RESEARCH REPORT SERIES
(Statistics #2012-06)

Uncorrelatedness and Other Correlation Options for Differenced Seasonal Decomposition Components of ARIMA Model Decompositions

David F. Findley

Center for Statistical Research & Methodology
Research and Methodology Directorate
U.S. Census Bureau
Washington, D.C. 20233

Report Issued: May 29, 2012

Uncorrelatedness and Other Correlation Options for Differenced Seasonal Decomposition Components of ARIMA Model Decompositions

David F. Findley
Center for Statistical Research and Methodology
U.S. Census Bureau

Abstract

Two gaps in the optimality theory supporting current ARIMA model-based seasonal adjustment are addressed. The main one concerns the requirement that decomposition components be uncorrelated with one another after they are minimally differenced to stationarity. Uncorrelatedness has been assumed but not formally verified. We verify it by introducing a model compatibility criterion fitting current practice that specifies how the ARIMA models of the seasonal decomposition components are to be compatible with the ARIMA model of the observed seasonal series. This criterion always supports the assumption of uncorrelated components for a stationary decomposition involving the differenced observed series and the similarly differenced and therefore overdifferenced stationary component series. We verify the requirement by proving that overdifferencing can be corrected for and that doing this preserves uncorrelatedness. Then we investigate whether correlated components are also allowed by the compatibility criterion and give a complete description of the allowed correlation structures for two-component decompositions. We also discuss their impracticality.

Contents

1 Introduction

This note addresses two gaps in the theoretical foundation of ARIMA model-based (AMB) seasonal adjustment and more general ARIMA signal extraction, as presented in Burman (1980), Hillmer and Tiao (1982), Bell (1984), Gómez and Maravall (2001), and McElroy (2008). The gaps have to do with correlation properties of the components of seasonal (or other) time series decompositions, after any ARIMA components have been differenced to stationarity in the way specified by their models. For simplicity, until we get to the main result, which is the Corollary of Section 4, we focus on two-component unobserved "signal plus noise" decompositions of an observable nonstationary ARIMA series Z_t,

$$Z_t = S_t + N_t. \tag{1}$$

The minimal degree differencing polynomials $\delta_S(B)$ and $\delta_N(B)$ that render S_t and N_t stationary are required to factorize the differencing polynomial $\delta(B)$ of Z_t's model,

$$\delta(B) = \delta_S(B)\,\delta_N(B)\,. \tag{2}$$

and to not have common zeroes: if $\delta_S(z_0) = 0$, then $\delta_N(z_0) \neq 0$ and if $\delta_N(z_0) = 0$, then $\delta_S(z_0) \neq 0$. If a component on the right in (1) is stationary, its differencing polynomial is the constant 1, e.g. $\delta_N(B) = 1$ if N_t is stationary, in which case $\delta_S(B) = \delta(B)$.

The series $x_t = \delta(B) S_t$ and $\tilde{x}_t = \delta(B) N_t$ are stationary and satisfy

$$w_t = x_t + \tilde{x}_t, t = 0, \pm 1, \ldots, \tag{3}$$

but at least one is overdifferenced because at least one factor on the right in (2) has lower degree than $\delta(B)$. In ARIMA model based signal extraction, ARIMA model pairs (one member of the pair could be ARMA) are determined or specified for S_t and N_t. We use the decomposition (3) to define a convenient *compatibility condition* to insure that any proposed ARIMA component model pair is compatible with the ARIMA model of the observed series Z_t: a model pair's implied autocovariances $\gamma_{x,j} = E x_t x_{t+j}$ and $\gamma_{\tilde{x},j} = E \tilde{x}_t \tilde{x}_{t+j}$ for x_t and $\tilde{x}_t$ are required to sum to the autocovariances $\gamma_j = E w_t w_{t+j}$ of w_t,

$$\gamma_j = \gamma_{x,j} + \gamma_{\tilde{x},j}, j = 0, \pm 1, \ldots \,. \tag{4}$$

The most important gap in the theory of AMB seasonal adjustment and signal extraction has to do with establishing that the minimally differenced stationary series $u_t = \delta_S(B) S_t$ and $\tilde{u}_t = \delta_N(B) N_t$ can be taken as uncorrelated with one another, $E u_t \tilde{u}_{t+j} = 0$, $t, j = 0, \pm 1, \ldots$, a property we denote by

$$\{u_t\} \perp \{\tilde{u}_t\}. \tag{5}$$

Bell (1984) showed that this is required to establish the linear mean square optimality property that provides an attractive theoretical justification for the standard procedures used for estimating S_t and N_t, assuming that the ARIMA model for $w_t = \delta(B) Z_t$ and its parameters are correct; see McElroy (2008) for an elementary finite-sample development[1]. The model-based seasonal adjustment estimation procedures subject to this requirement for optimality include those of TRAMO-SEATS (Gómez and Maravall, 1996), TSW (Caporello and Maravall, 2004), X-13ARIMA-SEATS (U.S. Census Bureau, 2012) and STAMP (Koopman, Harvey, Doornik, and Shephard, 2000).

What is immediately justifiable is the assumption

$$\{x_t\} \perp \{\tilde{x}_t\}, \tag{6}$$

because of its obvious compatibility with (4), see Section 3. Using (2), we can re-express (3) as

$$w_t = \delta_N(B) u_t + \delta_S(B) \tilde{u}_t, \tag{7}$$

and obtain a more revealing formulation of (6) as

$$\{\delta_N(B) u_t\} \perp \{\delta_S(B) \tilde{u}_t\}. \tag{8}$$

Our goal is to get from (8) to (5). To do this, we show how a differencing operation applied to an already stationary series can be inverted, and we verify

[1] Bell (1984) also showed that the "naive" historical assumption that S_t and N_t should be uncorrelateded requires more complex estimation formulas than those now used in order to achieve linear mean square optimality.

that this operation preserves uncorrelatedness. In the Corollary of Proposition 1 of Section 4, we establish for $K \geq 2$ component generalizations of (8) and (5), that if one holds, so does the other. Thus the mutual uncorrelatedness of minimally differenced component series can always be assumed for the correct model context.

The presentation of this general result, which covers all differencing polynomials used in AMB seasonal adjustment, is preceded by the treatment of a simple motivating trend plus irregular decomposition example in Section 3, after a brief review of spectral densities in Section 2. With $g_w = g_w(\lambda)$ denoting the spectral density of w_t, $g_w(\lambda) = \sum_{j=-\infty}^{\infty} \gamma_j e^{-i2\pi j\lambda}$, $-1/2 \leq \lambda \leq 1/2$ and $g_x(\lambda)$ and $g_{\tilde{x}}(\lambda)$ defined analogously for x_t, and $\tilde{x}_t$, it is revealing to re-express the compatibility requirement (4) as the spectral density decomposition property,

$$g_w = g_x + g_{\tilde{x}}. \tag{9}$$

As will be illustrated, such spectral density decompositions and their $K \geq 2$ component generalizations are algebraically equivalent to admissible pseudo-spectral density decompositions in the sense of Hillmer and Tiao (1982), and include those used by AMB seasonal adjustment programs. Hence (5) can be assumed for all AMB seasonal adjustments in the correct-model context.

The second gap we address concerns the question of whether an assumption different from (6), namely correlation between the series x_t and $\tilde{x}_t$ of (3), is also compatible with (9). The answer is yes: in Proposition 3 of Section 5, for fixed, everywhere continuous $g_x$and $g_{\tilde{x}}$, we characterize the infinitude of possible stationary correlation relations between x_t and $\tilde{x}_t$ that are compatible with (9), ranging from no correlation to complete correlation. The latter term means that each of the series x_t and $\tilde{x}_t$ is a linear filtered version of the other.

However, it will be seen that there is no information in g_w to favor a choice among the possible correlated decompositions. (We further show that the best-known correlated decomposition, that of Beveridge and Nelson (1981), which has no known optimality property, is not compatible with (9).) The uncorrelated decomposition is to be preferred as the most neutral one, in a specific sense to be described as well as in a general sense.

Our exposition assumes that the reader has some basic familiarity with ARIMA models and their decomposition for model-based seasonal adjustment. In this article, stationary always means covariance stationary and stationary series are taken to have mean zero.

2 Spectral Densities and ARMA Models

We start by reviewing basic material on spectral densities of stationary series and of ARMA models and their properties we will need. The reader familiar with the definition of a spectral density and the formula for the spectral density of an ARMA series can skip to Section 3 concerned with spectral density decompositions after noting the formula (15) which connects the spectral densities of the input and output of a linear (time-invariant) filter.

2.1 Review

We use a standard notation and terminology regarding complex numbers $z = a + ib$, with a and b real and with $i^2 = -1$. The number a is the real part of z, $a = \mathrm{Re}\,(z)$, and b the is the imaginary part, $b = \mathrm{Im}\,(z)$. The number $\bar{z} = a - ib$ is the complex conjugate of z. Its basic properties are $z + \bar{z} = 2\,\mathrm{Re}\,(z)$, $z - \bar{z} = 2i\,\mathrm{Im}\,(z)$ and $\sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$, which is the magnitude of z, denoted $|z|$ (the distance from (a, b) to $(0, 0)$ in the coordinate plane). Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$ shows that $e^{-i\theta}$ is the complex conjugate of $e^{i\theta}$. See Wikipedia Contributors (2012) for more information.

