Ann. Inst. Statist. Math.
Vol. 41, No. 4, 765–779 (1989)

A TEST FOR THE PRESENCE OF PURE FEEDBACK IN MULTIVARIATE DYNAMIC STOCHASTIC SYSTEMS

JUNJI NAKANO* AND SHIGEMI TAGAMI

Technical College, University of Tokushima, 2-1 Minamijosanjima, Tokushima 770, Japan

(Received June 22, 1987; revised October 7, 1988)

Abstract. This paper describes a procedure for testing the presence of a pure feedback loop in a transfer function model for a multivariate discrete dynamic stochastic system. A modification of the portmanteau statistic based on sample cross-covariance matrices of the prewhitened series is proposed. The statistic is shown to be asymptotically distributed according to a χ^2-distribution with certain degrees of freedom under some pure feedback assumptions. Some numerical results are given to show the behavior of the proposed method.

Key words and phrases: Feedback loop, FPE criterion, multivariate autoregressive moving average model, portmanteau statistics, transfer function model.

1. Introduction

Box and Jenkins (1976) proposed "transfer function model fitting" for describing a dynamic relationship between the input and output of a univariate system. They considered that the input and the noise are characterized by scalar autoregressive integrated moving average (ARIMA) models, and employed a three-step procedure (identification, estimation and diagnostic checking) for model building. Tee and Wu (1972) used this technique for investigating paper machine process data. Box and MacGregor (1974), however, pointed out that Box and Jenkins' "open-loop" identification procedure was not adequate for this kind of a "closed-loop" situation. Since the input was the stock gate opening controlled by an experienced operator observing the paper weight output, they suggested the use of a "pure" feedback closed-loop transfer function model, and also proposed a simple procedure for detecting the presence of a feedback loop in a system.

Besides the Box and Jenkins approach, several methods have been

---

*Now at Graduate School of Policy Science, Saitama University, Urawa, Saitama 338, Japan.

proposed for the analysis of multivariate input-output dynamic systems operating in closed loops. Akaike (1971), Phadke and Wu (1974) and Caines and Chan (1975) considered both the input and output processes jointly as the output of a system driven by noise only. The identifiability concerning this approach was extensively studied by Anderson and Gevers (1982). This method is useful for a wide class of general closed-loop systems. It is, however, sometimes too general to be used as a final model for finite data, particularly when the number of observations is not large. It is well known that the model should contain as few parameters as possible for the best accuracy of parameter estimation. Therefore, in cases where we have prior knowledge about a system and are confident that the system has a pure feedback, a parsimonious pure feedback model should be employed instead of other general models. We also note that the transfer function of the pure feedback model is not identifiable by this method because of the singularity of the noise process. Gustavsson *et al.* (1977) surveyed prediction error estimation methods and some identifiability results. Recently, techniques of analyzing linear systems using this method were explained in detail by Ljung (1987). By this method, if the form of the transfer function of a system is determined, parameters can be estimated regardless of whether the data have been collected in an open or a closed loop. However, some nonparametric methods, such as correlation analysis or spectral analysis, are used for specifying the model form. The results of these preliminary analyses should be carefully interpreted depending whether the process is operated under an open or a closed loop.

In any method stated above, it is important to know whether a system consists of an open or a closed loop, at least in the identification step, i.e., the stage to select tentative model forms. Caines and Chan (1975) gave a procedure to test the null hypothesis of an open-loop model against the alternative hypothesis of a general closed-loop model. In this paper, we propose a statistic to test for the presence of the (subset) pure feedback in a general closed-loop model by modifying the portmanteau statistic investigated by Box and Pierce (1970) and Hosking (1980).

In the next section, we describe definitions of three types of dynamic stochastic models and consider the difference among them. We also give the form of a test statistic. In Section 3, the asymptotic distribution of the statistic under the assumption of an autoregressive moving average (ARMA) model is shown to be a χ^2-distribution with certain degrees of freedom. A practical testing procedure and some comments on the usage are given in Section 4. Section 5 contains an analysis of Tee and Wu's paper machine data and some simulation results.

2. Properties of models and definition of a test statistic

We are concerned with systems in which both the input $\{y_t\}$ and the output $\{x_t\}$ are m-variate discrete stationary processes. Suppose that $\{x_t\}$ is dynamically related to $\{y_t\}$ by a transfer function model

$$x_t = H(B)y_t + w_t , \tag{2.1}$$

where $\{w_t\}$ is an m-variate random noise, the backward shift operator B is such that $By_t = y_{t-1}$ and $H(B) = H_0 + H_1B + H_2B^2 + \cdots$ is an $m \times m$ transfer function matrix describing the process characteristics. In general feedback control cases, $\{y_t\}$ is also considered to be written by a transfer function model

$$y_t = K(B)x_t + z_t , \tag{2.2}$$

where $\{z_t\}$ is an m-variate random noise and $K(B) = K_0 + K_1B + K_2B^2 + \cdots$ denotes an $m \times m$ transfer function matrix which characterizes the feedback controller. We assume that noises $\{w_t\}$ and $\{z_t\}$ are stationary processes generated from linear models

$$w_t = M(B)a_t , \tag{2.3}$$

$$z_t = N(B)b_t , \tag{2.4}$$

where $\{a_t\}$ and $\{b_t\}$ are mutually independent m-variate white noise sequences, $M(B) = I_m + M_1B + M_2B^2 + \cdots$ and $N(B) = I_m + N_1B + N_2B^2 + \cdots$ are stationary linear filter matrices, and I_m denotes an $m \times m$ identity matrix. This general closed-loop model, in which both a feedback loop and an input noise exist, is shown diagrammatically in Fig. 1, and has been studied by several authors, for example, Akaike (1971), Phadke and Wu

