*The Annals of Statistics*
1974, Vol. 2, No. 2, 353-358

# ESTIMATING THE KERNELS OF NONLINEAR ORTHOGONAL POLYNOMIAL FUNCTIONALS[1]

BY BENJAMIN KIMELFELD

*University of Haifa*

Let $(X(t), Y(t))$ be a complex vector process stationary of order $k$ for any $k$, $k = 1, 2, \cdots$, such that $Y(t)$ is expressed as a polynomial functional of degree 2 operating on $X(t)$. Then $Y(t)$ can be rewritten as a sum of orthogonal projections $G_j(K_j, Y(t))$, $j = 0, 1, 2$. It is shown that there is a set of functionals which approximate in mean square the projection $G_2(K_2, Y(t))$. Moreover, it is possible to determine the kernels associated with these functionals.

**1. Introduction.** Let $X(t)$, $-\infty < t < \infty$, be a zero mean continuous parameter complex stochastic process stationary of order $k$ for any $k$, $k = 1, 2, \cdots$. Second order stationarity implies that $X(t)$ admits the spectral representation with respect to a process of orthogonal increments $Z_X(\lambda)$, $-\infty < \lambda < \infty$, (e.g., see [3], page 527). By a polynomial functional of degree $n$ we mean a functional of the form:

$$(1.1) \qquad Y(t) = \int \cdots \int \exp[it(\lambda_1 + \cdots + \lambda_n)]H_n(\lambda_1, \cdots, \lambda_n)\, dZ_X(\lambda_1) \cdots dZ_X(\lambda_n) + \cdots + \int e^{it\lambda_1}H_1(\lambda_1)\, dZ_X(\lambda_1) + H_0\,, \qquad -\infty < t < \infty\,,^2$$

where $H_0$ is a constant, and $H_j$, $j = 1, \cdots, n$, are complex continuous and bounded functions (kernels). Integrals of this kind are discussed in [12], [13]. Let $\mathscr{H}$ be the subsequence of $L_2(\Omega, B_X, P)$, where $B_X$ is the $\sigma$-field generated by $X(t)$, which consists of all square integrable polynomial functionals of degree $n$, $n = 0, 1, 2$. In $\mathscr{H}$ we define an operator $T^t$ by $T^tY(S) = Y(S + t)$; see [11]. Every two polynomials $x$, $y$ are said to be orthogonal if $Ex\bar{y} = 0$.

Let $\mathscr{k}_0$ be the subspace of $\mathscr{H}$ of all constants, $\mathscr{k}_1$ be the subspace of all linear functionals (degree 1) which are orthogonal to every constant, and let $\mathscr{k}_2$ be the subspace of all second degree functionals which are orthogonal to every constant and every linear functional. Then $\mathscr{H}$ can be expressed as the direct sum

$$(1.2) \qquad \bigoplus_{j=0}^{2} \mathscr{k}_j = \mathscr{H}\,.$$

(See [4, page 109], [7], [15].) Now let $Y(t)$ be an element in $\mathscr{H}$. Then by (1.2)

$$(1.3) \qquad Y(t) = \sum_{j=0}^{2} G_j(K_j, Y(t))\,, \qquad -\infty < t < \infty\,,$$

Received September 1972; revised May 1973.

[1] This paper is based on the author's Ph. D. dissertation at Carnegie-Mellon University.

[2] Throughout this paper limits of integration are from $-\infty$ to $\infty$ unless otherwise specified.

*AMS* 1970 *subject classifications.* Primary 60G10; Secondary 62M10.

