Journal of Multivariate Analysis 101 (2010) 1532–1545

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Generating random AR(p) and MA(q) Toeplitz correlation matrices

Chi Tim Ng *, Harry Joe

Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong
Department of Statistics, University of British Columbia, Canada

ARTICLE INFO

Article history:
Received 26 September 2009
Available online 2 February 2010

AMS subject classifications:
15B05
15B52
62M10

Keywords:
Autoregressive process
Beta distribution
Longitudinal data
Moving average process

ABSTRACT

Methods are proposed for generating random $(p+1)\times(p+1)$ Toeplitz correlation matrices that are consistent with a causal AR(p) Gaussian time series model. The main idea is to first specify distributions for the partial autocorrelations that are algebraically independent and take values in $(-1, 1)$, and then map to the Toeplitz matrix. Similarly, starting with pseudo-partial autocorrelations, methods are proposed for generating $(q+1)\times(q+1)$ Toeplitz correlation matrices that are consistent with an invertible MA(q) Gaussian time series model. The density can be uniform or non-uniform over the space of autocorrelations up to lag p or q, or over the space of autoregressive or moving average coefficients, by making appropriate choices for the densities of the (pseudo)-partial autocorrelations. Important intermediate steps are the derivations of the Jacobians of the mappings between the (pseudo)-partial autocorrelations, autocorrelations and autoregressive/moving average coefficients. The random generating methods are useful for models with a structured Toeplitz matrix as a parameter.

1. Introduction

Methods for generating random correlation matrices (positive definite with 1s on the diagonal) of dimension $d\times d$ have been considered in the recent papers [1,2]. Toeplitz $d\times d$ correlation matrices are special cases with $d-1$ distinct correlations, one for each diagonal away from the main diagonal. There are statistical models with structured correlation matrices; for example, in longitudinal data analysis, Toeplitz matrices with AR(p) and MA(m) (or m-dependence) structure are used as model parameters—see [3], and PROC GEMMOD and PROC MIXED in SAS. Toeplitz matrices are also used in signal processing.

With motivation from signal processing, an early paper of generating random Toeplitz matrices is by Holmes [4], with a main approach of random Gram matrices. This approach does not allow analysis of distribution theory of the random autocorrelations. In the time series literature, Jones [5] has provided a simple algorithm for generating $\boldsymbol{\phi}_p=(\phi_1,\ldots,\phi_p)^T$ with uniform distribution over C_p, the space for the coefficients of a causal AR(p) Gaussian time series. By using the results of 1–1 correspondence between the autoregressive coefficients and the partial autocorrelations [6], Jones proposed an algorithm based on random partial autocorrelations with independent Beta random variables on the interval $(-1, 1)$. By mapping to the space of autocorrelation coefficients, this algorithm implies a distribution in the space of Toeplitz correlation matrices that is not uniform.

For signal processing applications, Makhoul [7] was interested in the size of the space of Toeplitz matrices and derived the volume of $d\times d$ Toeplitz matrices, but did not show how his theory can also be used to generate random Toeplitz matrices. By combining results in [7,2], one can generate random $d\times d$ Toeplitz matrices, uniform over the space of such matrices, based

* Corresponding author at: Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong.
E-mail address: machitim@inet.polyu.edu.hk (C.T. Ng).

0047-259X/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2010.01.013

on random partial autocorrelations that are independent Beta random variables on the interval $(-1, 1)$. The parameters of the Beta distributions are different from the case in the preceding paragraph.

In this paper, we propose methods of generating a random $\boldsymbol{\phi}_p$ over C_p and a random $\boldsymbol{\rho}_p = (\rho_1, \ldots, \rho_p)^T$ for the first p autocorrelations of an AR(p) Gaussian time series; the distributions can be uniform or have another simple density form. We show that the two approaches are quite different in the marginal moments of the $\{\phi_j\}$ or $\{\rho_j\}$. This means that the appropriate method depends on the context.

With motivation from statistical models that use structured correlation matrices of the MA(q) form, we also extend our results to generating random MA(q) Toeplitz matrices. This requires the derivation of some interesting connections of various parametrizations of Gaussian time series.

The key steps are (i) algebraic independence of partial autocorrelations in $(-1, 1)^p$ for AR(p), (ii) algebraic independence of pseudo-partial autocorrelations in $(-1, 1)^q$ for MA(q), (iii) derivation of the Jacobians of the transformations from the (pseudo)-partial autocorrelation space to the space of autocorrelation, autoregressive or moving average coefficients. That is, we take advantage of (pseudo)-partial autocorrelations being algebraically independent, because the constraints on autocorrelation, autoregressive or moving average coefficients are complicated for $p > 2$ and $q > 2$.

There are many applications for random Toeplitz matrices, especially for models or methods with a parameter that is an AR(p) or MA(q) correlation matrix. For simulation studies to compare estimation methods, it would be useful to be able to simulate a random structured correlation matrix that is (i) uniform over the relevant space or (ii) non-uniform with more concentration nearer the identity matrix or stronger dependence. Also cases (i) and (ii) would be relevant as possible priors for Bayesian inference in these models, and our algorithms can be implemented for MCMC. For Bayesian inference, knowing what the various distribution imply for marginal distributions is important for choosing a prior that matches historical information. Daniels and Pourahmadi [8] mention shrinkage of correlation and partial correlations toward zero near as being reasonable for Bayesian priors (in order to reduce Bayes risk); in their application, if stationarity is assumed, then a random Toeplitz matrix as generated in this paper is appropriate. For a different type of application in non-normal time series, an interesting question is the set of possible Toeplitz correlation matrices that are consistent with stationary processes with a fixed univariate margin F (which has finite variance). One could check on the range of $d\times d$ Toeplitz matrices for a parametric model of stationary processes with the margin F by sampling at random from the set of $d \times d$ Toeplitz matrices and checking if it is consistent with the given model.

With the theory described in Section 2, we can generate the partial autocorrelations to get other nice (non-uniform) distributions over Toeplitz correlation matrices consistent with causal AR(p) Gaussian time series models. Examples include (a) density of $\boldsymbol{\rho}_p$ proportional to a power of the determinant of the Toeplitz matrix, and (b) density of $\boldsymbol{\rho}_p$ uniform over strongly positive Toeplitz matrices with all partial autocorrelations being positive. Similarly, we can generate the partial autocorrelations so that the density of $\boldsymbol{\phi}_p$ is a power of the determinant of the Toeplitz matrix. In Section 3, we indicate how to generate uniformly random Toeplitz matrices that are consistent with invertible MA(q) Gaussian time series models. More general non-uniform distributions for AR(p) and MA(q) can be obtained via the Jacobians that are derived in these sections.

In Section 4, numerical results are given for the marginal distributions. Because the specifications for the random partial autocorrelations are quite different to get uniform in the $\boldsymbol{\phi}_p$ space and in the $\boldsymbol{\rho}_p$ space, to make comparisons, we obtain the moments of the resulting random $\boldsymbol{\phi}_p$ and $\boldsymbol{\rho}_p$ to show the differences. For example, one distinction is that uniform on the $\boldsymbol{\rho}_p$ space means each ϕ coefficient has a mean of 0, but this is not the case for uniform on the $\boldsymbol{\phi}_p$ space. In Section 5, we have an application of random Toeplitz correlation matrices to compute the probability that a Toeplitz matrix can be a Spearman rank correlation matrix for a marginally transformed AR(p) Gaussian time series. Section 6 concludes with some discussion.

2. Generating AR(p) parameters

In this section, methods are proposed for generating the autoregressive parameters in a causal AR(p) model, and a new family of joint density function is introduced. Suppose $\{Z_t : t = 0, \pm 1, \pm 2, \ldots\}$ is an independent and identically distributed $N(0, \sigma^2)$ sequence. For a positive integer p, with B for the backward operator, the AR(p) Gaussian time series or process is:

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + Z_t = \sum_{i=1}^{p} \phi_i B^i X_t + Z_t, \quad t = 1, 2, \ldots,$$

or

$$\phi(B)\, X_t = Z_t, \quad \text{where } \phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p.$$

The innovation Z_t is independent of $X_{t-1}, X_{t-2}, \ldots, X_{t-p}$. If the roots of $\phi(b) = 0$ are outside the unit circle, then

$$X_t = [\phi(B)]^{-1} Z_t$$

is a zero-mean (causal) stationary process.

