Online Companion to the Paper 'Traffic Modeling with Phase-Type Distributions and VARMA Processes'

Jan Kriege and Peter Buchholz

Department of Computer Science, TU Dortmund
{jan.kriege,peter.buchholz}@tu-dortmund.de

Abstract. This note includes some additional details as an extension to the paper "Traffic Modeling with Phase-Type Distributions and VARMA Processes" [1].

1 Introduction

This online companion adds some technical details for the ideas presented in the paper that could not be included due to space restrictions. In this respect it is only fully understandable with the knowledge of the original paper [1].
In Sect. 2 we summarize the steps to obtain an equivalent $VAR(1)$ representation for a $VARMA(p,q)$ process. The approach presented in [1] required the $VARMA(p,q)$ process to have some specific properties, e.g. to have standard normal marginal distributions. In Sect. 3 we explain how we can obtain this property by adjusting the covariance matrix of the innovations when constructing a $VARMA(p,q)$ process. Finally, in Sect. 4 we describe the initialization of $VARMA(p,q)$ processes for random number generation.

2 $\boldsymbol{VAR(1)}$ Representation of $\boldsymbol{VARMA(p,q)}$ Processes

A $VARMA(p,q)$ process

$$\boldsymbol{Z}_t = \boldsymbol{\alpha}_1 \boldsymbol{Z}_{t-1} + \boldsymbol{\alpha}_2 \boldsymbol{Z}_{t-2} + \ldots + \boldsymbol{\alpha}_p \boldsymbol{Z}_{t-p} + \boldsymbol{\beta_1 \epsilon_{t-1}} + \boldsymbol{\beta_2 \epsilon_{t-2}} + \ldots + \boldsymbol{\beta_q \epsilon_{t-q}} + \boldsymbol{\epsilon_t}$$

can be transformed into an equivalent $VAR(1)$ process $\tilde{\boldsymbol{Z}}_t = \tilde{\boldsymbol{\alpha}} \tilde{\boldsymbol{Z}}_{t-1} + \tilde{\boldsymbol{\epsilon}}_{\boldsymbol{t}}$ by setting [2]

$$\tilde{\boldsymbol{Z}}_t = \begin{bmatrix} \boldsymbol{Z}_t \\ \vdots \\ \boldsymbol{Z}_{t-p+1} \\ \boldsymbol{\epsilon_t} \\ \vdots \\ \boldsymbol{\epsilon_{t-q+1}} \end{bmatrix}_{(k(p+q)\times 1)} \qquad \tilde{\boldsymbol{\epsilon}}_{\boldsymbol{t}} = \begin{bmatrix} \boldsymbol{\epsilon_t} \\ 0 \\ \vdots \\ 0 \\ \boldsymbol{\epsilon_t} \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{(kp+kq\times 1)} \qquad \tilde{\boldsymbol{\alpha}} = \begin{bmatrix} \tilde{\boldsymbol{\alpha}}_{11} & \tilde{\boldsymbol{\alpha}}_{12} \\ \tilde{\boldsymbol{\alpha}}_{21} & \tilde{\boldsymbol{\alpha}}_{22} \end{bmatrix}_{(k(p+q)\times k(p+q))} \tag{1}$$

where

$$\tilde{\boldsymbol{\alpha}}_{11} = \begin{bmatrix} \boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_{p-1} & \boldsymbol{\alpha}_p \\ \boldsymbol{I}_k & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & \cdots & \boldsymbol{I}_k & 0 \end{bmatrix}_{(kp \times kp)} \qquad \tilde{\boldsymbol{\alpha}}_{12} = \begin{bmatrix} \boldsymbol{\beta}_1 & \cdots & \boldsymbol{\beta}_{q-1} & \boldsymbol{\beta}_q \\ 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix}_{(kp \times kq)} \tag{2}$$

$$\tilde{\boldsymbol{\alpha}}_{21} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}_{(kq \times kp)} \qquad \tilde{\boldsymbol{\alpha}}_{22} = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ \boldsymbol{I}_k & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & \cdots & \boldsymbol{I}_k & 0 \end{bmatrix}_{(kq \times kq)}$$