*Key words and phrases.* Kernel, stationary stochastic process, spectral representation, orthogonal polynomial functional, cumulant spectrum, lag process.

where $G_0(K_0, Y(t)) = K_0$ is a constant and for $j = 1, 2$, $G_j(K_j, Y(t))$ is the projection of $Y(t)$ in $\mathscr{H}_j$ and $K_j$ is the leading (fixed) kernel of the projection (see [7], [15]). For convenience assume $EY(t) = 0$ for all $t$. Then $K_0 = 0$ with probability one. Also by the orthogonality of $G_1$ and $G_2$, $K_1(\lambda) = f_{XY}(\lambda)/f_{XX}(\lambda)$, $-\infty < \lambda < \infty$, where $f_{XX}$ and $f_{XY}$ are the spectrum of $X(t)$ and the cross spectrum of $X(t)$ and $Y(t)$, respectively. Now it is difficult to solve for $K_2$ due to the complexity of the resulting equations unless $X(t)$ admits special properties; e.g., $X(t)$ is Gaussian (see [14]). In the following we suggest a way for getting around the difficulties encountered in the solution for $K_2$ by considering a special class of polynomial functionals.

The problem of kernels estimation drew much attention in the past twenty years or so (see [2], [5], [6], [8], [9], [10], [14], [15] among others). In some of the above references $X(t)$ was assumed to be Gaussian (e.g., in [2], [14]). However, in this paper we do not make this assumption.

**2. Estimating the projection $G_2(K_2, Y(t))$.** In this section we show that under some conditions there exists a subset of $\mathscr{H}_2$ of functionals whose kernels can be determined. These functionals are used in approximating $G_2(K_2, Y(t))$ in mean square.

Assume that $K_2$ admits the Fourier representation

$$K_2(\lambda_1, \lambda_2) = \int\int \exp[-i(t_1\lambda_1 + t_2\lambda_2)]b(t_1, t_2)\,dt_1\,dt_2\,, \tag{2.1}$$

where $b$ is continuous and absolutely integrable.

LEMMA 2.1. *Assume that the fourth order cumulant spectrum of $X(t)$, $g_{XXXX}(\lambda_1, \lambda_2, \lambda_3)$, is absolutely integrable. Also, in* (2.1) *let $b$ satisfy*

(1) $\int |b(\tau, \tau + u)|\,d\tau < \infty$ *for each fixed $u$, $-\infty < u < \infty$.*
(2) $|b(\tau, \tau + u)| < \psi(\tau)$, *integrable, $-\infty < u < \infty$.*

*Then there exist bounded and continuous functions, $B_k(\lambda)$, $k = 1, \cdots, n$, $-\infty < \lambda < \infty$, such that the quadratic functional*

$$\int\int K_2(\lambda_1, \lambda_2)\,dZ_X(\lambda_1)\,dZ_X(\lambda_2) \tag{2.2}$$

*can be approximated arbitrarily closely in mean square by*

$$y_n = \sum_{k=1}^{n} \int\int B_k(\lambda_1 + \lambda_2)e^{iu_k\lambda_2}\,dZ_X(\lambda_1)\,dZ_X(\lambda_2)\,, \tag{2.3}$$

*where $u_k$, $k = 1, \cdots, n$, are real numbers.*

PROOF. The homogeneous functional (2.2) may be thought of as if it were derived from the functional

$$\int\int b(-\tau_1, -\tau_2)X(\tau_1)X(\tau_2)\,d\tau_1\,d\tau_2\,. \tag{2.4}$$

Let $\tau = \tau_1$ and $u = \tau_2 - \tau_1$ and define

$$g(\tau, u) = b(-\tau, -\tau - u)\,. \tag{2.5}$$

Then (2.4) can be expressed as

$$\iint g(\tau, u)X(\tau)X(\tau + u)\, d\tau\, du\,, \tag{2.6}$$

(see Akaike (1966) for a similar transformation). Define

$$B(\lambda; u) = \int g(\tau, u)e^{i\tau\lambda}\, d\tau\,. \tag{2.7}$$

Then by (1) $B(\lambda; u_0)$ is a continuous and bounded function of $\lambda$ for each fixed $u_0$, and by (2) the family $\{B(\lambda; \cdot), -\infty < \lambda < \infty\}$ is equicontinuous. Also, $B(\lambda; u)$ is an absolutely integrable function of $u$ for all $\lambda$. Define $\varphi_1(\lambda_1, \lambda_2)$ by

$$\varphi_1(\lambda_1, \lambda_2) = \int_{-a_\varepsilon}^{a_\varepsilon} B(\lambda_1 + \lambda_2; u)e^{iu\lambda_2}\, du - \int_{-\infty}^{\infty} B(\lambda_1 + \lambda_2; u)e^{iu\lambda_2}\, du\,, \qquad -\infty < \lambda_1, \lambda_2 < \infty\,. \tag{2.8}$$

