
ECCE SIGNUM: AN R PACKAGE FOR MULTIVARIATE SIGNAL EXTRACTION AND TIME SERIES ANALYSIS

A PREPRINT

Tucker McElroy

U.S. Census Bureau
4600 Silver Hill Road, Washington, DC 20233
tucker.s.mcelroy@census.gov

James Livsey

U.S. Census Bureau
4600 Silver Hill Road, Washington, DC 20233
james.a.livsey@census.gov

ABSTRACT

The package provides multivariate time series models for structural analysis, allowing one to extract latent signals such as trends or seasonality. Models are fitted using maximum likelihood estimation, allowing for non-stationarity, fixed regression effects, and ragged-edge missing values. Simple types of extreme values can be corrected using the device of entropy maximization. Model adequacy is assessed through residual diagnostics, and model-based signal extraction filters can be assessed in time domain and frequency domain. Extracted signals are produced with uncertainty measures that account for sample edge effects and missing values, and the signals (as well as the original time series) can be forecasted.

Keywords Entropy maximization, Filtering, Missing values, Ragged edge, Seasonality, Signal extraction

1 Introduction

The R package **Ecce Signum**¹, or **sigex** for short, developed organically from R code for multivariate signal extraction required for the empirical applications of McElroy and Trimbur (2015). This methodology employed direct matrix formulas for signal extraction, which has certain advantages over the use of a state space smoother when the sample size is small. For further applications it became necessary to extend the models to handle diverse types of co-integration effects, such as common trends and common seasonality, and the resulting methods were summarized in McElroy (2017). However, at the same time it was recognized that extending the non-state space methodology to longer samples would be advantageous, given that general forms of integrated processes can be handled, including non-Markovian processes and long memory processes not amenable to a state space embedding. Algorithmic work to make feasible these extensions is discussed in McElroy (2018) and McElroy (2022); extensions to model estimation are discussed in McElroy and Roy (2018).

The result of these developments is a flexible set of tools for the analysis of multivariate time series. Time series data with some or all of the following attributes can be analyzed with **Ecce Signum**:

- Multivariate time series, or univariate time series with embedded multivariate structure
- Fixed effects, such as holiday patterns
- Aberrancies, such as additive outliers
- Complex dynamics, including trend, seasonality, and business cycles
- Seasonal cycles of diverse period (e.g., weekly and annual effects)
- Linked dynamics, such as common trends and common seasonality
- Ragged edge missing values
- Changes in sampling frequency

¹Latin for “Behold the Proof”; cf. Falstaff’s speech in Henry IV, Part 1 (Act 2, Scene 4).

Highly exotic time series cannot be handled appropriately with **Ecce Signum**. Data that can be easily linearized (i.e., rendered as a linear time series, possibly with Gaussian marginals) through transformations, regression, and differencing is appropriate for analysis. We have in mind several possible objectives of analysis, which the methodological tools can address:

- Modeling of the data via reduction to a maximum entropy residual
- Identification and estimation of fixed effects
- Linearization of the data via maximum entropy transformations
- Extraction of underlying signals, which may be common, such as trends, seasonals, and cycles
- Imputation of missing values
- Forecasting and aftcasting
- Quantification of uncertainty in extractions

The modeling philosophy is informed by the concept of entropy, and is discussed in McElroy and Penny (2019). Tools are available to compare fitted models and evaluate their performance, and two methods of fitting are presented: maximization of the Gaussian likelihood utilizing a proper factorization, and method-of-moments – which is faster and more appropriate for high-dimensional parameter spaces.

Modeling of univariate structural time series is well covered in the R package world. In fact, base R includes the function `StructTS`, which fits a univariate structural time series model via maximizing the Gaussian likelihood. If an analyst is willing to put their model formulation into the state-space framework and make use of the Kalman filter and smooth functions, there exist many packages that can estimate such a specification, viz., **dlm** (Petris, 2010), **dse** (focuses on ARMA models) (Gilbert, 2020), **bsts** (Bayesian implementation) (Scott, 2020), and **stsm** (López-de Lacalle, 2016). There do exist some packages for processing multivariate structural time series. The **MARSS** package (Holmes et al., 2012, 2020) adopts a state space approach to fitting and filtering algorithms. Additionally, **KFAS** (Kalman Filter and Smoother for Exponential Family State Space Models, Helske (2016)) can estimate multivariate time series models. **MARSS** uses **KFAS**, and implements more stable filtering/smoothing algorithms. **KFAS** on the other hand, allows users to consider non-Gaussian time series. The **MARSS** package currently can estimate a multivariate state-space model specification with seasonal components using either an EM algorithm or nonlinear optimization via BFGS. (However, it is noted in the user guide that the EM algorithm can be challenging for multivariate models.)

Ecce Signum does not rely on a state-space formulation, hence allowing for a quite flexible specification of structural components. A note in the documentation for `StructTS` alludes to the difficulty in performing optimization for even univariate structural time series models; this might explain the low number of R packages currently available for multivariate structural modeling. A crude search through R documentation using `RSiteSearch("structural time series")` (Graves et al., 2010) returns 114 documents matching our query. However, including the qualifier `RSiteSearch("multivariate structural time series")` returns 0 matches. This supports the contention that **Ecce Signum** can meaningfully add to R's ecosystem.

The other types of R packages that **Ecce Signum** overlaps with are those which perform signal extraction on high frequency time series. We use the term high frequency to mean series observed at the weekly, daily, or even higher frequencies. Here we discuss a few packages for performing high frequency seasonal adjustment. The **forecast** R package (Hyndman and Khandakar, 2008; Hyndman et al., 2020) incorporates multiple functions that can accomplish high frequency seasonal adjustment; `stlf()` uses an STL decomposition, `mstl()` is a variation for multiple seasonal periods, and `auto.arima()` can be utilized to perform dynamic harmonic regression (Hyndman and Athanasopoulos, 2018). **forecast** also includes an implementation of TBATS (Exponential Smoothing State Space Model With Box-Cox Transformation, ARMA Errors, Trend And Seasonal Components) (De Livera et al., 2011). **prophet**, a package developed at Facebook (Taylor and Letham, 2018), attempts to provide a black box seasonal adjustment method focused upon forecasting. For well-behaved series this works well, but in our experience the many curious features in most daily time series still generate challenges for these sorts of automated methods. **dsa**, a package developed for daily seasonal adjustment at Bundesbank (Ollech, 2018), implements a version of STL. A general overview of these methods can be found in Ladiray et al. (2018), and a comparison of different methods by their out-of-sample forecast accuracy can be found in Timmermans et al. (2017).

The framework of **Ecce Signum** consists of an N -dimensional multivariate time series, which may be observed with missing values at various times for various components. After the possible application of a log transform, the model is written additively in terms of fixed effects (given through user-specified time series regressors) and latent stochastic processes (which the user specifies during model declaration). The class of processes is fairly broad, including ARIMA and SARIMA models (Box et al., 2015), as well as popular structural models discussed in Harvey (1989) and Harvey

and Trimbur (2003). In broad strokes, the analyst proceeds from exploratory tools to a putative model, which can then be fitted using Gaussian maximum likelihood estimation – possibly accounting for additive outliers by entropy methods; if there are latent stochastic processes present, these are implicitly fitted as a result. If the modeling is satisfactory, the analyst can then proceed to applications such as forecasting or signal extraction.

The paper is organized into the following sections: Section 2 proceeds through the various stages of analysis that a user of **Ecce Signum** would utilize sequentially. Then Section 3 provides more statistical detail about model parameterizations, estimation methods, and vector embeddings. Several detailed examples are explored in Section 4; these are vignettes in the package. Appendix A treats cyclical processes, and Appendix B discusses issues related to weekly time series. Finally, Appendix C gives more granular detail on the functionality, with a topology of the function call structure given through look-up tables.

2 Stages of Analysis

In the analysis of multivariate time series in **Ecce Signum**, there are several stages that are interdependent. We discuss these in turn, in the sequence that they would typically be utilized in an application.

2.1 Loading, Transforming, and Exploring

We begin with entering the data into an R variable *data*, which is a $T \times N$ -dimensional matrix involving T time points and N series. If there are missing values, these can be encoded with “NA,” in which case the full sample size T is defined by taking the size of the union of all temporal indices for each component time series. A first stage of analysis is to call `sigex.load` with metadata such as the start date, period (i.e., number of seasons per year, if applicable), and epithets (i.e., shortened names or mnemonics) for each component series.

Secondly, a call to `sigex.prep` allows the analyst to apply a log transformation, aggregate component series (say if one needs to analyze the sum of all the series), restrict the temporal range, select a subset of component series for analysis, and plot. The output we shall denote as *data.ts*, and for convenience the cross-sectional dimension will be denoted N (even if a subset of series was selected). Note that this function call may also be utilized after some exploratory analysis to identify transformation, range, and subcollection; see Chapter 1 of McElroy and Politis (2020) for background.

Optionally, spectral plots can be viewed with a call by `sigex.specar`, which is just a wrapper for the native `spec.ar` of R. To the trained analyst, such plots provide insight into the salient dynamics of each component time series, and can assist one in positing unit root model structures in subsequent analyses; see Chapter 9 of McElroy and Politis (2020), or the discussion in McElroy and Roy (2021b). For instance, a sharp peak in the spectral density at a frequency $\lambda \in [0, \pi]$ indicates that a unit root of the form $e^{i\lambda}$ may be present in the autoregressive representation of the time series.

2.2 Model Declaration

Models in **Ecce Signum** have two facets: stochastic effects and fixed effects. Fixed effects determine the mean of the time series through specified regressors, one for each component series. Just as with RegARIMA models (Findley et al., 1998), once these fixed effects are subtracted out what remains is a mean zero stochastic process, which can be modeled with one or more latent processes². Let $\{X_t\}$ denote the transformed multivariate time series, for which our sample corresponds to *data.ts*. Our modeling framework is given by

$$X_t = z_t' \beta + Y_t. \quad (1)$$

Here $\{X_t\}$ is the data process, an N -dimensional time series with time-varying mean $\mathbb{E}[X_t] = z_t' \beta$ by our modeling assumption. The stochastic process $\{Y_t\}$ has mean zero. For each t in the sample period $\{1, 2, \dots, T\}$, the regressors can be placed in a matrix z_t of dimension $r \times N$, where r represents the union of all fixed effects for all component series. Because each component series has separate regression effects (even if an identical regressor is used for several component series, the parameters will be distinct by assumption), the structure of z_t is block diagonal. In particular, if $z_t(j)$ is an r_j -dimensional vector of regressors for component series j (for $1 \leq j \leq N$), then

$$z_t = \begin{bmatrix} z_t(1) & 0 & \dots & 0 \\ 0 & z_t(2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & z_t(N) \end{bmatrix}.$$

²Called “unobserved components” in the econometrics literature. We use the term latent process to avoid confusion with the components of a vector time series.

Then with $r = \sum_{j=1}^N r_j$, we have $\beta' = [\beta(1)', \beta(2)', \dots, \beta(N)']$ and the fixed effect for the j th component series is $z_t(j)' \beta(j)$. More details on adding fixed effects are given below, and in Section 3.1.3.

Our first step in modeling $\{X_t\}$ is to build the stochastic model for $\{Y_t\}$. Here we use *mdl* as the variable (a list object) storing all the model information. We initially declare *mdl* to be NULL, and then use repeated calls of `sigex.add` to insert latent processes into the model. The key inputs to these function calls are the scalar differencing operator $\delta(B)$ (where the backshift operator B acts on a time series via $BX_t = X_{t-1}$), the rank configuration for the driving white noise, the time series model type (e.g., VARMA or SARMA, etc.) with ancillary information such as the model order, and an epithet (e.g., trend, irregular, weekly seasonal, etc.). If a component is stationary, simply input $\delta(B) = 1$; otherwise, input various polynomials as coefficient strings, taking care that each non-stationary latent process has a unique polynomial³. In particular, suppose there are K vector latent processes $\{S_t^{(k)}\}$ (for $1 \leq k \leq K$) that are related to the stochastic portion of $\{X_t\}$ via

$$Y_t = \sum_{k=1}^K S_t^{(k)}. \quad (2)$$

For instance, $\{S_t^{(1)}\}$ might correspond to a trend, and $\{S_t^{(2)}\}$ to a transient stationary structure. Each latent component has an associated (scalar) differencing operator $\delta^{(k)}(B)$ such that

$$\delta^{(k)}(B)S_t^{(k)} = \underline{S}_t^{(k)}, \quad (3)$$

where $\{\underline{S}_t^{(k)}\}$ is a mean zero, covariance stationary vector stochastic process. The differencing in (3) is applied in the same way to each component of the vector process. The stationary model structure can be specified through ARMA/SARMA (the same model is used for each component), or a VARMA/SVARMA, or a cyclical model. Generally, all the models implemented in **Ecce Signum** are special cases of the causal moving average form

$$\underline{S}_t^{(k)} = \Psi^{(k)}(B)\epsilon_t^{(k)}, \quad (4)$$

where $\Psi^{(k)}(z)$ is a matrix power series. The coefficient matrices are multiples of the identity in the case of ARMA/SARMA models, or a cycle model, but for VARMA/SVARMA these matrices can be non-trivial; further details are in Section 3.1.2. The vector white noise process $\{\epsilon_t^{(k)}\}$ is independent and identically distributed (i.i.d.) with unspecified marginal distribution, and covariance matrix $\Sigma^{(k)}$ that can be reduced rank; further details are in Section 3.1.1.

The specification of these components takes some care and experience. Note that any unit roots deemed to be present in the data process' autoregressive representation must be featured among one of the latent processes, since the product of the differencing polynomials for the latent processes equals the differencing polynomial for the data process (Hillmer and Tiao, 1982). In particular, setting $\delta^{(-k)}(z) = \prod_{j \neq k} \delta^{(j)}(z)$, it follows from (2) and (3) that

$$\delta(B)Y_t = \sum_{k=1}^K \delta^{(-k)}(B)\underline{S}_t^{(k)}, \quad (5)$$

where $\delta(z) = \prod_{k=1}^K \delta^{(k)}(z)$. The right hand side of (5) is a sum of stationary processes, and hence $\delta(B)Y_t$ is stationary too, and yet if there are any common unit roots among any of the $\delta^{(-k)}(z)$ then $\{Y_t\}$ has been over-differenced – this will generate estimation problems, because the resulting process will be non-invertible (Chapter 6 of McElroy and Politis (2020)). Therefore it is important to specify the various $\delta^{(k)}(z)$ such that they are relatively prime. Henceforth $\underline{Y}_t = \delta(B)Y_t$ defines the mean zero stationary process given by (5).

The second stage in model declaration is building in fixed effects; at a minimum, one must call `sigex.meaninit`, which ensures that a polynomial trend is inserted with order corresponding to the number of specified trend differences in the first stage. This is important, because applying differencing to the data process will typically leave a non-zero constant mean, which is automatically associated with the leading coefficient of a polynomial trend effect. Next, if additional fixed effects are desired (e.g., calendar effects, level shifts, etc.) then these can be incorporated through calls to `sigex.reg`. This function requires that the regressors have already been constructed for the full sample span $\{1, 2, \dots, T\}$, so if a reduced range has been considered any regressors inputted to the program must be appropriately (and manually) culled. Because each component series can have different regressors, one must set up a loop over the N components to allocate the specified regressors.

³Identification problems in signal extraction occur if diverse latent components share common unit roots; see Bell (1984) and McElroy (2008) for discussion.

For moving holiday effects such as Easter, the utility `gethol` for daily time series can be used. One inputs the desired calendar dates, and a window of days before and after each occurrence of the holiday when it is deemed that economic activity will be increased relative to the rest of the year; this mimics the `Genhol` routine (Census Bureau, 2015) of X-13ARIMA-SEATS for constructing moving holiday regressors. In particular, if t is the index corresponding to a holiday, and f and a indicate the number of days before and after the holiday to be considered, then the window is $t - f, t - f + 1, \dots, t, \dots, t + a - 1, t + a$; the length of the window is $a + f + 1$. A corresponding sequence of regressors, computed using the start and end dates of the time series sample, is generated.

2.3 Model Fitting

By the fitting of a model, we mean that parameter estimates are obtained from the data. **Ecce Signum** has two methods for doing this: Method of Moments (MOM) and Maximum Likelihood (ML). Before these can be used, initial values of the parameters need to be set up, and the practitioner makes choices about whether a parameter is fixed or free. A detailed discussion of parameterization and fitting is given in Section 3.1, and MOM is expounded in Section 3.2.

Typically, the ML method is preferred. This is because the MOM method is limited to fairly simple latent process structures; essentially, the causal moving average portion $\Psi^{(k)}$ of the k th latent component model (4) is held fixed at an initial value specification, and only $\Sigma^{(k)}$ is estimated. Moreover, the MOM method is not equipped to handle missing values, and the estimation of fixed effects is through OLS. In contrast, ML allows for ragged edge missing values and the most general types of model structures encoded in **Ecce Signum**; moreover, the statistical estimates are efficient when the model is correctly specified and any non-Gaussian features are absent – see pp. 52–59 of Taniguchi and Kakizawa (2000), as well as the fourth appendix of Holan et al. (2017) for details. The parameters for fixed effects correspond to Generalized Least Squares (GLS) estimates based upon the time series model parameters (McElroy and Holan, 2014). Furthermore, we can impose parameter constraints (Section 3.1.4) and generate test statistics for parameter significance.

The demerits of ML are the larger computational resources required, as well as the usual pitfalls of nonlinear optimization. Each likelihood evaluation has an expense relative to the size of T and N , as well as the model complexity (McElroy, 2018). For models with many parameters, e.g., 100 or more, the parameter manifold is large and requires a time-consuming search. As usual, there are no guarantees that the algorithm’s termination corresponds to a global maximum. Our approach in **Ecce Signum** is to re-parameterize constraints into an unconstrained (so-called) pre-parameter space homeomorphic to Euclidean space, as advocated by Pinheiro and Bates (1996). Different initializations can be specified by the user, as well as what type of optimizer is desired; the pre-parameter corresponding to the last successful likelihood evaluation is saved to the global variables in the R workspace, so that if *optim* crashes the user can restart with a slight perturbation of this last value.

Upon termination of the optimization routine, the user can examine the parameter values and assess the model fit. This is possible by computing the time series residuals and measuring their entropy via a Shapiro-Wilks normality test (Shapiro and Wilk, 1965) as well as Ljung-Box statistics (Ljung and Box, 1978). More generally, through modeling we are attempting to map the data vector to a maximum entropy residual; see Chapter 8 of McElroy and Politis (2020) for background on this modeling philosophy. A serially uncorrelated Gaussian time series has maximum entropy, and therefore the user can assess departures of the model residuals from this ideal target.

More generally, the transformed time series may have extreme values present – see McElroy (2016b) for an overview of extremal time series, and the impact on signal extraction problems. As discussed in McElroy and Penny (2019), non-Gaussian processes can be handled through non-linear forecasting and signal extraction algorithms, or by first applying an extreme-value adjustment procedure followed by a linear algorithm. **Ecce Signum** follows the latter path, chiefly for simplicity and speed (non-linear methods typically require more nuanced modeling, and more expensive algorithms). This approach is in the tradition of intervention analysis: Chang et al. (1988) use an additive outlier (AO) regressor (Ljung, 1993), and Tsay (1986) considers more general types of outlier effects (e.g., level shifts).

We view the AO as an unknown stochastic additive effect, not a fixed effect; as shown in McElroy and Penny (2019), entropy is maximized by deleting the corresponding observation. Hence, the treatment of AOs in **Ecce Signum** is by inserting an NA for any putative additive extreme, and comparing the likelihood. The difference in log likelihoods has an interpretation as a test statistic for the significance of any collection of AOs. The identification of extremes is an important facet of model building, and can be accomplished in tandem with fitting a specified model; as missing values are introduced for possible extremes, the remaining sample attains to more typically Gaussian characteristics, and the ragged edge Durbin-Levinson algorithms allow for quick computation.