For a stationary time series w_t, the fact that $Ew_t w_{t+j} = Ew_{t-j}w_t$ shows that its autocovariances $\gamma_j = Ew_t w_{t+j}$ have the property $\gamma_{-j} = \gamma_j$ for all $j = 0, \pm 1, \ldots$. Because of this property the spectral density g_w of w_t can be defined by any of the three formulas

$$\begin{aligned} g_w(\lambda) &= \sum_{j=-\infty}^{\infty} \gamma_j e^{-i2\pi j\lambda} = \gamma_0 + \sum_{j=1}^{\infty} \gamma_j \left(e^{-i2\pi j\lambda} + e^{i2\pi j\lambda}\right) \\ &= \gamma_0 + 2\sum_{j=1}^{\infty} \gamma_j \cos 2\pi j\lambda,\ -1/2 \le \lambda \le 1/2. \end{aligned} \tag{10}$$

The simplest example, which we will build upon, is that of *white noise*, $w_t = a_t$, with a_t an uncorrelated mean zero series with variance σ_a^2. Its spectral density is its variance, a constant,

$$g_a(\lambda) = \sigma_a^2,\ \ -1/2 \le \lambda \le 1/2. \tag{11}$$

The second and third formulas of (10) show that a spectral density is always an even function, $g_w(\lambda) = g_w(-\lambda)$, and spectral densities can be proven to be non-negative for all λ, as the formula for the ARMA case (16) below illustrates. In the ARMA case, the γ_j converge to zero exponentially rapidly[2], making it easy to justify term by term integration in (10) in order to use the property

$$\int_{-1/2}^{1/2} e^{i2\pi(k-j)\lambda} d\lambda = \begin{cases} 1, & j = k \\ 0 & j \neq k \end{cases},$$

which can be verified using $e^{i2\pi(k-j)\lambda} = \cos 2\pi(k-j)\lambda + i\sin 2\pi(k-j)\lambda$, to obtain

$$\gamma_k = \int_{-1/2}^{1/2} e^{i2\pi k\lambda} g_w(\lambda)\, d\lambda,\ k = 0, \pm 1, \ldots\ , \tag{12}$$

see Theorem 4.3.2 of Brockwell and Davis (1991).

Remark. The formula (12) is the more versatile definition of the spectral density g_w of w_t because it avoids the issue of possible non-convergence of the

[2]Autocovariances of an ARMA processes converge exponentially rapidly to 0 at a rate connected to the smallest magnitude of a zero of the AR polynomial $\varphi(B)$. One has $|\gamma_k| \le Cr^{-k}$ for some constant $C = C(r)$ for any $1 < r < \min\{|z|;\ \varphi(z) = 0\}$, see (3.3.9) and §3.6 of Brockwell and Davis (1991).

series in (10). In the Appendix, Section 7.3, we introduce a weaker form of infinite series convergence which will be used to establish (5). When applied to the infinite series in (10), it yields convergence to $g_w(\lambda)$ for all λ, and shows the equivalence of both definitions of the spectral density, whenever g_w is continuous on $-1/2 \leq \lambda \leq 1/2$.

2.2 ARMA Spectral Densities

Recall that, for a zero-mean stationary ARMA(p,q) series w_t, with B denoting the backshift (or lag) operator, we have

$$\varphi(B) w_t = \vartheta(B) a_t, \tag{13}$$

with white noise a_t, with $\varphi(z) = 1 - \varphi_1 z - \cdots - \varphi_p z^p$ satisfying

$$\varphi(z) \neq 0 \text{ when } |z| \leq 1 \tag{14}$$

and having no zeroes in common with $\vartheta(z) = 1 - \vartheta_1 z - \cdots - \vartheta_q z^q$. The property (14) can be expressed as $\min\{|z| : \varphi(z) = 0\} > 1$. When w_t is a zero-mean seasonal ARMA series, $\phi(B)\Phi(B^s) w_t = \theta(B)\Theta(B^s) a_t$ for $s \geq 2$, then we set $\varphi(z) = \phi(z)\Phi(z^s)$ and $\vartheta(z) = \theta(z)\Theta(z^s)$. In the Gaussian case, there is no loss of generality in assuming that $\vartheta(z) \neq 0$ when $|z| < 1$. The ARMA model is said to be noninvertible if $\vartheta(z) = 0$ for some z with $|z| = 1$.

To obtain the spectral density of such a w_t, one uses the fundamental fact that when w_t is the output of a linear filter $A(B) = \sum_j \alpha_j B^j$ applied to some stationary x_t, i.e. $w_t = A(B) x_t = \sum_j \alpha_j x_{t-j}$ (where $B^j x_t = x_{t-j}$ with $j < 0$ is a forward shift $-j$ units in time), then the spectral densities of the input series x_t and the output w_t are related by

$$g_w(\lambda) = \left|A\left(e^{-i2\pi\lambda}\right)\right|^2 g_x(\lambda), \tag{15}$$

see §4.4 of Brockwell and Davis (1991). The function $A\left(e^{-i2\pi\lambda}\right) = \sum_j \alpha_j e^{-2i\pi j\lambda}$ is called the *transfer function* of the filter $A(B)$. When the coefficients α_j are real, its complex conjugate $\overline{A(e^{-i2\pi\lambda})}$ has the formula $\overline{A(e^{-i2\pi\lambda})} = A\left(e^{i2\pi\lambda}\right)$ and $\left|A\left(e^{-i2\pi\lambda}\right)\right|^2 = A\left(e^{-i2\pi\lambda}\right) A\left(e^{i2\pi\lambda}\right)$.

Using (15), (11) and (13) yield $\left|\varphi\left(e^{i2\pi\lambda}\right)\right|^2 g_w(\lambda) = \left|\vartheta\left(e^{i2\pi\lambda}\right)\right|^2 \sigma_a^2$ and therefore also

$$g_w(\lambda) = \sigma_a^2 \frac{\left|\vartheta\left(e^{i2\pi\lambda}\right)\right|^2}{\left|\varphi\left(e^{i2\pi\lambda}\right)\right|^2}, \; -1/2 \leq \lambda \leq 1/2, \tag{16}$$

where σ_a^2 is the variance of the white noise a_t. This formula shows that $g_w(\lambda)$ is non-negative and continuous for all λ due to (14).

2.2.1 The MA(1) Spectral Density and Autocovariances

For example, from (16), the spectral density of the MA(1) model, $w_t = (1 - \theta B)a_t,\ -1 \leq \theta \leq 1$ is

$$
\begin{aligned}
g_w(\lambda) &= \sigma_a^2 \left|1 - \theta e^{-i2\pi s\lambda}\right|^2 = \sigma_a^2 \left(1 - \theta e^{i2\pi s\lambda}\right)\left(1 - \theta e^{-i2\pi s\lambda}\right) \\
&= \sigma_a^2 \left(1 + \theta^2\right) - \sigma_a^2 \theta \left(e^{i2\pi\lambda} + e^{-i2\pi\lambda}\right),
\end{aligned} \tag{17}
$$

which provides an alternate derivation of the familiar fact that an MA(1) has $\gamma_0 = \sigma_a^2 \left(1 + \theta^2\right)$, $\gamma_{\pm 1} = -\sigma_a^2 \theta$ and $\gamma_j = 0$ for $|j| \geq 2$.

2.2.2 The Gap AR(s) Spectral Density and Its Autocovariances

It follows from (16) that the spectral density of a gap AR(s) model,
$w_t = \varphi w_{t-s} + a_t$, with $s \geq 1$, is

$$
g_w(\lambda) = \sigma_a^2 \left|1 - \varphi e^{i2\pi s\lambda}\right|^{-2}. \tag{18}
$$

We will need to know the coefficients of the bi-infinite series expansion (10). By using $\left(1 - \varphi B^s\right)^{-1} = \sum_{k=0}^{\infty} \varphi^k B^{ks}$, one obtains $w_t = \sum_{k=0}^{\infty} \varphi^k a_{t-ks}$, from which it easily follows that

$$
\gamma_j = \begin{cases} \sigma_a^2 \left(1 - \varphi^2\right)^{-1} \varphi^{|k|}, & j = ks \text{ for some integer } k \\ 0, & \text{otherwise.} \end{cases} \tag{19}
$$

3 Sums of Spectral Densities and Uncorrelated Decompositions

If two stationary mean zero times series x_t and $\tilde{x}_t$ are uncorrelated, $\{x_t\} \perp \{\tilde{x}_t\}$, then

$$
E\left(x_t + \tilde{x}_t\right)\left(x_{t+j} + \tilde{x}_{t+j}\right) = E x_t x_{t+j} + E \tilde{x}_t \tilde{x}_{t+j},\ j = 0, \pm 1, \ldots . \tag{20}
$$

Multiplying both sides of (20) by $e^{-i2\pi j\lambda}$ and summing over $-\infty < j < \infty$ yields that the spectral density of the sum series $w_t = x_t + \tilde{x}_t$ is the sum of the component spectral densities, i.e. (9) holds.

