Adv. & Appl. in Stat. **6**(2) (2006), 129-144

TIME SERIES DISCRIMINANT ANALYSIS OF $AR(p)$ PLUS NOISE PROCESSES: A TIME DOMAIN APPROACH

RAHIM CHINIPARDAZ and SEYED MAHMOUD LATIFI

Shahid Chamran University, Ahvaz, Iran

Jundishapour University of Medical Sciences, Ahvaz, Iran

Abstract

The problem of discrimination between two autoregressive processes of order p is considered when the observed time series is contaminated with an extra noise and the main discriminatory information is in the covariance structure rather than the mean. An analytic discrimination rule is given based on likelihood ratio and its performance is examined.

It is well known that the distribution of the discrimination can be expressed in terms of a weighted sum chi-square random variables of one degree of freedom. The weights in the sum have to be calculated numerically. The approximated weights are calculated. It is shown that they are very close to the true values.

1. Introduction

A number of practical problems in time series analysis reduce to classifying a stochastic process to one or other categories. These applications are in physical sciences as seismic records, medical sciences as recorded brain waves, audiology, archeology, engineering and even in biology and developmental psychology.

2000 Mathematics Subject Classification: 62H30, 62M10.

Keywords and phrases: quadratic discrimination, $AR(p)$ plus noise processes, band matrices, cumulants.

Received April 5, 2005

A comprehensive overview of applications has been given in Shumway [14, 15] as well as common methodologies to the discrimination time series analysis in both time domain and frequency domain approaches (see also Dargahi-Noubary [6]). Some more references in this area are Dargahi-Noubary and Laycock [8], Dargahi-Noubary [5, 7], Alagon [1], Chan et al. [2], Shumway and Unger [16], Kakizawa et al. [12] and Chinipardaz [4].

Majority of works in time series discrimination, however, is devoted to considering ARMA processes which can be expressed as a linear combination of white noise processes (see Fuller [10]). ARMA models have great success in engineering, business and economics applications. However, as Dargahi-Noubary [6] pointed out, despite their wide applicability, no real attempts have been made to find out the reason behind their success in any particular application area.

It should be noted that improved models can be constructed by incorporating more available information than linear models. For instance consider the tracking of a missile fired from a submarine using satellite measurements. The missile position at time t, x_t, and its position at same time as observed by radar, y_t, may be embedded in a white noise, say ε_t, i.e.,

$$y_t = x_t + \varepsilon_t. \quad (1)$$

Now, if for example the movement of the missile is an autoregressive of order p models, i.e.,

$$x_t = \theta_1 x_{t-1} + \theta_2 x_{t-2} + \cdots + \theta_p x_{t-p} + \eta_t \quad (2)$$

both (1) and (2) altogether can be expressed as $AR(p)$ plus noise model or be considered as a state-space model which include ARMA models as an especial case. Such examples viewed signal plus noise and are given in Zyweck and Bogner [18].

Clearly, these models are more complicated to be used in time series discrimination because of an extra noise.

Chinipardaz [4] obtained discrimination rule for two $AR(1)$ models with an extra noise and compared with other works given in time domain approach when various variance of ε_t is considered. In this article the work extended to an autoregressive process of order p, $AR(p)$. Throughout, it is assumed that data are observed on stationary $AR(p)$ denoted by $\mathbf{x} = (x_1, x_2, ..., x_T)$. However, the observed time series is subjected to an extra stationary noise.

The distribution theory of classification rule is extremely complicated and involves the weighted sum chi-squared random variables of one degree of freedom (see Shumway [15] and Chaudhuri and Borwanker [3]). However, these weights are very complicated to obtain and matrix manipulation is required.

In this study, an attempt has been made to give approximated analytic weights. The paper is organized as follows: In Section 2, approximation to discrimination is suggested. Section 3 is devoted to simulation works to obtain the performance of given approach. The analytic distribution of discrimination between two $AR(p)$ processes plus noise is given in Section 4. Finally, in the last section the cumulants of discriminant function are compared with those given in the literature.

2. Discrimination Between Two $AR(p)$ Plus Noise Processes

Consider that the T observed dimensional vector $\mathbf{y} = (y_1, y_2, ..., y_T)'$ is a stationary time series process subjected to an extra noise, i.e.,

$$H_1 : y_t = x_t + \varepsilon_t$$

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \cdots + \alpha_p x_{t-p} + \eta_t$$

and

$$H_2 : y_t = x_t + \varepsilon_t$$

$$x_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_p x_{t-p} + \eta_t,$$

where $\alpha'_j s$ and $\beta'_j s$ are constants. ε_t and η_t are serially uncorrelated disturbances with zero mean and variances σ^2_ε and σ^2_η, respectively. It is assumed that ε_t and η_s are uncorrelated for all t and $s(t,\ s = 1, 2, \ldots, T)$, i.e.,

$$\operatorname{cov}(\varepsilon_t,\ \eta_s) = 0$$

$\alpha'_j s$ and $\beta'_j s$ are so that x_t is stationary and invertible. It means that y_t is also stationary and invertible. If $\mathbf{x}_0 = (x_{-1}, x_{-2}, \ldots, x_{-1+p})$ and $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p)'$ under H_1, then we have