To allow greater degree of flexibility in the density functions, parameters $\eta > 0$ and $\delta > -1/p$ are introduced. We consider the following two families of density functions

$$f_{\boldsymbol{\rho}_p}(r_1, \ldots, r_p) \propto [\det(\mathbf{r})]^{\eta-1} \quad \text{and} \quad f_{\boldsymbol{\phi}_p}(v_1, \ldots, v_p) \propto [\det(\mathbf{r}(v_1, \ldots, v_p))]^{\delta}, \quad \mathbf{r} = (r_{jk}), \quad r_{jk} = r_{|j-k|}, r_{jj} = 1. \quad (2.1)$$

With $\boldsymbol{\rho}_p = (\rho_1, \ldots, \rho_p)^T$ being the column vector with the first p autocorrelations, the $(p+1)$-dimensional Toeplitz correlation matrix $\mathbf{R} = (\rho_{jk})$ based on $\boldsymbol{\rho}_p$ has form $\rho_{jk} = \rho_{|j-k|}$ for all $1 \le j, k \le p+1$. That is,

$$\mathbf{R} = \mathbf{R}(\boldsymbol{\rho}_p) = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_p \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{p-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_p & \rho_{p-1} & \rho_{p-2} & \cdots & 1 \end{pmatrix} = (\mathrm{Corr}(X_j, X_k))_{j,k=1,2,\ldots,p+1}.$$

Let $\boldsymbol{\alpha}_p = (\alpha_1, \ldots, \alpha_p)^T$ be the column vectors with the first p partial autocorrelations. The range of $\boldsymbol{\alpha}_p$ is $(-1, 1)^p$. Let C_p be the set of $\boldsymbol{\phi}_p = (\phi_1, \ldots, \phi_p)^T$ that correspond to causal AR(p) processes. We assume that the process has been standardized so that Var $(X_t) = 1$. Let M_p be the set of $\boldsymbol{\rho}_p$ that result in Toeplitz correlation matrices (equivalently, positive definite Toeplitz matrices with ones on the diagonal). When $\eta = 1$ in (2.1), $\boldsymbol{\rho}_p$ is uniformly distributed over M_p. When $\delta = 0$, $\boldsymbol{\phi}_p$ is uniformly distributed over C_p.

Theorem 1 of [2] has the form of the determinant of a correlation matrix in terms of partial correlations; from this result, the determinant of $\mathbf{R}(\boldsymbol{\rho}_p)$ is:

$$\det[\mathbf{R}(\boldsymbol{\rho}_p)] = \prod_{\ell=1}^{p} (1 - \alpha_l^2)^{p+1-\ell}. \tag{2.2}$$

This special case is also given as equation (5) in Barndorff-Nielsen and Schou [6].

2.1. Random $\boldsymbol{\phi}_p$

In this subsection, we indicate how to generate a random $\boldsymbol{\phi}_p \in C_p$ with joint density function

$$f_{\boldsymbol{\phi}_p}(v_1, \ldots, v_p) \propto [\det(\mathbf{r}(v_1, \ldots, v_p))]^{\delta}.$$

To generate $\boldsymbol{\phi}_p$, we consider the Levinson–Durbin formula (as given in equation (12) of [6], see also [9]) that gives the following 1–1 map between C_p and $(-1, 1)^p$ by expressing ϕ_j in terms of the partial autocorrelations α_j: for $k = 1, 2, \ldots, p-1$ and $j = 1, 2, \ldots, k$,

$$\begin{aligned} &\varphi_{k+1,k+1} = \alpha_{k+1}, \\ &\varphi_{k+1,j} = \varphi_{k,j} - \alpha_{k+1}\varphi_{k,k+1-j}, \\ &\phi_j = \varphi_{p,j}. \end{aligned} \tag{2.3}$$

In the above, note that $\phi_p = \alpha_p$.

The Jacobian of $\boldsymbol{\phi}_p$ to α_p is given in Barndorff-Nielsen and Schou [6],

$$\frac{\partial(\phi_1, \ldots, \phi_p)}{\partial(\alpha_1, \ldots, \alpha_p)} = \begin{cases} \prod_{k=2}^{p} (1-\alpha_k)^{[k/2]}(1+\alpha_k)^{[(k-1)/2]} & \text{when } p \ge 2, \\ 1 & \text{when } p = 1, \end{cases}$$

where $[\cdot]$ is the greatest integer function. Hence if the density of α_p is $f_{\alpha_p}(\mathbf{a}_p)$, then the density of $\boldsymbol{\phi}_p$ is

$$f_{\boldsymbol{\phi}_p}(\mathbf{v}_p) = f_{\alpha_p}(\mathbf{a}_p) \prod_{k=2}^{p} (1-a_k)^{-[k/2]}(1+a_k)^{-[(k-1)/2]}.$$

If $\alpha_j \sim \mathrm{Beta}([(j+1)/2] + \delta(p+1-j), [j/2] + 1 + \delta(p+1-j)), j = 1, 2, \ldots, p$, are chosen to be independent Beta random variables over the interval $(-1, 1)$, we have

$$f_{\boldsymbol{\phi}_p}(\mathbf{v}_p) \propto [\det(\mathbf{r})]^{\delta}.$$

Here, the smallest parameters are for α_1 and in order for these parameters to be positive, the condition on δ is $1 + \delta p > 0$ or $\delta > -1/p$. The effect of larger δ is obtaining distributions of α_j, ρ_j, ϕ_j that have smaller variances (or more concentration near 0).

In the special case of $\delta = 0$ where $\alpha_j \sim \mathrm{Beta}([(j+1)/2], [j/2] + 1)$ on the interval $(-1, 1)$, $\boldsymbol{\phi}_p$ is uniformly distributed in C_p; this result and extensions are given in Jones [5]. Note that the two parameters $[(j+1)/2]$ and $[j/2] + 1$ are the same if j is odd, and $[j/2] + 1 - [(j+1)/2] = 1$ is j is even. That is, α_j is symmetric about zero for j odd, and is negatively skewed (with negative expectation) for j even (but the skewness decreases as i increases for $j = 2i$). The volume of the space for $\boldsymbol{\phi}_p$ is

$$\int_{(-1,1)^p} \left| \frac{\partial \boldsymbol{\phi}_p}{\partial \boldsymbol{\alpha}_p} \right| \mathrm{d}\boldsymbol{\alpha}_p = \prod_{k=1}^{p} \int_{-1}^{1} (1-a)^{[k/2]}(1+a)^{[(k-1)/2]} \mathrm{d}a = \prod_{k=1}^{p} 2^k B([(k+1)/2], [k/2] + 1).$$

See [10] for a discussion of this result.

2.2. Random $\boldsymbol{\rho}_p$

For models, such as for longitudinal data, where the Toeplitz matrix is a parameter, it is simpler to consider generating a random $\boldsymbol{\rho}_p \in M_p$. Consider the joint density function

$$f_{\boldsymbol{\rho}_p}(r_1, \ldots, r_p) \propto [\det(\mathbf{r})]^{\eta-1}.$$

The 1–1 map between M_p and $(-1, 1)^p$ is given by the Levinson–Durbin formula (2.3) and the following recursive relationship,

$$\rho_k = \sum_{j=1}^{k} \varphi_{k,j}\rho_{k-j}. \tag{2.4}$$

Makhoul [7] has

$$\frac{\partial \rho_j}{\partial \alpha_j} = \prod_{\ell=1}^{j-1}(1-\alpha_\ell^2)$$

and a Jacobian of

$$\left|\frac{\partial(\rho_1, \ldots, \rho_p)}{\partial(\alpha_1, \ldots, \alpha_p)}\right| = \prod_{j=1}^{p-1}\prod_{i=1}^{j}(1-\alpha_i^2) = \prod_{j=1}^{p-1}(1-\alpha_j^2)^{p-j}.$$

Lemma 3 of [2] has the partial derivatives of partial correlations to correlations with the above transform; this lemma when applied to $\rho_{12}, \rho_{13}, \ldots, \rho_{1,p+1}$ also leads to the above Jacobian. Note that this Jacobian is just (2.2) with p replaced by $p-1$, i.e.,

$$\left|\frac{\partial(\rho_1, \ldots, \rho_p)}{\partial(\alpha_1, \ldots, \alpha_p)}\right| = \det(\mathbf{R}_{p-1}),$$

where $\mathbf{R}_{p-1} = \mathbf{R}(\boldsymbol{\rho}_{p-1})$ (dimension $p \times p$ with no ρ_p).

If we take independent densities $g_\ell(a_\ell)$ for α_l, $\ell = 1, \ldots, p$, then the joint density of $\rho_1, \ldots, \rho_p$ is

$$f_{\boldsymbol{\rho}_p}(r_1, \ldots, r_p) = g_1(a_1)\cdots g_p(a_p)\cdot\prod_{j=1}^{p-1}(1-a_j^2)^{-(p-j)}. \tag{2.5}$$

Suppose we take g_j to be a Beta(β_j, β_j) density on $(-1, 1)$ for $j = 1, \ldots, p$, that is,

$$g_j(u) = \frac{1}{2^{2\beta_j-1}B(\beta_j, \beta_j)}(1-u^2)^{\beta_j-1}, \quad -1 < u < 1.$$

Then (2.5) becomes

$$f_{\boldsymbol{\rho}_p}(r_1, \ldots, r_p) \propto \prod_{j=1}^{p}(1-a_j^2)^{\beta_j-1-p+j}. \tag{2.6}$$

By comparing with (2.2), the density in (2.6) is proportional to $[\det(\mathbf{r})]^{\eta-1}$ if $\beta_j = \eta(p+1-j)$. In particular, a uniform density obtains if $\eta = 1$ or $\beta_j = p+1-j$ for $j = 1, \ldots, p$. The effect of larger η are distributions of α_j, ρ_j, ϕ_j that have smaller variances (or more concentration near 0).

