

RESEARCH REPORT SERIES  
(*Statistics #2005-04*)

Matrix Formulas for Nonstationary Signal Extraction

Tucker McElroy

Statistical Research Division  
U.S. Census Bureau  
Washington, DC 20233

Report Issued: June 9, 2005

This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed are those of the author and not necessarily those of the U.S. Census Bureau.

# Matrix Formulas for Nonstationary Signal Extraction

Tucker McElroy  
U.S. Census Bureau

## Abstract

The paper provides general matrix formulas for extraction of a nonstationary signal in nonstationary noise, for a finitely sampled time series. These formulas are quite practical, in that they provide a ready intuition for the filtering operation, as well as being simple to implement on a computer. Applications to signal extraction diagnostics are discussed.

**Disclaimer** This report is released to inform interested parties of research and to encourage discussion. The views expressed on statistical issues are those of the author and not necessarily those of the U.S. Census Bureau.

## 1 Introduction

Signal extraction for nonstationary time series data has a long history, including Hannan (1967), Sobel (1967), Cleveland and Tiao (1976), and Bell (1984). Bell and Hillmer (1988) treats the finite sample case, presenting matrix formulas for the optimal time-varying filters, and McElroy and Sutcliffe (2004) provides certain relations between the various component filters. One drawback of Bell and Hillmer's approach is the separate estimation of initial values for nonstationary signals, resulting in formulas that are awkward to implement. McElroy and Sutcliffe (2004) furnishes an improvement, but only for a specific type of unobserved components model. The paper at hand provides general matrix formulas for extraction of a nonstationary signal from nonstationary noise; these formulas are quite practical, in that they produce a ready intuition for the filtering operation, as well as being simple to implement on a computer.

In the context of model-based signal extraction, one popular approach – e.g., utilized in SEATS of Gómez and Maravall (1997) – has been to use the bi-infinite filters of Bell (1984) on forecast and backcast extended data, with the aid of the algorithm of Tunnicliffe-Wilson (see Burman

1980). While being easy to implement, this method confounds any intuition about the time-varying filters. Moreover, it cannot produce correct finite-sample Mean Squared Errors (MSEs) for the signal estimates. Another approach is to formulate the model in State Space Form (Durbin and Koopman, 2001) and construct the appropriate state space smoother (Kailath, Sayed, and Hassibi 2000). Efficient algorithms exist to obtain the time-varying signal extraction filters from the Kalman smoother, if desired (Koopman and Harvey, 2003); of course, these methods do not provide formulas, only numbers, and thus also fail to provide intuition. Neither of the above approaches can provide the full covariance matrix of the signal error, which quantity is useful in diagnostic applications (Findley, McElroy, and Wills 2004). Hence, there is a definite need and appeal for having explicit, readily implemented matrix formulas for nonstationary signal extraction.

This paper first discusses background material and the main theoretical results in Section 2, and provides some applications in Section 3, including the topic of signal extraction diagnostics. Proofs are provided in a final appendix.

## 2 Matrix Formulas

Consider a nonstationary time series  $Y_t$  that can be written as the sum of two possibly nonstationary components  $S_t$  and  $N_t$ , the signal and the noise:

$$Y_t = S_t + N_t \tag{1}$$

Following Bell (1984), we let  $Y_t$  be an integrated process such that  $W_t = \delta(B)Y_t$  is stationary, where  $B$  is the backshift operator and  $\delta(z)$  is a polynomial with all roots located on the unit circle of the complex plane (also,  $\delta(0) = 1$  by convention). This  $\delta(B)$  is the differencing operator of the series, and we assume it can be factored into relatively prime polynomials  $\delta^S(z)$  and  $\delta^N(z)$  (i.e., they share no common zeroes), such that

$$U_t = \delta^S(B)S_t \quad V_t = \delta^N(B)N_t \tag{2}$$

are stationary time series. Note that included as special cases are  $\delta^S = 1$  and/or  $\delta^N = 1$ , in which case either the signal or the noise or both are stationary. We let  $d$  be the order of  $\delta$ , and  $d_S$  and  $d_N$  are the orders of  $\delta^S$  and  $\delta^N$ ; since the latter operators are relatively prime,  $\delta = \delta^S \cdot \delta^N$  and  $d = d_S + d_N$ .