It follows that for any arbitrarily small $\varepsilon > 0$ there exists an $a_\varepsilon$ such that

$$|\varphi_1(\lambda_1, \lambda_2)| \leqq \int_{|u|\geqq a_\varepsilon} \int_{-\infty}^{\infty} |g(\tau, u)|\, d\tau\, du < \varepsilon\,, \qquad -\infty < \lambda_1, \lambda_2 < \infty\,. \tag{2.9}$$

Partition (uniformly) the interval $[-a_\varepsilon, a_\varepsilon]$ by letting $-a_\varepsilon = u_0 < u_1' < u_1 < \cdots < u_{n-1} < u_n' < u_n = a_\varepsilon$ and define $\varphi_2(\lambda_1, \lambda_2)$ by

$$\varphi_2(\lambda_1, \lambda_2) = \sum_{k=1}^{n} B(\lambda_1 + \lambda_2; u_k')e^{iu_k'\lambda_2}(u_k - u_{k-1}) - \int_{-a_\varepsilon}^{a_\varepsilon} B(\lambda_1 + \lambda_2; u)e^{iu\lambda_2}\, du\,, \qquad -\infty < \lambda_1, \lambda_2 < \infty\,. \tag{2.10}$$

Obviously $\varphi_2$ is bounded:

$$|\varphi_2(\lambda_1, \lambda_2)| < M\,, \quad \text{constant}, \qquad -\infty < \lambda_1, \lambda_2 < \infty\,. \tag{2.11}$$

Moreover, for any $\varepsilon_1 > 0$ there exists $N(\varepsilon_1)$ such that whenever $n \geqq N(\varepsilon_1)$,

$$|\varphi_2(\lambda_1, \lambda_2)| < \varepsilon_1\,, \qquad \text{for all} \quad \lambda_2 \in \Lambda\,, \qquad -\infty < \lambda_1 < \infty\,, \tag{2.12}$$

where $\Lambda$ is any finite closed interval. In particular, choose $\Lambda$ such that on the complement of $\Lambda \times \Lambda \times \Lambda$, denoted by $(\Lambda \times \Lambda \times \Lambda)'$,

$$\iiint_{(\Lambda\times\Lambda\times\Lambda)'} |g_{XXXX}(\lambda_1, \lambda_2, \lambda_3)|\, d\lambda_1\, d\lambda_2\, d\lambda_3 < \varepsilon_2\,, \qquad \varepsilon_2 > 0\,, \tag{2.13}$$

and on the complement of $\Lambda \times \Lambda$, $(\Lambda \times \Lambda)'$, we have

$$\iint_{(\Lambda\times\Lambda)'} f_{XX}(\lambda_1)f_{XX}(\lambda_2)\, d\lambda_1\, d\lambda_2 < \varepsilon_3\,, \qquad \varepsilon_3 > 0\,. \tag{2.14}$$