Conversely, if a given spectral density g_w of a stationary series is found to have a decomposition into a sum of spectral densities, say $g_w = g_1 + g_2$, then as regards its autocovariance properties, one can treat w_t as admitting a decomposition $w_t = x_t + \tilde{x}_t$ with uncorrelated components having spectral densities $g_x = g_1$ and $g_{\tilde{x}} = g_2$. (Correlated decompositions also compatible with (9) are considered in Section 5.) Analogous conclusions hold for sums of more than two series.

An important example of a trend plus irregular decomposition of an IMA(1,1) Z_t,

$$
(1 - B) Z_t = (1 - \theta B)a_t,\ -1 \leq \theta < 1. \tag{21}
$$

is associated with the easily verified decomposition of the MA(1) spectral density

$$\sigma_a^2 \left|1 - \theta e^{i2\pi\lambda}\right|^2 = \sigma_b^2 \left|1 + e^{i2\pi\lambda}\right|^2 + \sigma_c^2 \left|1 - e^{i2\pi\lambda}\right|^2, \tag{22}$$

for which

$$\sigma_b^2 = \frac{1}{4}(1-\theta)^2 \sigma_a^2, \quad \sigma_c^2 = \frac{1}{4}(1+\theta)^2 \sigma_a^2. \tag{23}$$

In (22), the spectral density of the MA(1) process $w_t = (1 - \theta B)a_t$ has been decomposed into spectral densities of the components on the right of

$$w_t = (1 - B)\, T_t + (1 - B)\, I_t. \tag{24}$$

This yields the IMA(1,1) model

$$(1 - B)\, T_t = (1 + B)\, b_t$$

for the trend T_t and the overdifferenced MA(1) model $(1 - B)\, I_t = (1 - B)\, c_t$, which is equivalent to the white noise model

$$I_t = c_t$$

for the irregular component I_t, with the variances of the white noise process b_t and c_t given by (23). These are the models calculated (numerically) from (21) by TRAMO-SEATS, TSW and X-13ARIMA-SEATS to estimate the trend and irregular components of an IMA(1,1) Z_t. The spectral density decomposition (22) allows the assumption that the MA(1) processes $x_t = (1 + B)b_t$ and $\tilde{x}_t = (1 - B)\, c_t$ are uncorrelated, i.e. (6) is satisfied. The component $\tilde{x}_t$ of w_t is overdifferenced because I_t is stationary.

The spectral density decomposition (22) is algebraically equivalent to

$$\sigma_a^2 \frac{\left|1 - \theta e^{i2\pi\lambda}\right|^2}{\left|1 - e^{i2\pi\lambda}\right|^2} = \sigma_b^2 \frac{\left|1 + e^{i2\pi\lambda}\right|^2}{\left|1 - e^{i2\pi\lambda}\right|^2} + \sigma_c^2, \tag{25}$$

which is a decomposition of the *pseudo-spectral density* function of the model (21) as we now explain. The non-constant functions in (25) each have a form analogous to that of an ARMA spectral density (16), except their denominators are the squared magnitude of the transfer function $\delta\left(e^{-i2\pi\lambda}\right) = 1 - e^{-i2\pi\lambda}$ of the differencing polynomial $1 - B$ of (21) and therefore have the value 0 at $\lambda = 0$. This causes their integrals over $[-1/2, 1/2]$ to be infinite. This is in contrast to the finite value obtained from (12) with $k = 0$ for a spectral density. When $\left|\delta\left(e^{-i2\pi\lambda 2}\right)\right|^2$ from an ARIMA model is incorporated into the denominator of the ARMA spectral density formula (16), the resulting function is called the *pseudo-spectral density function* (p-sd for short) of the ARIMA model. Thus the function on the left in (25) is the p-sd of (21) and the similar function on the right is the p-sd of the trend's IMA(1,1) model, $(1 - B)\, T_t = (1 + B)\, b_t$. The formula (25) is the canonical p-sd decomposition of (21), meaning that every non-constant p-sd or spectral density belongs to a non-invertible model. The constant term on the right in (25) is the spectral density of the white noise irregular component. Such decompositions are used to identify the component models' ARIMA models in AMB seasonal adjustment, see Hillmer and Tiao (1982) for a variety of examples.

The algebraic equivalence of (22) and (25) illustrates a fundamental fact: the decompositions of Z_t's pseudo-spectral density that are the foundation of AMB seasonal adjustment, see Burman (1980), Hillmer and Tiao (1982) and Gómez and Maravall (2001) are algebraically equivalent to model compatibility criteria (9) or their K component generalizations (36) below.

4 Obtaining Uncorrelatedness after Overdifferencing

In this Section, we present and validate an elementary method for inverting the differencing polynomial $\delta(B) = 1 - B^s$, $s \geq 1$ after it has been applied to a stationary time series y_t whose autocovariances are $\gamma_j = Ey_t y_{t+j}$ are absolutely summable

$$\sum_{j=-\infty}^{\infty} |\gamma_j| = |\gamma_0| + 2\sum_{j=1}^{\infty} |\gamma_j| < \infty. \tag{26}$$

A direct attempt at inversion by applying $(1 - \beta B^s)^{-1} = \sum_{k=0}^{\infty} \beta^k B^{ks}$ with $\beta = 1$ to $(1 - B^s) y_t$, produces a non-convergent infinite series: the partial sums

$$\sum_{k=0}^{N} B^{ks}(1 - B^s) y_t = \sum_{k=0}^{N} \left(y_{t-ks} - y_{t-(k+1)s}\right) = y_t - y_{t-(N+1)s}$$

do not converge to y_t or to anything else. Their errors $\left(y_t - y_{t-(N+1)s}\right) - y_t = y_{t-(N+1)s}$ do not converge to zero. The mean square errors $Ey^2_{t-(N+1)s} = \gamma_0$ are constant and non-zero.

However, inversion can done with the aid of approximating inverses of the form $(1 - \beta B^s)^{-1}$, $0 < \beta < 1$. Proposition 1 shows that these yield the desired result in mean square limit as β increases to 1,

$$\lim_{\beta \uparrow 1} E\left\{(1 - \beta B^s)^{-1}(1 - B^s) y_t - y_t\right\}^2 = 0. \tag{27}$$

We also establish that a series x_t is uncorrelated with the stationary series y_t if and only if it is uncorrelated with the series $\delta(B) y_t$:

$$\{x_t\} \perp \{\delta(B) y_t\} \Longleftrightarrow \{x_t\} \perp \{y_t\}. \tag{28}$$

More generally, (28) will be established for all differencing polynomials that occur in ARIMA model-based seasonal adjustment, because the zeroes of these differencing polynomials all have the rational form $e^{\pm i2\pi(r/s)}$ for some $0 \leq r < s$. Such a zero is a zero of $1 - z^s$. Therefore a differencing polynomial whose zeroes have this form will be a divisor of a polynomial of the form $\Pi_s (1 - z^s)^{d_s}$ for certain $s > 0$, a property we can use to obtain the needed result.

4.1 Inverting Differencing of a Stationary Process: White Noise Case

To motivate the general results of this Section with an important example that requires only the simplest calculations, we return to the spectral density decomposition (22) and (23) of (24), resulting from differencing $Z_t = T_t + I_t$. In Section 3, it was explained that this decomposition permits us to assume that the stationary MA(1) processes $x_t = (1+B)\,b_t$ and $\tilde{x}_t = (1-B)\,I_t$ are uncorrelated $\{x_t\} \perp \{\tilde{x}_t\}$. From this, the series x_t must be shown to be uncorrelated with the undifferenced series I_t,

$$\{x_t\} \perp \{I_t\}\,. \tag{29}$$

We start by verifying (27) for the white noise case $y_t = I_t$ with $s = 1$. For $0 < \beta < 1$, define an approximation $I_t(\beta)$ to I_t by

$$\begin{aligned} I_t(\beta) &= (1-\beta B)^{-1}(1-B)\,I_t = \sum_{j=0}^{\infty} \beta^j (1-B)\,I_{t-j} && (30) \\ &= I_t + \sum_{j=1}^{\infty} \left(\beta^j - \beta^{j-1}\right) I_{t-j} \\ &= I_t + (\beta - 1) \sum_{j=1}^{\infty} \beta^{j-1} I_{t-j}. && (31) \end{aligned}$$

Next observe from (31) and the uncorrelatedness of the different I_{t-j} that this approximation's mean square error has the formula

$$\begin{aligned} E\{I_t(\beta) - I_t\}^2 &= (1-\beta)^2 \sum_{j=1}^{\infty} \beta^{2(j-1)} \sigma_I^2 \\ &= \frac{(1-\beta)^2}{1-\beta^2}\sigma_I^2 = \frac{1-\beta}{1+\beta}\sigma_I^2, && (32) \end{aligned}$$

uniformly in t. Letting β increase to 1, we conclude that

$$\lim_{\beta \uparrow 1} E\{I_t(\beta) - I_t\}^2 = 0. \tag{33}$$

That is, $I_t = \lim_{\beta\uparrow 1} I_t(\beta)$ in mean square, for all t.

