SF 298 MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
Dec 95

3. REPORT TYPE AND DATES COVERED
Final 15 Apr 91 - 14 Oct 95

4. TITLE AND SUBTITLE
Differential Geometrical Methods in Time Series

5. FUNDING NUMBERS
DAAL03-91-G-0104

6. AUTHOR(S)
Nalini Ravishanker

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)
Univ. of Connecticut
Storrs, CT 06268

19960212 072

ON

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
ARO 28743.11-MA

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Inference for time series processes was investigated using differential geometrical methods as well as sampling based Bayesian methods. For univariate autoregressive moving average (ARMA) processes and fractionally integrated ARMA processes, analytical forms of asymptotic, properties of inference such as bias in parameter estimates and improved test statistics were obtained from geometrical quantities. These terms provide corrections useful when the sample size is moderate or small. Markov chain Monte Carlo procedures facilitated modeling of univariate and multivariate ARMA and fractionally integrated ARMA processes in the Bayesian framework. This approach uses the exact likelihood function and is accurate even with small sample sizes. Outlier analysis, prediction and model selection were addressed.

14. SUBJECT TERMS
ARMA models, asymptotic inference, differential geometry, long-memory models, sampling-based Bayesian inference

15. NUMBER IF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

DIFFERENTIAL GEOMETRICAL METHODS IN TIME SERIES FINAL REPORT

Nalini Ravishanker

December 10, 1995

Grant No: DAAL03-91-G-0104

Institution: University of Connecticut, Storrs

THE VIEWS, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF THE AUTHOR(S) AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER DOCUMENTATION.

A. Statement of the Problem Studied.

Inference in parametric time series models was addressed with the goal of obtaining improved parameter estimates, test statistics and predictions. Research centered on two frameworks, viz.

a) application of differential geometrical methods in time series and

b) sampling based Bayesian methods in time series.

Four popular and useful classes of time series processes that find wide application in physical and social sciences were considered for study.

i) Univariate autoregressive moving average (ARMA) processes:

$$\phi(B)(x_t - \mu) = \theta(B)a_t,$$

where x_t is a time series, $Bx_t = x_{t-1}$, μ is the process mean, $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ are polynomials in B of degrees p and q respectively, $\{a_t\}$ is a sequence of i.i.d. normal variables with mean 0 and variance σ_a^2. These processes are useful in characterizing short-memory in time series.

ii) Univariate autoregressive fractionally integrated moving average (ARFIMA) processes:

$$\phi(B)(1-B)^d(x_t - \mu) = \theta(B)a_t,$$

where $(1-B)^d = \sum_{j=0}^{\infty} \binom{d}{j}(-B)^j$, $-\frac{1}{2} < d < \frac{1}{2}$ and the remaining terms are as defined in (i). These processes are useful in characterizing long-memory if $d > 0$ and intermediate-memory if $d < 0$.

iii) Multivariate ARMA processes:

$$\Phi(B)(\underset{\sim}{\mathbf{x}}_t - \underset{\sim}{\mu}) = \theta(B)\, \underset{\sim}{a}_t,$$

where $\underset{\sim}{\mu} = (\mu_1, ..., \mu_k)$ is the mean of the k-variate processes $\underset{\sim}{\mathbf{x}}_t$, $\Phi(B) = I - \Phi_1 B - \cdots - \Phi_p B^p$ and $\theta(B) = I - \theta_1 B - \cdots - \theta_q B^q$ are matrix polynomials in p and q respectively and $\underset{\sim}{a}_t$ are i.i.d. k-variate normal vectors with mean $\underset{\sim}{0}$ and covariance matrix Σ.

iv) Multivariate ARFIMA processes:

$$\Phi(B)D(B)(\underset{\sim}{\mathbf{x}}_t - \underset{\sim}{\mu}) = \theta(B)\, \underset{\sim}{a}_t,$$

where $D(B) = Diag[(1-B)^{d_1}, ..., (1-B)^{d_k}]$, a diagonal matrix containing the orders of fractional differencing for each series, and the remaining quantities are defined in (iii).

Application of differential geometrical methods to study asymptotic inference was investigated for (i) and (ii). All four processes were studied in the Bayesian framework.

B. Summary of the Most Significant Results.

For each of the four processes described in A, significant research findings under ARO support are reported here.

i) Univariate ARMA processes.