In the special case of $\eta = 1$, the proportional constant is

$$\prod_{j=1}^{p}\frac{1}{2^{2(p+1-j)-1}B(p+1-j, p+1-j)} = \prod_{j=1}^{p}\frac{1}{2^{2j-1}B(j,j)}.$$

Hence the volume of the $(p+1)\times(p+1)$ Toeplitz matrices in p-dimensional space is

$$V_p = \prod_{j=1}^{p}2^{2j-1}B(j,j) = \prod_{j=1}^{p}2^{2j-1}\frac{[(j-1)!]^2}{(2j-1)!} = V_{p-1}\times 2^{2p-1}\frac{[(p-1)!]^2}{(2p-1)!}.$$

This matches Makhoul's result of

$$V_p = 2^p\prod_{j=1}^{p-1}\left(\frac{2j}{2j+1}\right)^{p-j}.$$

By dividing V_p by 2^p, one gets the probabilities that a random Toeplitz matrix, with diagonals of 1 and other entries in $(-1, 1)$, is positive definite.

3. Generating MA(q) parameters

In this section, we derive a new result on generating random Toeplitz matrices that are uniform over the set of $(q+1)\times(q+1)$ correlation matrices that are consistent with an invertible MA(q) Gaussian time series for $q \geq 1$. Consider

$$X_t = Z_t - \theta_1 Z_{t-1} - \cdots - \theta_q Z_{t-q}, \quad Z_t \sim^{iid} N(0, \tilde{\tau}). \tag{3.1}$$

We are interested in one of the following:

1. $\boldsymbol{\theta}_q = (\theta_1, \ldots, \theta_q)^T$ uniformly distributed over C_q,
2. $\boldsymbol{\rho}_q = (\rho_1, \ldots, \rho_q)^T$ uniformly distributed over M_q^*,

where M_q^* consists of vectors $\boldsymbol{\rho}_q$ that are consistent with an invertible MA(q) model.

Uniformly distributed $\boldsymbol{\theta}_q$ can be obtained by virtue of the fact that there is symmetry between $\theta(B)$ for an invertible MA(q) Gaussian time series, and $\phi(B)$ for a causal stationary AR(q) process. To generate a random $\boldsymbol{\theta}_q \in C_q$, one can generate (pseudo-partial correlations) $\tilde{\boldsymbol{\alpha}}_q \in (-1, 1)^q$ and get $\boldsymbol{\theta}_q$ via the Levinson–Durbin algorithm. Then one can get the autocorrelations $\boldsymbol{\rho}_q$ based on $\boldsymbol{\theta}_q$; the autocorrelations of lags greater than q are zero. The results can be considered as a random coefficient vector and random autocorrelation vector for an MA(q) process. A uniform $\boldsymbol{\theta}_q \in C_q$ is covered in Jones [5], but not uniform $\boldsymbol{\rho}_q \in M_q^*$.

In the subsequent subsections, we consider the transform from $\tilde{\boldsymbol{\alpha}}_q$ to $\boldsymbol{\theta}_q$. The Jacobian is given and an algorithm for generating $\boldsymbol{\rho}_q$ is presented.

3.1. *Jacobian of $\boldsymbol{\theta}_q$ to $\tilde{\boldsymbol{\alpha}}_q$*

First, the 1–1 map between $\boldsymbol{\theta}_q$ and $\tilde{\boldsymbol{\alpha}}_q$ is given below. The first q (non-zero) autocorrelations are:

$$\begin{aligned}
\rho_1 &= \frac{-\theta_1 + \theta_1\theta_2 + \cdots + \theta_{q-1}\theta_q}{1 + \theta_1^2 + \cdots + \theta_q^2} \\
\rho_2 &= \frac{-\theta_2 + \theta_1\theta_3 + \cdots + \theta_{q-2}\theta_q}{1 + \theta_1^2 + \cdots + \theta_q^2} \\
\vdots &= \vdots \\
\rho_q &= \frac{-\theta_q}{1 + \theta_1^2 + \cdots + \theta_q^2}.
\end{aligned} \tag{3.2}$$

The mapping is 1–1 over $\boldsymbol{\theta}_q \in C_q$. From the Levinson–Durbin algorithm,

$$\begin{aligned}
&\theta_{k,k} = \tilde{\alpha}_k, \quad k = 1, \ldots, q, \\
&\theta_{k,j} = \theta_{k-1,j} - \theta_{k-1,k-j}\tilde{\alpha}_k, \quad j = 1, \ldots, k-1, k = 1, \ldots, q, \\
&\theta_j = \theta_{q,j}, \quad j = 1, \ldots, q.
\end{aligned} \tag{3.3}$$

In order to generate random $\boldsymbol{\rho}_q \in M_q^*$, we can start with a distribution for $\tilde{\boldsymbol{\alpha}}_q = (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_q)^T$ and apply (3.2) and (3.3). In order to get a distribution of $\boldsymbol{\rho}_q$ that is uniform (or non-uniform), we need the Jacobian of the transformation of $\tilde{\boldsymbol{\alpha}}_q$ to $\boldsymbol{\rho}_q$. If q is fixed and a small integer, the Jacobian can be obtained for an individual q using symbolic manipulation software. Below, we provide a proof of the general form of the Jacobian.

Proposition 3.1. *With the equations in (3.3) and (3.2), for an integer $q \geq 1$, the Jacobian of the transformation of $\tilde{\boldsymbol{\alpha}}_q$ to $\boldsymbol{\rho}_q$ is:*

$$\left|\frac{\partial(\rho_1, \ldots, \rho_q)}{\partial(\tilde{\alpha}_1, \ldots, \tilde{\alpha}_q)}\right| = \left\{\prod_{k=1}^{q}(1-\tilde{\alpha}_k)^{2[k/2]+1}(1+\tilde{\alpha}_k)^{2[(k-1)/2]+1}\right\}(1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)^{-(q+1)}. \tag{3.4}$$

Proof. Note that this Jacobian can be obtained for $q = 1$ from differentiating $\rho_1 = -\theta_1/(1+\theta_1^2)$ with respect to $\theta_1 = \tilde{\alpha}_1$.

The remainder of the proof is for $q \geq 2$. Consider the invertible MA(q) model:

$$X_t = (1-\theta_0^*)Z_t^* - \theta_1^* Z_{t-1}^* - \cdots - \theta_q^* Z_{t-q}^*, \quad Z_t^* \sim^{iid} N(0, 1),$$

with $\theta_0^* < 1$. It is equivalent to the model (3.1) with $\tilde{\tau} = (1-\theta_0^*)^2$, $\theta_1 = \theta_1^*(1-\theta_0^*)^{-1}, \ldots, \theta_q = \theta_q^*(1-\theta_0^*)^{-1}$. Suppose that $(\theta_1, \ldots, \theta_q)$ is obtained from the pseudo-partial autocorrelations $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_q$, via the equations in (3.3). The domain for

$(\theta_0^*, \ldots, \theta_q^*)$ is therefore the same as that of $(\phi_0^*, \ldots, \phi_q^*)$ in Lemma 3.3 in Section 3.3. We next obtain the Jacobian of the transformation from $(\theta_0^*, \ldots, \theta_q^*)$ to the autocovariance vector $(\gamma_0, \ldots, \gamma_1)$. The autocovariances satisfy:

$$\begin{aligned}
\gamma_0 &= (\theta_0^* - 1)^2 + (\theta_1^*)^2 + \cdots + (\theta_q^*)^2, \\
\gamma_1 &= (\theta_0^* - 1)\theta_1^* + \theta_1^*\theta_2^* + \cdots + \theta_{q-1}^*\theta_q^*, \\
\vdots &= \vdots \\
\gamma_q &= (\theta_0^* - 1)\theta_q^*.
\end{aligned}$$

In this case, we have

$$\left|\frac{\partial(\gamma_0, \ldots, \gamma_q)}{\partial(\theta_0^*, \ldots, \theta_q^*)}\right| = 2\det(\mathbf{I} - \Theta_U^* - \Theta_L^{*T}) = 2\det(\mathbf{I} - \Theta_U^* - \Theta_L^*), \tag{3.5}$$

where $\boldsymbol{\Theta}_U^*$ is symmetric,

$$\boldsymbol{\Theta}_U^* = \begin{pmatrix} \theta_0^* & \theta_1^* & \cdots & \theta_q^* \\ \vdots & & \nearrow & 0 \\ \theta_{q-1}^* & \theta_q^* & & \\ \theta_q^* & & & \end{pmatrix}, \qquad \boldsymbol{\Theta}_L^* = \begin{pmatrix} 0 & & & \\ 0 & \theta_0^* & & \\ \vdots & \vdots & \ddots & \\ 0 & \theta_{q-1}^* & \cdots & \theta_0^* \end{pmatrix}.$$

By Lemma 3.3 in Section 3.3, we have

$$\det(\mathbf{I} - \Theta_U^* - \Theta_L^*) = (1-\theta_0^*)^{q+1}\prod_{j=1}^{q}(1-\tilde{\alpha}_k)^{[k/2]+1}(1+\tilde{\alpha}_k)^{[(k-1)/2]+1}.$$

