The Annals of Statistics
1982, Vol. 10, No. 2, 630–633

A CHARACTERIZATION PROBLEM IN STATIONARY TIME SERIES

BY ERIC V. SLUD

University of Maryland

If a strictly stationary process $\{Z_k\}$ has residuals $Z_{k+1} - \sum_{j=1}^{k} a_{k,j}Z_j$ independent of $(Z_1, \cdots, Z_k)$ for all $k \geq m$, it is shown that the process is Gaussian or degenerate or m-step Markovian. Generalized (nonlinear) autoregressive stationary processes are defined and partially characterized.

1. Introduction. In standard presentations (e.g. Box and Jenkins, 1970) of prediction theory for stationary time series, the assumption that a series is Gaussian leads via the Wold Decomposition or Wiener's L^2 prediction theory (Grenander and Rosenblatt, 1957, pages 61–82) to the approximate correctness of the Autoregressive Moving Average (ARMA) modeling approach. Current work in stationary time series usually starts from an ARMA formulation (with or without Gaussian residuals) following the methods popularized by Box and Jenkins (1970). Relatively little prediction theory specific to non-Gaussian cases (as in Kanter, 1979) has been worked out. Even less work has gone into understanding the application of linear prediction methods to non-Gaussian non-ARMA series, but an interesting beginning has been made by Yaglom (1962). The present paper abstracts a special subclass (the "generalized autoregressive processes") of strictly stationary processes within which there is particularly well-defined non-linear prediction. Our results concern distributional characterizations of such processes satisfying additional assumptions motivated by ARMA theory.

We call a one-sided strictly stationary stochastic process $\{Z_k\}_{k=1}^{\infty}$ *generalized autoregressive* (GA) of order m if for every $k \geq m$ there is a measurable function $r_k(Z_1, \cdots, Z_k)$ such that $\varepsilon_{k+1} \equiv Z_{k+1} - r_k(Z_1, \cdots, Z_k)$ is independent of $(Z_1, \cdots, Z_k)$. The notion that trend functions can be subtracted away from (non-stationary) price series to arrive at independent (non-identically distributed) residuals suggests just such a definition. A priori, no special distributional form is associated with GA processes. But optimal nonlinear prediction (in any sense) for a GA process $\{Z_k\}$ is uniquely determined from the functions $r_k(\cdot)$ and the laws of residuals ϵ_{k+1}.

Gaussian stationary processes, the objects of study in standard ARMA theory, are GA with linear $r_k(\cdot)$. We call a stationary process $\{Z_k\}_{k=1}^{\infty}$ *generalized linear autoregressive* (GLA) of order m if it is GA with $r_k(Z_1, \cdots, Z_k) = \sum_{j=1}^{k} a_{j,k}Z_j$. The characterization problem of our title is simply: which non-Gaussian processes can be GLA?

We prove in Section 2 that GLA processes of order m are Gaussian, m-step Markovian, or linearly degenerate. Section 3 contains a brief discussion of non-Markovian GA processes, and of the bearing of the GA class on goodness of fit problems in time series.

2. Characterizing GLA processes. A fundamental observation on generalized autoregressive processes is provided in the following Lemma.

LEMMA 2.1. *Suppose $\{Z_k\}_{k=1}^{\infty}$ is stationary with $Z_{k+1} - r_k(Z_1, \cdots, Z_k) \equiv \varepsilon_{k+1}$ independent of $(Z_1, \cdots, Z_k)$ for $k \geq m$. Then for each $k \geq m$, there exists $\delta_k > 0$ such that*

$$E(\exp\{it[r_{k+1}(Z_1, \cdots, Z_{k+1}) - r_k(Z_2, \cdots, Z_{k+1})]\} \mid Z_2, \cdots, Z_{k+1})$$

almost surely does not depend on $Z_2, \cdots, Z_{k+1}$ for $|t| \leq \delta_k$.

Received November 1979; revised October 1981.

AMS 1979 *subject classifications.* Primary 62E10; secondary 62M10, 60G10, 60E10.

Key words and phrases. Stationary time series, generalized autoregressive process, characterization problem, nonlinear prediction.

