Nonminimum phase non-Gaussian autoregressive processes
(noncausal/maximum likelihood/asymptotic normality)

F. JAY BREIDT†, RICHARD A. DAVIS‡, KEH-SHIN LII†, AND MURRAY ROSENBLATT§

†Department of Statistics, Colorado State University, Fort Collins, CO 80523; ‡Department of Statistics, University of California, Riverside, CA 92521; and §Department of Mathematics, University of California, San Diego, La Jolla, CA 92039

Contributed by Murray Rosenblatt, September 29, 1989

ABSTRACT The structure of non-Gaussian autoregressive schemes is described. Asymptotically efficient methods for the estimation of the coefficients of the models are described under appropriate conditions, some of which relate to smoothness and positivity of the density function \( f \) of the independent random variables generating the process. The principal interest is in nonminimum phase models.

Let \( \{X_t\} \) be a stationary process satisfying the system of difference equations

\[
X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} = Z_t, \quad \phi_p \neq 0,
\]

where the random variables \( Z_t \) are independent, identically distributed with mean zero and variance \( \sigma^2 > 0 \). It is well known that there is a unique stationary solution of the system if and only if

\[
\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0
\]

for \( |z| = 1 \) (1–3). The stationary solution is given by

\[
X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j},
\]

where \( \psi_j \) is the coefficient of \( z^j \) in the Laurent expansion of \( \phi(z)^{-1} \)

\[
\phi(z)^{-1} = \sum_{j=-\infty}^{\infty} \psi_j z^j,
\]

which is valid in some annulus \( d < |z| < d^{-1}, d < 1 \). If \( \phi(z) \neq 0 \) for \( |z| \leq 1 \) then \( \psi_j = 0 \) for \( j < 0 \) and \( X_t \) is called minimum phase (or causal)

\[
X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}.
\]

Otherwise \( X_t \) is nonminimum phase (or noncausal). The case completely opposite to the minimum phase situation is that in which

\[
\phi(z) \neq 0 \quad \text{for} \quad |z| \geq 1,
\]

since then

\[
X_t = \sum_{j=0}^{\infty} \psi_j Z_{t+j}.
\]

Such a process is purely nonminimum phase or maximum phase (or purely noncausal). Then

\[
(1 - \phi_1 z - \cdots - \phi_p z^p)(\psi_0 + \psi_{-1} z^{-1} + \cdots) = 1,
\]

implying that

\[
\psi_0 = \psi_{-1} = \cdots = \psi_{-p} = 0, \quad \psi_{-p} = -\phi_{-1}^{-1}.
\]

In the minimum and maximum phase cases one can approximate the likelihood by the conditional likelihood given the first \( p \) and the last \( p \) observations respectively. The mixed case (which has not been treated previously) is analyzed by decomposing the polynomial \( \phi(z) \) into its minimum and maximum phase components and considering the two autoregressive models that arise from this decomposition.

Factor \( \phi(z) \)

\[
\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p = \phi^\ast(z)\phi^\ast(z),
\]

where

\[
\phi^\ast(z) = 1 - \theta_1 z - \cdots - \theta_r z^r \neq 0 \quad \text{for} \quad |z| \leq 1,
\]

\[
\phi^\ast(z) = 1 - \theta_{r+1} z - \cdots - \theta_s z^s \neq 0 \quad \text{for} \quad |z| \geq 1,
\]

with \( r, s \geq 0, r + s = p \). If \( m_1, \ldots, m_s, m_{s+1}, \ldots, m_p \) are the zeros of \( \phi^\ast(z) \) with \( |m_i| > 1, i = 1, \ldots, r \) and \( |m_i| < 1, i = r + 1, \ldots, p \) then

\[
\phi^\ast(z) = \prod_{i=1}^{r} (1 - m_i^{-1} z), \quad \phi^\ast(z) = \prod_{i=r+1}^{p} (1 - m_i^{-1} z).
\]

Consider the minimum and maximum phase autoregressive processes

\[
U_t = \phi^\ast(B)X_t, \quad V_t = \phi^\ast(B)X_t,
\]

with \( B \) the backward shift operator \((BX_t = X_{t-1})\). Then

\[
\phi^\ast(B)U_t = Z_t, \quad \phi^\ast(B)V_t = Z_t
\]

and so

\[
U_t = \sum_{j=0}^{\infty} \alpha_j Z_{t-j}, \quad V_t = \sum_{j=0}^{\infty} \beta_j Z_{t+j},
\]

where

\[
\phi^\ast(z)^{-1} = \sum_{j=0}^{\infty} \alpha_j z^j, \quad \phi^\ast(z)^{-1} = \sum_{j=0}^{\infty} \beta_j z^{-j}.
\]

