is a Brownian motion, formally writing,

$$\int_0^t \dot{B}(s)\,\mathrm{d}s = B(t), \tag{1.1}$$

the Brownian motion can be taken as an input signal.

Emission of the signal comes out at random times and is subject to the law of the exponential holding time with various intensities. It leads us to think of the emitting time series as a compound Poisson process; in particular an increasing stable process that enjoys self-similarity.

Applying the technique of subordination, the observed data can be expressed as $B(Y_\alpha(t))$, which is to be self-similar. The information loss shows the 'irreversibility' from the view point of information theory.

2. Prerequisite

In this section, we recall some well-known facts which we are going to apply in what follows.

2.1. COMPOUND POISSON PROCESSES

We start with a system of elementary stochastic processes for Lévy processes. Let $\{P_u(t,\omega), 0 < u < \infty\}$ be a system of independent Poisson processes with unit intensity, each of which is elementary. Then let us form a compound Poisson process of the following particular type:

$$\begin{aligned} Y_\alpha(t,\omega) &= \int_0^\infty u P_{\mathrm{d}u}(t) \frac{1}{u^{1+\alpha}}\,\mathrm{d}u, \quad 0 < \alpha < 1 \\ &= \int_{-\infty}^\infty u\{P_{\mathrm{d}u}(t) - E(P_{\mathrm{d}u}(t))\} \frac{1}{|u|^{1+\alpha}}\,\mathrm{d}u, \quad 1 \leqslant \alpha < 2. \end{aligned}$$

We see that the process $Y_\alpha(t)$ is a Lévy process and it enjoys a self-similar property, namely $\{Y_\alpha(at)\}$ and $\{a^{1/\alpha} Y_\alpha(t)\}$ have the same probability distribution.

In fact, $Y_\alpha(t)$ is a stable process of exponent α, initiated by P. Lévy. This property can be seen by its probability distribution. The characteristic function of which is

$$\varphi_t(z) = \exp(t\Psi_\alpha(z)), \tag{2.1}$$

where

$$\Psi_\alpha(z) = -|\Gamma(-\alpha)| \exp\left(\operatorname{sign}(z)\frac{i\alpha\pi}{2}|z|^\alpha\right). \tag{2.2}$$

In Section 3, we shall deal with a stable process $Y_{\frac{1}{2}}(t)$, with exponent $1/2$, for the choice of emission random time, the density function of which is given by

$$f(y(t)) = \frac{t}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2y}} y^{-\frac{3}{2}}. \tag{2.3}$$

For $\alpha = 1$, Y_1 is a symmetric stable process. We know precisely that it is a Cauchy process.

2.2. SUBORDINATION

We briefly revise the notion of subordination, due to S. Bochner.

Take an additive process $\{X(t, \omega)\}$ with independent stationary increments and form a new process $\{Z(t, \omega)\}$, such that

$$Z(t, \omega) = X(Y(t), \omega) \tag{2.4}$$

by changing the time variable t to $Y(t)$, where $Y(t)$ is an increasing random function with $Y(0) = 0$.

This process $\{X(Y(t), \omega)\}$ is said to be subordinate to $\{X(t)\}$ using the time $\{Y(t)\}$. It means that the original process is observed only at random times.

In our case, $X(t)$ and the time variable $Y(t)$ are taken to be Brownian motion $B(t)$ and $Y_{\frac{1}{2}}(t)$, respectively.

3. Stochastic Model

As was discussed in Section 1, observation of the emission will be made at a random time which is taken to be a stable process. We choose the stable process with index $\frac{1}{2}$, that is $Y_{\frac{1}{2}}(t)$, by noting that it is the inverse function of the maximum of a Brownian motion which can be considered as a driving force, i.e.

$$P(M(t) \geqslant y) = P(Y_{\frac{1}{2}}(y) \leqslant t), \tag{3.1}$$

where $M(t) = \max_{s \leqslant t} B(s)$.

Remark. This Brownian motion is a different one and is independent of the Brownian motion which is taken as an information source that we mentioned in Section 1.

The reason why we may consider $Y_{\frac{1}{2}}(t)$ is that there is a Brownian motion that attains its maximum value. Such a setup enables us to introduce $M(t)$ and $Y_{\frac{1}{2}}(t)$.

In addition, $Y_{\frac{1}{2}}(t)$ is a compound Poisson process so that each Poisson process is formed by a random waiting time with a probability distribution of the exponential type (note the lack of memory).

For computation of the entropy, we consider our model on a time interval $(0, T)$.

The above discussion yields a stochastic model, with the information source $\{B(t),\ t \in [0, T], Y_{\frac{1}{2}}(t),\ t \in [0, T]\}$, and the observation data $\{B(Y_{\frac{1}{2}}(t)),\ t \in [0, T]\}$.

