Invariance of the likelihood-ratio test for time-varying autoregressive (TVAR) model order estimation under the $H_0$ hypothesis

Yuri Abramovich$^1$ and Nicholas Spencer$^2$

$^1$Defence Science & Technology Organisation (DSTO), Adelaide, Australia
$^2$Adelaide Research & Innovation Pty Ltd (ARI), Adelaide, Australia

N. Spencer is supported by DSTO through the Centre of Expertise in Phased Array and Microwave Radar Systems (CEPAMiR), established in partnership between DSTO and the University of Adelaide, Australia.

1 Introduction

Methods for order estimation and parameter estimation of a stationary autoregressive (AR) model of order $m$, AR($m$), given a set of $T$ independent identically distributed (i.i.d.) $N$-variate Gaussian samples, are well established. These methods usually require a complete identification of the model, which can be achieved only approximately in the maximum-likelihood (ML) sense. In some applications, the strict stationarity of the observed training data is questionable, so a general problem is to select between a stationary and a time-varying model.

2 Order estimation of an AR/TVAR model

In Abramovich et al. (2007a) we demonstrated that the necessary and sufficient condition for an $N$-variate complex vector $x \equiv [x_1, \ldots, x_N]^T$ to be a sample of the TVAR($m$) process

$$x_j = \sum_{k=1}^m a_{kj} x_{j-k} + \eta_j \quad \text{for} \quad j = m+1, \ldots, N, \quad \mathcal{E}\{\eta_j \eta_k^*\} = \sigma_0^2 \delta_{jk}, \quad a_0 = 1 \quad (1)$$

is that its positive-definite (p.d.) Hermitian covariance matrix $R_N^{(m)} \equiv \mathcal{E}\{xx^H\}$ satisfies the “band-inverse” property

$$\{[R_N^{(m)}]^{-1}\}_{jk} = 0 \quad \text{for} \quad |j - k| > m \quad (2)$$

ie. the elements of its inverse are zero outside the $(2m + 1)$-wide diagonal band. Since the p.d. Toeplitz covariance matrix of the stationary AR($m$) model has this same property, a test for “band-inverseness” may be seen as a unified test for estimating the model order $m$, irrespective of its stationary or time-varying nature. Such a test can be constructed from the properties of the probability density function (p.d.f.) of a certain likelihood ratio (LR).

We consider a set of $T$ i.i.d. $N$-variate training data

$$x_j \equiv [x_1^{(j)}, \ldots, x_N^{(j)}]^T \quad \text{for} \quad j = 1, \ldots, T \quad (3)$$

that are samples of a complex Gaussian random process whose p.d.f. is $CN(0, R_N)$, where $R_N$ is an $N$-variate p.d. Hermitian matrix. The conventional sample covariance matrix $\hat{R} = \frac{1}{T} \sum_{j=1}^T x_j x_j^H$ is rank deficient for $T < N$, and the matrix $T \hat{R}$ is described by the anti-Wishart
complex distribution $ACW(T < N, N, R_N)$ (Janik & Nowak, 2003). Yet, for $T \geq m + 1$, all $(m + 1)$-variate central block matrices $\hat{R}_q^{(m)}$ of $\hat{R}$ are p.d. (Anderson, 1958), ie.

$$\hat{R}_q^{(m)} = \begin{bmatrix} \hat{r}_{qq} & \cdots & \hat{r}_{q,q+m} \\ \vdots & \ddots & \vdots \\ \hat{r}_{q+m,q} & \cdots & \hat{r}_{q+m,q+m} \end{bmatrix} > 0 \quad \text{for } q = 1, \ldots, N - m. \quad (4)$$

