Key Moments in the Rouwenhorst Method*

Damba Lkhagvasuren†
Concordia University and CIREQ

September 14, 2012

Abstract

This note characterizes the underlying structure of the autoregressive process generated by Rouwenhorst (1995) and calculates the key moments of the process for its general case. It also addresses the main issues that arise when targeting the shape of the distribution of the process. Moreover, the note finds a close link between skewness and kurtosis of the process. It proposes a simple technique to break the link and thus provides a flexible tool for targeting higher order moments in the presence of high persistence.

Keywords: Markov Chain, Autoregressive Processes, Numerical Methods, Moment Matching

JEL Codes: C15, C60

*The author thanks Gordon Fisher, Nikolay Gospodinov and Purevdorj Tuvaandorj for helpful comments.
†Department of Economics, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, Quebec H3G 1M8, Canada. Email: damba.lkhagvasuren@concordia.ca.

1 Introduction

Finite-state Markov chain approximation methods are widely used in solving functional equations where the state variables follow autoregressive processes. Recent work of Galindev and Lkhagvasuren (2010) and Kopecky and Suen (2010) shows that for a wide range of the parameter space the method of Rouwenhorst (1995) outperforms other commonly-used methods (e.g., Tauchen, 1986 and Tauchen and Hussey, 1991)[1] along the key first- and second-order moments. Although these lower order moments reveal the importance of the Rouwenhorst (1995) method for solving functional equations and dynamic models, they cannot capture the exact nature of the autoregressive processes generated by the method.

This note characterizes the underlying structure of the Markov chain generated by the Rouwenhorst method. It calculates the key moments of the Markov chain, including skewness and kurtosis,[2] and addresses the issue of how the parameters of the method affect the extent of asymmetry and tail thickness of the process generated by the method. This is important for the obvious reason that, unlike other commonly-used methods, the transition matrix in the Rouwenhorst method is not guided by a particular distribution function.[3]

The note also finds that there is a strong trade-off between targeting skewness and targeting kurtosis under the Rouwenhorst method. It proposes a simple technique to break the trade-off and thus provides a flexible tool for matching higher order moments in the presence of high persistence. Finally, the analysis contained in this note provides further insight into constructing autoregressive processes that are more flexible than those generated by the Rouwenhorst method alone.

The outline of the rest of the note is as follows. Sections 2 and 3 characterize the Markov chain constructed by the Rouwenhorst method. The key moments of the process are calculated in Section 4. Section 5 demonstrates how to choose the key parameters of the method when approximating continuous autoregressive processes. It also extends the application of the method to skewness and kurtosis. Section 6 summarizes the conclusions.

2 Two-state Markov chains

This section establishes the main properties of the well known two-state Markov process. The results will be useful in calculating the key moments of the Rouwenhorst method. Consider $n-1$ independent, identical, two-state Markov chains $\{x_1, x_2, \ldots, x_{n-1}\}$. Assume that each x_i takes on the value of 0 or 1. Let the probability that the process transits from 0 to 1 be

[1] See Adda and Cooper (2003), Floden (2008) and Terry and Knotek II (2011) for extensions of Tauchen (1986) and Tauchen and Hussey (1991).

[2] In this note, kurtosis refers to excess kurtosis.

[3] See Gospodinov and Lkhagvasuren (2012) for a formal discussion that calculating transition probabilities using the probability density function cannot always deliver a meaningful finite-state approximation.

$1-p$ and the probability that the process transits from 1 to 0 be $1-q$, where $0 \leq p < 1$ and $0 \leq q < 1$. Let Π_2 denote the transition matrix given by these probabilities:

$$\Pi_2 = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}. \quad (1)$$

2.1 Unconditional moments of x_i

Given Π_2, the stationary distribution of x_i is given by $\text{Prob}[x_i = 0] = 1-\alpha$ and $\text{Prob}[x_i = 1] = \alpha$, where $\alpha = \frac{1-p}{2-p-q}$. Then, it follows that, for each i, $E[x_i] = \alpha$ and $Var[x_i] = \alpha - \alpha^2$. Moreover, by letting x_i' denote the next period's value of x_i, the autocorrelation coefficient of x_i is given by $Corr[x_i, x_i'] = p + q - 1$.

Note that x_i follows the Bernoulli distribution which takes the value 1 with success probability α and the value 0 with failure probability $1-\alpha$. Accordingly, the skewness and kurtosis of x_i are given by $Skew[x_i] = (1-2\alpha)/\sqrt{\alpha(1-\alpha)} = (p-q)/\sqrt{(1-q)(p-q)}$ and $Kurt[x_i] = -6 + 1/(\alpha(1-\alpha)) = -2 + (p-q)^2/((1-q)(1-p))$, respectively.