If multiple satisfactory models have been obtained, we can perform nested comparisons using a difference of likelihoods (McElroy, 2016a). If particular pre-parameters are not significantly different from zero, they can be replaced by a zero (this is entered as a constraint) and the model re-fitted (as a nested model). Another possible restriction is through the

rank configuration of a latent process' white noise covariance matrix; see 3.1.1, although in this case a statistical test arising from the difference of likelihoods is not valid (without a non-standard asymptotic theory) because the restriction lies on the boundary of the parameter space.

Through use of these tools, we suppose the analyst has arrived at a final model, which is bundled using `sigex.bundle`: this includes `data.ts` (where any detected AOs have been marked by inserting NAs), the model `mdl`, the fitted pre-parameters ψ , and the transformation used.

2.4 Applications of the Fitted Model

At this point in the analysis we can generate output from the fitted model for applications. We focus on casting (forecasting, aftcasting, and midcasting) and signal extraction. The function `sigex.midcast` generates midcasts for all missing values in the data, which are marked as usual with the NA symbol; a number of forecasts and aftcasts to be generated as well. `sigex.midcast` is a "deluxe" version of the casting methodology, also generating the large covariance matrix between all casting errors; since this can be expensive to compute and store if there are hundreds of casts, a faster version is available when uncertainty is not needed – `sigex.cast` generates only forecasts and aftcasts, and presumes there are no missing values.

Neither of these two function calls is complete without accounting for regression effects: both `sigex.cast` and `sigex.midcast` presume that fixed effects have already been removed from `data.ts`, and these regression effects need to be extended and added back to the series. The reason this is not done automatically, is that there is no way to know future values of general regressors encoded in `mdl`; clearly, polynomial and sinusoidal functions can be extrapolated/interpolated, but for holiday regressors or calendar effects a human is required to do the modifications. One approach to analysis is to load the data with extended regressors entered into `mdl`, and `data.ts` padded fore and aft with NA, up to the horizons that are anticipated to be needed.

For signal extraction, there are two main options: Wiener-Kolmogorov (WK) filtering or *ad hoc* (AC) filtering, as discussed in McElroy (2022). In the first case, we have in mind a model that contains at least two latent processes, and it is of interest to extract one of them (or possibly a sum of a subset of these latent processes). For the specified latencies, **Ecce Signum** computes the minimum mean square linear estimate conditional on the (ragged edge) data, along with time-varying uncertainty. This uncertainty can be viewed as being composed of two portions: the so-called WK error and the casting error. In contrast, the AC case can be applied when there is only a single latency specified, but the analyst must define and supply the desired (non-model based) filter. Like the WK case, the optimal estimate is computed from the ragged edge data, but the time-varying uncertainty only depends on the casting error, as the WK error component is not defined. In either case, the extracted signal $\{U_t\}$ takes the form

$$U_t = \Upsilon(B)Y_t = \sum_{h=-\infty}^{\infty} \Upsilon_h Y_{t-h},$$

where $\Upsilon(B)$ is either a WK or AC filter. In the WK case, $\{U_t\}$ is viewed as an estimate of a stochastic process, such as a latent process $\{S_t^{(k)}\}$ (or aggregation of such). But for AC filtering, $\{U_t\}$ is viewed as the actual target.

We begin with the WK case, and there are two methods available: `sigex.signal` provides the exact filter matrix that yields the WK signal extraction at all time points (derived in McElroy and Trimbur (2015)), but requires there be no missing values. (Although we can apply `sigex.signal` to ragged edge data that has been "patched" with midcasts, we will not be able to account for midcasting uncertainty in the signal extraction uncertainty measure.) A secondary call to `sigex.extract` furnishes the actual extractions (along with point-wise confidence intervals defined as ± 2 standard deviations, based on the extraction error covariance matrix), but the signal extraction error covariances (assuming no missing values) as well as the filter weights are produced by `sigex.signal`. This method is not recommended for longer samples ($T > 500$), because a non-Toeplitz matrix of dimension NT needs to be inverted.

The broader method `sigex.wkextract` is generally preferred, except in cases where the time-varying filter weights are needed. The technique is described in McElroy (2022), and involves first patching the series with midcasts – and a set number of aftcasts and forecasts to handle the sample boundary – followed by filtering with the so-called WK filter. This method also obtains both components of the signal extraction uncertainty, accounting for missing values and the sample boundary. However, unlike `sigex.signal` it assumes a truncation of the WK filter. The truncation error can be reduced to any desired level by taking lag cutoff to be larger, but this induces a computation cost because there is a corresponding increase to the number of aftcasts and forecasts (with uncertainty) that are needed.

An additional feature is the ability to generate casts of signals; automatically the signal is extracted for time points $\{1, 2, \dots, T\}$, but for a specified horizon H we can obtain forecasts for times $\{T + 1, \dots, T + H\}$ and aftcasts for times $\{1 - H, \dots, 0\}$. A second auxiliary feature is the extraction of linear combinations of a signal: suppose that our

extraction is

$$\Xi(B)U_t = \sum_{j=0}^p \Xi_j U_{t-j}$$

for coefficient matrices Ξ_j . The matrix polynomial $\Xi(z)$ is passed into `sigex.wkextract`, and the extraction with appropriate covariance is returned. Note that such extended extractions can also be obtained through use of `sigex.signal`, by applying appropriate linear combinations to both the filter matrix and the extraction error covariance matrix, but in `sigex.wkextract` the error calculations are done using frequency domain instead of time domain. Mathematical details of these matters can be found in McElroy (2022).

A second option is to generate AC extractions. The desired filter $\Upsilon(z)$ is entered as an array object, in a truncated format:

$$\Upsilon(z) \approx \sum_{h=-c}^{m-1-c} \Upsilon_h z^h,$$

where m is the total number of nonzero coefficient matrices, and c is a shift parameter that indicates the proper centering. For instance, setting $c = 0$ indicates that the filter is concurrent, i.e., the output depends only on present and past data. Then `sigex.adhocextract` computes the filtered estimates, first patching and extending the series as needed; this routine functions much like `sigex.wkextract`, only the AC filter is used instead of the WK filter. Also, if we really desired a signal of the form $\Xi(B)U_t$, this amounts to

$$\Xi(B)U_t = \Xi(B)\Upsilon(B)Y_t,$$

so our AC filter would just be redefined as $\Xi(B)\Upsilon(B)$.

Plots of casts and extractions, with uncertainty indicated, is a final step in the analysis. Note that regression effects should be added back to casts, and may be added to certain extractions. For example, an Easter moving holiday regressor would be added to the final extracted seasonal component. To pair estimated parameters with the appropriate regressors, `sigex.fixed` can be used. For plotting, `sigex.graph` can be used to overlay extracted signals if the data has already been plotted; this function presumes the format of the output of `sigex.extract`. The second argument of `sigex.graph` accepts fixed effects that are added onto the extraction in the first argument. The starting date and period are also entered, and it is possible to vertically displace plotted signals (if this is visually desirable). For example, the seasonal component will be plotted around zero, and hence may be too far below the data and trend extraction to be easily visible – it can be shifted up in such a case. The extracted signals can also be assessed as to whether they contain only the proper dynamics. One can examine the spectral density through a call to `sigex.specar`, and of course the autocorrelations can be checked as usual.

It may be of interest to examine the WK filter coefficients, as well as the frequency response function. `sigex.wk` can be used to compute and plot the WK coefficients Υ_h . In particular, there are multiple panels and the jk th panel (for $1 \leq j, k \leq N$) plots the jk th entries of Υ_h for $h = -m, \dots, m$. For example, the weights on the second input series used to extract the first output series are given by $j = 1$ and $k = 2$. Similarly, `sigex.frf` computes (and can be plotted in multiple panels by calling `sigex.getfrf`) the frequency response function $\Upsilon(e^{-i\lambda})$. The output is an array, with elements corresponding to a mesh of frequencies λ taken in the interval $[0, \pi]$. Note that for purposes of inverting the frequency response function into coefficients, it is always sufficient to restrict to the interval $[0, \pi]$, because

$$\Upsilon_h = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\lambda h} \Upsilon(e^{-i\lambda}) d\lambda = \frac{1}{\pi} \int_0^{\pi} \Re [e^{i\lambda h} \Upsilon(e^{-i\lambda})] d\lambda$$

by change of variable. (This integration is discretized and calculated in `sigex.wk`.)

3 Detailed Aspects

3.1 Model Parameterization

Ecce Signum makes use of unconstrained parameterization, which means there exists a bijection between all model parameters and Euclidean space. Specifically, all model parameters are stored in a list object *param* (whose format depends on the specified model), and there exists a bijection that maps a real vector ψ to *param*. The vector ψ is further decomposed into three portions: ξ (which maps to covariance matrices for driving white noise of the latent processes), ζ (which maps to time series models, such as ARMA, for the latent processes), and β (which maps to the regression parameters for all component series). The functions `sigex.par2psi` and `sigex.psi2par` pass *param* to ψ and back.

Why do we do this? The format of *param* involves the Generalized Cholesky Decomposition (GCD) of covariance matrices as well as things such as AR coefficients, organized in a list that corresponds to component series; hence a user

can more easily interpret the quantities in *param*. But ψ consists of the unconstrained parameterization of covariance matrices and stable AR polynomials and so forth; it is not easy to understand what a particular component entry of ψ corresponds to. Typically the user will print *param*, which has a readable structure; the standard user will not need to understand the mapping of ψ to *param*.

3.1.1 Covariance Matrix Parameterization

First, the parameterization of a non-negative definite matrix Σ proceeds along the lines described in McElroy (2017). Specifically, the GCD (described in Golub and Van Loan (2012)) decomposes a possibly singular Σ into unit lower triangular L and diagonal D such that

$$\Sigma = L D L'. \quad (6)$$

For an N -dimensional matrix, there are $\binom{N}{2}$ lower triangular entries in L that can be any real number. The matrix D has N non-negative entries, which are all strictly positive if Σ is positive definite (invertible). Hence the total number of parameters required is $\binom{N+1}{2}$. If the rank of Σ is less than N , then at least one entry of D is zero, and the corresponding column of L becomes obsolete in the decomposition. To determine the requisite number of parameters, one must further specify which entries of D are zero; we call this a *rank configuration*. Clearly, if d_j (the j th diagonal entry of D) is zero, then $N - j$ free parameters in L become obsolete and can be removed from the needed parameterization. (In the case $j = N$, no reduction occurs.)

The complement of the rank configuration corresponds to entries of D that are positive, and the variable *vrank* captures this. If the rank is reduced, the number of parameters needs to be computed, and `sigex.param2gcd` constructs an L matrix from a vector of real numbers, where any columns corresponding to the rank configuration are omitted (i.e., L is a rectangular matrix with potentially more rows than columns). The positive entries of D are parameterized by taking the exponential of a real number. In this way ξ is mapped into the various covariance matrices of *param*, one for each latent process.

3.1.2 Serial Dependence Parameterization

Second, the parameterization of the serial dependence in the stochastic process is accomplished through `sigex.par2zeta` and `sigex.zeta2par`. The exact mapping is dictated by the specified model structures. For an ARMA model, both the AR and MA polynomials are stably parameterized through the *pacf* discussed in McElroy and Politis (2020). In particular, for $j = 1, 2, \dots, p$ for a degree p stable polynomial (with a minus convention, so that $\phi(x) = 1 - \sum_{j=1}^p \phi_j z^j$) we consider recursively parameterizing all polynomials of degree up to p , denoting the coefficients at the j th stage via $\phi_1^{(j)}, \dots, \phi_j^{(j)}$. For $p = 1$, we have $\phi_1^{(1)} \in (-1, 1)$, which is accomplished by taking $\phi_1^{(1)} = (\exp\{\zeta_1\} - 1)/(\exp\{\zeta_1\} + 1)$. Then for $2 \leq j \leq p$, parameterize $\phi_j^{(j)}$ in the same way and update $\phi_1^{(j)}, \dots, \phi_{j-1}^{(j)}$ via

$$\begin{bmatrix} \phi_1^{(j)} \\ \vdots \\ \phi_{j-1}^{(j)} \end{bmatrix} = \begin{bmatrix} \phi_1^{(j-1)} \\ \vdots \\ \phi_{j-1}^{(j-1)} \end{bmatrix} - \phi_j^{(j)} \begin{bmatrix} \phi_{j-1}^{(j-1)} \\ \vdots \\ \phi_1^{(1)} \end{bmatrix}.$$

This update can also be expressed as follows: multiply the old coefficient vector by the matrix $A(j-1)$ that is given by the sum of a $(j-1)$ -dimensional identity matrix and another matrix with $-\phi_j^{(j)}$ on the off-diagonal. In this way we map real numbers ζ_1, \dots, ζ_p to $\phi_1^{(p)}, \dots, \phi_p^{(p)}$. By applying the inverse of $A(j-1)$ recursively, we can also recover the original pre-parameters from a given stable polynomial. For a SARMA model (with seasonal period s),

$$\phi(B) \Phi(B^s) Y_t = \theta(B) \Theta(B^s) \epsilon_t,$$

and each polynomial is stably parameterized. There are four such polynomials – a regular AR polynomial ϕ , a seasonal AR polynomial Φ , a regular MA polynomial θ , and a seasonal MA polynomial Θ , all of which are defined using the minus convention of Box et al. (2015).

For a VAR model, the parameterization of Roy et al. (2019) is used, whereby real numbers are mapped into the space of stable matrix polynomials. The functions `var2.par2pre` and `var2.pre2par` provide the mapping from $\phi(B)$ to the real numbers, and back again. For VARMA and SVARMA, there are additional polynomials for the moving average and seasonal aspects.

Stochastic cycles come in four varieties. The basic n th order Butterworth cycle is described in Harvey and Trimbur (2003), and takes the form

$$(1 - 2\rho \cos(\omega)B + \rho^2 B^2)^n C_t = (1 - \rho \cos(\omega)B)^n \epsilon_t$$

for a cyclical latent process $\{C_t\}$, where $\{\epsilon_t\}$ is white noise (this can be multivariate). Here ρ corresponds to the persistency of the cycle, and is restricted to $(0, 1)$, while ω is the frequency (or 2π divided by the period of the cyclical effect). A higher order n ensures that the high frequency aspects of the cycle are damped, so that the resulting stochastic process more purely represents only cyclical phenomena. In the model specification, the user specifies n , and can also input range parameters for ρ and ω . This can be used to impose hard boundaries on the parameter values, although **Ecce Signum** considers an open interval dictated by the specified bounds and maps ζ to this interval using a logistic transform.

A second type of process called the Balanced cycle can also be used, and the spectral shape of the resulting process is slightly tilted – as compared to the Butterworth cycle – in such a way that the high frequency content is lowered. Harvey and Trimbur (2003) argue for the more favorable properties of the Balanced cycle, which can be specified in the same way as the Butterworth.

Finally, both cycles can be “stabilized,” which refers to a modification of the model that is non-invertible. Bell and Pugh (1989) first experimented with such models, describing them as canonical component models. Any stationary stochastic process has spectral density f that is non-negative, and if it is strictly positive we can consider subtracting off the minimum value (a positive number), thereby rendering f shifted downwards such that it now takes on a zero value. A process having such a spectral density is said to be non-invertible (Chapter 6 of McElroy and Politis (2020)); while slightly problematic from the standpoint of fitting and forecasting, such a process has the interpretative advantage that it is maximally stable in the sense that as much white noise as possible has been removed. The intuition here is that we can decompose a given stochastic process into the sum of a non-invertible process and an independent white noise process of variance equal to the minimum value of the spectral density – an operation sometimes called “canonization,” as it is at the core of the canonical decomposition routine of Hillmer and Tiao (1982). More details on the cycle models are given in Appendix A.

The last model available is the “damped trend.” This is essentially an AR(1) process where the user can place bounds on the range of the ϕ_1 parameter, in the same manner as with the persistency parameter in the cycles. Typically this specification would be used in conjunction with a trend differencing operator such as $\delta(B) = 1 - B$. The resulting model for a trend latent process $\{S_t\}$ would then be

$$(1 - B)(1 - \phi_1 B)S_t = \epsilon_t.$$

In Trimbur (2006) this type of trend interpolates continuously between the random walk and integrated random walk trends, allowing one to more finely model and extract a trend with a varying degree of smoothness (higher values of ϕ_1 correspond to smoother extracted trends).

3.1.3 Regression Effects Parameterization

Third, the regression parameters for each series are given by β , the final portion of ψ . This portion is straightforward; note that a mean effect for the differenced time series is always added through `sigex.meaninit`, so there will always be at least N parameters in β . When a user decides to add a regression effect to a component series, the program checks to see whether this is in the null space of $\delta(B)$, the full differencing operator for $\{Y_t\}$ described in (5). Consider applying $\delta(B)$ to the j th component (for $1 \leq j \leq N$) of (1). We obtain

$$\delta(B)X_{tj} = \delta(B)z_t(j)'\beta(j) + \underline{Y}_{tj},$$

where \underline{Y}_{tj} is the j th component of $\underline{Y}_t = \delta(B)Y_t$. As discussed below (5), the process \underline{Y}_t is invertible so long as the various $\delta^{(k)}(z)$ are relatively prime. If $\delta(B)$ completely annihilates $z_t(j)$, then there are no fixed effects at all; because the fitting of non-stationary processes relies upon a factorization of the Gaussian likelihood that features the likelihood of the differenced data $\{\underline{Y}_t\}$ (see McElroy and Monsell (2015) and McElroy (2022)), it follows that any covariate with corresponding regressor in the null space of $\delta(B)$ cannot be identified – therefore in **Ecce Signum** such a regressor is removed from the model. This is automatically done in `sigex.reg`, and the removal is done without further warning to the user. (The desired regression effect that the user enters is never incorporated into the model.) For regressors not in the null space of the full differencing operator, $\delta(B)z_t(j)$ will be non-zero and hence generates the time-varying mean for $\delta(B)X_{tj}$. This is accounted for in likelihood and casting calculations.

There is a type of weak converse to this discussion. If we have included differencing in our model, without loss of generality we can envision the inclusion of all regression effects that are in the null space of this $\delta(B)$, although such effects can never be identified. For instance, if $\delta(B) = 1 - B^4$ then a basis of functions composing the null space consists of 1 , i^t , $(-1)^t$, and $(-i)^t$ with $i = \sqrt{-1}$. This follows from the theory of ordinary difference equations, discussed in Chapter 5 of McElroy and Politis (2020). An individual latent process may contribute some (real) polynomial factors to $\delta(B)$. In the course of modeling a time series, we may find that the innovation variance of one latent process is very close to zero (relative to the overall scale of the data). In such a case, the stochastic process has

limiting form as the innovation variance tends to zero given by only the fixed time-varying mean function in the null space; in practice, we can replace such a latent process with the corresponding fixed effect.

For example, consider a quarterly seasonal process with differencing operator $1 + B + B^2 + B^3$. If this seasonal process were found to have innovation variance close to zero, we could eliminate this from the model and instead enter a fixed effect corresponding to the null space of $1 + B + B^2 + B^3$, i.e., i^t , $(-1)^t$, and $(-i)^t$. Because it is awkward to do regression with complex numbers – and using the fact that complex roots will always come in conjugate pairs if we are considering a real-coefficient polynomial – we can instead work with cosines and sines. In particular, if $e^{i\omega}$ (for $0 < \omega < \pi$) is a root of a differencing polynomial that is being eliminated from the model due to low innovation variance, then $e^{-i\omega}$ is also a root and a regression effect of the form

$$c e^{i\omega t} + \bar{c} e^{-i\omega t}$$

is in the null space, for a complex regression parameter c . However, this can be rewritten as

$$\beta_1 \cos(\omega t) + \beta_2 \sin(\omega t)$$

with $\beta_1 = c + \bar{c}$ and $\beta_2 = i(c - \bar{c})$, which are both real parameters. As a result, we can simply introduce the regressors $\cos(\omega t)$ and $\sin(\omega t)$ into the model. If instead the root of the differencing polynomial is -1 , we introduce the regressor $(-1)^t$.

