RESEARCH REPORT SERIES
(*Statistics #2002-04)*

Computation of Asymmetric Signal Extraction Filters and Mean Squared Error for ARIMA Component Models

William R. Bell[1] and Donald E. Martin[2]

Statistical Research Division
U.S. Bureau of the Census
Washington D.C. 20233

Report Issued: August 26, 2002

Disclaimer: This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This paper is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

[1]Methodology and Standards Directorate
[2]Statistical Research Division & Mathematics Department, Howard University

Computation of Asymmetric Signal Extraction Filters and Mean Squared Error for ARIMA Component Models

William R. Bell[1] and Donald E. K. Martin[1,2]
[1]Statistical Research Division, U.S. Census Bureau, and
[2]Howard University

June 26, 2002

Abstract

Standard signal extraction results for both stationary and nonstationary time series are expressed as linear filters applied to the observed series. Computation of the filter weights, and of the corresponding filter transfer function, is relevant for studying properties of the filter and of the resulting signal extraction estimates. Methods for doing such computations for symmetric, doubly infinite filters are well-established. This paper develops an algorithm for computing filter weights for asymmetric, semi-infinite signal extraction filters, including the important case of the concurrent filter (for signal extraction at the current time point.) The setting is where the time series components being estimated follow ARIMA (autoregressive-integrated-moving average) models. The algorithm provides expressions for the asymmetric signal extraction filters as rational polynomial functions of the backshift operator. The filter weights are then readily generated by simple expansion of these expressions, and the filter transfer function may be directly evaluated. Recursive expressions are also developed that relate the weights for filters that use successively increasing amounts of data. The results for the filter weights are then used to develop methods for computing mean squared error results for the asymmetric signal extraction estimates.

Disclaimer: This paper reports the results of research and analysis undertaken by Census Bureau staff and staff of Howard University. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

1 Introduction

Suppose an observed discrete time series Z_t is decomposed as

$$Z_t = S_t + N_t$$

and the objective is to use the data on Z_t to estimate the unobserved component series S_t and N_t. The component series might represent "signal plus noise," or "trend plus error," or "seasonal plus nonseasonal." Signal extraction results for optimal (minimum mean squared error, or MMSE) linear estimators of the components were given in the stationary case by Kolmogorov (1939,1941) and Wiener (1949); see also Whittle (1963). Extensions to the case of nonstationary S_t but stationary N_t were given by Hannan (1967), Sobel (1967), and Cleveland and Tiao (1976). Bell (1984a) gave a more general treatment that covered the case where both S_t and N_t are nonstationary. These papers dealt with estimation of S_t and N_t from an infinite realization of the series Z_t, a case that applies approximately when the observed time series Z_t is sufficiently long. Ansley and Kohn (1985), Kohn and Ansley (1987), and Bell and Hillmer (1988) gave results for MMSE linear estimators based on a finite sample of Z_t.

The results of the papers cited in the preceding paragraph express the MMSE linear estimator of S_t based on some set of consecutive observations on Z_t as

$$\widehat{S}_t = \sum_k \alpha_k Z_{t-k} \tag{1}$$

where $\{\alpha_k\}$ is the set of "filter weights." The summation in (1) extends over the set of k such that all available observations Z_{t-k} are included. As noted this summation could be finite or infinite; here we focus on the "semi-infinite" case where the summation is $\sum_{k=-m}^{+\infty}$ for some integer m (thus using data up to and including time $t+m$). It is convenient to rewrite (1) as

$$\widehat{S}_t = \alpha(B) Z_t$$

where

$$\alpha(B) = \sum_k \alpha_k B^k \tag{2}$$

and B is the backshift operator ($BZ_t = Z_{t-1}$). Results from the references cited above on infinite sample signal extraction give expressions for $\alpha(B)$ as functions of the models for Z_t, S_t, and N_t (or, equivalently, as functions of their spectral densities or autocovariance generating functions). From these expressions one may wish to compute the actual filter weights $\{\alpha_k\}$, or their frequency response function $\alpha(e^{i\lambda})$, since studying these quantities gives insight into the nature of the signal extraction estimate (1). (For examples of this sort of analysis, see Findley and Martin (2002).)

Direct computation of signal extraction weights from these expressions for asymmetric filters is awkward, however, since it requires that one manipulate truncations used to approximate infinite series expansions.

This paper develops an algorithm to obtain explicit expressions for asymmetric filters $\alpha(B)$ for signal extraction based on semi-infinite samples for the important case when Z_t, S_t, and N_t all follow ARIMA (autoregressive-integrated-moving average) models (Box and Jenkins 1970). The results express a given asymmetric signal extraction filter as a rational polynomial function of B, so the filter weights $\{\alpha_k\}$ can be obtained by direct expansion from these expressions, and $\alpha(e^{i\lambda})$ can also be calculated directly. This provides, for any m, the filter weights used to estimate S_t based on observations up to Z_{t+m} and extending back into the infinite past (i.e., using Z_{t+m-k} for $k = 0, 1, 2, \ldots$). Our algorithm generalizes an algorithm of Granville Tunnicliffe-Wilson (reported in Burman 1980) for computing weights for a symmetric signal extraction filter (using Z_{t-k} for $k = 0, \pm 1, \pm 2, \ldots$). Additional results of our paper relate the filter weights corresponding to different values of m, and provide methods for computing asymmetric signal extraction MSE. Trivial extensions of the results would accommodate decompositions that involve more than two components (e.g., seasonal plus trend plus irregular).

Koopman and Harvey (2000) give an algorithm for computing finite sample signal extraction weights for general linear state space models. To deal with the finite sample case, their approach uses results of the Kalman filter and an associated smoothing algorithm, and so is completely different from the approach presented here. For sufficiently long series following ARIMA component models their approach and ours will, of course, give approximately the same results.

Section 2 of the paper presents our algorithm for the important case $m = 0$, which produces the "concurrent" or "one-sided" signal extraction filter that estimates S_t using data up through the current observation Z_t. Section 3 gives the simple extension of the algorithm to the case of general m. Using these results, Section 4 notes some unit root properties of the resulting filters, and Section 5 derives expressions for the asymmetric signal extraction MSE. Sections 6 and 7 then provide examples illustrating the results of the previous sections, showing how to apply the algorithm to calculate the filter weights. Section 6 considers a simple trend estimation example, and Section 7 an example of canonical ARIMA model-based seasonal adjustment.

Some results of Sections 4-6 overlap with results of Pierce (1979,1980), and these instances of overlap will be noted. Pierce obtained results for the general difference stationary case where "differenced" S_t and N_t are assumed to be stationary time series though not necessarily following ARMA models. Where the results overlap, our results provide the simplifications that result for ARIMA models, in that they yield expressions of the signal extraction filters as rational polynomials in B, whereas Pierce's results give expressions that involve infinite series expansions.

2 Algorithm for Calculating Concurrent ("One-Sided") Filters ($m = 0$)

We assume that the time series Z_t, S_t, and N_t follow the ARIMA models

$$\begin{aligned} \varphi(B)Z_t &= \theta(B)a_t \\ \varphi_s(B)S_t &= \theta_s(B)b_t \\ \varphi_n(B)N_t &= \theta_n(B)e_t \end{aligned} \tag{3}$$

where $\varphi(B)$, $\varphi_s(B)$, $\varphi_n(B)$ have degrees p, p_s, p_n, and $\theta(B)$, $\theta_s(B)$, $\theta_n(B)$ have degrees q, q_s, q_n, respectively. The series a_t, b_t, and e_t are independent white noise series with variances σ^2, σ_b^2, σ_e^2, respectively. Note that $\varphi(B)$, $\varphi_s(B)$, $\varphi_n(B)$ are the products of any stationary and nonstationary autoregressive operators in the above three models for Z_t, S_t, and N_t, respectively. We let $\delta(B)$, $\delta_s(B)$, $\delta_n(B)$ denote the nonstationary AR operators. In typical applications these involve differencing operators (e.g., $(1-B)^d$ or $(1-B^{12})$) or "seasonal summation operators" (e.g., $1 + B + \cdots + B^{11}$).

We shall assume that the mean functions, $E(S_t)$ and $E(N_t)$, of the time series S_t and N_t are both zero. Hence, the mean function of Z_t is also zero. Equivalently, if the mean functions are not zero but have been modeled (e.g., by linear regression functions), they can be subtracted from the respective time series, the signal extraction performed, and the appropriate mean function added back to the signal extraction estimate.

We assume $\varphi_s(B)$ and $\varphi_n(B)$ have no common factors, so

$$\varphi(B) = \varphi_s(B)\varphi_n(B). \tag{4}$$

We also assume $\varphi_s(B)$ and $\varphi_n(B)$ have all their zeroes on or outside the unit circle. (This assumption could be relaxed to allow for explosive models. Also, the assumption of no stationary autoregressive factors common to $\varphi_s(B)$ and $\varphi_n(B)$ is for convenience of presentation and is not essential to the results.) For example, for seasonal adjustment of a monthly series, typically $\varphi_s(B) = 1 + B + \cdots + B^{11}$ and $\varphi_n(B) = (1-B)^d$ for $d = 1$ or 2, though $\varphi_n(B)$ could also contain a stationary AR operator. Finally, we assume $\varphi(B)$ and $\theta(B)$ have no common factors, and $\theta(B)$ has all its zeroes outside the unit circle (it is invertible). Consideration of the autocovariance generating function (ACGF) of $\varphi(B)Z_t$ shows that the component model AR and MA polynomials satisfy the constraint

$$\sigma^2\theta(B)\theta(F) = \sigma_b^2\varphi_n(B)\varphi_n(F)\theta_s(B)\theta_s(F) + \sigma_e^2\varphi_s(B)\varphi_s(F)\theta_n(B)\theta_n(F)$$

where $F = B^{-1}$ is the forward shift operator ($FZ_t = Z_{t+1}$).

Denote the concurrent signal extraction filter for estimating S_t by $\alpha(B)$. The concurrent filter for estimating N_t is $1 - \alpha(B)$. From the references cited in the introduction (e.g., Whittle (1963, ch. 6) or Hannan (1970, p. 168)), the filter $\alpha(B)$ is given by

$$\alpha(B) = \frac{1}{\sigma^2}\pi(B)\left[\pi(F)\gamma_s(B)\right]_+ \tag{5}$$

where $\pi(B) = \varphi(B)/\theta(B)$ is the infinite AR operator for Z_t, $\gamma_s(B)$ is the pseudo-autocovariance generating function of S_t ($\gamma_s(B) = \sigma_b^2\theta_s(B)\theta_s(F)/\varphi_s(B)\varphi_s(F)$), and the notation $[\bullet]_+$ indicates only terms with nonnegative powers of B are retained—those with positive powers of $F = B^{-1}$ are dropped. (Section 5 demonstrates that (5) actually provides the optimal signal extraction estimate in the general nonstationary case considered here.) In (5) we could use more explicit notation and write $\alpha_s^{(0)}(B)$ instead of just $\alpha(B)$, with the superscript (0) indicating that this is the concurrent filter ($m = 0$) and the subscript s indicating it is for estimating S_t (rather than N_t). In this section, to avoid cluttering the notation, we omit these notational details from $\alpha(B)$ and also from some quantities defined shortly (e.g., $c(F)$ and $d(B)$). Starting with Section 3, however, it becomes necessary to make the notation more explicit by including such details.

From (3)–(5) we have

$$\begin{aligned}\alpha(B) &= \frac{1}{\sigma^2}\frac{\varphi_s(B)\varphi_n(B)}{\theta(B)}\left[\frac{\varphi_s(F)\varphi_n(F)\theta_s(B)\theta_s(F)\sigma_b^2}{\theta(F)\varphi_s(B)\varphi_s(F)}\right]_+ \\ &= \frac{\sigma_b^2}{\sigma^2}\times\frac{\varphi_s(B)\varphi_n(B)}{\theta(B)}\left[\frac{\varphi_n(F)\theta_s(F)\theta_s(B)}{\theta(F)\varphi_s(B)}\right]_+ .\end{aligned} \tag{6}$$

Notice $\alpha(B)$ depends on the variances only through the variance ratio σ_b^2/σ^2.