For the computation, we take only a finite number of points $t_i,\ i = 1, \ldots, n$ in the interval $[0, T]$. Then the information source is taken to be

$$\left\{B(t_i), i = 1, \ldots, n,\ Y_{\frac{1}{2}}(t_i),\ i = 1, \ldots, n\right\}$$

and the observed data will be obtained in the form

$$\left\{B(Y_{\frac{1}{2}}(t_i)),\ i = 1, \ldots, n\right\}.$$

Since we wish to have independent systems, we transform them as follows:

$$\{B(t_i),\ i=1,\dots,n\} \ \to\ \{B(t_k)-B(t_{k-1}),\ k=1,\dots,n\},$$

$$\left\{B(Y_{\frac{1}{2}}(t_i)),\ i=1,\dots,n\right\} \to \left\{B(Y_{\frac{1}{2}}(t_k))-B(Y_{\frac{1}{2}})(t_{k-1}),\ k=1,\dots,n\right\}.$$

Then the observed data will be

$$\left\{B(Y_{\frac{1}{2}}(t_k))-B(Y_{\frac{1}{2}}(t_{k-1})),\ k=1,\dots,n\right\}.$$

Note that, for each case, according to the transformation, the Jacobian is 1.

Fact 1. First we evaluate the entropy for $\{B(t_1),\dots,B(t_n)\}$. Since the entropy for a Gaussian random variable with variance σ^2 is

$$\log\sqrt{2\pi e}+\log\sigma, \tag{3.2}$$

for $B(t)$ we have

$$\log\sqrt{2\pi e}+\log t. \tag{3.3}$$

Consequently, the entropy for $\{B(t_k)-B(t_{k-1}),\ k=1,\dots,n\}$ is

$$H(B)=n\log\sqrt{2\pi e}+\sum_{k=1}^{n}\log(t_k-t_{k-1}). \tag{3.4}$$

Fact 2. The entropy for $Y_{\frac{1}{2}}(t)$ is

$$\log\sqrt{2\pi}+\tfrac{1}{2}-\tfrac{3}{2}\{\log 2+\psi(\tfrac{1}{2})\}+2\log t, \tag{3.5}$$

where ψ is Euler's ψ-function.

So the entropy for $\{B(Y_{\frac{1}{2}}(t_i)),\ i=1,\dots,n\}$ is

$$H(Y)=n\log\sqrt{2\pi}+\frac{n}{2}-\tfrac{3}{2}n\{\log 2+\psi(\tfrac{1}{2})\}+2\sum_{k=1}^{n}\log(t_k-t_{k-1}). \tag{3.6}$$

Fact 3. $B(Y_{\frac{1}{2}}(t))$ has Cauchy distribution and its entropy is

$$\log 4\pi+\log t \tag{3.7}$$

and thus the entropy for $\{B(Y_{\frac{1}{2}}(t_k))-B(Y_{\frac{1}{2}}(t_{k-1})),\ k=1,\dots,n\}$ is

$$H(B(Y))=n\log 4\pi+\sum_{k=1}^{n}\log(t_k-t_{k-1}). \tag{3.8}$$

Then the entropy loss is obtained as

$$n\left\{1+2\sum_{k=1}^{n}\log(t_k-t_{k-1})-\tfrac{3}{2}\Psi(\tfrac{1}{2})-\tfrac{5}{2}\log 2\right\}. \tag{3.9}$$

If we take the time intervals with equal length 1, then the entropy loss is

$$n\{1-\tfrac{3}{2}\Psi(\tfrac{1}{2})-\tfrac{5}{2}\log 2\}. \tag{3.10}$$

THEOREM 3.1. *The entropy loss is propotional to n if we take the observations at times with unit intervals.*

4. Filtering

Following the idea of C. E. Shannon, we may introduce the quantity of the information contained in the stationary Gaussian process $X(t)$. However, the absolute entropy is infinite, so we are only interested in the difference of entropy between the white noise that is taken to be the input of the $X(t)$ and the output $X(t)$ itself. To make an actual comparison, we limit the band, say, to $[-W, W]$.

Let a stationary Gaussian process $X(t)$ have the canonical representation of the form

$$X(t) = \int_{-\infty}^{t} F(t-u)\dot{B}(u)\,\mathrm{d}u. \tag{4.1}$$

Then the spectral representation is

$$X(t) = \int \mathrm{e}^{it\lambda}\hat{F}(\lambda)\dot{Z}(\lambda)\,\mathrm{d}\lambda, \tag{4.2}$$

where

$$\dot{Z}(\lambda) = \frac{1}{\sqrt{2\pi}}\int \mathrm{e}^{it\lambda}\dot{B}(t)\,\mathrm{d}t \quad \text{with } E|\dot{Z}(\lambda)|^2 = \frac{1}{\mathrm{d}\lambda} \text{ formally.} \tag{4.3}$$

