The Annals of Statistics
1990, Vol. 18, No. 4, 1774–1783

MAXIMUM STANDARDIZED CUMULANT DECONVOLUTION OF NON-GAUSSIAN LINEAR PROCESSES

By Qiansheng Cheng

Peking University

A linear process is represented as a driving white noise convolved with a system response sequence. The concept of natural peakedness of a system response sequence is defined and its properties are investigated. Utilizing natural peakedness, the convergence theory of maximum standardized cumulant deconvolution is established and the uniqueness theorem of non-Gaussian linear process representations is proved. In addition, autoregressive models on a countable abelian group are defined and the relation between cumulant deconvolution and autoregressive models is given.

1. Introduction. Let G represent a countable abelian group, let $w = \{w_t\}_{t \in G}$ be a square-summable sequence and let $u = \{u_t\}_{t \in G}$ be an independent and identically distributed random series with $Eu_t = 0$, $Eu_t^2 = \sigma^2$ and $E|u_t|^m < \infty$ for some $m > 2$.

$$x_t = (w * u)_t = \sum_{s \in G} u_s w_{t-s} \tag{1.1}$$

is called a linear process; u and w are called the driving noise and the system response sequence, respectively.

Throughout this paper we shall make the following assumptions:

$$0 < \sum_{t \in G} |w_t|^2 < \infty \tag{1.2}$$

and the Fourier transform of w,

$$w(\gamma) = \sum_{s \in G} w_s \gamma(-s), \qquad \gamma \in \Gamma,$$

satisfies

$$w(\gamma) \neq 0, \qquad d\gamma \text{ a.s.}, \tag{1.3}$$

where Γ is the dual group of G which consists of all complex functions $\gamma(t)$ on G satisfying $|\gamma(t)| = 1$ and $\gamma(s + t) = \gamma(s)\gamma(t)$, $s, t \in G$; $d\gamma$ denotes the Haar measure on the group Γ [see Rudin (1962)].

We note that when G is the set of integers $\mathbb{Z}$, x is a linear time series, and when G is the set $\mathbb{Z}^2$ of pairs of integers, x is a linear random field.

The objective of linear process decomposition is to estimate the driving noise (deconvolution) and to estimate the system response sequence (system identification) from x_t. In this paper we shall study a kind of deconvolution which is called maximum standardized cumulant deconvolution.

Received February 1989; revised November 1989.

AMS 1980 *subject classifications*. Primary 62M10; secondary 60G10, 62M15.

Key words and phrases. Non-Gaussian linear processes, maximum standardized cumulant deconvolution.

The m-th cumulant of the random variable x_t is defined by

$$c_m(x_t) = \left(-i\frac{d}{dx}\right)^m \log(Ee^{isx_t})\Big|_{s=0}.$$

Set

$$\hat{x}_t = \frac{x_t - Ex_t}{(Ex_t^2)^{1/2}}.$$

The standardized m-th cumulant is defined by

$$k_m(x_t) = c_m(\hat{x}_t) = \frac{c_m(x_t)}{(c_2(x_t))^{m/2}}.$$

For x_t satisfying (1.1),

$$k_m(x_t) = k_m(u_t)\frac{\Sigma_t(w_t)^m}{(\Sigma_t w_t^2)^{m/2}}, \qquad m > 2. \tag{1.4}$$

Here we write Σ_t in place of $\Sigma_{t\in G}$. This result can be found in Granger (1976).

Now we define maximum standardized cumulant deconvolution operator.

Let $S_1 \subset S_2 \subset \cdots \subset S_n \subset \cdots$ denote an increasing sequence of finite subsets of G with the property that for every finite subset F of G, there exists a positive integer n_0 and a $t_0 \in G$ such that $F \subseteq t_0 + S_{n_0}$. Because G is countable, such sequences always exist. For instance, we can take $S_n = \{0, 1, \dots, n\}$ when $G = \mathbb{Z}$ and $S_n = \{(i, j)\colon 0 \le i, j \le n\}$ when $G = \mathbb{Z}^2$.

