The Annals of Statistics
1974, Vol. 2, No. 6, 1296–1301

CHARACTERIZATION OF THE PARTIAL AUTOCORRELATION FUNCTION

BY FRED L. RAMSEY

Institute for Mathematical Statistics, University of Copenhagen

The conditions $|\phi_k| \leqq 1$ for all $k = 1, 2, \cdots$ and $|\phi_k| = 1$ implies $\phi_{k+1} = \phi_k$ are both necessary and sufficient for a sequence of real numbers $\{\phi_k; k = 1, 2, \cdots\}$ to be the partial autocorrelation function for a real, discrete parameter, stationary time series. If all partial autocorrelations beyond the pth are zero, the series is an autoregression. If all beyond the pth have magnitude unity, the series satisfies a homogeneous stochastic difference equation. A stationary series is singular if and only if $\sum_1^N \phi_k^2$ diverges with N. The likelihood function for the partial autocorrelation function is produced, assuming normality.

1. Introduction and summary. Considerable attention has been given recently to the partial autocorrelation functions (PACF) of time series having particular model structure. (See [2] and [4].) This note provides necessary and sufficient conditions for a sequence of real numbers to be a PACF for a weakly stationary time series. One result in [2] is a special case of this theorem. The PACF provides an appealing vantage point from which to view the structure of time series because its own structure is so simple. It is unfortunate that parametrization of a time series by its PACF brings no apparent simplification of the difficult inference problems.

2. Preliminaries. Let $Z = \{0, \pm 1, \pm 2, \cdots\}$ and $Z_+ = \{1, 2, \cdots\}$. The discrete parameter time series $x = \{x_t, t \in Z\}$ is called a *second order* time series if all second moments are finite. x is said to be wide-sense stationary if it is a second order time series whose first and second order moments are independent of cardinal time.

Let χ be the set of all Gaussian, wide-sense stationary time series with zero mean and unit variance. (It is convenient to consider only χ but to think of an element $x \in \chi$ as being a typical member of a broad class of w.s. stationary series obtainable by location and scale changes and, possibly, distributional changes leaving the first and second order moments fixed.)

Let R be the set of all sequences $\rho = \{\rho_t, t \in Z\}$ which satisfy

(2.1) $$\rho_0 = 1\,;$$

$$\rho_t = \rho_{-t}\,, \quad \forall t\,, \qquad \text{(symmetry)};$$

and, for every $n \in Z_+$ and all choices of indices $t_1 < \cdots < t_n$ from Z and of real

Received May 1973; revised February 1974.

AMS 1970 *subject classifications.* 62M10, 62N15, 60G10.

Key words and phrases. Partial autocorrelations, autoregressions, time series models, stochastic difference equations, stationary random processes.

numbers $\delta_1, \cdots, \delta_n$,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \delta_i \delta_j \rho_{t_i - t_j} \geqq 0 \,, \qquad \text{(pos. semi-def.).} \tag{2.2}$$

It is convenient to consider

$$R = B(R) + I(R) \,,$$

where $I(R)$—the "interior"—consists of all sequences which are strictly positive definite, while $B(R)$—the "boundary"—consists of all sequences where equality in (2.2) is achievable for some non-trivial sets $\{\delta_i\}$, $\{t_i\}$.

It is well known that there is a 1–1 mapping $\Psi: \chi \to R$ for which $\Psi(x) = \rho$ if and only if $\rho_t = E\{x_t x_0\}$ for every $t \in Z$. (See [5].) The sequence $\rho = \Psi(x)$ is called the *autocorrelation function* (ACF) for the time series x.

Consider the Hilbert Space of real random variables with zero means and finite second order moments, with expected product as inner product (see [5]). Let $H_{s,t}$ be the subspace spanned by $\{x_{s+1}, \cdots, x_{t-1}\}$ for $t > s + 1$ and let $\hat{x}_s$ and $\hat{x}_t$ be the respective projections of x_s and x_t on $H_{s,t}$.