PROOF. If $\{Z_i\}_{i=1}^{\infty}$ is GA, then by stationarity $Z_{k+2} - r_k(Z_2, \cdots, Z_{k+1})$ is independent of $(Z_2, \cdots, Z_{k+1})$ for $k \geq m$, so that

$$E(\exp(it[r_{k+1}(Z_1, \cdots, Z_{k+1}) + \varepsilon_{k+2} - r_k(Z_2, \cdots, Z_{k+1})]) \mid Z_2, \cdots, Z_{k+1})$$
$$= E(\exp(it\varepsilon_{k+2})) \cdot E(\exp(it[r_{k+1}(Z_1, \cdots, Z_{k+1}) - r_k(Z_2, \cdots, Z_{k+1})]) \mid Z_2, \cdots, Z_{k+1})$$

is almost surely constant as a function of $Z_2, \cdots, Z_{k+1}$. However, since $E(\exp(it\varepsilon_{k+2}))$ is non-zero for t in a sufficiently small neighborhood $[-\delta_k, \delta_k]$, our Lemma follows. □

Using the polygonal characteristic functions described in Feller (1971, pages 503–505), one can easily construct random variables Z_1, Z_2, Z_3 with $\mathscr{L}(Z_1, Z_2) = \mathscr{L}(Z_2, Z_3)$ for which $Z_1, Z_2 - r_1(Z_1)$, and $Z_3 - r_2(Z_1, Z_2)$ are independent, but $r_2(Z_1, Z_2) - r_1(Z_2)$ is not independent of Z_2.

The main application of our Lemma 2.1 will be to GLA processes. First we write a functional equation for the joint ch.f. of $(Z_1, \cdots, Z_m)$ when $\{Z_k\}$ is GLA of order m.

LEMMA 2.2. *Suppose* $\{Z_i\}_{i=1}^{m+1}$ *is stationary, with* $\varepsilon_{m+1} \equiv Z_{m+1} - \sum_{j=1}^{m} a_j Z_j$ *independent of* $(Z_1, \cdots, Z_m)$. *Let* $g(\mathbf{s}) = E(\exp(i(s_1 Z_1 + \cdots + s_m Z_m)))$, *and* $f(t) = E(\exp(it\varepsilon_{m+1}))$. *Then*

$$(\circ) \qquad g(\mathbf{s}) = f(s_m) \cdot g(s_m a_1, s_m a_2 + s_1, \cdots, s_m a_m + s_{m-1}).$$

The proof is obvious, since ($\circ$) simply expresses the stationarity of the joint ch.f. of $(Z_1, \cdots, Z_m, \varepsilon_{m+1} + \sum_{j=1}^{m} a_j Z_j)$. It is also easy to check that if $\{Z_k\}_{k=1}^{m+1}$ is non-degenerate, the matrix

$$A = \begin{pmatrix} 0 & \cdots & & a_1 \\ 1 & & 0 & a_2 \\ & \ddots & & \vdots \\ 0 & & 1 & a_m \end{pmatrix}$$

must have maximum modulus (spectral radius) <1. In this case, iterating ($\circ$), we find for $n \geq 1$,

$$(\dagger) \qquad g(\mathbf{s}) = f(s_m) \cdot f(s_m a_m + s_{m-1}) \cdots f((A^{n-1}\mathbf{s})_m) \cdot g(A^n \mathbf{s})$$

and there is at most one $g(\cdot)$ satisfying ($\circ$) for fixed $\mathbf{a}$ and non-constant $f(\cdot)$.

THEOREM 2.3. *If* $\{Z_k\}_{k=1}^{\infty}$ *is* GLA *of order* $m \geq 1$, *then* $\{Z_k\}$ *is m-step Markovian or linearly degenerate or Gaussian.*

