THE HYPERGEOMETRIC PROCESS: A FLEXIBLE PARAMETERIZATION OF $MA(\infty)$ TIME SERIES

GRAHAM L. GILLER

ABSTRACT. We introduce an apparently new parameterisation of the Wold decomposition of a standard Box-Jenkins time series model in terms of the Hypergeometric Function ${}_pF_q(\cdot)$. This parameterization is sufficiently flexible to treat both finite standard $ARMA(p,q)$ models and also fractionally integrated models such as $FI(d)$ or the *Gegenbauer* process. We demonstrate the parametric domain over which the process is variance stationary and also over which it is ergodic. We compute the autocovariance function for the general process and illustrate how the parameterization proposed here may be used to model other well known parameterizations of the $MA(q)$ process.

1. INTRODUCTION

1.1. **Motivation.** The is a considerable body of evidence that common financial time series exhibit a "long memory" property. By this it is meant that the autocorrelation function has a rate of decay that is slower than that characterized by an exponential function, as would be expected for the standard $ARMA(p,q)$ formalism where the process may be written

$$ARMA(p,q) : \phi(L)y_t = \theta(L)\varepsilon_t; \tag{1}$$

where $\phi(L)$ and $\theta(L)$ are polynomial functions of the lag operator L of order p and q, respectively; y_t is the modeled process; and, ε_t is an i.i.d. innovation with variance σ^2. To obtain an autocorrelation function that decays at a rate slower than the exponential it is necessary to introduce a parameterization in terms of infinite polynomial functions of the lag operator. Clearly, a general infinite polynomial is not statistically tractable, so various authors have introduced restricted parameterizations that generate infinite moving average processes.

Perhaps the simplest functional form is the "fractionally integrated" process $FI(d)$ defined by

$$FI(d) : (1-L)^d y_t = \varepsilon_t. \tag{2}$$

For $0 < d < 1$, this process exhibits a hyperbolically decaying, strictly positive, autocorrelation function, and finds application in series such as —PROVIDE EXAMPLES HERE. (For $d \geq 1$ the process is not variance stationary.) A more complex functional form is used for the *Gegenbauer* process. Here the parameterization is

$$G(\eta, d) : (1 - 2\eta L + L^2)^d y_t = \varepsilon_t. \tag{3}$$

This parameterization has the property that the function on the L.H.S. is the reciprocal of the generating function for the Gegenbauer polynomials and Equation 3 has the Wold

Date: August 16, 2006. *Giller Investments Research Note*: 20060815/1.

decomposition

$$y_t = \sum_{n=0}^{\infty} C_n^{(d)}(\eta)\varepsilon_{t-n}, \tag{4}$$

where $C_n^{(d)}(\cdot)$ represents the Gegenbauer polynomial of order n. For $0 < d < \frac{1}{2}$, this process exhibits an autocorrelation function with hyperbolically damped oscillations with a characteristic frequency dependent on η. (For $d \geq \frac{1}{2}$ the process is not variance stationary.)

Despite the success of these simple functional forms, there are examples of time series that exhibit sample autocorrelation functions with similar properties to those given by the above models yet for which the hypothesis that the population process is given by Equation 2 or Equation 3 can be confidently rejected. Therefore, motivated by the remarkably simple parameterizations above, we propose a "hypergeometric" process with the Wold decomposition written in terms of hypergeometric functions.

1.2. **The Generalized Hypergeometric Series.** The generalized hypergeometric series, ${}_pF_q[a_1 \ldots a_p; b_1 \ldots b_q; z]$, is defined by

$${}_pF_q\begin{bmatrix} a_1 \cdots a_p \\ b_1 \cdots b_q \end{bmatrix}; z = \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n} \frac{z^n}{n!}. \tag{5}$$

The term $(a)_n$ is called the *Pochhammer Symbol* and represents the "rising factorial" $a(a+1)(a+2)\cdots(a+n-1)$. When the series is convergent, the series is referred to as the generalized hypergeometric function. For $p = 2$ and $q = 1$ the function is called Gauss' hypergeometric function, commonly written ${}_2F_1(a, b; c; z)$, and for $p = q = 1$ it is called the confluent hypergeometric function or the Kummer function ${}_1F_1(a; b; z)$. The hypergeometric function is a remarkably flexible power series and many common expressions or special functions may be represented in terms of it.

1.3. **Domain of Convergence.** If one of the a-parameters is a non-positive integer then the series terminates and the hypergeometric function is a simple polynomial. Otherwise, the radius of convergence depends on the relative number of "upper" and "lower" parameters, as illustrated in Table 1.

Relative Parameter Count	Radius of Convergence
$p < q + 1$	∞
$p = q + 1$	1 (parameter dependent)
$p > q + 1$	0

TABLE 1. Radius of Convergence for the Generalized Hypergeometric Series

When $p = q + 1$ the convergence is dependent on the value of the "parametric excess," defined as

$$s = \sum_{n=0}^{q} b_n - \sum_{n=0}^{p} a_n. \tag{6}$$

For a real argument, x, the parametric convergence criteria are summarized by Table 2 on the facing page.

