Kumamoto J. Sci. (Math.)
Vol, 17. 5~8 March (1986)

# A CHARACTERIZATION OF AUTOREGRESSIVE PROCESSES BY THE DISTRIBUTION WITH MAXIMUM ENTROPY

Isao HASHIMOTO

(Received Oct. 31, 1985)

## 1. Introduction

It was well known that $X$ is a random variable with mean $E[X]=0$ and given variance $\sigma^2$, then the entropy is a maximum if and only if $X$ is a normally distributed random variable. This fact is generalized to the form as in section 2 (e. g. see Mardia et al. [2], p. 46). The following also is well known as Burg's result [1]: to fit a stochastic process by the maximum entropy method is to fit an autoregressive process to the process (e. g. see Hino [3], pp. 86-89 and Ihara [4], pp. 65-83) In these books the analysis of spectrum is used in order to prove Burg's result. In this paper the proof of this result is given by using the theorem stated in section 2 instead of spectral analysis.

## 2. Maximum entropy property

The following theorem is well known (see Mardia et al. [2], p. 46).

THEOREM 1. *Let us consider the set of all probability densiy functions* $g(x)$ *with support* $S$ *satisfying the constraints*

$$E[b_i(X)]=c_i, \quad i=1, 2, \cdots, p \tag{1}$$

*where the* $c_i$ *are given constants. Then the entropy* $E[-\log g(X)]$ *is maximized by the probability density*

$$f(x; \vartheta)=exp\ [a_0(\vartheta)+b_0(x)+\sum_{i=1}^{p}\theta_i b_i(x)],\ x\in S, \tag{2}$$

*provided there exists* $\vartheta=(\theta_1, \cdots, \theta_p)$ *for which the constraints* (1) *are satisfied. If such a* $\vartheta$ *exists, the maximum is unique.*

## 3. Burg's result

The following statement of Burg's result follows from Ihara [4], p, 82.

THEOREM 2. *Let* $\{X_t;\ t=0,\ \pm 1,\ \pm 2,\ \cdots\}$ *be a real valued weakly stationary process with mean* $E[X_t]=0$ *and covariance function*

$$\gamma(i)=E[X_t X_{t+i}].$$

If

$$\gamma(i)=c_i,\quad i=0,\ 1,\ \cdots,\ p \tag{3}$$

*where the* $c_i$ *are given constants satisfying* $\sum_{j,k=0}^{p} d_j d_k c_{j-k}>0$ *fo rall nonzero vectors* $(d_0, \cdots, d_p)$ *and* $c_i=c_{-i}(\imath=1,\ 2,\ \cdots,\ p)$, *then the process with maximun entropy satisfying the constraints* (3) *is an autoregressive process of order* $p$.

## 4. Probability density of $AR$ (1)

An autoregressive process of order $p$ will be abbreviated to an $AR(p)$ process. Let $\{X_t:\ t=0,\ \pm 1,\ \cdots\}$ be a process defined by the difference equation

$$X_t=-aX_{t-1}+\varepsilon_t \tag{4}$$

where $|a|<1$ and $\{\varepsilon_t;\ t=0,\ \pm 1,\ \cdots\}$ is a sequence of independently and identically distributed random variables having a common normal distribution $N(0,\ \sigma^2)$.

Suppose that $x_1,\ x_2,\ \cdots,\ x_n$ are consecutive observations from (4). Since

$$E[X_t]=0,\quad t=0,\ 1,\ \cdots,\ n,\ \text{and}$$
$$E[X_t X_{t+i}]=\frac{(-a)^i\sigma^2}{1-a^2},\ t=1,\ 2,\ \cdots,\ n;\ i=0,\ 1,\ \cdots\ t,$$

the probability density of $x=(x_1,\ \cdots,\ x_n)$ is given by

$$f(x;\ a)=(2\pi)^{-\frac{n}{2}}\ |\textstyle\sum_a|^{-\frac{1}{2}}\ exp\left\{-\frac{1}{2\sigma^2}\ x\ \sum_n^{-1}\ x'\right\}$$

where

$$\Sigma_a = \frac{\sigma^2}{1-a^2}\begin{pmatrix} 1 & -a & (-a)^2 & \cdots & (-a)^{n-1} \\ -a & 1 & -a & \cdots & (-a)^{n-2} \\ & & \cdots\cdots & & \\ (-a)^{n-1} & (-a)^{n-2} & (-a)^{n-3} & \cdots & 1 \end{pmatrix}.$$