We denote by $g^{(n)} = \{g_t^{(n)}\}$ a sequence satisfying

$$g_t^{(n)} = 0, \qquad t \notin S_n. \tag{1.5}$$

$h^{(n)}$ is (not necessarily uniquely) defined as the maximum standardized m-th cumulant deconvolution operator of x if it satisfies

$$\left|k_m((h^{(n)} * x)_t)\right| = \max_{g^{(n)}}\left|k_m((g^{(n)} * x)_t)\right|. \tag{1.6}$$

We briefly explain why $h^{(n)}$ exists. By the formulas (1.4), (2.1) and (2.14), for any nonzero constant α, the m-th standardized cumulants of $g^{(n)}$ and $\alpha g^{(n)}$ have the same absolute value. So we can assume that $g^{(n)}$ satisfies

$$\sum_{t\in S_n}\left|g_t^{(n)}\right|^2 = 1.$$

Then $|k_m((g^{(n)} * x)_t)|$ is a continuous function with respect to $g_t^{(n)}$, $t \in S_n$ (note that S_n is a finite subset). Hence, there exists at least one $h^{(n)}$ such that (1.6) holds. Of course, it is possible that $h^{(n)}$ is not unique. We shall pay more attention to the limit property of such a deconvolution operator. When $m = 4$, by (1.4), we have

$$k_4(x_t) = k_4(u_t)\frac{\Sigma_t w_t^4}{\left(\Sigma_t w_t^2\right)^2}.$$

Then (1.6) becomes

$$k_4\big((h^{(n)} * x)_t\big) = \begin{cases} \max\limits_{g^{(n)}} k_4\big((g^{(n)} * x)_t\big), & \text{when } c_4(x_t) > 0, \\ \min\limits_{g^{(n)}} k_4\big((g^{(n)} * x)_t\big), & \text{when } c_4(x_t) < 0. \end{cases}$$

In this case, a maximum standardized cumulant deconvolution operator is just a kurtosis deconvolution operator [see Cheng (1988)]. In fact, minimum entropy deconvolution is a special case of the maximum standardized cumulant deconvolution [see Cheng (1988) and Wiggins (1978)].

The objective of this paper is to establish the convergence theory of maximum standardized cumulant deconvolution and to prove the uniqueness theorem of non-Gaussian linear process representations.

In Section 2 we present the concept of natural peakedness of a system response sequence and study its properties. Section 3 proves the uniqueness theorem of non-Gaussian linear processes. In Section 4 we give the convergence theorem of maximum standardized cumulant deconvolution; in addition, we define autoregressive models on a countable abelian group and discuss the relation between cumulant deconvolution and autoregressive models.

2. Natural peakedness of a system response sequence. We define the natural peakedness of system response sequence w,

$$q(w) = \left|\frac{\Sigma_t (w_t)^m}{\left(\Sigma_t w_t^2\right)^{m/2}}\right|, \qquad m > 2. \tag{2.1}$$

Let

$$\|w\|_2 = \left(\sum_t w_t^2\right)^{1/2} \tag{2.2}$$

$$\|w\|_m = \left(\sum_t |w_t|^m\right)^{1/m}, \qquad m > 2. \tag{2.3}$$

In order to derive the properties of natural peakedness, we introduce the absolute peakedness, defined as

$$p(w) = \left(\frac{\|w\|_m}{\|w\|_2}\right)^m. \tag{2.4}$$

LEMMA 2.1. *For any constant $a \neq 0$ and any $t_0 \in G$,*

$$p(w) = p(aw) = p(\delta^{(t_0)} * w), \tag{2.5}$$

where $\delta^{(t_0)} = \{\delta_t^{(t_0)}\}$, $\delta_t^{(t_0)} = 1$ for $t = t_0$ and zero for $t \neq t_0$.

The proof of Lemma 2.1 is immediate.

LEMMA 2.2.

$$0 < p(w) \le 1. \tag{2.6}$$

If $p(w) = 1$, *then* $w = a\delta^{(t_0)}$, *where* a *is a nonzero constant and* t_0 *is an element of* G.