The band limited representations are

$$X_W(t) = \int_{-W}^{W} \mathrm{e}^{it\lambda}\hat{F}(\lambda)\dot{Z}(\lambda)\,\mathrm{d}\lambda \tag{4.4}$$

and

$$\dot{B}_W(t) = \int_{-W}^{W} \mathrm{e}^{-it\lambda}\dot{Z}(\lambda)\,\mathrm{d}\lambda, \tag{4.5}$$

respectively. It is known that $\dot{Z}(\lambda)$ can be represented as

$$\dot{Z}(\lambda) = \frac{\mathrm{sgn}(\lambda)}{\sqrt{2}}\{\dot{Z}_1(\lambda) + i\dot{Z}_2(\lambda)\}; \quad \dot{Z}_1(\lambda), \dot{Z}_2(\lambda)\text{: real.}$$

Here we note that $\dot{Z}(\lambda) = -\overline{\dot{Z}(-\lambda)}$.

Based on white noise, we compute its relative entropy or entropy loss by approximation. Since white noise can be expressed as (4.5), we approximate it by $\sum_{i=1}^{N} \mathrm{e}^{it\lambda_k}\Delta_k Z$. Similarly, we can approximate $X(t)$ by

$$\sum_{i=1}^{N} \mathrm{e}^{it\lambda_k} G(\lambda_k)\Delta_k Z.$$

To compute the entropy of them we see the joint distribution of $\{\Delta_k Z\}$ and that of $\{G(\lambda_k)\Delta_k Z\}$.

Then we are given the information loss

$$\sum_K \log |G(\lambda_k)|^2 \Delta_k$$

by noting $(\Delta_k B_1, \Delta_k B_2)$ is a two-dimensional Gaussian random variable.

THEOREM 4.1. *Based on a band-limited white noise, the entropy loss of the output is given by*

$$\int_{-W}^{W} \log |G(\lambda)|^2 \, d\lambda. \tag{4.6}$$

We now consider a particular case of averaging. The integral operator G_a, $a > 0$ is defined by

$$(G_a f(t)) = a \int_{-\infty}^{t} e^{-a(t-u)} f(u) \, du. \tag{4.7}$$

It is

(i) causal (i.e. operating only on the past values),
(ii) averaging $(G_a 1)(t) = 1$,
(iii) stationarity holds, as is shown below.

Apply G_a to a stationary process, then we see some changing of the characteristic function

$$F(t) \to \hat{F}(\lambda)\hat{G}_a(\lambda), \quad \hat{G}_a(\lambda) = \frac{a}{a + i\lambda}, \tag{4.8}$$

where F is the cannonical kernel.

PROPOSITION 4.1. *The information loss is given by*

$$\int_{-W}^{W} \log \frac{a^2 + \lambda^2}{\lambda^2} \, d\lambda. \tag{4.9}$$

We recall that the transmission function of a stationary N-ple Markov Gaussian process $X(t)$ is expressed in the form

$$\hat{F}(\lambda) = \frac{Q(i\lambda)}{P(i\lambda)}, \tag{4.10}$$

where $\hat{F}$ is the Fourier transform of the canonical kernel F and where P and Q are polynomials with degree $Q <$ degree $P = N$. It is to be noted that all the roots of P and Q are in the lower half of the complex plane.

If the $X(t)$ passes through a filter with characteristic $\hat{G}(\lambda)$, then that of the output is given by $\hat{F} \cdot \hat{G}$.

PROPOSITION 4.2. *The information loss is given by*

$$\int_{-W}^{W} \log \frac{|Q|^2}{|P|^2}\, \mathrm{d}\lambda. \tag{4.11}$$

References

1. Bochner, S.: *Harmonic Analysis and the Theory of Probability*, Univ. of California Press, 1955.
2. Hida, T.: *Stationary Stochastic Processes*, Math. Notes, Princeton Univ. Press, 1970.
3. Hida, T.: *Brownian Motion*, Springer-Verlag, New York, 1980.
4. Hida, T. and Hitsuda, M.: *Gaussian Processes*, Transl. Math. Monogr. 12, Amer. Math. Soc., Providence, 1993.
5. Lévy, P.: *Théorie de l'addition des variables aléatoires*, Gauthier-Villars, Paris, 1937.
6. Lévy, P.: Sur certains processus stochastiques homogenes, *Compositio Math.* **7** (1939), 283–339.
7. Oda, M.: Fluctuation in astrophysical phenomena, In: T. Hida (ed.), *Proc. IIAS Workshop, Mathematical Approach to Fluctuations*, Vol. I, 1992, pp. 115–137.
8. Htay, Win Win: Optimalities for random functions: Lee–Wiener's network and non-canonical representations of stationary Gaussian processes, *Nagoya Math. J.* **149** (1998), 9–17.