The result (33) also yields (29), starting from $E\{x_t (1-B)\,I_{t+k}\} = 0$ for all integers k, which in turn yields

$$\begin{aligned} E\{x_t I_{t+k}\} &= \lim_{\beta\uparrow 1} E\{x_t I_{t+k}(\beta)\} \\ &= \lim_{\beta\uparrow 1} \sum_{j=0}^{\infty} \beta^j E\{x_t (1-B)\,I_{t+k-j}\} = \lim_{\beta\uparrow 1} \sum_{j=0}^{\infty} \beta^j \cdot 0 = 0, \end{aligned}$$

for all k, t. The first equality follows from the Cauchy-Schwarz inequality $|Cov(u,v)| \leq \sqrt{var(u)}\sqrt{var(v)}$ with $u = x_s$ and $v = I(\beta) - I_t$,

$$|E\{x_t(I_{t+k}(\beta) - I_{t+k})\}| \leq \sqrt{E\{x_t^2\}}\sqrt{E\{I_{t+k}(\beta) - I_{t+k}\}^2}, \tag{34}$$

and (33). The second follows from an analogous argument that justifies the interchange of expectation and infinite summation via a generalization of (32),

$$E\left\{\sum_{j=J}^{\infty}\beta^j\{(1-B)I_{t+k-j}\}\right\}^2 = \beta^{2J}\sigma_I^2 + \beta^{2J}\frac{1-\beta}{1+\beta}\sigma_I^2,$$

whose right hand side tends to 0 as $J \to \infty$ for any fixed $0 < \beta < 1$.

The result (33) and its generalization (27) established by (i) of Proposition 1 below can be regarded as mean square variants of convergence in the sense of Abel, which is discussed in the Appendix.

4.2 The Proposition 1 and its Corollary

The main results of this Section are Proposition 1 and its Corollary. The latter provides a generalization to any number of components of the needed uncorrelatedness result (5) discussed in the Introduction.

The differencing operators important for seasonal adjustment are factors of $1 - B^s$ for powers of such factors. For integers $s > 2$, the zeroes of $1 - z^s$ are $z = 1$ and $z = e^{\pm i2\pi(r/s)}$ for integers $0 < r < s/2$, together with $z = -1$ when s is even. Therefore $1 - B^s$ factors as

$$1 - B^s = (1-B)(1+B)^{n(s)}\Pi_{0<r<s/2}\left(1 - 2\cos(2\pi r/s)B + B^2,\right)$$

where $n(s) = 1$ or 0, according as s is even or odd. All rational λ with $-1/2 \leq \lambda \leq 1/2$ are associated with powers of these factors for some s.

Proposition 1 *Let y_t be a stationary series whose autocovariances $\gamma_j = Ey_ty_{t+j}$ are absolutely summable, i.e. (26) holds.*
(i) Then (27) holds for the differencing operators $1 - B^s$ for each $s \geq 1$.

(ii) More generally, let $\delta(B)$ be any differencing polynomial whose zeroes are rational (i.e. of the form $e^{\pm i2\pi(r/s)}$ for integers $0 \leq r \leq s/2$). Then a stationary series x_t is uncorrelated with the series $\delta(B)y_t$ if and only if it is uncorrelated with the series y_t,

$$\{x_t\} \perp \{\delta(B)y_t\} \Longleftrightarrow \{x_t\} \perp \{y_t\}. \tag{35}$$

The proof is given in the Appendix, see Subsection7.1

As a Corollary, whose proof is given in Subsection 7.2, we obtain the result required to establish the mean square optimality of the ARIMA model-based seasonal adjustment and signal extraction methods (under correct model assumptions) as described in the Introduction. It is given for decompositions

$Z_t = \sum_{k=1}^{K} Z_t^{(k)}$ into any number $K \geq 2$ of component series $Z_t^{(k)}$ whose stationarizing differencing polynomials $\delta^{(k)}(B)$ have no common zeroes (with $\delta^{(k)}(B) = 1$ when $Z_t^{(k)}$ is stationary) and are such that $\delta^{(\Pi)}(B) = \prod_{k=1}^{K} \delta^{(k)}(B)$ is the differencing polynomial for Z_t. For such decompositions, with g_k denoting the spectral density of $\delta^{(\Pi)}(B) Z_t^{(k)}$, the generalization of the model compatibility condition (9) for $w_t = \delta^{(\Pi)}(B) Z_t$ is

$$g_w = \sum_{k=1}^{K} g_k. \tag{36}$$

This is compatible with the assumption

$$\left\{\delta^{(\Pi)}(B) Z_t^{(j)}\right\} \perp \left\{\delta^{(\Pi)}(B) Z_t^{(k)}\right\}, \ j \neq k = 1, \ldots, K. \tag{37}$$

The desired property is

$$\left\{\delta^{(j)}(B) Z_t^{(j)}\right\} \perp \left\{\delta^{(k)}(B) Z_t^{(k)}\right\}, \ j \neq k = 1, \ldots, K. \tag{38}$$

Corollary 2 *Suppose that for some $K \geq 2$, the nonstationary ARIMA Z_t has a decomposition $Z_t = \sum_{k=1}^{K} Z_t^{(k)}$ whose component series $Z_t^{(k)}$ have stationarizing differencing polynomials $\delta^{(k)}(B)$ with rational zeroes such that no two polynomials have common zeroes and such that $\delta^{(\Pi)}(B) = \prod_{k=1}^{K} \delta^{(k)}(B)$ is the differencing polynomial for Z_t. Suppose too that the stationary series $\delta^{(k)}(B) Z_t^{(k)}$, $1 \leq k \leq K$ all have absolutely summable autocovariance sequences (as they do when they obey ARMA models). Then (38) holds if and only if (37) does.*

Remark. The stochastic integral representation of stationary time series can be used to give frequency domain definitions of filters that have no direct time domain representation. For the reader familiar with it, this advanced methodology offers a faster path to (35) for any differencing operator and under the weaker assumption that all stationary processes involved have continuous spectral densities. Theorem 4.10.1 of Brockwell and Davis (1991) shows how the inverse $\delta(B)^{-1}$ of any differencing filter $\delta(B)$ can be defined for application to a series differenced with $\delta(B)$ to yield $\delta(B)^{-1}\{\delta(B) y_t\} = y_t$. Basic mean square convergence results of §4.6-4.9 of this reference that are related to the definition of the stochastic integrals then yield the uncorrelatedness assertions of Proposition 1 via arguments based on the Cauchy-Schwarz inequality like those used above. One applies the inverse of $\tilde{\delta}^{(j)}(B) = \prod_{k=1, k\neq j}^{K} \delta^{(k)}(B)$ to $\delta^{(\Pi)}(B) Z_t^{(j)}$ to obtain $\delta^{(j)}(B) Z_t^{(j)}$, $1 \leq j \leq K$ and (38).

5 Regarding Correlated Decompositions

In the Proposition 3 of this Section, for a given pair of spectral densities g_x and $g_{\tilde{x}}$, we describe the infinitely many possible covariance possibilities between jointly stationary series x_t and $\tilde{x}_t$ with these spectral densities such that the spectral density of $x_t + \tilde{x}_t$ is $g_x + g_{\tilde{x}}$. The covariance information is expressed by the cross-spectral density functions defined below.

5.1 Joint Stationarity and Cross-Spectral Densities

Two stationary series x_t and $\tilde{x}_t$ are said to be jointly stationary when the *cross-covariances* $Ex_t\tilde{x}_{t+j}, j = 0, \pm 1, \ldots$ do not depend on t. In this case, in addition to the lag j autocovariances $\gamma_{x,j} = Ex_t x_{t+j}$ and $\gamma_{\tilde{x},j} = E\tilde{x}_t\tilde{x}_{t+j}$, we consider the lag j cross-covariances $\gamma_{x\tilde{x},j} = Ex_t\tilde{x}_{t+j}$ and $\gamma_{\tilde{x}x,j} = E\tilde{x}_t x_{t+j}$, observing from $Ex_t\tilde{x}_{t-j} = E\tilde{x}_{t-j}x_t$ that

$$\gamma_{x\tilde{x},-j} = \gamma_{\tilde{x}x,j} \; j = 0, \pm 1, \ldots \;. \tag{39}$$

For the sum $x_t + \tilde{x}_t$, with $\gamma_{x+\tilde{x},j} = E\left(x_t + \tilde{x}_t\right)\left(x_{t+j} + \tilde{x}_{t+j}\right)$, we have

$$\gamma_{x+\tilde{x},j} = \gamma_{x,j} + \gamma_{\tilde{x},j} + \gamma_{x\tilde{x},j} + \gamma_{\tilde{x}x,j}. \tag{40}$$