2.2 Conditional moments of x_i

It can be seen that, depending on whether the current state x_i is 0 or 1, the next state x_i' is Bernoulli distributed with success probability $1-p$ or q, respectively. Therefore, $E[x_i'|x_i = 0] = 1-p$, $E[x_i'|x_i = 1] = q$, $Var[x_i'|x_i = 0] = p(1-p)$ and $Var[x_i'|x_i = 1] = q(1-q)$.

Also, given the two conditional Bernoulli distributions, the conditional skewness is given by $Skew[x_i'|x_i = 0] = (2p-1)/\sqrt{p(1-p)}$ and $Skew[x_i'|x_i = 1] = (1-2q)/\sqrt{q(1-q)}$. Analogously, the conditional kurtosis is given by $Kurt[x_i'|x_i = 0] = -6 + 1/(p(1-p))$ and $Kurt[x_i'|x_i = 1] = -6 + 1/(q(1-q))$.

3 An n-state Markov chain

Next we construct an n-state Markov process using the above two-state processes.

3.1 An auxiliary n-state process

Let $\tilde{x}_n$ be defined as the sum of the above $n-1$ independent, identical two-state processes:

$$\tilde{x}_n = x_1 + x_2 + \cdots + x_{n-1}. \quad (2)$$

By construction, $\tilde{x}_n$ is a Markov process that takes n discrete values: $\{0, 1, \ldots, n-1\}$. Therefore, $\tilde{x}_n = k$ implies that k of the above $n-1$ two-state Markov processes are at state 1 and the remaining $n-k-1$ of them are at 0. It is important to keep this point in mind when calculating the moments of $\tilde{x}$.

3.2 An n-state Rouwenhorst process

Now consider the following n-state Markov chain:

$$y_n = a(n-1) + b\tilde{x}_n = \sum_{i=1}^{n-1}(a + bx_i), \tag{3}$$

where $b > 0$. Choose a and b so that y_n takes equispaced values on the interval $(m-\Delta, m+\Delta)$ for some $m > 0$ and $\Delta > 0$. It can be shown that $a = (m-\Delta)/(n-1)$ and $b = 2\Delta/(n-1)$. Consequently, the k-th grid point (in ascending order) of y_n is given by

$$\bar{y}_n^k = m - \Delta + 2\Delta(k-1)/(n-1), \tag{4}$$

where $k = \{1, 2, \ldots, n\}$.

Given the monotonic relationship in equation (3), the Markov chains $\tilde{x}_n$ and y_n share a common transition matrix. Let the common transition matrix of these two n-state Markov chains be Π_n. Clearly, when $n = 2$, the probability transition matrix is given by equation (1). For higher values of n, the transition probability matrix can be constructed recursively using the elements of matrix Π_2. In fact, this is done by Rouwenhorst (1995).[4] Thus, the transition matrix Π_n and the grid points $\{\bar{y}_n^1, \bar{y}_n^k, \cdots, \bar{y}_n^n\}$ give the Markov chain constructed Rouwenhorst (1995). More important, equations (1) and (2), along with the linear transformation in equation (3), characterize the underlying structure of the Markov chain constructed by the Rouwenhorst method.

4 Key moments

Now we establish the key moments of y_n using those of x_i.

4.1 Unconditional moments

It follows from equation (3) that $E[y_n] = (n-1)(a + bE[x_i])$ and $Var[y_n] = (n-1)b^2 Var[x_i]$. Then, using the results in Section 2,

$$E[y_n] = m + \Delta\frac{q-p}{2-p-q}$$

and

$$Var[y_n] = \frac{4\Delta^2}{n-1}\frac{(1-p)(1-q)}{(2-p-q)^2}.$$

By the linear dependence between y_n and $\tilde{x}_n$, $Corr[y_n, y_n'] = Corr[\tilde{x}_n, \tilde{x}_n']$. On the other hand, since $\{x_1, x_2, \ldots, x_{n-1}\}$ are independent and equally-persistent, $Corr[\tilde{x}_n, \tilde{x}_n'] = Corr[x_i, x_i']$. Therefore,

$$Corr[y_n, y_n'] = p + q - 1.$$

[4] A Matlab code of constructing the matrix is available upon request.