$$\begin{aligned}
p(\mathbf{x};\ \mathbf{x}_0,\ \alpha) &= (2\pi\sigma_\eta^2)^{-\frac{T}{2}} \exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T}(x_t - \alpha_1 x_{t-1} - \cdots - \alpha_p x_{t-p})^2\right\} \\
&= (2\pi\sigma_\eta^2)^{-\frac{T}{2}} \exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T}\sum_{i=0}^{p}\sum_{j=0}^{p}\alpha_i\alpha_j x_{t-i}x_{t-j}\right\} \\
&= (2\pi\sigma_\eta^2)^{-\frac{T}{2}} \exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{i=1}^{p}\sum_{j=0}^{p}\alpha_i\alpha_j\sum_{t=1}^{T}x_{t-i}x_{t-j}\right\} \\
&= (2\pi\sigma_\eta^2)^{-\frac{T}{2}} \exp\left\{-\frac{1}{2\sigma_\eta^2}\left[\sum_{i=0}^{p}\sum_{t=1}^{T}\alpha_i^2 x_t^2\right.\right. \\
&\qquad \left.\left. + 2\sum_{i=0}^{p-1}\alpha_i\sum_{j=i+1}^{p}\alpha_j\sum_{t=j+1}^{T+i}x_{t-i}x_{t-j} + g(\mathbf{x},\ \alpha)\right]\right\} \\
&= (2\pi\sigma_\eta^2)^{-\frac{T}{2}} \exp\left\{-\frac{1}{2\sigma_\eta^2}\left[\mathbf{x}' B_{2p+1,\,\alpha}\mathbf{x} + g(\mathbf{x},\ \alpha)\right]\right\},
\end{aligned}$$

where $\alpha_0 = 1$, $B_{2p+1,\,\alpha}$ is a band matrix with band width $2p+1$ and $g(\mathbf{x},\ \alpha)$ is a corrected term which depends on few first and last terms of $x_t \cdot p(\mathbf{x};\ \mathbf{x}_0,\ \beta)$ can be obtained similarly as,

$$p(\mathbf{x};\ \mathbf{x}_0,\ \beta) = (2\pi\sigma_\eta^2)^{-\frac{T}{2}} \exp\left\{-\frac{1}{2\sigma_\eta^2}\left[\mathbf{x}' B_{2p+1,\,\beta}\mathbf{x} + g(\mathbf{x},\ \beta)\right]\right\}.$$

Removing the corrected terms for two models the loglikelihood ratio is given by

$$LLR = \frac{\ln(\alpha, \mathbf{y})}{\ln(\beta, \mathbf{y})} = -\frac{1}{2}\ln\frac{|\Sigma_1|}{|\Sigma_2|} + \frac{1}{2}\mathbf{y}'(\Sigma_2^{-1} - \Sigma_1^{-1})\mathbf{y}$$

$$= \frac{1}{2}\ln\frac{|\sigma_\varepsilon^2 I + \sigma_\eta^2\Sigma_\beta|}{|\sigma_\varepsilon^2 I + \sigma_\eta^2\Sigma_\alpha|} + \frac{1}{2}\mathbf{y}'[(\sigma_\varepsilon^2 I + \sigma_\eta^2\Sigma_\beta)^{-1} - (\sigma_\varepsilon^2 I + \sigma_\eta^2\Sigma_\alpha)^{-1}]\mathbf{y}$$

$$= \frac{1}{2}\ln\frac{|\sigma_\varepsilon^2 I + \sigma_\eta^2 B_{2p+1,\beta}^{-1}|}{|\sigma_\varepsilon^2 I + \sigma_\eta^2 B_{2p+1,\alpha}^{-1}|} + \frac{1}{2}\mathbf{y}'[(\sigma_\varepsilon^2 I + \sigma_\eta^2 B_{2p+1,\beta}^{-1})^{-1}$$

$$-(\sigma_\varepsilon^2 I + \sigma_\eta^2 B_{2p+1,\alpha}^{-1})^{-1}]\mathbf{y},$$

where $\Sigma_\alpha(\Sigma_\beta)$ and $\Sigma_1(\Sigma_2)$ stand for the covariance matrix of $\mathbf{x}$ and $\mathbf{y}$ under $H_1(H_2)$, respectively. $B_{2p+1,\alpha}$ and $B_{2p+1,\beta}$ can be approximated by a polynomial of order p of B_3 (see Chinipardaz [4]). Then

$$B_{2p+1,\alpha} \approx \sum_{j=0}^{p} c_j B_3^j, \qquad B_{2p+1,\beta} \approx \sum_{j=0}^{p} d_j B_3^j, \tag{3}$$

where (i, j)th element of B_3 is given by

$$[B_3]_{ij} = \begin{cases} -1 & |i-j| = 1 \\ 0 & \text{otherwise} \end{cases}$$

and c_j and d_j are constant coefficients depend on α and β, respectively and have to be obtained from (3) and $B_3^0 = I$ is defined to be identity matrix.

The eigenvalue of jth element of $B_{3,\alpha}$ is given by $\lambda_j = 1 + \alpha^2 - 2\alpha\cos\left(\frac{j\pi}{T+1}\right)$ and normalized corresponding eigenvector is

$$\xi_j = \frac{2}{T+1}\left\{\sin\left(\frac{j\pi}{T+1}\right), \sin\left(\frac{2j\pi}{T+1}\right), \ldots, \sin\left(\frac{Tj\pi}{T+1}\right)\right\}$$

(see Chan et al. [2] and Chinipardaz [4]). Taking $L = \{\xi_1, \xi_2, \ldots, \xi_T\}$, L is symmetric and orthogonal matrix. Therefore, based on loglikelihood ratio $\mathbf{y}$ is classified to H_1 if

$$\mathbf{y}'[(\sigma_\varepsilon^2 I + \sigma_\eta^2 B_{2p+1,\beta}^{-1})^{-1} - (\sigma_\varepsilon^2 I + \sigma_\eta^2 B_{2p+1,\alpha}^{-1})^{-1}]y \geq \ln \frac{|\sigma_\varepsilon^2 I + \sigma_\eta^2 B_{2p+1,\beta}^{-1}|}{|\sigma_\varepsilon^2 I + \sigma_\eta^2 B_{2p+1,\alpha}^{-1}|}$$

and to H_2 otherwise. By defining $\mathbf{z} = L\mathbf{y}$

$$LLR = -\frac{1}{2}\ln\frac{|(\sigma_\varepsilon^2 I + \sigma_\eta^2 LB_{2p+1,\alpha}^{-1}L')|}{|(\sigma_\varepsilon^2 I + \sigma_\eta^2 LB_{2p+1,\beta}^{-1}L')|} + \frac{1}{2}\mathbf{z}'[(\sigma_\varepsilon^2 I + \sigma_\eta^2 LB_{2p+1,\alpha}^{-1}L')^{-1} - (\sigma_\varepsilon^2 I + \sigma_\eta^2 LB_{2p+1,\beta}^{-1}L')^{-1}]\mathbf{z}$$