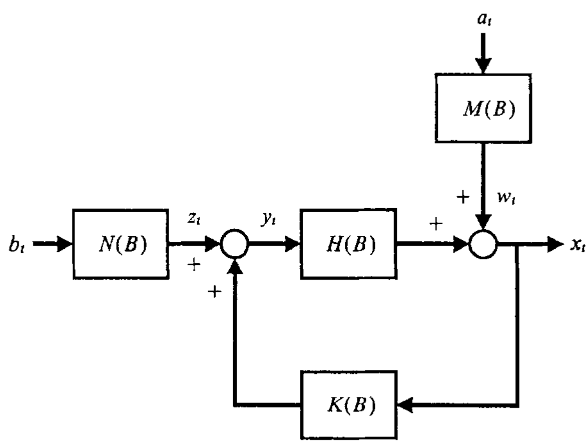

Fig. 1. Closed-loop model.

(1974), Caines and Chan (1975), Gustavsson *et al.* (1977) and Anderson and Gevers (1982).

Box and Jenkins (1976) studied mainly the case where $K(B) = 0$, i.e., the system which has no feedback loop. We define this model as an open-loop model. On the other hand, where $b_t = 0$, i.e., there is no added "dither" noise, is adequate for describing "pure" feedback operating data such as Tee and Wu's paper machine data. This model is defined as a pure feedback model. The open-loop model and the pure feedback model are two extreme cases of the closed-loop model shown by Fig. 1.

We assume that $I_m - H(B)K(B)$ and $I_m - K(B)H(B)$ are invertible. Then, the closed-loop model is expressed as

$$x_t = \{I_m - H(B)K(B)\}^{-1}\{M(B)a_t + H(B)N(B)b_t\}\,, \tag{2.5}$$

$$y_t = \{I_m - K(B)H(B)\}^{-1}\{K(B)M(B)a_t + N(B)b_t\}\,. \tag{2.6}$$

Under very general conditions, a stationary process $\{x_t\}$ can be "prewhitened" by an appropriate linear filter. We denote this prewhitened sequence as $\{u_t\}$. In other words, there exists a linear filter matrix $P(B) = I_m + P_1B + P_2B^2 + \cdots$ such that

$$P(B)x_t = u_t\,, \tag{2.7}$$

where $\{u_t\}$ is an m-variate white noise sequence. Considering equations (2.5) and (2.7), u_t is constructed by linear combinations of a_{t-k} and b_{t-k} $(k \geq 0)$. Similarly, $\{y_t\}$ can be prewhitened by a linear filter to a sequence $\{v_t\}$, and v_t is also represented by other linear combinations of a_{t-k} and b_{t-k} $(k \geq 0)$. Therefore, the covariance matrix between u_t and v_{t-k}, which is denoted by $\mathrm{Cov}\,[u_t, v_{t-k}]$, is generally not equal to zero for $k \geq 0$. The open-loop model also has the same property. The pure feedback model means that $N(B)b_t = 0$. Equations (2.5) and (2.6) then reduce to

$$x_t = \{I_m - H(B)K(B)\}^{-1}M(B)a_t\,, \tag{2.8}$$

$$y_t = \{I_m - K(B)H(B)\}^{-1}K(B)M(B)a_t\,. \tag{2.9}$$

If we define the prewhitened sequences $\{u_t\}$ and $\{v_t\}$ as above, both u_t and v_t are apparently products of constant matrices and a_t. Hence, for $k \geq 1$, $\mathrm{Cov}\,[u_t, v_{t-k}] = 0$. These facts suggest that if estimates of $\mathrm{Cov}\,[u_t, v_{t-k}]$ $(k \geq 1)$ are near to zero, the pure feedback model is appropriate, and otherwise, the closed-loop model or the open-loop model are adequate.

As the number of observations is finite, we usually use parsimonious models such as an ARMA model or an autoregressive (AR) model for prewhitening the data. Once the order of an ARMA or an AR model is decided, parameters are estimated by the maximum likelihood method under the assumption of normality, or by the least squares method. Using these estimated parameters, $\hat{u}_t$ and $\hat{v}_t$, which are estimates of u_t and v_t, can be calculated. Sample cross-covariance and covariance matrices are given by

$$\hat{C}(k) = \frac{1}{N} \sum_{t=1}^{N} \hat{u}_t \hat{v}'_{t-k} , \tag{2.10}$$

$$\hat{\Sigma} = \frac{1}{N} \sum_{t=1}^{N} \hat{u}_t \hat{u}'_t , \tag{2.11}$$

and

$$\hat{T} = \frac{1}{N} \sum_{t=1}^{N} \hat{v}_t \hat{v}'_t . \tag{2.12}$$

Here, $\hat{C}(k)$ is an estimate of Cov $[u_t, v_{t-k}]$.

As stated above, the hypothesis of the pure feedback model can be checked by examining whether $\hat{C}(k)$ is near to zero or not for $k \geq 1$. In a similar situation concerning univariate processes, Box and Jenkins (1976) proposed a quality-control-chart-type of approach, and Box and Pierce (1970) proposed the use of the "portmanteau" statistic. We employ the latter approach for testing the pure feedback model. Considering the multivariate portmanteau statistic introduced by Hosking (1980), we define a statistic

$$S = N \sum_{k=1}^{d} \mathrm{Tr}\, \{\hat{C}(k)' \hat{\Sigma}^{-1} \hat{C}(k) \hat{T}^{-1}\} , \tag{2.13}$$

where d is an adequately chosen integer and Tr C denotes the sum of the diagonal elements of a square matrix C. Following the lines of McLeod (1978), the asymptotic properties of this statistic are derived in the next section under the assumption that the pure feedback model equations (2.8) and (2.9) have stationary ARMA forms.