and $\boldsymbol{I}_k$ is the $(k \times k)$ identity matrix. For the special case $p = 0$ we set $p = 1$ and introduce an artificial matrix $\boldsymbol{\alpha}_1 = \mathbf{0}$. If $q = 0$ we have that $\tilde{\boldsymbol{\alpha}} = \tilde{\boldsymbol{\alpha}}_{11}$. Finally, the covariance matrix of the innovations is set to

$$\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} & 0 & \cdots & 0 & \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} \\ 0 & 0 & \vdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \vdots & 0 & 0 \\ \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} & 0 & \cdots & 0 & \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} \end{bmatrix}_{(k(p+q) \times k(p+q))} . \tag{3}$$

The covariance matrix at lag 0 of the $VAR(1)$ process $\tilde{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{Z}}}(0)$ then contains the covariance matrices of the original $VARMA(p,q)$ process and can be obtained by the relation [2]

$$\begin{aligned} & vec(\tilde{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{Z}}}(0)) \\ &= vec \begin{bmatrix} \boldsymbol{\Sigma}_Z(0) & \boldsymbol{\Sigma}_Z(1) & \cdots & \boldsymbol{\Sigma}_Z(p-1) & E[\boldsymbol{Z}_t \boldsymbol{\epsilon}_t'] & E[\boldsymbol{Z}_t \boldsymbol{\epsilon}_{t-1}'] & \cdots & E[\boldsymbol{Z}_t \boldsymbol{\epsilon}_{t-q+1}'] \\ \boldsymbol{\Sigma}_Z(-1) & \boldsymbol{\Sigma}_Z(0) & \cdots & \boldsymbol{\Sigma}_Z(p-2) & 0 & E[\boldsymbol{Z}_{t-1} \boldsymbol{\epsilon}_{t-1}'] & \cdots & E[\boldsymbol{Z}_{t-1} \boldsymbol{\epsilon}_{t-q+1}'] \\ \vdots & \vdots & \ddots & \vdots & \vdots & & & \vdots \\ \boldsymbol{\Sigma}_Z(-p+1) & \boldsymbol{\Sigma}_Z(-p+2) & \cdots & \boldsymbol{\Sigma}_Z(0) & 0 & 0 & & E[\boldsymbol{Z}_{t-p+1} \boldsymbol{\epsilon}_{t-q+1}'] \\ & & & & \boldsymbol{\Sigma}_\epsilon & 0 & \cdots & 0 \\ & & \cdot & & 0 & \boldsymbol{\Sigma}_\epsilon & 0 & \\ & & & \vdots & & \ddots & \vdots & \\ & & & 0 & & \vdots & \boldsymbol{\Sigma}_\epsilon & \end{bmatrix} \\ &= (\boldsymbol{I}_{k^2(p+q)^2} - \tilde{\boldsymbol{\alpha}} \otimes \tilde{\boldsymbol{\alpha}})^{-1} vec(\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}). \end{aligned}$$

3 Adjusting the Covariance Matrix of the Innovations for $VARMA(p,q)$ Fitting

When minimizing [1, Eq. 16] using the Nelder-Mead algorithm one obtains matrices $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_p$ and $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_q$ of dimension $k \times k$ with AR and MA

coefficients, respectively, in each step of the algorithm. To compute the value of the goal function from [1, Eq. 16] the autocorrelations of the $VARMA(p,q)$ process determined by those matrices have to be computed. For this the covariance matrix of the innovations $\boldsymbol{\Sigma}_\epsilon$ is needed. Additionally, we want the $VARMA(p,q)$ process to have standard normal marginal distributions, i.e. its autocovariance matrix $\boldsymbol{\Sigma}_Z(0)$ at lag 0 should have ones in the diagonal. In the following we construct $\boldsymbol{\Sigma}_\epsilon$ such that this requirement is fulfilled. Recall, that we restricted $\boldsymbol{\Sigma}_\epsilon$ to be a diagonal matrix.
In Sect. 2 we have seen that we can construct an equivalent $VAR(1)$ representation with coefficient matrix $\tilde{\boldsymbol{\alpha}}$ for the $VARMA(p,q)$ process and then we have

