Pergamon

Mathl. Comput. Modelling Vol. 22, No. 2, pp. 1–6, 1995
Copyright©1995 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0895-7177/95 $9.50 + 0.00

0895-7177(95)00106-9

The Simulation of Random Vector Time Series with Given Spectrum

M. J. CHAMBERS
Department of Economics, University of Essex
Wivenhoe Park, Colchester, Essex CO4 3SQ, England
mchamb@essex.ac.uk

(Received and accepted May 1995)

Abstract—A method is proposed for generating multivariate time series which are required to satisfy a given spectral density function, which extends previous work on univariate time series. The performance of the method is assessed in a small simulation exercise for a bivariate long memory model and is found to perform well.

Keywords—Simulation, Time series, Fourier transform, Spectrum, Long memory.

1. INTRODUCTION

In many disciplines, as diverse as economics, statistics, mathematics, engineering, physics, environmental science and medicine, there often arises the need to simulate a time series which possesses certain known properties. A common example in statistics and econometrics is where Monte Carlo simulations are used in order to investigate the finite sample performance of an estimator or test statistic. In many circumstances, the requirement is that the data be generated according to a given parametric model. When the model of interest is a member of the class of finite order autoregressive integrated moving average models, the generation of data is a relatively straightforward task. But in situations where the model is of infinite order in either the autoregressive or moving average operators, the finite truncation of the infinite order operator can result in the generated data only approximately satisfying the required properties. In other cases, it is more convenient to express the theoretical properties of the data in terms of an autocorrelation function or spectral density function. An area where this approach is common is in the analysis of long memory (or long-range dependent) time series which, when based on fractional differences, have a much neater representation in the frequency domain than in the time domain (this is also an example where an infinite order operator is involved in the time domain representation of the series).

This paper considers the problem of generating real-valued vector time series which are required to satisfy a known spectral density matrix. Since the spectral density matrix is simply the Fourier transform of the autocovariances, the method is also applicable to situations in which the data are required to satisfy a given autocovariance structure. The proposed method employs a discrete Fourier transform of standard normal complex-valued random vectors, weighted by an appropriate function of the known spectral density matrix. The method itself is related to the univariate approach of Thompson [1], and also to the algorithm of Davies and Harte [2], which

The work reported here was supported by the Economic and Social Research Council under grant number R000221195. All computations were performed using the matrix programming language GAUSS on a 486DX2-66MHz personal computer.

was designed to generate a scalar long memory time series from a given autocovariance function. The method described here, however, is applicable to the generation of vector time series, the elements of which may be contemporaneously, as well as serially, correlated. The algorithm does, of course, also specialize to the univariate case. While being particularly important in the examples mentioned in the previous paragraph for the generation of artificial data, the method may even be attractive computationally in finite-order autoregressive integrated moving average models where long time series are required. The computational attractions arise if the fast Fourier transform is employed, which can considerably speed up the computations, particularly in the generation of long time series. The method itself is detailed in the next section, and an application to the simulation of a bivariate system of long memory variables is presented in Section 3.

2. THE METHOD

The objective is to generate a real-valued q-dimensional time series $\{X_t\}_{t=1}^n$ which is required to satisfy (at least theoretically) the spectral density matrix function $F(\lambda)$ $(-\pi < \lambda \leq \pi)$. Emphasis is placed on the generation of a Gaussian series $\{X_t\}$, although this requirement could be relaxed if desired. Without loss of generality, assume that n is even, and define the real-valued q-dimensional random vectors $Z_{n/2} \sim N(0, 2I)$ and $Z_n \sim N(0, 2I)$, where I is the $q \times q$ identity matrix. In addition, for $1 \leq k \leq n/2-1$, define the complex-valued q-dimensional random vectors $Z_k = \mathrm{Re}(Z_k) + i\,\mathrm{Im}(Z_k)$, where $i^2 = -1$, such that

$$\begin{bmatrix} \mathrm{Re}(Z_k) \\ \mathrm{Im}(Z_k) \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}\right].$$