$p = q + 1$	$x < -1$	$x = -1$	$-1 < x < 1$	$x = 1$	$x > 1$
$s \leq 1$	d	d	C	d	d
$-1 < s \leq 0$	d	C	C	d	d
$s > 0$	d	C	C	C	d

TABLE 2. Domain of Convergence for the parametrically balanced generalized hypergeometric series ("C" indicates convergence; "d" indicates divergence)

2. THE HYPERGEOMETRIC PROCESS

2.1. Definition. We define the hypergeometric process by the expression

$$H(p, q; \phi) : y_t = {}_pF_q \left[\begin{matrix} a_1 \cdots a_p \\ b_1 \cdots b_q \end{matrix} ; \phi L \right] \varepsilon_t \tag{7}$$

2.2. Domain of Ergodicity. Writing Equation 7 in terms of the infinite series representation, we have

$$y_t = \sum_{n=0}^{\infty} \psi_n \varepsilon_{t-n} = \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n} \frac{\phi^n}{n!} \varepsilon_{t-n}. \tag{8}$$

The process will be ergodic if the series of lag coefficients in Equation 7 converges absolutely. i.e. If $\sum_{n=0}^{\infty} |\psi_n|$ converges. Clearly if the parameters of the hypergeometric function are positive, then

$$\sum_{n=0}^{\infty} |\psi_n| = {}_pF_q \left[\begin{matrix} a_1 \cdots a_p \\ b_1 \cdots b_q \end{matrix} ; \phi \right], \tag{9}$$

and the process is ergodic if $p < q + 1$; or if $p = q + 1$ then if $|\phi| < 1$ or if $\phi = 1$ and the parametric excess is positive or if $\phi = -1$ and the parametric excess is greater than -1.

If some of the parameters are negative but not integral, then the sequence of terms may initially alternate in sign but there exists a positive integer l such that $\operatorname{sgn} \psi_m = \operatorname{sgn} \psi_{m+1} \, \forall \, m > l$. Since the convergence of an infinite series is not affected by truncating the series by removing any finite number of initial terms, we see that the convergence is not due to alternating signs in the sequence of summands and therefore the series also converges absolutely for negative parameters under the same conditions as for positive parameters.

2.3. Domain of Stationarity in Variance. As the innovations are i.i.d., the unconditional variance of the process is given by

$$\gamma_0 = \sigma^2 \sum_{n=0}^{\infty} \left\{ \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n} \frac{\phi^n}{n!} \right\}^2 = \sigma^2 \sum_{n=0}^{\infty} \frac{(a_1)_n^2 \cdots (a_p)_n^2}{(b_1)_n^2 \cdots (b_q)_n^2 (1)_n} \frac{\phi^{2n}}{n!}. \tag{10}$$

Where σ^2 is the variance of the innovations and we have used the identity $(1)_n = n!$ to incorporate one of the factorial terms as a lower parameter. Therefore

$$\gamma_0 = \sigma^2 {}_{2p}F_{2q+1} \left[\begin{matrix} a_1 & a_1 & \cdots & a_p & a_p & \\ b_1 & b_1 & \cdots & b_q & b_q & 1 \end{matrix} ; \phi^2 \right]. \tag{11}$$

The hypergeometric series of Equation 11 is convergent as before or if $p = q+1$ and $\phi = 1$ and the parametric excess $1 + 2s$ is positive (with s as defined by Equation 6). i.e. The process is variance stationary whenever the process is ergodic or when the process has a parametric excess $s > -\frac{1}{2}$ for $\phi = 1$.

3. THE AUTOCOVARIANCE FUNCTION

The first lag autocovariance of the process is

$$\gamma_1 = \sigma^2 \sum_{n=0}^{\infty} \psi_n \psi_{n+1} = \sigma^2 \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n} \frac{\phi^n}{n!} \frac{(a_1)_{n+1} \cdots (a_p)_{n+1}}{(b_1)_{n+1} \cdots (b_q)_{n+1}} \frac{\phi^{n+1}}{(n+1)!}. \tag{12}$$

The Pochhammer symbol $(x)_n$ satisfies the shift identity $x(x+1)_n = (x)_{n+1}$, which allows us to write Equation 12 as

$$\gamma_1 = \sigma^2 \phi \frac{a_1 \cdots a_p}{b_1 \cdots b_q} \sum_{n=0}^{\infty} \frac{(a_1)_n (a_1+1)_n \cdots (a_p)_n (a_p+1)_n}{(b_1)_n (b_1+1)_n \cdots (b_q)_n (b_q+1)_n (2)_n} \frac{\phi^{2n}}{n!}. \tag{13}$$