Assume that the term inside the brackets in (6) can be written as

$$\frac{\varphi_n(F)\theta_s(F)\theta_s(B)}{\theta(F)\varphi_s(B)} = \frac{c(F)}{\theta(F)} + \frac{d(B)}{\varphi_s(B)} \tag{7}$$

where

$$\begin{aligned}c(F) &= c_0 + c_1F + \cdots + c_hF^h \\ d(B) &= d_0 + d_1B + \cdots + d_kB^k\end{aligned}$$

are determined to satisfy the relation

$$c(F)\varphi_s(B) + d(B)\theta(F) = \varphi_n(F)\theta_s(F)\theta_s(B). \tag{8}$$

The right-hand side of (8) is a polynomial in F and B of degree $(p_n + q_s, q_s)$, where the ordered pair gives the maximum powers of F and B that appear. Consideration of the left-hand side of (8) shows that, in general, we can set

$$\begin{aligned}h &= \max(q, p_n + q_s) \\ k &= \max(p_s, q_s).\end{aligned}$$

Now define

$$g(B) \equiv \sum_{j=-h}^{k} g_j B^j = \varphi_n(F)\theta_s(F)\theta_s(B) \tag{9}$$

where g_j is defined to be zero for values of $j > q_s$ or $j < -(p_n + q_s)$. The former occurs if $p_s > q_s$ and the latter if $q > p_n + q_s$.

Combining (8) and (9) gives

$$c(F)\varphi_s(B) + d(B)\theta(F) = g(B)$$

or more explicitly

$$\begin{aligned}\sum_{j=-h}^{k} g_j B^j \quad &= \quad (c_0 + c_1 F + \cdots + c_h F^h)(1 - \varphi_{s1}B - \cdots - \varphi_{s,p_s}B^{p_s}) \\ &\quad +(d_0 + d_1 B + \cdots + d_k B^k)(1 - \theta_1 F - \cdots - \theta_q F^q) \qquad (10)\end{aligned}$$

which provides $h + k + 1$ linear equations in $h + k + 2$ unknowns: $c_0, c_1, \ldots, c_h$ and $d_0, d_1, \ldots, d_k$. Since we have one more unknown than equations, we impose the constraint $c_0 = 0$. This implies that the bracketed term in (6) is

$$\begin{aligned}\left[\frac{\varphi_n(F)\theta_s(F)\theta_s(B)}{\theta(F)\varphi_s(B)}\right]_+ &= \left[\frac{c(F)}{\theta(F)} + \frac{d(B)}{\varphi_s(B)}\right]_+ \\ &= \frac{d(B)}{\varphi_s(B)} \qquad (11)\end{aligned}$$

since (i) the expansion of $c(F)/\theta(F)$ involves only terms in positive powers of F, which are terms in negative powers of B, and (ii) the expansion of $d(B)/\varphi_s(B)$ involves only terms in nonnegative powers of B. Combining (6) and (11) we see $\alpha(B)$ is given by

$$\begin{aligned}\alpha(B) &= \frac{\sigma_b^2}{\sigma^2} \times \frac{\varphi_s(B)\varphi_n(B)}{\theta(B)} \times \frac{d(B)}{\varphi_s(B)} \\ &= \frac{\sigma_b^2}{\sigma^2} \times \frac{\varphi_n(B)d(B)}{\theta(B)} . \qquad (12)\end{aligned}$$

From (12), once we have computed $d(B)$ we can compute the weights $\{\alpha_k\}$ in $\alpha(B)$ using standard computer routines for expanding rational polynomials. We now show how to compute $d(B)$ using (10).

Notice that with the constraint $c_0 = 0$ the equations (10) can be written in matrix form as follows:

$$
\begin{bmatrix} g_{-h} \\ g_{1-h} \\ \vdots \\ g_{-1} \\ g_0 \\ g_1 \\ \vdots \\ g_k \end{bmatrix}
=
\begin{bmatrix}
1 & & & \\
-\varphi_{s1} & 1 & & \\
\vdots & -\varphi_{s1} & \ddots & \\
\vdots & \vdots & \ddots & 1 \\
\vdots & \vdots & & -\varphi_{s1} \\
-\varphi_{s,p_s} & \vdots & & \vdots \\
 & -\varphi_{s,p_s} & & \vdots \\
 & & \ddots & \vdots \\
 & & & -\varphi_{s,p_s} \\
\hdashline
\multicolumn{4}{c}{\mathbf{0}_{(k+1-p_s)\times h}}
\end{bmatrix}
\begin{bmatrix} c_h \\ \vdots \\ c_1 \end{bmatrix}
$$

$$
+\begin{bmatrix}
\multicolumn{4}{c}{\mathbf{0}_{(h-q)\times(k+1)}} \\
\hdashline
-\theta_q & & & \\
\vdots & -\theta_q & & \\
\vdots & \vdots & \ddots & \\
\vdots & \vdots & & -\theta_q \\
-\theta_1 & \vdots & & \vdots \\
1 & -\theta_1 & & \vdots \\
 & 1 & \ddots & \vdots \\
 & & \ddots & -\theta_1 \\
 & & & 1
\end{bmatrix}
\begin{bmatrix} d_0 \\ d_1 \\ \vdots \\ d_k \end{bmatrix}
\tag{13}
$$

which can be rewritten as

$$
[A_1|A_2]\begin{bmatrix} c_h \\ \vdots \\ c_1 \\ d_0 \\ d_1 \\ \vdots \\ d_k \end{bmatrix} = g \tag{14}
$$

with g, A_1, and A_2 defined from the expressions in (13). The representation of the matrices A_1 and A_2 in (13) makes clear the pattern used in their construction.

First, for A_1 we construct the first column as shown in (13), and then construct $h-1$ additional columns by successively shifting down one position the entries of the previous column that correspond to 1, $-\varphi_{s1}$, ..., $-\varphi_{s,p_s}$. Note that since $k = \max(p_s, q_s) \geq p_s$, there is always at least one row of zeros at the bottom of A_1, with more than one row of zeros if $k = q_s > p_s$.

To construct A_2 we proceed in a similar fashion as for A_1. If $h > q$ we start by setting the first $h-q$ rows of A_2 to zero. If $h = q$ (because $q \geq p_n + q_s$), then we skip this step and there are no rows of zeros at the top of A_2. Then, immediately below any needed rows of zeros, we construct the first column as shown in (13), and then construct k additional columns by successively shifting down one position the entries of the previous column that correspond to $-\theta_q$, ..., $-\theta_1$, 1.

Appendix A proves that the matrix $[A_1|A_2]$ in (14) is nonsingular. Thus, (14) can be solved for $c_1, \ldots, c_h$ and $d_0, d_1, \ldots, d_k$, and we can then compute the expansion of $\alpha(B)$ from (12). We don't need to know $c_1, \ldots, c_h$ to compute $\alpha(B)$, but we get these as part of solving for $d_0, d_1, \ldots, d_k$. The quantities $c_1, \ldots, c_h$ will be used in later sections, however.

3 Extending the Algorithm to the Case of $m \neq 0$

We show two approaches to obtain the asymmetric signal extraction filters for $m \neq 0$. The first approach directly generalizes the approach for the concurrent filter ($m = 0$). The second obtains a recursive relation between the filters for successive values of m. Following presentation of these two approaches we discuss the relation between the signal extraction filters for S_t and N_t for any value of m.

3.1 Direct approach to calculating signal extraction filters for $m \neq 0$

For general m not necessarily zero, the asymmetric signal extraction filter, $\alpha_s^{(m)}(B)$, for estimating S_t from $Z_{t+m}, Z_{t+m-1}, \ldots$ is

$$\alpha_s^{(m)}(B) = \frac{F^m}{\sigma^2}\pi(B)\left[\pi(F)\gamma_s(B)B^m\right]_+ \tag{15}$$

which is a generalization of (5) that holds for all m. Results of this form are given for the stationary case by Whittle (1963, ch. 6), and for the case of a nonstationary signal with stationary noise by Hannan (1970, p. 168). Section 5 shows that this result is appropriate in the more general nonstationary case considered here.

Starting with (15), the resulting generalization of (6) is

$$\alpha_s^{(m)}(B) = F^m\frac{\sigma_b^2}{\sigma^2} \times \frac{\varphi_s(B)\varphi_n(B)}{\theta(B)}\left[\frac{\varphi_n(F)\theta_s(F)\theta_s(B)}{\theta(F)\varphi_s(B)}B^m\right]_+. \tag{16}$$

As before, we express the term in brackets in (16) as

$$\frac{\varphi_n(F)\theta_s(F)\theta_s(B)}{\theta(F)\varphi_s(B)}B^m = \frac{c_s^{(m)}(F)}{\theta(F)} + \frac{d_s^{(m)}(B)}{\varphi_s(B)} \tag{17}$$

where $c_s^{(m)}(F) = c_{s1}^{(m)}F + \cdots + c_{sh_m}^{(m)}F^{h_m}$ and $d_s^{(m)}(B) = d_{s0}^{(m)} + d_{s1}^{(m)}B + \cdots + d_{sk_m}^{(m)}B^{k_m}$ are determined to satisfy the relation

$$c_s^{(m)}(F)\varphi_s(B) + d_s^{(m)}(B)\theta(F) = \varphi_n(F)\theta_s(F)\theta_s(B)B^m. \tag{18}$$

Note we impose the same constraint as before, $c_{s0}^{(m)} = 0$. The right-hand side of (18) is now a polynomial in F and B of degree $(p_n + q_s - m, q_s + m)$, and considering the left-hand side of (18) we can set

$$\begin{aligned} h_m &= \max(q, p_n + q_s - m) \\ k_m &= \max(p_s, q_s + m). \end{aligned}$$

(Note: Strictly speaking we should write $h_s^{(m)}$ and $k_s^{(m)}$ for these quantities as analogous different quantities would be appropriate for computing the asymmetric filter $\alpha_n^{(m)}(B)$ for estimating N_t using data through $t+m$. We ignore this refinement to avoid overly complicated notation, particularly on the coefficients $c_{sh_m}^{(m)}$ and $d_{sk_m}^{(m)}$.)

We now define

$$g_s^{(m)}(B) \equiv \varphi_n(F)\theta_s(F)\theta_s(B)B^m = g(B)B^m \tag{19}$$

where $g(B) \equiv g_s^{(0)}(B) = \varphi_n(F)\theta_s(F)\theta_s(B)$ was given before in (9). Equation (19) shows that

$$g_s^{(m)}(B) \equiv \sum_{j=-h_m}^{k_m} g_{sj}^{(m)} B^j = \sum_{j=m-(p_n+q_s)}^{q_s+m} g_{j-m} B^j$$

where the weights $g_{sj}^{(m)} = g_{j-m}$ for $j = m - (p_n + q_s), \ldots, q_s + m$, and $g_{sj}^{(m)} = 0$ for $j < m - (p_n + q_s)$ or $j > q_s + m$.

Given h_m, k_m, and the $g_{sj}^{(m)}$, we can proceed exactly as before to solve for $c_{s1}^{(m)}, \ldots, c_{sh_m}^{(m)}$ and $d_{s0}^{(m)}, d_{s1}^{(m)}, \ldots, d_{sk_m}^{(m)}$ using equations of the same form as (13) and (14). That is, the fundamental quantities that change with m in (13) and (14) are just h_m and k_m. They determine the dimensions of the vectors and matrices in (13) and (14), and the positions in the vector g of the nonzero g_j. Once the versions of (13) and (14) appropriate for the given m are set up, this version of (14) can be solved for $c_{s1}^{(m)}, \ldots, c_{sh_m}^{(m)}$ and $d_{s0}^{(m)}, d_{s1}^{(m)}, \ldots, d_{sk_m}^{(m)}$.

The resulting solution for $\alpha_s^{(m)}(B)$ then follows as for (12):

$$\alpha_s^{(m)}(B) = \frac{\sigma_b^2}{\sigma^2} \times \frac{\varphi_n(B)d_s^{(m)}(B)}{\theta(B)}F^m. \tag{20}$$

Clearly a parallel derivation establishes that the asymmetric signal extraction filter for estimating N_t from $Z_{t+m}, Z_{t+m-1}, \ldots$ is

$$\alpha_n^{(m)}(B) = \frac{\sigma_e^2}{\sigma^2} \times \frac{\varphi_s(B) d_n^{(m)}(B)}{\theta(B)} F^m \tag{21}$$

where $d_n^{(m)}(B)$ is obtained in the analogous fashion to $d_s^{(m)}(B)$.

