Декомпозиция временных рядов

А.Г. Трофимов
к.т.н., доцент, НИЯУ МИФИ

lab@neuroinfo.ru
http://datalearning.ru

Курс "Статистическая обработка временных рядов"

Сентябрь 2018

Time Series Components

Three components of time series that are typically of interest:

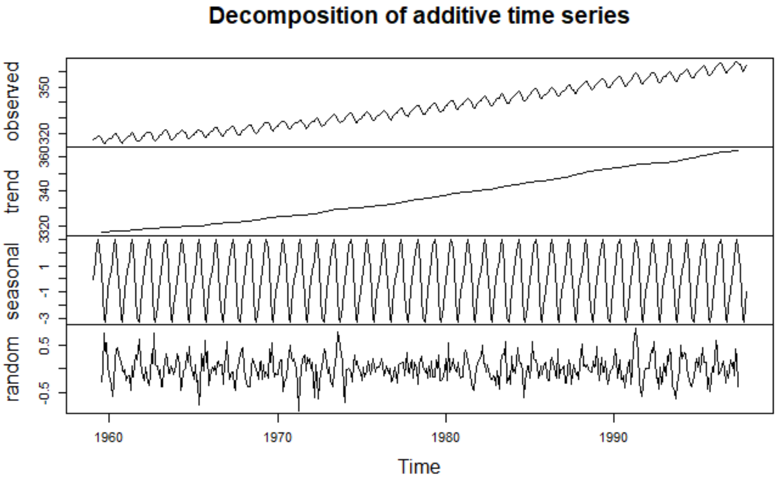
- Trend $\{\mu_t\}$
  It is a deterministic, non-seasonal secular component. Trend usually reflects the long-term evolution of the time series
- Seasonality $\{s_t\}$
  It is a deterministic component with known periodicity. Seasonality occurs over a fixed and known period (e.g., the quarter of the year, the month, or day of the week)
- Irregular stochastic component $\{I_t\}$ ("noise")
  It is a stochastic process (not necessarily a white noise process)

In some applications trend and seasonality can be aggregated and considered as "trend" (a deterministic component of time series)

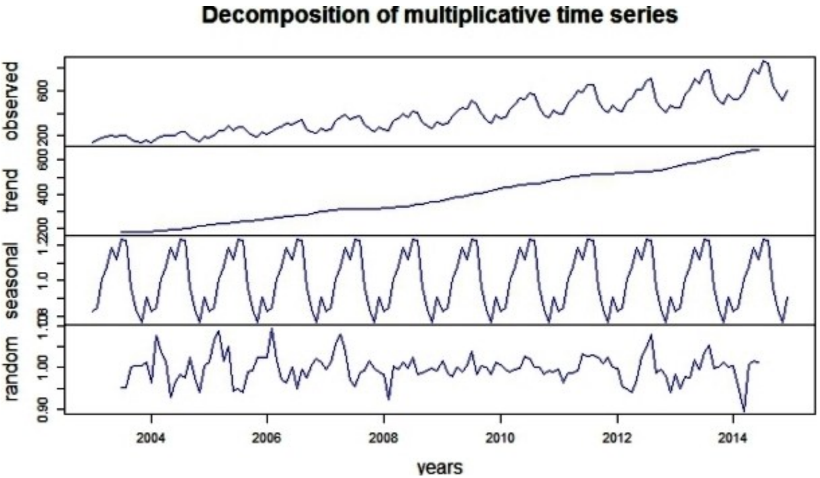
Models of Time Series Decomposition

- Additive decomposition: $Y_t = \mu_t + s_t + I_t$
  It is appropriate when there is no exponential growth in the series, and the amplitude of the seasonal component remains constant over time. The seasonal and irregular components are assumed to fluctuate around zero
- Multiplicative decomposition: $Y_t = \mu_t s_t I_t$
  It is appropriate when there is exponential growth in the series, and the amplitude of the seasonal component grows with the level of the series. The seasonal and irregular components are assumed to fluctuate around one
- Log-additive decomposition: $\log Y_t = \mu_t + s_t + I_t$
  It is an alternative to the multiplicative decomposition. If the original series has a multiplicative decomposition, then the logged series has an additive decomposition

Time Series Decomposition. Illustration 1



Time Series Decomposition. Illustration 2



Detrending and Deseasonalizing

Assume $\{\hat{\mu}_t\}$ and $\{\hat{s}_t\}$ are estimations of trend and seasonal component calculated for the sample path $y_1, ..., y_T$ of the time series $\{Y_t\}$