PROOF. Set

$$\tilde{w} = \frac{1}{w_{t_0}} w,$$

where $|w_{t_0}| = \max_{t \in G} |w_t|$. From Lemma 2.1, we have

$$p(w) = p(\tilde{w}).$$

Since

$$|\tilde{w}_t| \le 1, \qquad |\tilde{w}_{t_0}| = 1, \qquad \|\tilde{w}\|_m^m \le \|\tilde{w}\|_2^2,$$

we get

$$0 < p(w_t) \le \|\tilde{w}\|_2^{-(m-2)} \le 1.$$

If $p(w) = 1$, then $\|\tilde{w}\|_2 = 1$. This means $\tilde{w}_t = 0$ for $t \ne t_0$. Thus, $\tilde{w} = \delta^{(t_0)}$ and $w = w_{t_0}\delta^{(t_0)}$. □

LEMMA 2.3. *Let* $w^{(n)}$ *and* w *be system response sequences and let*

$$\|w^{(n)} - w\|_2 \to 0, \qquad n \to \infty. \tag{2.7}$$

Then

$$p(w^{(n)}) \to p(w), \qquad n \to \infty. \tag{2.8}$$

PROOF. By the property of norms,

$$\big| \|w^{(n)}\|_2 - \|w\|_2 \big| \le \|w^{(n)} - w\|_2 \to 0, \qquad n \to \infty.$$

It follows from Lemma 2.1 that

$$\|w^{(n)} - w\|_m \le \|w^{(n)} - w\|_2;$$

then

$$\big| \|w^{(n)}\|_m - \|w\|_m \big| \le \|w^{(n)} - w\|_m \to 0, \qquad n \to \infty.$$

Hence

$$p(w^{(n)}) = \frac{\|w^{(n)}\|_m^m}{\|w\|_2^m} \to \frac{\|w\|_m^m}{\|w\|_2^m} = p(w), \qquad n \to \infty. \qquad \square$$

LEMMA 2.4. *Let* $w^{(n)}$ *be system response sequences. In order that*

$$p(w^{(n)}) \to 1, \qquad n \to \infty, \tag{2.9}$$

it is necessary and sufficient that

$$\|\tilde{w}^{(n)} - \delta^{(t_n)}\|_2 \to 0, \qquad n \to \infty, \tag{2.10}$$

where

$$\tilde{w}_t^{(n)} = \frac{1}{w_{t_n}^{(n)}} w_t^{(n)}, \qquad \left|w_{t_n}^{(n)}\right| = \max_{t \in G} \left|w_t^{(n)}\right|. \tag{2.11}$$

PROOF. Necessity: It follows from the proof of Lemma 2.2 that

$$p(w^{(n)}) = p(\tilde{w}^{(n)}) \le \|\tilde{w}^{(n)}\|^{-(m-2)} \le 1.$$

Since (2.9) holds, we obtain

$$\|\tilde{w}^{(n)}\|_2^2 = 1 + \sum_{t \ne t_n} |\tilde{w}_t^{(n)}|^2 \to 1, \qquad n \to \infty.$$

Therefore,

$$\|\tilde{w}^{(n)} - \delta^{(t_n)}\|_2^2 = \sum_{t \ne t_n} |\tilde{w}_t^{(n)}|^2 \to 0, \qquad n \to \infty. \tag{2.12}$$

Sufficiency: Lemma 2.3 yields this immediately. □

We now discuss the property of natural peakedness. The following lemma is evident.

LEMMA 2.5.

$$0 \le q(w) \le p(w) \le 1, \tag{2.13}$$

$$q(aw) = q(w) = q(\delta^{(t_0)} * w), \tag{2.14}$$

where a is a nonzero constant and t_0 is an element of G.

LEMMA 2.6. *Let $w^{(n)}$ be system response sequences. In order that*

$$q(w^{(n)}) \to 1, \qquad n \to \infty, \tag{2.15}$$

it is necessary and sufficient that

$$p(w^{(n)}) \to 1, \qquad n \to \infty. \tag{2.16}$$

PROOF. From (2.13), necessity holds. Sufficiency follows from (2.1), (2.14) and (2.12). □

The following theorem follows immediately from Lemmas 2.5 and 2.6.