The *cross-spectral densities* can be defined by $g_{x\tilde{x}}\left(\lambda\right) = \sum_{j=-\infty}^{\infty}\gamma_{x\tilde{x},j}e^{-i2\pi j\lambda}$ and $g_{\tilde{x}x}\left(\lambda\right) = \sum_{j=-\infty}^{\infty}\gamma_{\tilde{x}x,j}e^{-i2\pi j\lambda}$, $-1/2 \le \lambda \le 1/2$. They complete the definition of the joint spectral density matrix,

$$g_{x,\tilde{x}} = \begin{bmatrix} g_x & g_{x\tilde{x}} \\ g_{\tilde{x}x} & g_{\tilde{x}} \end{bmatrix} = \begin{bmatrix} g_x & g_{x\tilde{x}} \\ \bar{g}_{x\tilde{x}} & g_{\tilde{x}} \end{bmatrix}, \tag{41}$$

whose second expression follows from

$$g_{\tilde{x}x}\left(\lambda\right) = \sum_{j=-\infty}^{\infty}\gamma_{x\tilde{x},-j}e^{-i2\pi j\lambda} = \sum_{j=-\infty}^{\infty}\gamma_{x\tilde{x},j}e^{i2\pi j\lambda} = \bar{g}_{x\tilde{x}}\left(\lambda\right), \tag{42}$$

where $\bar{g}_{x\tilde{x}}\left(\lambda\right)$ denotes the complex conjugate of $g_{x\tilde{x}}\left(\lambda\right)$. The spectral density matrix $g_{x,\tilde{x}}$ is Hermitian positive semi-definite, which in this bivariate case is equivalent (Sylvester's criterion) to $\det g_{x,\tilde{x}}\left(\lambda\right) \ge 0$ and therefore to

$$\left|g_{x\tilde{x}}(\lambda)\right|^2 \le g_x\left(\lambda\right)g_{\tilde{x}}\left(\lambda\right), \; -1/2 \le \lambda \le 1/2. \tag{43}$$

In the bivariate ARMA case, the autocovariances and cross-covariances of x_t and $\tilde{x}_t$ decay exponentially to 0, and the spectral and cross-spectral densities are continuous. The analogue of (12) holds for cross-covariances $\gamma_{x\tilde{x},j}$:

$$\gamma_{x\tilde{x},j} = \int_{-1/2}^{1/2} e^{i2\pi j\lambda}g_{x\tilde{x}}\left(\lambda\right)d\lambda, \; j = 0, \pm 1, \ldots \;. \tag{44}$$

From (40) and (42), the spectral density $g_{x+\tilde{x}}\left(\lambda\right) = \sum_{j=-\infty}^{\infty}\gamma_{x+\tilde{x},j}e^{i2\pi j\lambda}$ of the sum $x_t + \tilde{x}_t$ has the decompositions

$$\begin{aligned} g_{x+\tilde{x}}\left(\lambda\right) &= g_x\left(\lambda\right) + g_{\tilde{x}}\left(\lambda\right) + g_{x\tilde{x}}\left(\lambda\right) + g_{\tilde{x}x}\left(\lambda\right) && (45) \\ &= g_x\left(\lambda\right) + g_{\tilde{x}}\left(\lambda\right) + 2\operatorname{Re} g_{x\tilde{x}}\left(\lambda\right), \; -1/2 \le \lambda \le 1/2. && (46) \end{aligned}$$

5.2 Characterizations of $g_{x+\tilde{x}} = g_x + g_{\tilde{x}}$

Here we characterize the structure of spectral density matrices of jointly stationary time series x_t and $\tilde{x}_t$ with the property that

$$g_{x+\tilde{x}} = g_x + g_{\tilde{x}}. \tag{47}$$

This enables us to describe all possible cross-spectral densities compatible with (47) for the continuous case.

It follows from (46) that (47) holds if and only if

$$\operatorname{Re} g_{x\tilde{x}}(\lambda) = 0, \ -1/2 \le \lambda \le 1/2,$$

that is, $g_{x\tilde{x}}(\lambda)$ is purely imaginary where it is nonzero. This can be expressed as

$$\bar{g}_{x\tilde{x}}(\lambda) = -g_{x\tilde{x}}(\lambda), \ -1/2 \le \lambda \le 1/2, \tag{48}$$

which yields

$$g_{x\tilde{x}}(0) = 0. \tag{49}$$

From (42), (48) is equivalent to

$$g_{\tilde{x}x} = -g_{x\tilde{x}} \tag{50}$$

yet another formulation of the model compatibility condition (9). More conveniently, (47) holds if and only if the joint spectral density matrix $g_{x,\tilde{x}}$ has the form

$$g_{x,\tilde{x}} = \begin{bmatrix} g_x & g_{x\tilde{x}} \\ -g_{x\tilde{x}} & g_{\tilde{x}} \end{bmatrix}. \tag{51}$$

Finally, combining (39) with $\gamma_{\tilde{x}x,j} = -\gamma_{x\tilde{x},j}$ from (40), we obtain the characterization

$$\gamma_{x\tilde{x},-j} = -\gamma_{x\tilde{x},j}, j = 0, \pm 1, \ldots \ . \tag{52}$$

The following Proposition characterizes all continuous $g_{x\tilde{x}}$, and therefore all cross-correlation possibilities, that are compatible with specified spectral densities g_x and $g_{\tilde{x}}$ in the continuous case.

Proposition 3 *For any given continuous spectral densities g_x and $g_{\tilde{x}}$, the continuous cross-spectral densities $g_{x\tilde{x}}$ associated with jointly stationary time series x_t and $\tilde{x}_t$ for which (47) holds are all functions of the form*

$$g_{x\tilde{x}}(\lambda) = \begin{cases} -if(-\lambda) & , -1/2 < \lambda < 0 \\ 0 & , \lambda = 0 \\ if(\lambda) & , 0 < \lambda \le 1/2 \end{cases}, \tag{53}$$

where f is any continuous real-valued function on $[0, 1/2]$ satisfying

$$|f(\lambda)|^2 \le g_x(\lambda)\, g_{\tilde{x}}(\lambda), 0 \le \lambda \le 1/2. \tag{54}$$

and

$$f(0) = 0. \tag{55}$$

Proof. This result follows from (50), (49), and the fact that, due to (54), each matrix function

$$g_{x,\tilde{x}} = \begin{bmatrix} g_x & g_{x\tilde{x}} \\ -g_{x\tilde{x}} & g_{\tilde{x}} \end{bmatrix} \tag{56}$$

with purely imaginary $g_{x\tilde{x}}$ defined by (53) is Hermitian positive semi-definite and therefore defines a joint spectral density matrix.

The cross-covariances associated with a given f can be obtained from (44). The case of no correlation is associated with $f = 0$.

Tucker McElroy provided the following example of a bivariate vector MA(1) whose spectral density matrix has the form (51).

Example. Let ε_t and $\tilde{\varepsilon}_t$ denote white noise processes with variances equal to 1 that are mutually uncorrelated: $\{\varepsilon_t\} \perp \{\tilde{\varepsilon}_t\}$. Define

$$\begin{bmatrix} x_t \\ \tilde{x}_t \end{bmatrix} = \begin{bmatrix} \varepsilon_t \\ \tilde{\varepsilon}_t \end{bmatrix} + \begin{bmatrix} 1 & \beta \\ -\beta & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{t-1} \\ \tilde{\varepsilon}_{t-1} \end{bmatrix}.$$

Then with I denoting the 2×2 identity matrix and

$$\Theta = \begin{bmatrix} 1 & \beta \\ -\beta & 1 \end{bmatrix}, \Theta^T = \begin{bmatrix} 1 & -\beta \\ \beta & 1 \end{bmatrix}$$

the joint spectral density matrix is given by

$$\begin{aligned} \begin{bmatrix} g_x(\lambda) & g_{x\tilde{x}}(\lambda) \\ g_{\tilde{x}x}(\lambda) & g_{\tilde{x}}(\lambda) \end{bmatrix} &= \left[I + \Theta e^{i2\pi\lambda}\right]\left[I + \Theta^T e^{-i2\pi\lambda}\right] \\ &= I + \Theta\Theta^T + \Theta e^{i2\pi\lambda} + \Theta^T e^{-i2\pi\lambda} \\ &= \begin{bmatrix} 2 + \beta^2 + 2\cos 2\pi\lambda & \beta\left(e^{i2\pi\lambda} - e^{-i2\pi\lambda}\right) \\ -\beta\left(e^{i2\pi\lambda} - e^{-i2\pi\lambda}\right) & 2 + \beta^2 + 2\cos 2\pi\lambda \end{bmatrix}, \end{aligned}$$

see Example 11.8.1 of Brockwell and Davis (1991) for the general vector ARMA generalization of (16). Thus the compatibility condition (50) holds.

We now describe several broad classes of examples.