3.2 Recursive approach to calculating signal extraction filters for $m \neq 0$

First we consider the case $m > 0$. Let

$$\eta(F) \equiv \sum_{j=1}^{\infty} \eta_j F^j = \frac{c(F)}{\theta(F)}$$

where $c(F) \equiv c_s^{(0)}(F)$ is calculated for the concurrent signal extraction filter for S_t in Section 2. Considering (7) given there and (16) above, we see that

$$\begin{aligned}
\alpha_s^{(m)}(B) &= F^m \frac{\sigma_b^2}{\sigma^2} \times \pi(B) \left[\left(\eta(F) + \frac{d(B)}{\varphi_s(B)} \right) B^m \right]_+ \\
&= \frac{\sigma_b^2}{\sigma^2} \times \pi(B) \left[(\eta_1 F + \cdots + \eta_m F^m) + \frac{d(B)}{\varphi_s(B)} \right] \\
&= \alpha_s^{(0)}(B) + \frac{\sigma_b^2}{\sigma^2} \pi(B)(\eta_1 F + \cdots + \eta_m F^m) \qquad (22)
\end{aligned}$$

where $d(B) \equiv d_s^{(0)}(B)$ is calculated for the concurrent filter in Section 2. Applying $\alpha_s^{(m)}(B)$ as given in (22) to Z_t gives $\widehat{S}_{t|t+m}$, the signal extraction estimate of S_t using data through time $t+m$, and shows that (note $\pi(B) Z_t = a_t$)

$$\widehat{S}_{t|t+m} = \widehat{S}_{t|t} + \frac{\sigma_b^2}{\sigma^2}(\eta_1 F + \cdots + \eta_m F^m) a_t. \tag{23}$$

As $m \to \infty$, $\alpha_s^{(m)}(B)$ becomes the symmetric filter $\alpha_s^{(\infty)}(B)$, and (23) becomes

$$\widehat{S}_{t|\infty} = \widehat{S}_{t|t} + \frac{\sigma_b^2}{\sigma^2}(\eta_1 F + \eta_2 F^2 + \cdots) a_t \tag{24}$$

which shows that $\frac{\sigma_b^2}{\sigma^2}(\eta_1 F + \eta_2 F^2 + \cdots) a_t$ is the "total revision" from the concurrent estimate $\widehat{S}_{t|t}$ to the "final" estimate $\widehat{S}_{t|\infty}$ that is obtained from the symmetric filter. Equation (23) also makes it clear that for any $m' > m$

$$\begin{aligned}
\widehat{S}_{t|t+m'} &= \widehat{S}_{t|t+m} + \frac{\sigma_b^2}{\sigma^2}(\eta_{m+1} F^{m+1} + \cdots + \eta_{m'} F^{m'}) a_t \qquad (25) \\
\widehat{S}_{t|\infty} &= \widehat{S}_{t|t+m} + \frac{\sigma_b^2}{\sigma^2}(\eta_{m+1} F^{m+1} + \eta_{m+2} F^{m+2} + \cdots) a_t \qquad (26)
\end{aligned}$$

so that $\frac{\sigma_b^2}{\sigma^2}(\eta_{m+1}F^{m+1} + \cdots + \eta_{m'}F^{m'})a_t$ is the revision from $\widehat{S}_{t|t+m}$ to $\widehat{S}_{t|t+m'}$, and $\frac{\sigma_b^2}{\sigma^2}(\eta_{m+1}F^{m+1} + \eta_{m+2}F^{m+2} + \cdots)a_t$ is the revision from $\widehat{S}_{t|t+m}$ to $\widehat{S}_{t|\infty}$. Pierce (1980) gave general results on how signal extraction revisions depend on the one-step-ahead forecast errors a_t.

We can now see the fundamental importance of equation (7) and the constraint $c_0 = 0$. Equation (7) breaks $\frac{1}{\sigma_b^2}\pi(F)\gamma_s(B) = \varphi_n(F)\theta_s(F)\theta_s(B)/\theta(F)\varphi_s(B)$ into the two parts $c(F)/\theta(F)$ and $d(B)/\varphi_s(B)$. The second part gives rise to the concurrent estimate, $\widehat{S}_{t|t} = [\sigma_b^2\varphi_n(B)d(B)/\sigma^2\theta(B)]Z_t$ (derived in Section 2), which is a linear function of current and past Z_t. The first part, involving $\eta(F) = c(F)/\theta(F)$, gives rise to the revision from $\widehat{S}_{t|t}$ to $\widehat{S}_{t|t+m}$ in (23) (for $m > 0$) and to the revision from $\widehat{S}_{t|t}$ to $\widehat{S}_{t|\infty}$ in (24), as given by the second terms on the right hand sides of these equations. Because of the constraint $c_0 = 0$, these revisions are linear functions of only the future innovations $a_{t+1}, a_{t+2}, \ldots$ As long as S_t is nonstationary ($\delta_s(B) \neq 1$), neither $\widehat{S}_{t|t}$ nor Z_t can be expressed as a linear function of just the a_t; one must account for the effects of starting values (Bell 1984).

Returning to calculation of the asymmetric filters, considering (22) for $m+1$ (or (25) for $m' = m+1$) gives

$$\begin{aligned} \alpha_s^{(m+1)}(B) &= \alpha_s^{(0)}(B) + \frac{\sigma_b^2}{\sigma^2}\pi(B)(\eta_1 F + \cdots + \eta_{m+1}F^{m+1}) \\ &= \alpha_s^{(m)}(B) + \frac{\sigma_b^2}{\sigma^2}\pi(B)\eta_{m+1}F^{m+1}. \end{aligned} \tag{27}$$

We write (27) more explicitly as

$$\sum_{j=-(m+1)}^{\infty} \alpha_{sj}^{(m+1)}B^j = \sum_{j=-m}^{\infty} \alpha_{sj}^{(m)}B^j - \frac{\sigma_b^2}{\sigma^2}\eta_{m+1} \sum_{j=-(m+1)}^{\infty} \pi_{j+m+1}B^j \tag{28}$$

where we define

$$\pi_0 = -1.$$

Equation (28) shows the relation between the filter weights at $m+1$ and m:

$$\alpha_{sj}^{(m+1)} = \begin{cases} \frac{\sigma_b^2}{\sigma^2}\eta_{m+1} & j = -(m+1) \\ \alpha_{sj}^{(m)} - \frac{\sigma_b^2}{\sigma^2}\eta_{m+1}\pi_{j+m+1} & j = -m, -m+1, \ldots \end{cases} \tag{29}$$

Once $\alpha_s^{(0)}(B)$ has been computed as shown in Section 2, and $\eta(F) = c(F)/\theta(F)$ has also been computed, (29) can be used to compute the filter weights for $\alpha_s^{(m)}(B)$ for $m = 1, 2, \ldots$

We now obtain a similar expression for the filter weights for the case of $m < 0$ (for prediction of future S_t using data through Z_{t+m}). We relate the signal extraction estimates $\widehat{S}_{t|t+m} = \alpha_s^{(m)}(B)Z_t$ and $\widehat{S}_{t|t+m-1} = \alpha_s^{(m-1)}(B)Z_t$. In general, to get $\widehat{S}_{t|t+m-1}$ we can apply $\alpha_s^{(m)}(B)$ to the series $\widehat{Z}_{t+m|t+m-1}, Z_{t+m-1}, Z_{t+m-2}, \ldots$, i.e., to the observed series up to Z_{t+m-1} with the optimal one-step-ahead forecast $\widehat{Z}_{t+m|t+m-1} = \sum_{j=1}^{\infty} \pi_j Z_{t+m-j}$ of Z_{t+m} appended to it. This produces the following:

$$\begin{aligned} \widehat{S}_{t|t+m-1} &= \alpha_{s,-m}^{(m)} \widehat{Z}_{t+m|t+m-1} + \sum_{j=-m+1}^{\infty} \alpha_{sj}^{(m)} Z_{t-j} \\ &= \alpha_{s,-m}^{(m)} \sum_{j=1}^{\infty} \pi_j Z_{t+m-j} + \sum_{j=-m+1}^{\infty} \alpha_{sj}^{(m)} Z_{t-j} \\ &= \sum_{j=-m+1}^{\infty} \left(\alpha_{sj}^{(m)} + \alpha_{s,-m}^{(m)} \pi_{j+m} \right) Z_{t-j}. \end{aligned}$$

Writing

$$\widehat{S}_{t|t+m-1} = \alpha_s^{(m-1)}(B) Z_t \equiv \sum_{j=-m+1}^{\infty} \alpha_{sj}^{(m-1)} Z_{t-j}$$

we see that

$$\alpha_{sj}^{(m-1)} = \alpha_{sj}^{(m)} + \alpha_{s,-m}^{(m)} \pi_{j+m} \qquad j = -m+1, -m+2, \ldots \tag{30}$$

Given the concurrent filter $\alpha_s^{(0)}(B)$, the relation (30) can be used to compute the weights for the asymmetric filters $\alpha_s^{(m-1)}(B)$ for $m = 0, -1, \ldots$

Note that the second part of (29) can be written

$$\alpha_{sj}^{(m)} = \alpha_{sj}^{(m+1)} + \alpha_{s,-(m+1)}^{(m+1)} \pi_{j+m+1} \qquad j = -m, -m+1, \ldots$$

Comparing with (30), we see that these are really the same equations. Thus, the only difference between the cases $m > 0$ and $m < 0$ is that when $m > 0$ we need to compute $\alpha_{s,-(m+1)}^{(m+1)} = \frac{\sigma_b^2}{\sigma^2} \eta_{m+1}$ to start the calculations.

3.3 Relation between signal extraction filters for S_t and N_t

In the case where Z_t is observed ($m \geq 0$), $\alpha_n^{(m)}(B) = 1 - \alpha_s^{(m)}(B)$. For $m < 0$ (Z_t is in the future) let $\widehat{S}_{t|t+m} = \alpha_s^{(m)}(B) Z_t$ and $\widehat{N}_{t|t+m} = \alpha_n^{(m)}(B) Z_t$ be the signal extraction estimates of S_t and N_t based on $Z_{t+m}, Z_{t+m-1}, \ldots$ For clarity let $\ell = -m$ (note $\ell > 0$) and note that $\widehat{S}_{t|t-\ell}$ and $\widehat{N}_{t|t-\ell}$ satisfy

$$\widehat{Z}_{t|t-\ell} = \widehat{S}_{t|t-\ell} + \widehat{N}_{t|t-\ell}$$

where $\widehat{Z}_{t|t-\ell}$ is the optimal (minimum MSE) ℓ-step-ahead forecast of Z_t from time $t - \ell$. We can write $\widehat{Z}_{t|t-\ell} = \pi^{(\ell)}(B) Z_t$ where $\pi^{(\ell)}(B) = \sum_{j=1}^{\infty} \pi_j^{(\ell)} B^{j+\ell-1}$ and the $\pi_j^{(\ell)}$ are the ℓ-step-ahead forecast weights. Box and Jenkins (1970, pp. 160-162) discuss computation of the $\pi_j^{(\ell)}$. (Note that $\pi(B) = 1 - \pi^{(1)}(B)$ where $\pi_j^{(1)} = \pi_j$.) From these results we see that for $\ell > 0$

$$\pi^{(\ell)}(B) = \alpha_s^{(-\ell)}(B) + \alpha_n^{(-\ell)}(B) \tag{31}$$

so that, in terms of $m = -\ell$, $\alpha_{nj}^{(m)} = \pi_j^{(-m)} - \alpha_{sj}^{(m)}$ and vice-versa. If we interpret $\pi^{(-m)}(B)$ as 1 for $m \geq 0$, then (31) holds for all m.

4 Unit Root Properties of the Asymmetric Filters

Recall that the differencing or, more generally, nonstationary AR operators for S_t and N_t are denoted by $\delta_s(B)$ and $\delta_n(B)$, and that these are contained in $\varphi_s(B)$ and $\varphi_n(B)$. From (20) and (21) we see that $\alpha_s^{(m)}(B)$ contains $\varphi_n(B)$ and hence contains $\delta_n(B)$, while $\alpha_n^{(m)}(B)$ contains $\varphi_s(B)$ and hence contains $\delta_s(B)$. Pierce (1979, pp. 1312-1313) established this result for general difference stationary models.

The fact that $\alpha_s^{(m)}(B)$ contains $\delta_n(B)$ and $\alpha_n^{(m)}(B)$ contains $\delta_s(B)$ (for any m) has important implications. For example, in model-based seasonal adjustment of a monthly series typically $\delta_s(B) = U(B) \equiv 1 + B + \cdots + B^{11}$ and $\delta_n(B) = (1-B)^d$ for some $d > 0$ (note Burman 1980, Hillmer and Tiao 1982, Gersch and Kitagawa 1983, Harvey 1989). It then follows that

- $\alpha_s^{(m)}(B)$ annihilates polynomials up to degree $d-1$ (because $(1-B)^d$ does).
- $\alpha_n^{(m)}(B)$ annihilates fixed seasonal effects (because $U(B)$ does). Fixed seasonal effects are defined as any deterministic sequence ξ_t such that $U(B)\xi_t = 0$. This includes sums of trigonometric functions at the seasonal frequencies ($2\pi j/12$ for $j = 1, \ldots, 6$), and also fixed effects defined from monthly indicator variables with the average effect over 12 consecutive months subtracted off. (See Bell 1984b.)