Definition. The PACF for the second order time series x is the doubly infinite sequence $\{\Phi_{s,t};\ t > s \in Z\}$ defined by

$$\Phi_{s,s+1} = \text{Correlation}\,\{x_s, x_{s+1}\} \,, \qquad (s \in Z) \,,$$

and for $t > s + 1$,

$$\Phi_{s,t} = \text{Correlation}\,\{(x_s - \hat{x}_s), (x_t - \hat{x}_t)\} \,.$$

Thus ([7], page 424) $\Phi_{s,t}$ is the partial correlation between x_s and x_t eliminating linear regressions on $x_{s+1}, \cdots, x_{t-1}$.

The PACF of an $x \in \chi$ is determined by a singly infinite sequence $\Phi = \{\Phi_t, t \in Z_+\}$ where $\Phi_t = \Phi_{s,s+t}$ for all $s \in Z$ and $t \in Z_+$. Supposing that $\rho \in I(R)$ the sequence Φ can be determined by solving the sequence of matrix equations

$$\mathbf{R}_k \boldsymbol{\alpha}^{(k)} = \boldsymbol{\rho}_k \,, \qquad \text{for} \quad k \in Z_+ \tag{2.3}$$

for $\boldsymbol{\alpha}^{(k)\prime} = (\alpha_1^{(k)}, \cdots, \alpha_k^{(k)})$, where $\mathbf{R}_k = (\rho_{|i-j|})$, $i, j = 1, \cdots, k$, and $\boldsymbol{\rho}_k' = (\rho_1, \cdots, \rho_k)$. The PACF is the sequence given by

$$\Phi_k = \alpha_k^{(k)} \,, \qquad \text{for} \quad k \in Z_+ \,.$$

Durbin [6] gave a method (usable for $\rho \in I(R)$) for sequentially solving (2.3). The relevant equations are:

$$\Phi_1 = \alpha_1^{(1)} = \rho_1 \tag{D.1}$$

$$\sigma_1^2 = 1 - \Phi_1^2 \tag{D.2}$$

$$\Phi_{k+1} = \alpha_{k+1}^{(k+1)} = \{\rho_{k+1} - \textstyle\sum_{j=1}^{k} \alpha_j^{(k)} \rho_{k+1-j}\}/\sigma_k^2 \tag{D.3}$$

$$\alpha_j^{(k+1)} = \alpha_j^{(k)} - \Phi_{k+1} \alpha_{k+1-j}^{(k)} \,, \qquad (j = 1, \cdots, k) \tag{D.4}$$

$$\sigma_{k+1}^2 = \sigma_k^2 (1 - \Phi_{k+1}^2) \,. \tag{D.5}$$

The first two equations of $D = \{(D.1) \to (D.5)\}$ give starting values and the remaining three explain how to go from stage k to $(k+1)$. Physically, $\alpha_j^{(k)}$ may be interpreted as the coefficient of x_{k+1-j} in the linear regression of x_{k+1} on $\{x_1, \cdots, x_k\}$. The value of σ_k^2 is the variance of the residual from that regression. Equation (D.5) can be used to show further that

$$|\mathbf{R}_{k+1}| = \prod_{j=1}^{k} (1 - \Phi_j^2)^{k+1-j} . \tag{2.4}$$

What seems to have been overlooked is that any sequence of constants Φ having $|\Phi_k| < 1$ also defines via D a unique sequence ρ which is positive definite because of (2.4). This is the essence of the proof for Theorem 1.

3. Characterization of a PACF. Necessary and sufficient conditions for a sequence of numbers to be a PACF are given. Some consequences are noted.

Definition. Let the set S consist of all real-valued sequences $s = \{s_k, k \in Z_+\}$ which satisfy

$$\begin{aligned} &\text{(a)} \quad |s_k| \leqq 1 , \qquad \text{for all} \quad k \in Z_+ ; \qquad \text{and} \\ &\text{(b)} \quad |s_k| = 1 \quad \text{implies} \quad s_{k+1} = s_k . \end{aligned} \tag{3.1}$$

It is again convenient to decompose S as

$$S = B(S) + I(S) ,$$

where $I(S)$ consists of all $s \in S$ for which (a) holds as a strict inequality for all k. Thus $B(S)$ consists of sequences which have $|s_k| = 1$ for some k.