As an application of *differential geometry*, relative curvature measures of nonlinearity were derived for nonseasonal and multiplicative seasonal autoregressive moving average (ARMA) models to be used as diagnostic tools to assess the degree of model nonlinearity. To illustrate, the maximum and root mean square curvatures were computed for 16 time series data sets modeled in the literature by ARMA models; it was seen that more than half the sets exhibited significantly large intrinsic or parameter-effects curvature. The effect of curvature on the confidence regions for parameters was discussed and illustrated by examples. In particular, the maximum relative intrinsic curvature

$$\gamma^N_{\max} = m\sigma_a^2 \max_h \frac{\| h'\widehat{F}^N_{..}h \|}{\| \widehat{F}_{.}h \|^2}$$

and the maximum relative parameter-effects curvature

$$\gamma^T_{\max} = m\sigma_a^2 \max_h \frac{\| h'\widehat{F}^T_{..}h \|}{\| \widehat{F}_{.}h \|^2}$$

where h is an $(m \times 1)$ direction vector were computed; both the curvatures must be small for the validity of the linear approximation. Whereas $\gamma^N_{\max}$ does not depend on the particular model parametrization, $\gamma^T_{\max}$ does and can be reduced by a suitable reparametrization. For the same setting, the intrinsic and parameter-effects mean square curvatures were also computed (see Ravishanker, 1994 and Bates and Watts, 1980). For time series models, the relative intrinsic curvature is large, which implies that reparameterization would not alleviate the nonlinearity. Hence, corrections to confidence regions for curvature is useful and was studied. These curvature measures would serve as indicators of nonlinearity for a model/data combination in practical modeling.

We considered (Ravishanker and Tsai, 1993) tests of the autocorrelation parameter in a linear regression model with first-order autoregressive, AR(1), errors, viz.

$$Y = X\beta + e,$$

where Y and e are $n \times 1$ vectors of observations and errors respectively, X is an $n \times p$ matrix of deterministic regressors, β is a $p \times 1$ unknown parameter vector. We also assumed that e has a normal distribution with mean 0 and covariance matrix $\sigma^2\psi$ where

$$\psi = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & \rho & \cdots\cdot\rho^{n-1} \\ \rho & 1 & \cdots\cdot\rho^{n-2} \\ \cdot & \cdot & \cdot\cdot \\ \cdot & \cdot & \cdot\cdot \\ \cdot & \cdot & \cdot\cdot \\ \rho^{n-1} & \rho^{n-2} & \cdots\cdot 1 \end{bmatrix}$$

and $|\rho| < 1$. We derived modified profile likelihood ratio and modified score tests fpr $H_0 : \rho = 0$ by using the modified profile likelihood of Cox and Reid (1987). We showed via Monte Carlo simulations that these two tests are powerful and reliable by comparison to other existing tests such as the likelihood ratio test, the likelihood ratio test with Bartlett's adjustment and the most popular test, the Durbin-Warson test. We also generalized these tests to nonlinear regression models with AR(1) errors. We illustrated the behavior of these tests through examples.

For a general stationary and invertible $ARMA(p,q)$ process, we showed how to carry out an exact *Bayesian analysis* (Marriott, Ravishanker, Gelfand and Pai, 1995). Our approach was through the use of sampling based methods involving three novel aspects. First the constraints on the parameter space arising from the stationarity and invertibility conditions were handled by a convenient reparametrization to all of Euclidean $(p+q)$-space. Second, required sampling was facilitated by the introduction of latent variables which, though increasing the dimensionality of the problem, greatly simplified the evaluation of the likelihood. Third, the particular sampling based approach used was a Markov chain Monte Carlo method which is a hybrid of the Gibbs sampler (Gelfand and Smith, 1990) and the Metropolis algorithm. We also showed how straightforwardly the sampling based approach accommodates missing observations, outlier detection, prediction and model determination.

ii) Univariate ARFIMA processes.

The *differential geometry* of ARFIMA processes was studied (Ravishanker, 1994). Properties of Toeplitz forms associated with the spectral density functions of these long memory processes were used to compute the geometric quantities associated with these problems. The role of these geometric quantities on the asymptotic bias and efficiency of the maximum likelihood estimates of the model parameters and on the Bartlett correction to the likelihood ratio test statistic for the fractional difference parameter was discussed. The study of asymptotic inference for the $ARFIMA$ process is of considerable current research interest (Dahlhaus, 1989). This research is an extension of previous work by Ravishanker et.al. (1990) where the geometry of the ARMA process was discussed in relation to asymptotic inference. Analytical expressions for various geometric properties for $ARFIMA$ processes were computed by utilizing results on Toeplitz forms (see Dahlhaus, 1989). Due to the unboundedness of $s_\beta(w)$ at $w = 0$, these computations are not just trivial extensions of those for $ARMA(p,q)$ processes. They provide useful information on leading terms associated with asymptotic bias of the fractional difference parameter d asymptotic efficiency of d. So far, these quantities were only assessed through simulation.

We presented *Bayesian inference* for ARFIMA models using the sampling based approach (Pai and Ravishanker, 1994, 1996). We derived a form of the posterior density based on the exact likelihood function that is suitable for repetitive evaluation under the Metropolis-within-Gibbs algorithm. We presented a form of the posterior density based on the exact likelihood and which has greater computational feasibility, through the incorporation of latent variables $\underset{\sim 0}{a} = (a_{1-q}, ..., a_0)$ and $\underset{\sim 0}{z} = (z_{1-p}, ..., z_0)$ corresponding to the unknown historical portion of the $ARFIMA(0, d, 0)$ processes $\{a_t\}$ and the $ARFIMA(p, d, q)$ process $\{z_t\}$ respectively. We derived the joint distribution of the latent variables and data and thus a simple form for the posterior distribution. This work has generated considerable interest when presented at statistical meetings.

iii) Multivariate ARMA processes.