THEOREM 2.1. *Let $w^{(n)}$ be system response sequences. In order that*

$$q(w^{(n)}) \to 1, \qquad n \to \infty, \tag{2.17}$$

it is necessary and sufficient that

$$\|\tilde{w}^{(n)} - \delta^{(t_n)}\|_2 \to 0, \qquad n \to \infty, \tag{2.18}$$

where $\tilde{w}$ *and* t_n *are defined in* (2.11).

3. The uniqueness theorem of non-Gaussian linear processes. In this section we prove the uniqueness theorem of non-Gaussian linear processes.

Let H denote the Hilbert space of all random variables with finite variances and with inner product defined by covariance. Let x, u, and u' be random processes on G, and let H_x, H_u and $H_{u'}$ denote the linear closed subspaces of H generated by x_t, u_t and u'_t, $t \in G$, respectively.

Lemma 3.1. *Let* x_t *satisfy* (1.1). *Then*

$$H_x = H_u.$$

Proof. It is obvious that $H_x \subset H_u$. So it suffices to show that any $y = \sum_{s \in G} v_s u_s$ in H_u with the property

$$\sigma^{-2} E x_t y = \sum_{s \in G} v_{t-s} v_s = (w * v)_t = 0, \qquad t \in G, \tag{3.1}$$

must be 0. We denote by $W(\gamma)$ and $V(\gamma)$ the Fourier transforms or the Plancherel transforms of w and v, respectively.

As the Plancherel transform, we have

$$(w * v)_t = \int_\Gamma \gamma(t) W(\gamma) V(\gamma)\, d\gamma. \tag{3.2}$$

(3.1) and (3.2) yield

$$W(\gamma)V(\gamma) = 0, \qquad (d\gamma \text{ a.s.})$$

[see Rudin (1962) pages 26 and 27]. Applying (1.3), it follows that $V(\gamma) = 0$, $d\gamma$ a.s. Hence $v_s = 0$, for all $s \in G$, and $y = 0$. □

Theorem 3.1 (The uniqueness theorem). *Let*

$$x_t = (w * u)_t = (w' * u')_t, \qquad t \in G, \tag{3.3}$$

where $\{u_t\}$ *and* $\{u'_t\}$ *are i.i.d. and* w *and* w' *are system response sequences satisfying* (1.2) *and* (1.3). *If* $c_m(x_t) \neq 0$ *for some* $m > 2$, *then*

$$u'_t = a u_{t-t_0}, \qquad w'_t = \frac{1}{a} w_{t+t_0}, \tag{3.4}$$

where a *is a nonzero constant and* t_0 *is an element of* G.

Proof. By Lemma 3.1, we have

$$H_u = H_x = H_{u'}.$$

Hence, there exist sequences $c = \{c_t\}$ and $d = \{d_t\}$ such that

$$u' = c * u, \qquad u = d * u'. \tag{3.5}$$

According to the relation (1.4) and the definition (2.1), we have

$$|k_m(u'_t)| = |k_m(u_t)|q(c), \qquad |k_m(u_t)| = |k_m(u'_t)|q(d).$$

Thus,

$$q(c)q(d) = 1. \tag{3.6}$$

It follows from (3.6) and (2.13) that

$$q(c) = p(c) = 1.$$

Applying Lemma 2.2 yields

$$c = a\delta^{(t_0)}. \tag{3.7}$$

(3.7), (3.5) and (3.3) imply (3.4). □

Theorem 3.1 shows that if we ignore the scale and shift, the representation (1.1) of non-Gaussian linear processes is essentially unique.

COROLLARY 3.1. *Let $\{u_t\}$ and $\{x_t\}$ be i.i.d and $x = w * u$. If $k_m(x_t) \neq 0$ for some $m > 2$, then*

$$w = a\delta^{(t_0)},$$

where a is a nonzero constant and t_0 is an element of G.