There are many examples from econometrics and engineering that fit into this scheme. The signal could be a trend plus irregular component, in which case  $\delta^S(z)$  could be  $(1 - z)^2$ , and the noise is a seasonal with  $\delta^N(z) = 1 + z + z^2 + \dots + z^{11}$  for monthly data. This is essentially the case explored in McElroy and Sutcliffe (2004). Alternatively, the signal could be a business cycle (typically modelled as stationary with  $\delta^S(z) = 1$ ) and the noise is trend plus irregular. In a third application, the noise could be a stationary sampling error with  $\delta^N(z) = 1$ , and the signal is seasonal plus trend plus irregular.

As in Bell and Hillmer (1988), we assume Assumption A of Bell (1984) holds on the component decomposition, and we treat the case of a finite sample with  $t = 1, 2, \dots, n$ . Assumption A states that the initial  $d$  values of  $Y_t$ , i.e., the variables  $Y_1, Y_2, \dots, Y_d$ , are independent of  $\{U_t\}$  and  $\{V_t\}$ . For a discussion of the implications of this assumption, see Bell (1984) and Bell and Hillmer (1988). Note that mean square optimal signal extraction filters derived under Assumption A agree exactly with the filters implicitly used by the Kalman smoother, see Kohn and Ansley (1986, 1987). A further assumption that we make is that  $\{U_t\}$  and  $\{V_t\}$  are uncorrelated time series.

Now we can write (2) in a matrix form, as follows. Let  $\Delta$  be a  $n - d \times n$  matrix with entries given by  $\Delta_{ij} = \delta_{i-j+d}$  (the convention being that  $\delta_k = 0$  if  $k < 0$  or  $k > d$ ). The matrices  $\Delta_S$  and  $\Delta_N$  have entries given by the coefficients of  $\delta^S(z)$  and  $\delta^N(z)$ , but are  $n - d_S \times n$  and  $n - d_N \times n$  dimensional respectively. This means that each row of these matrices consists of the coefficients of the corresponding differencing polynomial, horizontally shifted in an appropriate fashion. Hence

$$W = \Delta Y \quad U = \Delta_S S \quad V = \Delta_N N$$

where  $Y$  is the transpose of  $(Y_1, Y_2, \dots, Y_n)$ , and  $W, U, V, S,$  and  $N$  are also column vectors. We will denote the transpose of  $Y$  by  $Y'$ . It follows from the equation

$$W_t = \delta^N(B)U_t + \delta^S(B)V_t \tag{3}$$

that we need to define further differencing matrices  $\underline{\Delta}_N$  and  $\underline{\Delta}_S$  with row entries given by the coefficients of  $\delta^N(z)$  and  $\delta^S(z)$  respectively, which are  $n - d \times n - d_S$  and  $n - d \times n - d_N$  dimensional. It then follows from Lemma 1 of McElroy and Sutcliffe (2004) that

$$\Delta = \underline{\Delta}_N \Delta_S = \underline{\Delta}_S \Delta_N. \tag{4}$$

Then we can write down the matrix version of (3):

$$W = \underline{\Delta}_N U + \underline{\Delta}_S V \quad (5)$$

The minimum mean squared error signal extraction estimate is  $\hat{S}_t = \mathbb{E}[S_t|Y]$ , which is expressed as some linear function of the data vector  $Y$  when the data is Gaussian; putting this together for each time  $t$ , we obtain the various rows of a matrix  $F$ :

$$\hat{S} = FY = \mathbb{E}[S|Y].$$

We note that the various rows of  $F$  differ (unlike in the bi-infinite filtering case), since only a finite number of  $Y_t$ 's are available to filter. The last row of  $F$ , for example, corresponds to the concurrent filter, i.e., a one-sided filter used to extract a signal at “time present.”