For $m \geq 0$ these results and the expression $1 = \alpha_s^{(m)}(B) + \alpha_n^{(m)}(B)$ imply that $\alpha_n^{(m)}(B)$ reproduces polynomials up to degree $d-1$ and $\alpha_s^{(m)}(B)$ reproduces fixed seasonal effects. In fact, from (31) we see that these results also hold for $m < 0$ because $\pi^{(-m)}(B)$ (i.e., $\ell = -m$ step-ahead forecasting) can be shown to reproduce both polynomials up to degree $d-1$ and fixed seasonal effects.

The preceding results can be contrasted with those for the corresponding symmetric signal extraction filters for S_t and N_t. The symmetric filters are (Hillmer, Bell, and Tiao 1993, p. 75)

$$\alpha_s^{(\infty)}(B) = \frac{\sigma_b^2}{\sigma^2} \times \frac{\varphi_n(B)\varphi_n(F)\theta_s(B)\theta_s(F)}{\theta(B)\theta(F)} \tag{32}$$

$$\alpha_n^{(\infty)}(B) = \frac{\sigma_e^2}{\sigma^2} \times \frac{\varphi_s(B)\varphi_s(F)\theta_n(B)\theta_n(F)}{\theta(B)\theta(F)}. \tag{33}$$

For model-based seasonal adjustment we see from (32) and (33) that $\alpha_s^{(\infty)}(B)$ contains $\delta_n(B)\delta_n(F) = (1-B)^d(1-F)^d = (1-B)^{2d}(-F)^d$ and $\alpha_n^{(\infty)}(B)$ contains $\delta_s(B)\delta_s(F) = U(B)U(F) = U(B)^2F^{11}$. So $\alpha_s^{(\infty)}(B)$ annihilates (and $\alpha_n^{(\infty)}(B)$ reproduces) polynomials up to degree $2d-1$, and $\alpha_n^{(\infty)}(B)$ annihilates (and $\alpha_s^{(\infty)}(B)$ reproduces) not just fixed seasonal effects but some deterministic effect ζ_t that requires application of $U(B)$ twice to be removed. Thus, the symmetric filters reproduce functions of higher order than the asymmetric filters.

Unit root results for asymmetric trend estimation filters are now obvious. Thus, if S_t is a trend component requiring differencing by $(1-B)^d$ and N_t is a stationary noise component, then the trend estimation filter $\alpha_s^{(m)}(B)$ reproduces, and the trend removal filter $\alpha_n^{(m)}(B)$ annihilates, polynomials up to degree $d-1$. The corresponding symmetric trend estimation filter and symmetric trend removal filters reproduce and annihilate, respectively, polynomials up to degree $2d-1$.

5 Asymmetric Signal Extraction Mean Squared Error (MSE)

We shall obtain the MSE of the asymmetric signal extraction estimate in several alternative ways. Let

$$\epsilon_{t|t+m} = S_t - \widehat{S}_{t|t+m}$$

be the error in the estimate of S_t using data through time $t+m$. For $m \geq 0$ we can write

$$\begin{aligned} \epsilon_{t|t+m} &= S_t - \alpha_s^{(m)}(B)[S_t + N_t] \\ &= [1-\alpha_s^{(m)}(B)]S_t - \alpha_s^{(m)}(B)N_t \\ &= \alpha_n^{(m)}(B)S_t - \alpha_s^{(m)}(B)N_t \end{aligned} \tag{34}$$

and from (20) and (21) this is

$$\begin{aligned} \epsilon_{t|t+m} &= \frac{\sigma_e^2}{\sigma^2} \times \frac{\varphi_s(B)d_n^{(m)}(B)}{\theta(B)} F^m S_t - \frac{\sigma_b^2}{\sigma^2} \times \frac{\varphi_n(B)d_s^{(m)}(B)}{\theta(B)} F^m N_t \\ &= \frac{\sigma_e^2}{\sigma^2} \times \frac{d_n^{(m)}(B)\theta_s(B)}{\theta(B)} b_{t+m} - \frac{\sigma_b^2}{\sigma^2} \times \frac{d_s^{(m)}(B)\theta_n(B)}{\theta(B)} e_{t+m} \end{aligned} \tag{35}$$

From (35) we can see that (for fixed m) $\epsilon_{t|t+m}$ is stationary with ACGF

$$\gamma_{\epsilon,m}(B) = \frac{\sigma_e^2\sigma_b^2}{\sigma^4\theta(B)\theta(F)} \left[\sigma_e^2 d_n^{(m)}(B)d_n^{(m)}(F)\theta_s(B)\theta_s(F) + \sigma_b^2 d_s^{(m)}(B)d_s^{(m)}(F)\theta_n(B)\theta_n(F)\right]. \tag{36}$$

Each of the two parts of (36) can be computed using Tunnicliffe-Wilson's algorithm (Burman 1980). Alternatively, variances and autocovariance for the first part can be computed by applying standard results on computing ARMA model autocovariances (McLeod 1975,1977; Wilson 1979) to the "pseudo-model"

$$\theta(B)u_t = \left\{ \left[d_{n0}^{(m)} \right]^{-1} d_n^{(m)}(B) \right\} \theta_s(B)\xi_t \tag{37}$$

where ξ_t is white noise with variance $[d_{n0}^{(m)}]^2 \sigma_e^4 \sigma_b^2 / \sigma^4$, and u_t is a place holder for any time series following the model (37). (Note that the constant term, $d_{n0}^{(m)}$, in the $d_n^{(m)}(B)$ MA operator in (37) is not 1, so $d_{n0}^{(m)}$ is factored out of $d_n^{(m)}(B)$ and $\text{Var}(\xi_t)$ includes $[d_{n0}^{(m)}]^2$ to compensate.) Autocovariances for the second part are computed the same way. The analogous derivation for the error $\epsilon_{t|\infty}$ in the symmetric signal extraction estimate yields a result analogous to (36), which simplifies to (Bell 1984)

$$\gamma_{\epsilon,\infty}(B) = \frac{\sigma_e^2 \sigma_b^2 \theta_s(B)\theta_s(F)\theta_n(B)\theta_n(F)}{\sigma^2 \theta(B)\theta(F)} \ . \tag{38}$$

Pierce (1979), working with general difference stationary models, showed that the two terms in (34) are stationary by showing that $\alpha_n^{(m)}(B)$ and $\alpha_s^{(m)}(B)$ contain $\varphi_s(B)$ and $\varphi_n(B)$, respectively. He then directly obtained an expression for the spectral density of the signal extraction error in terms of his general expressions for $\alpha_n^{(m)}(B)$ and $\alpha_s^{(m)}(B)$. The result (36) shows how the results for ARIMA component models simplify to rational polynomial expressions that are easily evaluated for given m. These simplifications result from using the decomposition of the term in $\alpha_s^{(m)}(B)$ within the $[\bullet]_+$ notation into the two parts shown in (17), and similarly for $\alpha_n^{(m)}(B)$.

Another approach to computing the MSE, which works for all m, starts by writing

$$\begin{aligned} \epsilon_{t|t+m} &= (S_t - \widehat{S}_{t|\infty}) + (\widehat{S}_{t|\infty} - \widehat{S}_{t|t+m}) && (39) \\ &= \epsilon_{t|\infty} + \left[\alpha_s^{(\infty)}(B) - \alpha_s^{(m)}(B) \right] Z_t. && (40) \end{aligned}$$

Now from (15) and (32)

$$\begin{aligned} \alpha_s^{(\infty)}(B) - \alpha_s^{(m)}(B) &= \frac{1}{\sigma^2}\pi(B)\pi(F)\gamma_s(B) - \frac{F^m}{\sigma^2}\pi(B)\left[\pi(F)\gamma_s(B)B^m\right]_+ \\ &= \frac{F^m}{\sigma^2}\pi(B)\left[\pi(F)\gamma_s(B)B^m\right]_- && (41) \end{aligned}$$

where $[\bullet]_-$ retains only those terms involving negative powers of B (positive powers of $F = B^{-1}$). From (15)-(17) we can write (41) as

$$\alpha_s^{(\infty)}(B) - \alpha_s^{(m)}(B) = F^m \frac{\sigma_b^2}{\sigma^2}\pi(B)\frac{c_s^{(m)}(F)}{\theta(F)}$$

from which it follows that (40) can be written as

$$\epsilon_{t|t+m} = \epsilon_{t|\infty} + \frac{\sigma_b^2}{\sigma^2} \times \frac{c_s^{(m)}(F)}{\theta(F)} a_{t+m} \ . \tag{42}$$

The error $\epsilon_{t|\infty}$ in the optimal symmetric signal extraction estimate is uncorrelated with a_{t+m} for all t and m, so the ACGF of $\epsilon_{t|t+m}$ is the sum of the ACGFs of the two parts, that is

$$\gamma_{\epsilon,m}(B) = \gamma_{\epsilon,\infty}(B) + \frac{\sigma_b^4 c_s^{(m)}(B) c_s^{(m)}(F)}{\sigma^2 \theta(B)\theta(F)} \tag{43}$$

where $\gamma_{\epsilon,\infty}(B)$ is given in (38). As above, the two terms in (43) can be evaluated by computing autocovariances for appropriate pseudo-models.

As an aside we note that the above derivation shows that $\widehat{S}_{t|t+m} = \alpha_s^{(m)}(B) Z_t$, with $\alpha_s^{(m)}(B)$ defined by (20), actually is the optimal signal extraction estimate of S_t based on the data $(\ldots, Z_{t+m-1}, Z_{t+m})$. This result is well-established in the stationary case, as noted earlier, and Hannan (1967) and Sobel (1967) gave such results for a nonstationary signal observed with stationary noise. However, this result has not previously been explicitly demonstrated in the more general nonstationary case considered here. The result follows since both terms on the right hand side of (42) are orthogonal to the differenced observed data $(\ldots, w_{t+m-1}, w_{t+m})$ where $w_t = \delta(B) Z_t$. The time series $\epsilon_{t|\infty}$ is orthogonal to the complete series $\{w_t\}$ since it is the error in the optimal symmetric signal extraction estimate (Bell 1984). The term $(\sigma_b^2/\sigma^2)[c_s^{(m)}(F)/\theta(F)]a_{t+m}$ is orthogonal to $(\ldots, w_{t+m-1}, w_{t+m})$ since it is a linear function of the innovations a_{t+j} for $j > m$. Hence $\epsilon_{t|t+m} = S_t - \widehat{S}_{t|t+m}$ is orthogonal to $(\ldots, w_{t+m-1}, w_{t+m})$ which implies that $\widehat{S}_{t|t+m}$ is the optimal estimate. This line of argument also applies with general difference stationary models as considered by Pierce (1979,1980) if we start from the more general expression (15) and use (41).

A final approach to computing asymmetric signal extraction MSE for $m \geq 0$ starts from (26), which using (39) leads to

$$\epsilon_{t|t+m} = \epsilon_{t|\infty} + \frac{\sigma_b^2}{\sigma^2}(\eta_{m+1}F^{m+1} + \eta_{m+2}F^{m+2} + \cdots)a_t \tag{44}$$

where $\eta(F) = c(F)/\theta(F)$ was defined in Section 3.2. Again by virtue of $\epsilon_{t|\infty}$ being uncorrelated with a_j for all j, the ACGF of $\epsilon_{t|t+m}$ is the sum of the ACGFs of the two parts of the right-hand-side of (44):

$$\gamma_{\epsilon,m}(B) = \gamma_{\epsilon,\infty}(B) + \frac{\sigma_b^4}{\sigma^2}\left(\sum_{j=m+1}^{\infty} \eta_j F^j\right)\left(\sum_{j=m+1}^{\infty} \eta_j B^j\right) \tag{45}$$

The signal extraction MSE is thus obviously

$$\mathrm{Var}(\epsilon_{t|t+m}) = \mathrm{Var}(\epsilon_{t|\infty}) + \frac{\sigma_b^4}{\sigma^2}\left(\sum_{j=m+1}^{\infty} \eta_j^2\right) \tag{46}$$

This involves an infinite sum, which would need to be truncated as an approximation, but has the advantage (relative to (36) and (43)) that the only dependence on m is in the limit of the summation. In fact, we see that as m increases to $m+1$ (we get one more observation) the signal extraction variance decreases by $(\sigma_b^4/\sigma^2)\eta_{m+1}^2$, ultimately decreasing to $\mathrm{Var}(\epsilon_{t|\infty})$ in the limit. Starting from (25) leads to the corresponding expression for any $m' > m$

$$\mathrm{Var}(\epsilon_{t|t+m}) = \mathrm{Var}(\epsilon_{t|t+m'}) + \frac{\sigma_b^4}{\sigma^2}\left(\sum_{j=m+1}^{m'} \eta_j^2\right) \tag{47}$$

From (25) we note that $(\sigma_b^4/\sigma^2)\left(\sum_{j=m+1}^{m'} \eta_j^2\right)$ is the variance of the revision from the estimate at time $t+m$ to the estimate at time $t+m'$.