$$LLR \propto \mathbf{z}'\left\{\left[\sigma_\varepsilon^2 I + \sigma_\eta^2 L\left(\sum_{j=0}^{p} d_j B_3^j\right)^{-1} L'\right]^{-1} - \left[\sigma_\varepsilon^2 I + \sigma_\eta^2 L\left(\sum_{j=0}^{p} c_j B_3^j\right)^{-1} L'\right]^{-1}\right\}\mathbf{z}$$

$$= \mathbf{z}'\left\{\left[\sigma_\varepsilon^2 I + \sigma_\eta^2\left(\sum_{j=0}^{p} d_j (LB_3^j L')\right)^{-1}\right]^{-1} - \left[\sigma_\varepsilon^2 I + \sigma_\eta^2\left(\sum_{j=0}^{p} c_j (LB_3^j L')\right)^{-1}\right]^{-1}\right\}\mathbf{z}$$

$$= \mathbf{z}'\left\{\left[\sigma_\varepsilon^2 I + \sigma_\eta^2\left(\sum_{j=0}^{p} d_j \Lambda_3^j\right)^{-1}\right]^{-1} - \left[\sigma_\varepsilon^2 I + \sigma_\eta^2\left(\sum_{j=0}^{p} c_j \Lambda_3^j\right)^{-1}\right]^{-1}\right\}\mathbf{z}$$

$$= \sum_{r=1}^{T}\left\{\left[\frac{1}{\sigma_\varepsilon^2 + \sigma_\eta^2\left(\sum_{j=0}^{p} d_j \lambda_r^j\right)^{-1}}\right] - \left[\frac{1}{\sigma_\varepsilon^2 + \sigma_\eta^2\left(\sum_{j=0}^{p} c_j \lambda_r^j\right)^{-1}}\right]\right\} z_r^2$$

$$= \sum_{r=1}^{T}\left\{\frac{1}{\sigma_\varepsilon^2 + \sigma_\eta^2\left[\sum_{j=0}^{p} d_j\left(-2\cos\frac{\pi r}{T+1}\right)^j\right]^{-1}} - \frac{1}{\sigma_\varepsilon^2 + \sigma_\eta^2\left[\sum_{j=0}^{p} c_j\left(-2\cos\frac{\pi r}{T+1}\right)^j\right]^{-1}}\right\} z_r^2$$

and therefore $\mathbf{y}$ is classified to H_1 if

$$\sum_{r=1}^{T}\left\{\frac{1}{\sigma_\varepsilon^2 + \sigma_\eta^2\left[\sum_{j=0}^{p} d_j\left(-2\cos\frac{\pi r}{T+1}\right)^j\right]^{-1}} - \frac{1}{\sigma_\varepsilon^2 + \sigma_\eta^2\left[\sum_{j=0}^{p} c_j\left(-2\cos\frac{\pi r}{T+1}\right)^j\right]^{-1}}\right\} z_r^2$$

$$\geq \ln\left\{\frac{\prod_{r=1}^{T}\left|\sigma_\varepsilon^2 + \sigma_\eta^2\left[\sum_{j=0}^{p} c_j\left(-2\cos\frac{\pi r}{T+1}\right)^j\right]^{-1}\right|}{\prod_{r=1}^{T}\left|\sigma_\varepsilon^2 + \sigma_\eta^2\left[\sum_{j=0}^{p} d_j\left(-2\cos\frac{\pi r}{T+1}\right)^j\right]^{-1}\right|}\right\}.$$

It is a closed form of discrimination. Considering $\sigma_\varepsilon^2 = 0$, the problem leads to discriminant between two $AR(p)$ processes and $\mathbf{y}$ is classified to H_1 if

$$\frac{1}{\sigma_\eta^2}\sum_{r=1}^{T}\sum_{j=0}^{p}(d_j - c_j)\left(-2\cos\frac{\pi r}{T+1}\right)^j z_r^2 \geq \ln\frac{\prod_{r=1}^{T}\sum_{j=0}^{p} d_j\left(-2\cos\frac{\pi r}{T+1}\right)^j}{\prod_{r=1}^{T}\sum_{j=0}^{p} c_j\left(-2\cos\frac{\pi r}{T+1}\right)^j}.$$

In special case $p = 1$

$$d_0 = (1 + \beta^2), \qquad d_1 = \beta$$

$$c_0 = (1 + \alpha^2), \qquad c_1 = \alpha \tag{4}$$

and the discrimination rule leads to classified $\mathbf{y}$ to H_1 if

$$\sum_{r=1}^{T} \left\{ \frac{\sigma_\eta^2(\beta - \alpha)\left(\beta + \alpha - 2\cos\frac{\pi r}{T+1}\right)}{\left[\sigma_\varepsilon^2 + \sigma_\eta^2\left(1 + \alpha^2 - 2\alpha\cos\frac{\pi r}{T+1}\right)\right]\left[\sigma_\varepsilon^2 + \sigma_\eta^2\left(1 + \beta^2 - 2\beta\cos\frac{\pi r}{T+1}\right)\right]} \right\} z_r^2$$

$$\geq \ln \frac{\prod_{r=1}^{T}\left|\sigma_\varepsilon^2 + \sigma_\eta^2\left[1 + \alpha^2 - 2\alpha\cos\left(\frac{\pi r}{T+1}\right)\right]^{-1}\right|}{\prod_{r=1}^{T}\left|\sigma_\varepsilon^2 + \sigma_\eta^2\left[1 + \beta^2 - 2\beta\cos\left(\frac{\pi r}{T+1}\right)\right]^{-1}\right|} \tag{5}$$

and to H_2 otherwise. For $AR(2)$ plus noise processes

$$LLR = -\frac{1}{2}\ln\frac{|\Sigma_1|}{|\Sigma_2|} + \frac{1}{2}\mathbf{y}'(\Sigma_2^{-1} - \Sigma_1^{-1})\mathbf{y}$$

$$= \frac{1}{2}\ln\frac{|\sigma_\varepsilon^2 I + B_{5,\beta}^{-1}|}{|\sigma_\varepsilon^2 I + B_{5,\alpha}^{-1}|} + \frac{1}{2}\mathbf{y}'[(\sigma_\varepsilon^2 I + \sigma_\eta^2 B_{5,\beta}^{-1})^{-1} - (\sigma_\varepsilon^2 I + \sigma_\eta^2 B_{5,\alpha}^{-1})^{-1}]\mathbf{y}.$$