For any random vector  $X$ , let  $\Sigma_X$  denote its covariance matrix. With these notations in hand, we can now state the signal extraction formulas.

**Theorem 1** *Assume that Assumption A holds on the model decomposition (1), and that  $\{U_t\}$  and  $\{V_t\}$  are independent and purely nondeterministic. Then the minimum mean square error linear estimate of  $S$  is given by  $\hat{S} = FY$ , where*

$$F = \left( \Delta_S' \Sigma_U^{-1} \Delta_S + \Delta_N' \Sigma_V^{-1} \Delta_N \right)^{-1} \Delta_N' \Sigma_V^{-1} \Delta_N. \quad (6)$$

*It also follows from our assumptions that all matrix inverses exist.*

**Remark 1** As demonstrated in the proof, the invertibility of

$$M = \Delta_S' \Sigma_U^{-1} \Delta_S + \Delta_N' \Sigma_V^{-1} \Delta_N \quad (7)$$

depends on  $\delta^S$  and  $\delta^N$  being relatively prime.

**Theorem 2** *Under the same assumptions as Theorem 1, the covariance matrix of  $\hat{S} - S$  is given by  $M^{-1}$ , where  $M$  is defined in (7).*

### 3 Applications

The formula (6) for the signal extraction filter  $F$  is important for providing intuition. One first applies  $\Delta_N$  to the data  $Y$ , which reduces the noise component  $N$  to stationarity; the signal will not

be affected in this way, since no zeroes of  $\delta^S$  are in common with  $\delta^N$ . In particular, (6) establishes that each row of  $F$  is a convolution of a certain filter with the coefficients of the polynomial  $\delta^N(z)$ , and hence that  $\delta^N(e^{-i\lambda})$  is a factor in the transfer function in *every* row of  $F$ . See Findley and Martin (2003) for applications of this observation.

### 3.1 Estimating Component Models

In order to implement these results, it is necessary to estimate the component models for  $S_t$  and  $N_t$  from the data. Together with (2), specifying ARMA models for  $U_t$  and  $V_t$  will fully specify an ARIMA model for  $S_t$  and  $N_t$ . In the structural approach of Harvey (1989), the ARMA model for  $W_t$  is determined by the component models. The Gaussian likelihood for  $W_t$  depends directly on  $\Sigma_W$ , which in turn depends on the ARMA models for  $U_t$  and  $V_t$ . Therefore maximization of the likelihood provides estimates of the ARMA parameters for both  $U_t$  and  $V_t$ . This approach differs from the canonical decomposition technique of Hillmer and Tiao (1982), where an ARMA model for  $W_t$  is specified first; one estimates the ARMA parameters for  $W_t$ , and then one tries to decompose the model for  $W_t$  into component models, using (3). Below we provide some details on these two approaches.

Suppose that ARMA models are specified for  $U_t$  and  $V_t$  such that

$$\phi^U(B)U_t = \theta^U(B)\epsilon_t^U \quad \phi^V(B)V_t = \theta^V(B)\epsilon_t^V,$$

and denote their spectral densities by  $f_U$  and  $f_V$ . Then it follows from (3) that the spectral density for  $W_t$  is

$$\begin{aligned} f_W(\lambda) &= |\delta^N(e^{-i\lambda})|^2 f_U(\lambda) + |\delta^S(e^{-i\lambda})|^2 f_V(\lambda) \\ &= \frac{\sigma_{\epsilon^U}^2 |\delta^N(e^{-i\lambda})|^2 |\theta^U(e^{-i\lambda})|^2 |\phi^V(e^{-i\lambda})|^2 + \sigma_{\epsilon^V}^2 |\delta^S(e^{-i\lambda})|^2 |\theta^V(e^{-i\lambda})|^2 |\phi^U(e^{-i\lambda})|^2}{|\phi^U(e^{-i\lambda})|^2 |\phi^V(e^{-i\lambda})|^2} \end{aligned} \quad (8)$$