3. The asymptotic distribution of the test statistic

In this section, we deal with the case where the pure feedback model equations (2.8) and (2.9) are reduced to ARMA models

$$A(B)x_t = B(B)Ua_t , \tag{3.1}$$

$$E(B)y_t = Z(B)Va_t , \tag{3.2}$$

where $A(B)=\sum_{i=0}^{p} A_i B^i$, $B(B)=\sum_{i=0}^{q} B_i B^i$, $E(B)=\sum_{i=0}^{r} E_i B^i$ and $Z(B)=\sum_{i=0}^{s} Z_i B^i$ with $A_0=B_0=E_0=Z_0=I_m$, U and V are upper triangular regular matrices, and $\{a_t\}$ is a mutually independent m-variate white noise sequence with mean 0, covariance matrix I_m and bounded fourth moments. We assume that the orders p, q, r and s are known and all the roots of $|A(z)|=0$, $|B(z)|=0$, $|E(z)|=0$ and $|Z(z)|=0$ lie outside the unit circle $|z|=1$, i.e., $\{x_t\}$ and $\{y_t\}$ are stationary and invertible processes. We also assume that identifiability conditions such as rank $[A_p\ \ B_q]=$ rank $[E_r\ \ Z_s]=m$ are satisfied; see Hannan (1969).

Suppose that the data $\{x_t, y_t\}$ are observed at $t=1,\dots,N$. We define

$$\dot{u}_t = x_t + \dot{A}_1 x_{t-1} + \cdots + \dot{A}_p x_{t-p} - \dot{B}_1 \dot{u}_{t-1} - \cdots - \dot{B}_q \dot{u}_{t-q}\,, \tag{3.3}$$

where $\dot{A}_1,\dots,\dot{A}_p,\dot{B}_1,\dots,\dot{B}_q$ are arbitrary $m\times m$ matrices in the admissible parameter space and $\dot{u}_k=x_k=0$ for $k\le 0$. We can estimate parameters $A_1,\dots,A_p,B_1,\dots,B_q,U$ by $\hat{A}_1,\dots,\hat{A}_p,\hat{B}_1,\dots,\hat{B}_q,\hat{U}$, which minimize the value of $\log|\dot{\Sigma}|+(1/N)\sum_{t=1}^{N}\dot{u}_t'\dot{\Sigma}^{-1}\dot{u}_t$ where $\dot{\Sigma}=\dot{U}\dot{U}'$. These estimates are approximate maximum likelihood estimates if we assume the normality of $\{a_t\}$. The estimate $\hat{u}_t$ of Ua_t is given by replacing $\dot{A}_1,\dots,\dot{A}_p,\dot{B}_1,\dots,\dot{B}_q$ in (3.3) by $\hat{A}_1,\dots,\hat{A}_p,\hat{B}_1,\dots,\hat{B}_q$. Parameters $E_1,\dots,E_r,Z_1,\dots,Z_s$ and V are also estimated by minimizing $\log|\dot{T}|+(1/N)\sum_{t=1}^{N}\dot{v}_t'\dot{T}^{-1}\dot{v}_t$, where $\dot{T}=\dot{V}\dot{V}'$,

$$\dot{v}_t = y_t + \dot{E}_1 y_{t-1} + \cdots + \dot{E}_r y_{t-r} - \dot{Z}_1 \dot{v}_{t-1} - \cdots - \dot{Z}_s \dot{v}_{t-s} \tag{3.4}$$

and $\dot{v}_k=y_k=0$ for $k\le 0$. The estimate $\hat{v}_t$ of Va_t is similarly defined. Then $\mathrm{Cov}\,[Ua_t, Va_{t-k}]$ is estimated by (2.10).

We introduce some definitions and properties of Kronecker products to express the results concisely. If C is an $m\times n$ matrix whose (i,j)-th element is c_{ij}, vec C is defined by vec $C=[c_{11},\dots,c_{m1},c_{12},\dots,c_{m2},\dots,c_{1n},\dots,c_{mn}]'$. If D is a $p\times q$ matrix, $C\otimes D$ denotes the $mp\times nq$ Kronecker product whose (i,j)-th submatrix is $c_{ij}D$. If A, B and C are matrices such that the matrix product ABC is defined, $\mathrm{vec}\,(ABC)=(C'\otimes A)\,\mathrm{vec}\,B$ can be proved.

We investigate the asymptotic distribution of $\hat{c}_d=\mathrm{vec}\,[\hat{C}(1)\cdots\hat{C}(d)]$. Let $\dot{\theta}=\mathrm{vec}\,[\dot{A}_1\cdots\dot{A}_p\ \ \dot{B}_1\cdots\dot{B}_q]$, $\dot{\lambda}=\mathrm{vec}\,[\dot{E}_1\cdots\dot{E}_r\ \ \dot{Z}_1\cdots\dot{Z}_s]$, $\dot{\xi}=[\dot{\theta}',\dot{\lambda}']'$ and $c_d,\dot{c}_d,\theta,\hat{\theta},\lambda,\hat{\lambda},\xi,\hat{\xi}$ are similarly defined. Using Taylor's theorem and the result that $\hat{\xi}-\xi=O_p(1/\sqrt{N})$ (Hosoya and Taniguchi (1982)), we can obtain an approximate linear expansion of $\hat{c}_d$ as

$$\hat{c}_d = c_d + \left.\frac{\partial \dot{c}_d}{\partial \dot{\theta}'}\right|_{\dot{\xi}=\xi}(\hat{\theta}-\theta) + \left.\frac{\partial \dot{c}_d}{\partial \dot{\lambda}'}\right|_{\dot{\xi}=\xi}(\hat{\lambda}-\lambda) + O_p\left(\frac{1}{N}\right), \tag{3.5}$$

where O_p denotes order in probability; see Fuller (1976).