We make a parenthetical remark here about levels and rates. All regression effects are specified for levels (the original, undifferenced data process), and when moving to rates (or a series where $1 - B$ has been applied) the regressors are also differenced in a like manner, so the interpretation of the regression coefficients is unaltered. Sometimes economists enter fixed effects into data that has already been differenced, but don't difference the regressors; then the resulting parameter estimates no longer correspond to that particular fixed effect as it is manifested in levels. Since this is confusing (one can find transformations for the regression parameters), we avoid this and always specify the model for the time series (in levels) before any differencing has been applied.

So far, we have discussed the incorporation of fixed effects that a user thinks may be appropriate. Because a time series after applying full differencing will rarely have mean zero, even if it is stationary, it is prudent to include a constant regressor for each component in the equations for $\delta(B)X_{tj}$. But what regression effect does this imply for the original series $\{X_{tj}\}$? Again, any regressors in the null space of $\delta(B)$ cannot be identified. But we associate the constant regressor for the differenced series with a polynomial trend effect in the original series, which is explained as follows. We say a unit root of order d is present in $\delta(B)$ if $(1 - B)^d$ is a factor, where $d \geq 0$. The action of $(1 - B)^d$ on a polynomial of order d reduces it to a constant (Chapter 3 of McElroy and Politis (2020)). Hence we can insert the regressor t^d in the model, and its coefficient is proportional to the mean effect. In particular, if $\delta(B) = (1 - B)^d \delta_S(B)$ is the factorization, then

$$\delta(B) [\beta t^d] = \delta_S(B) d! \beta = \delta_S(1) d! \beta.$$

The last equality follows from the fact that passing a constant through a filter yields the sum of the filter coefficients; $\delta_S(1) \neq 0$ because 1 is not a root of $\delta_S(B)$, as all the unit roots were already accounted for in the factorization of $\delta(B)$. Hence the mean is $\delta_S(1) d! \beta$, and the coefficient of the trend polynomial equals the mean effect divided by $\delta_S(1) d!$. In the case that $d = 0$, `sigex.meaninit` allows the user to input a trend polynomial of any desired order, but if $d > 0$ then this input is ignored and only the regressor t^d is incorporated.

3.1.4 Fixed and Free Parameters

It is possible to impose general linear constraints on the pre-parameters ψ . For example, one can fix specific elements of ψ to a number, such as zero, or one can impose that a linear combination of components equals a constant. We can enforce that two components of ψ take the same value by imposing a linear constraint consisting of 1 and -1 as the combination, and 0 as the value that the linear combination is enforced to equal. In general these constraints are encoded with the equation

$$C \psi = b, \tag{7}$$

where b is a vector with length given by the number of constraints. So each row of C corresponds to one linear constraint, providing the linear combinations of ψ . The row dimension of C must be less than the column dimension, and the difference between these is the number of free variables.

A limitation of this framework is that constraints are enforced on pre-parameters instead of the original, interpretable parameters; in applied work a researcher might be interested in enforcing a constraint on the parameters, but apart from the regression parameters it is difficult to see how elements of ψ should be bounded to yield such a constraint. Because the mapping of regression parameters to the pre-parameter is the identity (they are mapped to the subvector β

of ψ), this case is easily handled. Otherwise, constraints for ξ or ζ could arise from a prior fit of the model where it was subsequently determined that certain pre-parameters were not significantly different from zero.

Here we determine a real vector of free variables, called η , and a real vector of constrained variables ν , such that together η and ν make up the vector ψ . Then likelihood evaluation is in terms of η , and this is the variable that is optimized over for maximum likelihood estimation. To determine η and ν from a given ψ , and vice versa, consider the QR decomposition (Golub and Van Loan, 2012) of the matrix C :

$$C = Q R \Pi,$$

where Q is an orthogonal matrix, R is unit upper triangular (with dimensions the same as those of C), and Π is a permutation matrix. Let $R = [R_1 | R_2]$ be partitioned such that R_1 is square; it is invertible because it is upper triangular with ones on the diagonal. Then we define η to be the bottom portion of $\Pi \psi$, and we solve for ν in terms of η as follows:

$$\nu = R_1^{-1} Q^{-1} b - R_1^{-1} R_2 \eta. \quad (8)$$

Finally, ψ is constructed from η and ν via

$$\psi = \Pi^{-1} \begin{bmatrix} \nu \\ \eta \end{bmatrix}. \quad (9)$$

It follows that (using 1 and 0 to denote either an identity matrix or a zero matrix of appropriate dimension)

$$\psi = \Pi^{-1} \left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} R_1^{-1} R_2 \right) \eta + \Pi^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} R_1^{-1} Q^{-1} b, \quad (10)$$

which expresses ψ as $A\eta$ plus a constant vector. These calculations are performed by `sigex.eta2psi`, and `sigex.psi2eta` produces η and ν from a given ψ . Default values of *param* are determined through `sigex.default`, which occurs by setting η equal to the zero vector and computing ψ via (8) and (9). This requires the matrix *constraint* to be defined, which has the format $[b|C]$. This can be set to NULL if there are no constraints.

The model can then be fitted using `sigex.mlefit` and the output stored to a variable *fit*. Setting the last argument to “bfgs” tells the computer to use the BFGS (Golub and Van Loan, 2012) algorithm for nonlinear optimization of the Gaussian likelihood, based on initial values dictated by *param*, but fixing parameters according to *constraint*, and an overall model structure given by *mdl*. However, we begin by obtaining ψ from *param*, and ψ should satisfy (7); the code checks this condition. At the end of fitting a model, we will typically have the MLE for η (denoted *eta.mle*). Then a call to `sigex.eta2psi` will yield *psi.mle*, and we can apply `sigex.psi2par` to obtain *par.mle*.

One can also extract the Hessian matrix if BFGS was used for model fitting. Note that uncertainty for the model parameters only pertains to η , so the dimension of *hess* is given accordingly. The inverse of *hess* is the asymptotic covariance matrix V of the MLEs for η ; to get the asymptotic covariance matrix for ψ , we compute $A V A'$, where A is the matrix described below (10).

3.2 Method of Moments

The goal of the MOM procedure is to obtain quick, rough estimates that might be used as initial values in the ML routine, or alternatively as final estimates for higher-dimensional time series data where ML is not computationally feasible. Background for MOM for latent process models is given in McElroy and Roy (2021a), although some more specific implementation details are discussed here. If $\Psi^{(k)}(z)$ in (4) is known, then by applying (5) we can express the autocovariances of $\{Y_t\}$ in terms of certain known matrix polynomials and the covariance matrices $\Sigma^{(k)}$; by substituting sample autocovariances, we can solve for estimators of the covariance matrices. These are called the MOM estimates.

In practice several provisos mitigate the usefulness of this device. First, the presence of fixed effects interferes with the population equations. Second, the solutions to the equations yield symmetric matrices, but these are not guaranteed to be positive definite (pd). Third, there may be model parameters involved in the $\Psi^{(k)}(z)$. Fourth, missing values in the data generate obstacles to constructing the sample autocovariances.

To handle the first obstacle, we begin with a simple OLS in `sigex.momfit`, which is fast to compute and consistent notwithstanding serial correlation present in $\{Y_t\}$. Subtracting off the fitted regression effects, we proceed and now presume the de-meaned series has no fixed effects present at all. We can then apply the MOM equations so long as there are no missing values (the routine will not work if there are NAs); if parameters ζ are present and enter into the $\Psi^{(k)}(z)$, these are fixed at the values dictated by the default settings.

The second obstacle is important to rectify. In **Ecce Signum** a covariance matrix can be assessed in terms of condition numbers discussed in McElroy (2017), which are direct functions of the diagonal entries of D in (6). In fact, it can be

shown that the condition numbers equal D_{jj}/Σ_{jj} , and they are typically viewed in log scale. After taking the logarithm, a value of $-\infty$ corresponds to a singularity in Σ that occurs at a particular index in the rank configuration (certain partial correlations of Σ are equal to ± 1 in this case). Moreover, negative values of large magnitude (in log scale) indicate the matrix is close to singularity; it is possible to alter the condition numbers to positive quantities such that the entries of D are altered and the whole matrix becomes better conditioned. This can be done without altering the various partial correlations embedded in L , and is implemented in `sigex.renderpd`, as shown below.

Let a j -dimensional covariance matrix be denoted Σ_j , which can be decomposed as $\Sigma_j = L_j D_j L_j'$ for a unit lower-triangular L_j and a diagonal matrix D_j . Letting the lower row of L_j be denoted $[\ell_j', 1]$, it is known (McElroy, 2017) that

$$g_{j+1} = d_{j+1} + \ell_{j+1}' D_j \ell_{j+1},$$

where g_{j+1} denotes the lower right entry of Σ_{j+1} . Moreover, the $(j+1)$ th condition number can be expressed as

$$\tau_{j+1} = -\log(g_{j+1}/d_{j+1}) = -\log(1 + \ell_{j+1}' D_j \ell_{j+1}/d_{j+1}).$$

These are recursive relations, so that for a given N -dimensional Σ we can examine the upper left submatrix Σ_j for $1 \leq j \leq N$ and recursively compute the condition numbers τ_j . Hence, if we wish to modify a given Σ such that the condition numbers exceed some minimum threshold, then we must modify either ℓ_{j+1} or D_j . Whereas the former is related to linkages between variables, each d_j is a partial variance (of the j th variable given the preceding variables 1 through $j-1$); choosing to modify these to some new \tilde{d}_j , we can see how the condition number changes accordingly. Thus, if we set τ_{j+1} equal to a desired threshold α , and given that we have already modified in a recursive fashion the preceding D_j to \tilde{D}_j , we have the formula

$$\tilde{d}_{j+1} = \left(\frac{\exp\{-\alpha\} - 1}{\ell_{j+1}' \tilde{D}_j \ell_{j+1}} \right)^{-1}. \quad (11)$$

This suggests the following algorithm to produce a positive definite modification of a given Σ matrix: compute the GCD, and iteratively check whether $1/d_{j+1} < (\exp\{-\alpha\} - 1)/\ell_{j+1}' \tilde{D}_j \ell_{j+1}$; if so, we replace d_{j+1} by \tilde{d}_{j+1} in (11) and increment j . This yields a new

$$\tilde{\Sigma} = L \tilde{D} L',$$

where all the condition numbers are bounded below by α . In practice, the thresholds for α can be set according to the criteria discussed in McElroy (2017). For instance, $\alpha = -1.66$ corresponds to a partial correlation of .90; this is sufficiently distinct from unity, so as to ensure there is no signal leakage. Calling `sigex.reduce` with a specified value of α (denoted by *thresh* in the code) replaces low logged condition numbers with *thresh*, rendering the matrix pd through `sigex.renderpd`.

However, this discussion pertains to matrices that are non-negative definite. The covariance matrices estimated by MOM can potentially have both positive and negative eigenvalues. In this case a second usage of `sigex.reduce` occurs (setting *modelFlag* equal to *TRUE*) whereby logged condition numbers lower than *thresh* result in the corresponding component index being added to the rank configuration. In other words, a low condition number for component j will result in j being added to the rank configuration. Now the GCD algorithm will still run on a symmetric matrix where the entries of D are possibly zero or negative; in such a case `sigex.reduce` will automatically assign a non-positive value of D to being below the threshold, and hence the rank configuration will be incremented. The outcome for MOM is that matrices that are not estimated as being pd get replaced by reduced rank approximations – essentially any negative values of the diagonal of D get replaced by a zero.

Once a reduced rank pd covariance matrix has been obtained, one can then determine *param* accordingly. However, the process of reducing the rank changes the model structure, and hence `sigex.reduce` updates *mdl*. This is a side-effect of MOM estimation, and may be undesirable in some applications (say, where we wish to keep all the covariance matrices as being full rank initially).

3.3 Casting and Fixed Effects

The basic model structure (1) yields a $\{Y_t\}$ time series that is mean zero. Optimal extraction means that we should filter $\{Y_t\}$ after having removed fixed effects, and then add back those fixed effects that pertain to the particular signal. However, in practice we do not know β exactly, and we instead subtract $z_t' \hat{\beta}$ from the data to get the de-meanned data:

$$\hat{Y}_t = X_t - z_t' \hat{\beta}.$$

If there are missing values, or extreme-value adjustment is being performed for an AO, then there is already an NA in \hat{Y}_t , and this will be estimated with a midcast. Importantly, the casting algorithms are predicated on the time series

having mean zero, which is approximately true of \widehat{Y}_t . Let \widetilde{Y}_t denote the modification arising from casting, i.e., imputing missing values and adjusting extremes. The discrepancy between \widetilde{Y}_t and \widehat{Y}_t is

$$\varepsilon_t = \widehat{Y}_t - \widetilde{Y}_t,$$

which is also the casting error. It is important to keep track of this quantity, because in many applications of signal extraction we want all the extracted latencies to exactly add back up again to the original data. The form this typically takes is the following:

$$\begin{aligned} X_t &= \widehat{Y}_t + z'_t \widehat{\beta} \\ &= \widetilde{Y}_t + \varepsilon_t + z'_t \widehat{\beta} \\ &= \Upsilon(B) \widetilde{Y}_t + (1 - \Upsilon(B)) \widetilde{Y}_t + \varepsilon_t + z'_t \widehat{\beta}. \end{aligned}$$

In the first line we break out the estimated fixed effects. In the second line we correct for missing values and extreme values. In the last line we apply two complementary filters – such as seasonal adjustment and seasonal filtering. The casting error must be added to one of these two extractions in order that all sums back to X_t . The fixed effects might be split up between the extractions. For example, if $\Upsilon(B)$ is a seasonal adjustment filter (so that $1 - \Upsilon(B)$ is a seasonal filter), then the casting error for extreme-value adjustment should be added to the seasonal adjustment (because it corresponds to an outlier, and is therefore not seasonal); fixed effects that correspond to moving holidays can be added to the seasonal component, but the trend polynomial should be combined with the seasonal adjustment.

The only case where this breaks down is if the original date is missing, i.e., X_t is NA. In such a case, we replace the left hand side with its imputation – essentially by subtracting the casting error from both sides of the equation, along with the fixed effects. For instance, if X_t is missing but we wish to publish a seasonal adjustment, the sum of the adjustment and seasonal factor equals the imputation of X_t given by $\widetilde{Y}_t + z'_t \widehat{\beta}$.

3.4 Embedding

A vector embedding of a univariate process consists of rewriting the process as an N -variate time series of lower sampling frequency, such that the new sampling frequency equals the original sampling frequency divided by N . Sometimes embedding facilitates the modeling of a scalar time series. For example, a univariate daily time series might be embedded as a weekly time series with $N = 7$, so that day-of-week specific effects can be modeled in the mean or variance directly. The function `sigex.daily2weekly` provides such an embedding given a specification of the day-of-week corresponding to the first component in the vector, and adds missing values to pad out the possibly incomplete first and last weeks. An inverse is provided by `sigex.weekly2daily`, which can be used on extracted components obtained from the weekly series.

3.4.1 Embedding Filters

Scalar filters can also be embedded, turning them into a multivariate filter with Toeplitz structure. We can also apply embedding results to the moving average representation of a scalar series, allowing us to re-express it as a vector moving average. We can embed a given high-frequency scalar process $\{X_t\}$ as an s -variate (where s is a positive integer) low-frequency process $\{\mathbf{X}_n\}$, where the j th component (for $1 \leq j \leq s$) of \mathbf{X}_n is X_{ns+j} . This j th component is denoted as the j th *season's series*, and the vector time series is the *seasonal vector series*, which is denoted with a bold font. Gladyshev (1961), Tiao and Grupe (1980), Osborn (1991), and Franses (1994) have studied this type of embedding.

Define the backshift operator B , which acts on the high-frequency time index, via $BX_t = X_{t-1}$. Similarly, the low-frequency lag operator is defined to be $L = B^s$, so that $L\mathbf{X}_n = \mathbf{X}_{n-1}$. The $s \times s$ identity matrix is denoted I_s . Given a scalar Laurent series $v(z)$, its embedded matrix Laurent series (cf. equation (2.7) of Tiao and Grupe (1980)) is defined to be the $s \times s$ -dimensional matrix Laurent series $\Upsilon(L) = \sum_h \Upsilon_h L^h$, where

$$\text{the } j, k \text{ coefficients of } \Upsilon_h \text{ are given by } v_{j-k+sh}. \quad (12)$$

It follows that if $v(B)$ is a polynomial of degree q , then the degree of the matrix polynomial $\Upsilon(L)$ is $\lfloor q/s \rfloor + 1$. Consider the following relations of filter input to filter output (expressed as operating on the de-measured process):

$$\begin{aligned} U_t &= v(B) Y_t \\ \mathbf{U}_n &= \Upsilon(L) \mathbf{Y}_n. \end{aligned}$$

We verify that these two equations are equivalent, if we utilize the rule (12). The j th component (for $1 \leq j \leq s$) of the second equation (with e_j the j th unit vector) indicates

$$\mathbf{U}_{n,j} = e'_j \sum_{h \in \mathbb{Z}} \Upsilon_h \mathbf{Y}_{n-h} = \sum_{h \in \mathbb{Z}} \sum_{k=1}^s v_{j-k+sh} Y_{(n-h)s+k} = \sum_{\ell \in \mathbb{Z}} v_\ell Y_{ns+j-\ell} = v(B)Y_t,$$

which uses the change of variable $\ell = j - k + s * h$.

Next, consider the case of a finite filter $v(B)$ that has a total of m coefficients, and such that v_0 is the $(c+1)$ th coefficient, where $0 \leq c < m$ is the integer shift parameter. (In the documentation for `sigex.hi2low` and `sigex.adhocextract`, L is used for the filter length.) This means that the non-zero coefficients are $v_{-c}, v_{1-c}, \dots, v_0, \dots, v_{m-1-c}$. So $c = 0$ corresponds to a causal filter, where only present and past observations get weighted; if m is odd and $c = (m-1)/2$, then the filter weights an equal number of past and future observations. The function `sigex.hi2low` calculates the corresponding non-zero filter matrices Υ_h , and determines the appropriate shift parameter for the low-frequency multivariate filter. From (12), we have

$$\dots, \Upsilon_{-1} = \begin{bmatrix} v_{-s} & \dots & v_{1-2s} \\ \vdots & \ddots & \vdots \\ v_{-1} & \dots & v_{-s} \end{bmatrix}, \Upsilon_0 = \begin{bmatrix} v_0 & \dots & v_{1-s} \\ \vdots & \ddots & \vdots \\ v_{s-1} & \dots & v_0 \end{bmatrix}, \Upsilon_1 = \begin{bmatrix} v_s & \dots & v_1 \\ \vdots & \ddots & \vdots \\ v_{2s-1} & \dots & v_s \end{bmatrix}, \dots$$

We need to determine the length M of the multivariate filter, and its shift parameter C . Clearly, if any Υ_h has at least one non-zero entry then it must be included in the array object for the multivariate filter. We begin by padding v_{-c}, \dots, v_{m-1-c} with ℓ zeroes on the left and r zeroes on the right, such that: (i) $\ell + c \equiv 0 \pmod{s}$ and $0 \leq \ell < s$, (ii) $m + \ell + r \equiv 0 \pmod{s}$, and such that r is the smallest integer $\geq s$ with this property.

First, (i) guarantees that the number of coefficients v_k with $k < 0$ is divisible by s , and is the smallest such number. Examining the bottom (s th) row of the Υ_h matrix, we have the values $v_{sh+s-1}, \dots, v_{sh}$, so if all $v_k = 0$ for $k < -(\ell+c)$, then $\Upsilon_h = 0$ if $h < -(\ell+c)/s$. This means that the non-zero matrices are $\Upsilon_{-(\ell+c)/s}, \dots, \Upsilon_{M-1-(\ell+c)/s}$, with M to be determined. Hence the low-frequency shift parameter is $C = (\ell + c)/s$.