5.3 Examples with Correlated Components after Differencing

Consider signal extraction from an ARIMA $Z_t = S_t + N_t$ whose differencing polynomial $\delta(B)$ has a factor of $(1 - B)$. Let g_x and $g_{\tilde{x}}$ denote the ARMA spectral densities of $x_t = \delta(B) S_t$ and $\tilde{x}_t = \delta(B) N_t$, respectively. Then either $g_x(0) = 0$ or $g_{\tilde{x}}(0) = 0$. In this case, $f = \sqrt{g_x g_{\tilde{x}}}$ satisfies both (54) and (55), and the joint spectral density matrix (56) with $g_{x\tilde{x}}$ defined by (53) is always singular, $\det g_{x,\tilde{x}}(\lambda) = 0$, for all λ. In this case, the correlation between components of the decomposition $w_t = x_t + \tilde{x}_t$ is *complete*: each component is a linear filtered version of the other. For example $\tilde{x}_t = \eta_{\tilde{x}}(B) x_t$, where the filter transfer function has the formula[3] $\eta_{\tilde{x}}\left(e^{-i\lambda}\right) = -g_{x\tilde{x}}(\lambda) g_x(\lambda)^{-1}$ (set equal to 0 where $g_x(\lambda) = 0$ if this occurs, which can happen for only finitely many λ). The factor $-g_{x\tilde{x}}(\lambda) = g_{\tilde{x}x}(\lambda)$ will cause $\eta_{\tilde{x}}\left(e^{-i\lambda}\right)$ to be purely imaginary, with $\bar{\eta}_{\tilde{x}}\left(e^{-i\lambda}\right) = \eta_{\tilde{x}}\left(e^{i\lambda}\right) = -\eta_{\tilde{x}}\left(e^{-i\lambda}\right)$, so the filter $\eta_{\tilde{x}}(B)$ will have the form $\sum_{j=1}^{\infty} \eta_j \left(B^{-j} - B^j\right)$.

For the example (22), where $g_x(\lambda) = \sigma_b^2 \left|1 + e^{i2\pi\lambda}\right|^2 = 2\sigma_b^2 (1 + \cos 2\pi\lambda)$ and $g_{\tilde{x}}(\lambda) = \sigma_c^2 \left|1 - e^{i2\pi\lambda}\right|^2 = 2\sigma_c^2 (1 - \cos 2\pi\lambda)$, we have

$$f(\lambda) = \sqrt{g_x(\lambda) g_{\tilde{x}}(\lambda)} = 2\sigma_b\sigma_c\sqrt{1 - \cos^2 2\pi\lambda} = 2\sigma_b\sigma_c \left|\sin 2\pi\lambda\right|.$$

To obtain a different f for (22) with the property that the cross-spectral density $g_{\tilde{x}x}$ defined by (53) is such that (56) is non-singular except when $g_x(\lambda) g_{\tilde{x}}(\lambda) = 0$ (at $\lambda = 0, \pm 1/4, 1/2$), one could define

$$\begin{aligned} f(\lambda) &= \sqrt{2}\sigma_b\sigma_c \min_{0\le\lambda\le 1/2} \left\{\sqrt{1 + \cos 2\pi\lambda}, \sqrt{1 - \cos 2\pi\lambda}\right\} \\ &= \begin{cases} \sqrt{2}\sigma_b\sigma_c\sqrt{1 - \cos 2\pi\lambda} & , 0 \le \lambda \le 1/4 \\ \sqrt{2}\sigma_b\sigma_c\sqrt{1 + \cos 2\pi\lambda} & , 1/4 < \lambda \le 1/2 \end{cases}. \end{aligned}$$

More generally, for all decompositions $g_w(\lambda) = g_x(\lambda) + g_{\tilde{x}}(\lambda)$ of ARMA spectral densities such that $g_x(0) g_{\tilde{x}}(0) = 0$, the definition $f(\lambda) = \sqrt{\min_{0\le\lambda\le 1/2}\{g_x(\lambda), g_{\tilde{x}}(\lambda)\}}$ will result in (56) being non-singular for all but finitely many λ. As a result, there will be correlation, but not complete correlation, between x_t and $\tilde{x}_t$.

5.4 No Correlation versus Correlation

In the ARIMA signal extraction context, to the extent that all of the model-relevant statistical information in the data is expressed in the model spectral density of $\delta(B) Z_t$, there is no statistical information available to favor one of the function f in the infinite collection described by Proposition 3 over another. The choice $f = 0$ can be viewed as the neutral choice, also in the following specific sense: with any non-zero f that satisfies the conditions of Proposition 3,

[3] In general, for jointly stationary x_t and $\tilde{x}_t$, the mean square optimal linear approximation to $\tilde{x}_t$ from $x_{t\pm j}$, $0 \le j < \infty$ is given by $\widehat{x}_t = \eta(B) x_t$, where $\eta(B)$ has transfer function $\eta\left(e^{-i2\pi\lambda}\right) = g_{\tilde{x}x}(\lambda) g_x(\lambda)^{-1}$. This well-known result follows from the fact that mean square optimality is characterized by the property that the errors $e_t = \tilde{x}_t - \eta(B) x_t$ satisfy $\{e_t\} \perp \{x_t\}$, which is equivalent to $g_{ex} = 0$ for the cross spectral density of e_t and x_t. We have $g_{ex}(\lambda) = g_{\tilde{x}x}(\lambda) - \eta\left(e^{-i2\pi\lambda}\right) g_x(\lambda) = g_{\tilde{x}x}(\lambda) - g_{\tilde{x}x}(\lambda) g_x(\lambda)^{-1} g_x(\lambda) = 0$. It follows from the uncorrelatedness of the decomposition $\tilde{x}_t = \widehat{x}_t + e_t$ that

$$\begin{aligned} g_{\tilde{x}}(\lambda) &= g_{\widehat{x}}(\lambda) + g_e(\lambda) \\ &= \left|\eta\left(e^{-i2\pi\lambda}\right)\right|^2 g_x(\lambda) + g_e(\lambda) \\ &= \left|g_{\tilde{x}x}(\lambda)\right|^2 / g_x(\lambda) + g_e(\lambda). \end{aligned}$$

When $\left|g_{\tilde{x}x}(\lambda)\right|^2 = g_{\tilde{x}}(\lambda) g_x(\lambda)$, the final formula for $g_{\tilde{x}}(\lambda)$ becomes $g_{\tilde{x}}(\lambda) = g_{\tilde{x}}(\lambda) + g_e(\lambda)$, showing that $g_e(\lambda) = 0$ for all λ. Therefore $e_t = 0$ for all t, so $\tilde{x}_t = \widehat{x}_t = \eta(B) x_t$.

its opposite function $-f$ could just as well be used to generate cross-correlations, and $-f$ would give the exact opposite cross-correlations from those obtained with f. Of course, the choice of no correlation has always had the important practical advantage that it provides the simplest linear mean square optimal signal estimation formulas with the least amount of information required.

Remark. The ARIMA process decomposition with correlated components after minimal differencing that has received considerable attention is the "permanent plus transitory" decomposition for ARIMA (p,1,q) processes Z_t, due Beveridge and Nelson (1981), also see Gómez and Breitung (1999). It is based on the Wold decomposition $w_t = a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j} = \psi(B) a_t$ of $w_t = (1-B) Z_t$, where a_t is the white noise innovations process of w_t. Its differenced components are completely correlated. The Beveridge-Nelson decomposition lies outside the scope of our discussion because its implied decomposition of w_t,

$$w_t = \psi(1) a_t + \{\psi(B) - \psi(1)\} a_t, \tag{57}$$

does not satisfy the compatibility condition (9) with $x_t = \psi(1) a_t$ and $\tilde{x}_t = \{\psi(B) - \psi(1)\} a_t$ except in the degenerate case $\psi(1) = 0$. This follows from cross-covariance calculation, which yields

$$\begin{aligned} g_{x\tilde{x}}(\lambda) &= \sigma_a^2 \psi(1)(1-\psi(1)) + \sigma_a^2 \psi(1) \sum_{j=1}^{\infty} \psi_j e^{-i2\pi\lambda} \\ &= \sigma_a^2 \psi(1)(1-\psi(1)) + \sigma_a^2 \psi(1) \left(\psi\left(e^{-i2\pi\lambda}\right) - 1\right). \end{aligned}$$

The decomposition (57) is compatible with the ARIMA (p,1,q) in the different sense that (57) uniquely identifies cross-covariances. Hence (45) is automatically satisfied in a meaningful way. It does not have a mean square optimality property.

5.5 More Than Two Components

We now briefly comment on correlation possibilities for ARIMA model decompositions of a time series $Z_t = \sum_{k=1}^{K} Z_t^{(k)}$ with $K > 2$ components and differencing polynomial $\delta(B)$. With $x_t^{(k)} = \delta(B) Z_t^{(k)}$, let g_{jk} denote the cross-spectral density of $x_t^{(j)}$ and $x_t^{(k)}$, for $j \neq k = 1, \ldots K$. The model compatibility condition (36) is equivalent to $\mathrm{Re}\left(\sum_{j\neq k:,j,k=1}^{K} g_{jk}\right) = 0$, which we have not undertaken to characterize for $K > 2$.