The Likelihood Function. At this point we assume \( A_0 \). The distribution of \( Z_t \) is absolutely continuous with density function

\[
f_x(x) = \sigma^{-1} f(x/\sigma).
\]

The probability density function of \((U_1, \ldots, U_t, V_{n-s+1}, \ldots, V_n)\) can then be seen to be

\[
h_U(U_1, \ldots, U_t) \prod_{i=r+1}^{p} f_x(U_t - \theta_i U_{t-1} - \cdots - \theta_s U_{t-r})
\]

\[
h_V(V_{n-s+1}, \ldots, V_n),
\]

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. 81734 solely to indicate this fact.
where \( h_U \) and \( h_V \) are the joint probability density functions of \((U_1, \ldots, U_r)' \) and \((V_{n-1}, \ldots, V_n)' \), respectively. The joint probability density function of \((U_1, \ldots, U_s, X_1, \ldots, X_n)' \) is then

\[
h_U(U_1, \ldots, U_r) \prod_{i=r+1}^{n} f_p(X_i - \phi_1 X_{i-1} - \cdots - \phi_p X_{i-p})
\]

\[= \prod_{i=r+1}^{n} f_p(X_i - \phi_1 X_{i-1} - \cdots - \phi_p X_{i-p})
\]

with \( \log |\det T| \sim \log |\theta_p|^{-p} \). This suggests approximating the log-likelihood by

\[
L(\theta_1, \ldots, \theta_p, \sigma) = \sum_{i=p+1}^{n} \{ \log f_p(U_i - \theta_1 U_{i-1} - \cdots - \theta_p U_{i-p}) + \log |\theta_p| \}
\]

\[
= \sum_{i=p+1}^{n} g_i(\theta), \quad [14]
\]

with

\[
g_i(\theta) = \log f_p(U_i - \theta_1 U_{i-1} - \cdots - \theta_p U_{i-p}) + \log |\theta_p| = \log f_p(V_i - \theta_1 V_{i-1} - \cdots - \theta_p V_{i-p}) + \log |\theta_p|
\]

\[\text{and } \theta = (\theta_1, \ldots, \theta_p, \sigma)' \text{ is the parameter vector. In the minimum phase case, the } \log |\theta_p| \text{ term does not appear in Eq. 14.}
\]

We now make the following additional assumptions on \( f \).

\[f(x) > 0 \text{ for all } x, \quad [A1]\]

\[f \in C^2, \quad [A2]\]

\[f' \in L^1 \text{ with } \int f'(x)dx = f(x) \big|_{-\infty}^{\infty} = 0, \quad [A3]\]

\[xf'(x)dx = xf(x) \big|_{-\infty}^{\infty} - \int f(x)dx = -1, \quad [A4]\]

\[f'(x)dx = f'(x) \big|_{-\infty}^{\infty} = 0, \quad [A5]\]

\[xf'(x)dx = xf'(x) \big|_{-\infty}^{\infty} - \int f'(x)dx = 0, \quad [A6]\]

\[xf'(x)dx = 2xf'(x) \big|_{-\infty}^{\infty} - 2 \int f'(x)dx = 0, \quad [A7]\]

\[x^2 f'(x)dx = (x^2 f'(x)) \big|_{-\infty}^{\infty} - 2 \int xf'(x)dx = 2, \quad [A8]\]

\[\int (1 + x^2)(f'(x))^2/f(x)dx < \infty, \quad [A9]\]

The functions \( \gamma_U(\cdot) \) and \( \gamma_V(\cdot) \) denote the covariance functions of \( \{U_i\} \) and \( \{V_i\} \). Let \( \Sigma \) be the symmetric matrix with \( (i,j) \)-th element \( \sigma_{ij}(i \leq j) \)

\[\begin{bmatrix}
I \gamma_U(i - j) & 1 \leq i \leq j \leq r \\
I \gamma_V(i - j) & r < i \leq j \leq p, i \neq p \\
I \gamma_U(0) + \sigma_p^2 \gamma_U(j - 1) & i = j = p \\
-\sigma_p^{-2} \gamma_V(i - j) & i = p, j = p + 1 \\
J & i = j = p + 1 \\
0 & \text{otherwise.} \\
\end{bmatrix}
\]

This is the covariance matrix in the case \( s > 0 \). Otherwise in the minimum phase case, lines 2–5 of Eq. 16 are omitted.