From

$$\left|\frac{\partial(\theta_0^*, \ldots, \theta_q^*)}{\partial(\tilde{\tau}, \theta_1 \ldots, \theta_q)}\right| = (1-\theta_0^*)^{q-1}/2, \tag{3.6}$$

and

$$\left|\frac{\partial(\gamma_0, \rho_1, \ldots, \rho_q)}{\partial(\gamma_0, \gamma_1, \ldots, \gamma_q)}\right| = \gamma_0^{-q}, \tag{3.7}$$

after some manipulations, we have from a product of (3.7), (3.5) and (3.6):

$$\left|\frac{\partial(\gamma_0, \rho_1, \ldots, \rho_q)}{\partial(\tilde{\tau}, \theta_1 \ldots, \theta_q)}\right| = (\tilde{\tau}/\gamma_0)^q\prod_{k=1}^{q}(1-\tilde{\alpha}_k)^{[k/2]+1}(1+\tilde{\alpha}_k)^{[(k-1)/2]+1}.$$

Note that $\rho_1, \ldots, \rho_q$ is determined by $\theta_1, \ldots, \theta_q$ only and that

$$\gamma_0 = \tilde{\tau}(1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2),$$

so that we obtain

$$\left|\frac{\partial(\gamma_0, \rho_1, \ldots, \rho_q)}{\partial(\tilde{\tau}, \theta_1 \ldots, \theta_q)}\right| = \frac{\gamma_0}{\tilde{\tau}}\left|\frac{\partial(\rho_1, \ldots, \rho_q)}{\partial(\theta_1 \ldots, \theta_q)}\right|,$$

and therefore

$$\begin{aligned}
\left|\frac{\partial(\rho_1, \ldots, \rho_q)}{\partial(\theta_1 \ldots, \theta_q)}\right| &= (\tilde{\tau}/\gamma_0)^{q+1}\prod_{k=1}^{q}(1-\tilde{\alpha}_k)^{[k/2]+1}(1+\tilde{\alpha}_k)^{[(k-1)/2]+1} \\
&= \left\{\prod_{k=1}^{q}(1-\tilde{\alpha}_k)^{[k/2]+1}(1+\tilde{\alpha}_k)^{[(k-1)/2]+1}\right\}(1+\theta_1^2+\theta_2^2+\cdots+\theta_q^2)^{-(q+1)}.
\end{aligned}$$

Since, from Section 2.1,

$$\left|\frac{\partial(\theta_1 \ldots, \theta_q)}{\partial(\tilde{\alpha}_1, \ldots, \tilde{\alpha}_q)}\right| = \prod_{k=2}^{q}(1-\tilde{\alpha}_k)^{[k/2]}(1+\tilde{\alpha}_k)^{[(k-1)/2]},$$

from a product of the above two equations, (3.4) obtains. □

3.2. Algorithm

In this subsection, we derive the algorithm for generating a uniform $\boldsymbol{\rho}_q \in M_q^*$.

Let c_q be the volume of M_q^* as a subset of q-dimensional Euclidean space. For $q = 1$, ρ_1 lies in $(-\frac{1}{2}, \frac{1}{2})$ so that $c_1 = 1$.

Using arguments $t_{k,j}$, $\tilde{a}_j$ in place of the random variables $\theta_{k,j}$, $\tilde{\alpha}_j$ in (3.3), then from (3.4), we get the uniform density $f_{\boldsymbol{\rho}_q}(r_1, \dots, r_q) = c_q^{-1}$ if

$$f_{\tilde{\alpha}_q}(\tilde{a}_1, \dots, \tilde{a}_q) = c_q^{-1} \prod_{k=1}^{q} (1-\tilde{a}_k)^{2[k/2]+1}(1+\tilde{a}_k)^{2[(k-1)/2]+1}(1+t_1^2+t_2^2+\cdots+t_q^2)^{-(q+1)}. \tag{3.8}$$

We can decompose (3.8) as a product

$$f_{\tilde{\alpha}_q}(\tilde{a}_1, \dots, \tilde{a}_q) = \frac{(1-\tilde{a}_1)(1+\tilde{a}_1)}{(1+t_{1,1}^2)^2} \times \prod_{k=2}^{q} \frac{c_{k-1}}{c_k}(1-\tilde{a}_k)^{2[k/2]+1}(1+\tilde{a}_k)^{2[(k-1)/2]+1} \frac{[1+\sum_{j=1}^{k-1} t_{k-1,j}^2]^k}{[1+\sum_{j=1}^{k} t_{k,j}^2]^{k+1}}.$$

Note that $t_{k,j}$ is a function of $\tilde{a}_1, \dots, \tilde{a}_k$ for $j = 1, \dots, k$. With this functional form, by induction, we identify

$$f_{\tilde{\alpha}_1}(\tilde{a}_1) = (1-\tilde{a}_1)(1+\tilde{a}_1)/(1+\tilde{a}_1^2),$$

and conditional densities (for $k \geq 2$):

$$\begin{aligned} f_{\tilde{\alpha}_k|\tilde{\alpha}_1,\dots,\tilde{\alpha}_{k-1}}(\tilde{a}_k|\tilde{a}_1,\dots,\tilde{a}_{k-1}) &= \frac{c_{k-1}}{c_k}(1-\tilde{a}_k)^{2[k/2]+1}(1+\tilde{a}_k)^{2[(k-1)/2]+1} \frac{[1+\sum_{j=1}^{k-1} t_{k-1,j}^2]^k}{[1+\sum_{j=1}^{k} t_{k,j}^2]^{k+1}} \\ &= \frac{c_{k-1}}{c_k}(1-\tilde{a}_k)^{2[k/2]+1}(1+\tilde{a}_k)^{2[(k-1)/2]+1}(1-2\eta_{k-1}\tilde{a}_k+\tilde{a}_k^2)^{-(k+1)}\left[1+\sum_{j=1}^{k-1} t_{k-1,j}^2\right]^{-1}, \end{aligned} \tag{3.9}$$

where for $i \geq 1$, with $\tilde{a}_k = t_{k,k}$ and $t_{k,j} = t_{k-1,j} - t_{k-1,k-j}\tilde{a}_k$ $(1 \leq j < k)$,

$$\eta_i = \frac{t_{i,1}t_{i,i} + t_{i,2}t_{i,i-1} + \cdots + t_{i,i}t_{i,1}}{1 + t_{i,1}^2 + \cdots + t_{i,i}^2}. \tag{3.10}$$

In (3.9), after omitting the terms that do not involve $\tilde{a}_k$, let

$$g_k(\tilde{a}_k|\tilde{a}_1, \dots, \tilde{a}_{k-1}) = (1-\tilde{a}_k)^{2[k/2]+1}(1+\tilde{a}_k)^{2[(k-1)/2]+1}(1-2\eta_{k-1}\tilde{a}_k+\tilde{a}_k^2)^{-(k+1)}. \tag{3.11}$$

We have $|\eta_i| < 1$ for all $i = 1, \dots, q$, and η_{k-1} depends on $\tilde{a}_1, \dots, \tilde{a}_{k-1}$ only. Let $\nu_{k1} = 2[k/2]+1$ and $\nu_{k2} = 2[(k-1)/2]+1$. An upper bound on $(1-a)^{\nu_{k1}}(1+a)^{\nu_{k2}}$ is $(2\nu_{k1})^{\nu_{k1}}(2\nu_{k2})^{\nu_{k2}}/(\nu_{k1}+\nu_{k2})^{\nu_{k1}+\nu_{k2}}$ when $1-a = 2\nu_{k1}/(\nu_{k1}+\nu_{k2})$. An upper bound on $(1-2\eta_{k-1}a+a^2)^{-1}$ is $(1-\eta_{k-1}^2)^{-1}$ when $a = \eta_{k-1}$. Therefore, an upper bound for the right-hand side of (3.11) is

$$N_k(\eta_{k-1}) = \frac{(2\nu_{k1})^{\nu_{k1}}(2\nu_{k2})^{\nu_{k2}}}{(\nu_{k1}+\nu_{k2})^{\nu_{k1}+\nu_{k2}}(1-\eta_{k-1}^2)^{k+1}}. \tag{3.12}$$

To get uniform $\boldsymbol{\rho}_q$, the random variables $\tilde{\alpha}_1, \dots, \tilde{\alpha}_q$ can be simulated sequentially by using method of rejection as follows.

1. Let $U \sim \text{Unif}(-\frac{1}{2}, \frac{1}{2})$, set $\tilde{\alpha}_1 \leftarrow [-1+\sqrt{1-4U^2}\,]/(2U)$ [or $U = -\tilde{\alpha}_1/(1+\tilde{\alpha}_1^2)$], set $\eta_1 \leftarrow \tilde{\alpha}_1^2/(1+\tilde{\alpha}_1^2)$, and compute $N_2(\eta_1)$ from (3.12) with $\nu_{21} = 3$, $\nu_{22} = 1$.
2. For $k = 2, \dots, q$:
   a. Generate $U_k \sim \text{Unif}(0, 1)$ and $V \sim \text{Unif}(-1, 1)$.
   b. If $U_k < [N_k(\eta_{k-1})]^{-1} g_k(V|\tilde{\alpha}_1, \dots, \tilde{\alpha}_{k-1})$, set $\tilde{\alpha}_k \leftarrow V$ and go to (c), otherwise repeat step (a).
   c. Compute $\theta_{k,j}$ $(1 \leq j \leq k-1)$ with (3.3), η_k from (3.10) with $t_{k,j} = \theta_{k,j}$, and $N_{k+1}(\eta_k)$ from (3.12) with $\nu_{k+1,1} = 2[(k+1)/2]+1$ and $\nu_{k+1,2} = 2[k/2]+1$.
3. Compute $\rho_j, j = 1, \dots, q$ from (3.2).