Theorem 1. *The real, discrete parameter, second order time series x is wide sense stationary if and only if its* PACF $\{\Phi_{s,t}, t > s \in Z\}$ *satisfies*

$$\begin{aligned} &\text{(A)} \quad \Phi_{s,s+k} = \Phi_k \quad \textit{for all} \quad s \in Z \quad \textit{and} \quad k \in Z_+ ; \qquad \textit{and} \\ &\text{(B)} \quad \Phi = \{\Phi_k, k \in Z_+\} \in S . \end{aligned} \tag{3.2}$$

Furthermore, $\Phi \in I(S)$ if and only if $\Psi(x) = \rho \in I(R)$.

Equivalently,

Theorem 1. *There exists a one-to-one mapping $\xi: R \to S$ such that if $\rho = \Psi(x)$ for $x \in \chi$, then $\Phi = \xi(\rho)$ is the* PACF *of x. Furthermore, $\rho \in I(R)$ if and only if $\xi(\rho) \in I(S)$.*

Proof. *Case* I. *Necessity for $\rho \in I(R)$.* D has a unique solution for Φ. Each Φ_k is the correlation between two well-defined random variables and thus $|\Phi_k| \leqq 1$. However, (2.4) implies strict inequality must hold so that $\Phi \in I(S)$.

Case II. *Sufficiently for $\Phi \in I(S)$.* D has a unique solution for $\{\rho_k, k \in Z_+\}$, which is extended to ρ by (2.1). (2.4) implies $|\mathbf{R}_k|$ is strictly positive for all k. So for each k, all principal minorants of $\mathbf{R}_k$ have positive determinants. This implies $\mathbf{R}_k$ is positive definite for every $k \in Z_+$ which implies that ρ is itself positive definite.

Case III. *Necessity for $\rho \in B(R)$.* There exists a positive integer p for which

$|\mathbf{R}_k| = 0$ for all $k > p$ and $|\mathbf{R}_k| > 0$ for all $k \leqq p$. D has a unique solution for $\{\Phi_1, \cdots, \Phi_p\}$ with $|\Phi_k| < 1$ for $k < p$ and with $|\Phi_p| = 1$, by (2.4). There exists a unique vector $\boldsymbol{\lambda}' = (1, -\lambda_1, \cdots, -\lambda_p)$ such that

$$0 = \boldsymbol{\lambda}'\mathbf{R}_{p+1}\boldsymbol{\lambda} = E(x_t - \textstyle\sum_{j=1}^{p} \lambda_j x_{t-j})^2 \tag{3.3}$$

for all t, where $x = \Psi^{-1}(\rho)$. (Indeed $\lambda_j = \alpha_j^{(p)}$.) Stationarity implies the residuals from regression of x_s and x_t on $x_{s+1}, \cdots, x_{t-1}$ are zero with probability one for all s and $t > s + p$. Hence in the sense that zero predicts itself perfectly, it is natural to set $\Phi_{p+k} = \Phi_p$ for $k = 1, 2, \cdots$, arriving at a full sequence $\Phi \in B(S)$.

Case IV. *Sufficiency for* $\Phi \in B(S)$. There exists a p for which $|\Phi_k| = 1$ for all $k \geqq p$ and $|\Phi_k| < 1$ for all $k < p$. D has a unique solution for $\{\rho_0, \rho_1, \cdots, \rho_p\}$ where $\rho_0 = 1$, $|\mathbf{R}_k| > 0$ for $k < p$ and $|\mathbf{R}_{p+1}| = 0$. Let $\{x_1, \cdots, x_{p+1}\}$ be defined as having a multivariate Gaussian distribution with means zero and covariance matrix $\mathbf{R}_{p+1}$ (of rank p). The residuals from regressions of x_1 on $\{x_2, \cdots, x_{p+1}\}$ and of x_{p+1} on $\{x_1, \cdots, x_p\}$ are zero with probability one. That is, with probability one,

$$x_t = \textstyle\sum_{j=1}^{p} \beta_j x_{t+j} \qquad (t = 1) \tag{3.4}$$

and

$$x_t = \textstyle\sum_{j=1}^{p} \alpha_j x_{t-j} \qquad (t = p + 1)\,. \tag{3.5}$$