Second, (ii) ensures that the entire length of the padded scalar filter is divisible by s , given that with the padding from (i) the filter is currently of length $\ell + m$. We need to add at least s zeroes on in this stage so that the bottom row of Υ_h with highest index is zero except possibly for the bottom right entry v_{sh} ; if v_{sh+1} were non-zero, then it would appear in Υ_{h-1} , indicating that the matrix is non-zero. It follows that $M = (m + \ell + r)/s$.

3.4.2 Embedding From Daily to Weekly

We discuss some of the details on the weekly embedding. Dates are stored in a month-day-year format as a three-component vector, and the utility `date2day` computes the day index (between 1 and 366) corresponding to any such date. The inverse of this routine is `day2date`, and `day2week` yields the day-of-week (1 for Sunday, 2 for Saturday, etc.) corresponding to a given calendar date. To compute the week index corresponding to a given date, `sigex.daily2weekly` uses the following procedure: the day index is first computed from the date, and we determine the day index corresponding to the first day of the first week. Specifically, between 0 and 6 NAs are pre-pended to form the first week, with this number of NAs determined by what day-of-week corresponds to the calendar date, and what type of day corresponds to the first component (this is called *first.day*, and is chosen by the user). Let ℓ denote the day index for the first day of the first week in the embedded series; note that ℓ can be negative (e.g., the calendar date is very early in the year, and we pad with enough NAs). Then the day indices for that first week are $\{\ell, \ell+1, \dots, \ell+6\}$. Our rule is that these indices corresponds to week w of that year if and only if $7(w-1) + 1 \in \{\ell, \ell+1, \dots, \ell+6\}$. The logic is that if $1 \in \{\ell, \ell+1, \dots, \ell+6\}$ then this must be the first week (as it includes the first day index), and hence $w = 1$; thus the formula follows. An equivalent condition is that

$$w = \lceil (\ell - 1)/7 \rceil + 1.$$

The lowest value that ℓ can take is -5 , in the case that 6 NAs were padded on and the calendar date is the first day of the year; in that case $w = 1$, so we always get $1 \leq w \leq 53$. For example, January 1, 2020 was a Wednesday. Suppose we wish to have Saturday as our first series, but our time series has a start date of February 28, 2020. Hence, there will be 6 NAs added onto the first week; the day index for the start date is 59, so $\ell = 53$ and $w = 9$. The first “custom” week covers December 28, 2019 through January 3, 2020, and the ninth custom week includes February 22, 2020 through February 28, 2020. Note that week one will typically have a few days from the prior year; here the start year for the weekly series is 2020, even though four days of the first week are from 2019.

When passing back to a daily series via `sigex.weekly2daily`, the original start date can be recovered as follows. Compute the difference between the day-of-week of January 1 of the start year (for the weekly series) and the given

first.day, and add seven if this difference is negative; the result, called *day.lead*, gives the number of days prior to January 1 that are included in the first week, and hence

$$\ell = 7(w - 1) - \text{day.lead} + 1.$$

From this day index ℓ , we can determine the corresponding start date via `day2date`; however, if $\ell \leq 0$ (this can happen if $w = 1$) then we need to decrement the year by one and add to ℓ the day index of December 31 for the prior year (i.e., either 365 or 366). If we started by embedding a daily time series of a certain start date, use `sigex.daily2weekly`, and then invert with `sigex.weekly2daily`, we will not recover the exact start date unless *first.day* corresponds to the day-of-week of that calendar date – because otherwise, additional NAs are added and the daily series gets lengthened (so that the total sample size is a multiple of 7).

4 Illustrations

We consider several numerical illustrations of the capabilities. These are meant to be consecutive, in the sense that commands documented in a prior example are not explicitly re-visited in subsequent examples.

4.1 Non-Defense Capital Goods

We first examine data for monthly shipments and new orders from the Manufacturing, Shipments, and Inventories survey,⁴ or *ndc* for short. The data for New Orders is not available at January 1992, since this series starts at February 1992 – so this value is entered as an NA. This is an example of ragged edge data.

We perform a bivariate analysis by fitting a VAR model. After loading the data into a variable named *ndc*, preliminary metadata is assigned as follows:

```
dataALL.ts <- sigex.load(data      = ndc,
                        start.date = c(1992, 1),
                        period     = 12,
                        epithets   = c("Shipments", "NewOrders"),
                        plot       = TRUE )
```

These commands attach a starting date and frequency, as well as names for the series. The last argument of `sigex.load` generates a time series plot. On the basis of this plot, we decide (below) to utilize no transformation for the data.

Next, data transformations and subsets of series can be selected. Here, we will examine both series, so *subseries* is set equal to `c(1,2)`. The Boolean *aggregate* can be used to define a cross-sectional sum corresponding to *subseries*, but in this example that option is not utilized. The *range* argument can be used to select a window of times, but is set to NULL if the full span of the data is to be used.

```
data.ts <- sigex.prep(data.ts. = dataALL.ts,
                     transform = "none",
                     aggregate = FALSE,
                     subseries = 1:2,
                     range     = NULL,
                     plot      = TRUE )
```

Next, we move to model declaration, first recording the dimensions via

```
N <- dim(data.ts)[2]
T <- dim(data.ts)[1]
```

We will fit a VAR model to the differenced data, and we use Yule-Walker estimation to identify the model order p and get preliminary estimates.

```
ar.fit    <- ar.yw(diff(ts(ndc[2:T, ])))
p.order   <- ar.fit$order
```

⁴Seasonally adjusted monthly data covering the sample period is January 1992 to April 2020, downloaded July 21, 2020 (4:45 PM), U.S. Census Bureau, obtained from <https://www.census.gov/mtis/index.html> by selecting Non-Defense Capital Goods, and either Value of Shipments or New Orders.

```

par.yw    <- aperm(ar.fit$ar, c(2, 3, 1))
covmat.yw <- getGCD(ar.fit$var.pred, 2)
var.out   <- var.par2pre(par.yw)
psi.init  <- as.vector(c(covmat.yw[[1]][2, 1], log(covmat.yw[[2]]),
                        var.out, colMeans(diff(ts(ndc[2:T, ])))) )

```

The first line fits a VAR model using the native R function `ar.yw`. Note that we exclude the first observation because it is missing for the `NewOrders` series. We also apply differencing, as there seems to be some change in level for the series. The second line extracts the identified model order ($p = 7$ in this case), and the third line stores the fitted parameter in an array of dimension $2 \times 2 \times 7$. The fourth line takes the estimated innovation covariance matrix, and produces the GCD via `getGCD`. The output is a list object, where the lower triangular entry of L and the log of the diagonal entries of D are entered into the appropriate positions of `psi.init` in line seven. Line five maps the array of VAR parameters via `var.par2pre` to a pre-parameter vector `var.out` of length 28. Lines six and seven assemble `psi.init` from the Yule-Walker fit, including the sample means for the last two pre-parameters.

Next, we declare the model. In this example there are no regressors specified, and there is only one latent process:

```

mdl <- NULL
mdl <- sigex.add(mdl      = mdl,
                vrank    = seq(1,N),
                class     = "varma",
                order     = c(p.order,0),
                bounds    = NULL,
                name      = "process",
                delta     = c(1,-1) )

# regressors:
mdl <- sigex.meaninit(mdl = mdl, data.ts = data.ts, d = 0)

```

In the second line the main process is defined through a call to `sigex.add`. The list object `mdl` stores all our specifications about the model, but none of the parameter estimates. The second argument indicates the rank configuration for the trend; setting `vrank` equal to `seq(1,N)` ensures a full rank covariance matrix. The third argument gives the model type as `varma`; the order $p = 7$, $q = 0$ of the VARMA is given in the fourth argument. The fifth argument provides parameter bounds, and does not apply to this type of model, and hence is set to `NULL`. The sixth argument associates a label, and the last argument is for the scalar differencing operator $\delta^{(1)}(z)$, expressed as a vector of coefficients `c(1, -1)`.

Now we move on to the next stage: model fitting. Here we demonstrate MLE using BFGS. We begin by setting up parameter constraints; there are none in this case.

```

constraint <- NULL
psi.mle    <- psi.init
par.mle    <- sigex.psi2par(psi = psi.mle, mdl = mdl, data.ts = data.ts)

```

The second line assigns `psi.mle` to the initial values of the pre-parameters – this will speed up optimization. The third line converts this pre-parameter to the corresponding value of `param` via `sigex.psi2par`. Next, we run the MLE routine as follows:

```

fit.mle <- sigex.mlefit(data.ts    = data.ts,
                       param      = par.mle,
                       constraint   = constraint,
                       mdl         = mdl,
                       method      = "bfgs",
                       debug       = TRUE )

```

With the debug option set to `TRUE`, the values of the divergence are printed to the console; this is a way to track progress and understand what parts of the likelihood function are being searched. Setting debug to `FALSE` turns off this screen output. When the routine finally terminates (this can take a few minutes, depending on the machine), we manage the output using

```

psi.mle <- sigex.eta2psi(eta = fit.mle[[1]]$par, constraint = constraint)
hess    <- fit.mle[[1]]$hessian
par.mle <- fit.mle[[2]]

```


The first line ensures that if constraints have been used, the final η is mapped to the corresponding ψ via `sigex.eta2psi`. In this case $\eta = \psi$ because there are no constraints. We also extract the Hessian matrix and the *param* corresponding to ψ in the second and third lines. Next, we can check our model fit by examining the residuals.

```
resid.mle <- sigex.resid(psi = psi.mle, mdl = mdl, data.ts = data.ts)[[1]]
resid.mle <- sigex.load(data      = t(resid.mle),
                      start.date = start(data.ts),
                      period     = frequency(data.ts),
                      epithets   = colnames(data.ts),
                      plot       = TRUE )
resid.acf <- acf(x = resid.mle, lag.max = 4*period, plot = TRUE)$acf
```

The first line extracts the residuals corresponding to the MLE. We want the first item in the list output of `sigex.resid`; the second item is an array of the fitted model's autocovariances. The second call to `sigex.load` does formatting, and the last line computes sample autocorrelations of the residuals. Next, we can examine the condition numbers via

```
log(sigex.conditions(data.ts = data.ts, psi = psi.mle, mdl = mdl))

      [,1]      [,2]
eta2      0 -0.1349298
```

The results indicate that the innovation covariance matrix is well-conditioned. The model goodness-of-fit can be checked with

```
sigex.portmanteau(resids = resid.mle, lag = 4*period, nump = length(psi.mle))

[1] 165.1150040    0.3534992
```

This code calls `sigex.portmanteau`, which yields the Portmanteau statistic for residual serial correlation discussed in Lütkepohl (2005); the second argument gives the number of lags to be used (48 in this case), and the third argument is the number of estimated parameters. The output indicates there is little residual autocorrelation, which is confirmed by inspection of the sample autocorrelations. The call to `sigex.gausscheck` yields p-values for both series from the Shapiro-Wilks normality test; although there is some non-normality indicated, this does not invalidate the model.

```
sigex.gausscheck(resids = resid.mle)

[1] 1.211720e-07 2.548542e-20
```

We can examine the Hessian matrix, and compute the t statistics for pre-parameters via

```
print(eigen(hess)$values)
tstats <- sigex.tstats(mdl,psi.mle,hess,constraint)
print(tstats)
```

If the Hessian is not positive definite (this can happen if nonlinear optimization terminates wrongfully at a saddlepoint), then the standard error is set to zero, and `sigex.tstats` returns $\pm\infty$ for each t statistic. In this case the Hessian is positive definite, and some of the pre-parameters are not significantly different from zero – we could refine the model by imposing zero constraints. Finally, we bundle the analysis with the call

```
analysis.mle <- sigex.bundle(data.ts = data.ts,
                           transform = "none",
                           mdl       = mdl,
                           psi       = psi.mle )
```

As an application, we generate forecasts and afts with uncertainty; also a midcast at time $t = 1$ is generated for the New Orders series, since this value is missing. Typically we unload the bundled model via

```
data.ts <- analysis.mle[[1]]
mdl     <- analysis.mle[[3]]
psi     <- analysis.mle[[4]]
param   <- sigex.psi2par(psi = psi, mdl = mdl, data.ts = data.ts)
```

This amounts to a relabeling. Fifty forecasts and afts are generated via

```
window <- 50
data.casts <- sigex.midcast(psi      = psi,
                           mdl      = mdl,
                           data.ts  = data.ts,
                           castspan = window )
extract.casts <- sigex.castextract(data.ts  = data.ts,
                                  data.casts = data.casts,
                                  mdl      = mdl,
                                  castspan = window,
                                  param    = param )
```

The output of `sigex.midcast` is a list with two items: the matrix of casts, and a 4-array of casting error covariances. The t th column of the cast matrix contains a vector of casts for the t th index in the time series that has any missing values; this takes into account the specified number of afts and forecasts with the `castspan` variable. Also note that these casts are for the series with all fixed effects removed. This output can be formatted for use via `sigex.castextract`, which inserts the casts in the correct places in the time series and adds an extended trend regression effect (it is extended forward and backward by the `castspan` value). The output of `sigex.castextract` is a list with three items, in the same format as for `sigex.extract` output; therefore the output can be used by `sigex.graph`:

```
dataPad.ts <- rbind(matrix(NA, nrow = window, ncol = N),
                   data.ts,
                   matrix(NA, nrow = window, ncol = N))
for(i in 1:N){
  plot(x      = ts(dataPad.ts[, i], start = c(1992, 1), frequency = 4),
       xlab = "Year",
       ylab = "",
       lwd  = 1,
       col  = 1)
  sigex.graph(extract  = extract.casts,
              reg      = NULL,
              start.date = c(1992,1),
              period    = 4,
              series    = i,
              displace  = 0,
              castcol   = "black",
              fade      = 60 )
}
```

This code sets up grey shading for cast uncertainty in `castcol` and `fade`. For each component time series, the raw data is plotted after first padding with NA before and after. The first argument of `sigex.graph` takes the extracted list object `extract.casts`; other fixed effects could be added in the second argument (these are typically not known at the times covered by the fore and aft windows, unless the regressors are simple to extrapolate). The sixth argument allows for a vertical displacement in case one does not wish to overlay the extraction on top of the data – setting to zero indicates no displacement. The last argument determines the darkness of shading corresponding to the signal extraction uncertainty. This concludes the example.

4.2 Petrol

We next examine two monthly petroleum time series: “Industrial Petroleum Consumption and OPEC Oil Imports,” or *petrol* for short.⁵ We perform a bivariate analysis by fitting two simple models, both of which are based on a trend-irregular specification referred to as the Local Level Model (LLM) in Harvey (1989). Here there are no regressors specified, and $K = 2$, with $\{S_t^{(1)}\}$ corresponding to a trend process and $\{S_t^{(2)}\}$ a so-called irregular process, which is a

⁵Source: Department of Energy, Energy Information Administration. Both series cover the period January 1973 through December 2016, and are seasonally adjusted (by DOE). The first series, “Industrial Petroleum Consumption,” is U.S. Product Supplied of Crude Oil and Petroleum Products, Monthly, in Thousand Barrels per Day. The second series, “OPEC Oil Imports,” is Petroleum Imports from OPEC, in Thousand Barrels per Day.

bivariate white noise. The trend is a bivariate random walk, i.e., $\delta^{(1)}(z) = (1 - z)$ with $\{\mathcal{S}_t^{(1)}\}$ a white noise process. These same series were analyzed with this model in McElroy and Wildi (2020).

After loading the data into a variable named *petrol*, the preliminary metadata is assigned as discussed in the prior example; a log transformation is used. We perform exploratory analysis by generating spectral density estimates; the function `sigex.specar` generates vertical lines at frequencies that are multiples of 2π divided by the *period* argument. The second argument of `sigex.specar` is a Boolean variable, indicating whether the data should be first differenced, i.e., apply $1 - B$ to each series. Here we set *period* to 12 in `sigex.specar` because we are interested in the monthly frequencies:

```
par(mfrow = c(2, 1))
for(i in subseries){
  sigex.specar(data.ts = data.ts,
               diff    = FALSE,
               subseries = i,
               period   = period)
}
```

The basic LLM is referred to as a “related trends model,” in contrast to a “common trends model” considered subsequently, where it is enforced that $\Sigma^{(1)}$ has rank one. Consider the code

```
mdl <- NULL
mdl <- sigex.add(mdl = mdl,
                 vrank = seq(1, N),
                 class = "arma",
                 order = c(0, 0),
                 bounds = NULL,
                 name = "trend",
                 delta = c(1, -1)
                )
mdl <- sigex.add(mdl = mdl,
                 vrank = seq(1, N),
                 class = "arma",
                 order = c(0, 0),
                 bounds = NULL,
                 name = "irregular",
                 delta = 1
                )
mdl <- sigex.meaninit(mdl = mdl, data.ts = data.ts, d = 0)
```

The first call to `sigex.add` defines the first latent process, labeled as “trend,” and the second such call gives the “irregular” latent process. Both latent processes have full rank. Next, the “common trends model” is specified via

```
mdl2 <- NULL
mdl2 <- sigex.add(mdl = mdl2,
                 vrank = 1,
                 class = "arma",
                 order = c(0, 0),
                 bounds = NULL,
                 name = "trend",
                 delta = c(1, -1)
                )
mdl2 <- sigex.add(mdl = mdl2,
                 vrank = seq(1, N),
                 class = "arma",
                 order = c(0, 0),
                 bounds = NULL,
                 name = "irregular",
                 delta = 1
                )
mdl2 <- sigex.meaninit(mdl = mdl2, data.ts = data.ts, d = 0)
```

where the second argument of `sigex.add` in the second line has *vrank* set to 1 instead of `seq(1, N)`, as in *mdl*.

Next, we fit the models, beginning with the related trends specification. The default parameter setting involves setting the pre-parameter vector to zero, via:

```
constraint <- NULL
par.mle <- sigex.default(mdl = mdl, data.ts = data.ts, constraint = constraint)
psi.mle <- sigex.par2psi(param = par.mle, mdl = mdl)
```

The call to `sigex.default` creates a *param* object *par.mle* corresponding to a pre-parameter *psi.mle* of all zeroes – there are 8 components in this case.