The time series $\{Y_t - \hat{\mu}_t\}$ (for additive model) or $\left\{\frac{Y_t}{\hat{\mu}_t}\right\}$ (for multiplicative model) is called as detrended time series

The time series $Y_t - \hat{s}_t$ (or $\frac{Y_t}{\hat{s}_t}$) is called as deseasonalized time series

The irregular component $\{I_t\}$ is estimated as:

$$\hat{I}_t = Y_t - \hat{\mu}_t - \hat{s}_t \quad \text{(for additive model)}$$

$$\hat{I}_t = \frac{Y_t}{\hat{\mu}_t \hat{s}_t} \quad \text{(for multiplicative model)}$$

$\{\hat{I}_t\}$ is detrended and deseasonalized time series $\{Y_t\}$

Trend Estimation

Approaches to trend estimation:

- Parametric (using regression models)
  Choose a class of regression models (linear, polynomial, etc.) and fit the model's parameters to the data
  Result: trend model $\mu_t = F(t)$
- Non-parametric (using trend filters)
  Choose a filter (e.g. smoothing) and apply it to the data
  Result: filtered values $z_1, ..., z_T$

Trend estimation is also used in prediction tasks

The future trend values can be predicted with according to:

- estimated trend model: $\mu_{T+k} = F(T+k)$
- filtered values: $\mu_{T+k} = z_T$, $k = 1, 2, \ldots$

Smoothing Filters for Trend Estimation

- Simple moving average
  Parameter: window half-size r

$$z_t = \frac{1}{2r+1} \sum_{i=-r}^{r} y_{t+r}, \quad t = r+1, ..., T-r$$

- Weighted moving average
  Parameters: window half-size r, weights $w_{-r}, ..., w_r$

$$z_t = \sum_{i=-r}^{r} w_i y_{t+r}, \quad t = r+1, ..., T-r$$

- Exponentially weighted moving average (EWMA)
  Parameters: smoothing parameter α, initial value z_1

$$z_t = \alpha y_t + (1-\alpha) z_{t-1}, \quad t = 2, ..., T$$

- Holt's EWMA
- ...

Exponentially weighted moving average (EWMA)

EWMA smoothing filter:

$$z_t = \alpha y_t + (1-\alpha) z_{t-1}, \quad t = 2, ..., T$$

Parameters: smoothing parameter α ($0 \leq \alpha \leq 1$), initial value z_1

$\alpha \approx 0 \Rightarrow$ very inertial filter, low sensibility to new data
$\alpha \approx 1 \Rightarrow$ static, memoryless filter

k-step-ahead forecast:

$$\mu_{T+k} = z_T, \quad k = 1, 2, ...$$

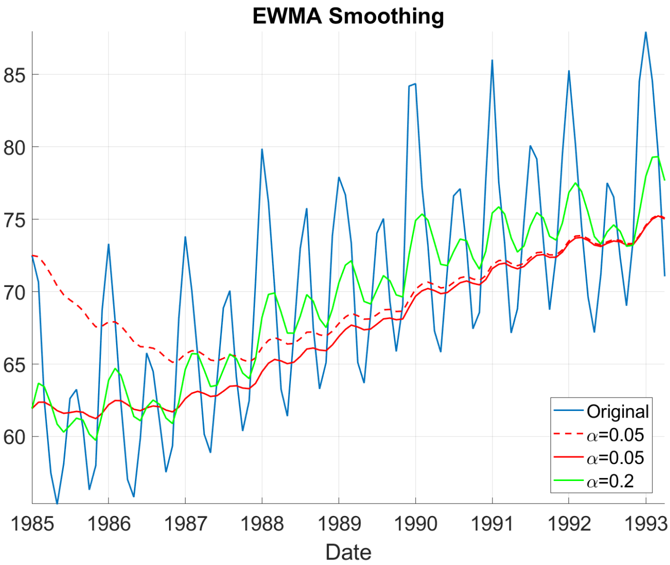
Initial value is usually:

$$z_1 = y_1$$

For sharp time series:

$$z_1 = \frac{1}{r} \sum_{i=1}^{r} y_r$$

EWMA. Illustration



Holt's EWMA (1957)