It follows from simple computations that

$$|\Sigma_a| = \frac{(\sigma^2)^n}{1-a^2}, \text{and}$$

$$\Sigma_a^{-1} = \frac{1}{\sigma^2}\begin{pmatrix} 1 & a & 0 & 0 & \cdots & 0 \\ a & 1+a^2 & a & 0 & \cdots & 0 \\ & & \cdots\cdots & & & \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Thus $f(x;\, a)$ can alternatively be written by

(5)
$$f(x;a) = (2\pi\sigma^2)^{-\frac{1}{2}} \sqrt{1-a^2}\, exp\left[\frac{1}{2\sigma^2}\{x_1^2 + x_n^2 + (1+a^2)\sum_{t=2}^{n-1} x_t^2 + 2a\sum_{t=2}^{n} x_{t-1}x_t\}\right].$$

**5. A simple proof of Theorem 2**

Let $X=(X_1,\, X_2,\, \cdots,\, X_n)$ be a random variable having the probability density $f(x)$ with $f(x)>0$ for $x \in R^n$. Let, further, $f(x)$ satisfy the constraints

(6)
$$\begin{aligned} &E[X_t]=0, && t=1,\,2,\,\cdots,\,n \\ &E[X_t^2]=c_0, && t=1,\,2,\,\cdots,\,n \\ &E[X_tX_{t+1}]=c_1, && t=1,\,2,\,\cdots,\,n-1 \end{aligned}$$

where $c_0 \neq 0$ and $c_0 \neq c_1$.

Then we have from Theorem 1 that the maximum entropy is attained by the distribution with probability density of the form

$$f(x:\,\vartheta) = exp\,[a_0(\vartheta) + b_0(x) + (\theta_1x_1 + \cdots + \theta_nx_n) + (\theta_{n+1}x_1^2 + \cdots + \theta_{2n}x_n^2) + (\theta_{2n+1}x_1x_2 + \cdots + \theta_{3n-1}x_{n-1}x_n)].$$

Comparing to (5), we choose $\vartheta=(\theta_1\, \cdots,\, \theta_{3n-1})$ such that

$$exp\ a_0(\vartheta) = (2\pi)^{-\frac{n}{2}} \left(\frac{c_0^2 - c_1^2}{c_0}\right)^{-\frac{n}{2}+1}$$

$$b_0(x) \equiv 0$$

$$\theta_1 = \cdots = \theta_n = 0$$

$$\theta_{n+1} = -\frac{c_0}{2(c_0^2 - c_1^2)},\quad \theta_{n+2} = \cdots = \theta_{2n-1} = -\frac{1}{2c_0},\ \theta_{2n} = -\frac{c_0}{2(c_0^2 - c_1^2)}$$

$$\theta_{2n+1} = \cdots = \theta_{3n-1} = -\frac{c_1}{c_0^2 - c_1^2}.$$

For $AR(p)$, we compute the constraints corresponding to (6) and we may apply Theorem 1 as in the above discussion.

## REFERENCES


[1] Burg, J. P. (1967). Maximum entropy spectral analysis. *Th 37th Ann. Mtg., Soc. of Explor.Geophys.*

[2] Mardia, K. V., Kent, J. T. and Bibby, J. M. (1979). *Multivariate Analysis*. Academic Press.

[3] Hino, M. (1979). *Spectral Analysis* (in Japanese). Asakura.

[4] Ihara, S. (1984). *Stochastic Processes and Entropy* (in Japanese). Iwanami.



Institute of Mathematics
Faculty of Engineering
Kumamoto University