3.3. Technical Lemmae

Lemma 3.2. *Let $\boldsymbol{\alpha} \in (-1, 1)^p$ and $\boldsymbol{\phi}_p$ be defined as in (2.3). We have*

$$\left|\frac{\partial(\phi_1, \ldots, \phi_p)}{\partial(\rho_1, \ldots, \rho_p)}\right| = \left|\frac{\partial(\phi_1, \ldots, \phi_p)}{\partial(\alpha_1, \ldots, \alpha_p)}\right| \cdot \left|\frac{\partial(\alpha_1, \ldots, \alpha_p)}{\partial(\rho_1, \ldots, \rho_p)}\right| = \left\{\prod_{k=2}^{p}(1-\alpha_k)^{[k/2]}(1+\alpha_k)^{[(k-1)/2]}\right\} \cdot \det(\mathbf{R}_{p-1}^{-1}), \tag{3.13}$$

where $\mathbf{R}_{p-1} = \mathbf{R}(\boldsymbol{\rho}_{p-1})$ (dimension $p \times p$ with no ρ_p). Note that when $p = 1$, the Jacobian is 1.

Proof. The proposition is a direct consequence of the results in Sections 2.1 and 2.2. □

Lemma 3.3. *Let $\boldsymbol{\alpha} \in (-1, 1)^p$ and $\phi_0^* < 1$, Define $\boldsymbol{\phi}_p$ by (2.3). Let*

$$\boldsymbol{\Phi}_U^* = \begin{pmatrix} \phi_0^* & \phi_1^* & \cdots & \phi_p^* \\ \vdots & & \nearrow & 0 \\ \phi_{p-1}^* & \phi_p^* & & \\ \phi_p^* & & & \end{pmatrix}, \qquad \boldsymbol{\Phi}_L^* = \begin{pmatrix} 0 & & & \\ 0 & \phi_0^* & & \\ \vdots & \vdots & \ddots & \\ 0 & \phi_{p-1}^* & \cdots & \phi_0^* \end{pmatrix},$$

where $\phi_1 = \phi_1^*(1-\phi_0^*)^{-1}, \ldots, \phi_p = \phi_p^*(1-\phi_0^*)^{-1}$. Then,

$$\det(\mathbf{I} - \boldsymbol{\Phi}_U^* - \boldsymbol{\Phi}_L^*) = (1-\phi_0^*)^{p+1}\prod_{k=1}^{p}(1-\alpha_k)^{[k/2]+1}(1+\alpha_k)^{[(k-1)/2]+1}. \tag{3.14}$$

Proof. Consider the AR(p) model of the form:

$$X_t = \phi_0^* X_t + \phi_1^* X_{t-1} + \cdots + \phi_p^* X_{t-p} + \epsilon_t^*, \tag{3.15}$$

where $\phi_0^* < 1$ and $\epsilon_t^* \sim N(0, 1)$. It is equivalent to the model with $\tau = \sigma^2 = (1-\phi_0^*)^{-2}$, $\phi_1 = \phi_1^*(1-\phi_0^*)^{-1}, \ldots, \phi_p = \phi_p^*(1-\phi_0^*)^{-1}$.

In the following, we find $\det(\mathbf{I} - \boldsymbol{\Phi}_U^* - \boldsymbol{\Phi}_L^*)$ via the identity

$$\mathbf{J}^* = (\Gamma + \tau \mathbf{e}_0 \mathbf{e}_0^T)^{-1}(\mathbf{I} - \boldsymbol{\Phi}_U^* - \boldsymbol{\Phi}_L^*), \tag{3.16}$$

where $\mathbf{J}^*$ is the Jacobian of transforming $\phi_0^*, \phi_1^*, \ldots, \phi_p^*$ to the autocovariances $\gamma_0, \gamma_1, \ldots, \gamma_p$.

To establish (3.16), we consider a modified Yule–Walker equation

$$\begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_p \end{pmatrix} = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_p \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_p & \gamma_{p-1} & \cdots & \gamma_0 \end{pmatrix} \begin{pmatrix} \phi_0^* \\ \phi_1^* \\ \vdots \\ \phi_p^* \end{pmatrix} + \begin{pmatrix} (1-\phi_0^*)^{-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

With indices of the vectors and matrices from 0 to p, and $\mathbf{e}_i$ ($i = 0, \ldots, p$) representation the vector with 1 in the position i and 0 elsewhere, the matrix form of the above equation is:

$$\boldsymbol{\gamma} = \boldsymbol{\Gamma}\boldsymbol{\phi}^* + (1-\phi_0^*)^{-1}\mathbf{e}_0. \tag{3.17}$$

Differentiate both sides of (3.17) with respect to γ_i, we obtain

$$\boldsymbol{\Gamma}\begin{pmatrix} \dfrac{\partial \boldsymbol{\phi}_0^*}{\partial \gamma_i} \\ \vdots \\ \dfrac{\partial \boldsymbol{\phi}_p^*}{\partial \gamma_i} \end{pmatrix} + \begin{pmatrix} \phi_i^* \\ \vdots \\ \phi_p^* \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \phi_0^* \\ \vdots \\ \phi_{p-i}^* \end{pmatrix} + (1-\phi_0^*)^{-2}\begin{pmatrix} \dfrac{\partial \boldsymbol{\phi}_0^*}{\partial \gamma_i} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{e}_i,$$

for $i = 1, 2, \ldots, p$, and

$$\boldsymbol{\Gamma}\begin{pmatrix}\dfrac{\partial\boldsymbol{\phi}_0^*}{\partial\gamma_0}\\ \vdots\\ \dfrac{\partial\boldsymbol{\phi}_p^*}{\partial\gamma_0}\end{pmatrix} + \begin{pmatrix}\phi_0^*\\ \vdots\\ \phi_p^*\end{pmatrix} + (1-\phi_0^*)^{-2}\begin{pmatrix}\dfrac{\partial\boldsymbol{\phi}_0^*}{\partial\gamma_0}\\ 0\\ \vdots\\ 0\end{pmatrix} = \mathbf{e}_0.$$

Then, we have

$$(\boldsymbol{\Gamma} + \tau\mathbf{e}_0\mathbf{e}_0^T)\mathbf{J}^* + \boldsymbol{\Phi}_U^* + \boldsymbol{\Phi}_L^* = \mathbf{I},$$

and the identity (3.16) follows.

Next, we find $\det(\boldsymbol{\Gamma} + \tau\mathbf{e}_0\mathbf{e}_0^T)$ and $\det(\mathbf{J}^*)$. We have

$$\begin{aligned}\det(\boldsymbol{\Gamma} + \tau\mathbf{e}_0\mathbf{e}_0^T) &= \gamma_0^{p+1}\det[\mathbf{R}_p] + \tau\gamma_0^p\det(\mathbf{R}_{p-1})\\ &= \gamma_0^{p+1}\prod_{j=1}^{p}(1-\alpha_j^2)^{p+1-j} + \gamma_0^{p+1}\prod_{j=1}^{p}(1-\alpha_j^2)\prod_{j=1}^{p-1}(1-\alpha_j^2)^{p-j}\\ &= 2\gamma_0^{p+1}\prod_{j=1}^{p}(1-\alpha_j^2)^{p+1-j},\end{aligned} \tag{3.18}$$

where, in the second equality, we have used equation (8) from Barndorff-Nielsen and Schou [6]:

$$\gamma_0 = \tau(1-\alpha_1^2)^{-1}\ldots(1-\alpha_p^2)^{-1}. \tag{3.19}$$

Also

$$\det(\mathbf{J}^*) = \left|\frac{\partial(\phi_0^*,\ldots,\phi_p^*)}{\partial(\gamma_0,\ldots,\gamma_p)}\right| = \left|\frac{\partial(\phi_0^*,\ldots,\phi_p^*)}{\partial(\tau,\phi_1\ldots,\phi_p)}\right|\cdot\left|\frac{\partial(\tau,\phi_1\ldots,\phi_p)}{\partial(\gamma_0,\rho_1,\ldots,\rho_p)}\right|\cdot\left|\frac{\partial(\gamma_0,\rho_1,\ldots,\rho_p)}{\partial(\gamma_0,\gamma_1,\ldots,\gamma_p)}\right|.$$

It is straightforward to show that

$$\left|\frac{\partial(\phi_0^*,\ldots,\phi_p^*)}{\partial(\tau,\phi_1\ldots,\phi_p)}\right| = (1-\phi_0^*)^{p+3}/2, \qquad \left|\frac{\partial(\gamma_0,\rho_1,\ldots,\rho_p)}{\partial(\gamma_0,\gamma_1,\ldots,\gamma_p)}\right| = \gamma_0^{-p}.$$

Note that $\rho_1, \ldots, \rho_p$ are determined solely by $\phi_1, \ldots, \phi_p$. Using (3.19),

$$\left|\frac{\partial(\tau,\phi_1\ldots,\phi_p)}{\partial(\gamma_0,\rho_1,\ldots,\rho_p)}\right| = \frac{\tau}{\gamma_0}\cdot\left|\frac{\partial(\phi_1\ldots,\phi_p)}{\partial(\rho_1,\ldots,\rho_p)}\right|.$$