Extend the sequence $x_1, \cdots, x_{p+1}$ according to the difference equation (3.4) for $t = 0, -1, -2, \cdots$ and according to (3.5) for $t = p + 2, p + 3, \cdots$. The result is a wide sense stationary time series for which Φ is the PACF. Clearly the corresponding ρ, derivable from the series, is in $B(R)$. This completes the proof. □

One advantage to characterizing χ by S is that the structure of S is so simple. Each partial autocorrelation is free to vary over the open interval $(-1, 1)$ independently of the others. Another is the simplicity with which singular (purely deterministic, perfectly predictable) series may be described.

COROLLARY 1. *A stationary time series is singular if and only if its* PACF *satisfies*

$$\lim_{N\to\infty} \textstyle\sum_{k=1}^{N} \Phi_k^2 = +\infty\,.$$

The validity of this corollary follows from re-writing (D.5) as

$$\sigma_N^2 = \textstyle\prod_{k=1}^{N} (1 - \Phi_k^2)\,.$$

A third is that the likelihood function is easily derived.

THEOREM 2. *Let* $x = \{x_t, t \in Z\}$ *be such that, for every* $t \in Z$,

$$x_t = \mu + \gamma_0^{\frac{1}{2}} y_t\,,$$

where $y = \{y_t, t \in Z\}$ *is in* χ. *Then* x_1 *has a Gaussian distribution with mean* μ *and variance* γ_0; *and for* $k = 1, 2, \cdots$ *the conditional distribution of* x_{k+1} *given* $x_1, \cdots, x_k$ *is Gaussian with mean and variance given respectively by*

$$E(x_{k+1} \mid x_1, \cdots, x_k) = \mu + \textstyle\sum_{j=1}^{k} \alpha_j^{(k)}(x_{k+1-j} - \mu) \tag{3.6}$$

and

$$(3.7) \qquad \operatorname{Var}(x_{k+1} \mid x_1, \cdots, x_k) = \gamma_0(1 - \Phi_1^2) \cdots (1 - \Phi_k^2) .$$

Here the $\alpha_j^{(k)}$ coefficients are defined in (2.3).

The proof consists simply of remarking that $x_{k+1} - E(x_{k+1} \mid x_1, \cdots, x_k)$ is uncorrelated with $x_1, x_2, \cdots, x_k$ and has variance σ_k^2 and then extending (D.5) to (3.7).

Unfortunately, however, (3.6) is a rather complicated function of the PACF.

4. Stochastic difference equations. A time series x is called an autoregression of order p, denoted AR(p), if it satisfies the pth order stochastic difference equation

$$(4.1) \qquad x_t = \sum_{j=1}^{p} \alpha_j x_{t-j} + y_t$$

for all t, where $y = \{y_t, t \in Z\}$ is a "white noise" sequence of uncorrelated shocks having mean zero and variances $\sigma^2 > 0$. The following result, the necessity of which is well known, is used in [4] for identifying autoregressions.

THEOREM 3.[1] *The stationary time series x is an* AR(p) *if and only if its* PACF *is zero beyond p.*

Proof of sufficiency. Assume $x \in X$ and for each $t \in Z$ let (4.1) denote the unique decomposition of X_t into the sum of its projection on an orthogonal distance, y_t, to $H_{t-p-1,t}$. Stationarity and Corollary 1 imply $E(y_t) = 0$ and $E(y_t^2) = \sigma^2 > 0$ for all t. To establish that $y_t \perp y_s$ for all $s \neq t$, define H_t^* as the subspace spanned by all y_s for $s < t$. Then $H_t^* = \bigcup_{k=1}^{\infty} H_{t-k,t}$. We have $y_t \perp H_{t-p-1,t}$ by construction. So if, according to $H_{t-p-1,t}$ we write $x_{t-p-1} = \hat{x}_{t-p-1} + w_{t-p-1}$, then $y_t \perp \hat{x}_{t-p-1} (\in H_{t-p-1,t})$ and $y_t \perp w_{t-p-1}$ because $\Phi_{p+1} = 0$. Thus $y_t \perp x_{t-p-1}$, implying that $y_t \perp H_{t-p-2,t}$. This argument may be iterated when it is noted that this implies that (4.1) also represents the decomposition of x_t according to $H_{t-p-2,t}$. □