The latter shows that the serial correlation is independent of the number of grid points n. This is a highly desirable feature of the method because the accuracy of the other commonly used methods is highly sensitive to the number of grid points.

Next we calculate the skewness and kurtosis of y_n as those of a sum of independent random variables. For this purpose, we note that, for M independent random variables $\{z_1, z_2, \cdots, z_M\}$, the following identities hold:

$$Skew\left[\sum_{i=1}^{M} z_i\right] = \sum_{i=1}^{M} \omega_i^S Skew[z_i] \tag{5}$$

and

$$Kurt\left[\sum_{i=1}^{M} z_i\right] = \sum_{i=1}^{M} \omega_i^K Kurt[z_i], \tag{6}$$

where $\omega_i^S = Var[z_i]^{3/2} / \left(\sum_{i=1}^{M} Var[z_i]\right)^{3/2}$ and $\omega_i^K = Var[z_i]^2 / \left(\sum_{i=1}^{M} Var[z_i]\right)^2$. Using these equations and noting that the distributions of y_n and $\tilde{x}_n$ have the same skewness and kurtosis,

$$Skew[y_n] = \frac{Skew[x_i]}{\sqrt{n-1}} = \frac{p-q}{\sqrt{(n-1)(1-p)(1-q)}} \tag{7}$$

and

$$Kurt[y_n] = \frac{Kurt[x_i]}{n-1} = \frac{1}{n-1}\left(-2 + \frac{(p-q)^2}{(1-q)(1-p)}\right). \tag{8}$$

4.2 Conditional moments

Recall that $\tilde{x}_n = k-1$ (or, equivalently, $\tilde{x}_n$ is in its k-th lowest state) means that $k-1$ of $n-1$ two-state Markov processes are at state 1 while the remaining $n-k$ of them are at 0. Therefore, $E[\tilde{x}'_n|\tilde{x}_n = k-1] = (k-1)E[x'_i|x_i = 1] + (n-k)E[x'_i|x_i = 0]$ and $Var[\tilde{x}'_n|\tilde{x}_n = k-1] = (k-1)Var[x'_i|x_i = 1] + (n-k)Var[x'_i|x_i = 0]$. Combining these with the results of Section 2 and the linear transformation in equation (3),

$$E[y'_n|y_n = \bar{y}_n^k] = m - \Delta + \frac{2\Delta}{n-1}\left[(k-1)q + (n-k)(1-p)\right] \tag{9}$$

and

$$Var[y'_n|y_n = \bar{y}_n^k] = \frac{4\Delta^2}{(n-1)^2}\left[(k-1)q(1-q) + (n-k)p(1-p)\right]. \tag{10}$$

Analogously, using equations (5) and (6),

$$Skew[y'_n|y_n = \bar{y}_n^k] = \frac{(k-1)q(1-q)(1-2q) - (n-k)p(1-p)(1-2p)}{\left[(k-1)q(1-q) + (n-k)p(1-p)\right]^{\frac{3}{2}}} \tag{11}$$

and

$$Kurt[y'_n|y_n = \bar{y}_n^k] = \frac{(k-1)q(1-q)(1-6q(1-q)) + (n-k)p(1-p)(1-6p(1-p))}{[(k-1)q(1-q) + (n-k)p(1-p)]^2}. \quad (12)$$

5 Matching moments

Next we show how to choose the parameters of the method to approximate an autoregressive process and how these affect the targeted moments. For this purpose, consider a random variable y which follows the following AR(1) process:[5] $y_t = \rho(1-\mu) + \rho y_{t-1} + \varepsilon_t$ in which $0 \leq \rho < 1$, ε_t is white noise with zero mean and $\mu = E[y_t]$. Let $\sigma^2 = Var[y_t]$.

Given p, q and n, one can choose the parameters m and Δ by matching the unconditional mean and variance of y:

$$m + \Delta\frac{q-p}{2-p-q} = \mu$$

and

$$\frac{4\Delta^2}{n-1}\frac{(1-p)(1-q)}{(2-p-q)^2} = \sigma^2.$$

Moreover, one can match the autocorrelation coefficient of the underlying process by setting

$$p + q - 1 = \rho.$$

Note that these three moment conditions do not provide clear guidance on how to choose p and q subject to $p + q - 1 = \rho$. One way to pin down the parameters p (or q) subject to the above three conditions is to target the skewness of y to capture the extent of asymmetry of the AR(1) process.