LEMMA 3.1. *The terms* $\partial \dot{c}_d / \partial \dot{\theta}' |_{\dot{\xi}=\xi}$ *and* $\partial \dot{c}_d / \partial \dot{\lambda}' |_{\dot{\xi}=\xi}$ *converge in probability to* $\{I_d \otimes (VU') \otimes I_m\} X$ *and* 0, *respectively, where* $X = [X_1' X_2' \cdots X_d']'$,

$$X_j = [\Psi_{j-1} \cdots \Psi_{j-p} \quad -I_m \otimes \Pi_{j-1} \cdots -I_m \otimes \Pi_{j-q}] , \tag{3.6}$$

$$B(B)^{-1} = \Pi(B) = \Pi_0 + \Pi_1 B + \Pi_2 B^2 + \cdots , \tag{3.7}$$

$$A(B)^{-1} B(B) = \Phi(B) = \Phi_0 + \Phi_1 B + \Phi_2 B^2 + \cdots , \tag{3.8}$$

$$\Psi_i = \sum_{j=0}^{\infty} \Phi_{i-j}' \otimes \Pi_j , \tag{3.9}$$

and all the elements of Π_j, Φ_j *and* Ψ_j *are defined to zero for* $j < 0$.

PROOF. Let $\dot{\alpha} = \operatorname{vec} [\dot{A}_1 \cdots \dot{A}_p]$, $\dot{\beta} = \operatorname{vec} [\dot{B}_1 \cdots \dot{B}_q]$ and $\dot{c}(k) = \operatorname{vec} \dot{C}(k)$. We can write

$$\frac{\partial \dot{c}_d}{\partial \dot{\theta}'} = \begin{bmatrix} \dfrac{\partial \dot{c}(1)}{\partial \dot{\alpha}'} & \dfrac{\partial \dot{c}(1)}{\partial \dot{\beta}'} \\ \vdots & \vdots \\ \dfrac{\partial \dot{c}(d)}{\partial \dot{\alpha}'} & \dfrac{\partial \dot{c}(d)}{\partial \dot{\beta}'} \end{bmatrix} , \tag{3.10}$$

and one of submatrices is written as

$$\frac{\partial \dot{c}(k)}{\partial \dot{\alpha}'} = \frac{1}{N} \sum_{t=1}^{N} \dot{v}_{t-k} \otimes \frac{\partial \dot{u}_t}{\partial \dot{\alpha}'} , \tag{3.11}$$

for $k = 1, \ldots, d$. After some calculations, we have

$$\frac{\partial \dot{u}_t}{\partial \dot{\alpha}'} = \sum_{j=0}^{\infty} [x_{t-1-j}' \cdots x_{t-p-j}'] \otimes \dot{\Pi}_j . \tag{3.12}$$

Therefore, we obtain

$$\begin{aligned} \left. \frac{\partial \dot{c}(k)}{\partial \dot{\alpha}'} \right|_{\dot{\xi}=\xi} &= \frac{1}{N} \sum_{t=1}^{N} v_{t-k} \otimes \left\{ \sum_{j=0}^{\infty} [x_{t-1-j}' \cdots x_{t-p-j}'] \otimes \Pi_j \right\} \\ &= \sum_{j=0}^{\infty} \left[\frac{1}{N} \sum_{t=1}^{N} v_{t-k} \otimes x_{t-1-j}' \cdots \frac{1}{N} \sum_{t=1}^{N} v_{t-k} \otimes x_{t-p-j}' \right] \otimes \Pi_j . \end{aligned} \tag{3.13}$$

Using the equation $x_{t-h-j} = \sum_{l=0}^{\infty} \Phi_l u_{t-h-j-l}$, we have

$$(3.14)\qquad \frac{1}{N}\sum_{t=1}^{N} v_{t-k}\otimes x'_{t-h-j}=\sum_{l=0}^{\infty}\left(\frac{1}{N}\sum_{t=1}^{N} v_{t-k}u'_{t-h-j-l}\right)\Phi'_l\,.$$

Since it is easily verified that $u_t=\dot{u}_t|_{\dot{\theta}=\theta}$ and $v_t=\dot{v}_t|_{\dot{\lambda}=\lambda}$ have the same asymptotic properties as Ua_t and Va_t, respectively, we have

$$(3.15)\qquad \frac{1}{N}\sum_{t=1}^{N} v_{t-k}u'_t=\delta_k VU'+O_p\left(\frac{1}{\sqrt{N}}\right),$$

where

$$(3.16)\qquad \delta_k=\begin{cases}1, & k=0\,,\\ 0, & \text{otherwise}\,.\end{cases}$$

Thus, we obtain

$$(3.17)\qquad \left.\frac{\partial\dot{c}(k)}{\partial\dot{\alpha}'}\right|_{\dot{\xi}=\xi}=\sum_{j=0}^{\infty}[VU'\Phi'_{k-1-j}\cdots VU'\Phi'_{k-p-j}]\otimes\Pi_j+O_p\left(\frac{1}{\sqrt{N}}\right)$$

$$=\{(VU')\otimes I_m\}[\Psi_{k-1}\cdots\Psi_{k-p}]+O_p\left(\frac{1}{\sqrt{N}}\right).$$

Note that Φ_h and Π_h decay exponentially to zero as h goes to infinity.