The collation and examination of results can be done in the manner described for the *ndc* example in Section 4.1; for the related trends model, results are bundled into *analysis.mle*. In the case of the common trends model, we initialize at a given ψ value obtained from the first fitted model as follows: the pre-parameters associated with the trend innovation covariance matrix are altered to correspond to a full-correlation matrix (but with the same variances). In this case

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \mapsto \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix},$$

and this mapping is accomplished by setting the lower entry of L in the GCD equal to σ_2/σ_1 , and d_1 is set equal to σ_1^2 . This is encoded with

```
cov.mat <- par.mle[[1]][[1]] %*% diag(exp(par.mle[[2]][[1]])) %*%
  t(par.mle[[1]][[1]])
l.trend <- sqrt(cov.mat[2,2]/cov.mat[1,1])
d.trend <- cov.mat[1,1]
```

Then the initialization to the common trends model is

```
constraint <- NULL
psi.mle2 <- c(l.trend, log(d.trend), psi.mle[4:8])
par.mle2 <- sigex.psi2par(psi = psi.mle2, mdl = mdl2, data.ts = data.ts)
```

After fitting and managing output, we see that the common trends model is quite a bit worse (not surprising, given that the correlation ρ is estimated at .29 in the related trends model), having a much higher divergence. We can compare the fits via

```
test.glr <- sigex.glr(data.ts = data.ts,
  psi.nested = psi.mle2,
  psi.nesting = psi.mle,
  mdl.nested = mdl2,
  mdl.nesting = mdl)
print(c(test.glr[1], 1-pchisq(test.glr[1], df = test.glr[2])))

[1] -21084.29      1.00
```

A likelihood comparison statistic (McElroy, 2016a) for nested models is computed using `sigex.glr`, and the p-value is reported in the second line. (However, this p-value is not correct when the nested model arises at the boundary of the parameter space, as in this case.) Finally, we consider trend extraction using either model. For example, to load the second model we use

```
data.ts <- analysis.mle2[[1]]
mdl <- analysis.mle2[[3]]
psi <- analysis.mle2[[4]]
param <- sigex.psi2par(psi = psi, mdl = mdl, data.ts = data.ts)
```

Here we use the matrix formula approach to signal extraction. The first step is to get the matrix filters:

```
signal.trend <- sigex.signal(data.ts = data.ts,
  param = param,
  mdl = mdl,
  sigcomps = 1 )
signal.irr <- sigex.signal(data.ts = data.ts,
```

```

param    = param,
mdl      = mdl,
sigcomps = 2 )

```

There are two latent processes (trend and irregular), and the filters for these are obtained through the two calls to `sigex.signal`; the last argument `sigcomps` of `sigex.signal` is a vector of indices denoting the composition of the desired signal. In this case, `signal.trend` is composed of just the first latent process, so the value of `sigcomps` is 1. The second step is to compute the extractions:

```

extract.trend <- sigex.extract(data.ts = data.ts,
                             filter   = signal.trend,
                             mdl      = mdl,
                             param    = param )
extract.irr    <- sigex.extract(data.ts = data.ts,
                             filter   = signal.irr,
                             mdl      = mdl,
                             param    = param )

```

In the first line, `extract.trend` is defined as a list object, with first item corresponding to a $T \times N$ matrix of trend extractions. The second and third items of the list give upper and lower bounds of confidence intervals, based on ± 2 square root signal extraction MSE. Next, it is important to re-integrate fixed regression effects, and this is done via

```

reg.trend <- NULL
for(i in 1:N) {
  reg.trend <- cbind(reg.trend, sigex.fixed(data.ts = data.ts,
                                           mdl      = mdl,
                                           series   = i ,
                                           param    = param,
                                           type     = "Trend" ) )
}

```

The function `sigex.fixed` finds a specified regressor (named in the last argument) and multiplies by the corresponding parameter estimate, for the component specified in the third `series` argument. Here, the default linear trend line regressor is passed into the variable `reg.trend`. The results can be displayed using calls to `sigex.graph`, as in the following code:

```

trendcol <- "tomato"
fade     <- 40
par(mfrow=c(2,1),mar=c(5,4,4,5)+.1)
for (i in 1:N){
  plot(x      = data.ts[,i],
       xlab   = "Year",
       ylab   = colnames(data.ts)[i],
       ylim   = c(min(data.ts[, i]), max(data.ts[, i])),
       lwd    = 2,
       yaxt   = "n",
       xaxt   = "n")
  sigex.graph(extract    = extract.trend,
              reg         = reg.trend,
              start.date  = begin.date,
              period      = period,
              series       = i,
              displace     = 0,
              color        = trendcol,
              fade         = fade)
  axis(1, cex.axis = 1)
  axis(2, cex.axis = 1)
}

```

This code sets up a red color for the trend in `trendcol`, and a shading percentage `fade`. For each component time series, the raw data is plotted (in logarithms) and overlaid with the trend; the first argument of `sigex.graph` takes the extracted list object `extract.trend`, and any fixed effects that should be added must be included in the second argument.

The sixth argument allows for a vertical displacement in case one does not wish to overlay the extraction on top of the data – setting to zero indicates no displacement. The last argument determines the darkness of shading corresponding to the signal extraction uncertainty.

Finally, one can examine the frequency response functions for WK filters through `sigex.getfrf`, which can also plot the real portions.

```
grid <- 1000
frf.trend <- sigex.getfrf(data.ts = data.ts,
                          param    = param,
                          mdl      = mdl,
                          sigcomps = 1,
                          plotit   = TRUE,
                          grid     = grid )
```

The `grid` argument corresponds to a mesh of frequencies in $[0, \pi]$, and the fourth argument of `sigex.getfrf` indicates the combination of components desired. One can also examine the WK filter coefficients:

```
len      <- 50
target   <- array(diag(N), c(N, N, 1))
wk.trend <- sigex.wk(data.ts = data.ts,
                    param    = param,
                    mdl      = mdl,
                    sigcomps = 1,
                    target   = target,
                    plotit   = TRUE,
                    grid     = grid,
                    len      = len)
```

The `target` argument indicates that no linear combination of the trend is being considered here; `len` is set to 50, so that the indices of the filter run from -50 up to 50 . This finishes the example.

4.3 Housing Starts

Here we study “New Residential Construction (1964-2012), Housing Units Started, Single Family Units,” or *starts* for short,⁶ corresponding to the four regions of South, West, NorthEast, and MidWest. This example illustrates modeling a multivariate series with 8 components, estimated with both MOM and MLE methods, and the extraction of signals with two methods (direct matrix and truncated WK). We also show how functions of a signal, such as growth rates, can be obtained.

The initial loading scripts are similar to the previous example. Exploratory analysis indicates we should consider a model with trend and seasonal effects; here we consider a structural model with $K = 8$ latent processes:

```
mdl <- NULL
mdl <- sigex.add(mdl, seq(1, N), "arma", c(0, 0), NULL, "trend", c(1, -2, 1))
mdl <- sigex.add(mdl, seq(1, N), "arma", c(0, 0), NULL, "first seasonal", c(1, -sqrt(3), 1))
mdl <- sigex.add(mdl, seq(1, N), "arma", c(0, 0), NULL, "second seasonal", c(1, -1, 1))
mdl <- sigex.add(mdl, seq(1, N), "arma", c(0, 0), NULL, "third seasonal", c(1, 0, 1))
mdl <- sigex.add(mdl, seq(1, N), "arma", c(0, 0), NULL, "fourth seasonal", c(1, 1, 1))
mdl <- sigex.add(mdl, seq(1, N), "arma", c(0, 0), NULL, "fifth seasonal", c(1, sqrt(3), 1))
mdl <- sigex.add(mdl, seq(1, N), "arma", c(0, 0), NULL, "sixth seasonal", c(1, 1, 1))
mdl <- sigex.add(mdl, seq(1, N), "arma", c(0, 0), NULL, "irregular", 1)
mdl <- sigex.meaninit(mdl, data.ts, 0)
```

Each of the eight latent processes is labeled, and has a distinct differencing polynomial, whose application produces a full-rank white noise; details of this model are discussed in McElroy (2017). Next, we illustrate the MOM to obtain initial estimates.

```
mdl.mom <- mdl
```

⁶The four series are from the Survey of Construction of the U.S. Census Bureau, available at http://www.census.gov/construction/nrc/how_the_data_are_collected/soc.html.

```

constraint <- NULL
par.default <- sigex.default mdl = mdl.mom, data.ts = data.ts, constraint = constraint)
par.mom <- sigex.momfit(data.ts = data.ts, param = par.default, mdl = mdl.mom)
psi.mom <- sigex.par2psi(param = par.mom, mdl = mdl.mom)
sigex.lik(psi = psi.mom,
          mdl = mdl.mom,
          data.ts = data.ts,
          debug = FALSE )

[1] 6329.107

```

First, we define a new model object *mdl.mom*, and the subsequent initialization is standard in the second and third lines. The MOM estimates are obtained through *sigex.momfit* in the fourth line. In the last call we evaluate the divergence, obtaining 6329.107.

Examination of the residuals indicates that a substantial degree of serial dependence remains. We can also examine the condition numbers, and observe that all the six seasonal latent processes seem to have reduced rank covariance matrices, indicating various types of seasonal co-integration. We can modify the initial model to one that has reduced rank components, by using *sigex.reduce*:

```

thresh <- -6.22
reduced.mom <- sigex.reduce(data.ts = data.ts,
                           param = par.mom,
                           mdl = mdl.mom,
                           thresh = thresh,
                           modelflag = TRUE )

mdl.mom <- reduced.mom[[1]]
par.mom <- reduced.mom[[2]]
psi.mom <- sigex.par2psi(param = par.mom, mdl = mdl.mom)
sigex.lik(psi = psi.mom,
          mdl = mdl.mom,
          data.ts = data.ts,
          debug = FALSE )

[1] 6292.841

```

Here the threshold $\alpha = -6.22$, and the updated model is stored in *reduced.mom*, a list object. The first item in the list contains the updated information for the model – in particular, the new rank configurations. These are entered into *mdl.mom* by extraction from *reduced.mom*. Then the parameters are updated, and the last call just returns the divergence for the MOM estimates: the value is 6292.841. Curiously, this is less than the value obtained previously for a nesting model; however, the previous divergence of 6329.107 corresponds to a modification of the initial MOM fit where the covariance matrices have been rendered non-negative definite. The reduced rank specification utilizes the same cholesky factors *L*, only with columns omitted corresponding to the rank configuration; hence it is possible to obtain a lower value of the divergence.

We move to the MLE fitting at this point; initial values from the MOM fit are loaded through

```

data.ts <- analysis.mom[[1]]
mdl <- analysis.mom[[3]]
psi.mom <- analysis.mom[[4]]
par.mom <- sigex.psi2par(psi = psi.mom, mdl = mdl, data.ts = data.ts)

```

Fitting this reduced rank model can take a substantial amount of time, because there are 67 pre-parameters to be estimated. Optimization terminates at the divergence of 6152.607, which is substantially better than that of the MOM estimate. Now when we examine the condition numbers, we see values of $-\infty$ (or very large negative values) corresponding to the complement of the rank configuration.

The next stage of analysis is signal extraction, and both the matrix method and the WK method are illustrated in the script. Here we focus on three latent processes: trend, seasonal, and seasonal adjustment. Since the seasonal process is composed of the six atomic seasonals that occur at indices two through seven, *sigcomps* equals *seq(2, 7)*. Similarly, the setting is *c(1, 8)* for the seasonal adjustment. In the display we illustrate how the seasonal component can be plotted with a vertical displacement. In order to use *sigex.wkextract*, one sets the *grid* parameter to determine the

number of mesh points in the interval $[-\pi, \pi]$ used as a Riemann approximation of the integral; the *horizon* parameter corresponds to H (the number of desired afts and forecasts), whereas *target* is the array storing the coefficients of $\Xi(B)$. Here we set up the WK signal extraction via

```
grid    <- 7000
window  <- 50
horizon <- 0
target  <- array(diag(N), c(N, N, 1))
```

which specifies that the filter will be truncated to length 101 (equal to twice *window* plus one) at the sample boundaries – only 50 forecasts and afts will be generated. Setting *horizon* to zero means that we will not forecast the signal ahead. Finally, in the fourth line the *target* variable is set to the identity (as an array). The calls to `sigex.wkextract` are then given by

```
extract.trend2 <- sigex.wkextract(psi    = psi,
                                mdl      = mdl,
                                data.ts  = data.ts,
                                sigcomps = 1,
                                target   = target,
                                grid      = grid,
                                window    = window,
                                horizon    = horizon,
                                needMSE   = TRUE )
extract.seas2  <- sigex.wkextract(psi    = psi,
                                mdl      = mdl,
                                data.ts  = data.ts,
                                sigcomps = seq(2,7),
                                target   = target,
                                grid      = grid,
                                window    = window,
                                horizon    = horizon,
                                needMSE   = TRUE )
extract.sa2    <- sigex.wkextract(psi    = psi,
                                mdl      = mdl,
                                data.ts  = data.ts,
                                sigcomps = c(1,8),
                                target   = target,
                                grid      = grid,
                                window    = window,
                                horizon    = horizon,
                                needMSE   = TRUE )
```

The fourth argument in these function calls is *sigcomps*, and parallels the prior calls of `sigex.signal`. Note that whereas in the matrix approach we used two calls in succession (`sigex.signal` followed by `sigex.extract`), here there is a single call with output in the same list format as that provided by `sigex.extract`. The following plot in the script just compares the square root MSE for the trend components arising from both signal extraction methods; with *window* size of 50 there is no discrepancy. (The user can re-run the script with *window* set to 10, to observe some discrepancy between the exact matrix results and the approximate WK results.)

The final stage of this script is extraction of the trend growth rate, i.e., we extract $(1 - B)S_t^{(1)}$. Note that this is an I(2) trend, so the growth rate is a random walk – however, the dynamics of an extracted component in general do not match those of the specified process (see discussion in McElroy (2012)). The first code block extracts the trend growth rate and MSE directly from the output of `sigex.signal` stored in *signal.trend*, and we do not cover this in detail, instead focusing on the second code block which produces approximately the same results via the WK filter. Noting that the growth rate polynomial *gr.poly* has already been set to `c(1, -1)`, we proceed to

```
grid    <- 70000 # need large grid value to get accuracy
window  <- 100
horizon <- 0
gr.array <- array(t(gr.poly) \%x\% diag(N), c(N, N, p+1))
reg.gr   <- array(0, c(T, N))
```



```

for (k in 1:N) {
  reg.gr[, k] <- filter(x      = reg.trend[, k],
                        filter = gr.poly,
                        method = "convolution",
                        sides  = 1)
}

```

In the first line we take a large *grid* value to get decent accuracy, and take a larger truncation *window* of 100. In the fourth line, the target is defined by *gr.array* by writing *gr.poly* in an array. Finally, we need to modify the trend regressors so that they correctly correspond to the trend growth rate; this is accomplished by filtering the polynomial trend with *gr.poly*. (Recall that *reg.trend* was defined above, and already incorporates the regression parameters.) Next, we call

```

extract.trendgr2 <- sigex.wkextract(psi      = psi,
                                   mdl       = mdl,
                                   data.ts   = data.ts,
                                   sigcomps  = 1,
                                   target    = gr.array,
                                   grid      = grid,
                                   window    = window,
                                   horizon    = horizon,
                                   needMSE   = TRUE )

```

which can take some time because of the high value of *grid*. Note that *gr.array* is now in the *target* argument. The output is plotted in the usual way, displaying the trend growth rate for the four series, with appropriate uncertainty. The script repeats the exercise for the seasonal adjustment, and the output can be compared to the direct matrix output. This concludes the example.

4.4 Business Formation Statistics

Next, we study a weekly univariate series: the business applications component of the “Business Formation Statistics,”⁷ or *bfs* for short. We consider the non-seasonally adjusted data at the national level. This component is the main business applications series, which captures the weekly flow of IRS applications for Employer Identification Numbers, mainly for business purposes. For the metadata specifications we enter

```

begin <- c(2006, 1)
end   <- c(2020, 27)
period <- 52

```

which uses the fallacious *period* of 52. The actual period should be approximately $365.25/7 \approx 52.17857$, but *Ecce Signum* requires integer values of *period*. This is not entirely satisfactory, because a 53rd week is present every five to seven years; we present some ways below to circumvent this difficulty.

In loading the series, we select one of four national series corresponding to business applications, and choose a log transformation. In order to specify the holiday regressors, we need to know the calendar date (in month-day-year format) for the beginning and end of the series, i.e., the first day of the first week and the last day of the last week. The function *weekly2date* provides this:

```

first.day <- 1
all.date  <- weekly2date(first.day = first.day, begin = begin, T = T)
start.date <- all.date[[1]]
end.date  <- all.date[[2]]

```

Weekly data typically corresponds to measurements over a seven-day period beginning with Sunday, in which case *first.day* is set equal to 1, but non-standard week configurations can be accommodated as well. Having extracted the output in lines three and four, we can load the holiday regressors. One supplies an ASCII file whose every row contains a calendar date for the relevant holiday, covering a broad range of years. (Because we later remove a long-term mean of the holiday regressor to avoid confounding with seasonal effects, we recommend providing these dates for several

⁷Weekly data as of week 27 of 2020, downloaded on July 14, 2020 (11:27 AM), U.S. Census Bureau, obtained from <https://www.census.gov/econ/bfs/index.html>.

centuries.) These date files can be constructed from the Internet, or the `datefinder` program can be used. For Easter we utilize the following code; note `easter500`⁸ comes preloaded with the package.

```
easter.reg <- gethol(hol.dates = easter500,
                   hol.fore  = 7,
                   hol.aft   = 0,
                   start.date = start.date,
                   end.date  = end.date )
```

Regressors can simply be created as a daily time series by `gethol`, and then later converted to a weekly regressor. Here we consider a window of 7 days before and 0 days after, so that the holiday effect is presumed to constitute eight days, including the entire prior week. Note that in the call to `gethol`, both `start.date` and `end.date` are required.

For *bfs* we consider several holiday effects: all the ten federal holidays as well as Easter and Black Friday. In each case we take the actual date of the holiday rather than its observed date (the observation dates are more difficult to determine, but this could be done). Note that holidays that do not have moving dates (such as Independence Day, Veteran's Day, and Christmas) constitute a deterministic annual seasonal effect; because `gethol` removes a long-term mean over many years (so as to avoid confounding with seasonal effects), such regressors become zero. (We still include these three holidays in our analysis to demonstrate below that they are automatically removed from the model.)

Next, the daily regressors are converted to weekly flow regressors by first embedding with `sigex.daily2weekly` followed by averaging over each week:

```
easter.reg <- sigex.daily2weekly(data.ts = easter.reg,
                                first.day = first.day,
                                start.date = start.date)
easter.reg <- rowSums(easter.reg)/7
```

The averaging in the third line corresponds to the construction of monthly and quarterly regressors in X-12-ARIMA described in Findley et al. (1998), whereby the proportion of the effect (over the entire window) that falls in any given week is declared to be the value of the regressor for that week.