In addition, let $Z_{n-k} = \overline{Z}_k$ for $1 \leq k \leq n/2 - 1$, where $\overline{Z}_k = \mathrm{Re}(Z_k) - i\,\mathrm{Im}(Z_k)$ is the complex conjugate of Z_k. Note that, for $k \neq n/2, n$, $\mathrm{E}Z_k Z_k' = 0$ and $\mathrm{E}Z_k Z_k^* = 2I$, where Z_k^* denotes the operations of conjugation and transposition applied to Z_k, so that $Z_k^* = \overline{Z}_k'$. Since many computer packages and programming languages have facilities for the generation of standard normal variates, it is an easy task to generate $\{Z_k\}_{k=1}^{n/2}$ and Z_n.

An implication of X being real-valued is that the matrix $F(\lambda)$ is Hermitian, so that $F(\lambda) = F(\lambda)^*$, with real elements on the diagonal. The spectral density matrix can also be written $F(\lambda) = U(\lambda)M(\lambda)U(\lambda)^*$, where $M(\lambda)$ is a $q \times q$ diagonal matrix with the real-valued eigenvalues of $F(\lambda)$ on the diagonal, and $U(\lambda)$ is the complex-valued $q \times q$ matrix of associated eigenvectors, which is unitary, so that $U(\lambda)U(\lambda)^* = U(\lambda)^*U(\lambda) = I$. This decomposition of $F(\lambda)$ enables the 'square root matrix' to be defined, and is given by $F^{1/2}(\lambda) = U(\lambda)M^{1/2}(\lambda)U(\lambda)^*$, where $M^{1/2}(\lambda)$ is the diagonal matrix with the square roots of the eigenvalues on the diagonal. It is trivial to show that $F^{1/2}(\lambda)F^{1/2}(\lambda) = F(\lambda)$. Furthermore, $F(\lambda)$ is symmetric, so that $F(-\lambda) = F(\lambda)'$.

The series $\{X_t\}_{t=1}^n$ is generated using a discrete Fourier transform of linear combinations of the elements of $\{Z_k\}_{k=1}^{n/2}$ and Z_n, and is motivated by the properties of $F(\lambda)$ discussed above. Let $\{\lambda_k = 2\pi k/n; k = 1, \ldots, n\}$ denote the set of Fourier frequencies, and define the random variables

$$V_k = U_k M_k^{1/2} Z_k, \qquad k = 1, \ldots, \frac{n}{2}, n,$$

$$V_{n-k} = \overline{U}_k M_k^{1/2} \overline{Z}_k = \overline{V}_k, \qquad 1 \leq k \leq \frac{n}{2} - 1,$$

where $U_k = U(\lambda_k)$ and $M_k^{1/2} = M^{1/2}(\lambda_k)$. The result of interest is stated in the following proposition.

PROPOSITION 1. *The series*

$$X_t = \left(\frac{\pi}{n}\right)^{1/2} \sum_{k=1}^{n} V_k e^{it\lambda_k}, \qquad t = 1, \ldots, n, \tag{1}$$

is real-valued and satisfies the spectral density matrix $F(\lambda)$ for large n.

PROOF. To show that each X_t is real-valued, first note that

$$V_{n/2}e^{it\lambda_{n/2}} = V_{n/2}e^{it\pi} = V_{n/2}\cos(t\pi),$$
$$V_n e^{it\lambda_n} = V_n e^{2it\pi} = V_n,$$

are both real, and for $1 \le k \le n/2$,

$$\begin{aligned} V_k e^{it\lambda_k} + V_{n-k}e^{it\lambda_{n-k}} &= V_k e^{it\lambda_k} + \overline{V}_k e^{it(2\pi-\lambda_k)} \\ &= V_k e^{it\lambda_k} + \overline{V}_k e^{-it\lambda_k} \\ &= 2\,\mathrm{Re}\left(V_k e^{it\lambda_k}\right), \end{aligned}$$

which is also real. Hence, the sum in (1) is real for each $t = 1, \ldots, n$.

