Comparison of Methods Accounting for Outliers during Seasonal Adjustment

J. Aston
Comparison of Methods Accounting for Outliers during Seasonal Adjustment

J. Aston

Institute of Statistical Science
Academia Sinica

Eurostat Seasonal Adjustment Conference 2006
Outline

Outliers

Seasonal Adjustment
   Structural Component Models
   ARIMA models

Outliers
   Gaussian Methods
   Non-Gaussian Methods

Simulations

Real Data
   UK Housing Transactions
   US Automobile Retail Series
   US Material Handling Equipment Manufacturing Series

Summary
Model Based Seasonal Adjustment

Use a component model to define a seasonal component and then remove this estimated component from the data.

\[ y_{t}^{sa} = y_{t} - S_{t} \]

\( S_{t} \) can be defined in many ways and is also dependent on all the other components in the model. The Unobserved Component models often define \( S_{t} \) in terms of trigonometric functions, while ARIMA decomposition models give constrained ARIMA model representations.
Structural Component Models

The Trig-1 seasonal model, with seasonal period $s$, is defined as

$$S_t = \sum_{j=1}^{[s/2]} S_{j,t}$$

$$S_{j,t} = S_{j,t-1} \cos \lambda_j + S_{j,t-1}^* \sin \lambda_j + \omega_{j,t}$$

$$S_{j,t}^* = -S_{j,t-1} \sin \lambda_j + S_{j,t-1}^* \cos \lambda_j + \omega_{j,t}^*$$

and $\omega_{j,t}, \omega_{j,t}^* \sim N(0, \sigma_\omega^2)$ and $\lambda_j = \frac{2\pi j}{s}$. In addition, the local level model,

$$T_t = T_{t-1} + \eta_t$$

where $\eta_t \sim N(0, \sigma_\eta^2)$, for the Trend specification, and a white noise Gaussian irregular component (with variance $\sigma_\xi^2$).

$$\cdots$$
ARIMA models
Airline Model Basics

\[(1 - B)(1 - B^s)y_t = (1 - \theta B)(1 - \Theta B^s)\xi_t\]

Using the Hillmer-Tiao decomposition we can find (for suitable parameters)

\[ (1 - B)^2 T_t = \theta_T(B)\omega_t \]
\[ U(B)S_t = \theta_S(B)\eta_t \]
\[ I_t = \epsilon_t \]

where \( U(B) = (1 + B + \ldots + B^{s-1}) \) and \( \omega_t, \eta_t, \epsilon_t \) are white noise processes

then

\[ y_t = S_t + T_t + I_t \]
Gaussian Outlier Adjustment - Simple Procedure

1. Set Outlier Threshold
2. Check Residuals - perform t-test
3. Include most significant term in model
4. Repeat 2. and 3. until no further significant terms
5. Check current outliers are still applicable
Gaussian Outlier Identification Simulation
Variance Estimation for the differenced process

100 Simulated series (143 observations of monthly data) with 3 outliers. Histogram of estimated variance parameters (true value is 1.0) based on number of outliers detected by automatic Gaussian procedure (threshold value is 3.0)
Non Gaussian Components

\[ I_t^* \sim t(0, \sigma^2_\xi, \nu), \quad t = 1, \ldots, n, \]

where \( \nu > 2 \) is the number of degrees of freedom and \( \sigma^2_\xi \) is the variance, which is constant for any \( \nu \).

The decomposition model with a t-distributed irregular term can be expressed in its canonical form by

\[ y_t = S_t + T_t + I_t^*, \quad I_t^* \sim t(0, \sigma^2_\xi, \nu), \quad t = 1, \ldots, n. \]

where \( t(0, \sigma^2_\xi, \nu) \) refers to the t density.

Structural Component Model: Estimate \( \sigma_\xi, \nu \)

AMB Component Model: Estimate \( \nu \).

Overall model gives \( \sigma_\xi \)
Simulation Setup

- 144/156 Series for both SC and AMB models
- Three outliers at 5 times sd of Irr / No outliers added
- Gaussian / Non-Gaussian Model Estimation
- Mean Squared Seasonal Difference between 144/156 data lengths measured for 144 coincident points.
Comparison Histograms

Trig-1 Three Outliers

Trig-1 No Outliers

AMB Three Outliers

AMB No Outliers
England and Wales Housing Transactions Series
Trig-1 Model

Graph showing seasonal data with two lines, one for Seasonal - Orig + 12 months and another for Seasonal - Orig. There is also a graph showing the difference between the two seasonals.
AMB Model

[Graphs showing seasonal data and differences between Seasonals]
US Automobile Retail Series
Trig-1 Model
AMB Model

Graph showing the comparison of Seasonal - Orig + 12months and Seasonal - Orig, with a difference between Seasonals graph also included.
US Material Handling Equipment Manufacturing Series
Trig-1 Model

Seasonal - Orig + 12months vs. Seasonal - Orig

Difference between Seasonals

Seasonal - Orig + 12months vs. Seasonal - Orig

Difference between Seasonals

Seasonal - Orig + 12months vs. Seasonal - Orig

Difference between Seasonals

Seasonal - Orig + 12months vs. Seasonal - Orig

Difference between Seasonals
AMB Model
Summary

- Critical values can lead to unstable seasonals
- Non-Gaussian models give reasonable seasonals when no outliers present
- Non-Gaussian models reduce instability when outliers present
- Caveats:
  - More Parameters needed
  - Computationally more intensive
References I


References II