For convenience let

$$B_k(\lambda_1 + \lambda_2) = B(\lambda_1 + \lambda_2; u_k')(u_k - u_{k-1})\,, \qquad k = 1, \cdots, n\,. \tag{2.15}$$

$$\begin{aligned}
&E\{\textstyle\iint [\sum_{k=1}^{n} B_k(\lambda_1 + \lambda_2)e^{iu_k'\lambda_2}]\, dZ_X(\lambda_1)\, dZ_X(\lambda_2) \\
&\quad - \textstyle\iint [\int B(\lambda_1 + \lambda_2; u)e^{iu\lambda_2}\, du]\, dZ_X(\lambda_1)\, dZ_X(\lambda_2)\}^2 \\
&\quad = E\,[\textstyle\iint \varphi_2(\lambda_1, \lambda_2)\, dZ_X(\lambda_1)\, dZ_X(\lambda_2) - \iint \varphi_1(\lambda_1, \lambda_2)\, dZ_X(\lambda_1)\, dZ_X(\lambda_2)]^2 \\
&\quad \leqq M^2\varepsilon_2 + \varepsilon_1^2 \textstyle\iiint_{\Lambda\times\Lambda\times\Lambda} |g_{XXXX}(\lambda_2, \lambda_3, \lambda_4)|\, d\lambda_2\, d\lambda_3\, d\lambda_4 \\
&\qquad + 3M^2\varepsilon_3 + 3\varepsilon_1^2 \textstyle\iint_{\Lambda\times\Lambda} f_{XX}(\lambda_1)f_{XX}(\lambda_2)\, d\lambda_1\, d\lambda_2 \\
&\qquad + \varepsilon^2 \textstyle\iiint |g_{XXXX}(\lambda_2, \lambda_3, \lambda_4)|\, d\lambda_2\, d\lambda_3\, d\lambda_4 + 3\varepsilon^2 R_{XX}^2(0) \\
&\qquad + 2(M\varepsilon \textstyle\iint |g_{XXXX}(\lambda_2, \lambda_3, \lambda_4)|\, d\lambda_2\, d\lambda_3\, d\lambda_4 + 3M\varepsilon R_{XX}^2(0))\,.
\end{aligned} \tag{2.16}$$

Now interchange the order of summation and integration in (2.16), and note that (2.2) can be rewritten as

(2.17) $$\iint [\int B(\lambda_1 + \lambda_2; u)e^{iu\lambda_2}\, du]\, dZ_X(\lambda_1)\, dZ_X(\lambda_2)\,. \qquad \square$$

It should be noted that polynomial functionals of the form

(2.18) $$\iint B_k(\lambda_1 + \lambda_2)e^{iu_k\lambda_2}\, dZ_X(\lambda_1)\, dZ_X(\lambda_2) + \int A_k(\lambda)\, dZ_X(\lambda) - B_k(0)R_{XX}(u_k)\,,$$

with the condition

(2.19) $$A_k(\lambda)f_{XX}(\lambda) + B_k(\lambda)\int e^{iu_k\lambda_1}f_{XXX}(\lambda - \lambda_1, \lambda_1)\, d\lambda_1 = 0\,,$$

are also elements of $\mathscr{h}_2$. In fact, sums of these functionals constitute a dense set in $\mathscr{h}_2$.

THEOREM 2.1. *Let the hypothesis of the lemma hold and let* $y_n$ *be as in* (2.3). *Define*

(2.20) $$y_n^* = \sum_{k=1}^{n} [\iint B_k(\lambda_1 + \lambda_2)e^{iu_k\lambda_2}\, dZ_X(\lambda_1)\, dZ_X(\lambda_2) + \int A_k(\lambda)\, dZ_X(\lambda) - B_k(0)R_{XX}(u_k)]\,,$$

*where for each k.* $k = 1, \cdots, n$, $A_k$ *is related to* $B_k$ *by* (2.19). *Then* $y_n^* \to G_2(K_2, Y(0))$ *in mean square as* $n \to \infty$.