If we were to specify a spectral density for  $W_t$  using an ARMA model, without reference to the component models, it would have the form

$$f_W(\lambda) = \sigma_{\epsilon^W}^2 \frac{|\theta^W(e^{-i\lambda})|^2}{|\phi^W(e^{-i\lambda})|^2}. \quad (9)$$

In the structural approach, the additional structure of the specification (8) for the spectral density of  $W_t$  allows for the structure of the component models. In contrast, the canonical decomposition procedure of Hillmer and Tiao (1982) would start by estimating parameters for the model (9),

and hope that the resulting spectral density can be decomposed in the form (8). By imposing the structure of (8) from the outset, we avoid admissibility problems. The expense to be paid is that the structural approach generally involves more parameters, and in practice often has poor goodness of fit properties.

The most efficient method of estimating the structural model (8) is to map the component models into State Space Form, construct the overall state space of  $W_t$ , and utilize a state space approach to maximizing the resulting likelihood. This is described in Durbin and Koopman (2001), and implemented in the SsfPack software package (Koopman, Shephard, and Doornik, 1999). The State Space approach is generally preferred, because the matrix operations typically involve matrices of dimension equal to the size of the state vector, which can be considerably less than the sample size  $n$ . The alternative is to directly compute  $\Sigma_W^{-1}$ , which involves order  $n^2$  operations – Golub and Van Loan (1996). In comparison, the State Space approach to maximum likelihood estimation is computationally efficient.

### 3.2 Signal Extraction Diagnostics

Findley, McElroy, and Wills (2004) describe model-based signal extraction diagnostics, based on the original diagnostic for “underestimation” and “overestimation” of the components of a seasonal time series, which is described in Maravall (2003). The diagnostic of Findley et al. (2004) is based on the sample second moment of an estimated differenced signal of interest, i.e., the sample variance of  $\hat{U} = \Delta_S \hat{S}$ . In the original approach of Maravall (2003),  $S$  was taken to be the seasonal component, in which case  $\delta_S(z) = 1 + z + z^2 + \dots + z^{11}$ . However, Findley et al. (2004) focuses on  $S$  being a stationary irregular component, in which case  $\Delta_S$  is the identity matrix. The expected value of the sample variance is given by  $trace(\Sigma_{\hat{U}})/n$ , which utilizes (10) below; the formula for the standard error of the signal extraction diagnostic requires all of the entries of  $\Sigma_{\hat{U}}$ . Below we provide the formula for  $\hat{U}$  and  $\Sigma_{\hat{U}}$ .

**Proposition 1** *Under the same assumptions as Theorem 1,*

$$\begin{aligned}\hat{U} &= \Sigma_U \underline{\Delta}'_N \Sigma_W^{-1} W \\ \Sigma_{\hat{U}} &= \Sigma_U \underline{\Delta}'_N \Sigma_W^{-1} \underline{\Delta}_N \Sigma_U\end{aligned}\tag{10}$$

The formula (10) is the most expedient for the calculation of the standardized signal extraction diagnostic of Findley et al. (2004), since the Kalman smoother only produces the diagonal entries

of  $M^{-1}$ . The work of de Jong and MacKinnon (1988) provides an approach to this problem, but is far less direct. If  $b_t$  denotes the state vector at time  $t$ , and  $\hat{b}_t$  is its estimate from the data, then some linear combination of the entries of  $\hat{b}_t$  will produce  $\hat{U}_t$ . De Jong and MacKinnon (1988) present formulas for the covariance matrix of  $\hat{b}_t$  and  $\hat{b}_s$  for any  $t$  and  $s$ , from which one could obtain the covariance of  $\hat{U}_t$  and  $\hat{U}_s$ . Since each such calculation involves matrix multiplications and inversions, with matrix dimension equal to the size of the state vector, and since order  $n^2$  such calculations are needed to produce  $\Sigma_{\hat{U}}$ , the procedure is cumbersome and indirect. In contrast, the matrix formulas (10) are simple and involve fewer operations; although computation of  $\Sigma_W^{-1}$  involves the inversion of a  $n \times n$  matrix,  $\Sigma_W$  has a Toeplitz structure that can be utilized to make an order  $n^2$  calculation – see Golub and Van Loan (1996). So asymptotically, use of (10) is at least as fast as the approach of de Jong and MacKinnon (1988), and is easier to implement.