PROOF. If we write $Z_{m+1} = \sum_{j=1}^{m} a_j Z_j + \varepsilon_{m+1}$, $Z_{m+2} = \sum_{j=1}^{m+1} b_j Z_j + \varepsilon_{m+2}$, then by Lemma 2.1, the conditional characteristic function of $\sum_{j=1}^{m+1} b_j Z_j - \sum_{j=2}^{m+1} a_{j-1} Z_j$ given $(Z_2, \cdots, Z_{m+1})$ does not depend on $(Z_2, \cdots, Z_{m+1})$ for $|t| \leq \delta_m$. Now $b_1 = 0$ would already imply that $(Z_2, \cdots, Z_{m+1})$ is linearly degenerate unless $b_j = a_{j-1}$ for $j = 2, \cdots, m+1$ (in which case $(Z_1, \cdots, Z_{m+2})$ is m-step Markovian). Otherwise $b_1 \neq 0$, and

$$(*) \qquad \begin{array}{l} E\{\exp(it[Z_1 + \sum_{j=2}^{m+1} (b_j - a_{j-1}) b_1^{-1} Z_j]) \mid Z_2, \cdots, Z_{m+1}\} \\ \text{does not depend on } Z_2, \cdots, Z_{m+1} \text{ for } |t| \leq \delta_m, \end{array}$$

while $Z_{m+1} - \sum_{j=1}^{m} a_j Z_j$ is by hypothesis independent of $(Z_1, \cdots, Z_m)$. If $b_{m+1} = a_m$ and $m \geq 2$, then locally near 0 the conditional characteristic function of Z_{m+1} given $Z_1, \cdots, Z_m$ does not depend on Z_1, i.e., the conditional characteristic function of $\sum_{j=1}^{m} a_j Z_j$ given Z_1 locally does not depend on Z_1. It follows using (*) when $b_{m+1} = a_m$ and $m \geq 2$ that $(Z_1, \cdots, Z_m)$ must be linearly degenerate unless $a_1 = 0$, in which case $(Z_1, \cdots, Z_{m+1})$ is $m-1$ step Markovian, $\{Z_k\}_{k=1}^{\infty}$ is GLA of order $m-1$, and the proof starts over inductively with m replaced by $m-1$. If $b_{m+1} = a_m$ and $m = 1$, then $E(\exp(itZ_1) \mid Z_2)$ a.s. does not depend on Z_2 for small enough t, and again $a_1 = 0$ unless Z_1, Z_2 are linearly degenerate. Finally, if b_{m+1}

$\neq a_m$, then the Skitovich-Darmois Theorem (Kagan, Linnik, and Rao, 1973, page 89) implies $Z_{m+1} - a_1 Z_1$ is conditionally normal given $Z_2, \cdots, Z_m$. Since ε_{m+1} is independent of $(Z_1, \cdots, Z_m)$, ε_{m+1} is a Gaussian variable. Now, by relation (†) above, if $\{Z_i\}_{i=1}^{m+1}$ is not linearly degenerate then $(Z_1, \cdots, Z_m)$ is jointly Gaussian.

The GLA process $\{Z_k\}_{k=1}^{\infty}$ of order m is a fortiori GLA of order n for each $n > m$. Applying the proof from the previous paragraph successively for $n = m + 1, m + 2, \cdots$, we conclude that $\{Z_k\}_{k=1}^{\infty}$ is linearly degenerate, m-step Markovian, or Gaussian. □

As has been pointed out by an anonymous referee, the special case of our Theorem where $m = 1$ can be made to follow from Theorem 10.3.1 of Kagan, Linnik, and Rao (1973).

3. Non-Markovian GA examples and time series. It is easy to find m-step Markovian examples of nonlinear Generalized Autoregressive processes. In fact, if $\rho : \mathbb{R}^m \to \mathbb{R}$ is any nonlinear contraction, and ε_{m+1} a random variable with $E(\max(\log|\varepsilon_{m+1}|, 0)) < \infty$, then there exists a unique stationary law for $(X_1, \cdots, X_{m+1})$ such that $X_{m+1} - \rho(X_1, \cdots, X_m)$ is independent of $(X_1, \cdots, X_m)$ and has the same law as ε_{m+1}.