In Abramovich et al. (2007a) we demonstrated that this condition is necessary and sufficient for the existence of an accurate nondegenerate ML estimate of a TVAR($m$) covariance matrix that is calculated directly from the blocks $\hat{R}_q^{(m)}$ using the Dym–Gohberg formula (Dym & Gohberg, 1981):

$$\hat{R}_q^{(m)} = [\hat{V}^{(m)H}]^{-1}[\hat{V}^{(m)}]^{-1} \quad (5)$$

where $\hat{V}^{(m)}$ is a lower-triangular matrix whose elements are defined as

$$\hat{V}_{ij}^{(m)} \equiv \begin{cases} \hat{v}_{ij}^{(m)} \hat{v}_{jj}^{(m)} & \text{for } j \leq i \leq L(j) \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where

$$\begin{bmatrix} \hat{v}_{qq}^{(m)} \\ \vdots \\ \hat{v}_{L(q),q}^{(m)} \end{bmatrix} = \begin{bmatrix} \hat{r}_{qq} & \cdots & \hat{r}_{q,L(q)} \\ \vdots & \ddots & \vdots \\ \hat{r}_{L(q),q} & \cdots & \hat{r}_{L(q),L(q)} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \equiv [\hat{R}_q^{(L)}]^{-1} e_{L(q)-q+1} \quad (7)$$

are the time-varying AR coefficients, with $L(q) = \min\{N, q + m\}$, and $e_z$ is the $z$-variate unit vector. This ML TVAR($m$) covariance matrix is uniquely specified by the remarkable properties

$$\left\{\begin{array}{c} \hat{R}_q^{(m)} \\ \hat{R}_q^{(m)} \end{array}\right\}_{ij} = \hat{r}_{ij} \quad \text{for } |i - j| \leq m$$

$$\left\{\begin{array}{c} \hat{R}_q^{(m)} \\ \hat{R}_q^{(m)} \end{array}\right\}_{ij}^{-1} = 0 \quad \text{for } |i - j| > m. \quad (8)$$

For a stationary AR($m$) model, the inverse of its ML-optimal Toeplitz covariance matrix estimate is also a band matrix, like $[\hat{R}_q^{(m)}]^{-1}$, and of the same bandwidth, but its elements cannot be directly and simply obtained from $\hat{R}$; only numerical solutions are currently available for ML Toeplitz covariance matrix estimation, and so suboptimal solutions are usually suggested (Grigoriadis et al., 1994).

For the ML estimate $\hat{R}_q^{(m)}$, the Gaussian likelihood function

$$LF(X, R) = \frac{\exp[- \text{tr}(TR^{-1}\hat{R})]}{[\det \hat{R}]^T} \quad (9)$$

evaluates to

$$\max_{\hat{R}} LF(X, R) = LF(X, \hat{R}_q^{(m)}) = \exp[-NT] \left( \prod_{q=1}^{N} \hat{v}_{qq}^{(m)} \right)^T \quad (10)$$

where

$$\hat{v}_{qq}^{(m)} = e_{L(q)-q+1}^T [\hat{R}_q^{(L)}]^{-1} e_{L(q)-q+1}. \quad (11)$$
In fact, (8) follows directly from the ML equation \( \partial \log LF(X, R) / \partial (R^{-1})_{ij} = 0 \) subject to the TVAR\((m)\) constraint \( (R^{-1})_{ij} = 0 \) for \( |i - j| > m \). According to (8), for \( \hat{B} \equiv [\hat{R}_{TVAR}^{(m)}]^{-1} \) we get

\[
\text{tr}[\hat{B} \hat{R}] = \text{tr}[\hat{B} \hat{R}_{TVAR}^{(m)}] = NT, \quad \therefore \quad \det \hat{B} = \det[\hat{V}^{(m)} H \hat{V}^{(m)}] = \prod_{q=1}^{N} \hat{v}_{qq}^{(m)} \tag{12}
\]

by (5). Let \( m_{\text{max}} \) be the maximum admissible order of a TVAR\((m)\) model that is identifiable for the sample volume \( T \), then \( m_{\text{max}} + 1 \leq T \). From the “nested” property of the model-order testing problem, and directly from (10), it is evident that

\[
LF(X, \hat{R}_{TVAR}^{(m)}) \geq LF(X, \hat{R}_{TVAR}^{(m_2)}) \quad \text{for} \quad m_1 > m_2 \tag{13}
\]

and so our hypothesis test for the TVAR\((m)\) order can be based on the likelihood ratio

\[
LR(m) = \frac{\max_{\mu \leq m_{\text{max}}} LF[X, \hat{R}_{TVAR}^{(\mu)}]}{LF[X, \hat{R}_{TVAR}^{(m_{\text{max}})}]} = \left( \prod_{q=1}^{N} \frac{\hat{v}_{qq}^{(m_{\text{max}})}}{\hat{v}_{qq}^{(m)}} \right)^T \tag{14}
\]