5.1 Special case

Equation (7) indicates that by setting $p = q$ one can obtain a symmetric distribution, i.e. $Skew[y_n] = 0$. At the same time, $Kurt[y_n] = -\frac{2}{n-1}$. This means that when $p = q$, the excess kurtosis is negative, but approaches zero as n goes to infinity, which is not surprising given the central limit theorem. Therefore, by choosing a sufficiently large n and setting $p = q = (1+\rho)/2$, one can generate a Markov chain that is arbitrarily close to a Gaussian AR(1) process.

Moreover, given the continuous AR(1) process y, $E[y_{t+1}|y_t = \bar{y}_n^k] = \rho(1-\mu) + \rho\bar{y}_n^k$ and $Var[y_{t+1}|y_t = \bar{y}_n^k] = (1-\rho^2)Var[y_t]$. On the other hand, using equations (9) and (10), it can be seen that, when $p = q = (1+\rho)/2$, $E[y'_n|y_n = \bar{y}_n^k] = \rho(1-\mu) + \rho\bar{y}_n^k$ and $Var[y'_n|y_n = \bar{y}_n^k] = (1-\rho^2)Var[y_n]$. These imply that, under symmetry, the conditional mean and conditional

[5] In the interest of clarity, we focus on a continuous AR(1) process. In general, the method can be used to approximate a less restrictive Markov process whose transition function takes the following form: $Q(y'|y) = \text{Prob}[y_t < y'|y_{t-1} = y]$.

variance of the AR(1) process are perfectly matched.[6]

When $p = q$, equations (13) and (14) convert into

$$Skew[y'_n | y_n = \bar{y}^k_n] = \frac{(2k - n - 1)(1 - 2p)}{\sqrt{(n-1)^3 p(1-p)}} \tag{13}$$

and

$$Kurt[y'_n | y_n = \bar{y}^k_n] = \frac{1}{(n-1)p(1-p)} - \frac{6}{n-1}. \tag{14}$$

These results imply that, skewness and kurtosis are highly sensitive to p when the value of that parameter is close to unity. Put differently, the distribution of the innovation of the Markov chain changes as one switches to a higher or lower time frequency.[7] Therefore, to preserve the accuracy of the approximation of conditional skewness and kurtosis, the number of grid points must increase with p. Specifically, equation (14) shows that when p goes to 1, the number of grids must increase to keep $(n - 1)(1 - p)$ sufficiently large.

5.2 General case

Suppose that $p \neq q$. Then, using equations (7) and (8), it can be seen that, given n, one can obtain arbitrarily large values for $Skew[y_n]$ and $Kurt[y_n]$, in absolute terms, by setting p (or q) to a number close to unity. However, it could be impossible to target the two moments simultaneously as they are determined by common parameters. In fact, equations (7) and (8) imply that

$$Kurt[y_n] = -\frac{2}{n-1} + Skew[y_n]^2. \tag{15}$$

Given this equation, we make the following two observations:

a) When $Skew[y_n] = 0$, $Kurt[y_n] < 0$. Therefore, under symmetry, the method can not generate commonly occurring leptokurtic processes.
b) There is a strong trade-off between targeting kurtosis and targeting symmetry. Put differently, the absolute value of skewness and kurtosis move together under the Rouwenhorst method.

On the other hand, common values of the absolute value of skewness are below 1 whereas those of the excess kurtosis are an order of magnitude greater (Tsay, 2005). Thus, the above

[6]The converse is also true: when $p \neq q$, the two conditional moments are not matched. In fact, equations (9) and (10) show that the two moments are strictly increasing with the order of the grid point, k. However, it can be seen that the impact of this gradient on the two conditional moments can be reduced sufficiently by increasing both n and Δ while preserving $4\Delta^2/(n-1)$ so that these two moments do not exhibit substantial variation over the range of several (unconditional) standard deviations around the unconditional mean.

[7]For example, when ρ increases, the distribution of the innovations must become more asymmetric to capture the mean-preserving property of y_n. Moreover, to capture persistence, the probability that the current state repeats itself must increase, which will raise kurtosis of the innovation.

restrictions summarized by equation (15) are highly severe.