In order to show that $\{X_t\}_{t=1}^n$ is consistent with $F(\lambda)$, we proceed by showing that $\mathrm{E}X_t X'_{t-\tau} = (2\pi/n)\sum_{k=1}^n F(\lambda_k)e^{i\tau\lambda_k}$ for $\tau \ge 0$, which is a Riemann sum approximation to the autocovariance matrix $G_\tau = \int_0^{2\pi} F(\lambda)e^{i\tau\lambda}\,d\lambda$. Clearly

$$\mathrm{E}X_t X'_{t-\tau} = \frac{\pi}{n}\mathrm{E}\left[\sum_{k=1}^n V_k e^{it\lambda_k}\right]\left[\sum_{l=1}^n V'_l e^{i(t-\tau)\lambda_l}\right], \tag{2}$$

which, selecting the nonzero expectations, can be written

$$\mathrm{E}X_t X'_{t-\tau} = \frac{\pi}{n}\left[\sum_{k\neq n/2,n} \mathrm{E}V_k V'_{n-k} e^{it\lambda_k}e^{i(t-\tau)\lambda_{n-k}} + \sum_{k=n/2,n} \mathrm{E}V_k V'_k e^{it\lambda_k}e^{i(t-\tau)\lambda_k}\right]. \tag{3}$$

It can be shown that

$$\mathrm{E}V_k V'_{n-k} = U_k M_k^{1/2}\mathrm{E}Z_k Z_k^* M_k^{1/2}U_k^* = 2U_k M_k U_k^* = 2F(\lambda_k), \qquad k \neq \frac{n}{2}, n,$$
$$\mathrm{E}V_k V'_k = U_k M_k^{1/2}\mathrm{E}Z_k Z'_k M_k^{1/2}U'_k = 2U_k M_k U'_k = 2F(\lambda_k), \qquad k = \frac{n}{2}, n,$$

and, furthermore,

$$e^{it\lambda_k}e^{i(t-\tau)\lambda_{n-k}} = e^{i\tau\lambda_k}, \qquad k \neq \frac{n}{2}, n,$$
$$e^{it\lambda_k}e^{i(t-\tau)\lambda_k} = e^{i\tau\lambda_k}, \qquad k = \frac{n}{2}, n.$$

Substituting these expressions into (3) yields

$$\mathrm{E}X_t X'_{t-\tau} = \left(\frac{2\pi}{n}\right)\sum_{k=1}^n F(\lambda_k)e^{i\tau\lambda_k}$$

as required. ∎

Inspection of the proof of Proposition 1 reveals that the series $\{X_t\}_{t=1}^n$ only approximately satisfies $F(\lambda)$ for finite n. This is most easily seen by examining the autocovariance matrices given by $\mathrm{E}X_t X'_{t-\tau} = (2\pi/n)\sum_{k=1}^n F(\lambda_k)e^{i\tau\lambda_k}$ from which it follows that

$$\lim_{n\to\infty}\frac{2\pi}{n}\sum_{k=1}^n F(\lambda_k)e^{i\tau\lambda_k} = \int_0^{2\pi} F(\lambda)e^{i\tau\lambda}\,d\lambda,$$

the expression on the right-hand side denoting the true autocovariance consistent with $F(\lambda)$. A comparison of the Riemann sum approximation for finite n with the integral stated above would give some idea as to the accuracy of the approximation. Percival [3] provides some expressions in the univariate case which could be extended to the multivariate case if desired.

The periodic nature of the Fourier transform means that X_1 and X_n will be close in value, which is a less desirable feature of the proposed method. Mitchell and McPherson [4] and Percival [3] suggest simulating a series of length $m = kn$ for some integer $k \ge 2$ and then choosing a subsample from the simulated series of length n. The resulting series can be shown to also satisfy (at least approximately) the required spectral density function.