$$\underbrace{(\boldsymbol{I}_{k^2(p+q)^2} - \tilde{\boldsymbol{\alpha}} \otimes \tilde{\boldsymbol{\alpha}})^{-1}}_{\boldsymbol{A}} \underbrace{vec(\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}})}_{\boldsymbol{x}} = \underbrace{vec(\tilde{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{Z}}}(0))}_{\boldsymbol{b}} \tag{4}$$

which defines a system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}$. Matrix $\boldsymbol{A}$ is known, i.e. it is defined by the coefficients $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_p$ and $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_q$ (cf. Eqs. 1 and 2). For some entries of vector $\boldsymbol{b}$ the values are known, i.e. the entries corresponding to the k diagonal elements of $\boldsymbol{\Sigma}_Z(0)$ should be 1. Vector $\boldsymbol{x}$ is unknown and has to be determined, but since most of its entries are 0 it can be reduced in size. In fact, according to Eq. 3 vector $\boldsymbol{x}$ has only k unknown elements (i.e. the diagonal elements of $\boldsymbol{\Sigma}_\epsilon$), each appearing four times. Consequently, we can define a system of k linear equations with k unknowns by using a subset of rows and columns from $\boldsymbol{Ax} = \boldsymbol{b}$ that correspond to the diagonal elements in $\boldsymbol{\Sigma}_Z(0)$ and the non-zero elements in $\boldsymbol{x}$. We will denote this system $\boldsymbol{A}_\epsilon \boldsymbol{x}_\epsilon = \boldsymbol{b}_\epsilon$. The solution vector $\boldsymbol{x}_\epsilon$ then provides the diagonal elements of $\boldsymbol{\Sigma}_\epsilon$ that result in standard normal marginal distributions and complete our $VARMA(p,q)$ description.

3.1 Example:

We demonstrate the construction of $\boldsymbol{A}_\epsilon \boldsymbol{x}_\epsilon = \boldsymbol{b}_\epsilon$ by means of a small example. Let $k = 2, p = q = 1$ and

$$\boldsymbol{\alpha}_1 = \begin{bmatrix} 0.12 & 0.21 \\ 0.15 & 0.22 \end{bmatrix}, \qquad \boldsymbol{\beta}_1 = \begin{bmatrix} 0.1 & 0.2 \\ 0.25 & 0.3 \end{bmatrix}.$$

Then we have to determine

$$\boldsymbol{\Sigma}_\epsilon = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}.$$

According to Eqs. 1, 2 and 3 we have the following equivalent $VAR(1)$ representation:

$$\tilde{\boldsymbol{\alpha}} = \left[\begin{array}{cc|cc} \multicolumn{2}{c|}{\overbrace{\hphantom{0.12 \quad 0.21}}^{\tilde{\boldsymbol{\alpha}}_{11}}} & \multicolumn{2}{c}{\overbrace{\hphantom{0.10 \quad 0.20}}^{\tilde{\boldsymbol{\alpha}}_{12}}} \\ 0.12 & 0.21 & 0.10 & 0.20 \\ 0.15 & 0.22 & 0.25 & 0.30 \\ \hline 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 \\ \multicolumn{2}{c|}{\underbrace{\hphantom{0.00 \quad 0.00}}_{\tilde{\boldsymbol{\alpha}}_{21}}} & \multicolumn{2}{c}{\underbrace{\hphantom{0.00 \quad 0.00}}_{\tilde{\boldsymbol{\alpha}}_{22}}} \end{array}\right] \quad \text{and} \quad \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} = \begin{bmatrix} \sigma_1 & 0 & \sigma_1 & 0 \\ 0 & \sigma_2 & 0 & \sigma_2 \\ \sigma_1 & 0 & \sigma_1 & 0 \\ 0 & \sigma_2 & 0 & \sigma_2 \end{bmatrix}$$

The autocovariance matrix of the $VAR(1)$ process has the form

$$\tilde{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{Z}}}(0) = \begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

since we require the two elements corresponding to the diagonal of $\boldsymbol{\Sigma}_Z(0)$ to be 1. After vectorization the elements are at positions 1 and 6. Consequently, we take the first and sixth row of $\boldsymbol{A} = (\boldsymbol{I}_{k^2(p+q)^2} - \tilde{\boldsymbol{\alpha}} \otimes \tilde{\boldsymbol{\alpha}})^{-1}$ as they are used for the computation of those elements in Eq. 4, i.e.