Then, from Lemma 3.2, we have

$$\det(\mathbf{J}^*) = \frac{1}{2}(1-\phi_0^*)^{p+1}\gamma_0^{-(p+1)}\left\{\prod_{k=2}^{p}(1-\alpha_k)^{[k/2]}(1+\alpha_k)^{[(k-1)/2]}\right\}\det(\mathbf{R}_{p-1}^{-1}). \tag{3.20}$$

Eq. (3.14) now follows from Eqs. (3.16) and (2.2) with the product of (3.20) and (3.18). □

4. Marginal distributions of ρ_j and ϕ_j for AR(p) and MA(q)

For simulation studies or Bayesian inference for statistical models with a Toeplitz matrix as a parameter, the behavior of the marginal distributions of ρ_j can help in choosing among the various generating methods for $\boldsymbol{\rho}_p \in M_p$ or $\boldsymbol{\rho}_q \in M_q^*$. In this section, we present numerical results on the first two moments of the marginal distributions of ρ_j and ϕ_j for AR(p), or ρ_j and θ_j for MA(q), when $\boldsymbol{\alpha}_p$ or $\tilde{\boldsymbol{\alpha}}_q$ is generated with one of the distributions in Section 2 or 3. Some theoretical properties of the expected values, motivated from the computational results, of ρ_j, ϕ_j and θ_j are also given.

First, we consider moments for AR(p). For AR(p), the expressions for the ϕ_j's in terms of α_j's can be obtained by means of the Levinson–Durbin formula in (2.3). For $p = 3$, this leads to:

$$\phi_1 = \alpha_1 - \alpha_1\alpha_2 - \alpha_2\alpha_3, \qquad \phi_2 = \alpha_2 - \alpha_1\alpha_3 + \alpha_1\alpha_2\alpha_3, \qquad \phi_3 = \alpha_3.$$

A general pattern is given in the next proposition.

Proposition 4.1. *$\phi_1, \ldots, \phi_p$ are polynomials in $\alpha_1, \ldots, \alpha_p$ where the power of each α_j in any term is* 0*or* 1*. In addition, for all $k = 1, 2, \ldots, p-1$, and $j = 1, 2, \ldots, p$, ϕ_k includes at least one term involving α_j.*

Proof. This can be seen inductively from the Levinson–Durbin formula (2.3). Note that both $\varphi_{k,j}$ and $\varphi_{k,k+1-j}$ do not depend on α_{k+1}. Therefore, we have the power of each α_j in any term is 0 or 1. $\square$

The expressions for ρ_j in terms of α_j's can be obtained with properties of conditional distributions of multivariate normal, or recursively by (2.4) and the Levinson–Durbin formula (2.3). We list the first four ρ_k:

$$\begin{aligned}
\rho_1 &= \alpha_1,\\
\rho_2 &= \alpha_1^2 + \alpha_2(1-\alpha_1^2),\\
\rho_3 &= \alpha_3(1-\alpha_2^2)(1-\alpha_1^2) + \alpha_1^3(1-\alpha_2)^2 + \alpha_1(2\alpha_2-\alpha_2^2),\\
\rho_4 &= \alpha_4 + \left[\alpha_1^2+\alpha_2^2+\alpha_3^2-\alpha_1^2\alpha_2^2-\alpha_1^2\alpha_3^2-\alpha_2^2\alpha_3^2+\alpha_1^2\alpha_2^2\alpha_3^2\right](1-\alpha_4)\\
&\quad + (1-\alpha_1^2)(1-\alpha_2)\big(\alpha_2\alpha_3+\alpha_3-\alpha_1+\alpha_1\alpha_2\big)^2.
\end{aligned}$$

The next proposition explains a general pattern of the ρ_k's in terms of the α_j's. This and the previous proposition help with the evaluation, via symbolic manipulation software, of the moments of the random ρ_k's or ϕ_k's when the α_j's are independent random variables.

Proposition 4.2. *For $k = 1, 2, \ldots, p$, ρ_k is a polynomial in $\alpha_1, \ldots, \alpha_k$; the highest degree of α_i in ρ_k is $k+1-i$, where $i = 1, \ldots, k$. For example, in ρ_4, the highest power of α_1 is α_1^4, the highest power of α_2 is α_2^3, the highest power of α_3 is α_3^2, the highest power of α_4 is α_4^1.*

Proof. This can be shown inductively from Proposition 4.1 and (2.4). Assume that for all $j = 1, 2, \ldots, k-1$, when $m \le k-j$, the highest power of α_m in ρ_{k-j} is $\alpha_m^{k+1-j-m}$. In what follows, we consider the highest power of α_i, $i = 1, 2, \ldots, k$, for each terms $\varphi_{k,j}\rho_{k-j}$.

Case: $i = 1, 2, \ldots, k-1$. From Proposition 4.1, the highest power of α_i in $\varphi_{k,j}$ is α_i^1. By our inductive assumptions, the highest power of α_i in ρ_{k-j} is $\alpha_i^{(k+1-j-i)}$. We see that on the right-hand side of (2.4), the highest power of α_i appears in the term with $j = 1$, and the power is $1 + (k-i) = k+1-i$.

Case: $i = k$. Note that α_k appears only in the term with $j = k$ and we have $\varphi_{k,k}\rho_{k-k} = \alpha_k = \alpha_k^{k+1-k}$.

This completes the induction. $\square$

If the partial correlations α_j have independent Beta distributions as given in Sections 2.1 and 2.2, then $\rho_1 = \alpha_1$ has a Beta distribution. However there are no simple marginal distributions for $\rho_2, \ldots, \rho_p$. For example for AR(2), it can be checked that there is no simple distribution for $\rho_2 = \alpha_1^2 + \alpha_2(1-\alpha_1^2)$ when α_1, α_2 are random. Because $\phi_p = \alpha_p$ for AR(p), $p \ge 2$, α_p has a Beta distribution. However there are no simple marginal distributions for $\phi_1, \ldots, \phi_{p-1}$.

This differs from the case of an unstructured correlation matrix, as studied in Joe [2]; with appropriate Beta distributions on some partial correlations, the joint distribution of all correlations is proportional to a power of the determinant, which then implies that the marginal distribution of each correlation is the same Beta distribution.

Therefore for comparing the distributions of ρ_j and ϕ_j for based on uniform over the $\boldsymbol{\rho}_p$ space or $\boldsymbol{\phi}_p$ space, we compare the expected values and variances. Table 1 show the numerical values for $p = 3, 4, 5$; the results were obtained with symbolic manipulation programs in Maple (and compared with simulation results as a check); the properties of Propositions 4.1 and 4.2 were embedded in the programs. The last two columns contain another interesting case, corresponding to uniform over the $\boldsymbol{\rho}_p$ space with the restriction to positive α values. This is interpretation as a case of strong positive serial dependence, which one might be useful in simulation studies. Probabilistically, α_j are independent random variables with density functions

$$f_{\alpha_j}(a) \propto (1-a^2)^{p-j} \quad \text{for } 0 < a < 1, \tag{4.1}$$

that is, $\alpha_j = |B_j|$ where $B_j \sim \text{Beta}(p+1-j, p+1-j)$ on $(-1, 1)$ or equivalently $\alpha_j^2 \sim \text{Beta}(\frac{1}{2}, p+1-j)$; the even moments of α_j can be obtained from the Beta distribution and the odds moments can be obtained through a recursion.

Note that uniform over the $\boldsymbol{\phi}_p$ space leads to ρ_j that have higher variance than uniform over the $\boldsymbol{\rho}_p$ space. Also the expected values of the ϕ_j are all zero only in the latter case of uniform $\boldsymbol{\rho}_p$. For the case of positive α_j in (4.1), note that the condition implies all ρ_j are positive, but the ϕ_j can be negative. The pattern from larger values of $p \ge 5$ is the expected value of ϕ_1 is negative and then increases to 0.5 for ϕ_p. Some of the patterns appearing in Table 1 are explained in the following proposition.

Proposition 4.3. (I) *For $\alpha_1, \ldots, \alpha_p$ leading to uniform in the $\boldsymbol{\phi}_p$ space, $\mathrm{E}[\phi_j] = 0$ for j odd, and $\mathrm{E}[\phi_j] < 0$ for j even; in addition, when p is odd, $\mathrm{E}[\phi_1], \ldots, \mathrm{E}[\phi_{p-1}]$ are the same as those for $p-1$. For $\alpha_1, \ldots, \alpha_p$ leading to uniform in the $\boldsymbol{\rho}_p$ space, $\mathrm{E}[\phi_j] = 0$ for all j;*

(II) *For $\alpha_1, \ldots, \alpha_p$ leading to uniform in the $\boldsymbol{\phi}_p$ or $\boldsymbol{\rho}_p$ space, $\mathrm{E}[\rho_j] = 0$ for j odd.*

Table 1
Means and variances for AR(p) with $p = 3$ and $p = 4$.