The proof of the corollary is immediate.

Donoho (1981) discusses the problem of uniqueness, using the concept of a partial order which describes the relation between probability distributions of random variables. Rosenblatt (1985, 1986) studies the uniqueness under the additional assumption that $\Sigma |t|\,|w_t| < \infty$ (when $G = \mathbb{Z}$). When G is any countable abelian group, we cannot make the additional assumption. Under the condition that x_t has moments of all orders, Findley (1986) gives a different proof of the uniqueness result in the case $G = \mathbb{Z}$. He seems to have payed more attention to the property of Gaussian distributions (the m-th cumulant is zero, for all $m \geq 3$) and overlooked the fact that his proof only needs the condition, as does ours. At any rate, from the proof of Theorem 3.1, we see that natural peakedness is a simple and powerful instrument.

4. Maximum standardized cumulant deconvolution. We have defined the maximum standardized m-th cumulant deconvolution operator in Section 1. Now we give the convergence theorem of maximum standardized cumulant deconvolution.

THEOREM 4.1. *Let $c_m(x_t) \neq 0$, for some $m > 2$, and let $h^{(n)}$ be maximum standardized m-th cumulant deconvolution operators of x. Let t_n be an element of G such that $|(h^{(n)} * w)_{t_n}| = \max_{t \in G}|(h^{(n)} * w)_t|$, and set $a_n = 1/(h^{(n)} * w)_{t_n}$.*

Then

$$\lim_{n\to\infty} E\left|a_n(\delta^{(-t_n)} * h^{(n)} * x)_t - u_t\right|^2 = 0. \tag{4.1}$$

PROOF. From (1.4), (1.6) and (2.1), $h^{(n)}$ satisfies

$$q(h^{(n)} * w) = \max_{g(n)} q(g^{(n)} * w). \tag{4.2}$$

It follows from Lemma 3.1 that there exist $l_n \in G$ and sequences $\tilde{g}^{(n)} = \{\tilde{g}_t^{(n)}\}_{t\in G}$ satisfying $\tilde{g}_t^{(n)} = 0$ if $t \notin l_n + S_n$, such that

$$\begin{aligned} E\left|(\tilde{g}^{(n)} * x)_t - u_t\right|^2 &= E\left|(\tilde{g}^{(n)} * w * u)_t - (\delta^{(0)} * u)_t\right|^2 \\ Eu_t^2\|\tilde{g}^{(n)} * w - \delta^{(0)}\|_2^2 &\to 0, \qquad n \to \infty. \end{aligned} \tag{4.3}$$

Applying Lemma 2.3, we have

$$\lim_{n\to\infty} p(\tilde{g}^{(n)} * w) = 1.$$

By Lemma 2.6,

$$\lim_{n\to\infty} q(\tilde{g}^{(n)} * w) = 1.$$

Note that $\delta^{(l_n)} * \tilde{g}^{(n)}$ satisfies $(\delta^{(l_n)} * \tilde{g}^{(n)})_t = 0$ if $t \notin G$. From (2.13) and (4.2) we get

$$q(\tilde{g}^{(n)} * w) = q(\delta^{(l_n)} * \tilde{g}^{(n)} * w) \le q(h^{(n)} * w) \le 1.$$

Therefore,

$$\lim_{n\to\infty} q(h^{(n)} * w) = 1.$$

From Theorem 2.1,

$$\lim_{n\to\infty} \left\|a_n(\delta^{(-t_n)} * h^{(n)} * w) - \delta^{(0)}\right\|_2 = 0. \tag{4.4}$$

(4.4) and (4.3) yield (4.1). □

Theorem 4.1 shows that when some m-th cumulant of the process is not equal to zero, we can extract the driving noise and the system response sequence only from a non-Gaussian linear process.

We now turn to the autoregressive model on G.

A linear process x satisfying (1.1) is regarded as obeying an autoregressive model on G if there exists a finite set $F \subset G$ and a sequence $a = \{a_t\}$ satisfying $a_t = 0$ if $t \notin F$ such that $a * w = \delta^{(0)}$. Such an x is denoted by $AR(F)$; $a = \{a_t\}$ is said to be the sequence of autoregressive coefficients of x.