3. AN APPLICATION TO LONG MEMORY TIME SERIES

Although it is beyond the scope of this paper to conduct an extensive Monte Carlo experiment for a particular estimator or test statistic, some evidence of the usefulness of the proposed method can be obtained by repeatedly simulating a vector long memory time series and comparing the resulting estimated autocovariances with their required theoretical values. In order to achieve this, consider the simple q-vector fractionally integrated model defined by

$$D(L)X_t = u_t, \qquad t = 1, \ldots, n, \tag{4}$$

where L denotes the lag operator such that $L^j X_t = X_{t-j}$, and

$$D(L) = \begin{bmatrix} (1-L)^{d_1} & 0 & \cdots & 0 \\ 0 & (1-L)^{d_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & (1-L)^{d_q} \end{bmatrix},$$

where $-1/2 \leq d_j \leq 1/2$ for $j = 1, \ldots, q$. For values of the fractional differencing parameter d_j in this range, the series $\{X_t\}_{t=1}^n$ is stationary and invertible, provided that $\{u_t\}_{t=1}^n$ is also a stationary process; see [5]. With regard to u_t, it will be assumed that, for all $t = 1, \ldots, n$,

$$\mathrm{E}(u_t) = 0, \quad \mathrm{E}(u_t u_t') = \Omega, \quad \mathrm{E}(u_t u_s') = 0, \qquad \text{for } s \neq t.$$

The generation of a sequence of values for long memory X_t can be difficult due to the fractional differencing operator $(1-L)^{d_j}$. Although this operator has the binomial representation

$$(1-L)^{d_j} = 1 + \sum_{k=1}^{\infty} \frac{\Gamma(k-d_j)}{\Gamma(-d_j)\Gamma(k+1)} L^k,$$

using this expression to generate an artificial series of finite length will, by necessity, have to use a finite truncation which can potentially distort the properties of the resulting series. An alternative approach, used by Geweke and Porter-Hudak [6] in the univariate case, is to derive the autocovariance function for X_t and then to use a Cholesky decomposition of the autocovariance matrix of $\{X_1, \ldots, X_n\}$ with which to transform a sequence of independent random variates. This method also relies on a finite truncation, this time of the autocovariance function, and can result in departures of the simulated series from its desired properties.

The method proposed in this paper would, therefore, appear to be useful in generating a sequence of values satisfying (4). The spectral density function for X is given by

$$F(\lambda) = \frac{1}{2\pi} D\left(e^{-i\lambda}\right)^{-1} \Omega D\left(e^{i\lambda}\right)^{-1}, \qquad -\pi < \lambda < \pi, \tag{5}$$

and, consistent with long memory time series, possesses a singularity at the origin ($\lambda = 0$). For $-1/2 < d_j < 0$, this singularity is a zero, so that $F_{jj}(\lambda) \to 0$ as $\lambda \to 0$, where $F_{jj}(\lambda)$ is the j^{th} diagonal element of $F(\lambda)$. Alternatively, for $0 < d_j < 1/2$, the singularity is a pole, and $F_{jj}(\lambda) \to \infty$ as $\lambda \to 0$. In fact, $F_{jj}(\lambda)$ behaves like λ^{-2d_j} in the neighbourhood of $\lambda = 0$. Due to this singularity, the term $V_n e^{it\lambda_n}$ should be omitted from (1) in the generation of long memory series.