After these specifications, we fit a baseline model that involves no fixed effects, but models the serial dependence with a SARMA specification with order one for the AR, MA, seasonal AR, and seasonal MA polynomials:

```
mdl <- NULL
mdl <- sigex.add(mdl = mdl,
                vrank = seq(1,N),
                class = "sarima",
                order = c(1,1,1,1,52),
                bounds = NULL,
                name = "process",
                delta = 1)
mdl <- sigex.meaninit(mdl = mdl, data.ts = data.ts, d = 0)
```

In particular, we do not specify any differencing even though the level of the series seems to change – gradual (but transitory) level changes are not inconsistent with a stationary specification. The divergence is -2076.881 , and the fit seems reasonably good. Our next model incorporates the ten federal holidays along with Easter and Black Friday, although Independence Day, Veteran's Day, and Christmas are not included by the call to `sigex.reg` (which checks for regressors that are numerically zero after differencing). For instance,

```
mdl <- sigex.reg(mdl = mdl,
                series = 1,
                reg = ts(data = as.matrix(xmas.reg),
                        start = start(xmas.reg),
                        frequency = period,
                        names = "Xmas") )
mdl <- sigex.reg(mdl = mdl,
                series = 1,
```

⁸This is a conversion of the U.S.Census Bureau's list of Easter dates for a 500 year period (1600-2099): <https://www.census.gov/srd/www/genhol/easter500.html>

```

reg = ts(data      = as.matrix(black.reg),
         start      = start(black.reg),
         frequency  = period,
         names      = "BlackFriday") )

```

will make no modification to *mdl* in the first call, whereas the second call adds the Black Friday regressor. So this specification has seven federal holidays, plus two other holidays and the trend constant, for a total of 10 regressors. This improved model is fitted, producing a divergence of -2089.925 , but the routine terminates at a saddlepoint (the Hessian has a negative eigenvalue). We refine the model by restricting ourselves to the New Year's (NY) and Martin Luther King (MLK) holidays. At this point, we observe in the fitted residuals that there is a substantial outlier in the first week of 2012, which is not explained by the New Year's regressor. We can treat this as an additive outlier, which is handled via

```

AO.times <- 314
dataNA.ts <- data.ts
dataNA.ts[AO.times] <- NA

```

This defines a new time series object *dataNA.ts* that is identical to *data.ts* but has NA inserted wherever an additive outlier occurs. Our third refined model is therefore

```

mdl <- NULL
mdl <- sigex.add(mdl = mdl,
               vrank = seq(1,N),
               class = "sarma",
               order = c(1, 1, 1, 1, 52),
               bounds = NULL,
               name = "process",
               delta = 1)
mdl <- sigex.meaninit(mdl = mdl, data.ts = dataNA.ts, d= 0)
mdl <- sigex.reg(mdl = mdl,
               series = 1,
               reg = ts(data      = as.matrix(nyd.reg),
                       start      = start(nyd.reg),
                       frequency  = period,
                       names      = "NewYearDay") )
mdl <- sigex.reg(mdl = mdl,
               series = 1,
               reg = ts(data      = as.matrix(mlk.reg),
                       start      = start(mlk.reg),
                       frequency  = period,
                       names      = "MLK") )

```

This model is fitted and a lower divergence (-2329.284) is obtained – this is theoretically impossible (because the model is nested), but occurs due to wrongful termination of the optimization search in the larger model. This is not an uncommon situation in our experience, when working with time series models with many parameters, and seems worth mentioning. For this final model, we get a positive definite Hessian and hence *t* statistics for the parameters, confirming the significance of all the variables. Residual analysis indicates the model seems to be a good fit. We record the midcast for time $t = 314$ via

```

data.casts <- sigex.midcast(psi      = psi.mle,
                          mdl        = mdl,
                          data.ts    = dataNA.ts,
                          castspan  = 0 )

```

and bundle the analysis using *dataNA.ts*. The final stage of analysis is to filter the data. The relevant fixed effects are loaded into variables via

```

reg.trend <- sigex.fixed(data.ts = data.ts,
                       mdl       = mdl,
                       series    = 1,
                       param     = param,

```

```

                                type    = "Trend" )
reg.nyd    <- sigex.fixed(data.ts = data.ts,
                        mdl      = mdl,
                        series   = 1,
                        param    = param,
                        type     = "NewYearDay" )
reg.mlk    <- sigex.fixed(data.ts = data.ts,
                        mdl      = mdl,
                        series   = 1,
                        param    = param,
                        type     = "MLK" )
dataLIN.ts <- data.ts - ts(data      = reg.trend + reg.nyd + reg.mlk,
                        start      = start(data.ts),
                        frequency  = period)

```

The last instruction creates *dataLIN.ts*, consisting of the original time series with the fixed effects removed, i.e., \hat{Y}_t . This de-measured series shall be adjusted for the AO, and then filtered. We utilize a set of nonparametric filters to extract trend, seasonal, and non-seasonal (or seasonally adjusted) latent processes, implemented in *x11filters*. Details of these filters are given in Appendix B.

```

week.period <- 365.25 / 7
half.len    <- floor(week.period / 2)
x11.filters <- x11filters(period = week.period, p.seas = 1)
trend.filter <- x11.filters[[1]]
seas.filter  <- x11.filters[[2]]
sa.filter    <- x11.filters[[3]]
shift       <- (dim(sa.filter)[3] - 1) / 2

```

The first line defines the non-integer period, which *x11filters* is designed to accommodate. The call in the third line returns a three-item list, containing the three filters, each of which is a symmetric sequence of numbers. Next, we call *sigex.adhocextract*:

```

trend.comp <- sigex.adhocextract(psi    = psi,
                                mdl     = mdl,
                                data.ts = dataNA.ts,
                                adhoc   = trend.filter,
                                shift   = half.len,
                                horizon = 0,
                                needMSE = TRUE )
sa.comp    <- sigex.adhocextract(psi    = psi,
                                mdl     = mdl,
                                data.ts = dataNA.ts,
                                adhoc   = sa.filter,
                                shift   = shift,
                                horizon = 0,
                                needMSE = TRUE )
AO.errs    <- dataLIN.ts[AO.times] - data.casts[[1]]
sa.comp[[1]][AO.times] <- sa.comp[[1]][AO.times] + AO.errs

```

In the second call, note that we apply the method to *dataNA.ts*, which was defined previously as the de-measured series with an NA inserted for the AO. The method will automatically generate the appropriate cast for the NA, filter the resulting casted series, and return the output along with appropriate uncertainty. The AC filter *trend.filter* is inputted, along with the *shift* parameter set to the middle position of the filter vector – this corresponds to a symmetric filter. The penultimate line computes the casting error ε_t at the AO time $t = 314$, and this is added onto the seasonal adjustment in the last line – because an AO effect belongs with the irregular in the seasonal adjustment extraction.

The final parts of the example display the extractions. The illustration is complete.

4.5 New Zealand Arrivals

Our last illustration considers a daily immigration series (*nz* for short): New Zealand residents arriving in New Zealand after an absence of less than 12 months, covering the period January 1, 2008 through July 31, 2012.⁹ The most obvious dynamics are the annual movements due to seasonality; more subtle are the weekly dynamics. Related time series (of longer sample size) were investigated in McElroy and Jach (2019); here we find that the sample autocorrelation functions for the series is somewhat different for each of the seven weekly series corresponding to each day-of-the-week. This phenomenon indicates that the daily series – analogously to the daily retail series analyzed in McElroy (2022) – can be modeled by a vector embedding as a 7-variate weekly series. Some initial processing of the daily data is needed:

```
n.months <- dim(imm)[1]/32
imm <- imm[-seq(1,n.months)*32,] # strip out every 32nd row (totals)
imm <- matrix(imm[imm != 0],ncol=6) # strip out 31st False days
```

The original data is grouped into “months” of 31 days, with extra days of value zero added for months of shorter length; our code removes these features, as well as a monthly summation. Next, a set of customized regressors, *NZregs* comes with package. The first column is a regressor that captures a moving holiday effect for Easter, and the next six columns correspond to school holidays for public schools in New Zealand: an effect is estimated for the start and end of the holiday for the first three school vacations in each year.¹⁰ The last three columns correspond to Y2K effects, and will not be used. For metadata we enter

```
start.date <- c(9, 1, 1997)
end.date <- day2date(day = dim(imm)[1] - 1, start.date = start.date)
period <- 365
```

The full data begins on September 1, 1997 and ends July 31, 2012, though we will be interested in a restricted range. Following are some calendrical calculations, and a call to *sigex.load*:

```
dataALL.ts <- sigex.load(data      = imm,
                        start.date = begin,
                        period     = period,
                        epithets   = c("NZArr", "NZDep", "VisArr",
                                      "VisDep", "PLTArr", "PLTDep"),
                        plot       = TRUE )
```

There are six time series in the data set, though we will be focused on just the first one “NZArr,” described above. We restrict the data span and focus on the first series via

```
transform <- "log"
aggregate <- FALSE
subseries <- 1
range <- list(c(2008, 1), end)
dataONE.ts <- sigex.prep(data.ts   = dataALL.ts,
                        transform = transform,
                        aggregate  = aggregate,
                        subseries  = subseries,
                        range      = range,
                        plot       = TRUE )
```

We can examine spectral plots to get an idea of the time series dynamics. Also we set the first day to be Sunday, and embed as a 7-variate weekly time series via a call to *sigex.daily2weekly*:

```
first.day <- 1
```

⁹The series consists of New Zealand residents arriving in New Zealand after an absence of less than 12 months. These are public use data produced by Stats New Zealand via a customized extract, and correspond to a portion of the “daily border crossings - arrivals” tab (Total) of the Travel category in the Covid-19 portal: <https://www.stats.govt.nz/experimental/covid-19-data-portal>.

¹⁰The school holiday regressor for the first day of the vacation period is set to one for that day and one for the day after, zero otherwise. The school holiday regressor for the last day of the vacation period is set to one for that day and one for the day before, zero otherwise. The school holiday regressors are centered by removing the daily mean of each of the regressors over the entire series.

```
data.ts    <- sigex.daily2weekly(data.ts = dataONE.ts,
                                first.day = first.day,
                                start.date = start.date)

plot(data.ts)
```

As a result $N = 7$ and $T = 240$, which yields an initial NA (because the first data value is on Monday, but the week begins on Sunday). Also the data ends on a Sunday, so the last week has six NA values. This is a classic example of a ragged edge data pattern. Next, we define appropriate subspans of the holiday regressors via

```
s <- (range[[1]][1] - begin[1]) * period + (range[[1]][2] - begin[2]) + 1
e <- (range[[2]][1] - begin[1]) * period + (range[[2]][2] - begin[2]) + 1
times <- s:e
easter.reg    <- NZregs[times,1]
school1.reg   <- NZregs[times,2]
school1e.reg  <- NZregs[times,3]
school2.reg   <- NZregs[times,4]
school2e.reg  <- NZregs[times,5]
school3.reg   <- NZregs[times,6]
school3e.reg  <- NZregs[times,7]
```

We have labeled the regressors corresponding to Easter and the three school holiday effects (both the beginning and end of each holiday period). These are the regressors for daily data, and need to be embedded as weekly series:

```
easter.reg <- sigex.daily2weekly(data.ts    = easter.reg,
                                first.day   = first.day,
                                start.date  = start.date )
easter.reg[is.na(easter.reg)] <- 0
```

The last line replaces any ragged edge NA values (arising from embedding) with zero, which is appropriate because the corresponding data values are also missing.

Now we come to the model specification. The data has a slight upwards trend drift, and there is strong annual correlation present. Although a stationary model can be entertained, we found a fairly good model fit given by a SVARMA for data differenced to stationarity by applying annual differencing $1 - L^{52}$. (Recall L is the weekly lag operator, and there are approximately 52 weeks per year.) After trying different possibilities, we consider nonseasonal $q = 1$ and seasonal $P = 1$, which is specified by the code

```
mdl <- NULL
mdl <- sigex.add(mdl    = mdl,
                vrank   = seq(1,N),
                class   = "svarma",
                order   = c(0, 1, 1, 0, 52),
                bounds  = NULL,
                name    = "process",
                delta   = c(1, rep(0, 51), -1) )
mdl <- sigex.meaninit(mdl = mdl, data.ts = data.ts, d = 0)
```

The specification of the regressors takes more coding, and we show the code for Easter (the index i loops over the days of the week):

```
mdl <- sigex.reg(mdl    = mdl,
                series  = i,
                reg     = ts(data    = as.matrix(easter.reg[,i]),
                            start    = start(easter.reg),
                            frequency = frequency(easter.reg),
                            names    = "Easter-day") )
```

Next we consider model fitting. Here we make use of parameter constraints for the regressors: because the same holiday effect is present for each day of the week (recall that the weekly regressors are derived from a single daily regressor), the parameter estimates should be constrained to be identical. We use the `sigex.constrainreg` function to do this: the `regindex` argument specifies indices J_1, J_2, \dots, J_N that delineate which regressors are being considered for each series.

For instance, $J_1 = \{1, 3, 4\}$ would correspond to the trend, school1, and school1e regressors for the first (Sunday) series. The *combos* argument provides the linear combinations of these regressors, but when set to NULL instead it is enforced that all the regression parameters are the same.

```
constraint <- rbind(constraint,
  sigex.constrainreg(mdl,data.ts,list(2,2,2,2,2,2,2),NULL))
constraint <- rbind(constraint,
  sigex.constrainreg(mdl,data.ts,list(3,3,3,3,3,3,3),NULL))
constraint <- rbind(constraint,
  sigex.constrainreg(mdl,data.ts,list(4,4,4,4,4,4,4),NULL))
constraint <- rbind(constraint,
  sigex.constrainreg(mdl,data.ts,list(5,5,5,5,5,5,5),NULL))
constraint <- rbind(constraint,
  sigex.constrainreg(mdl,data.ts,list(6,6,6,6,6,6,6),NULL))
constraint <- rbind(constraint,
  sigex.constrainreg(mdl,data.ts,list(7,7,7,7,7,7,7),NULL))
```

So the first call here identifies in *regindex* the Easter regressor for each of the seven series, and *combos* set to NULL indicates that these parameters will be equal. The next call does the same for the school1 regressor, and so forth; the first regressor is the trend, and this remains unconstrained. The output of *sigex.constrainreg* is a *constraint* matrix, which can be then passed into the MLE fitting routine. In this case the nonlinear optimization can take a long time – several days – and we just report the output in the script. The remaining code lines review diagnostics, as discussed in previous illustrations.

Moving on to signal extraction, we consider applying a daily seasonal adjustment filter that we embed as a 7-variate weekly filter. Consider the code

```
sa.hifilter <- c(1, rep(2, 365), 1) / (2 * 365)
len <- 183
hi.freq <- 7
low.freq <- 1
shift.hi <- len
out <- sigex.hi2low(filter.hi = sa.hifilter,
  hi.freq = hi.freq,
  low.freq = low.freq,
  shift.hi = shift.hi )
sa.lowfilter <- out[[1]]
shift.low <- out[[2]]
```

The daily seasonal adjustment filter *sa.hifilter* is an extension of the classic crude trend filter (known as the 2×12 filter) of X11 (discussed in McElroy and Politis (2020)) from monthly to daily frequency; it should suppress all daily frequencies of the form $2\pi j/365$ for $1 \leq j \leq 183$, while also linear trends. This is embedded with a call to *sigex.hi2low*, where the high and low frequencies are indicated and the *shift.hi* parameter is $c = 183$, corresponding to a symmetric filter. The output is the 7-variate weekly filter *sa.lowfilter*, with shift parameter $C = 27$ given by *shift.low*.

Next, this filter is applied using a standard call of *sigex.adhocextract*. Recall that this automatically supplies missing value imputations for the ragged edges, and does forecast and aforecast extension for application of the filter. However, this should be transformed back to a daily series, so that we can view the daily series with its seasonal adjustment:

```
daily.ts <- ts(data = sa.low[[1]],
  start = start(dataONE.ts),
  frequency = frequency(dataONE.ts) )
sa.hi.daily[[1]] <- sigex.weekly2daily(data.ts = daily.ts,
  first.dat = first.day )
```

The function *sigex.weekly2daily* transforms the series, using the indicated value of *first.day*. We repeat this for the lower and upper values of the confidence intervals (not shown). This can now be displayed, but we first generate another signal for comparison – we remove only the weekly dynamics. This is done by considering a simple filter annihilating weekly effects from the daily data, namely $(1 + B + \dots + B^6)/7$, which we call the Trading Day (TD) filter. The relevant code is

```

td.hifilter <- rep(1, 7) / 7
len         <- 3
hi.freq     <- 7
low.freq    <- 1
shift.hi    <- len
out <- sigex.hi2low(filter.hi = td.hifilter,
                    hi.freq   = hi.freq,
                    low.freq   = low.freq,
                    shift.hi   = shift.hi )
td.lowfilter <- out[[1]]
shift.low    <- out[[2]]

```

Now $c = 3$ and $C = 1$ for the TD filter. Before plotting, the trend must be computed. A subtlety is that we have 7 trends, one for each day of week, which therefore combine long-term trend effects and day-of-week effects. The mean of the 7 parameter values corresponds to the overall mean of each week, and the residual is the day-of-week effect in excess of the mean. So we define the daily trend to be the output of applying *td.hifilter* to the regression mean, i.e.,

```

reg.td <- NULL
for (i in 1:N){
  reg.td <- cbind(reg.td, sigex.fixed(data.ts = data.ts,
                                     mdl      = mdl,
                                     series   = i,
                                     param    = param,
                                     type     = "Trend") )
}
reg.td <- rbind(rep(0, N), reg.td)
reg.td <- ts(data      = sigex.weekly2daily(data.ts = reg.td, first.day = first.day),
             start     = start(dataONE.ts),
             frequency = period)
reg.trend <- filter(x      = reg.td,
                   filter  = td.hifilter,
                   method  = "convolution",
                   sides   = 1)
reg.trend <- as.matrix(reg.trend[8:length(reg.td)])

```

First *reg.td* is computed, which is the trend means for each day-of-week expressed as a daily time series. We augment by one prior week's values, which would be all zero, and then convert to a daily series. Next, *reg.trend* is obtained by filtering with $(1 + B + \dots + B^6)/7$, and finally remove the NA values (resulting from using *filter*, which does not extend the input series). The illustration is completed with a display: the seasonal adjustment contains some mild weekly effects (because the filter is designed to remove daily, and not weekly frequencies), whereas the TD-adjusted component has annual movements but no weekly effects.

Acknowledgments

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

References

- Bell, W. (1984). Signal extraction for nonstationary time series. *The Annals of Statistics*, 12:646–664.
- Bell, W. R. and Pugh, M. G. (1989). Alternative approaches to the analysis of time series components. In *Proceedings of the Statistics Canada Symposium on Analysis of Data in Time*, Ottawa.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons, New York.
- Census Bureau, U. S. (2015). *Genhol Documentation*.
- Chang, I., Tiao, G. C., and Chen, C. (1988). Estimation of time series parameters in the presence of outliers. *Technometrics*, 30(2):193–204.

- De Livera, A. M., Hyndman, R. J., and Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American statistical association*, 106(496):1513–1527.
- Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C., and Chen, B. C. (1998). New capabilities of the x-12-arima seasonal adjustment program (with discussion). *Journal of Business and Economic Statistics*, 16.
- Franses, P. H. (1994). A multivariate approach to modeling univariate seasonal time series. *Journal of Econometrics*, 63:133–151.
- Gilbert, P. (2020). *dse: Dynamic Systems Estimation (Time Series Package)*.
- Gladyshev, E. G. (1961). Periodically correlated random sequences. *Soviet Mathematics*, 2:385–388.
- Golub, G. H. and Van Loan, C. F. (2012). *Matrix computations (Vol. 3)*. JHU press.
- Graves, S., Dorai-Raj, S., and François, R. (2010). *sos: Search Contributed R Packages, Sort by Package*.
- Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.
- Harvey, A. C. and Trimbur, T. M. (2003). General model-based filters for extracting cycles and trends in economic time series. *Review of Economics and Statistics*, 85((2)):244–255.
- Helske, J. (2016). Kfas: Exponential family state space models in r. *arXiv preprint arXiv:1612.01907*.
- Hillmer, S. C. and Tiao, G. C. (1982). An arima-model-based approach to seasonal adjustment. *Journal of the American Statistical Association*, 77:63–70.
- Holan, S., McElroy, T. S., and Wu, G. (2017). The cepstral model for multivariate time series: the vector exponential model. *Statistica Sinica*, 27:23–42.
- Holmes, E., Ward, E., and Wills, K. (2012). Marss: Multivariate autoregressive state-space models for analyzing time-series data. *R journal*, 4(1).
- Holmes, E., Ward, E., and Wills, K. (2020). *MARSS: Multivariate Autoregressive State-Space Modeling*.
- Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O’Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., and Yasmeen, F. (2020). *forecast: Forecasting functions for time series and linear models*. R package version 8.13.
- Hyndman, R. J. and Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.
- Hyndman, R. J. and Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 26(3):1–22.
- Ladiray, D., Palate, J., Mazzi, G., and Proietti, T. (2018). Seasonal adjustment of daily data. *16th Conference of the International Association of Official Statisticians (IAOS), Paris, France*.
- Ljung, G. (1993). On outlier detection in time series. *Journal of the Royal Statistical Society, Series B*, 55:559–567.
- Ljung, G. M. and Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65:297–303.
- López-de Lacalle, J. (2016). *stsm: Structural Time Series Models*.
- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer, Science & Business Media.
- McElroy, T. S. (2008). Matrix formulas for nonstationary arima signal extraction. *Econometric Theory*, 24:1–22.
- McElroy, T. S. (2012). An alternative model-based seasonal adjustment that reduces residual seasonal autocorrelation. *Taiwan Economic Forecast and Policy*, 43.
- McElroy, T. S. (2016a). Nonnested model comparisons for time series. *Biometrika*, 103:905–914.
- McElroy, T. S. (2016b). On the measurement and treatment of extremes in time series. *Extremes*, 19:467–490.
- McElroy, T. S. (2017). Multivariate seasonal adjustment, economic identities, and seasonal taxonomy. *Journal of Business & Economic Statistics*, 35(4):611–625.
- McElroy, T. S. (2018). Recursive computation for block-nested covariance matrices. *Journal of Time Series Analysis*, 39(3):299–312.
- McElroy, T. S. (2022+). Casting vector time series with the information filter: Algorithms for forecasting, imputation, and signal extraction. Mimeo.
- McElroy, T. S. and Holan, S. H. (2014). Asymptotic theory of cepstral random fields. *The Annals of Statistics*, 42:64–86.
- McElroy, T. S. and Jach, A. (2019). Testing collinearity of vector time series. *The Econometrics Journal*, 22:97–116.