Holt's EWMA is a modification of EWMA smoothing filter:

$$z_t = \alpha y_t + (1-\alpha)(z_{t-1} + b_{t-1}), \quad t = 2, ..., T$$

$$b_t = \beta(z_t - z_{t-1}) + (1-\beta) b_{t-1}, \quad t = 2, ..., T$$

Parameters: smoothing parameter α $(0 \leq \alpha \leq 1)$, trend smoothing parameter β $(0 \leq \beta \leq 1)$, initial values z_1, b_1

Holt's EWMA smoothing is useful for time series with linear trend

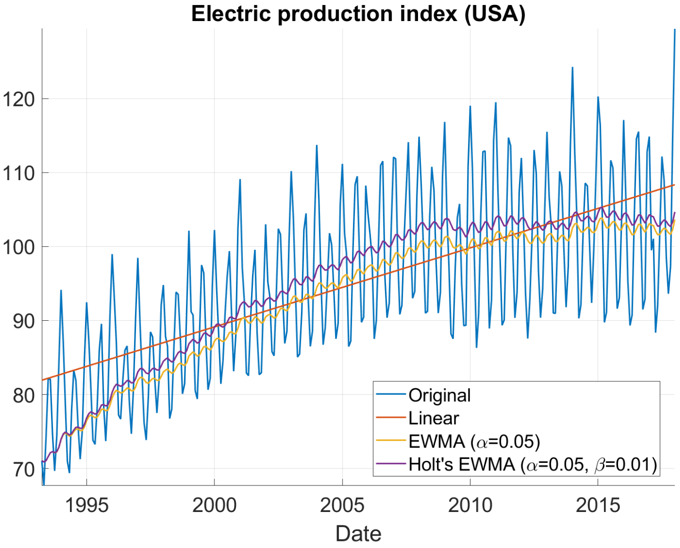
k-step-ahead forecast:

$$\mu_{T+k} = z_T + k b_T, \quad k = 1, 2, ...$$

Initial values are usually:

$$z_1 = y_1, \quad b_1 = 0$$

Trend Estimation. Illustration



Seasonal Adjustment

In some applications, the trend is of primary interest and seasonal component is considered as a nuisance periodic component

Seasonal adjustment is the process of removing a seasonal component

The result of a seasonal adjustment is a deseasonalized time series. Deseasonalized data is useful for exploring the trend and remaining irregular component

To best estimate the seasonal component of a series, you should first estimate and remove the trend component. Conversely, to best estimate the trend component, you should first estimate and remove the seasonal component. Thus, seasonal adjustment is typically performed as an iterative process

Seasonality Estimation

Approaches to seasonality estimation:

- Parametric (using seasonality models)
  Choose a class of seasonality models and fit the model's parameters to the data
  Result: seasonality model $s_t = F(t)$
- Non-parametric (using seasonal filters)
  Choose a seasonal filter and apply it to the data
  Result: seasonal filtered values $s_1, ..., s_T$

Seasonality estimation is also used in prediction tasks

The future seasonal values can be predicted with according to:

- estimated seasonal model: $s_{T+k} = F(T+k)$
- seasonal filtered values: $s_{T+k} = s_{T+k-\Delta}$, $k = 1, 2, ...$, where Δ is a period

Seasonal Filters

Definition

Seasonal moving average filter is a convolution of weights and observations made during past and future periods:

$$\tilde{s}_{k+i\Delta} = \sum_{j=-r}^{r} w_j y_{k+(i+j)\Delta}, \quad k = 1, ..., \Delta, \quad i = 0, 1, ...$$

where $\tilde{s}_{k+i\Delta}$ is a smoothed value at k-th point (day of week, month of year, etc.) in i-th cycle, r is a smoothing window half-size, $(2r+1)$ is a number of periods to smooth

There is a problem of end-point estimates when using seasonal filters. The end-point values can be estimated using asymmetric weights or other special filters

Stable Seasonal Filter

Stable seasonal filter is a seasonal filter with equal weights and maximal number of smoothed periods (N_Δ):

$$\tilde{s}_k = \frac{1}{n_k} \sum_{j=0}^{N_\Delta} y_{k+j\Delta}, \quad k = 1, ..., \Delta$$

where $\tilde{s}_k$ is a mean value at k-th point over all cycles, $k = 1, ..., \Delta$

The output of stable seasonal filter:

$$\tilde{s}_1, ..., \tilde{s}_\Delta, \tilde{s}_1, ..., \tilde{s}_\Delta, ...$$

This time series has length T and only Δ different values

The stable seasonal filter is useful when the seasonal levels does not change over cycles or the time series is too short (up to 5 cycles)