PROOF. We see that $y_n = y_n^* + y_n^{*\perp}$ and

$$\iint K_2(\lambda_1, \lambda_2)\, dZ_X(\lambda_1)\, dZ_X(\lambda_2) = G_2(K_2, Y(0)) + G_2(K_2, Y(0))^{\perp}\,,$$

where $y_n^{*\perp}$, $G_2(K_2(Y(0))^{\perp} \in \mathscr{h}_2^{\perp}$, $\mathscr{h}_2^{\perp}$ being the orthogonal complement of $\mathscr{h}_2$. But $y_n \to \iint K_2(\lambda_1, \lambda_2)\, dZ_X(\lambda_1)\, dZ_X(\lambda_2)$, and therefore by continuity $y_n^* \to G_2(K_2, Y(0))$, $n \to \infty$. $\square$

COROLLARY 2.1. *Under the hypothesis of the lemma and with the same notation, there exists a sequence* $y_n^*(t)$, $-\infty < t < \infty$, *of functionals in* $\mathscr{h}_2$ *given by*

(2.21) $$y_n^*(t) = \sum_{k=1}^{n} [\iint \exp[it(\lambda_1 + \lambda_2)] \exp[iu_k\lambda_2]B_k(\lambda_1 + \lambda_2)\, dZ_X(\lambda_1)\, dZ_X(\lambda_2) + \int e^{it\lambda}A_k(\lambda)\, dZ_X(\lambda) - B_k(0)R_{XX}(U_k)]\,,$$

*such that* $y_n^*(t) \to G_2(K_2, Y(t))$ *in mean square as* $n \to \infty$.

PROOF. By Theorem 2.1, there exists $y_n^* = y_n^*(0)$ such that $y_n^*(0) \to G_2(K_2, Y(0))$ in mean square as $n \to \infty$. Therefore by the continuity of $T^t$ we have for each $t$, $-\infty < t < \infty$,

$$T^t y_n^*(0) \to T^t G_2(K_2, Y(0))$$

or

$$y_n^*(t) \to G_2(K_2, Y(t))\,. \qquad \square$$

We shall now determine the kernels $B_1(\lambda), \cdots, B_n(\lambda)$. It is not difficult to see that (2.21) can be rewritten as

(2.22) $$\sum_{k=1}^{n} [\int e^{it\lambda}B_k(\lambda)\, dZ_{U_k}(\lambda) + \int e^{it\lambda}A_k(\lambda)\, dZ_X(\lambda)]\,,$$

where $U_k(t)$ for each $k$, $k = 1, \cdots, n$, is a stationary lag process defined by

$$U_k(t) = X(t)X(t + u_k) - R_{XX}(u_k)\,, \tag{2.23}$$

and

$$A_k(\lambda) = -\frac{B_k(\lambda) f_{XU_k}(\lambda)}{f_{XX}(\lambda)}\,, \qquad -\infty < \lambda < \infty\,. \tag{2.24}$$

Clearly, $f_{XU_k}(\lambda) = \int \exp[iu_k \lambda_1] f_{XXX}(\lambda - \lambda_1, \lambda_1)\, d\lambda_1$.

Now, for a sufficiently large $n$, $Y(t)$ in (1.3) may be expressed by

$$\begin{aligned} Y(t) = {} & G_1\left(\frac{f_{XY}(\lambda)}{f_{XX}(\lambda)},\, Y(t)\right) \\ & + \sum_{k=1}^{n} \left[\int e^{it\lambda} B_k(\lambda)\, dZ_{U_k}(\lambda) + \int e^{it\lambda} A_k(\lambda)\, dZ_X(\lambda)\right], \\ & \qquad -\infty < t < \infty\,. \end{aligned} \tag{2.25}$$

Multiply both sides of (2.25) by

$$\int \exp[i(t + h)\lambda] B_\mu(\lambda)\, dZ_{U_\mu}(\lambda) + \int \exp[i(t + h)\lambda] A_\mu(\lambda)\, dZ_X(\lambda)\,, \qquad \mu = 1, \cdots, n\,. \tag{2.26}$$