- McElroy, T. S. and Monsell, B. C. (2015). Model estimation, prediction, and signal extraction for nonstationary stock and flow time series observed at mixed frequencies. *Journal of the American Statistical Association*, 110:1284–1303.
- McElroy, T. S. and Penny, R. (2019). Maximum entropy extreme-value seasonal adjustment. *Australian New Zealand Journal of Statistics*, 61:152–174.
- McElroy, T. S. and Politis, D. N. (2020). *Time Series: A First Course with Bootstrap Starter*. CRC Press, New York.
- McElroy, T. S. and Roy, A. (2018). The inverse kullback–leibler method for fitting vector moving averages. *Journal of Time Series Analysis*, 39(2):172–191.
- McElroy, T. S. and Roy, A. (2021a). Model identification via total frobenius norm of multivariate spectra. *Journal of the Royal Statistical Society, Series B*, pages 1–23.
- McElroy, T. S. and Roy, A. (2021b). Testing for adequacy of seasonal adjustment in the frequency domain. *Journal of Statistical Planning and Inference*, 221:241–255.
- McElroy, T. S. and Trimbur, T. (2015). Signal extraction for non-stationary multivariate time series with illustrations for trend inflation. *Journal of Time Series Analysis*, 36(2):209–227.
- McElroy, T. S. and Wildi, M. (2020). The multivariate linear prediction problem: model-based and direct filtering solutions. *Econometrics and Statistics*.
- Ollech, D. (2018). Seasonal adjustment of daily time series. *Deutsche Bundesbank Discussion Paper*.
- Osborn, D. R. (1991). The implications of periodically varying coefficients for seasonal time-series processes. *Journal of Econometrics*, 28:373–384.
- Petris, G. (2010). An R package for dynamic linear models. *Journal of Statistical Software*, 36(12):1–16.
- Pinheiro, J. C. and Bates, D. M. (1996). Unconstrained parameterization for variance-covariance matrices. *Statistics and computing*, 6:289–296.
- Roy, A., McElroy, T. S., and Linton, P. (2019). Estimation of causal invertible varma models. *Statistica Sinica*, 29:455–478.
- Scott, S. L. (2020). *bsts: Bayesian Structural Time Series*.
- Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, pages 591–611.
- Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic theory of statistical inference for time series*. Springer, New York.
- Taylor, S. J. and Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1):37–45.
- Tiao, G. C. and Grupe, M. R. (1980). Hidden periodic autoregressive-moving average models in time series data. *Biometrika*, 67:365–373.
- Timmermans, M., Heijmans, R., and Daniels, H. (2017). Cyclical patterns in risk indicators based on financial market infrastructure transaction data. De Nederlandsche Bank Working Paper 558.
- Trimbur, T. S. (2006). Properties of higher order stochastic cycles. *Journal of Time Series Analysis*, 27:1–17.
- Tsay, R. S. (1986). Time series model specification in the presence of outliers. *Journal of the American Statistical Association*, 81:132–141.

SUPPLEMENTAL MATERIAL FOR “ECCE SIGNUM: AN R PACKAGE FOR MULTIVARIATE SIGNAL EXTRACTION AND TIME SERIES ANALYSIS”

A PREPRINT

Tucker McElroy
U.S. Census Bureau
4600 Silver Hill Road, Washington, DC 20233
tucker.s.mcelroy@census.gov

James Livsey
U.S. Census Bureau
4600 Silver Hill Road, Washington, DC 20233
james.a.livsey@census.gov

ABSTRACT

The package provides multivariate time series models for structural analysis, allowing one to extract latent signals such as trends or seasonality. Models are fitted using maximum likelihood estimation, allowing for non-stationarity, fixed regression effects, and ragged-edge missing values. Simple types of extreme values can be corrected using the device of entropy maximization. Model adequacy is assessed through residual diagnostics, and model-based signal extraction filters can be assessed in time domain and frequency domain. Extracted signals are produced with uncertainty measures that account for sample edge effects and missing values, and the signals (as well as the original time series) can be forecasted.

Keywords Entropy maximization, Filtering, Missing values, Ragged edge, Seasonality, Signal extraction

A Definition of Cyclical Processes

A.1 Butterworth Cycle

The Butterworth cyclical process is described in Harvey and Trimbur (2003), and in the first order case is an ARMA(2,1) process whose parameters are governed by a persistence ρ and a frequency ω . The order n *Butterworth cyclical process* is defined via

$$(1 - 2\rho \cos(\omega)B + \rho^2 B^2)^n x_t = (1 - \rho \cos(\omega)B)^n \epsilon_t,$$

for $\epsilon_t \sim \text{WN}(0, \sigma^2)$. It is known that the causal moving average representation corresponds to $x_t = \Psi(B)^n \epsilon_t$ with

$$\Psi(B) = \sum_{j \geq 0} \psi_j B^j = \frac{1 - \rho \cos(\omega)B}{1 - 2\rho \cos(\omega)B + \rho^2 B^2} \quad \text{and} \quad \psi_j = \rho^j \cos(\omega j).$$

Clearly, as ρ increases towards one the cyclical dynamics become increasingly regular, and the impulse response function when $n = 1$ becomes that of perfect sinusoid. Moreover, the AR(2) operator matches the case of an atomic process with complex conjugate root $e^{\pm i\omega}$, as $\rho \rightarrow 1$. The period of the cyclical effect is approximately $2\pi/\omega$.

Stabilization is determined by minimizing $f(\lambda) = |\Psi(e^{-i\lambda})|^{2n}$. Taking the logarithm and differentiating yields

$$\frac{-2n\rho \sin(\omega + \lambda)}{1 + \rho^2 - 2\rho \cos(\omega + \lambda)} - \frac{-2n\rho \sin(\omega - \lambda)}{1 + \rho^2 - 2\rho \cos(\omega - \lambda)} + \frac{2n\rho \cos(\omega) \sin(\lambda)}{1 - 2\rho \cos(\omega) \cos(\lambda) + \rho^2 \cos(\omega)^2}.$$

Setting this equal to zero and simplifying, we must solve

$$\begin{aligned}
0 &= (2\rho \cos(\omega) \sin(\lambda)) \left(1 + 4\rho^2 \cos(\omega)^2 + \rho^4 - 4\rho(1 + \rho^2) \cos(\omega) \cos(\lambda) + 2\rho^2 \cos(2\lambda) \right) \\
&\quad - \left(1 + \rho^2 \cos(\omega)^2 - 2\rho \cos(\omega) \cos(\lambda) \right) (4\rho(1 + \rho^2) \cos(\omega) \sin(\lambda) - 4\rho^2 \sin(2\lambda)) \\
&= 2\rho \cos(\omega) \left(1 + 4\rho^2 \cos(\omega)^2 + \rho^4 \right) \sin(\lambda) + 4\rho^3 \cos(\omega) \sin(\lambda) \cos(2\lambda) \\
&\quad - 4\rho(1 + \rho^2 \cos(\omega)^2) (1 + \rho^2) \cos(\omega) \sin(\lambda) + 8\rho^2 (1 + \rho^2 \cos(\omega)^2) \sin(\lambda) \cos(\lambda) \\
&\quad - 16\rho^3 \cos(\omega) \sin(\lambda) \cos(\lambda)^2.
\end{aligned}$$

Hence $\lambda = 0, \pi$ are critical points; but if $\lambda \in (0, \pi)$ we can divide out by $\sin(\lambda)$, obtaining

$$\begin{aligned}
0 &= a_0 + a_1 z + a_2 z^2 \\
a_0 &= 2\rho \cos(\omega) \left(1 + 4\rho^2 \cos(\omega)^2 + \rho^4 \right) - 4\rho^3 \cos(\omega) - 4\rho(1 + \rho^2 \cos(\omega)^2) (1 + \rho^2) \cos(\omega) \\
&= 2\rho \cos(\omega) \left(2\rho^2 \cos(\omega)^2 (1 - \rho^2) - 1 - 4\rho^2 \right) \\
a_1 &= 8\rho^2 (1 + \rho^2 \cos(\omega)^2) \\
a_2 &= -8\rho^3 \cos(\omega)
\end{aligned}$$

with $z = \cos(\omega)$. We apply the quadratic formula: the discriminant simplifies to

$$\sin(\omega)^2 \left(\sin(\omega)^2 + (1 - \rho^2)^2 \cos(\omega)^2 \right),$$

and the cosine of the critical point is

$$z = \frac{1 + \rho^2 \cos(\omega)^2 \pm \sin(\omega) \sqrt{\sin(\omega)^2 + (1 - \rho^2)^2 \cos(\omega)^2}}{2\rho \cos(\omega)}.$$

Hence, a critical point is given by $\arccos(z)$ when $|z| \leq 1$; in this case, we can evaluate $f(\lambda) = |\Psi(e^{-i\lambda})|^{2n}$ at this point and compare to the values at $\lambda = 0, \pi$, thereby determining λ^* as the minimizer:

$$c = |\Psi(e^{-i\lambda^*})|^{2n}.$$

The stabilized Butterworth cycle spectral density is then

$$f^*(\lambda) = \frac{|1 - \rho \cos(\omega) e^{-i\lambda}|^{2n} - c |1 - 2\rho \cos(\omega) e^{-i\lambda} + \rho^2 e^{-i2\lambda}|^{2n}}{|1 - 2\rho \cos(\omega) e^{-i\lambda} + \rho^2 e^{-i2\lambda}|^{2n}} \sigma^2,$$

which corresponds to an ARMA(2n, 2n) process. The moving average polynomial can be determined by spectral factorization; the autocovariances are simply given by those of the original Butterworth cycle process, only with the variance being decreased by c .

A.2 Balanced Cycle

The Balanced cyclical process is described in Harvey and Trimbur (2003). This is described through a formulation in state space, but can be expressed in ARMA form as follows. The order n *Balanced cyclical process* is defined via

$$\begin{aligned}
(1 - 2\rho \cos(\omega)B + \rho^2 B^2)^n x_t &= .5 B^{n-1} \left[(1 - \rho e^{i\omega} B)^n + (1 - \rho e^{-i\omega} B)^n \right] \kappa_t \\
&\quad + .5 i B^{n-1} \left[(1 - \rho e^{i\omega} B)^n - (1 - \rho e^{-i\omega} B)^n \right] \kappa_t^*,
\end{aligned}$$

for $\kappa_t \sim \text{WN}(0, \sigma^2)$ and $\kappa_t^* \sim \text{WN}(0, \sigma^2)$, independent of one another. The moving averages of $\{\kappa_t\}$ and $\{\kappa_t^*\}$ are each real polynomials in B of degree $2n - 1$; writing them in terms of the complex number $e^{i\omega}$ produces compact expressions that facilitate other calculations. The expressions for the autocovariance function given in Trimbur (2006) can be rewritten as

$$\begin{aligned}
\gamma_h &= \rho^h \cos(h\omega) \left(\sum_{j=0}^{n-1} \binom{h}{j} \alpha_{n-j} \right) (1 - \rho^2)^{1-2n} \sigma^2 \\
\alpha_k &= (1 - \rho^2)^{n-k} \sum_{r=0}^{k-1} \binom{k-1}{r} \binom{n-1}{r+n-k} \rho^{2r}
\end{aligned}$$

for $h \geq 0$, where we interpret α_k for $k \leq 0$ as zero (and $\binom{h}{j} = 0$ if $j > h$). A simple expression for the spectral density can be obtained as follows: the spectral density of the moving average portion is obtained by substituting $e^{-i\lambda}$ for B in both moving average polynomials, followed by taking the squared magnitude and summing both terms. This yields

$$\begin{aligned} & .25 \left| (1 - \rho e^{i\omega} e^{-i\lambda})^n + (1 - \rho e^{i\omega} e^{-i\lambda})^n \right|^2 + .25 \left| (1 - \rho e^{i\omega} e^{-i\lambda})^n - (1 - \rho e^{i\omega} e^{-i\lambda})^n \right|^2 \\ & = .5 (1 - 2\rho \cos(\omega - \lambda) + \rho^2)^n + .5 (1 - 2\rho \cos(\omega + \lambda) + \rho^2)^n. \end{aligned}$$

Noting that the two summands comprise the two terms in the squared magnitude of the AR(2) polynomial's frequency response function, we see that the spectral density is

$$f(\lambda) = .5 \left[(1 - 2\rho \cos(\omega - \lambda) + \rho^2)^{-n} + (1 - 2\rho \cos(\omega + \lambda) + \rho^2)^{-n} \right] \sigma^2.$$

These results can be used to determine the ARMA(2n,n) representation of the process: the autocovariance function corresponding to the moving average portion can be expressed as

$$v_h = \cos(\omega h) \sum_{k=0}^{n-h} \binom{n}{k+h} \binom{n}{k} (-\rho)^{2k+h}$$

for $0 \leq h \leq n$, by using the binomial formula on $(1 - \rho e^{i\omega} B)^n$ to determine the autocovariance generating function. Also $v_h = 0$ for $h > n$, so we can compute the moving average polynomial using spectral factorization.

Another consequence of the simple expression for the spectral density is that the stabilized balanced cycle is easy to determine. Let $g(\lambda) = (1 - 2\rho \cos(\omega - \lambda) + \rho^2)^{-n}$, which is uni-modal with minimum value over $[0, \pi]$ occurring at $\lambda = \pi$ so long as $\omega \in (0, \pi/2)$. (Otherwise, if $\omega \in [\pi/2, \pi)$, there is a minimum value at $\lambda = 0$ – but this case is of less interest in practice.) The maximizer of g is ω (without loss of generality, suppose this is positive), and hence $g(\pi) < g(\lambda)$ for all $\lambda \in [0, \pi]$ and $g(-\pi) < g(\lambda)$ for all $\lambda \in [-\pi, \omega]$. It follows that $f(\lambda) = .5 (g(\lambda) + g(-\lambda)) \sigma^2$ satisfies $f(\pi) < f(\lambda)$ for all $\lambda \in [0, \pi]$. Hence $c = f(\pi)$, and the stabilized Balanced cycle spectral density is

$$f^*(\lambda) = \frac{.5 g(\lambda)^{-1} + .5 g(-\lambda)^{-1} - f(\pi) |1 - 2\rho \cos(\omega) e^{-i\lambda} + \rho^2 e^{-i2\lambda}|^{2n}}{|1 - 2\rho \cos(\omega) e^{-i\lambda} + \rho^2 e^{-i2\lambda}|^{2n}} \sigma^2.$$

Again, spectral factorization can be used to determine the moving average corresponding to the numerator. But the autocovariance function is the same as the original process, but with the variance decremented by $c = f(\pi)$.

B Fractional Seasonal Period

B.1 Seasonal Adjustment of Weekly Time Series

Cleveland and Scott (2007) discuss a technique for adjusting weekly series based upon locally weighted regression, motivated by the varying presence of either 52 or 53 weeks in a year. Starting with the SABL method of Cleveland et al. (1978), this non-integer period of seasonality has been handled without classical filters, as in the X-11 procedure. The linearized X-11 filter for monthly time series is described in Ladiray and Quenneville (2012); a simplified, heuristic discussion is given in Chapter 3 of McElroy and Politis (2020)). One proceeds in four steps, letting s be the seasonal period:

1. Crude Trend: apply the symmetric filter $\Upsilon_{\text{trend}}(B) = (1 + B)(1 + B + \dots + B^{s-1})B^{-s/2}/(2s)$ to get an initial trend with both seasonality and high-frequency noise suppressed.
2. De-trend: subtract the crude trend from the data. Equivalently, apply $1 - \Upsilon_{\text{trend}}(B)$.
3. Seasonal: now apply to the de-trended data an order p seasonal moving average of the form $\Upsilon_{\text{seas}}(B) = (2p + 1)^{-1} \sum_{j=-p}^p B^{sj}$.
4. De-seasonalize: finally, subtract the extracted seasonal from the data.

Altogether, the seasonal adjustment filter is

$$\Upsilon_{\text{sa}}(B) = 1 - \Upsilon_{\text{seas}}(B) (1 - \Upsilon_{\text{trend}}(B)).$$

This simple approach, which is predicated on an additive decomposition of the data in terms of trend, seasonal, and irregular components, is effective in many cases, but requires s to be an even integer. Actually, the technique can

be adapted to the case that s is an odd integer (e.g., when removing weekly effects from a daily time series), but for non-integer s it is unclear how to proceed.

For weekly time series this generalization is needed, because the number of weeks occurring in a year is not an integer: on average, there are 365.25 days in a year (this includes leap year, but excludes the corrections occurring every 100 years and every 400 years, for simplicity), and dividing by 7 yields $s = 52.1786$. To handle non-integer s , we first observe that in step 1 the crude trend filter is constructed from

$$(1 + B)(1 + B + \dots + B^{s-1}) = \prod_{k=1}^{s/2} (1 - 2 \cos(2\pi k/s)B + B^2)$$

when s is an even integer. This ensures annihilation of all sinusoids of frequency $2\pi k/s$ for $k = 1, \dots, s/2$ (McElroy and Politis, 2020). So for a generalization, consider

$$U_{n,s} = \prod_{k=1}^n (1 - 2 \cos(2\pi k/s)B + B^2) \quad (1)$$

for integer $n \leq [s/2]$, where $[\cdot]$ denotes the floor function. Then define $\Upsilon_{\text{trend}}(B) = c^{-1} B^{-n} U_{[s/2],s}(B)$, where c is a normalization constant such that $\Upsilon_{\text{trend}}(1) = 1$. It is apparent that such a filter annihilates the desired harmonics, and maintains the level of an input time series. To modify the seasonal filter, it is useful to study $1 - \Upsilon_{\text{seas}}(B)$. Observe that for even integer s ,

$$\begin{aligned} 1 - \Upsilon_{\text{seas}}(B) &= (2p+1)^{-1} \sum_{j=-p}^p (1 - B^{sj}) = (2p+1)^{-1} \sum_{j=1}^p (2 - B^{sj} - B^{-sj}) \\ 1 - B^{sj} &= (1 - B)(1 + B + \dots + B^{sj-1}) = (1 - B^2) \prod_{k=1}^{(sj-2)/2} (1 - 2 \cos(2\pi k/s)B + B^2), \end{aligned}$$

for $j \geq 1$. This indicates that for non-integer s we can define the filter as

$$\Upsilon_{\text{seas}}(B) = 1 - (2p+1)^{-1} \sum_{j=1}^p ((1 - B^2) U_{[(s|j|-2)/2],s}(B) + (1 - B^{-2}) U_{[(s|j|-2)/2],s}(B^{-1})).$$

When s is large, direct computation of the coefficients of $U_{[(s|j|-2)/2],s}(z)$ can generate numerical distortions. For better numerical stability, we use the cepstral method for computing the product of polynomials; see pp.232-237 of McElroy and Politis (2020) for background. The ℓ th cepstral coefficients τ_ℓ for $\ell \geq 1$ corresponding to the polynomial $1 - 2 \cos(2\pi k/s)z + z^2$ are given by

$$\tau_\ell = -2 \cos(2\pi k\ell/s)/\ell,$$

and hence the cepstral coefficients of $\prod_{k=1}^{[sj-2]/2} (1 - 2 \cos(2\pi k/s)z + z^2)$ are given by summing τ_ℓ over $k = 1, \dots, [sj-2]/2$ for any $j = 1, \dots, p$. It is possible to construct weekly seasonal adjustment filters in this way.

B.2 Weekly SARIMA Models

A drawback of constructing AC seasonal adjustment filters, as described above, is that the user has little control over the relative degree of smoothing over trend and seasonal dynamics. The only parameter is p , which offers little flexibility. In contrast, WK filtering offers a wider class of filters corresponding to model parameter MLEs, which are adapted to the empirical features in the data. A classical decomposition posits trend, seasonal, and irregular latent processes (McElroy and Politis, 2020), where these models are either estimated through the structural approach or are algebraically derived from a process model via a so-called canonical decomposition (Hillmer and Tiao, 1982).