More generally, we can give a typical inductive step in the construction of continuous-valued nonlinear autoregressions. Suppose $(Z_1, \cdots, Z_k)$ has stationary law, with $\varepsilon_k \equiv Z_k - r_{k-1}(Z_1, \cdots, Z_{k-1})$ independent of $(Z_1, \cdots, Z_{k-1})$. If $\mathscr{L}(Z_1 | Z_2, \cdots, Z_k)$ (with d.f. $F_{Z_1|Z_2,\cdots,Z_k}(\cdot | \cdot)$) is a.s. nonatomic and if $\mathscr{L}(\varepsilon_k)$ is not prime with respect to convolution, then we can define $Z_{k+1} \equiv r_{k-1}(Z_2, \cdots, Z_k) + \varphi(F_{Z_1|Z_2,\cdots,Z_k}(Z_1 | Z_2, \cdots, Z_k)) + \varepsilon_{k+1}$, where ε_{k+1} is independent of $(Z_1, \cdots, Z_k)$ and φ is a monotone function such that $\mathscr{L}(\varepsilon_k) = \mathscr{L}(\varphi(F_{Z_1|Z_2,\cdots,Z_k}(Z_1 | Z_2, \cdots, Z_k)) + \varepsilon_{k+1})$. Automatically from this definition $\mathscr{L}(Z_1, \cdots, Z_k) = \mathscr{L}(Z_2, \cdots, Z_{k+1})$, so that $(Z_1, \cdots, Z_{k+1})$ is stationary. Special nonlinear non-Markovian GA processes are readily constructed with this inductive step. For example, if $\beta > 0$ and $\alpha_i > 0$ are fixed for $i = 1, 2, \cdots$, so that $\alpha \equiv \sum_{i=1}^{\infty} \alpha_i < \infty$, then there is an essentially unique way to construct a GA stationary process $\{Z_k\}$ with $\mathscr{L}(Z_k) = \Gamma(\alpha, \beta)$ and $\mathscr{L}(\varepsilon_k) = \Gamma(\sum_{j=k}^{\infty} \alpha_j, \beta)$. Since the ch.f.'s of ε_k are non-vanishing, Lemma 2.1 implies $r_{k+1}(Z_1, \cdots, Z_{k+1}) - r_k(Z_2, \cdots, Z_{k+1})$ must actually be independent of $Z_2, \cdots, Z_{k+1}$, and the only freedom in the inductive definition of $r_{k+1}(\cdot)$ lies in the choice of (not necessarily monotone) function $\varphi(\cdot)$.

The form of the general inductive step of the previous paragraph renders it unlikely that there are any simple examples (with all $r_k(\cdot)$ defined in few recursive steps from elementary functions) of non-Gaussian and non-Markovian GA processes. Moreover, Kagan, Linnik and Rao (1973, Chapter 4) have many deep results restricting the types of $\mathscr{L}(Z_1)$ or functions $r_{k+1}(Z_1, \cdots, Z_{k+1}) - r_k(Z_2, \cdots, Z_{k+1}) \equiv R_k(\mathbf{Z})$ for which $R_k(\mathbf{Z})$ and $(Z_2, \cdots, Z_{k+1})$ can be independent. In the (m-step) Markovian case as well, (∘) of Section 2 is a very special relation subsisting between the law of ε_{m+1} and the law of $(Z_1, \cdots, Z_m)$.

We take the GA processes to be relevant to statistical practice wherever the goodness of fit of time series models or hypothetical trend lines is assessed by testing residuals for independence (cf. Box and Pierce, 1970). Further work along the lines initiated by Yaglom (1962) is needed to see how bad linear model identification and prediction could be for nonlinear autoregressive processes. The richness of the GA class may make it the proper arena for examining the robustness of time-series prediction methods.

REFERENCES

BOX, G. and JENKINS, G. (1970). *Time Series Analysis: Forecasting and Control.* Holden-Day, San Francisco.

BOX, G. and PIERCE, D. (1970). Distribution of residual auto-correlations in auto-regressive integrated moving average time series models. *J. Amer. Statist. Assoc.* **65** 1509–1526.

FELLER, W. (1971). *An Introduction to Probability Theory and its Applications*, vol. II, 2nd ed. Wiley, New York.

GRENANDER, U. and ROSENBLATT, M. (1957). *Statistical Analysis of Stationary Time Series.* Wiley, New York.

KAGAN, A., LINNIK, YU. and RAO, C. R. (1973). *Characterization Problems in Mathematical Statistics*. Wiley, New York.

KANTER, M. (1979). Lower bounds for nonlinear prediction error in moving average processes. *Ann. Probability* **7** 128–138.

YAGLOM, A. M. (1962). Examples of optimum nonlinear extrapolation of stationary random processes. *Proc. of Sixth All-Union Conf. of Prob. Theory and Math. Statist, Vilnius*. 273.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742