The preceding approach extends to $m < 0$ by using (42) as a starting point. To see this, suppose $M < 0$ is the smallest value of m of interest ($-M$ is the maximum forecast lead of interest) and define

$$\eta^{(M)}(F) \equiv \sum_{j=1}^{\infty} \eta_j^{(M)} F^j = \frac{c_s^{(M)}(F)}{\theta(F)} \ . \tag{48}$$

Then from (42) $\epsilon_{t|t+M} = \epsilon_{t|\infty} + \frac{\sigma_b^2}{\sigma^2}[\eta_1^{(M)} a_{t+M+1} + \eta_2^{(M)} a_{t+M+2} + \cdots]$ and clearly

$$\mathrm{Var}(\epsilon_{t|t+M}) = \mathrm{Var}(\epsilon_{t|\infty}) + \frac{\sigma_b^4}{\sigma^2}\left(\sum_{j=1}^{\infty}\left[\eta_j^{(M)}\right]^2\right).$$

Furthermore, for any $m > M$, $\frac{\sigma_b^2}{\sigma^2}[\eta_1^{(M)} a_{t+M+1} + \cdots + \eta_{m-M}^{(M)} a_{t+m}]$ is the revision from $\widehat{S}_{t|t+M}$ to $\widehat{S}_{t|t+m}$ so that

$$\mathrm{Var}(\epsilon_{t|t+m}) = \mathrm{Var}(\epsilon_{t|\infty}) + \frac{\sigma_b^4}{\sigma^2}\left(\sum_{j=m-M+1}^{\infty}\left[\eta_j^{(M)}\right]^2\right). \tag{49}$$

Also, for $m' > m$

$$\mathrm{Var}(\epsilon_{t|t+m}) = \mathrm{Var}(\epsilon_{t|t+m'}) + \frac{\sigma_b^4}{\sigma^2}\left(\sum_{j=m-M+1}^{m'-M}\left[\eta_j^{(M)}\right]^2\right). \tag{50}$$

Equations (49) and (50) are analogous to (46) and (47) which effectively used (42) with $M = 0$ as a starting point. If MSEs for $m < 0$ are of interest, then after computing the $\eta_j^{(M)}$ from (48) we can use (49) and (50) to compute these MSEs for any $m \geq M$.

An alternative way to calculate the weights η_j that appear in (44)–(47) is to expand $\varphi_n(F)\theta_s(F)\theta_s(B)/\theta(F)\varphi_s(B)$, the term within the $[\bullet]_+$ notation in equation

(6) for the concurrent filter $\alpha(B)$, and then pick out the coefficients of $F, F^2, \ldots$ This would replace the intermediate calculation of $c(F)$ followed by expansion of $c(F)/\theta(F)$. For another alternative, note that the general expression (5) for $\alpha(B)$ can be written

$$\begin{aligned}\alpha(B) &= \pi(B)\left[\frac{1}{\sigma^2}\pi(F)\pi(B)\gamma_s(B)\psi(B)\right]_+ \\ &= \pi(B)\left[\alpha_s^{(\infty)}(B)\psi(B)\right]_+ \end{aligned} \quad (51)$$

where $\alpha_s^{(\infty)}(B) = \sigma^{-2}\pi(F)\pi(B)\gamma_s(B)$ is the general expression for the symmetric signal extraction filter and $\psi(B) = \theta(B)/\varphi(B) = \pi(B)^{-1}$. Equation (51) shows $\eta(F)$ can be obtained by computing the symmetric filter $\alpha_s^{(\infty)}(B)$ and multiplying it by $\psi(B)$ (and taking the terms in powers of F). Pierce (1980, p. 99 and p. 104) and Hillmer (1985, p. 62) do just this and obtain expressions analogous to some of (44)–(47). Pierce obtains results on the MSE of revisions in signal extraction estimates for general difference stationary models, including non-optimal signal extraction estimates. Hillmer derives approximate expressions for signal extraction MSE based on finite data. (The approximation comes from assuming approximately zero covariance between the contribution to error from having no data before the first observation and that from having no data after the last observation.) Letting the time point of the first observation recede to $-\infty$, Hillmer's MSE result becomes exact and agrees with (46).

6 Example: Random Walk Trend Plus Error

Perhaps the simplest "trend plus error" model in common use assumes that $Z_t = T_t + e_t$, where the trend T_t follows the random walk model

$$(1 - B)T_t = b_t.$$

Here b_t is white noise with variance σ_b^2, and e_t is white noise with variance σ_e^2. Applying $(1 - B)$ to Z_t gives

$$(1 - B)Z_t = b_t + (1 - B)e_t. \quad (52)$$

The right hand side of (52) has variance $\sigma_b^2 + 2\sigma_e^2$, lag-1 autocovariance $-\sigma_e^2$, and all other autocovariances zero. It is thus an MA(1) model and Z_t follows the ARIMA(0,1,1) model

$$(1 - B)Z_t = (1 - \theta_1 B)a_t \quad (53)$$

where θ_1 and $\sigma^2 = \mathrm{Var}(a_t)$ are determined to yield the same variance and lag-1 autocovariance for the right hand side of (53), i.e.,

$$\begin{aligned}(1 + \theta_1^2)\sigma^2 &= \sigma_b^2 + 2\sigma_e^2 \\ -\theta_1\sigma^2 &= -\sigma_e^2.\end{aligned}$$

Notice from these two equations that

$$\begin{aligned}\frac{\sigma_b^2}{\sigma^2} &= (1+\theta_1^2) - \frac{2\sigma_e^2}{\sigma^2} \\ &= (1+\theta_1^2) - 2\theta_1 \\ &= (1-\theta_1)^2. \end{aligned} \tag{54}$$

We now show how to calculate the optimal filter $\alpha(B)$ for estimating T_t from $Z_t, Z_{t-1}, \ldots$. Relative to our previous notation, we identify S_t with T_t and N_t with e_t. For simplicity of notation, we omit the superscript $(m = 0)$ and subscript s from $\alpha(B)$ and related quantities ($d(B)$ and $c(F)$). We add this detail in later material as needed.

We now identify

$$\begin{aligned} \varphi(B) &= \varphi_s(B) = 1 - B \qquad & \varphi_n(B) = 1, \\ \theta(B) &= 1 - \theta_1 B, & \theta_s(B) = \theta_n(B) = 1, \\ g(B) &= \varphi_n(F)\theta_s(F)\theta_s(B) = 1, \\ h &= \max(q, p_n + q_s) = \max(1, 0+0) = 1, \\ k &= \max(p_s, q_s) = \max(1, 0) = 1. \end{aligned} \tag{55}$$

Given these identifications, for this example (10) becomes

$$\begin{aligned} g(B) &= 1 \\ &= c_1 F(1-B) + (d_0 + d_1 B)(1 - \theta_1 F) \\ &= (c_1 - d_0\theta_1)F + (d_0 - d_1\theta_1 - c_1) + d_1 B. \end{aligned} \tag{56}$$

The solution for c_1, d_0, d_1 can easily be found to be

$$\begin{aligned} c_1 &= \theta_1/(1-\theta_1), \\ d_0 &= 1/(1-\theta_1), \\ d_1 &= 0. \end{aligned} \tag{57}$$

We could alternatively have obtained this by solving the equations set up as in (13), which for this example are

$$g = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} c_1 + \begin{bmatrix} -\theta_1 & 0 \\ 1 & -\theta_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d_0 \\ d_1 \end{bmatrix} = \begin{bmatrix} 1 & -\theta_1 & 0 \\ -1 & 1 & -\theta_1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ d_0 \\ d_1 \end{bmatrix}.$$

Now from (12), (54), (57), and the identifications in (55), the desired filter $\alpha(B)$ is

$$\alpha(B) = \frac{\sigma_b^2}{\sigma^2} \times \frac{\varphi_n(B) d(B)}{\theta(B)}$$

$$
\begin{aligned}
&= (1-\theta_1)^2 \frac{1[1/(1-\theta_1)]}{1-\theta_1 B} \\
&= \frac{1-\theta_1}{1-\theta_1 B} \\
&= (1-\theta_1)[1+\theta_1 B + \theta_1^2 B^2 + \cdots]. \qquad (58)
\end{aligned}
$$

(This result agrees with an expression given for this example by Pierce (1979, p. 1315) who also gave results for other values of m.) We see that $\widehat{T}_t = \alpha(B) Z_t$ is an exponentially weighted moving average of Z_t (Box and Jenkins 1970). The exponential weighting is obvious, and since the sum of the filter weights is

$$
\sum_{k=0}^{\infty} \alpha_k = (1-\theta_1)[1+\theta_1+\theta_1^2+\cdots] = 1 \qquad (59)
$$

(58) produces an average of the current and past values of Z_t.

The result (59) implies that the trend estimation filter $\alpha(B)$ reproduces constants, which implies that the detrending filter $\alpha_n(B) = 1-\alpha(B)$ annihilates constants. The latter result could also be seen by noting from (21) that $\alpha_n(B)$ contains $\varphi_s(B) = 1-B$.

7 Example: Canonical Seasonal Adjustment with the Airline Model

We now illustrate computation of asymmetric seasonal and nonseasonal signal extraction filters when Z_t follows the quarterly "airline model"

$$
(1-B)(1-B^4)Z_t = (1-\theta B)(1-\Theta B^4)a_t.
$$

For more concrete illustration we shall use the values $\theta = .4$, $\Theta = .8$, $\sigma^2 = 1$. Letting S_t and N_t denote the canonical seasonal and nonseasonal components from thel decomposition approach of Burman (1980) and Tiao and Hillmer (1982), we identify

$$
\begin{aligned}
\varphi(B) &= (1-B)(1-B^4) = (1-B)^2(1+B+B^2+B^3), \\
\varphi_n(B) &= (1-B)^2 = 1-2B+B^2, \qquad \varphi_s(B) = (1+B+B^2+B^3), \\
\theta(B) &= (1-\theta B)(1-\Theta B^4).
\end{aligned}
$$

Furthermore, the moving average operators in the models for S_t and N_t are of the general form $\theta_s(B) = 1-\theta_{s1}B-\theta_{s2}B^2-\theta_{s3}B^3$ and $\theta_n(B) = 1-\theta_{n1}B-\theta_{n2}B^2$, respectively. For the particular parameter values noted we computed the model decomposition using the program SEATS (Maravall and Gomez 1997), getting

$$
\begin{aligned}
\theta_s(B) &= 1-.0464B-.4959B^2-.4578B^3, \qquad \sigma_b^2 = .00482, \qquad (60) \\
\theta_n(B) &= 1-1.3463B+.3788B^2, \qquad \sigma_e^2 = .8506.
\end{aligned}
$$

To set up the calculation of the concurrent filter $\alpha(B)$ for estimating S_t we identify

$$\begin{aligned} h &= \max(q, p_n + q_s) = \max(5, 2+3) = 5 \\ k &= \max(p_s, q_s) = \max(3, 3) = 3 \ . \end{aligned}$$

The vector $g = (g_{-5}, g_{-4}, \ldots, g_3)'$, where prime denotes transpose, is determined via (9) by multiplying out

$$\begin{aligned} g(B) &\equiv \sum_{j=-5}^{3} g_j B^j = \varphi_n(F)\theta_s(F)\theta_s(B) \\ &= (1-F)^2(1-\theta_{s1}F - \theta_{s2}F^2 - \theta_{s3}F^3)(1-\theta_{s1}B - \theta_{s2}B^2 - \theta_{s3}B^3) \end{aligned}$$

yielding for the $\theta_s(B)$ given in (60)

$$g = (-.4578, .4409, .6951, .5758, -2.5079, .5758, .6951, .4409, -.4578)'.$$

Note that since $p_s \leq q_s$ and $q \leq p_n + q_s$, none of the g_j are set to zero. The vector $(c_5, c_4, c_3, c_2, c_1, d_0, d_1, d_2, d_3)'$ is determined by solving (14) with