Following a similar argument, we obtain

$$(3.18)\qquad \left.\frac{\partial\dot{c}(k)}{\partial\dot{\beta}'}\right|_{\dot{\xi}=\xi}=\{(VU')\otimes I_m\}[-I_m\otimes\Pi_{k-1}\cdots-I_m\otimes\Pi_{k-q}]$$

$$+O_p\left(\frac{1}{\sqrt{N}}\right).$$

These results show that $\partial\dot{c}(k)/\partial\dot{\theta}'|_{\dot{\xi}=\xi}$ converges in probability to

$$\{(VU')\otimes I_m\}[\Psi_{k-1}\cdots\Psi_{k-p}\quad -I_m\otimes\Pi_{k-1}\cdots-I_m\otimes\Pi_{k-q}]$$
$$=\{(VU')\otimes I_m\}X_k\,.$$

$\partial\dot{c}(k)/\partial\dot{\lambda}'|_{\dot{\xi}=\xi}$ is shown to converge to 0 in the same way.

LEMMA 3.2. *Asymptotically,*

$$(3.19)\qquad \sqrt{N}(\hat{\theta}-\theta)=F^{-1}w\,,$$

where

$$F = \sum_{j=1}^{\infty} X_j'(\Sigma \otimes \Sigma^{-1})X_j, \tag{3.20}$$

X_j is as Lemma 3.1, *and*

$$w = -\frac{1}{\sqrt{N}} \sum_{t=1}^{N} \frac{\partial \dot{u}_t'}{\partial \dot{\theta}} \bigg|_{\dot{\theta}=\theta} \Sigma^{-1} u_t \tag{3.21}$$

are asymptotically distributed with mean 0 *and covariance matrix F.*

PROOF. Estimates $\hat{\theta}$ and $\hat{U}$ must satisfy the equations

$$\frac{1}{N} \sum_{t=1}^{N} \frac{\partial \dot{u}_t'}{\partial \dot{\theta}} \dot{\Sigma}^{-1} \dot{u}_t \bigg|_{\dot{\theta}=\hat{\theta},\, \dot{U}=\hat{U}} = 0 \tag{3.22}$$

and

$$\hat{U}\hat{U}' = \hat{\Sigma} = \frac{1}{N} \sum_{t=1}^{N} \hat{u}_t \hat{u}_t', \tag{3.23}$$

from their definitions. As it is verified that $\hat{\Sigma} = \Sigma + O_p(1/\sqrt{N})$ (Hosoya and Taniguchi (1982)), we have

$$\begin{aligned} &\sqrt{N}(\hat{\theta} - \theta) \\ &= \left\{ \frac{1}{N} \sum_{t=1}^{N} \frac{\partial}{\partial \dot{\theta}'} \left(\frac{\partial \dot{u}_t'}{\partial \dot{\theta}} \Sigma^{-1} \dot{u}_t \right) \bigg|_{\dot{\theta}=\theta} \right\}^{-1} \left\{ -\frac{1}{\sqrt{N}} \sum_{t=1}^{N} \frac{\partial \dot{u}_t'}{\partial \dot{\theta}} \bigg|_{\dot{\theta}=\theta} \Sigma^{-1} u_t \right\} \\ &\quad + O_p\left(\frac{1}{\sqrt{N}} \right) \end{aligned} \tag{3.24}$$

by expanding (3.22).

We can show that $(1/N) \sum_{t=1}^{N} \partial\{(\partial \dot{u}_t'/\partial \dot{\theta})\Sigma^{-1}\dot{u}_t\}/\partial \dot{\theta}'|_{\dot{\theta}=\theta}$ converges to F. The (i,j)-th element of $\partial\{(\partial \dot{u}_t'/\partial \dot{\theta})\, \Sigma^{-1}\dot{u}_t\}/\partial \dot{\theta}'$ is written as $(\partial^2 \dot{u}_t'/\partial \dot{\theta}_j \partial \dot{\theta}_i)\Sigma^{-1}\dot{u}_t + (\partial \dot{u}_t'/\partial \dot{\theta}_i)\Sigma^{-1}(\partial \dot{u}_t/\partial \dot{\theta}_j)$, where $\dot{\theta}_i$ denotes the i-th element of $\dot{\theta}$. From the proof of Lemma 3.1, elements of $\partial^2 \dot{u}_t'/\partial \dot{\theta}_j \partial \dot{\theta}_i$ are written by linear combinations of x_{t-j} and u_{t-j} for $j \geq 1$. Hence, summing over t and dividing by N, we have asymptotically that the (i,j)-th element of $(1/N) \sum_{t=1}^{N} \partial\{(\partial \dot{u}_t'/\partial \dot{\theta})\Sigma^{-1}\dot{u}_t\}/\partial \dot{\theta}'|_{\dot{\theta}=\theta}$ is $(1/N) \sum_{t=1}^{N} (\partial \dot{u}_t'/\partial \dot{\theta}_i)\Sigma^{-1}(\partial \dot{u}_t/\partial \dot{\theta}_j)|_{\dot{\theta}=\theta}$. Also from the proof of Lemma 3.1, we have