### 3.3 Connections to the Bi-infinite Case

The minimum mean square error signal extraction filter under Assumption A is given by (see Bell 1984)

$$H(B) = \frac{\delta^N(B)\delta^N(F)\gamma_U(B)}{\gamma_W(B)}$$

with

$$\begin{aligned}\gamma_U(B) &= \frac{\theta^U(B)\theta^U(F)}{\phi^U(B)\phi^U(F)}\sigma_{\epsilon^U}^2 \\ \gamma_W(B) &= \frac{\theta^W(B)\theta^W(F)}{\phi^W(B)\phi^W(F)}\sigma_{\epsilon^W}^2.\end{aligned}$$

Defining

$$\gamma_V(B) = \frac{\theta^V(B)\theta^V(F)}{\phi^V(B)\phi^V(F)}\sigma_{\epsilon^V}^2,$$

the filter  $H(B)$  can be rewritten as

$$H(B) = \frac{\delta^N(B)\delta^N(F)/\gamma_V(B)}{\delta^N(B)\delta^N(F)/\gamma_V(B) + \delta^S(B)\delta^S(F)/\gamma_U(B)}$$

which can be compared with (6), with  $1/\gamma_V(B)$  playing the role of  $\Sigma_V^{-1}$ , etc. The Fourier Transform of the filter is

$$H(e^{-i\lambda}) = \frac{1/f_N(\lambda)}{1/f_N(\lambda) + 1/f_S(\lambda)}$$

where  $f_N$  and  $f_S$  denote the pseudo-spectral densities of  $N_t$  and  $S_t$ , i.e.,

$$f_N(\lambda) = \frac{f_V(\lambda)}{|\delta_N(e^{-i\lambda})|^2} \quad f_S(\lambda) = \frac{f_U(\lambda)}{|\delta_S(e^{-i\lambda})|^2}.$$



Thus by analogy, the quantity  $\Delta_N' \Sigma_V^{-1} \Delta_N$  plays the role of the inverse covariance matrix of  $N$ . In the case that the noise is stationary, this is exactly correct, since then  $\delta^N = 1$  and  $V = N$ . This provides an interpretation for the formula (6).

## 4 Appendix

**Proof of Theorem 1.** First note that all covariance matrices for  $U$ ,  $V$ , and  $W$  are invertible, since the processes are purely nondeterministic – see Proposition 5.1.1 of Brockwell and Davis (1991). Let  $1_n$  denote the  $n$ -dimensional identity matrix. A formula for  $F$  is given in Bell and Hillmer (1988) – also see Bell (2004) – which we reproduce here in part. Let  $\tilde{\Delta}_S$  be  $n \times n$  dimensional, with the first  $d_S$  rows given by  $[1_{d_S} \ 0]$  and the bottom  $n - d_S$  rows given by  $\Delta_S$ ; this is an invertible extension of  $\Delta_S$ . Then

$$F = \tilde{\Delta}_S^{-1} \begin{bmatrix} P \\ \Sigma_U \underline{\Delta}_N' \Sigma_W^{-1} \Delta \end{bmatrix}$$

follows from Bell and Hillmer (1988), for some  $d_S \times n$  dimensional matrix  $P$ . Thus

$$\Delta_S F = [0 \ 1_{n-d_S}] \tilde{\Delta}_S F = \Sigma_U \underline{\Delta}_N' \Sigma_W^{-1} \Delta. \quad (11)$$