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} \theta\Theta & 0 & 0 & 0 \\ -\Theta & \theta\Theta & 0 & 0 \\ 0 & -\Theta & \theta\Theta & 0 \\ 0 & 0 & -\Theta & \theta\Theta \\ -\theta & 0 & 0 & -\Theta \\ 1 & -\theta & 0 & 0 \\ 0 & 1 & -\theta & 0 \\ 0 & 0 & 1 & -\theta \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} .32 & 0 & 0 & 0 \\ -.8 & .32 & 0 & 0 \\ 0 & -.8 & .32 & 0 \\ 0 & 0 & -.8 & .32 \\ -.4 & 0 & 0 & -.8 \\ 1 & -.4 & 0 & 0 \\ 0 & 1 & -.4 & 0 \\ 0 & 0 & 1 & -.4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The resulting $(c_5, c_4, c_3, c_2, c_1, d_0, d_1, d_2, d_3)'$ is

$$(-7.2827, 18.2929, 1.7480, -1.6132, -12.7708, 21.3279, 20.2905, 13.0286, -.4578)'.$$

Given $d(B) = d_0 + d_1 B + d_2 B^2 + d_3 B^3$ we can expand (12) to compute $\alpha(B)$. The first eleven filter weights $(\alpha_0, \alpha_1, \ldots, \alpha_{10})$ are

$$(.1028, -.0667, -.0567, -.0527, .1284, -.0371, -.0388, -.0395, .1037, -.0293, -.0309).$$

Figure 1, discussed shortly, shows the first 60 weights of this filter and, along with Figure 2, filter weights for other values of m.

The concurrent signal extraction filter for estimating N_t can be simply computed as $1-\alpha(B)$. As an alternative, and as a check on the above calculations, we computed this filter directly using the same computer program that produced the above results, but with the roles of S_t and N_t reversed. Denote this filter $\widetilde{\alpha}(B)$ $(= \alpha_n^{(0)}(B))$, where the tilde here and below denotes quantities analogous to those above but obtained with the roles of S_t and N_t reversed. In this case (12) becomes

$$\widetilde{\alpha}(B) = \frac{\sigma_e^2}{\sigma^2} \times \frac{\varphi_s(B)\widetilde{d}(B)}{\theta(B)} \tag{61}$$

where $\widetilde{d}(B)$ is obtained by solving the analogue to (14) for the filter for N_t. For this we identify

$$\begin{aligned} \widetilde{h} &= \max(q, p_s + q_n) = \max(5, 3+2) = 5 \\ \widetilde{k} &= \max(p_n, q_n) = \max(2, 2) = 2 \end{aligned}$$

and

$$\begin{aligned} \widetilde{g}(B) &\equiv \sum_{j=-5}^{2} \widetilde{g}_j B^j = \varphi_s(F)\theta_n(F)\theta_n(B) \\ &= (1 + F + F^2 + F^3)(1 - \theta_{n1}F - \theta_{n2}F^2)(1 - \theta_{n1}B - \theta_{n2}B^2) \end{aligned}$$

which gives

$$\widetilde{g} = (.3788, -1.4774, 1.4785, -.3777, -.3777, 1.4785, -1.4774, .3788)'.$$

Analogous to before, since $p_n \leq q_n$ and $q \leq p_s + q_n$, none of the $\widetilde{g}_j$ are set to zero. The vector $(\widetilde{c}_5, \widetilde{c}_4, \widetilde{c}_3, \widetilde{c}_2, \widetilde{c}_1, \widetilde{d}_0, \widetilde{d}_1, \widetilde{d}_2, \widetilde{d}_3)'$ is determined by solving (14) with

$$\widetilde{A}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \widetilde{A}_2 = \begin{bmatrix} \theta\Theta & 0 & 0 \\ -\Theta & \theta\Theta & 0 \\ 0 & -\Theta & \theta\Theta \\ 0 & 0 & -\Theta \\ -\theta & 0 & 0 \\ 1 & -\theta & 0 \\ 0 & 1 & -\theta \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} .32 & 0 & 0 \\ -.8 & .32 & 0 \\ 0 & -.8 & .32 \\ 0 & 0 & -.8 \\ -.4 & 0 & 0 \\ 1 & -.4 & 0 \\ 0 & 1 & -.4 \\ 0 & 0 & 1 \end{bmatrix}$$

Note that $\widetilde{A}_2$ has the same entries as A_2 but it has one less row and column. Carrying through the computations, the result for $(\widetilde{c}_5, \widetilde{c}_4, \widetilde{c}_3, \widetilde{c}_2, \widetilde{c}_1, \widetilde{d}_0, \widetilde{d}_1, \widetilde{d}_2, \widetilde{d}_3)'$ is

$$(.04126, -.10365, -.00990, .00914, .07236, 1.05474, -1.39825, .37878)'$$

and when (61) is expanded using this $\widetilde{d}(B)$, the resulting $\widetilde{\alpha}(B)$ checks with $1 - \alpha(B)$ to within rounding error.

Figure 1 ($m = 0, -1, -2, -4$) and Figure 2 ($m = 1, 2, 3, 4$) show the canonical seasonal filter weights $\alpha_{sk}^{(m)}$ obtained from (20) for $k = m, \ldots, 60$, for various values of m for this case where $\theta = .4$ and $\Theta = .8$. The seasonal pattern of the filter weights is evident, with positive weights occurring every 4 lags and compensating negative weights occurring at lags in between. Also evident is the exponential decay of the weights over years, the rate of the decay from year-to-year being governed by $\Theta = .8$. One other result worth noting is that for the concurrent ($m = 0$) filter the largest weight occurs not at lag 0 (for the current observation) but rather at lag 4 (for the observation one year ago). This is something of an artifact that occurs for the seasonal filter. If we looked at the corresponding nonseasonal filter $1 - \alpha(B)$ (the concurrent seasonal adjustment filter), the largest weight by far would indeed be on the current observation. Also, for all the positive values of m shown the largest weight in both the seasonal and nonseasonal filters occurs at lag 0.

Figure 3 contains four graphs showing the signal extraction MSE for estimating the airline model canonical seasonal component for values of m from -12 to $+12$. Each of the four graphs show MSEs for one of four sets of values of the airline model parameters (θ, Θ): $(.4, .4)$, $(.4, .8)$, $(.8, .4)$, and $(.8, .8)$. For all cases $\sigma^2 = 1$. To compute the MSEs we used equation (38) to compute $\gamma_{\epsilon,\infty}(B)$ and the approach of Section 3.1 to compute the polynomials $c_s^{(m)}(F)$, then expanded both terms on the right hand side of (43) and added them together to get $\gamma_{\epsilon,m}(B)$. The MSE is the coefficient of B^0 in $\gamma_{\epsilon,m}(B)$. Figure 3 shows how the MSEs decrease as m increases (more observed data lowers MSE) for the different sets of parameters. In all four cases there is a seasonal pattern with the largest decreases in MSE occurring when adding another observation for the same quarter as the one of interest (e.g., adding another first quarter observation for estimating the first quarter seasonal in some year). The overall magnitude of the MSEs is considerably smaller when $\Theta = .8$ than when $\Theta = .4$ — note the differences in the vertical scales of the graphs on the left versus those on the right. This occurs because as Θ increases the canonical seasonal innovations variance decreases, which leads to lower signal extraction MSE. (That increasing Θ decreases the canonical seasonal innovations variance follows by extending a result of Hillmer and Tiao (1982, p. 67) to show that the airline model canonical seasonal ACGF depends on Θ only through the multiplicative factor $(1 - \Theta)^2$.) For a given value of Θ the MSEs are larger when $\theta = .8$ than when $\theta = .4$, though they also decrease faster with increasing m when $\theta = .8$ (though this may appear so partly because they are decreasing from larger values).

8 Appendix A: Proof that the matrix $[A_1|A_2]$ of (14) is nonsingular

First we note that from (13), the matrix $[A_1|A_2]$ of (14) can generally be partitioned as shown below. The row and column dimensions of the partitioned blocks are shown around the margins of the array.

$$[A_1|A_2] = \begin{array}{c} \left[\begin{array}{ccc} G_{11} & 0 & 0 \\ G_{21} & G_{22} & G_{23} \\ 0 & 0 & G_{33} \end{array}\right] \begin{array}{c} h-q \\ q+p_s \\ k+1-p_s \end{array} \\ \begin{array}{ccc} h-q & q+p_s & k+1-p_s \end{array} \end{array} \tag{62}$$

(In (62) and following we let 0 denote a matrix or vector of appropriate dimensions.) If G_{11}, G_{22}, and G_{33} are nonsingular then $[A_1|A_2]$ in (62) is nonsingular. First note from (13) that

$$G_{11} = \begin{bmatrix} 1 & & & & \\ -\varphi_{s1} & 1 & & & \\ -\varphi_{s2} & -\varphi_{s1} & \ddots & & \\ \vdots & \vdots & \ddots & 1 & \\ \vdots & \vdots & \cdots & -\varphi_{s1} & 1 \end{bmatrix} \qquad (h-q) \times (h-q),$$

is a lower triangular matrix with ones on the diagonal, so it is nonsingular. Similarly,

$$G_{33} = \begin{bmatrix} 1 & -\theta_1 & -\theta_2 & \cdots & \vdots \\ & 1 & -\theta_1 & & \vdots \\ & & 1 & \ddots & \vdots \\ & & & \ddots & -\theta_1 \\ & & & & 1 \end{bmatrix} \qquad (k+1-p_s) \times (k+1-p_s),$$

is an upper triangular matrix with ones on the diagonal, so it is also nonsingular. Furthermore,

$$G_{22} = \left[\begin{array}{cccc|cccc} 1 & & & & -\theta_q & & & \\ -\varphi_{s1} & 1 & & & \vdots & -\theta_q & & \\ \vdots & -\varphi_{s1} & \ddots & & \vdots & \vdots & \ddots & \\ \vdots & \vdots & \ddots & 1 & -\theta_1 & \vdots & & -\theta_q \\ -\varphi_{s,p_s} & \vdots & & -\varphi_{s1} & 1 & -\theta_1 & & \vdots \\ & -\varphi_{s,p_s} & & \vdots & & 1 & \ddots & \vdots \\ & & \ddots & \vdots & & & \ddots & -\theta_1 \\ & & & -\varphi_{s,p_s} & & & & 1 \end{array}\right] \begin{array}{c} q+p_s \\ \text{rows} \end{array} \tag{63}$$

$$\qquad q \text{ columns} \qquad\qquad\qquad p_s \text{ columns}$$

Below we state and prove a Lemma that we can use to prove that G_{22} is nonsingular.

Before proceeding to the Lemma we clarify some special cases. First, if $p_s = 0$ then $G_{11} = I_{h-q}$, the $h-q$ dimensional identity matrix. Also the right half of G_{22} in (63) is omitted and in fact $G_{22} = I_q$ which is nonsingular. If $q = 0$ then $G_{33} = I_{k+1-p_s}$, the left half of G_{22} in (63) is omitted, and $G_{22} = I_{p_s}$ which is nonsingular. Finally, if $h = q$ then the top block row of (62), $[G_{11}|0|0]$, is omitted from $[A_1|A_2]$, and we only need show that G_{22} and G_{33} are nonsingular. When $h = q$ the results just discussed for G_{33} still apply, as does the Lemma below for G_{22}, so this case need not be treated separately.

We state and prove our Lemma in a more general form and using more general notation than actually needed for the application to G_{22} of the particular form given in (63). The notation used here is unique to this Lemma in that we reuse some notational symbols (such as α) that appear earlier in the paper, but here they have a different meaning.