Letting $F_{jk}(\lambda)$ denote the j,k^{th} element of $F(\lambda)$, it can be shown from (5) that

$$F_{jk}(\lambda) = \frac{\omega_{jk}}{2\pi} \left(1 - e^{-i\lambda}\right)^{-d_j} \left(1 - e^{i\lambda}\right)^{-d_k}, \qquad j,k = 1, \ldots, q, \tag{6}$$

where ω_{jk} is the j,k^{th} element of Ω. The autocovariance matrices are given by

$$G_\tau = \mathrm{E} X_t X_{t-\tau}' = \int_0^{2\pi} F(\lambda) e^{i\tau\lambda}\, d\lambda, \tag{7}$$

the j,k^{th} element of which, using (6), is given by

$$G_{\tau,jk} = \frac{\omega_{jk}}{2\pi}\int_0^{2\pi}\left(1-e^{-i\lambda}\right)^{-d_j}\left(1-e^{i\lambda}\right)^{-d_k}e^{i\tau\lambda}\,d\lambda. \tag{8}$$

A compact expression for $G_{\tau,jk}$ is given in the following proposition.

PROPOSITION 2. *The elements of the autocovariance matrices G_τ are given by*

$$G_{\tau,jk} = \frac{(-1)^\tau\omega_{jk}\Gamma(1-d_j-d_k)}{(1-d_j-d_k)\Gamma(1+\tau-d_k)\Gamma(1-\tau-d_j)}, \qquad j,k=1,\ldots,q.$$

PROOF. It is required to find an expression for the integral in (8). Note first that

$$1-e^{\pm i\lambda} = 2\sin\left(\frac{\lambda}{2}\right)e^{\pm i(x-\pi)/2}$$

and hence,

$$\left(1-e^{-i\lambda}\right)^{-d_j}\left(1-e^{i\lambda}\right)^{-d_k} = 2^{-d_j-d_k}e^{i\pi(d_k-d_j)/2}\sin^{-d_j-d_k}\left(\frac{\lambda}{2}\right)e^{i\lambda(d_j-d_k)/2},$$

so that

$$G_{\tau,jk} = \frac{\omega_{jk}e^{i\pi(d_k-d_j)/2}}{2\pi 2^{d_j+d_k}}\int_0^{2\pi}\sin^{-d_j-d_k}\left(\frac{\lambda}{2}\right)e^{i\lambda(2\tau+d_j-d_k)/2}\,d\lambda.$$

Using the change of variable $\mu=\lambda/2$ in the integral, this becomes

$$G_{\tau,jk} = \frac{\omega_{jk}e^{i\pi(d_k-d_j)/2}}{\pi 2^{d_j+d_k}}\int_0^{\pi}\sin^{-d_j-d_k}(\mu)e^{i\mu(2\tau+d_j-d_k)}\,d\mu.$$

Equation 3.893.1 of Gradshteyn and Ryzhik [7] gives

$$\int_0^\pi e^{i\beta x}\sin^{\nu-1}x\,dx = \frac{\pi e^{i\beta\pi/2}}{2^{\nu-1}\nu B\left(\frac{\nu+\beta+1}{2},\frac{\nu-\beta+1}{2}\right)}, \qquad [\operatorname{Re}\nu>-1],$$

where $B(x,y)$ is the beta function $B(x,y)=\Gamma(x)\Gamma(y)/\Gamma(x+y)$ and $\Gamma(x)$ is the gamma function $\Gamma(x)=\int_0^\infty t^{x-1}e^{-t}\,dt$. In the integral of interest, $\beta=2\tau+d_j-d_k$, $\nu=1-d_j-d_k$ and $x=\mu$. Using the above result gives the expression for $G_{\tau,jk}$ given in the proposition, which also uses the fact that $e^{i\pi\tau}=(-1)^\tau$ for integer τ. ∎