Ghysels (1987) has proposed, for illustrative purposes, a correlated three-component model for first differences of demand for money as the sum of the first differences of cycle, seasonal and trend. In this decomposition, the first differences of the seasonal component are assumed to follow a first-order stationary seasonal autoregressive model (a nonstandard model). This three component model is outside our framework because the same differencing operator $\delta(B) = 1 - B$ is used for the money supply and its components, in conflict with our requirements for $\delta(B)$, e.g. $\delta(B) \neq \prod_{k=1}^{3} \delta^{(k)}(B) = \delta(B)^3$. The first differences of the cycle are correlated with those of the seasonal and trend in quite complex ways. Ghysels (1987) provides a state space formulation that can be used to estimate the stationary first-differenced components in a mean square optimal way.

Acknowledgements. We are indebted to Agustin Maravall and William Bell for influential comments and suggestions from an earlier draft of this note, and similarly to Tucker McElroy, who commented helpfully on several drafts and provided the bivariate MA(1) example of Subsection 5.2. Brian Monsell is thanked for his careful reading of two earlier drafts.

6 Conclusions

In Proposition 1 and Corollary 2, by showing the inherent compatibility of the requirement that minimally differenced decomposition components be uncorrelated, we have completed the theoretical linear mean square optimality theory supporting the currently implemented methods for calculating model-based seasonal adjustments. With Proposition 3 and the subsequent discussion, we have provided further support for these methods by showing that the assumption of uncorrelated components after minimal differencing is the most neutral among an infinite set of correlated alternatives compatible with the ARIMA model for the observed series. The choice of no correlation has the important advantages that it yields the simplest linear mean square optimal signal estimation formulas requires the least amount of information.

References

[1] Bell, W. R. (1984) Signal Extraction for Nonstationary Time Series. *Annals of Statistics* 12, 646–664.

[2] Beveridge, S. and C. R. Nelson (1981) A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle. Journal of Monetary Economics, 7, 151-174.

[3] Brockwell, P. J. and R. A. Davis (1991) *Time Series: Theory and Methods* (2nd Ed.) Springer-Verlag: New York.

[4] Burman, J. P. (1980) Seasonal Adjustment by Signal Extraction. *Journal of the Royal Statistical Society* A 143, 321–337.

[5] Caporello, G. and A. Maravall (2004) Program TSW: Revised Manual, Documentos Ocasionales 0408, Bank of Spain.

[6] Ghysels, E. (1987) Seasonal Extraction in the Presence of Feedback. Journal of Business & Economic Statistics, 5, 191-194.

[7] Gómez, V. and J. Breitung (1999) The Beveridge-Nelson Decomposition: A Different Perspective with New Results. Journal of Time Series Analysis, 20, 527-536.

[8] Gómez, V. and A. Maravall (1996) Programs TRAMO and SEATS : Instructions for the user (beta version:June 1997). Banco de España, Servicio de Estudios, DT 9628. (Updates and additional documentation can be found at
http://www.bde.es/webbde/es/secciones/servicio/software
/econom.html.

[9] Gómez, V. and A. Maravall (2001) Seasonal Adjustment and Signal Extraction in Economic Time Series. In Peña, D., G. C. Tiao and R. S. Tsay, (eds.): *A Course in Time Series Analysis*, Chapter 8. New York: J. Wiley and Sons.

[10] Hillmer, S. C. and G. C. Tiao, (1982) An ARIMA-Model-Based Approach to Seasonal Adjustment. *Journal of the American Statistical Association,* 77 63–70.

[11] Koopman S. J., A. C. Harvey, J. A. Doornik, and N. Shephard, (2000) *STAMP 6.0: Structural Time Series Analyser, Modeller and Predictor.* Timberlake Consultants Press: London.

[12] McElroy, T. S. (2008) Matrix Formulas for Nonstationary ARIMA Signal Extraction. *Econometric Theory* 28, 988–1009. Preprint available at http://www.census.gov/ts/papers/matform3.pdf

[13] U.S. Census Bureau (2012) X-13ARIMA-SEATS Reference Manual.

[14] Wikipedia Contributors (2011) Partial Fraction, Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/wiki/Partial_fraction . (accessed October 21, 2011)

[15] Wikipedia Contributors (2012) Complex Number. *Wikipedia, The Free Encyclopedia* http://en.wikipedia.org/wiki/Complex_number (Accessed January 29, 2012).

[16] Zygmund, A. (1968) *Trigonometric Series.* Cambridge University Press: Cambridge.

7 Appendix

7.1 Proof of Proposition 1

(i) Suppose $\delta(B) = 1 - B^s$ for $s \geq 1$. We will show for $y_t(\beta) = (1 - \beta B^s)^{-1}(1 - B^s)$ that

$$\lim_{\beta \uparrow 1} E\{y_t(\beta) - y_t\}^2 = 0, \tag{58}$$

which is a re-expression of (27).

Observe that

$$\begin{aligned} y_t(\beta) - y_t &= (1-\beta B^s)^{-1}\{(1-B^s) - (1-\beta B^s)\} y_t \\ &= (\beta - 1)\left[(1-\beta B^s)^{-1} B^s\right] y_t, \end{aligned} \tag{59}$$

and that the squared gain of the filter $A_\beta(B) = (1-\beta B^s)^{-1} B^s$ has the formula $\left|A_\beta\left(e^{-i2\pi\lambda}\right)\right|^2 = \left|1-\beta e^{-i2\pi s\lambda}\right|^{-2}$. Therefore, making the appropriate changes of notation in (15) and in (12) with $k=0$, we obtain from (59) that

$$E\{y_t(\beta) - y_t\}^2 = (\beta-1)^2 \int_{-1/2}^{1/2} \left|1-\beta e^{-i2\pi s\lambda}\right|^{-2} g_y(\lambda)\, d\lambda, \tag{60}$$

with $g_y(\lambda) = \sum_{j=-\infty}^{\infty} \gamma_j e^{-i2\pi j\lambda}$. The function $\left|1-\beta e^{-i2\pi s\lambda}\right|^{-2}$ has the form of the spectral density of a gap AR(s) process (18) with $\varphi = \beta$ and and $\sigma_a^2 = 1$. Using (19) in Parseval's identity, see Theorem 2.4.2 of Brockwell and Davis (1991) or (8.2) of Zygmund (1968, p. 157), we obtain

$$\begin{aligned} (\beta-1)^2 \int_{-1/2}^{1/2} \left|1-\beta e^{-i2\pi\lambda}\right|^{-2} g_y(\lambda)\, d\lambda &= \frac{(\beta-1)^2}{1-\beta^2} \sum_{k=-\infty}^{\infty} \beta^{|k|}\gamma_{ks} \\ &\le \frac{1-\beta}{1+\beta}\left\{\gamma_0 + 2\sum_{j=0}^{\infty} |\gamma_{ks}|\right\} \end{aligned} \tag{61}$$

with the inequality following from $\left|\sum_{k=-\infty}^{\infty} \beta^{|k|}\gamma_{ks}\right| \le \gamma_0 + 2\sum_{j=0}^{\infty} |\gamma_{ks}| < \infty$ for any $0 < \beta < 1$. Now (58) follows from (61) and (26).

(ii) It is always the case that $\{x_t\} \perp \{y_t\}$ implies $\{x_t\} \perp \{\delta(B) y_t\}$ because $\delta(B) y_t$ is a finite linear combination of variates y_{t-j}. Thus it remains to establish the forward assertion in (35),

$$\{x_t\} \perp \{\delta(B) y_t\} \Longrightarrow \{x_t\} \perp \{y_t\}. \tag{62}$$

From (58), as in (34), the Cauchy-Schwarz inequality yields

$$\{x_t\} \perp \left\{(1-B^s)^d y_t\right\} \Longrightarrow \{x_t\} \perp \{y_t\} \tag{63}$$

for $d=1$. We can obtain (63) for $d \ge 2$ by induction. Define $y_t^{(1)} = y_t$ and $y_t^{(d)} = (1-B^s)\, y_t^{(d-1)}$ for $d \ge 2$, and note that each $y_t^{(d)}$ has an absolutely summable autocovariance sequence because it is a fixed finite linear combination of variates y_{t-j}. Thus we can apply the $(1-B^s)$ result to obtain $\{x_t\} \perp \left\{y_t^{(d)}\right\} \Longrightarrow \{x_t\} \perp \left\{y_t^{(d-1)}\right\}$ and then apply the induction hypothesis for $y_t^{(d-1)}$ to obtain (63) for all $d \ge 2$.

An analogous induction argument shows that

$$\{x_t\} \perp \left\{ \prod_{j=1}^{J} \left(1 - B^{s(j)}\right)^{d(j)} y_t \right\} \Longrightarrow \{x_t\} \perp \{y_t\}, \tag{64}$$

for any integers $s(j) \geq 1, d(j) \geq 1$, $d = 1, \ldots, J$ and $J \geq 1$.