Lemma: Let

$$\begin{aligned} \alpha(x) &= \alpha_0 + \alpha_1 x + \cdots + \alpha_r x^r \\ \beta(x) &= \beta_0 + \beta_1 x + \cdots + \beta_s x^s \end{aligned}$$

be polynomials in x of degrees $r > 0$ and $s > 0$. Assume that $\alpha_0, \alpha_r, \beta_0$, and β_s are all nonzero. Consider the $(r+s) \times (r+s)$ matrix $[C_1|C_2]$ where

$$[C_1|C_2] = \underbrace{\left[\begin{array}{cccc|cccc} \alpha_0 & & & & \beta_0 & & & \\ \alpha_1 & \alpha_0 & & & \beta_1 & \beta_0 & & \\ \vdots & \alpha_1 & \ddots & & \vdots & \beta_1 & \ddots & \\ \vdots & \vdots & \ddots & \alpha_0 & \vdots & \vdots & \ddots & \beta_0 \\ \alpha_r & \vdots & & \alpha_1 & \beta_s & \vdots & & \beta_1 \\ & \alpha_r & & \vdots & & \beta_s & & \vdots \\ & & \ddots & \vdots & & & \ddots & \vdots \\ & & & \alpha_r & & & & \beta_s \end{array}\right]}_{s \text{ columns} \qquad r \text{ columns}} . \quad \begin{array}{c} r+s \\ \text{rows} \end{array} \tag{64}$$

Assume that $\alpha(x)$ and $\beta(x)$ have no common zeros. Then $[C_1|C_2]$ is nonsingular.

Proof: Consider the homogeneous difference equation

$$\alpha(B)\beta(B)z_t = 0. \tag{65}$$

Let the zeros of $\alpha(x)$ be $\xi_1, \ldots, \xi_m$ with multiplicities $\mu_1, \ldots, \mu_m$ where $\mu_1 + \cdots + \mu_m = r$. Let the zeros of $\beta(x)$ be $\zeta_1, \ldots, \zeta_n$ with multiplicities $\nu_1, \ldots, \nu_n$ where

$\nu_1 + \cdots + \nu_n = s$. It is well-known (see Henrici 1974, pp. 584-587) that the space of solutions to (65) has dimension $r+s$ (the order of $\alpha(B)\beta(B)$) and that the sequences (defined for $t \geq 0$)

$$u(t) = t^j \xi_\ell^{-t} \qquad j = 0, \ldots, \mu_\ell - 1 \qquad \ell = 1, \ldots, m \tag{66}$$
$$v(t) = t^j \zeta_\ell^{-t} \qquad j = 0, \ldots, \nu_\ell - 1 \qquad \ell = 1, \ldots, n \tag{67}$$

provide a (linearly independent) basis for this space. Take the r sequences $u_1(t), \ldots, u_r(t)$ given by (66), truncate each to $r+s$ elements, and put each truncated sequence into a vector. Label these vectors $u_1, \ldots, u_r$. Similarly construct vectors $v_1, \ldots, v_s$ corresponding to the first $r+s$ elements of the s sequences $v_1(t), \ldots, v_s(t)$ given by (67). We note two facts about these vectors:

1. The u_i and v_i are constructed so that $C_1' u_i = 0$ for $i = 1, \ldots, r$ and $C_2' v_i = 0$ for $i = 1, \ldots, s$.
2. The $r+s$ vectors $u_1, \ldots, u_r, v_1, \ldots, v_s$ are linearly independent.

We establish the second fact just noted by induction. Suppose that

$$a_1 u_1 + \cdots + a_r u_r + a_{r+1} v_1 + \cdots + a_{r+s} v_s = 0 \tag{68}$$

for some set of coefficients a_i. Then we can show that the same linear combination of the sequences from (66) and (67) is zero for all $t \geq 0$. Equation (68) covers $0 \leq t \leq r+s$. Let $t > 0$ and assume the relation as in (68) holds up through $t-1$. We take the linear combination of $u_1(t), \ldots, v_s(t)$ as in (68) and reexpress it using the difference equation (65). Letting $\delta(B) = 1 - \delta_1 B - \cdots - \delta^{r+s} B^{r+s} = [\alpha_0 \beta_0]^{-1} \alpha(B)\beta(B)$ we have

$$\begin{aligned}
& a_1 u_1(t) + \cdots + a_r u_r(t) + a_{r+1} v_1(t) + \cdots + a_{r+s} v_s(t) \\
= \; & a_1[\delta_1 u_1(t-1) + \cdots + \delta_{r+s} u_1(t-r-s)] + \cdots \\
& \quad + a_{r+s}[\delta_1 v_s(t-1) + \cdots + \delta_{r+s} v_s(t-r-s)] \\
= \; & \delta_1[a_1 u_1(t-1) + \cdots + a_{r+s} v_s(t-1)] + \cdots \\
& \quad + \delta_{r+s}[a_1 u_1(t-r-s) + \cdots + a_{r+s} v_s(t-r-s)] \\
= \; & \delta_1 \cdot 0 + \cdots + \delta_{r+s} \cdot 0 \\
= \; & 0
\end{aligned}$$

where the next to last line above follows from the induction hypothesis. Since the $u_i(t)$ and $v_i(t)$ sequences are linearly independent the a_is must all be zero, which shows that the vectors $u_1, \ldots, u_r, v_1, \ldots, v_s$ are also linearly independent.

Now we define matrices U and V with columns given by the vectors $u_1, \ldots, u_r$, and $v_1, \ldots, v_s$:

$$U = [u_1 | \ldots | u_r] \qquad V = [v_1 | \ldots | v_s].$$

The first fact noted above for the vectors $u_1, \ldots, u_r, v_1, \ldots, v_s$ shows that $C_1'U = 0$ and $C_2'V = 0$. Let $S(C_1)$ denote the linear subspace of R^{r+s} spanned by the columns of C_1, and similarly for $S(C_2)$, $S(U)$, and $S(V)$. Let $S(C_1)^{\perp}$ denote the orthogonal complement of $S(C_1)$ and $S(C_2)^{\perp}$ that of $S(C_2)$. Since C_1 has full rank s, $S(C_1)^{\perp}$ has rank r. Since $u_1, \ldots, u_r$ are all orthogonal to C_1 ($C_1'u_i = 0$) they are all in $S(C_1)^{\perp}$, and since they are linearly independent they span $S(C_1)^{\perp}$, i.e., $S(C_1)^{\perp} = S(U)$. Similarly, $S(C_2)^{\perp} = S(V)$.

Now we can show that $[C_1|C_2]$ has full row rank and thus is nonsingular. Suppose that

$$d'[C_1|C_2] = 0$$

for some $(r+s) \times 1$ vector d. Then $d \in S(C_1)^{\perp} = S(U)$ and also $d \in S(C_2)^{\perp} = S(V)$. This implies that $d = Ub_1$ for some vector b_1 and that also $d = Vb_2$ for some vector b_2. But then

$$0 = d - d = Ub_1 - Vb_2$$

and by the linear independence of the columns of $[U|V]$ this implies that $b_1 = 0$ and $b_2 = 0$, which implies that $d = 0$. Hence, $[C_1|C_2]$ has full row rank and thus is nonsingular. This proves the Lemma.

To complete the proof that the matrix $[A_1|A_2]$ of (14) is nonsingular note that we assume $\varphi_{s,p_s} \neq 0$ and $\theta_q \neq 0$ (otherwise the actual orders of the operators $\varphi_s(B)$ and $\theta(B)$ would be less than p_s and q, respectively). Thus, the matrix $[A_1|A_2]$ is of the form required by the Lemma. Hence, $[A_1|A_2]$ is nonsingular if

$$\begin{aligned} \varphi_s(x) &= 1 - \varphi_{s1}x - \cdots - \varphi_{s,p_s}x^{p_s} \qquad \text{and} \\ \beta(x) &= -\theta_q - \theta_{q-1}x - \cdots - \theta_1 x^{q-1} + x^q \end{aligned}$$

have no common zeros. Note that $\beta(x) = x^q\theta(x^{-1})$. The zeros of $\varphi_s(x)$ are assumed to lie on or outside the unit circle. The zeros of $\theta(x)$ are assumed to lie outside the unit circle, which implies that those of $\beta(x) = x^q\theta(x^{-1})$ lie inside the unit circle. Hence, $\varphi_s(x)$ and $\beta(x)$ have no common zeros and the Lemma establishes that $[A_1|A_2]$ of (14) is nonsingular.

Acknowledgment: We thank David Findley for useful suggestions that improved the presentation.

Disclaimer: This paper reports the results of research and analysis undertaken by Census Bureau staff and staff of Howard University. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

References

[1] Ansley, Craig F. and Kohn, Robert (1985), “Estimation, Filtering, and Smoothing in State Space Models with Incompletely Specified Initial Conditions,” *Annals of Statistics*, **13**, 1286-1316.

[2] Bell, William R. (1984a), “Signal Extraction for Nonstationary Time Series,” *Annals of Statistics*, **12**, 646-664.

[3] Bell, William R. (1984b), “Seasonal Decomposition of Deterministic Effects,” Research Report Number 84/01, Statistical Research Division, Bureau of the Census.

[4] Bell, William R. and Hillmer, Steven C. (1988), “A Matrix Approach to Signal Extraction and Likelihood Evaluation for ARIMA Component Time Series Models,” Research Report Number 88/22, Statistical Research Division, Bureau of the Census.

[5] Box, G.E.P. and Jenkins, G. M. (1970), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden Day.

[6] Burman, J. P. (1980), “Seasonal Adjustment by Signal Extraction,” *Journal of the Royal Statistical Society Series A*, **143**, 321-337.

[7] Cleveland, William P. and Tiao, George C. (1976), “Decomposition of Seasonal Time Series: A Model for the X-11 Program,” *Journal of the American Statistical Association*, **71**, 581-587.

[8] Findley, David F. and Martin, Donald E. K. (2002), “Frequency Domain Analysis of SEATS and X-11/12 ARIMA Seasonal Adjustment Filters for Short and Moderate-Length Time Series,” working paper, Statistical Research Division, U.S. Census Bureau.

[9] Gersch, Will and Kitagawa, Genshiro (1983), “The Prediction of Time Series With Trends and Seasonalities,” *Journal of Business and Economic Statistics*, **1**, 253-264.

[10] Gomez, Victor and Maravall, Agustin (1997), “Programs TRAMO and SEATS: Instructions for the User (Beta Version: June 1997),” Working Paper 97001, Ministerio de Economía y Hacienda, Dirección General de Análisis y Programación Presupuestaria, Madrid, Spain.

[11] Hannan, Edward J. (1967), “Measurement of a Wandering Signal Amid Noise,” *Journal of Applied Probability*, **4**, 90-102.

[12] Hannan, Edward J. (1970), *Multiple Time Series*, New York: Wiley.

[13] Harvey, Andrew C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge, U. K.: Cambridge University Press.

[14] Henrici, P. (1974), *Applied and Computational Complex Analysis, Vol. 1*, New York: Wiley.

[15] Hillmer, Steven C. (1985), "Measures of Variability for Model-Based Seasonal Adjustment Procedures," *Journal of Business and Economic Statistics*, **3**, 60-68.

[16] Hillmer, Steven C., Bell, William R., and Tiao, George C. (1983), "Modeling Considerations in the Seasonal Adjustment of Economic Time Series," in *Applied Time Series Analysis of Economic Data*, ed. Arnold Zellner, U.S. Department of Commerce, Bureau of the Census, 74-100.

[17] Hillmer, Steven C., and Tiao, George C. (1982), "An ARIMA-Model-Based Approach to Seasonal Adjustment," *Journal of the American Statistical Association*, **77**, 63-70.

[18] Kohn, Robert and Ansley, Craig F. (1987), "Signal Extraction for Finite Nonstationary Time Series," *Biometrika*, **74**, 411-421.

[19] Kolmogorov, A. N. (1939), "Sur L interpolation et Extrapolation des Suites Stationnaires," *C. R. Acad. Sci. Paris*, **208**, 2043-2045.

[20] Kolmogorov, A. N. (1941), "Interpolation und Extrapolation von stationären zufälligen Folgen," *Bull. Acad. Sci. U.R.S.S., Ser. Math.*, **5**, 3-14.

[21] Koopman, Siem Jan and Harvey, Andrew C. (2000), "Computing Observation Weights for Signal Extraction and Filtering," working paper, Department of Economics, Free University Amsterdam.

[22] McLeod, I. (1975), "Derivation of the Theoretical Autocovariance Function of Autoregressive-Moving Average Time Series," *Applied Statistics*, **24**, 255-256.

[23] McLeod, I. (1977), "Correction to Derivation of the Theoretical Autocovariance Function of Autoregressive-Moving Average Time Series," *Applied Statistics*, **26**, 194.

[24] Pierce, David A. (1979), "Signal Extraction Error in Nonstationary Time Series," *Annals of Statistics*, **7**, 1303-1320.

[25] Pierce, David A. (1980), "Data Revisions With Moving Average Seasonal Adjustment Procedures," *Journal of Econometrics*, **14**, 95-114.