$$(3.25)\quad \frac{1}{N}\sum_{t=1}^{N}\frac{\partial\dot{u}_t'}{\partial\dot{\theta}}\Sigma^{-1}\frac{\partial\dot{u}_t}{\partial\dot{\theta}'}\bigg|_{\dot{\theta}=\theta}$$

$$=\frac{1}{N}\sum_{t=1}^{N}\left(\sum_{j=0}^{\infty}[x_{t-1-j}'\cdots x_{t-p-j}' \quad -u_{t-1-j}'\cdots -u_{t-q-j}']\otimes\Pi_j\right)'$$

$$\cdot\Sigma^{-1}\left(\sum_{l=0}^{\infty}[x_{t-1-l}'\cdots x_{t-p-l}' \quad -u_{t-1-l}'\cdots -u_{t-q-l}']\otimes\Pi_l\right)$$

$$=\sum_{j=0}^{\infty}\sum_{l=0}^{\infty}\left\{\frac{1}{N}\sum_{t=1}^{N}[x_{t-1-j}'\cdots x_{t-p-j}' \quad -u_{t-1-j}'\cdots -u_{t-q-j}']'\right.$$

$$\left.\cdot[x_{t-1-l}'\cdots x_{t-p-l}' \quad -u_{t-1-l}'\cdots -u_{t-q-l}']\right\}\otimes(\Pi_j'\Sigma^{-1}\Pi_l)\,.$$

After some algebra, this expression is shown to converge to

$$(3.26)\quad \sum_{j=1}^{\infty}[\Psi_{j-1}\cdots\Psi_{j-p} \quad -I_m\otimes\Pi_{j-1}\cdots -I_m\otimes\Pi_{j-q}]'(\Sigma\otimes\Sigma^{-1})$$

$$\cdot[\Psi_{j-1}\cdots\Psi_{j-p} \quad -I_m\otimes\Pi_{j-1}\cdots -I_m\otimes\Pi_{j-q}]$$

$$=\sum_{j=1}^{\infty}X_j'(\Sigma\otimes\Sigma^{-1})X_j=F\,.$$

Considering the equation

$$(3.27)\quad -\frac{1}{\sqrt{N}}\sum_{t=1}^{N}\frac{\partial\dot{u}_t'}{\partial\dot{\theta}}\bigg|_{\dot{\theta}=\theta}\Sigma^{-1}u_t$$

$$=-\frac{1}{\sqrt{N}}\sum_{t=1}^{N}\left\{\sum_{j=0}^{\infty}\begin{bmatrix}x_{t-1-j}\\ \vdots\\ x_{t-p-j}\\ -u_{t-1-j}\\ \vdots\\ -u_{t-q-j}\end{bmatrix}\otimes\Pi_j'\right\}\Sigma^{-1}u_t\,,$$

the asymptotic normality of this term is proved by an argument similar to that of the proof of central limit theorem for m-dependent random variables (see Fuller (1976)), but details are omitted here. We note that the expectation of this term is 0 and the covariance matrix is shown to be equal to F similar to the above calculations.

From Lemmas 3.1 and 3.2, we have asymptotically

$$(3.28)\quad \hat{c}_d=[\{I_d\otimes(VU')\otimes I_m\}XF^{-1} \quad I_{m^2d}]\begin{bmatrix}\frac{1}{\sqrt{N}}w\\ c_d\end{bmatrix}.$$

LEMMA 3.3. *The joint asymptotic distribution of* $\begin{bmatrix} w \\ \sqrt{N}\, c_d \end{bmatrix}$ *is normal with mean* 0 *and covariance matrix*

$$\begin{bmatrix} F & -X'\{I_d \otimes (UV') \otimes I_m\} \\ -\{I_d \otimes (VU') \otimes I_m\}X & I_d \otimes T \otimes \Sigma \end{bmatrix},$$

where F *and* X *are as in Lemmas* 3.1 *and* 3.2.

PROOF. The asymptotic normality of c_d was essentially shown by Chitturi (1976). Considering the relation

$$\begin{bmatrix} w \\ \sqrt{N}\, c_d \end{bmatrix} = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} \begin{bmatrix} \left(-\sum_{j=0}^{\infty} \begin{bmatrix} x_{t-1-j} \\ \vdots \\ x_{t-p-j} \\ -u_{t-1-j} \\ \vdots \\ -u_{t-q-j} \end{bmatrix} \otimes \Pi_j' \right) \Sigma^{-1} \\ \begin{bmatrix} v_{t-1} \\ \vdots \\ v_{t-d} \end{bmatrix} \otimes I_m \end{bmatrix} u_t, \tag{3.29}$$

we can calculate the covariance matrix of this random vector after some similar algebra as given above.

THEOREM 3.1. *The statistics* S *defined by* (2.13) *is asymptotically distributed according to a* χ^2*-distribution with* $m^2(d-p-q)$ *degrees of freedom for sufficiently large* d.