We take an integer n_0 and $t_0 \in G$ such that $F \subseteq t_0 + S_{n_0}$. The following theorem gives the relation between cumulant deconvolution operator and autoregressive coefficients.

THEOREM 4.2. *Let x be $AR(F)$, let $a = \{a_t\}$ be the autoregressive coefficients of x and let t_0 and n_0 satisfy $F \subseteq t_0 + S_{n_0}$. Let $c_m(x) \neq 0$ for some $m > 2$ and let $h^{(n_0)}$ be the maximum standardized m-th cumulant deconvolution operator. Then*

$$h^{(n_0)} = \lambda \delta^{(t_1)} * a, \tag{4.5}$$

where t_1 is an element of G and λ is a nonzero constant.

PROOF. From the definition of $AR(F)$, we know that $a * w = \delta^{(0)}$. Then,

$$q(a * w) = 1.$$

We note that $\delta^{(t_0)} * a$ satisfies $(\delta^{(t_0)} * a)_t = 0$ if $t \notin S_{n_0}$. By (4.2) and (2.13),

$$q(a * w) = q(\delta^{(t_0)} * a * w) \leq q(h^{(n_0)} * w) \leq 1.$$

Hence,

$$q(h^{(n_0)} * w) = 1.$$

It follows from (2.13) and Lemma 2.2 that

$$h^{(n_0)} * w = \lambda \delta^{(t_1)}.$$

Since

$$\delta^{(t_1)} = \delta^{(t_1)} * \delta^{(0)} = \delta^{(t_1)} * a * w,$$

we have

$$h^{(n_0)} * w = \lambda \delta^{(t_1)} * a * w. \tag{4.6}$$

The condition (1.3) and the relation (4.6) yield (4.5). □

Theorem 4.2 shows that maximum standardized cumulant deconvolution operators for autoregressive processes are just rescaled and shifted versions of the autoregressive coefficients.

Finally, we point out that maximum standardized cumulant deconvolution is a nonlinear problem. It is possible to find a good algorithm by combining maximum standardized cumulant deconvolution and an autoregressive model and choosing a suitable initialization procedure.

Acknowledgments. The author is grateful to a referee for his valuable comments and constructive suggestions. The results in the original manuscript were proved for time series ($G = \mathbb{Z}$) and random fields ($G = \mathbb{Z}^2$). The referee pointed out that only slight modification to the original proofs is needed to cover the situation in which G is any countable abelian group and presented the version of the deconvolution result and the proof of Lemma 3.1.

REFERENCES

CHENG, Q. (1988). Convolution decomposition of 1-D and 2-D linear stationary signals. *Proc. IEEE Internat. Conf. Acoustics, Speech and Signal Processing* **2** 898–899.

DONOHO, D. L. (1981). On minimum entropy deconvolution. In *Applied Time Series Analysis II* (D. F. Findley, ed.) 565–608. Academic, New York.

FINDLEY, D. (1986). The uniqueness of moving average representation with independent and identically distributed random variables for non-Gaussian stationary time series. *Biometrika* **73** 520–521.

GRANGER, C. W. J. (1976). Tendency towards normality of linear combinations of random variables. *Metrika* **23** 237–238.

ROSENBLATT, M. (1985). *Stationary Sequences and Random Fields*. Birkhäuser, Basel.

ROSENBLATT, M. (1986). Higher order spectral methods and deconvolution. Third Acoustics, Speech and Signal Processing workshop on spectrum estimation and modeling. Nov. 17–18, Boston.

RUDIN, W. (1962). *Fourier Analysis on Groups*. Interscience, New York.

WIGGINS, R. A. (1978). Minimum entropy deconvolution. *Geoexploration* **16** 21–35.

DEPARTMENT OF MATHEMATICS
AND INSTITUTE OF MATHEMATICS
PEKING UNIVERSITY
BEIJING
CHINA