[26] Sobel, Eugene L. (1967), "Prediction of a Noise-distorted, Multivariate, Nonstationary Signal," *Journal of Applied Probability*, **4**, 330-342.

[27] Wiener, Norbert (1949). *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*, New York: Wiley.

[28] Whittle, P. (1963), *Prediction and Regulation by Linear Least-Square Methods*, Princeton: Van Nostrand.

[29] Wilson, G. Tunnicliffe (1979), "Some Efficient Computational Procedures for High Order ARMA Models," *Journal of Statistical Computation and Simulation*, **8**, 301-309.

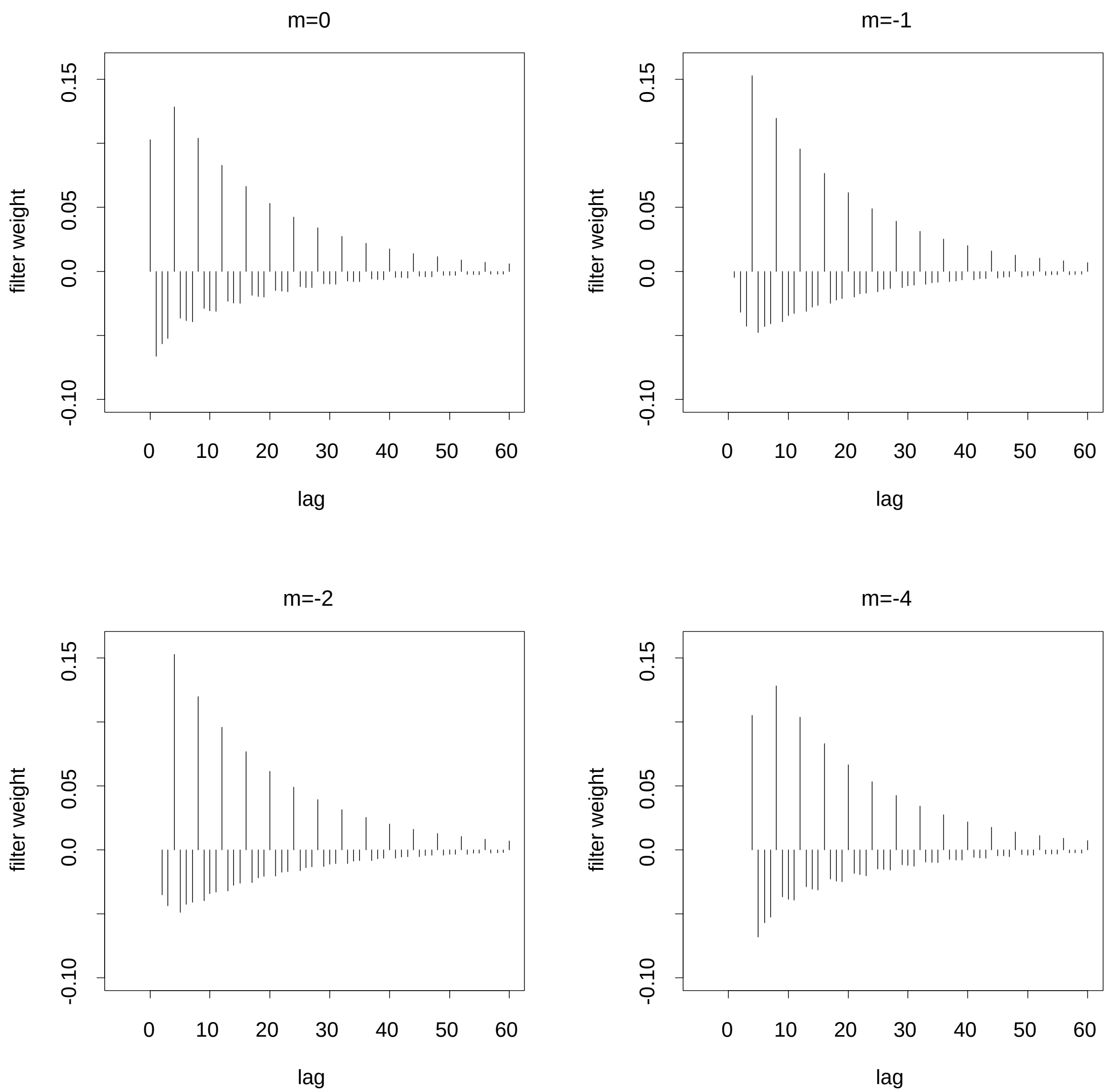

Figure 1. Canonical Seasonal Filter Weights for m = 0, -1, -2, and -4
(airline model with $\theta = .4$, $\Theta = .8$)

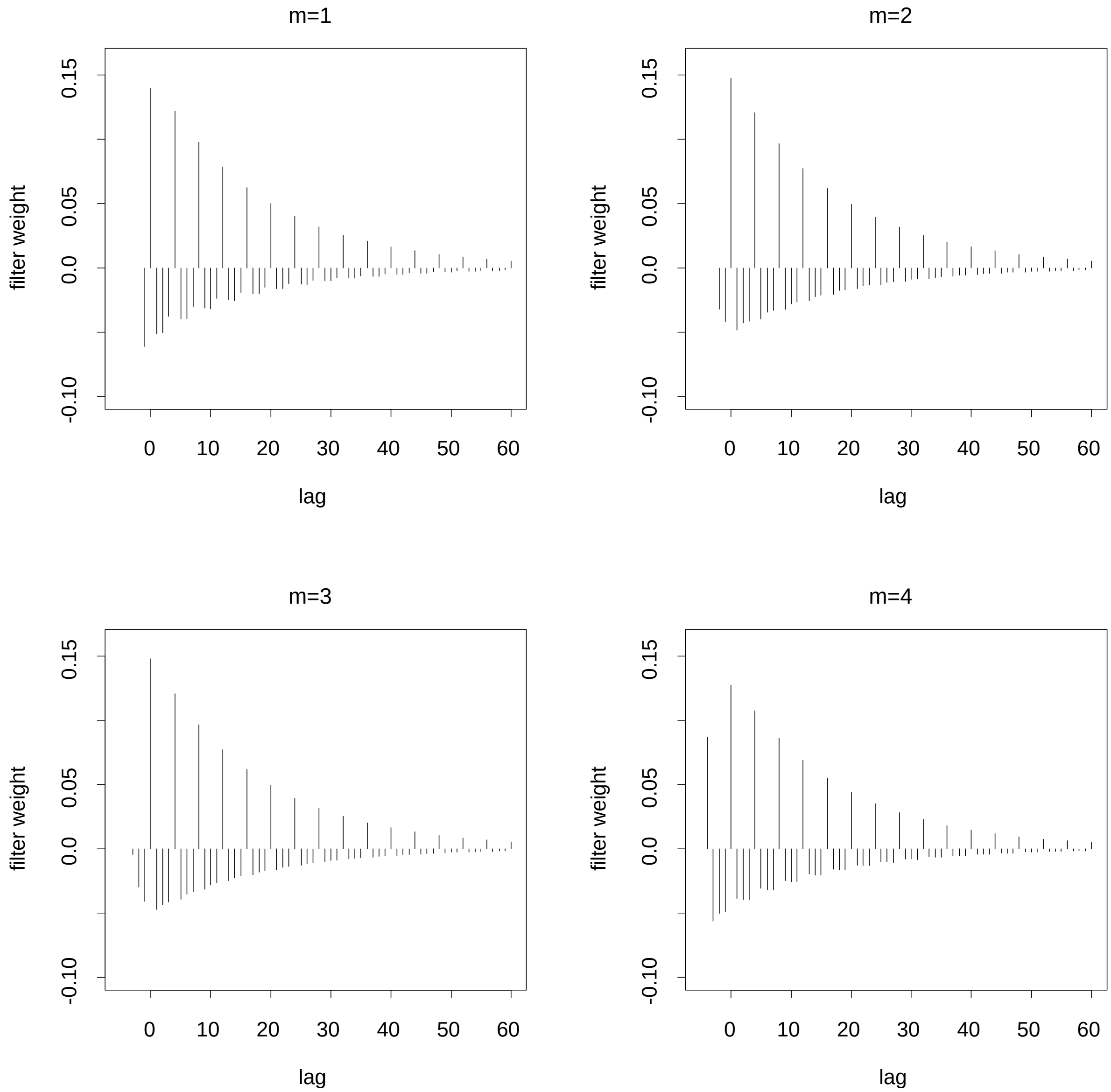

Figure 2. Canonical Seasonal Filter Weights for m = 1, 2, 3, and 4
(airline model with $\theta = .4$, $\Theta = .8$)

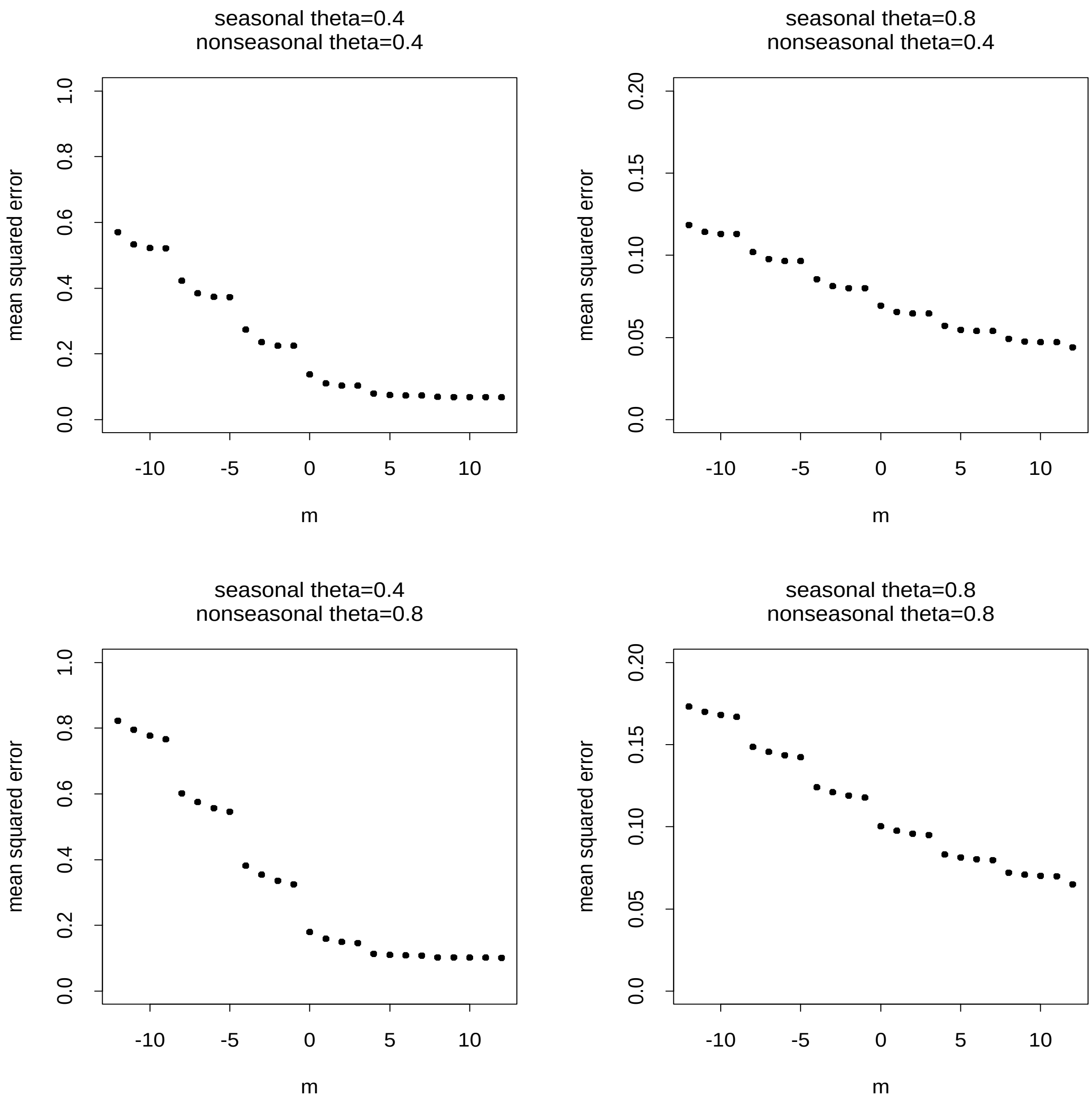

Figure 3. Signal Extraction MSE for Various m and Various Airline Model Parameter Values