5.3 Targeting skewness and kurtosis

One way to break the above tight link between the skewness and excess kurtosis is to consider a sum of two or more Markov chains. To illustrate the idea, consider J independent Markov chains $y^1_{n_1}, y^2_{n_2}, \ldots, y^J_{n_J}$. Suppose that the parameters p and q of $y^j_{n_j}$ are denoted by p_j and q_j respectively. Suppose that these J processes share the same persistence ρ, i.e., $p_j + q_j - 1 = \rho$ for all j. Consider the following sum: $v = \sum_{j=1}^J y^j_{n_j}$. Clearly, the persistence of v is ρ. Now, using equations (5) and (6), it can be seen that

$$Skew[v] = \sum_{j=1}^{J} \omega^S_j Skew[y^j_{n_j}] \tag{16}$$

and

$$Kurt[v] = \sum_{j=1}^{J} \omega^K_j Kurt[y^j_{n_j}], \tag{17}$$

where $\omega^S_j = Var[y^j_{n_j}]^{3/2} / \left(\sum_{i=1}^J Var[y^j_{n_j}]\right)^{3/2}$ and $\omega^K_j = Var[y^j_{n_j}]^2 / \left(\sum_{i=1}^J Var[y^j_{n_j}]\right)^2$. These suggest that by appropriately choosing $\{p_j, Var[y^j_{n_j}]\}$, $j = \{1, 2, \ldots, J\}$, one can generate high kurtosis and low skewness simultaneously.

Example 1

Let $J = 2$. Consider the following restrictions: $p_1 = q_2$, $q_1 = p_2$, $n_1 = n_2$ and $Var[y^1_{n_1}] = Var[y^2_{n_2}]$. Then, it follows that $Skew[y^1_{n_1}] = -Skew[y^2_{n_2}]$. Therefore, there will be zero skewness, i.e. $Skew[v] = 0$. However,

$$Kurt[v] = \frac{1}{n_1 - 1}\left(-1 + \frac{(p_1 - q_1)^2}{2(1 - q_1)(1 - p_1)}\right).$$

Then, given that p_1 is sufficiently close to 1 (and $q_1 = 1 + \rho - p_1$), the autoregressive process v exhibits symmetric, leptokurtic distribution with persistence ρ.

Example 2

Suppose now that one needs to generate a leptokurtic autoregressive process with negative skewness. This can be achieved by reducing $Var[y^1_{n_1}]$ through the parameter Δ_1 (while adjusting m_1 to preserve the mean) and choosing p_1 to satisfy equations (16) and (17).

6 Conclusion

This note analyzes the Markov chain generated by the method of Rouwenhorst (1995). Using the underlying structure of the process, it calculates key moments of the process, including higher order moments such as skewness and kurtosis. It is shown that the absolute magnitude of skewness and kurtosis of the Markov chain move together and thus the method cannot approximate the commonly occurred processes with high kurtosis and low skewness. It is also shown that using a sum of Markov chains constructed by the method, one can target skewness and kurtosis simultaneously without affecting lower-order moments of the sum. The moments calculated in this note can provide useful guidelines to construct a richer set of autoregressive processes by mixing Markov chains constructed by the Rouwenhorst method.

References

Adda, Jerome and Russel Cooper, *Dynamic Economics*, MIT Press, Cambridge, MA, 2003.

Floden, Martin, "A Note on the Accuracy of Markov-Chain Approximations to Highly Persistent AR(1) Processes," *Economics Letters*, 2008, *99* (3), 516–520.

Galindev, Ragchaasuren and Damba Lkhagvasuren, "Discretization of Highly Persistent Correlated AR(1) Shocks," *Journal of Economic Dynamics and Control*, 2010, *34* (7), 1260–1276.

Gospodinov, Nikolay and Damba Lkhagvasuren, "A Moment-Matching Method for Approximating Vector Autoregressive Processes by Finite-State Markov Chains," 2012. Working Paper 11-005, Concordia University.

Kopecky, Karen A. and Richard M.H. Suen, "Finite State Markov-Chain Approximations to Highly Persistent Processes," *Review of Economic Dynamics*, 2010, *13* (3), 701–714.

Rouwenhorst, Geert K., "Asset Pricing Implications of Equilibrium Business Cycle Models," in Thomas Cooley, ed., *Structural Models of Wage and Employment Dynamics*, Princeton: Princeton University Press, 1995.

Tauchen, George, "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions," *Economics Letters*, 1986, *20* (2), 177–181.

_ **and Robert Hussey**, "Quadrature-Based Methods for Obtaining Approximate Solutions to Linear Asset Pricing Models," *Econometrica*, 1991, *59* (2), 371–396.

Terry, Stephen J. and Edward S. Knotek II, "Markov-Chain Approximations of Vector Autoregressions: Application of General Multivariate-Normal Integration Techniques," *Economics Letters*, 2011, *110* (1), 4–6.

Tsay, Ruey S., *Analysis of Financial Time Series, Second Edition*, New Jersey: John Wiley and Sons, Inc, 2005.