Time series texts state conditions on the structural coefficients in (4.1) which guarantee a stationary solution of (4.1) for x given that y is as stated. The conditions are that no roots of the polynomial equation

$$(4.2) \qquad \alpha(u) = 1 - \alpha_1 u - \cdots - \alpha_p u^p = 0$$

lie on the unit circle. If all of the roots (call them $g_1, \cdots, g_p$) lie outside the unit circle, a stationary solution exists where x_t is an infinite moving average of past and present shocks. Furthermore, if some roots lie inside the unit circle, a stationary solution exists whose autocorrelation function is identical to that of a stationary solution to a pth order stochastic difference equation where all roots do lie outside the unit circle. Therefore, Theorems 1 and 3 are seen to provide a statistical proof to the Barndorff-Nielsen and Schou theorem which states: the mapping of the (complicated) parametric region $\{(\alpha_1, \cdots, \alpha_p); |g_j| > 1$

[1] A similar statement, not using the PACF directly, was proved in [3].

for $j = 1, \cdots, p\}$ into the PACF of the corresponding AR(p) is a one-to-one map onto the cube $(-1, +1)^p$.

An interesting situation arises when $\sigma^2 = 0$ in (4.1). Proofs of the foregoing statements (see e.g. Anderson [1]) show that the roots of (4.2) need not be off the unit circle. The following corollary to Theorem 1, whose proof is Case IV, identifies these situations as boundary cases.

COROLLARY 2. *If, according to the equivalent decompositions of R and S, we write*

$$\chi = B(\chi) + I(\chi),$$

then $x \in B(\chi)$ if and only if it is a solution to (4.1) *with $y_t \equiv 0$ for some finite integer p.*

And the final result is a counterpart to the discussion above.

THEOREM 4. *The structural coefficients $\{\alpha_1, \cdots, \alpha_p\}$ admit a non-trivial, stationary solution to* (4.1) *with $y_t \equiv 0$ only if some roots of* (4.2) *lie on the unit circle. Furthermore, the spectral measure, ν, for such a series has its support limited to those frequencies $\{f_k, k = 1, \cdots, q \leqq p\}$ for which $\exp(i2\pi f_k)$ is a root of* (4.2). (That is, *only* those roots exactly on the unit circle are relevant to the structure of the series.)

PROOF. A stationary series has a spectral representation (see [5]), about which our assumptions imply the condition

$$\sigma^2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} |\alpha(e^{i2\pi f})|^2 \nu(df) = 0 \tag{4.3}$$

holds for the spectral measure ν. This implies that $\nu(\{f \mid \alpha(e^{i2\pi f}) \neq 0\}) = 0$. So if no root of (4.2) is on the circle, the only series possible is the trivial one. But if q roots are on the circle, then (4.3) will still hold if the support for ν is as stated.

REFERENCES

[1] ANDERSON, T. W. (1972). *The Statistical Analysis of Time Series.* Wiley, New York.

[2] BARNDORFF-NIELSEN, O. and SCHOU, G. (1973). On the parametrization of autoregressive models by partial autocorrelations. *J. Multivar. Anal.* 3.

[3] BLAESILD, P. (1972). Om autoregressive processer. Unpublished Masters thesis report. Department of Theoretical Statistics, Aarhus Univ., Denmark.

[4] BOX, G. E. P. and JENKINS, G. M. (1970). *Time Series Analysis: Forecasting and Control.* Holden-Day, San Francisco.

[5] CRAMÉR, H. and LEADBETTER, M. R. (1967). *Stationary and Related Stochastic Processes.* Wiley, New York.

[6] DURBIN, J. (1960). The fitting of time series models. *Rev. Inst. Internat. Statist.* **28** 233–244.

[7] KENDALL, M. G. and STUART, A. (1966). *Advanced Theory of Statistics*, **3**. Hafner, New York.

DEPARTMENT OF STATISTICS
OREGON STATE UNIVERSITY
CORVALLIS, OREGON 97331