**Proposition 1.** Under conditions A0–A8

\[(n - p)^{-1} \operatorname{cov} \left( \frac{\partial L}{\partial \theta_j} \right) \rightarrow \Sigma \quad [17]\]

as \( n \to \infty \). If \( f \) is a nonnormal probability density satisfying these conditions the matrix \( \Sigma \) is positive definite (strictly).

**Asymptotics.** A simple central limit theorem for finite step-dependent stationary sequences can be used to obtain Proposition 2.

**Proposition 2.** If \( f \) satisfies assumptions A1–A8

\[
(n - p)^{-1/2} \sum_{i=p+1}^{n} \frac{\partial L}{\partial \theta_j} \rightarrow N(0, \Sigma), \quad [18]\]

converges in distribution to \( N(0, \Sigma) \), the \((p + 1)\)-dimensional normal distribution with mean zero and covariance matrix \( \Sigma \).

Consider the likelihood equations

\[
\frac{\partial L(\theta)}{\partial \theta_j} = 0, \quad j = 1, \ldots, p + 1. \quad [18]
\]

Initially we assume that \( L \) does not depend on \( s \), the number of zeros of \( \phi(z) \) that lie inside \( |z| < 1 \); i.e., \( s \) is given. Also assume that

\[
h'(x) = \frac{d}{dx} \left( \frac{f'(x)}{f(x)} \right) = h_1(x) - h_2(x), \quad [19]
\]

with \( h_1, h_2 \) nondecreasing and

\[E(Z_j^2 | h_j(Z_4)) < \infty, \quad j = 1, 2. \quad [20]\]

The following result can then be obtained.

**Theorem 1.** Let \( \{X_t\} \) be the stationary autoregressive process satisfying Eq. 1 where the polynomial \( \phi(Z) \) has factorization 6 and \( \{Z_t\} \) is independent, identically distributed with mean zero, variance \( \sigma^2 > 0 \), and density function \( \sigma^{-2}f(x/\sigma) \). Let \( f \) be a nonnormal density function satisfying conditions A1–A8 and Eqs. 19 and 20. Then there is a sequence of solutions \( \theta_n \) to the likelihood Eq. 18 that is asymptotically normal with mean \( \hat{\theta}_0 \) (the true parameter vector) and asymptotic covariance matrix \( n^{-1} \Sigma^{-1} \) with \( \Sigma \) given by Eq. 16.

By using Theorem 1 it is possible to obtain the corresponding result for estimates of the original autoregressive param-
The \( \theta \) values can be determined from the \( \theta \) values by using the equations

\[
\phi_j = \begin{cases} 
\theta_j - \sum_{i=1}^{j} \theta_{r+i}, & j = 1, \ldots, r \\
- \sum_{i=j-r}^{j} \theta_{r+i}, & j = r + 1, \ldots, p.
\end{cases}
\]  \[21\]

where \( \theta_0 = -1 \) and \( \theta_j = 0 \) if \( j \notin \{0, \ldots, p\} \). The estimate is computed by replacing the \( \theta \) values in Eq. 21 by their estimated values. One can then show that

\[ n^{1/2}(\hat{\phi}_n - \phi) \]

converges in distribution to \( N(0, R \Sigma^{-1} R') \), where \( \Sigma^{-1} \) is the \( p \times p \) submatrix of \( \Sigma^{-1} \) corresponding to \( \theta_1, \ldots, \theta_p \) and

\[ R = \left( \frac{\partial \phi}{\partial \theta} \right)^p_{j,k=1}. \]  \[22\]

Theorem 1 remains valid when \( s \) is included as an unknown parameter of the likelihood function. Notice that the parameter space for the model is taken to be

\[ \Omega = \{ \phi \in R^p, \sigma > 0: 1 - \phi_1 z \cdots - \phi_p z^p \neq 0 \text{ for } |z| = 1, \phi_p \neq 0 \}. \]  \[23\]

We remark that ref. 4 provides a nonparametric inefficient procedure for parameter estimation in a broader context. The paper of Kreiss (5) considers an adaptive procedure in the minimum phase case. We have not yet considered adaptive procedures for the nonminimum phase case.

This research was supported by National Science Foundation Grants 8802559 and DMS83-12106 and by Office of Naval Research Contracts N00014-85-k-0468 and N00014-81-0003.