$$\begin{aligned} \boldsymbol{A}(1,:) = (&1.0168, 0.0289, 0.0154, 0.0286, 0.0289, 0.0499, 0.0263, 0.0491, 0.0154, \\ &0.0263, 0.0147, 0.0264, 0.0286, 0.0491, 0.0264, 0.0486) \\ \boldsymbol{A}(6,:) = (&0.0255, 0.0377, 0.0416, 0.0506, 0.0377, 1.0557, 0.0614, 0.0748, 0.0416, \\ &0.0614, 0.0681, 0.0827, 0.0506, 0.0748, 0.0827, 0.1006)\,. \end{aligned}$$

Then we have

$$\boldsymbol{A}(1,:) \cdot \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} = 1 \quad \text{and} \quad \boldsymbol{A}(6,:) \cdot \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} = 1.$$

This can be further simplified, since we have various entries equal to 0 in $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}$ and because each σ_i appears several times in $\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}}$. We get

$$\begin{aligned} &(1.0168 + 0.0154 + 0.0154 + 0.0147)\sigma_1 + (0.0499 + 0.0491 + 0.0491 + 0.0486)\sigma_2 = 1 \\ &(0.0255 + 0.0416 + 0.0416 + 0.0681)\sigma_1 + (1.0557 + 0.0748 + 0.0748 + 0.1006)\sigma_2 = 1. \end{aligned}$$

Finally, we have

$$\begin{gathered} \boldsymbol{A}_\epsilon \boldsymbol{x}_\epsilon = \boldsymbol{b}_\epsilon \\ \begin{bmatrix} 1.0623 & 0.1967 \\ 0.1768 & 1.3059 \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{gathered}$$

which yields the solution $\sigma_1 = 0.8201$ and $\sigma_2 = 0.6548$.

4 Initialization of $VARMA(p, q)$ Processes for Random Number Generation

Since for the generation of random samples from the $VARMA(p, q)$ (cf. [1, Sect. 3.5]) it is assumed that p previous observations and q previous innovations already exist, it might be desirable to initialize the process with reasonable values to start in a stationary state. E.g. for generating the first sample $\boldsymbol{z}_1$ we require realizations for the (virtual) previous samples $\boldsymbol{z}_0, \boldsymbol{z}_{-1}, \ldots, \boldsymbol{z}_{-p+1}$ and the innovations $\boldsymbol{\epsilon}_0, \boldsymbol{\epsilon}_{-1}, \ldots, \boldsymbol{\epsilon}_{-q+1}$. The $\boldsymbol{\epsilon}_i$ are independent and can be determined as described in [1, Sect. 3.5], i.e. by setting $\boldsymbol{\epsilon_i} = \boldsymbol{S}\boldsymbol{v}_i$ where $\boldsymbol{S}$ and $\boldsymbol{v}_i$ are defined as in [1, Sect. 3.5]. Since the $\boldsymbol{z}_i$ are correlated we use matrix $\boldsymbol{\Sigma}_{\boldsymbol{Z}}$ from

[1, Eq. 15] that contains the covariance values of the $VARMA(p, q)$ process for the determination of the $\boldsymbol{z}_i$. Again, we apply a Cholesky decomposition to get $\boldsymbol{P_Z P'_Z} = \boldsymbol{\Sigma_Z}$ and obtain $(\boldsymbol{z}'_0, \boldsymbol{z}'_{-1}, \ldots, \boldsymbol{z}'_{-p+1}) = \boldsymbol{P_Z}(v_1, \ldots, v_{kp})'$ where the v_i are again random numbers with standard normal distribution [2].

References

1. J. Kriege and P. Buchholz. Traffic Modeling with Phase-Type Distributions and VARMA Processes. In *Proc. of the 13th International Conference on Quantitative Evaluation of SysTems (QEST)*, 2016.
2. H. Lütkepohl. *Introduction to Multiple Time Series Analysis*. Springer, 1993.