where, according to (7) and (11),

\[
\hat{v}_{qq}^{(m_{\text{max}})} \equiv e_{K(q)-q+1}^{T} [\hat{R}_{K}^{(m)}]^{-1} e_{K(q)-q+1} \tag{15}
\]

with \( K(q) \equiv \min\{N, q + m_{\text{max}}\} \). Now since \( \hat{v}_{qq}^{(m)} = \hat{v}_{qq}^{(m_{\text{max}})} \) for \( q \geq N - m \), the LR is

\[
LR(m) = \left( \prod_{q=1}^{N-m-1} \frac{\hat{v}_{qq}^{(m)}}{\hat{v}_{qq}^{(m_{\text{max}})}} \right)^T. \tag{16}
\]

Note that the dimension of the matrix \( \hat{R}_{K}^{(m)} \) is

\[
\text{dim} \hat{R}_{K}^{(m)} = \begin{cases} m_{\text{max}} + 1 & \text{for} \quad q \leq N - m_{\text{max}} \\ m + 2 & \text{for} \quad q = N - m_{\text{max}} - 1. \end{cases} \tag{17}
\]

We introduce the notation

\[
\mu \equiv \begin{cases} m_{\text{max}} & \text{for} \quad q < N - m_{\text{max}} \\ N - q & \text{for} \quad N - m_{\text{max}} \leq q \leq N - m - 1. \end{cases} \tag{18}
\]

Instead of \( LR(m) \), we can deal with \( LR_0 \equiv [LR]^{1/T} \); and so we investigate the p.d.f. of the LR

\[
LR_0(m) = \prod_{q=1}^{N-m-1} \frac{\hat{v}_{qq}^{(m)}}{\hat{v}_{qq}^{(\mu)}}. \tag{19}
\]

**Theorem 1.** Let \( m_0 \) be the true order of the AR or TVAR input data, then for all \( m \geq m_0 \), the p.d.f. of \( LR_0(m) \) does not depend on scenario, i.e. is completely defined by the parameters \( \{N, T, m_{\text{max}}, m\} \). Specifically,

\[
f(x) = C(N, T, m_{\text{max}}, m) x^{(T-m_{\text{max}}-1)} G_{(N-m-1),(N-m-1)}^{(N-m-1)} \left( x \left| \begin{array}{c} \mu \max_{m} \vdots \max_{m} \\ m_{\max} - 1 \vdots 0, \ldots, 0 \end{array} \right. \right) \tag{20}
\]
where $G^{a,b}_{c,d}(\cdot)$ is Meijer’s $G$-function (Gradshteyn & Ryzhik, 2000), and

$$C(N,T,m_{\text{max}},m) = \prod_{q=1}^{N-m-1} \frac{\Gamma(T-m)}{\Gamma(T-\mu)}.$$  \hspace{1cm} (21)

The $p^{th}$ moment of $LR_0(m \geq m_0)$ is

$$\mathcal{E}\{x^p\} = \prod_{q=1}^{N-m-1} \frac{\Gamma(T-m)\Gamma(T-\mu+p)}{\Gamma(T-\mu)\Gamma(T-m+p)}.$$  \hspace{1cm} (22)

Moreover, the p.d.f. of $LR_0(m)$ can be expressed as the p.d.f. of a product of $(N - m - 1)$ independent random numbers $\beta_q$:

$$LR_0(m) = \prod_{q=1}^{N-m-1} \beta_q, \quad \text{with} \quad \beta_q \sim \frac{\beta_q^{(T-\mu-1)}(1-\beta_q)^{(\mu-m-1)}}{B[\mu-m,T-\mu]}.$$  \hspace{1cm} (23)

For the proof, see Appendix I of Abramovich et al. (2007b).

### 3 Final remarks

The “scenario-invariant” nature of the LR allows us to derive a test to estimate the order of an AR or TVAR model et al. (2007b). Computationally, it may be easier to deal with the expression (23) than the somewhat obscure analytic formula (20).

### References