PROOF. Lemma 3.3 and (3.28) show that $\sqrt{N}\,\hat{c}_d$ is asymptotically normally distributed with mean 0 and covariance matrix $(I_d \otimes V \otimes U) \cdot (I_{m^2d} - YF^{-1}Y')(I_d \otimes V' \otimes U')$ where $Y = (I_d \otimes U' \otimes U^{-1})X$. We assume that d is sufficiently large so that $F = \sum_{j=1}^{\infty} X_j'(\Sigma \otimes \Sigma^{-1})X_j$ is approximated by $\sum_{j=1}^{d} X_j'(\Sigma \otimes \Sigma^{-1})X_j = Y'Y$. Then $\sqrt{N}\,(I_d \otimes V^{-1} \otimes U^{-1})\hat{c}_d$ is asymptotically normally distributed with mean 0 and covariance matrix $I_{m^2d} - Y(Y'Y)^{-1}Y'$. Considering that $I_{m^2d} - Y(Y'Y)^{-1}Y'$ is a symmetric idempotent matrix with rank $m^2(d-p-q)$, the asymptotic distribution of

$$\begin{aligned} &\{\sqrt{N}\,(I_d \otimes V^{-1} \otimes U^{-1})\hat{c}_d\}'\{\sqrt{N}\,(I_d \otimes V^{-1} \otimes U^{-1})\hat{c}_d\} \\ &\quad = N\hat{c}_d'(I_d \otimes T^{-1} \otimes \Sigma^{-1})\hat{c}_d = N\sum_{k=1}^{d} \mathrm{Tr}\,\{\hat{C}(k)'\Sigma^{-1}\hat{C}(k)T^{-1}\} \end{aligned} \tag{3.30}$$

is a χ^2-distribution with $m^2(d-p-q)$ degrees of freedom. Σ and T can be replaced by $\hat{\Sigma}$ and $\hat{T}$ without changing the asymptotic distribution.

Note that the degrees of freedom of the χ^2-distribution do not depend on r and s, which are the ARMA orders of the input $\{y_t\}$.

4. Practical procedure and some remarks

The results of Section 3 are directly applicable to testing for the presence of pure feedback if data are known to be generated from ARMA models. However, we rarely have such information about a process. Most real data are considered to be generated from general linear stationary processes, which are represented by infinite order AR models. We also know that the identification and the estimation of multivariate ARMA or MA models require complicated calculations. For these reasons, appropriate finite order AR models may be used for approximating real processes.

Once AR models are assumed, we can use some order selection criteria, for example, the FPE criterion (Akaike (1971)) or the AIC criterion (Akaike (1973)). If we consider the estimated order of a model to be the true order, we can use the results of the above theorem. Therefore, the test for the presence of a pure feedback loop in process dynamics may be achieved by the following procedure:

1. For each sequence of the m-variate observed input $\{y_t\}$ and output $\{x_t\}$, fit an AR model by the FPE (or AIC, etc.) criterion. We denote the estimated autoregressive order of $\{x_t\}$ as p.
2. Using the fitted AR models, prewhiten the input and output sequences.
3. Calculate the statistic S using sample cross-covariance matrices of prewhitened series. We choose the value of d appropriately.
4. Compare the value of S with a significance point (5% or 1%, say) of a χ^2-distribution with $m^2(d-p)$ degrees of freedom. If the value of S is greater than the significance point, we can conclude that a pure feedback model is inappropriate. Otherwise, the model is not inappropriate.

In some environments including many process control cases whose feedback controller is integrating, the input and/or the output are not stationary but homogeneous nonstationary; i.e., some suitable differences of the data are stationary. Then we should use $\{(1-B)^{d_1}y_t\}$ and $\{(1-B)^{d_2}x_t\}$, instead of the original $\{y_t\}$ and $\{x_t\}$, for the above procedure. Here, nonnegative integers d_1 and d_2 may be determined in order to make all the elements of $\{(1-B)^{d_1}y_t\}$ and $\{(1-B)^{d_2}x_t\}$ stationary, following the methodology proposed by Box and Jenkins (1976).

Our procedure is also applicable to testing a subset of pure feedback loops in a general closed-loop model. Let the input $\{y_t\}$ and the output $\{x_t\}$

be stationary and partitioned into subvectors such as $x_t = [x'_{1t}, x'_{2t}]'$ and $y_t = [y'_{1t}, y'_{2t}]'$, where dimensions of x_{2t} and y_{2t} are identical, say, r. The hypothesis of pure feedback from $\{x_{2t}\}$ to $\{y_{2t}\}$ is given by equations (2.1) and

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} K_{11}(B) & K_{12}(B) \\ 0 & K_{22}(B) \end{bmatrix} \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} + \begin{bmatrix} z_{1t} \\ 0 \end{bmatrix}, \tag{4.1}$$

where $K_{22}(B)$ is an $r \times r$ transfer function matrix and the dimension of $\{z_{1t}\}$ is as same as that of $\{y_{1t}\}$. Note that (4.1) is a restricted version of the closed-loop model (2.2). The hypothesis of subset pure feedback can be tested by applying the above procedure to the input $\{y_{2t}\}$ and the output $\{x_{2t}\}$.

5. Numerical results

We first analyzed Tee and Wu's paper machine data. Their data consist of 160 observations on the stock gate opening $\{y_t\}$ and the paper weight deviation $\{x_t\}$. We used the FPE criterion with maximum autoregressive order 10. For the (mean deleted) input $\{y_t\}$, the AR(1) model

$$(1 - 0.895B)y_t = v_t \tag{5.1}$$

was selected. The output $\{x_t\}$ was decided to be a white noise sequence. Setting $d = 30$, the value of S is 38.94. As this value is less than the 5% significance point of the χ^2-distribution with degrees of freedom 30 ($= 30 - 0$), we can conclude that the pure feedback model is not inappropriate for this data.