It leads to classify $\mathbf{y}$ to H_1 if

$$\sum_{r=1}^{T}\left\{\left[\sigma_\varepsilon^2 + \sigma_\eta^2\left[\left(d_0 + d_1\left(-2\cos\frac{\pi r}{T+1}\right) + d_2\left(-2\cos\frac{\pi r}{T+1}\right)^2\right)\right]^{-1}\right]^{-1}\right.$$

$$\left. - \left[\sigma_\varepsilon^2 + \sigma_\eta^2\left[\left(c_0 + c_1\left(-2\cos\frac{\pi r}{T+1}\right) + c_2\left(-2\cos\frac{\pi r}{T+1}\right)^2\right)\right]^{-1}\right]^{-1}\right\} z_r^2$$

$$\geq \ln\frac{\prod_{r=1}^{T}\left|\sigma_\varepsilon^2 + \sigma_\eta^2\left[c_0 + c_1\left(-2\cos\frac{\pi r}{T+1}\right) + c_2\left(-2\cos\frac{\pi r}{T+1}\right)^2\right]^{-1}\right|}{\prod_{r=1}^{T}\left|\sigma_\varepsilon^2 + \sigma_\eta^2\left[d_0 + d_1\left(-2\cos\frac{\pi r}{T+1}\right) + d_2\left(-2\cos\frac{\pi r}{T+1}\right)^2\right]^{-1}\right|}.$$

With substituting $c_0,\ c_1,\ c_2,\ d_0,\ d_1$ and d_2 obtained from (3) as

$$c_0 = \alpha_1^2 + (1+\alpha_2)^2,\ c_1 = \alpha_1(1-\alpha_2),\ c_2 = -\alpha_2$$

$$d_0 = \beta_1^2 + (1+\beta_2)^2,\ d_1 = \beta_1(1-\beta_2),\ d_2 = -\beta_2$$

we have

$$\sum_{r=1}^{T}\left\{\left[\sigma_\varepsilon^2 + \sigma_\eta^2\left[\beta_1^2 + (1+\beta_2)^2 - 2\beta_1(1-\beta_2)\left(\cos\frac{\pi r}{T+1}\right) - 4\beta_2\left(\cos^2\frac{\pi r}{T+1}\right)\right]^{-1}\right]^{-1}\right.$$

$$\left.-\left[\sigma_\varepsilon^2 + \sigma_\eta^2\left[\alpha_1^2 + (1+\alpha_2)^2 - 2\alpha_1(1-\alpha_2)\left(\cos\frac{\pi r}{T+1}\right) - 4\alpha_2\left(\cos^2\frac{\pi r}{T+1}\right)\right]^{-1}\right]^{-1}\right\} z_r^2$$

$$\geq \ln\frac{\prod_{r=1}^{T}\left|\sigma_\varepsilon^2 + \sigma_\eta^2\left[\alpha_1^2 + (1+\alpha_2)^2 - 2\alpha_1(1-\alpha_2)\left(\cos\frac{\pi r}{T+1}\right) - 4\alpha_2\left(\cos^2\frac{\pi r}{T+1}\right)\right]^{-1}\right|}{\prod_{r=1}^{T}\left|\sigma_\varepsilon^2 + \sigma_\eta^2\left[\beta_1^2 + (1+\beta_2)^2 - 2\beta_1(1-\beta_2)\left(\cos\frac{\pi r}{T+1}\right) - 4\beta_2\left(\cos^2\frac{\pi r}{T+1}\right)\right]^{-1}\right|}. \quad (6)$$

3. Simulation Study

The performance of this method was investigated via computer simulations. Two hundred data sets of size 500 were generated for H_1 and H_2 using SPLUS/2000 package. The considered time series models were both $AR(1)$ plus noise and $AR(2)$ plus noise with $\sigma_\eta^2 = 1.0$ and different values of σ_ε^2 and various values of α and β according to H_1 and H_2 models. Then equation (5) or (6) was used, based on which $AR(1)$ or $AR(2)$ had been used, to classify observed time series to one of two models.

Reprehensive results of these simulations are given in Table 1. As can be seen from the results of the table the method works well and the performance would be improved if observation noise variance, σ_η^2, take smaller value.

Table 1. The percentage of misclassification for $AR(1)$ plus noise processes for various parameter values $\alpha(\alpha_1, \alpha_2)$ and $\beta(\beta_1, \beta_2)$. Time series of length 500 generated from $H_1(H_2)$ with $\sigma_\eta^2 = 1.0$ and different values of σ_ε^2