Given a $\delta(B)$ whose zeroes are rational, let J denote the number of distinct real zeroes ($e^{i2\pi(r/2)}$, $r = 0, 1$) plus the number of distinct complex conjugate pairs $e^{\pm i2\pi(r/s)}$, $r/s \neq 0, 1/2$. Indexing these in some way, let $d(j)$ denote the multiplicity of the j-th root or root pair and let $s(j)$ denote the denominator of its ratio $r(j)/s(j)$. Then there is a polynomial $\tilde{\delta}(B)$ such that $\tilde{\delta}(B)\delta(B) = \prod_{j=1}^{J} \left(1 - B^{s(j)}\right)^{d(j)}$. Since $\{x_t\} \perp \{\delta(B) y_t\}$ implies $\{x_t\} \perp \left\{\tilde{\delta}(B)\delta(B) y_t\right\}$, we can conclude from (64) that $\{x_t\} \perp \{y_t\}$.

7.2 Proof of Corollary 2

First consider the case $K = 2$. If $\delta^{(1)}(B) \neq 1$, apply (62) with $x_t = \delta^{(2)}(B)\delta^{(1)}(B) Z_t^{(1)}$, $y_t = \delta^{(2)}(B) Z_t^{(2)}$ and $\delta(B) = \delta^{(1)}(B)$, to obtain $\left\{\delta^{(2)}(B)\delta^{(1)}(B) Z_t^{(1)}\right\} \perp \left\{\delta^{(2)}(B) Z_t^{(2)}\right\}$. Next, if $\delta^{(2)}(B) \neq 1$, apply (62) with $x_t = \delta^{(2)}(B) Z_t^{(2)}$, $y_t = \delta^{(1)}(B) Z_t^{(1)}$, and $\delta(B) = \delta^{(2)}(B)$ to conclude that $\left\{\delta^{(1)}(B) Z_t^{(1)}\right\} \perp \left\{\delta^{(2)}(B) Z_t^{(2)}\right\}$. Thus the desired result holds for $K = 2$.

For decompositions of Z_t with $K > 2$ components, $Z_t = \sum_{k=1}^{K} Z_t^{(k)}$, it suffices to consider the case $K = 3$, because we only need to show that the members of any pair of properly differenced components are mutually uncorrelated and the third component can be the sum of the components not in the chosen pair. The three series $\delta^{(\Pi)}(B) Z_t^{(k)}$, $1 \leq k \leq 3$, can be written as $\delta^{(3)}(B)\left[\delta^{(2)}(B)\left(\delta^{(1)}(B) Z_t^{(1)}\right)\right]$, $\delta^{(3)}(B)\left[\delta^{(1)}(B)\left(\delta^{(2)}(B) Z_t^{(2)}\right)\right]$ and $\delta^{(\Pi)}(B) Z_t^{(3)}$ respectively. They are assumed to be mutually uncorrelated. We need to show that the stationary series $\delta^{(1)}(B) Z_t^{(1)}$and $\delta^{(2)}(B) Z_t^{(2)}$ are mutually uncorrelated. Assuming $\delta^{(3)}(B) \neq 1$, we can apply (62) twice with $\delta(B) = \delta^{(3)}(B)$ to conclude that $\left\{\delta^{(2)}(B)\left[\delta^{(1)}(B) Z_t^{(1)}\right]\right\} \perp \left\{\delta^{(1)}(B)\left[\delta^{(2)}(B) Z_t^{(2)}\right]\right\}$, and then the $K = 2$ result (assuming $\delta^{(1)}(B) \neq 1$ or $\delta^{(2)}(B) \neq 1$) to obtain $\left\{\delta^{(1)}(B) Z_t^{(1)}\right\} \perp \left\{\delta^{(2)}(B) Z_t^{(2)}\right\}$.

7.3 Abel Convergence

We now return to the issue raised in the Remark of Subsection 2.1 regarding the convergence of the infinite series formula for the spectral density g_w. This issue applies as well to cross-spectral densities. For continuous $g_w(\lambda)$, ordinary convergence of the series can fail on an infinite set of λ (that has measure

0), but *convergence in the sense of Abel* to $g_w(\lambda)$, a more general concept of convergence introduced here, holds for all λ. This approach to convergence was used in Subsection 4.1 to establish important cases of Proposition 1 in a very concrete way. We present it here in its original infinite series context, in a way that covers both the real-valued spectral densities (10) and the complex-valued cross-spectral densities (42) that are the focus of Proposition 3.

Let the real- or complex-valued function g be such that $\int_{-1/2}^{1/2} |g(\lambda)| \, d\lambda < \infty$. Its Fourier coefficients c_j, $j = 0, \pm 1, \ldots$ are defined by

$$c_j = \int_{-1/2}^{1/2} e^{i2\pi k\lambda} g(\lambda) \, d\lambda, j = 0, \pm 1, \ldots \; .$$

These are real-valued for the g we consider, which are spectral and cross-spectral densities, so only this case will be discussed. For the *symmetric sum* $\sum_{j=-n}^{n} c_j e^{-i2\pi j\lambda}$, substituting $e^{\pm i2\pi j\lambda} = \cos 2\pi j\lambda \pm i \sin 2\pi j\lambda$ and rearranging yields the identity

$$\sum_{j=-n}^{n} c_j e^{-i2\pi j\lambda} = c_0 + \sum_{j=1}^{n} (c_{-j} + c_j) \cos 2\pi j\lambda + i \sum_{j=1}^{n} (c_{-j} - c_j) \sin 2\pi j\lambda, \quad (65)$$

for all finite integers $n \geq 1$ and for $n = \infty$. Thus $c_0 + \sum_{j=1}^{n} (c_{-j} + c_j) \cos 2\pi j\lambda$ is the real part of $\sum_{j=-n}^{n} c_j e^{-i2\pi j\lambda}$ and $\sum_{j=1}^{n} (c_{-j} - c_j) \sin 2\pi j\lambda$ is its imaginary part. Convergence of an infinite sum of the form $\sum_{j=-\infty}^{\infty} c_j e^{-i2\pi j\lambda}$ will always mean convergence of the symmetric partial sums (65) for finite n tending to ∞. Note that $|c_{\pm j}| \leq \int_{-1/2}^{1/2} |g(\lambda)| \, d\lambda$ for all $j \geq 0$. This makes clear that if we replace the coefficients c_j in (65) by the exponentially decaying coefficients, $c_j \beta^{|j|}$ for any fixed $0 < \beta < 1$, all three of the resulting series will be absolutely convergent, e.g.

$$\begin{aligned}
\sum_{j=1}^{n} \left| (c_{-j} - c_j) \beta^{|j|} \sin 2\pi j\lambda \right| &\leq \sum_{j=1}^{n} (|c_{-j}| + |c_j|) \beta^{|j|} \\
&\leq 2 \int_{-1/2}^{1/2} |g(\lambda)| \, d\lambda \sum_{j=1}^{n} \beta^{|j|} \\
&= 2\beta (1 - \beta)^{-1} \int_{-1/2}^{1/2} |g(\lambda)| \, d\lambda < \infty.
\end{aligned}$$

Therefore series with the replaced coefficients are convergent uniformly in λ.

Definition 4 *The series $\sum_{j=-\infty}^{\infty} c_j e^{-i2\pi j\lambda}$ is said to converge in the sense of Abel to the number g_λ at frequency λ if, for $0 < \beta < 1$, the two sums on the right in (66) are convergent for each β and their sum plus c_0 converges to g_λ as β increases* 1,

$$g_\lambda = c_0 + \lim_{\beta \uparrow 1} \sum_{j=1}^{\infty} \beta^j (c_{-j} + c_j) \cos 2\pi j\lambda + i \lim_{\beta \uparrow 1} \sum_{j=1}^{\infty} \beta^j (c_{-j} - c_j) \sin 2\pi j\lambda. \quad (66)$$

The classical Abel convergence results, see Theorem (6.11) of Zygmund (1968), include the facts that (i) convergence of $\sum_{j=-\infty}^{\infty} c_j e^{-i2\pi j\lambda}$ in the usual sense to g_λ implies (66). (ii) If g is continuous at λ_0, then (66) holds for $\lambda = \lambda_0$, with $g_{\lambda_0} = g(\lambda_0)$, and the convergence is uniform over every interval $[\lambda_1, \lambda_2]$ on which $g$is continuous. Further (iii) if $g(\lambda)$ has a jump discontinuity at $\lambda = \lambda_0$, but the left and right hand limits $g(\lambda_0-) = \lim_{\lambda\uparrow\lambda_0} g(\lambda)$ and $g(\lambda_0+) = \lim_{\lambda\downarrow\lambda_0} g(\lambda)$ exist (with the same sign if they are infinite), then

$$\frac{g(\lambda_0-) + g(\lambda_0+)}{2} = c_0 + \sum_{j=-\infty}^{\infty} c_j e^{-i2\pi j\lambda} \tag{67}$$

holds in the sense of Abel.