We presented a general framework for simultaneous modeling and fitting of multivariate and concurrent ARMA processes (Pai, Ravishanker and Gelfand,

1994) using a framework of an exchangeable hierarchical Bayesian model, incorporating dependence among the time series. Our motivating data set consisted of regional IBM revenue available monthly for several geographic regions. A modified Gibbs sampling algorithm was used to carry out the fitting and to enable all subsequent inference. Graphical techniques using predictive distributions were employed to assess model adequacy and to select among models. Outlier estimation and prediction under the chosen model were used for planning and to measure the effect of special promotional events.

We discussed forecasting, planning and contemporaneous outlier analysis for concurrent time series based on shrinkage estimation (Ravishanker, Wu and Dey, 1994a). We constructed a bootstrap estiamte of the covariance of the parameter estimates vector to facilitate the shrinkage and showed the improvement derived (see also Ravishanker, Wu and Dey, 1995).

iv) Multivariate ARFIMA processes.

We presented a general framework for Bayesian inference of multivariate time series exhibiting both long and short memory behavior (Ravishanker and Ray, 1995). The series were modeled using a multivariate ARFIMA process, which can capture both the short and long memory characteristics of the individual series, as well as interdependence and feedback relationships between the series. To facilitate a sampling-based Bayesian approach, we derived the exact joint posterior density for the parameters in a form that is computationally feasible and used a modified Gibbs sampling algorithm to generate samples from the complete conditional distribution associated with each parameter. We also showed how an approximate form of the joint posterior density may be used for long time series. Scatter plots constructed from the samples from the Markov chain Monte Carlo approach enable the characterization of properties of the estimates.

C. List of Publications and Technical Reports.

1. Ravishanker, N. (1994). Relative curvature measures of nonlinearity for time series models. Communications in Statistics, Simulation and Computation, 23, 415-430.

2. Marriott, J.M., Ravishanker, N., Gelfand, A.E. and Pai. J.S. (1995). Bayesian analysis of ARMA processes: complete sampling based inference under exact likelihoods. Bayesian Statistics and econometrics: Essays in honor of Arnold Zellner, eds., D. Berry, K. Chaloner and J. Geweke, John Wiley, New York.

3. Pai, J.S., Ravishanker, N. and Gelfand, A.E. (1994). Bayesian analysis of concurrent time series with application to regional IBM revenue data. J. Forecasting, 13, 463-479.

4. Ravishanker, N., Wu, L.S.-Y. and Dey, D.K. (1995). Shrinkage estimation in Time Series using a bootstrapped covariance estimate. J. Statistical Computation and Simulation (in press).

5. Ravishanker, N., Wu, L.S.-Y. and Dey, D.K. (1994). Forecasting, planning and contemporaneous outlier analysis for IBM regional revenue based on shrinkage estimation. ASA 1994 Proceedings of the Business and Economic Statistics Section.

6. Pai, J.S. and Ravishanker, N. (1994). Bayesian analysis of fractionally differenced ARIMA processes using Gibbs sampling. ASA 1994 Proceedings of the Bayesian Statistics Section.

7. Pai, J.S. and Ravishanker, N. (1996). Bayesian modeling of ARFIMA processes by Markov chain Monte Carlo methods. J. Forecasting (in press).

8. Ravishanker, N. and Ray, B.K. (1995). Bayesian analysis of multivariate ARFIMA processes. Technical Report No. 95-26, University of Connecticut.

9. Ravishanker, N. (1994). Differential geometry of autoregressive fractionally integrated moving average processes. Technical Report No. 94-23, University of Connecticut.

10. Ravishanker, N. and Tsai, C.L. (1993). Comparison of tests for AR(1) parameter in regression models with autocorrelated errors. Technical Report No. 93-01, University of Connecticut.

D. Participating Scientific Personnel

My graduate student, Jeffrey S. Pai, was supported by this grant. He received his Ph.D. degree from the University of Connecticut in June, 1994. The title of his thesis was 'Bayesian analysis of ARIMA processes'.

E. Bibliography

Bates, D.M. and Watts, D.G. (1980). Relative curvature measures of nonlinearity (with discussion). J. Roy. Statist. Soc., B, 42, 1-25.

Cox, D.R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference (with discussion). J. Roy. Statist. Soc., B, 49, 1-39.

Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. Ann. Statist., 17, 1749-1766.

Gelfand, A.E. and Smith, A.F.M. (1990). Sampling based approaches to calculating marginal densities. J. Amer. Statist. Assoc., 85, 398-409.

Ravishanker, N., Melnick, E.L. and Tsai, C.L. (1990). Differential geometry of ARMA models. J. Time Ser. Anal., 11, 259-274.