The general form of such latent processes is as follows: let $\{S_t^{(1)}\}$ be trend, $\{S_t^{(2)}\}$ be seasonal, and $\{S_t^{(3)}\}$ the irregular (a white noise), specified as

$$\begin{aligned} (1 - B)^d S_t^{(1)} &= \Psi^{(1)}(B) \epsilon_t^{(1)} \\ U_{[s/2],s}(B) S_t^{(2)} &= \Psi^{(2)}(B) \epsilon_t^{(2)}. \end{aligned}$$

The right hand side specification of $\Psi^{(1)}(z)$ and $\Psi^{(2)}(z)$ can correspond to ARMA or SARMA models (for example), and typically $d = 1, 2$. Then from standard theory (Bell, 1984) the WK filter for the trend-irregular component has

frequency response function equal to a bounded function times $|U_{[s/2],s}(e^{-i\lambda})|^2$, indicating that all frequencies of the form $\lambda = 2\pi k/s$ for integer k will be annihilated.

Although such models can be estimated structurally, there may be an advantage to fitting a SARIMA process model and obtaining the latent models by decomposition. To do so, we need an extension of the SARIMA model to non-integer periods. Ladiray and Mazzi (2018) provide a so-called fractional SARIMA model by a Taylor series approximation to the differencing operator $1 - B^s$. To explain their approach, let $s = [s] + \langle s \rangle$, so that $\langle s \rangle$ is the fractional remainder. By Taylor series expansion around unity, $z^r \approx 1 + r(1 - z)$ for $0 \leq r \leq 1$. Therefore

$$\begin{aligned} 1 - z^s &= 1 - z^{[s]}z^{\langle s \rangle} \approx 1 - z^{[s]}(1 + \langle s \rangle(1 - z)) \\ &= 1 - (1 + \langle s \rangle)z^{[s]} + \langle s \rangle z^{[s]+1}. \end{aligned}$$

This operator does not annihilate dynamics of frequency $\lambda = 2\pi k/s$, except in the basic case that $\langle s \rangle = 0$, so this does not seem to be the correct approach. The deficiency can be rectified by taking infinitely many terms in the Taylor series expansion, but this is of no practical use because the differencing operator will no longer be a polynomial. In contrast, the approach adopted above in constructing $U_{n,s}(z)$ is to begin with the frequencies one seeks to annihilate, and then incorporate all the minimal degree polynomials that accomplish this nullification.

We describe an alternative approach to constructing fractional period SARIMA models, which we hope to implement in future versions of **Ecce Signum**. Recall that

$$(1 - z)U_{[s/2],s}(z) = 1 - z^s$$

when s is an even integer. Plugging in $z = \vartheta^{1/s}B$ for some $\vartheta \in (0, 1]$, we see that operators of the form $1 - \vartheta B^s$ appearing in SARIMA model specifications can be replaced, for fractional s , by

$$(1 - \vartheta^{1/s}B)U_{[s/2],s}(\vartheta^{1/s}B). \quad (2)$$

For instance, a SARIMA(p,q,1,1) could be written as

$$(1 - B)^2 U_{[s/2],s}(B) \phi(B) (1 - \Phi^{1/s}B) U_{[s/2],s}(\Phi^{1/s}B) X_t = \theta(B) (1 - \Theta^{1/s}B) U_{[s/2],s}(\Theta^{1/s}B) \epsilon_t.$$

This does not work if Φ or Θ take negative values. Also, for higher order specifications, we need to include more and more terms of type (2); however, these would be used only rarely.

C Functionality

There are over 70 functions in the **Ecce Signum** package, which fall into three categories: (i) general calculation functions, (ii) algorithms for multivariate time series, and (iii) functions particular to **sigex** model structures. In the first category are functions for performing algebraic operations, as well as some generic time series functionality; these have no prefix in their nomenclature. In the second category are functions written for multivariate time series to compute likelihoods or generate forecasts, and do not presume any particular model forms but instead are based on the autocovariance function. These have an “mvar” prefix in their nomenclature. The third category works with the presumed model and parameter structures of **sigex**, and have a “sigex” prefix in their nomenclature. Tables 1, 2, 3, 4, 5, and 6 provide an alphabetical list (prefix is ignored in the alphabetization) of the package’s functions, along with a brief description as well as which functions it calls and is called by.

Acknowledgments

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

References

- Bell, W. (1984). Signal extraction for nonstationary time series. *The Annals of Statistics*, 12:646–664.
- Cleveland, W. P. and Scott, S. (2007). Seasonal adjustment of weekly time series with application to unemployment insurance claims and steel production. *Journal of Official Statistics*, 23((2)):209–221.
- Cleveland, W. S., Dunn, D. M., and Terpenning, I. J. (1978). Sabl: A resistant seasonal adjustment procedure with graphical methods for interpretation and diagnosis. *Seasonal analysis of economic time series, NBER*, pages 201–241.

-
- Harvey, A. C. and Trimbur, T. M. (2003). General model-based filters for extracting cycles and trends in economic time series. *Review of Economics and Statistics*, 85((2)):244–255.
- Hillmer, S. C. and Tiao, G. C. (1982). An arima-model-based approach to seasonal adjustment. *Journal of the American Statistical Association*, 77:63–70.
- Ladiray, D. and Mazzi, G. (2018). X11-like seasonal adjustment of daily data. In *2nd International Conference on Advanced Research Methods and Analytics*.
- Ladiray, D. and Quenneville, B. (2012). *Seasonal adjustment with the X-11 method (Vol. 158)*. Springer, Science & Business Media.
- McElroy, T. S. and Politis, D. N. (2020). *Time Series: A First Course with Bootstrap Starter*. CRC Press, New York.
- Trimbur, T. S. (2006). Properties of higher order stochastic cycles. *Journal of Time Series Analysis*, 27:1–17.

Function	Purpose	Calls	Called By
<code>sigex.acf</code>	Compute the autocovariance function of a differenced latent component	<code>ARMAauto</code> <code>sigex.canonize</code> <code>sigex.getcycle</code> <code>polymulMat</code> <code>polymult</code> <code>polysum</code> <code>specFact</code> <code>specFactmvar</code> <code>VARMAauto</code>	<code>sigex.cast</code> <code>sigex.lik</code> <code>sigex.midcast</code> <code>sigex.resid</code> <code>sigex.signal</code> <code>sigex.whittle</code>
<code>sigex.add</code>	Adds a latent component to model		<code>sigex.reduce</code>
<code>sigex.adhocextract</code>	Compute signal extractions and MSE from an ad hoc filter	<code>sigex.cast</code> <code>sigex.midcast</code> <code>sigex.psi2par</code>	
<code>ARMAauto</code>	Compute autocovariances of ARMA		<code>sigex.acf</code> <code>sigex.momfit</code>
<code>sigex.blocktoep</code>	Generates lower triangular array for block Toeplitz matrix		<code>sigex.signal</code>
<code>sigex.bundle</code>	Writes fitted model information to a list		
<code>sigex.canonize</code>	Computes canonization of a given ARMA process	<code>polymult</code> <code>polysum</code> <code>sigex.specFact</code>	<code>sigex.acf</code> <code>sigex.momfit</code> <code>sigex.spectra</code>
<code>sigex.cast</code>	Computes forecasts and aftcasts without uncertainty	<code>sigex.acf</code> <code>sigex.delta</code> <code>mvar.forecast</code> <code>sigex.param2gcd</code> <code>sigex.zeta2par</code> <code>sigex.zetalen</code>	<code>sigex.wkextract</code>
<code>sigex.castextract</code>	Formats casts for plotting		
<code>sigex.conditions</code>	Computes condition number for a covariance matrix	<code>getGCD</code> <code>sigex.param2gcd</code>	<code>sigex.reduce</code>
<code>sigex.constrainreg</code>	Determines constraint matrix for regressors	<code>sigex.default</code> <code>sigex.par2psi</code>	
<code>daily2monthly</code>	Converts a daily time series to a monthly time series	<code>date2day</code> <code>day2date</code> <code>day2week</code>	
<code>sigex.daily2weekly</code>	Embeds a daily time series as a weekly time series	<code>date2day</code> <code>day2week</code>	

Table 1: **Ecce Signum** Function Taxonomy.

Function	Purpose	Calls	Called By
date2day	Converts month-day-year date to a day index		day2date day2week sigex.daily2weekly
datefinder	Finds a holiday date in month-day-year format	date2day day2date day2week date2day	
day2date	Converts a day index to a date in month-day-year format		
day2week	Converts a date in month-day-year format and finds the day of week	date2day	sigex.daily2weekly
sigex.default	Initializes param with zeroes	sigex.eta2psi sigex.psi2par sigex.zetalen	sigex.mvar2uvar
sigex.delta	Computes differencing polynomial with factors omitted	polymult	sigex.cast sigex.frf sigex.lik sigex.lpmse sigex.meaninit sigex.midcast sigex.momfit sigex.reg sigex.resid sigex.signal sigex.whittle sigex.wkmse sigex.default sigex.mlefit sigex.tstats
sigex.eta2psi	Transforms η to ψ		
sigex.extract	Computes signal extraction estimates with two standard errors		
sigex.fixed	Computes all non-trend fixed regression effects		
mvar.forecast	Computes forecasts and predictors		sigex.cast
sigex.frf	Computes signal extraction filter frequency response function	sigex.delta sigex.spectra	sigex.wk
sigex.gausscheck	Tests for normality		
sigex.getcycle	Computes AR and MA polynomials for Butterworth and Balanced cycles	polymult	sigex.acf sigex.momfit sigex.spectra

Table 2: **Ecce Signum** Function Taxonomy.

Function	Purpose	Calls	Called By
<code>sigex.getfrf</code>	Computes and plots signal frequency response function	<code>sigex.frf</code>	
<code>getGCD</code>	Computes the generalized Cholesky decomposition		<code>sigex.conditions</code> <code>sigex.momfit</code>
<code>gethol</code>	Generates holiday regressors from holiday dates	<code>date2day</code>	
<code>sigex.glr</code>	Computes log likelihood ratio statistic for nested models	<code>sigex.lik</code>	
<code>sigex.graph</code>	Graphs signal extraction estimates to an existing time series plot		
<code>sigex.hi2low</code>	Embeds a high frequency filter as a low frequency filter		
<code>sigex.i2rag</code>	Generate list of missing value indices		<code>sigex.midcast</code>
<code>sigex.lik</code>	Computes Gaussian divergence of model	<code>sigex.acf</code> <code>sigex.delta</code> <code>mvar.midcast</code> <code>sigex.param2gcd</code> <code>sigex.zeta2par</code> <code>sigex.zetalen</code>	<code>sigex.glr</code> <code>sigex.mlefit</code>
<code>sigex.load</code>	Loads data into a time series object		
<code>sigex.lpfiltering</code>	Computes trend and cycle extractions by combining with a low-pass filter	<code>sigex.lpmse</code> <code>sigex.lpwk</code>	
<code>sigex.lpmse</code>	Computes trend and cycle MSE by combining with a low-pass filter	<code>sigex.delta</code> <code>sigex.spectra</code> <code>sigex.wkmse</code>	<code>sigex.lpfiltering</code>
<code>sigex.lpwk</code>	Computes trend and cycle filters by combining with a low-pass filter	<code>sigex.wk</code>	<code>sigex.lpfiltering</code>
<code>sigex.meaninit</code>	Adds trend regressors to an existing model	<code>sigex.delta</code> <code>sigex.whichtrend</code>	<code>sigex.reduce</code>
<code>mvar.midcast</code>	Computes multi-step imputations and predictors of a multivariate process via Levinson-Durbin algorithm with missing values		<code>sigex.lik</code> <code>sigex.midcast</code> <code>sigex.resid</code>

Table 3: **Ecce Signum** Function Taxonomy.

Function	Purpose	Calls	Called By
<code>sigex.midcast</code>	Computes predictors for variables at various indices	<code>sigex.acf</code> <code>sigex.delta</code> <code>sigex.i2rag</code> <code>mvar.midcast</code> <code>sigex.param2gcd</code> <code>sigex.zeta2par</code> <code>sigex.zetalen</code>	<code>sigex.wkextract</code>
<code>sigex.mlefit</code>	Fits model to the data via MLE	<code>sigex.eta2psi</code> <code>sigex.lik</code> <code>sigex.par2psi</code> <code>sigex.psi2eta</code> <code>sigex.psi2par</code>	
<code>sigex.momfit</code>	Computes initial parameter estimates by method of moments	<code>ARMAauto</code> <code>sigex.canonize</code> <code>sigex.delta</code> <code>sigex.getcycle</code> <code>getGCD</code> <code>polymult</code> <code>polysum</code>	
<code>sigex.mvar2uvar</code>	Transform multivariate parameters to implied univariate form	<code>sigex.default</code>	<code>sigex.precision</code>
<code>var2.par2pre</code>	Computes VAR pre-parameters	<code>VARMAauto</code>	<code>sigex.par2zeta</code>
<code>sigex.par2psi</code>	Transforms param to ψ	<code>sigex.par2zeta</code>	<code>sigex.mlefit</code> <code>sigex.reduce</code> <code>sigex.par2psi</code> <code>sigex.cast</code> <code>sigex.conditions</code> <code>sigex.lik</code> <code>sigex.midcast</code> <code>sigex.psi2par</code> <code>sigex.resid</code> <code>sigex.whittle</code> <code>sigex.zeta2par</code> <code>sigex.acf</code>
<code>sigex.par2zeta</code>	Transforms param to ζ	<code>var2.par2pre</code>	
<code>sigex.param2gcd</code>	Maps real vector to a unit lower triangular matrix		
<code>polymulMat</code>	Computes the product of two matrix polynomials		
<code>polymult</code>	Computes the product of two polynomials		<code>sigex.acf</code> <code>sigex.canonize</code> <code>sigex.delta</code> <code>sigex.getcycle</code> <code>sigex.momfit</code> <code>sigex.spectra</code> <code>sigex.acf</code> <code>sigex.canonize</code> <code>sigex.momfit</code> <code>sigex.spectra</code>
<code>polysum</code>	Compute the sum of two polynomials		
<code>sigex.portmanteau</code>	Portmanteau tests for autocorrelation		
<code>var2.pre2par</code>	Generates a stable VAR process		<code>sigex.zeta2par</code>

Table 4: **Ecce Signum** Function Taxonomy.

Function	Purpose	Calls	Called By
<code>sigex.precision</code>	Compares signal extraction MSE arising from multivariate fit with the implied univariate fit	<code>sigex.mvar2uvar</code> <code>sigex.signal</code> <code>sigex.transform</code>	
<code>sigex.prep</code> <code>sigex.psi2eta</code> <code>sigex.psi2par</code>	Applies preliminary transformations to the data Transforms ψ to η Transforms ψ to param	<code>sigex.param2gcd</code> <code>sigex.zeta2par</code> <code>sigex.zetalen</code>	<code>sigex.mlefit</code> <code>sigex.default</code> <code>sigex.mlefit</code> <code>sigex.wkextract</code>
<code>sigex.reduce</code>	Obtain a reduced rank model	<code>sigex.add</code> <code>sigex.conditions</code> <code>sigex.meaninit</code> <code>sigex.par2psi</code> <code>sigex.renderpd</code> <code>sigex.delta</code>	
<code>sigex.reg</code> <code>sigex.renderpd</code> <code>sigex.resid</code>	Adds regressors to an existing model Modifies a covariance matrix so that it is pd Computes residuals from model's Gaussian likelihood	<code>sigex.acf</code> <code>sigex.delta</code> <code>mvar.midcast</code> <code>sigex.param2gcd</code> <code>sigex.zeta2par</code> <code>sigex.zetalen</code>	<code>sigex.reduce</code>
<code>sigex.signal</code>	Computes signal extraction matrix and error covariance matrix	<code>sigex.acf</code> <code>sigex.blocktoep</code> <code>sigex.delta</code>	<code>sigex.precision</code>
<code>sigex.signalcheck</code> <code>sigex.sim</code>	Computes model-based signal extraction diagnostics Simulates a stochastic process	<code>sigex.acf</code> <code>sigex.delta</code> <code>sigex.acf</code> <code>sigex.delta</code> <code>sigex.param2gcd</code> <code>sigex.zeta2par</code> <code>sigex.zetalen</code>	
<code>sigex.specar</code> <code>specFact</code>	Plot AR spectrum of a time series Compute the spectral factorization of a given univariate pd sequence		<code>sigex.acf</code> <code>sigex.canonize</code> <code>sigex.momfit</code> <code>sigex.spectra</code>
<code>specFactmvar</code>	Compute the spectral factorization of a given multivariate pd sequence		<code>sigex.acf</code> <code>sigex.spectra</code>
<code>sigex.spectra</code>	Computes scalar part of spectrum of a differenced latent component	<code>sigex.canonize</code> <code>sigex.getcycle</code> <code>polymult</code> <code>polysum</code> <code>specFact</code> <code>specFactmvar</code>	<code>sigex.frf</code> <code>sigex.lpmse</code> <code>sigex.whittle</code> <code>sigex.wkmse</code>

Table 5: **Ecce Signum** Function Taxonomy.

Function	Purpose	Calls	Called By
<code>sigex.transform</code>	Applies aggregation and transformations		<code>sigex.prep</code>
<code>sigex.tstats</code>	Computes t statistics for parameter estimates	<code>sigex.eta2psi</code>	
<code>ubgenerator</code>	Computes the product of unit root differencing operators		<code>x11filters</code>
<code>VARMAauto</code>	Computes autocovariances of VARMA		<code>sigex.acf</code> <code>var2.par2pre</code>
<code>sigex.weekly2daily</code>	De-embeds a weekly time series as a daily time series	<code>date2day</code> <code>day2week</code> <code>sigex.load</code>	
<code>weekly2date</code>	Obtains start and end dates for a weekly time series	<code>date2day</code> <code>day2date</code> <code>day2week</code>	
<code>sigex.whichtrend</code>	Determines which component is for trend		<code>sigex.meaninit</code>
<code>sigex.whittle</code>	Computes Whittle likelihood of model	<code>sigex.acf</code> <code>sigex.delta</code> <code>sigex.param2gcd</code> <code>sigex.spectra</code> <code>sigex.zeta2par</code> <code>sigex.zetalen</code>	<code>sigex.mlefit</code>
<code>sigex.wk</code>	Computes signal extraction filter coefficients and MSE	<code>sigex.frf</code> <code>sigex.wkmse</code>	<code>sigex.lpwk</code> <code>sigex.wkextract</code>
<code>sigex.wkextract</code>	Computes signal extractions and MSE via WK method	<code>sigex.cast</code> <code>sigex.midcast</code> <code>sigex.psi2par</code> <code>sigex.wk</code>	
<code>sigex.wkmse</code>	Computes signal extraction error spectrum from bi-infinite sample	<code>sigex.delta</code> <code>sigex.spectra</code>	<code>sigex.lpmse</code> <code>sigex.wk</code>
<code>x11filters</code>	Generates x11 trend, seasonal, and seasonal adjustment filters	<code>polymult</code> <code>ubgenerator</code>	
<code>sigex.zeta2par</code>	Transform ζ to param	<code>var2.pre2par</code> <code>sigex.param2gcd</code>	<code>sigex.cast</code> <code>sigex.lik</code> <code>sigex.midcast</code> <code>sigex.psi2par</code> <code>sigex.resid</code> <code>sigex.whittle</code> <code>sigex.cast</code> <code>sigex.default</code> <code>sigex.lik</code> <code>sigex.midcast</code> <code>sigex.psi2par</code> <code>sigex.resid</code> <code>sigex.whittle</code>
<code>sigex.zetalen</code>	Computes the length of ζ		

Table 6: **Ecce Signum** Function Taxonomy.