User Guide for MFSS: Mixed Frequency State Space Models

November 6, 2018

Contents

1 Defining Accumulators

MFSS supports two types of accumulators: sum and averages. For details on how these are set up, see "A Practitioner's Guide and Matlab Toolbox for Mixed Frequency State Space Models".

This section describes how to implement these accumulators with a model. To do so, set up the model (likely a `StateSpaceEstimation` or `MFVAR` object, but possibly a `StateSpace` object if all parameters are known) as if the data were observed at the base frequency, define an `Accumulator` object, and use the accumulator to augment the model to respect the mixed-frequency nature of the observe data.

An accumulator has 3 defining properties:

1. The *index* of the observation in y_t that is observed at a lower frequency than the state (an integer)
2. The *calendar* of when observations occur (a vector with an integer for each period)
3. The *horizon* over which observed values have been determined (a vector with an integer for each period)

The manner in which these properties are set depends on the nature of the mixed-frequency data.

1.1 Regular Accumulators

Most mixed-frequency models are built around data observed at regular intervals. For example, in a monthly model where quarterly data is observed, every quarterly observation occurs every 3rd month.

For regular accumulators, the only things that needs to be specified are the type of accumulator (sum or average) and the horizon over which the data are determined. For a panel of four series where the first is a sum over 3 periods, the second is a simple average, the third is a triangle average of 3 periods, and the last is observed at the base frequency, the `Accumulator` would be defined as:

```
accum = Accumulator.GenerateRegular(y, {'sum', 'avg', 'avg', ''}, [3 1 3 0]);
```

1.2 Custom Accumulators

Accumulators where the observation period changes throughout the sample are called "custom" accumulators. Most commonly, these cases arise due to either the nature of the data - for example, monthly observations from a weekly model occur every 4th or 5th week, depending on the month. It's also possible to handle data that changes frequency mid-sample - for example, data that was collected quarterly at one point but is now reported monthly.

Both sum and average accumulators treatment of *index* are similar but *calendar* and *horizon* are different.

For sum accumulators, the *calendar* ought to be either 0 or 1. In the first high-frequency period of a low-frequency period (i.e., the first month of a quarter), the *calendar* ought to be 0, otherwise it should be 1. The *horizon* of a sum accumulator is not used - it is ignored by `Accumulator`.

For average accumulators, the *calendar* runs from 1 to the number of high-frequency periods within the low-frequency period before restarting at 1 (i.e., cycling through $\{1, 2, 3, 1, 2, 3, \ldots\}$ for a quarterly observation from a monthly model). The horizon of the average accumulator should be a vector of ones unless the data are a "long-difference." For such series (i.e., observed quarterly log-differences from a monthly model), the horizon is set equal to the number of base-frequency periods occurring between differences of the data. For example, in a monthly model with observed Q4/Q4 data, the *calendar* would cycle as earlier and the *horizon* would be specified as 12 for each period.

To create a custom accumulator, use the `Accumulator` constructor. For a sample with `T` periods, we can create an average accumulator for the first series in y_t as a quarterly difference with the following 4 lines:

```
index = 1;
calendar = repmat([1; 2 3], [T/3 1]);
horizon = repmat(3, [T 1]);
accum = Accumulator(index, calendar, horizon);
```

2 Examples

Each example should be self-contained and able to be run from the `examples` subfolder. The examples detailed here come from "A Practitioner's Guide and Matlab Toolbox for Mixed Frequency State Space Models" by Scott Brave, Andrew Butters, and David Kelley. Other examples included with MFSS in the `examples` folder include

- The local-level model of the Nile data
- An ARMA(2,1) model of GDP

These examples hopefully show how to use the toolbox more fully. For details on how to

2.1 Example 1: Dynamic Factor Model

See `examples/pgmtmfss1_dfm.m`.

We estimate a mixed-frequency dynamic factor model on a panel of 5 time series:

- The quarterly log-difference of Real Gross Domestic Product
- The monthly log-difference of All Employees: Total Nonfarm Payrolls
- The monthly log-difference of Real personal income excluding current transfer receipts
- The monthly log-difference of Industrial Production Index
- The monthly log-difference of Real Manufacturing and Trade Industries Sales

The monthly series are normalized have a mean of zero and a standard deviation of one. Data retreived from FRED as of August 2, 2018. Identification requires a sign and scale normalization which is accomplished through setting the loading on GDP (in the Z matrix) to 1.

2.2 Example 2: Vector Autoregression

See `examples/pgmtmfss2_var.m`.

We run 3 VAR models, each containing 4 series. In the quarterly and mixed-frequency VARs, these are

- The log-level of Real Gross Domestic Product
- The log-level of Consumer Price Index for All Urban Consumers: All Items
- The log-level of Producer Price Index for All Commodities
- The level of the effective federal funds rate.

For a monthly VAR, GDP is replaced by the log-level of All Employees: Total Nonfarm Payrolls. Data retrieved from FRED as of October 10, 2018.

2.3 Example 3: Trend-Cycle Decomposition

See `examples/pgmtmfss3_trend_cycle.m`.