Now clearly  $1_n - F$  is the extraction matrix for  $N$ , so by symmetry

$$\Delta_N (1_n - F) = [0 \ 1_{n-d_N}] \tilde{\Delta}_N (1_n - F) = \Sigma_V \underline{\Delta}_S' \Sigma_W^{-1} \Delta,$$

where  $\tilde{\Delta}_N$  is defined analogously to  $\tilde{\Delta}_S$ . If we write  $G = M^{-1} \Delta_N' \Sigma_V^{-1} \Delta_N$ , then

$$G(1_n - F) = M^{-1} \Delta_N' \Sigma_V^{-1} \Delta_N (1_n - F) = M^{-1} \Delta' \Sigma_W^{-1} \Delta$$

which uses (4). Similarly, we obtain

$$(1_n - G)F = M^{-1} \Delta_S' \Sigma_U^{-1} \Delta_S F = M^{-1} \Delta' \Sigma_W^{-1} \Delta$$

which implies that  $(1_n - G)F = G(1_n - F)$ , or  $F = G$ . It remains to check the invertibility of  $M$ . If a vector  $x$  were in the Null Space of  $M$ , then due to its symmetry and non-negative definiteness,  $x$  would have to be in the Null Space of both  $\Delta_S$  and  $\Delta_N$ . As demonstrated in Lemma 2 of McElroy and Sutcliffe (2004),  $x$  can be nonzero only if  $\delta^S(z)$  and  $\delta^N(z)$  share a common zero. It follows that only the zero vector is in the Null Space of  $M$ .  $\square$

**Proof of Theorem 2.** Note that  $\hat{S} - S$  and  $\hat{N} - N$  have the same covariance matrix, since they are negatives of each other. This is because  $\hat{N} = (1_n - F)Y$ , where  $1_n - F = M^{-1}\Delta_S'\Sigma_U^{-1}\Delta_S$  using (6). So we have

$$\begin{aligned} S - \hat{S} &= S - FY = (1_n - F)S - FN \\ &= M^{-1}\Delta_S'\Sigma_U^{-1}\Delta_S S - M^{-1}\Delta_N'\Sigma_V^{-1}\Delta_N N \\ &= M^{-1}\left(\Delta_S'\Sigma_U^{-1}U - \Delta_N'\Sigma_V^{-1}V\right). \end{aligned}$$

Since  $U$  and  $V$  are orthogonal, the covariance of this vector is

$$\Sigma_{S-\hat{S}} = M^{-1}\left(\Delta_S'\Sigma_U^{-1}\Delta_S + \Delta_N'\Sigma_V^{-1}\Delta_N\right)M^{-1} = M^{-1} \quad \square$$

**Proof of Proposition 1.** The formula for  $\hat{U}$  follows from (11). Hence

$$\Sigma_{\hat{U}} = \mathbb{E}[\hat{U}\hat{U}'] = \Sigma_U \underline{\Delta}_N' \Sigma_W^{-1} \underline{\Delta}_N \Sigma_U. \quad \square$$

## References

- [1] Bell, W. (1984), Signal Extraction for Nonstationary Time Series, *The Annals of Statistics* **12**, 646 – 664.
- [2] Bell, W. (2004). “On RegComponent Time Series Models and Their Applications,” in *State Space and Unobserved Component Models: Theory and Applications*, eds. Andrew C. Harvey, Siem Jan Koopman, and Neil Shephard, Cambridge, UK: Cambridge University Press.
- [3] Bell, W. and Hilmer, S. (1988), A Matrix Approach to Likelihood Evaluation and Signal Extraction for ARIMA Component Time Series Models. SRD Research Report No. *RR-88/22*, Bureau of the Census.
- [4] Brockwell, P. and Davis, R. (1991), *Time Series: Theory and Methods*, New York: Springer-Verlag.
- [5] Burman, J. (1980), Seasonal Adjustment by Signal Extraction. *Journal of the Royal Statistical Society. Series A (General)* **143**, 321 – 337.
- [6] Cleveland, W. and Tiao, G. (1976), Decomposition of Seasonal Time Series: A Model for the Census X-11 Program, *Journal of the American Statistical Association* **71**, 581 – 587.