α, β	0	1	2	3	4	5	6	7	8	9	10
(-0.2, -0.6)	1	2.5	3.8	6.5	11.8	12.5	18	22.5	23.8	27.8	40.8
(0.3, 0.5)	0	0	0	0.8	3	4.3	6.5	6.8	12.3	13.8	11.5
(-0.3, -0.5)	0	0	0.5	0.8	2.5	3.3	7.3	6.5	14.5	14	18.8
(-0.1, -0.6)	0	0	0.8	1.5	2.8	4.5	9.5	8.5	12.3	13.5	13.3
(0.2, 0.5)	0.8	1.5	6.2	2	2.8	7.5	9	11.5	12.8	20.5	19.3
(0.1, -0.5)	0	0.5	4.3	6	11.5	15.3	16.5	21.3	25.8	25.5	26
(0.2, -0.4)	7.5	13	19	23.5	25.8	30	30.5	34.5	37	38	40.3
$(\alpha_1, \alpha_2), (\beta_1, \beta_2)$	0	1	2	3	4	5	6	7	8	9	10
(-0.1, -0.2), (-0.2, -0.2)	0.5	12	23	30.3	32	31.8	36.8	46.3	41.8	45.5	46.5
(-0.2, -0.1), (-0.2, 0.2)	0.5	7.8	7	24.5	26.8	33.3	32	32.3	38.5	38.8	41.3
(-0.2, -0.05), (0.2, 0.5)	32.5	37.5	40	44.3	42.3	44	46.5	46	47.8	44.8	48.8
(-0.3, -0.2), (0.3, 0.2)	22	24.8	28.3	29	31.3	33.8	32	38.5	38	38.5	42.5
(-0.4, -0.2), (0.4, 0.2)	13.8	19.5	17.5	19.5	21.8	23	24.5	29.5	21.5	27.8	26.8
(0.3, -0.3), (-0.2, -0.4)	1.3	1.8	4	5.3	10	12.5	7.8	14.5	19.3	23.5	26

4. Distribution of the Discriminant Function

As was mentioned the distribution of the discriminant function when the variances of two populations are different can be expressed in terms of a weighted sum of random variables, each having a chi-square distribution with one degree of freedom, i.e.,

$$d_Q(\mathbf{x}) = \sum_{j=1}^{T} \lambda_j \chi_{1j}^2. \tag{7}$$

So far the weights in the sum are calculated numerically. To give analytically weights for $AR(p)$ plus noise consider $\mathbf{y}$ is from H_1. It means that

$$\mathbf{Y} \sim N(\mathbf{0}, \Sigma_\alpha + \sigma_\varepsilon^2 I)$$

and $\mathbf{z} = L\mathbf{y}$ has multivariate normal distribution with zero mean vector and diagonal covariance matrix with (r, r)th element

$$\sigma_\varepsilon^2 + \sigma_\eta^2 \left[\sum_{j=1}^{p} c_j \left(-2 \cos \frac{\pi r}{T+1} \right)^j \right]^{-1}. \tag{8}$$

The discriminant function between two populations in favor of H_1 after removing constant term is

$$d_Q(\mathbf{z}) = \sum_{r=1}^{T} \left\{ \frac{1}{\sigma_\varepsilon^2 + \sigma_\eta^2 \left[\sum_{j=0}^{p} d_j \left(-2 \cos \frac{\pi r}{T+1} \right)^j \right]^{-1}} - \frac{1}{\sigma_\varepsilon^2 + \sigma_\eta^2 \left[\sum_{j=0}^{p} c_j \left(-2 \cos \frac{\pi r}{T+1} \right)^j \right]^{-1}} \right\} z_r^2.$$

Using (8)

$d_Q(\mathbf{z})$

$$\approx \sum_{r=1}^{T} \left[\left(\sigma_\varepsilon^2 + \frac{\sigma_\eta^2}{\sum_{j=1}^{p} d_j \left(-2\cos \frac{\pi r}{T+1} \right)^j} \right)^{-1} \left(\sigma_\varepsilon^2 + \frac{\sigma_\eta^2}{\sum_{j=1}^{p} c_j \left(-2\cos \frac{\pi r}{T+1} \right)^j} \right) - 1 \right] \chi_{1r}^2, \tag{9}$$

where χ_{1r}^2 are nearly independent chi-squared random variables each with one degree of freedom. Therefore, the rth coefficient of linear combination of χ^2_{1r} is given by

$$\lambda_r = \left(\sigma_\varepsilon^2 + \frac{\sigma_\eta^2}{\sum_{j=1}^{p} c_j\left(-2\cos\frac{\pi r}{T+1}\right)^j}\right)\left(\sigma_\varepsilon^2 + \frac{\sigma_\eta^2}{\sum_{j=1}^{p} d_j\left(-2\cos\frac{\pi r}{T+1}\right)^j}\right)^{-1} - 1.$$

The coefficient when $\mathbf{y}$ is from H_2 is given by

$$\lambda_r = \left[\sigma_\varepsilon^2 + \frac{\sigma_\eta^2}{\sum_{j=1}^{p} c_j\left(-2\cos\frac{\pi r}{T+1}\right)^j}\right]\left[\sigma_\varepsilon^2 + \frac{\sigma_\eta^2}{\sum_{j=1}^{p} d_j\left(-2\cos\frac{\pi r}{T+1}\right)^j}\right]^{-1} - 1. \quad (10)$$

Some Special Results

In the case of discrimination between two pure $AR(p)$ processes, $\sigma_\varepsilon^2 = 0$, the rth coefficient of distribution of discriminant function is

$$\lambda_r = \begin{cases} \dfrac{\sum_{j=0}^{p} d_j\left(-2\cos\frac{\pi r}{T+1}\right)^j}{\sum_{j=0}^{p} c_j\left(-2\cos\frac{\pi r}{T+1}\right)^j} - 1 & \mathbf{y} \in H_1, \\[2ex] 1 - \dfrac{\sum_{j=0}^{p} c_j\left(-2\cos\frac{\pi r}{T+1}\right)^j}{\sum_{j=0}^{p} d_j\left(-2\cos\frac{\pi r}{T+1}\right)^j} & \mathbf{y} \in H_2. \end{cases} \quad (11)$$