Variable	Uniform $\boldsymbol{\rho}_p$		Uniform $\boldsymbol{\phi}_p$		Uniform Toeplitz, pos. α	
	Exp. value	Variance	Exp. value	Variance	Exp. value	Variance
$p = 3$						
ρ_1	0	0.143	0	0.333	0.312	0.045
ρ_2	0.143	0.180	0.111	0.277	0.464	0.056
ρ_3	0	0.247	0	0.263	0.550	0.059
ϕ_1	0	0.238	0	0.733	0.008	0.076
ϕ_2	0	0.257	−0.333	0.356	0.277	0.090
ϕ_3	0	0.333	0	0.200	0.500	0.083
$p = 4$						
ρ_1	0	0.111	0	0.333	0.273	0.036
ρ_2	0.111	0.133	0.111	0.277	0.389	0.045
ρ_3	0	0.168	0	0.263	0.446	0.048
ρ_4	0.123	0.215	0.076	0.244	0.570	0.055
ϕ_1	0	0.222	0	0.800	−0.117	0.096
ϕ_2	0	0.224	−0.400	0.587	0.121	0.025
ϕ_3	0	0.252	0	0.373	0.340	0.094
ϕ_4	0	0.333	−0.200	0.160	0.500	0.083
$p = 5$						
ρ_1	0	0.091	0	0.333	0.246	0.030
ρ_2	0.091	0.106	0.111	0.277	0.339	0.037
ρ_3	0	0.127	0	0.263	0.383	0.039
ρ_4	0.099	0.153	0.076	0.244	0.477	0.045
ρ_5	0	0.195	0	0.238	0.501	0.054
ϕ_1	0	0.212	0	0.828	−0.211	0.108
ϕ_2	0	0.206	−0.400	0.640	−0.003	0.050
ϕ_3	0	0.216	0	0.480	0.210	0.070
ϕ_4	0	0.248	−0.200	0.274	0.387	0.097
ϕ_5	0	0.333	0	0.143	0.500	0.083

Proof of (I). Uniform $\boldsymbol{\rho}_p$. The required results can be established by induction using the Levinson–Durbin formula (2.3), the facts that $\mathrm{E}[\varphi_{j,j}] = \mathrm{E}[\alpha_j] = 0$, and $\varphi_{k,k+1-j}$ do not depend on α_{k+1}.

Uniform $\boldsymbol{\phi}_p$. Note that when k is odd, the two Beta parameters of α_k are the same, and $\mathrm{E}[\varphi_{k,k}] = \mathrm{E}[\alpha_k] = 0$. When k is even, the two Beta parameters differ by one, and $\mathrm{E}[\varphi_{k,k}] = \mathrm{E}[\alpha_k] < 0$. Next, we apply mathematical induction with the Levinson–Durbin formula. Assume that $\mathrm{E}[\varphi_{k,j}] = 0$ for j odd, and $\mathrm{E}[\varphi_{k,j}] < 0$ for j even. Below, we show that $\mathrm{E}[\varphi_{k+1,j}] = 0$ for j odd, and $\mathrm{E}[\varphi_{k+1,j}] < 0$ for j even.

Case: j is odd. Note that $k+1$ and $k+1-j$ cannot be both odd or both even. Therefore, either $\mathrm{E}[\varphi_{k+1,k+1}] = 0$ or $\mathrm{E}[\varphi_{k,k+1-j}] = 0$. Then, we have $\mathrm{E}[\varphi_{k+1,j}] = 0$.

Case: j is even. When k is even, both $\mathrm{E}[\varphi_{k+1,k+1}]$ and $\mathrm{E}[\varphi_{k,k+1-j}]$ equal zero, whereas when k is odd, both of the above expectations. are less than zero. In both situations, we have $\mathrm{E}[\varphi_{k+1,j}] < 0$ by the Levinson–Durbin formula.

From the Levinson–Durbin formula (2.3), when p is odd and $j < p$, we have

$$\mathrm{E}[\phi_j] = \mathrm{E}[\varphi_{pj}] = \mathrm{E}[\varphi_{p-1,j}] - \mathrm{E}[\alpha_p]\mathrm{E}[\varphi_{p-1,p-j}] = \mathrm{E}[\varphi_{p-1,j}],$$

and the latter is $\mathrm{E}(\phi_j)$ for AR($p-1$). Here, we have used $\mathrm{E}[\alpha_p] = 0$.

Proof of (II). Uniform $\boldsymbol{\rho}_p$: This can be seen by noting that the one to one transformation

$$(\rho_1, \rho_2, \rho_3, \ldots, \rho_p) \longmapsto (-\rho_1, \rho_2, -\rho_3, \ldots, (-1)^p\rho_p)$$

preserves the positive definiteness of $\mathbf{R}_p$. Therefore, $\mathrm{E}[\rho_j] = 0$ is a result of symmetry. In addition, we have $\mathrm{E}[h(\rho_j)] = 0$ for any odd function $h(\cdot)$.

Uniform $\boldsymbol{\phi}_p$. Note that the above transformation is equivalent to

$$(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_p) \longmapsto (-\alpha_1, \alpha_2, -\alpha_3, \ldots, (-1)^p\alpha_p).$$

The density function of ρ remains unchange under this transformation, because α_k have symmetric Beta distributions on $(-1, 1)$ for k odd. $\square$

We next show some results for $p = 3$ to illustrate the effect of η in Section 2.2. By making the joint density of $\boldsymbol{\rho}_p$ proportional to $|\mathbf{r}|^{\eta-1}$, increasing η leads to decreasing variance in the ρ_j and ϕ_j. The equations for the variances are obtained via Maple as:

- $\mathrm{Var}(\rho_1) = 1/(6\eta + 1)$;
- $\mathrm{Var}(\rho_2) = 4\eta(9\eta + 2)/[(4\eta + 1)(6\eta + 1)^2]$;

Table 2
Means and variances for MA(q) with $q = 3$ and $q = 4$.

Variable	Uniform $\boldsymbol{\rho}_q$		Uniform $\boldsymbol{\theta}_q$	
	Exp. value	Variance	Exp. value	Variance
$q = 3$				
ρ_1	0	0.104	0	0.242
ρ_2	0.084	0.054	0.104	0.097
ρ_3	0	0.039	0	0.049
θ_1	0	0.167	0	0.733
θ_2	−0.140	0.094	−0.333	0.356
θ_3	0	0.076	0	0.200
$q = 4$				
ρ_1	0	0.101	0	0.267
ρ_2	0.096	0.065	0.137	0.143
ρ_3	0	0.041	0	0.070
ρ_4	0.057	0.029	0.067	0.031
θ_1	0	0.155	0	0.800
θ_2	−0.140	0.106	−0.400	0.587
θ_3	0	0.071	0	0.373
θ_4	−0.088	0.062	−0.200	0.160

- $\mathrm{Var}\,(\rho_3) = (72\eta^2 + 18\eta + 5)/[(4\eta + 1)(6\eta + 1)(6\eta + 5)]$;
- $\mathrm{Var}\,(\phi_1) = (2\eta + 3)/[(2\eta + 1)(6\eta + 1)]$;
- $\mathrm{Var}\,(\phi_2) = 3(2\eta + 1)/[(4\eta + 1)(6\eta + 1)]$;
- $\mathrm{Var}\,(\phi_3) = 1/(2\eta + 1)$.

The non-zero mean is $\mathrm{E}\,(\rho_2) = 1/(6\eta + 1)$.

Finally, we go to the MA(q) model as discussed in Section 3. The MA coefficients can be simulated to be uniform in the $\boldsymbol{\theta}_q$ space, from the pseudo-partial autocorrelations $\tilde{\alpha}_j \sim \mathrm{Beta}([(j + 1)/2], [j/2] + 1)$ on the interval $(-1, 1)$. Also, random MA coefficients corresponding to uniform in the $\boldsymbol{\rho}_q$ space can be obtained. Table 2 has some simulation results for MA(3) and MA(4) models; the means and variances are obtained from 10^6 replications. Note that the mean and variance of θ_j for uniform $\boldsymbol{\theta}_q$ coincide with the results of $\boldsymbol{\phi}_j$ for uniform $\boldsymbol{\phi}_p$. As before, there is smaller variances for uniform in the $\boldsymbol{\rho}_q$ space.

Some patterns shown in Table 2 are summarized in the next proposition.

Proposition 4.4. (I) *For pseudo-partial autocorrelations $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_q$ leading to uniform in the $\boldsymbol{\theta}_q$ space or the $\boldsymbol{\rho}_q$ space, we have $\mathrm{E}\,[\theta_j] = 0$ and $\mathrm{E}\,[\rho_j] = 0$ for j odd.*
(II) *For pseudo-partial autocorrelations $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_q$ leading to uniform in the $\boldsymbol{\theta}_q$ space, we have $\mathrm{E}\,[\theta_j] < 0$ for j even.*

Proof. The conclusion (I) can be seen by considering the following three bijections,

$$(\rho_1, \rho_2, \rho_3, \ldots, \rho_q) \longmapsto (-\rho_1, \rho_2, -\rho_3, \ldots, (-1)^q\rho_q)$$
$$(\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \ldots, \tilde{\alpha}_q) \longmapsto (-\tilde{\alpha}_1, \tilde{\alpha}_2, -\tilde{\alpha}_3, \ldots, (-1)^q\tilde{\alpha}_q)$$
$$(\theta_1, \theta_2, \theta_3, \ldots, \theta_q) \longmapsto (-\theta_1, \theta_2, -\theta_3, \ldots, (-1)^q\theta_q).$$

The above transformations are equivalent to each other. Therefore, the conclusions are a result of symmetry. The conclusion (II) is a consequence of Proposition 4.3 by considering the equivalence of $\boldsymbol{\theta}_q$ space and $\boldsymbol{\phi}_q$ space. □

5. Application of random Toeplitz matrix

In this section, we give an example to illustrate the usefulness of random Toeplitz matrices for non-normal time series. We apply the simulation methods developed in Section 2.2 to check if a Toeplitz matrix can be a Spearman rank correlation for a given time series model, and estimate the proportion of such Toeplitz matrices relative to all Toeplitz correlation matrices of a fixed dimension.