The results of applying the method of data generation to the model (4) are summarized in Table 1 for a bivariate ($q=2$) model with $d_1=0.1$, $d_2=0.4$ and an assumed innovations covariance matrix Ω given by

$$\Omega = \begin{bmatrix} 1.0 & -0.5 \\ -0.5 & 0.1 \end{bmatrix}.$$

Data samples of size $n=256$ were simulated 1000 times and a subsample of 128 observations was selected at each replication in accordance with the recommendation of Percival [3] discussed earlier. Table 1 contains the mean values of selected short- and long-range autocorrelations from the simulated series as well as the theoretical values of the autocorrelations for comparison. The theoretical values of the autocorrelation matrices are calculated as

$$R_\tau = C_0 G_\tau C_0,$$

with G_τ defined in Proposition 2, $C_0=\operatorname{diag}\{1/G_{0,kk}^{1/2}\}$, and $G_{0,kk}$ denoting the k^{th} diagonal element of G_0. The values of these matrices are estimated from the simulated series using

$$\widehat{R}_\tau = \widehat{C}_0\widehat{G}_\tau\widehat{C}_0,$$

Table 1. Mean estimated autocorrelation matrices from 1000 replications; sample size 128.

Lag length	Mean	Theoretical
1	$\begin{bmatrix} 0.13 & -0.12 \\ -0.39 & 0.47 \end{bmatrix}$	$\begin{bmatrix} 0.11 & -0.16 \\ -0.43 & 0.67 \end{bmatrix}$
5	$\begin{bmatrix} 0.03 & -0.09 \\ -0.17 & 0.41 \end{bmatrix}$	$\begin{bmatrix} 0.03 & -0.07 \\ -0.21 & 0.49 \end{bmatrix}$
10	$\begin{bmatrix} 0.01 & -0.06 \\ -0.13 & 0.35 \end{bmatrix}$	$\begin{bmatrix} 0.02 & -0.05 \\ -0.15 & 0.42 \end{bmatrix}$
25	$\begin{bmatrix} 0.01 & -0.03 \\ -0.07 & 0.27 \end{bmatrix}$	$\begin{bmatrix} 0.01 & -0.03 \\ -0.09 & 0.35 \end{bmatrix}$
50	$\begin{bmatrix} 0.00 & -0.02 \\ -0.05 & 0.21 \end{bmatrix}$	$\begin{bmatrix} 0.00 & -0.02 \\ -0.07 & 0.31 \end{bmatrix}$

where $\widehat{G}_\tau = (n-\tau)^{-1}\sum_{t=\tau+1}^{n} X_t X_{t-\tau}'$ and $\widehat{C}_0 = \mathrm{diag}\{1/\widehat{G}_{0,kk}^{1/2}\}$. Autocorrelation matrices for lag lengths of $\tau = 1, 5, 10, 25$ and 50 are reported in Table 1, all figures being reported to two decimal places.

Although the mean values of the autocorrelations do show some sampling deviations from the theoretical values, the proposed method of data generation is certainly able to mimic the broad pattern of autocorrelation required of the series on average.

In conclusion, the method of simulating random vector time series subject to a known spectral density function appears to perform well in practice, at least in the simple long memory example given. The proposed algorithm can exploit the fast Fourier transform to gain greatest computational efficiency, and has potential applications in Monte Carlo experimentation in all areas of time series analysis.

REFERENCES

1. R. Thompson, Generation of stochastic processes with given spectrum, *Utilitas Mathematica* **3**, 127–137 (1973).
2. R.B. Davies and D.S. Harte, Tests for Hurst effect, *Biometrika* **74**, 95–101 (1987).
3. D.B. Percival, Simulating Gaussian random processes with specified spectra, *Computing Science and Statistics* **24**, 534–538 (1992).
4. R.L. Mitchell and D.A. McPherson, Generating nonstationary random sequences, *IEEE Trans. Aero. Elec. Sys.* **17**, 553–560 (1981).
5. J.R.M. Hosking, Fractional differencing, *Biometrika* **68**, 165–176 (1981).
6. J. Geweke and S. Porter-Hudak, The estimation and application of long memory time series models, *Journal of Time Series Analysis* **4**, 221–238 (1983).
7. I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series and Products*, 5th edition, Academic Press, London, (1994).