For $AR(1)$ plus noise

$$H_1 : x_t + \varepsilon_t, \quad x_t = \alpha x_{t-1} + \eta_t$$

$$H_2 : x_t + \varepsilon_t, \quad x_t = \beta x_{t-1} + \eta_t$$

leads to

$$\lambda_j = \begin{cases} \dfrac{\sigma_\eta^2(\beta-\alpha)\left(\alpha+\beta-2\cos\frac{\pi r}{T+1}\right)}{\left(1+\alpha^2-2\alpha\cos\frac{\pi r}{T+1}\right)\left[\sigma_\eta^2\left(1+\beta^2-2\beta\cos\frac{\pi r}{T+1}\right)+\sigma_\varepsilon^2\right]} & \mathbf{y} \in H_1, \\[2ex] \dfrac{\sigma_\eta^2(\beta-\alpha)\left(\alpha+\beta-2\cos\frac{\pi r}{T+1}\right)}{\left(1+\beta^2-2\beta\cos\frac{\pi r}{T+1}\right)\left[\sigma_\eta^2\left(1+\alpha^2-2\alpha\cos\frac{\pi r}{T+1}\right)+\sigma_\varepsilon^2\right]} & \mathbf{y} \in H_2 \end{cases} \quad (12)$$

and reduces to

$$\lambda_j = \begin{cases} \dfrac{(\beta - \alpha)\left(\alpha + \beta - 2\cos\dfrac{\pi r}{T+1}\right)}{\left(1 + \alpha^2 - 2\alpha\cos\dfrac{\pi r}{T+1}\right)} & \mathbf{y} \in H_1 \\[2ex] \dfrac{(\beta - \alpha)\left(\alpha + \beta - 2\cos\dfrac{\pi r}{T+1}\right)}{\left(1 + \beta^2 - 2\beta\cos\dfrac{\pi r}{T+1}\right)} & \mathbf{y} \in H_2 \end{cases} \tag{13}$$

when discrimination is for two pure $AR(1)$ processes. It should be mentioned that (12) and (13) are the same as given in Chinipardaz [4] and in Chan et al. [2], respectively. In $AR(2)$ plus noise

$$\lambda_j = \begin{cases} \dfrac{\sigma_\eta^2[f(\beta) - f(\alpha)]}{[\sigma_\eta^2 f(\beta) + \sigma_\varepsilon^2] f(\alpha)} & \mathbf{y} \in H_1, \\[2ex] \dfrac{\sigma_\eta^2[f(\beta) - f(\alpha)]}{[\sigma_\eta^2 f(\alpha) + \sigma_\varepsilon^2] f(\beta)} & \mathbf{y} \in H_2, \end{cases} \tag{14}$$

where

$$f(\beta) = \beta_1^2 + (1 + \beta_2^2)^2 - 2\beta_1(1 - \beta_2)\cos\frac{\pi r}{T+1} - 4\beta_2\cos^2\frac{\pi r}{T+1},$$

$$f(\alpha) = \alpha_1^2 + (1 + \alpha_2^2)^2 - 2\alpha_1(1 - \alpha_2)\cos\frac{\pi r}{T+1} - 4\alpha_2\cos^2\frac{\pi r}{T+1}.$$

5. Numerical Comparison of Classical Method and New Method for Weights

The distribution of the discriminant function has been studied by some authors (see for examples Johnson and Kotz [11] and Davies [9]). For example Johnson and Kotz [11] tabulated $\sum_{j=1}^{T} \lambda_j \chi_{1j}^2$ for $T = 5$. This approach has been followed of fitting Pearson curve to intractable distribution of quadratic forms. A Pearson curve is fitted by the using the first four cumulants (see Krishnaiah et al. [13]). Solomon and Stephens [17] showed that the sth cumulants, $\mathcal{K}_s(d_Q(\mathbf{y}))$ under H_i, $i = 1, 2$ is given by

$$\mathcal{K}_s(d_Q(\mathbf{y})) = 2^{s-1}(s-1)!\sum_{j=1}^{T}\lambda_j^s.$$

The weights can be obtained numerically by finding the eigenvalue of $\frac{1}{T}\sum_{j=1}^{T}\left(\sum_2^{-1} - \sum_1^{-1}\right)$, where it depends on whether $\mathbf{y}$ comes from the first or second population. In general, the eigenvalues are very difficult to obtain and require the cumbersome manipulation especially in time series because T has very large dimension. An analytical formula has now been suggested for the weights. The performance of the analytic method is investigated by comparison between two methods. For $T = 100$ and different α and β in $AR(1)$ plus noise process is considered. The various values of σ_ε^2 also is selected to find the effect of observation noise.

Table 2 compares the first four cumulants of $\mathcal{K}_s(d_Q(\mathbf{y}))$ calculating by using (I): λ_j given in (12) and (II): $\lambda_j = j$th eigenvalue of $\frac{1}{T}\sum_{j=1}^{T}\left(\sum_2^{-1} - \sum_1^{-1}\right)$. From the tables it was found that the explicit approximation to the eigenvalues gave the cumulants very close to those using the true values of the cumulants. The cumulants are closer if the variance of the observation noise σ_ε^2 takes small value and two models are more different.

Table 2. Comparison of the first four cumulants of the discriminant function obtained by analytical method and numerical method, given in patronesses, for $AR(1)$ plus noise processes

$\alpha = 0.2, \quad \beta = 0.4$

σ_ϵ^2	$\mathcal{K}_1$	$\mathcal{K}_2$	$\mathcal{K}_3$	$\mathcal{K}_4$
0.0	$4.25 \times 10^{-2} (4.00 \times 10^{-2})$	$1.55 \times 10^{-3} (1.55 \times 10^{-3})$	$5.27 \times 10^{-7} (2.82 \times 10^{-7})$	$4.37 \times 10^{-7} (4.39 \times 10^{-7})$
0.5	$-4.98 \times 10^{-3} (-6.69 \times 10^{-3})$	$7.13 \times 10^{-4} (7.18 \times 10^{-4})$	$-4.20 \times 10^{-6} (-4.30 \times 10^{-6})$	$1.27 \times 10^{-7} (1.29 \times 10^{-7})$
1.0	$-1.66 \times 10^{-2} (-1.79 \times 10^{-2})$	$4.55 \times 10^{-4} (4.58 \times 10^{-4})$	$-3.12 \times 10^{-6} (-3.18 \times 10^{-6})$	$6.37 \times 10^{-8} (6.46 \times 10^{-8})$
2.5	$-2.04 \times 10^{-2} (-2.02 \times 10^{-2})$	$1.94 \times 10^{-4} (1.96 \times 10^{-4})$	$-1.17 \times 10^{-6} (-1.19 \times 10^{-6})$	$1.50 \times 10^{-8} (1.52 \times 10^{-8})$
5.0	$-1.6 \times 10^{-2} (-1.68 \times 10^{-2})$	$8.14 \times 10^{-5} (8.24 \times 10^{-5})$	$-3.56 \times 10^{-7} (-3.62 \times 10^{-7})$	$2.99 \times 10^{-9} (3.03 \times 10^{-9})$