Let F be a continuous univariate cumulative distribution function. One example of a model with stationary margin F is constructed as follows. Let Φ be the standard normal cumulative distribution function. Let $\{X_t\}$ is a stationary zero-mean Gaussian time series, and let $Y_t = F^{-1}(\Phi(X_t))$, so that $\{Y_t\}$ is a stationary time series with univariate margin F. This model is used in Biller and Nelson [11].

The serial correlation $\mathrm{Corr}(Y_t, Y_{t+k})$ depends on F, but the rank correlation $\mathrm{Corr}(F(Y_t), F(Y_{t+k}))$ does not. The rank correlation corresponds to the correlation when F is the $U(0, 1)$ cumulative distribution function. For the bivariate normal distribution with correlation parameter ρ, the rank correlation is $\rho_r = 6\pi^{-1} \arcsin(\rho/2)$.

For a stationary time series model with non-normal margin, we might be interested in the possible Toeplitz matrices of order $p+1$. We can do the following to get a proportion. Start with a $(p+1) \times (p+1)$ Toeplitz matrix with entries $\rho_{r1}, \ldots, \rho_{rp}$

Table 3
Proportion of Toeplitz matrices that can be served as a Ranked Correlation matrix of $\Phi(X_t)$ with $\{X_t\}$ being stationary zero-mean AR(p) Gaussian time series.

p	Proportion	p	Proportion
3	0.943	11	0.667
4	0.908	12	0.637
5	0.870	13	0.610
6	0.835	14	0.582
7	0.798	15	0.557
8	0.764	16	0.532
9	0.730	17	0.509
10	0.699	18	0.485

for lags 1 to p. Let $\rho_j = 2\sin(\pi\rho_{rj}/6)$ for $j = 1, \ldots, p$. If the Toeplitz matrix with $\boldsymbol{\rho}_p = (\rho_1, \ldots, \rho_p)$ is positive definite, then it corresponds to a stationary zero-mean AR(p) Gaussian time series $\{X_t\}$ and $\rho_{r1}, \ldots, \rho_{rp}$ are the serial correlations for $\{\Phi(X_t)\}$. By simulating random $(\rho_{r1}, \ldots, \rho_{rp})$ uniform in M_p, and determining the proportion for which the Toeplitz matrix with $\boldsymbol{\rho}_p$ is positive definite, we will get the proportion of Toeplitz correlation matrices that can be serial correlations of some stationary time series of the form $\{\Phi(X_t)\}$.

The simulation results are summarized in Table 3. Each estimated value is obtained from 10^6 replications and the standard errors are from 0.0002 to 0.0005 so we report three decimal places. The proportions do not decrease to 0 as fast as in the case of the correlation matrix based on the general normal-to-anything method (compare with page 82 of [12]).

More generally in dependence modelling (see [13]), multivariate models are compared in the range of dependence that they cover. For models for non-normal time series and longitudinal data for which autocorrelations are a reasonable dependence measure, the inequalities for the autocorrelations might be complicated. Examples are integer-valued moving average time series in Al-Osh and Alzaid [14], integer-valued autoregressive time series in Al-Osh and Alzaid [15], methods for generating binary longitudinal data with m-dependence in Lunn and Davies [16]. The range of dependence of different models can be assessed via the proportion of the relevant Toeplitz matrix space that is covered.

6. Discussion

The main idea in this paper has been to derive distributions, include uniform, of (a) $\boldsymbol{\rho}_p$ and $\boldsymbol{\phi}_p$ for AR(p) based on the partial autocorrelation vector $\boldsymbol{\alpha}_p \in (-1, 1)^p$, and (b) $\boldsymbol{\rho}_q$ and $\boldsymbol{\theta}_q$ for MA(q), based on the pseudo-partial autocorrelation vector $\tilde{\boldsymbol{\alpha}}_q \in (-1, 1)^q$. The (pseudo)-partial autocorrelations are algebraically independent, whereas the constraints of $\boldsymbol{\rho}_p$, $\boldsymbol{\phi}_p$, $\boldsymbol{\rho}_q$ and $\boldsymbol{\theta}_q$ are non-linear.

We have shown that there are big differences in the behavior of the autocorrelations and AR or MA coefficients when generating at random uniformly from the autocorrelation $\boldsymbol{\rho}$ space versus uniformly from the coefficient vector $\boldsymbol{\phi}$ or $\boldsymbol{\theta}$ space. Researchers who want to simulate random parameters for a Gaussian time series model or a model with a structured Toeplitz correlation matrix need to think carefully about the choice of distributions for the partial autocorrelations.

The main results that we derive include the Jacobians in Sections 2.1 and 2.2 and Proposition 3.1. Combining the Jacobian with any distribution for $\boldsymbol{\alpha}_p$ or $\tilde{\boldsymbol{\alpha}}_q$ and we can get non-uniform distributions for $\boldsymbol{\rho}$ consistent with AR(p) or MA(q). For applications such as random effects for $\boldsymbol{\rho}$ or Bayesian inference with historical information, one might want to choose distributions for $\boldsymbol{\alpha}_p$ or $\tilde{\boldsymbol{\alpha}}_q$ so that the distribution of $\boldsymbol{\rho}$ is centered at a Toeplitz matrix with positive dependence, rather than having a simple density that is proportional to a power of $\det(\mathbf{r})$. More specifically, consider a large data set with longitudinal data for many subjects or multivariate time series (univariate time series for many different related financial/economic variables). The model can be simplified if we assume that the autocorrelation parameters of the marginal time series are random with a parametric mixing distribution. Our theory provides a way to specify the mixing distribution by choosing appropriate distributions for $\boldsymbol{\alpha}_p$ or $\tilde{\boldsymbol{\alpha}}_q$. If a mixing density that is proportional to a power of $\det(\mathbf{r})$, is appropriate, then the parameter η or δ in (2.1) could be estimated with an empirical Bayesian approach.

Acknowledgments

This research has been supported by an NSERC Canada Discovery Grant and a Start-up grant from Hong Kong Polytechnic University. We are grateful to the referees for their valuable comments.

References

[1] S. Ghosh, S.G. Henderson, Behavior of the NORTA method for correlated random vector generation as the dimension increases, ACM Trans. Model. Comput. Simul. 13 (2003) 276–294.
[2] H. Joe, Generating random correlation matrices based on partial correlations, J. Multivariate Anal. 97 (2006) 2177–2189.
[3] P.X.-K. Song, Correlated Data Analysis, Springer, New York, 2007.
[4] R.B. Holmes, On random correlation matrices. II. The Toeplitz case, Commun. Stat. -Simul. 18 (1989) 1511–1537.
[5] M.C. Jones, Randomly choosing parameters for the stationarity and invertibility region of autoregressive-moving average models, Appl. Stat. 36 (1987) 134–148.

[6] O. Barndorff-Nielsen, G. Schou, On the parametrization of autoregressive model by partial autocorrelations, J. Multivariate Analysis 3 (1973) 408–419.
[7] J. Makhoul, Volume of the space of positive definite sequences, IEEE Trans. Acoust. Speech Signal Process. 38 (1990) 506–511.
[8] M.J. Daniels, M. Pourahmadi, Modeling covariance matrices via partial autocorrelations, J. Multivariate Analysis 100 (2009) 2352–2363.
[9] N. Levinson, The Wiener RMS (root mean square) error criterion in filter design and prediction, J. Math. Phys. Mass. Inst. Tech. 25 (1947) 261–278.
[10] D. Piccolo, The size of the stationary and invertibility region of an autoregressive-moving average process, J. Time Ser. Anal. 3 (1982) 245–247.
[11] B. Biller, B.L. Nelson, Fitting time-series input processes for simulation, Open. Res. 53 (2005) 549–559.
[12] D. Kurowicka, R. Cooke, Uncertainty Analysis with High Dimensional Dependence Modelling, Wiley, Chichester, 2006.
[13] H. Joe, Multivariate Models and Dependence Concepts, Chapman & Hall, London, 1997.
[14] M. Al-Osh, A.A. Alzaid, Integer-valued moving average (INMA) process, Statist. Papers 29 (1988) 281–300.
[15] A.A. Alzaid, M. Al-Osh, An integer-valued *p*th order autoregressive structure (INAR(p)) process, J. Appl. Probab. 27 (1990) 314–324.
[16] A.D. Lunn, S.J. Davies, A note on generating correlated binary variables, Biometrika 85 (1998) 487–490.