We estimate a trend-cycle decomposition on the log-level of GDP. The stochastic trend-cycle decomposition model is as follows:

$$
\begin{aligned}
y_t &= \mu_t + \psi_t \\
\mu_t &= \mu_{t-1} + \phi_{t-1} \\
\phi_t &= \phi_{t-1} + \xi_t \\
\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} &= \rho \begin{bmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix},
\end{aligned}
$$

where ξ_t is normally distributed and κ_t and κ_t^* are independently normally distributed with a common variance. The structural parameters ρ and λ are restricted so that the cyclical component remains stationary with an expected period between 1.5 and 12 years. Casting the system into state space form, we have

$$
\begin{aligned}
y_t &= \begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix} \\
\alpha_t &= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \rho\cos\lambda & \rho\sin\lambda \\ 0 & 0 & -\rho\sin\lambda & \rho\cos\lambda \end{bmatrix} \alpha_{t-1} + \begin{bmatrix} 0 \\ \xi_t \\ \kappa_t \\ \kappa_t^* \end{bmatrix}.
\end{aligned}
$$

2.4 Example 4: Natural Rate of Interest

See `examples/pgmtmfss4_r_star.m`.

The data used in this model are

- The log-level of Real Gross Domestic Product (GDP_t)
- The percent change in the PCE Price Index (π_t), its lag (π_{t-1}), the 2nd lag of its 3-month average ($\pi_{t-2,4}$), and the 5th lag of its 4-month average ($\pi_{t-5,8}$)
- Import price inflation ($\pi_{I,t}$)
- Energy import price inflation ($\pi_{O,t}$)
- The level of the effective federal funds rate less 1-year expected inflation derived from rolling autoregressive forecasts (r_t)

The model is comprised of several parts:

- Trend GDP (y_t^*) growing according to a smooth trend growth rate (g_t)
- A trend real interest rate (r_t^*) that follows trend growth and other determinants of the natural rate of interest (z_t)
- Actual GDP ($\tilde{y}_t$) determined by its own lags and the difference of the actual real interest rate from the natural rate of interest
- Actual inflation, determined by its own lags, adjustments for import and energy inflation, and the difference of GDP from its trend

In equations:

$$
\begin{aligned}
\tilde{y}_t &= a_1\tilde{y}_{t-1} + a_2\tilde{y}_{t-2} + \frac{a_r}{2}\sum_{j=1}^{2}\left(r_{t-j} - r_{t-j}^*\right) + \varepsilon_t^y \\
\pi_t &= b_1\pi_{t-1} + b_2\pi_{t-2,4} + (1 - b_1 - b_2)\pi_{t-5,8} + b_y\tilde{y}_{t-1} \\
&\quad + b_I(\pi_{I,t-1} - \pi_{t-1}) + b_O(\pi_{O,t-1} - \pi_{t-1}) + \epsilon_t^\pi \\
\tilde{y}_t &= 100 \times (\mathrm{GDP}_t - y_t^*) \\
r_t^* &= 12cg_t + z_t \\
z_t &= z_{t-1} + \epsilon_t^z \\
y_t^* &= y_{t-1}^* + g_{t-1} + \epsilon_t^{y^*} \\
g_t &= g_{t-1} + \epsilon_t^g
\end{aligned}
$$

Defining the data and state as

$$
\begin{aligned}
y_t &= \begin{bmatrix}\mathrm{GDP}_t & \pi_t\end{bmatrix}^\top \\
x_t &= \begin{bmatrix}\pi_{t-1} & \pi_{t-2,4} & \pi_{t-5,8} & \pi_{I,t-1} - \pi_{t-1} & \pi_{O,t-1} - \pi_{t-1}\end{bmatrix}^\top \\
w_t &= \begin{bmatrix}r_t & r_{t-1}\end{bmatrix}^\top \\
\alpha_t &= \begin{bmatrix}y_t & y_{t-1} & y_t^* & y_{t-1}^* & g_t & g_{t-1} & z_t & z_{t-1} & r_t^*\end{bmatrix}^\top
\end{aligned}
$$

the parameters in state space notation (prior to augmentation for mixed-frequency consistency) are

$$Z = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -b_y & 0 & b_y & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\beta = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ b_1 & b_2 & 1 - b_1 - b_2 & b_I & b_O \end{bmatrix}$$

$$H = \begin{bmatrix} 0 & 0 \\ 0 & \sigma_\pi^2 \end{bmatrix}$$

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 - a_1 & -a_2 & a_1 & a_2 & 1 - 12ca_r/2 & -12ca_r/2 & -a_r/2 & -a_r/2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 12c & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\gamma = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ a_r/2 & a_r/2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \lambda_g/3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_z/(3a_r) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$Q = \begin{bmatrix} \sigma_y^2 & 0 & 0 & 0 \\ 0 & \sigma_y^2 & 0 & 0 \\ 0 & 0 & \sigma_\pi^2 & 0 \\ 0 & 0 & 0 & \sigma_\pi^2 \end{bmatrix}$$

where λ_g and λ_z are the ratio of variances set according to the Stock & Watson median-unbiased treatment estimated by Laubach & Williams, available at `https://www.newyorkfed.org/research/policy/rstar/overview`.