$\alpha = -0.1, \beta = 0.1$

σ_ϵ^2	$\mathcal{K}_1$	$\mathcal{K}_2$	$\mathcal{K}_3$	$\mathcal{K}_4$
0.0	$4.00 \times 10^{-2} (4.00 \times 10^{-2})$	$1.70 \times 10^{-3} (1.70 \times 10^{-3})$	$1.22 \times 10^{-5} (1.22 \times 10^{-5})$	$5.95 \times 10^{-7} (5.96 \times 10^{-7})$
0.5	$1.77 \times 10^{-2} (1.77 \times 10^{-2})$	$7.29 \times 10^{-4} (7.28 \times 10^{-4})$	$2.33 \times 10^{-6} (2.33 \times 10^{-6})$	$1.02 \times 10^{-7} (1.03 \times 10^{-7})$
1.0	$9.97 \times 10^{-3} (9.98 \times 10^{-3})$	$4.05 \times 10^{-4} (4.05 \times 10^{-4})$	$7.33 \times 10^{-7} (7.33 \times 10^{-7})$	$3.10 \times 10^{-8} (3.10 \times 10^{-8})$
2.5	$3.27 \times 10^{-3} (3.27 \times 10^{-3})$	$1.31 \times 10^{-5} (1.31 \times 10^{-5})$	$7.82 \times 10^{-8} (7.82 \times 10^{-8})$	$3.19 \times 10^{-9} (3.19 \times 10^{-9})$
5.0	$1.12 \times 10^{-3} (1.12 \times 10^{-3})$	$4.47 \times 10^{-5} (4.47 \times 10^{-5})$	$9.10 \times 10^{-9} (9.11 \times 10^{-9})$	$3.66 \times 10^{-10} (3.67 \times 10^{-10})$

$\alpha = -0.8, \beta = 0.8$

σ_ϵ^2	$\mathcal{K}_1$	$\mathcal{K}_2$	$\mathcal{K}_3$	$\mathcal{K}_4$
0.0	6.80 (7.04)	6.34 (6.69)	14.92 (15.87)	58.29 (62.35)
0.5	2.49 (2.59)	$9.38 \times 10^{-1} (9.92 \times 10^{-1})$	$8.36 \times 10^{-1} (8.92 \times 10^{-1})$	1.25 (1.34)
1.0	1.47 (1.53)	$3.61 \times 10^{-1} (3.82 \times 10^{-1})$	$1.98 \times 10^{-1} (2.11 \times 10^{-1})$	$1.82 \times 10^{-1} (1.96 \times 10^{-1})$
2.5	$6.07 \times 10^{-1} (6.35 \times 10^{-1})$	$8.02 \times 10^{-2} (8.50 \times 10^{-2})$	$2.00 \times 10^{-2} (2.14 \times 10^{-2})$	$8.59 \times 10^{-3} (9.24 \times 10^{-2})$
5.0	$2.75 \times 10^{-1} (2.88 \times 10^{-1})$	$2.34 \times 10^{-2} (2.48 \times 10^{-2})$	$2.92 \times 10^{-3} (3.14 \times 10^{-3})$	$6.75 \times 10^{-4} (7.27 \times 10^{-4})$

$\alpha = -0.5, \beta = 0.5$

σ_ϵ^2	$\mathcal{K}_1$	$\mathcal{K}_2$	$\mathcal{K}_3$	$\mathcal{K}_4$
0.0	1.31 (1.32)	$1.80 \times 10^{-1} (1.81 \times 10^{-1})$	$4.38 \times 10^{-2} (4.42 \times 10^{-2})$	$1.78 \times 10^{-2} (1.80 \times 10^{-2})$
0.5	$5.38 \times 10^{-1} (5.42 \times 10^{-1})$	$4.32 \times 10^{-2} (4.36 \times 10^{-2})$	$4.65 \times 10^{-3} (4.70 \times 10^{-3})$	$9.05 \times 10^{-4} (9.17 \times 10^{-4})$
1.0	$3.09 \times 10^{-1} (3.12 \times 10^{-1})$	$1.97 \times 10^{-2} (1.99 \times 10^{-2})$	$1.28 \times 10^{-3} (1.30 \times 10^{-3})$	$1.69 \times 10^{-4} (1.71 \times 10^{-4})$
2.5	$1.11 \times 10^{-1} (1.12 \times 10^{-1})$	$5.41 \times 10^{-3} (5.47 \times 10^{-3})$	$1.39 \times 10^{-4} (1.40 \times 10^{-4})$	$1.03 \times 10^{-5} (1.05 \times 10^{-5})$
5.0	$4.15 \times 10^{-2} (4.19 \times 10^{-2})$	$1.80 \times 10^{-3} (1.82 \times 10^{-3})$	$1.83 \times 10^{-5} (1.86 \times 10^{-5})$	$9.81 \times 10^{-7} (9.95 \times 10^{-7})$