Then on taking the expectations we have

$$\mathbf{B}(\lambda) = \left(\mathbf{f}_{UU}(\lambda) - \frac{1}{f_{XX}(\lambda)} \mathbf{f}_{UX}(\lambda)\mathbf{f}_{XU}(\lambda)\right)^{-1} \left(\mathbf{f}_{UY}(\lambda) - \frac{f_{XY}(\lambda)}{f_{XX}(\lambda)} \mathbf{f}_{UX}(\lambda)\right), \tag{2.27}$$

where

$$\mathbf{f}_{UU}(\lambda) = (f_{U_iU_j}(\lambda))\,,$$

$\mathbf{B}(\lambda) = (B_1(\lambda), \cdots, B_n(\lambda))'$, $\mathbf{f}_{UX}(\lambda) = (f_{U_1X}(\lambda), \cdots, f_{U_nX}(\lambda))'$, $\mathbf{f}_{UY}(\lambda) = (f_{U_1Y}(\lambda), \cdots, f_{U_nY}(\lambda))'$, and $\mathbf{f}_{XU}(\lambda)$ is the conjugate transpose of $\mathbf{f}_{UX}(\lambda)$.

**3. Summary.** The basic philosophy which characterizes the handling of a nonlinear problem in this paper is that of orthogonalization and linearization. Our linearization of quadratic functionals turned out to be fruitful due to the fact that we were able to orthogonalize these linearized functionals and consequently solve for their kernels. These functionals were used in the estimation of $G_2(K_2, Y(t))$.

Suppose we let $\mathscr{H}$ contain all polynomial functionals of degree up to and including $n$, $n > 2$. Then an extension of the proposed procedure to the estimation of higher degree projections seems to be difficult as the number of lags needed for an approximation increases rapidly and the orthogonality conditions become complicated.

**Acknowledgments.** I wish to thank Professor Paul Shaman for suggesting this area of research and for many helpful discussions. I also thank Professor Victor Mizel and the anonymous referee for their useful remarks.

## REFERENCES

[1] AKAIKE, H. (1966). Note on higher order spectra. *Ann. Inst. Statist. Math.* **18** 123-126.

[2] BRILLINGER, D. R. (1970). The identification of polynomial systems by means of higher order spectra. *J. Sound Vibr.* **12** 301-313.

[3] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[4] HANNAN, E. J. (1970). *Multiple Time Series.* Wiley, New York.

[5] KATZENELSON, J. and GOULD, L. A. (1962). The design of nonlinear filters and control systems, I. *Information and Control* **5** 108–143.

[6] KIČATOV, JU. F. (1970). The Gauss method of eliminating unknowns as applied to problems of nonlinear optimization. *Selected Translations in Mathematical Statistics and Probability*, **8** 159–166. I.M.S. and A.M.S., Providence.

[7] KIMELFELD, B. (1972). An orthogonal representation of nonlinear systems. Ph. D. thesis, Carnegie-Mellon Univ.

[8] KUZNETSOV, P. I., STRATONOVICH, R. L. and TIKHONOV, V. I. (1965). *Nonlinear Transformations of Stochastic Processes.* (English translation). Pergamon Press, Oxford.

[9] LEE, Y. W. and SCHETZEN, M. (1965). Measurement of the Wiener kernels of a nonlinear system by closs-correlation. *Internat. J. Control* **2** 237–254.

[10] MARMARELIS, P. Z. and NAKA, K. (1972). White-noise analysis of a neuron chain: An application of the Wiener theory. *Science* **175** No. 4027, 1276–1278.

[11] ROZANOV, YU. A. (1967). *Stationary Random Processes.* (English translation). Holden-Day, San Francisco.

[12] SHIRYAEV, A. N. (1960). Some problems in the spectral theory of higher-order moments, I. *Theor. Probability Appl.* **5** 265–284.

[13] SINAI, YA. G. (1963). On higher order spectral measures of ergodic stationary processes. *Theor. Probability Appl.* **8** 429–436.

[14] TICK, L. J. (1961). The estimation of transfer functions of quadratic systems. *Technometrics* **3** 563–567.

[15] WIENER, N. (1958). *Nonlinear Problems in Random Theory.* M.I.T. Press, Cambridge.

DEPARTMENT OF STATISTICS
UNIVERSITY OF HAIFA
MOUNT CARMEL, HAIFA, ISRAEL