Some simulation experiments were done to see the behavior of the proposed statistic. We considered the closed-loop model defined by

$$x_t = \begin{bmatrix} 1+0.7B & -0.2B \\ 0.4B & 1+0.8B \end{bmatrix}^{-1} \begin{bmatrix} 0.4-0.24B & -0.16B \\ 0.12B & 0.4+0.2B \end{bmatrix} y_t + \begin{bmatrix} 1-0.7B & -0.2B \\ -0.3B & 1-0.6B \end{bmatrix}^{-1} \begin{bmatrix} 1+0.9B & 0.7B \\ 0 & 1-0.8B \end{bmatrix} a_t, \tag{5.2}$$

$$y_t = \begin{bmatrix} 1-0.6B & 0.3B \\ -0.6B & 1-0.7B \end{bmatrix}^{-1} \begin{bmatrix} 0.2-0.16B & -0.1B \\ 0.04B & 0.2-0.14B \end{bmatrix} x_t + vb_t, \tag{5.3}$$

where $\{a_t\}$ and $\{b_t\}$ were normally distributed with mean 0 and covariance matrix I_2. The stationary nature of processes can be easily verified. For some values of ν, observations of length 200 were generated for 100 replications. For each replication, S and its degrees of freedom were calculated for $d = 30$ using the FPE criterion with maximum autoregressive order 10. The results are summarized in Table 1.

Table 1. Results of simulation.

ν	S		D.F.		Number	
	Mean	Variance	Mean	Variance	5%	1%
0	98.13	285.35	103.24	36.43	4	2
0.2	108.01	204.42	103.60	34.42	8	1
0.4	117.84	234.20	103.04	24.93	27	6
0.6	124.90	213.33	102.44	33.58	39	19
0.8	130.63	195.33	102.20	34.06	63	27
1.0	135.17	198.45	101.96	34.10	74	40
2.0	147.22	196.08	103.48	17.67	94	65

First column of Table 1 shows the value of ν. The sample mean of values of S and their sample variance are given in the second and third columns. The sample mean of values of degrees of freedom and their sample variance were given in the fourth and fifth columns. The last two columns show the numbers of S whose values were greater than 5% or 1% significance points of the χ^2-distributions with corresponding degrees of freedom. In the case of $\nu = 0$, the process follows a pure feedback model. Simulation results show that the asymptotic result described in Section 3 holds well in this case. As the value of ν becomes large, the hypothesis of pure feedback tends to be rejected. These results show an example of the effect of feedback noise $\{z_t\}$ on the test procedure.

The same experiments were carried out for an open-loop model written by (5.2) and

$$y_t = \begin{bmatrix} 1 - 0.8B & 0.3B \\ -0.3B & 1 - 0.7B \end{bmatrix}^{-1} \begin{bmatrix} 1 - 0.7B & 0 \\ -0.5B & 1 - 0.8B \end{bmatrix} b_t \,. \tag{5.4}$$

The sample mean of values of S and their sample variance were 152.23 and 279.47 and the sample mean of values of degrees of freedom and their sample variance were 101.40 and 30.51. The numbers of S which were greater than 5% and 1% significance points of χ^2-distribution were 98 and 82. We can see that there is little possibility of wrongly accepting the hypothesis of a pure feedback model for this open-loop model.

Acknowledgements

The authors are grateful to the two referees for their helpful comments on an earlier version of this paper.

REFERENCES

Akaike, H. (1971). Autoregressive model fitting for control, *Ann. Inst. Statist. Math.*, **23**, 163–180.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle, *2nd International Symposium on Information Theory*, (eds. B. N. Petrov and F. Csáki), 267–281, Akadémiai Kiadó, Budapest.

Anderson, B. D. O. and Gevers, M. R. (1982). Identifiability of linear stochastic systems operating under linear feedback, *Automatica—J. IFAC*, **18**, 195–213.

Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*, 2nd ed., Holden-Day, San Francisco.

Box, G. E. P. and MacGregor, J. F. (1974). The analysis of closed-loop dynamic-stochastic systems, *Technometrics*, **16**, 391–398.

Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models, *J. Amer. Statist. Assoc.*, **65**, 1509–1526.

Caines, P. E. and Chan, C. W. (1975). Feedback between stationary stochastic processes, *IEEE Trans. Automat. Control*, **20**, 498–508.

Chitturi, R. V. (1976). Distribution of multivariate white noise autocorrelations, *J. Amer. Statist. Assoc.*, **71**, 223–226.

Fuller, W. A. (1976). *Introduction to Statistical Time Series*, Wiley, New York.

Gustavsson, I., Ljung, L. and Söderström, T. (1977). Identification of processes in closed loop—Identifiability and accuracy aspects, *Automatica—J. IFAC*, **13**, 59–75.

Hannan, E. J. (1969). The identification of vector mixed autoregressive-moving average systems, *Biometrika*, **56**, 223–225.

Hosking, J. R. M. (1980). The multivariate portmanteau statistic, *J. Amer. Statist. Assoc.*, **75**, 602–608.

Hosoya, Y. and Taniguchi, M. (1982). A central limit theorem for stationary processes and the parameter estimation of linear processes, *Ann. Statist.*, **10**, 132–153.

Ljung, L. (1987). *System Identification: Theory for the User*, Prentice-Hall, Englewood Cliffs, New Jersey.

McLeod, A. I. (1978). On the distribution of residual autocorrelations in Box-Jenkins models, *J. Roy. Statist. Soc. Ser. B*, **40**, 296–302.

Phadke, M. S. and Wu, S. M. (1974). Identification of multiinput-multioutput transfer function and noise model of a blast furnace from closed-loop data, *IEEE Trans. Automat. Control*, **19**, 944–951.

Tee, L. H. and Wu, S. M. (1972). An application of stochastic and dynamic models for the control of a papermaking process, *Technometrics*, **14**, 481–496.