$\alpha = 0.2, \beta = 0.7$

σ_ϵ^2	$\mathcal{K}_1$	$\mathcal{K}_2$	$\mathcal{K}_3$	$\mathcal{K}_4$
0.0	$2.62 \times 10^{-1} (2.53 \times 10^{-1})$	$9.64 \times 10^{-3} (9.63 \times 10^{-3})$	$1.82 \times 10^{-4} (1.77 \times 10^{-4})$	$16.66 \times 10^{-6} (16.61 \times 10^{-6})$
0.5	$4.29 \times 10^{-2} (3.65 \times 10^{-2})$	$3.49 \times 10^{-3} (3.53 \times 10^{-3})$	$-3.29 \times 10^{-5} (-3.51 \times 10^{-5})$	$2.85 \times 10^{-6} (2.93 \times 10^{-6})$
1.0	$-1.37 \times 10^{-2} (-1.91 \times 10^{-2})$	$2.35 \times 10^{-3} (2.40 \times 10^{-3})$	$-3.89 \times 10^{-5} (-4.05 \times 10^{-5})$	$1.92 \times 10^{-6} (1.98 \times 10^{-6})$
2.5	$-5.28 \times 10^{-2} (-5.68 \times 10^{-2})$	$1.34 \times 10^{-3} (1.38 \times 10^{-3})$	$-2.64 \times 10^{-5} (-2.74 \times 10^{-5})$	$9.82 \times 10^{-7} (1.02 \times 10^{-6})$
5.0	$-5.54 \times 10^{-2} (-5.84 \times 10^{-2})$	$7.94 \times 10^{-4} (8.20 \times 10^{-4})$	$-1.41 \times 10^{-5} (-1.46 \times 10^{-5})$	$4.29 \times 10^{-7} (4.47 \times 10^{-7})$

$\alpha = -0.3, \beta = -0.6$

σ_ϵ^2	$\mathcal{K}_1$	$\mathcal{K}_2$	$\mathcal{K}_3$	$\mathcal{K}_4$
0.0	$1.01 \times 10^{-1} (9.49 \times 10^{-2})$	$3.37 \times 10^{-3} (3.39 \times 10^{-3})$	$5.27 \times 10^{-6} (3.94 \times 10^{-6})$	$2.07 \times 10^{-9} (2.09 \times 10^{-9})$
0.5	$-2.38 \times 10^{-3} (-6.52 \times 10^{-3})$	$1.56 \times 10^{-3} (1.59 \times 10^{-3})$	$-1.87 \times 10^{-5} (-1.94 \times 10^{-5})$	$7.77 \times 10^{-7} (7.99 \times 10^{-7})$
1.0	$-2.86 \times 10^{-2} (-3.21 \times 10^{-2})$	$1.11 \times 10^{-3} (1.13 \times 10^{-3})$	$-1.61 \times 10^{-5} (-1.66 \times 10^{-5})$	$5.25 \times 10^{-7} (5.41 \times 10^{-7})$
2.5	$-4.27 \times 10^{-2} (-4.51 \times 10^{-2})$	$6.14 \times 10^{-4} (6.29 \times 10^{-4})$	$-8.63 \times 10^{-6} (-8.90 \times 10^{-6})$	$2.18 \times 10^{-7} (2.24 \times 10^{-7})$
5.0	$-3.81 \times 10^{-2} (-3.98 \times 10^{-2})$	$3.28 \times 10^{-4} (3.37 \times 10^{-4})$	$-3.73 \times 10^{-6} (-3.85 \times 10^{-6})$	$7.21 \times 10^{-8} (7.49 \times 10^{-8})$

Acknowledgement

The authors wish to thank Shahid Chamran University for financial support.

References

[1] J. Alagon, Spectral discrimination for two groups of time series, J. Time Ser. Anal. 10(3) (1989), 203-214.

[2] H. T. Chan, R. Chinipardaz and T. F. Cox, Discrimination of AR, MA and ARMA time series models, Comm. Statist. Theory Methods 25(6) (1996), 1247-1260.

[3] G. Chaudhuri and J. D. Borwanker, Bhattacharyya distance-based linear discrimination, J. Indian Statist. Assoc. 29 (1991), 47-56.

[4] R. Chinipardaz, Discrimination of $AR(1)$ plus noise, Iran. J. Sci. Technol. Trans. A 24(2) (2000), 165-172.

[5] G. R. Dargahi-Noubary, Discrimination between Gaussian time series based on their spectral differences, Comm. Statist. Theory Methods 21(9) (1992), 2439-2458.

[6] G. R. Dargahi-Noubary, A linear discriminant function for Gaussian time series, J. Time Ser. Anal. 20(2) (1999a), 144-153.

[7] G. R. Dargahi-Noubary, Time Series with Applications to Seismology, Nova Science Publishers, Inc., 1999b.

[8] G. R. Dargahi-Noubary and P. J. Laycock, Spectral ratio discriminants and information theory, J. Time Ser. Anal. 2 (1981), 71-86.

[9] R. B. Davies, Numerical inversion of a characteristic function, Biometrika 60 (1973), 415.

[10] W. A. Fuller, Introduction to Statistical Time Series, 2nd ed., John Wiley, New York, 1996.

[11] N. L. Johnson and S. Kotz, Distributions in Statistics: Continuous Multivariate Distributions, Vol. 1, John Wiley, New York, 1972.

[12] Y. R. Kakizawa, R. H. Shumway and M. Taniguchi, Discrimination and clustering for multivariate time series, J. Amer. Statist. Assoc. 93 (1998), 328-340.

[13] P. R. Krishnaiah, J. C. Lee and T. C. Chang, The distribution of the likelihood ratio statistics for tests of certain covariance structures of complex multivariate normal populations, Biomertika 63 (1976), 543-549.

[14] R. H. Shumway, Discriminant analysis for time series, Handbook of Statistics, Vol. 2, P. R. Krishnaiah and L. N. Kanal, eds., North-Holland, Amsterdam, 1982, pp. 1-46.

[15] R. H. Shumway, Applied Statistical Time Series Analysis, Englewood Cliffs, Prentice-Hall, 1988.

[16] R. H. Shumway and A. N. Unger, Linear discriminant functions for stationary time series, J. Amer. Statist. Assoc. 72 (1974), 881-885.

[17] H. Solomon and M. A. Stephens, Approximations to density functions using Pearson curves, J. Amer. Statist. Assoc. 73(361) (1978), 153-160.

[18] A. Zyweck and R. Bogner, Radar target classification of commerical aircraft, IEEE Transactions in Aerospace Electronic Systems 32 (1996), 598-606.

■