Equation 13 may be recognized as

$$\gamma_1 = \sigma^2 \phi \frac{a_1 \cdots a_p}{b_1 \cdots b_q} {}_{2p}F_{2q+1} \begin{bmatrix} a_1 & (a_1+1) & \cdots & a_p & (a_p+1) & \\ b_1 & (b_1+1) & \cdots & b_q & (b_q+1) & 2 \end{bmatrix} ; \phi^2 \Bigg], \tag{14}$$

which is convergent for under the same conditions given earlier. Similarly, the k-lag autocovariance, γ_k, of the process is

$$\gamma_k = \sigma^2 \sum_{n=0}^{\infty} \psi_n \psi_{n+k} = \sigma^2 \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n} \frac{\phi^n}{n!} \frac{(a_1)_{n+k} \cdots (a_p)_{n+k}}{(b_1)_{n+k} \cdots (b_q)_{n+k}} \frac{\phi^{n+k}}{(n+k)!}. \tag{15}$$

To reduce this expression we use the more general identities for the Pochhammer symbol: $(x)_k(x+k)_n = (x)_{n+k}$ and $(n+k)! = k!(k+1)_n$. Equation 15 may then be written as (16)

$$\gamma_k = \sigma^2 \frac{\phi^k}{k!} \frac{(a_1)_k \cdots (a_p)_k}{(b_1)_k \cdots (b_q)_k} {}_{2p}F_{2q+1} \left[\begin{matrix} a_1 & (a_1+k) & \cdots & a_p & (a_p+k) & \\ b_1 & (b_1+k) & \cdots & b_q & (b_q+k) & (k+1) \end{matrix} ; \phi^2 \right].$$

This hypergeometric series has the same domain of convergence as Equation 14. The autocorrelation function is then trivially given by $\rho_k = \gamma_k/\gamma_0$.

4. REPRESENTATION OF WELL KNOWN PROCESSES

The flexibility of the hypergeometric series as a power series representation of many well known common and special functions allows us to represent well known time series processes in terms of specific parameterizations of the hypergeometric process. Here we illustrate how to construct some of these processes.

4.1. The $MA(1)$ Process. The standard $MA(1)$ process is written $y_t = \varepsilon_t + \theta\varepsilon_{t-1}$. We may truncate the hypergeometric series at order q by setting an upper parameter to $-q$ since $(-q)_{q+1} = 0$. Therefore we choose $a_1 = -1$, $b_1 = -\frac{1}{\theta}$ and $\phi = 1$ to obtain a confluent hypergeometric process

$$y_t = {}_1F_1(-1; -\frac{1}{\theta}; L)\varepsilon_t = (1 + \theta L)\varepsilon_t. \tag{17}$$

As the a_1 parameter is a negative integer, the series terminates and convergence criteria are not relevant. We may use Equation 16 to obtain:

$$\gamma_0 = \sigma^2 {}_2F_3 \left[\begin{matrix} -1 & -1 & \\ -\frac{1}{\theta} & -\frac{1}{\theta} & 1 \end{matrix} ; 1 \right] = \sigma^2(1+\theta^2) \tag{18}$$

$$\text{and}\, \gamma_1 = \sigma^2 \frac{(-1)_1}{(-\frac{1}{\theta})_1} {}_2F_3 \left[\begin{matrix} -1 & 0 & \\ -\frac{1}{\theta} & 1-\frac{1}{\theta} & 2 \end{matrix} ; 1\right] = \sigma^2\theta. \tag{19}$$

4.2. **The $AR(1)$ Process.** The standard $AR(1)$ process is written $y_t = \varepsilon_t + \phi y_{t-1}$, which is often represented as the $MA(\infty)$ process $y_t = (1-\phi L)^{-1}\varepsilon_t = 1 + \phi L + \phi^2 L^2 + \ldots$. This series can be represented by any hypergeometric series with $p = q+1$ and all upper and lower coefficients equal except for one upper coefficient which must be unity. e.g.

$$y_t = {}_1F_0(1; \phi L)\varepsilon_t. \tag{20}$$

$$\Rightarrow \gamma_k = \sigma^2 \phi^k {}_2F_1(1, k+1; k+1; \phi^2) = \sigma^2 \frac{\phi^k}{1-\phi^2}. \tag{21}$$

4.3. **The $FI(d)$ Process.** The fractionally integrated process $FI(d)$ is developed in a manner very similar to that for the $AR(1)$, on the current page. In this case we write $y_t = (1-L)^{-d}\varepsilon_t$, which also generates an $MA(\infty)$ process. We may represent the operator by the hypergeometric series of Equation 20 with the unity coefficient replaced by d. Thus

$$y_t = {}_1F_0(d; L)\varepsilon_t. \tag{22}$$

$$\Rightarrow \gamma_k = \sigma^2 \frac{(d)_k}{k!} {}_2F_1(d, d+k; 1+k; 1). \tag{23}$$

As this representation of the process[1] is parametrically balanced ($p = q+1$) with $\phi = 1$, we immediately see from the discussion of Section 2.3 that the parametric excess of Equation 22 $s = -d > -\frac{1}{2} \Rightarrow d < \frac{1}{2}$.

REFERENCES

[1]The relationship between the $FI(d)$ process and the hypergeometric series is well known in the literature; however, we believe that this presentation is of merit as it is analytically very compact.