- [7] de Jong, P. and MacKinnon, M. (1988), Covariances for Smoothed Estimates in State Space Models, *Biometrika* **75**, 601 – 602.
- [8] Durbin, J. and Koopman, S. (2001), *Time Series Analysis by State Space Methods*, Oxford: Oxford University Press.
- [9] Findley, D. and Martin, D. (2003), Frequency Domain Analyses of SEATS and X-12-ARIMA Seasonal Adjustment Filters for Short and Moderate-Length Time Series, Statistical Research Division Research Report Series, *Statistics #2003-02*, Census Bureau, Washington, D.C. Also [www.census.gov/srd/papers/pdf/rrs2003-02.pdf](http://www.census.gov/srd/papers/pdf/rrs2003-02.pdf)
- [10] Findley, D., McElroy, T., Wills, K. (2004), Diagnostics for ARIMA-Model-Based Seasonal Adjustment, *2004 Proceedings American Statistical Association* [CD-ROM]: Alexandria, VA. <http://www.census.gov/ts/papers/findleymcelroywills2005.pdf>
- [11] Gómez, V. and Maravall, A. (1997), Program SEATS Signal Extraction in ARIMA Time Series: Instructions for the User (beta version: June 1997), Working Paper 97001, Ministerio de Economía y Hacienda, Dirección General de Análisis y Programación Presupuestaria, Madrid.
- [12] Golub, G. and Van Loan, C. (1996), *Matrix Computations*, Baltimore and London: The Johns Hopkins University Press.
- [13] Kailath, T., Sayed, A., and Hassibi, B. (2000), *Linear Estimation*, Upper Saddle River, New Jersey: Prentice Hall.
- [14] Hannan, E. (1967), Measurement of a Wandering Signal Amid Noise, *Journal of Applied Probability* **4**, 90 – 102.
- [15] Harvey, A. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press.
- [16] Hilmer, S. and Tiao, G. (1982), An ARIMA-Model-Based Approach to Seasonal Adjustment, *Journal of the American Statistical Association* **77**, 377, 63 – 70.
- [17] Kohn, R. and Ansley, C. (1986), Estimation, Prediction, and Interpolation for ARIMA Models with Missing Data, *Journal of the American Statistical Association* **81**, 751 – 761.

- [18] Kohn, R. and Ansley, C. (1987), Signal Extraction for Finite Nonstationary Time Series, *Biometrika* **74**, 411 – 421.
- [19] Koopman, S. and Harvey, A. (2003), Computing observation weights for signal extraction and filtering, *Journal of Economic and Dynamic Control* **27**, 1317 – 1333.
- [20] Koopman, S., Shephard, N., and Doornik, J. (1999), Statistical Algorithms for Models in State Space Using SsfPack 2.2, *Econometrics Journal* **2**, 113 – 166.
- [21] Maravall, A. (2003), A Class of Diagnostics in the ARIMA-Model-Based Decomposition of a Time Series, Memorandum, Bank of Spain. <http://www.bde.es/servicio/software/tramo/diagnosticsamb.pdf>
- [22] McElroy, T. and Sutcliffe, A. (2004), An Iterated Parametric Approach to Nonstationary Signal Extraction, Statistical Research Division Research Report Series, *Statistics #2004-05*, Census Bureau, Washington, D.C. Also [www.census.gov/srd/papers/pdf/rrs2004-05.pdf](http://www.census.gov/srd/papers/pdf/rrs2004-05.pdf)
- [23] Sobel, E. (1967), Prediction of a Noise-distorted, Multivariate, Non-stationary Signal, *Journal of Applied Probability* **4**, 330 – 342.