$S_{n\times m}$ Seasonal Filter

$S_{n\times m}$ seasonal filter is a symmetric n-term moving average of m-term averages. This is equivalent to taking a symmetric, unequally weighted moving average with $n+m-1$ terms

Example: $S_{3\times 5}$ filter is a seasonal filter with weights

$$(1/15, 2/15, 1/5, 1/5, 1/5, 2/15, 1/15)$$

To illustrate, suppose you have monthly data over 10 years. Let Jan_{yy} denote the value observed in January, 20yy. The $S_{3\times 5}$-filtered value for January 2015 is

$$\tilde{Jan}_{15} = \frac{1}{3}\Bigg(\frac{1}{5}(Jan_{12} + Jan_{13} + Jan_{14} + Jan_{15} + Jan_{16}) + \frac{1}{5}(Jan_{13} + Jan_{14} + Jan_{15} + Jan_{16} + Jan_{17}) + \frac{1}{5}(Jan_{14} + Jan_{15} + Jan_{16} + Jan_{17} + Jan_{18})\Bigg)$$

Seasonality Estimation Using Seasonal Filters

The seasonal component $\{\hat{s}_t\}$ is calculated as centred seasonally filtered process $\{\tilde{s}_t\}$

- For stable seasonal filter

$$\hat{s}_t = \tilde{s}_t - \bar{s} \quad \text{(for additive model)}$$

$$\hat{s}_t = \tilde{s}_t / \bar{s} \quad \text{(for multiplicative model)}$$

where $\bar{s} = \frac{1}{\Delta} \sum\limits_{k=1}^{\Delta} \tilde{s}_k$ is an average value in period

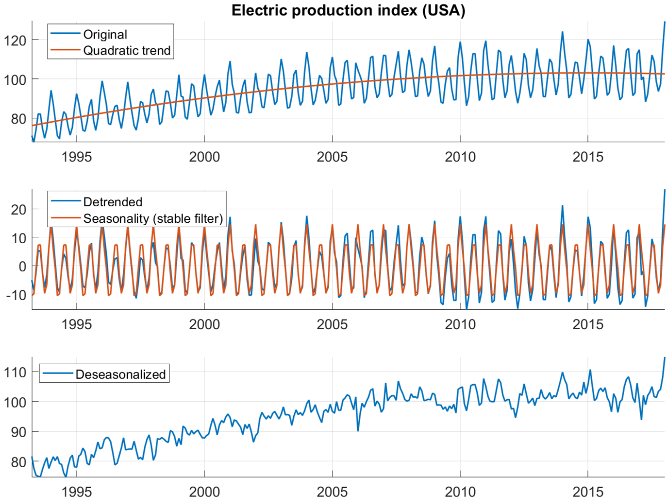
- For $S_{n \times m}$ seasonal filter

$$\hat{s}_t = \tilde{s}_t - \bar{s}_t \quad \text{(for additive model)}$$

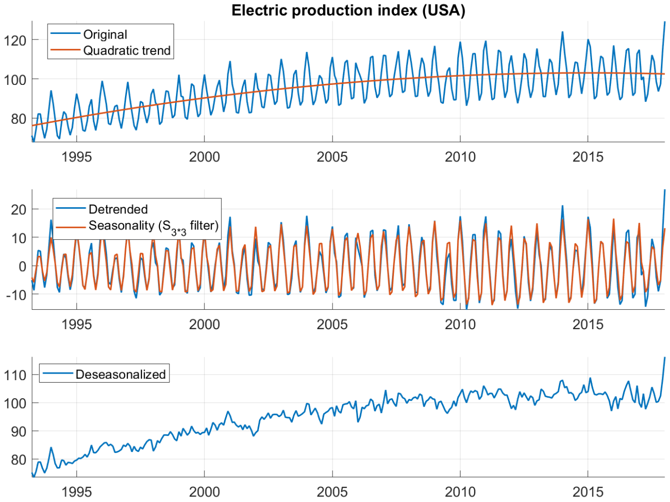
$$\hat{s}_t = \tilde{s}_t / \bar{s}_t \quad \text{(for multiplicative model)}$$

where $\bar{s}_t = \sum\limits_{j=-q}^{q} v_j \tilde{s}_{t+j}$ is a moving average of $\{\tilde{s}_t\}$,
$v_{-q}, \dots, v_q$ are weights, $(2q+1)$ is window size

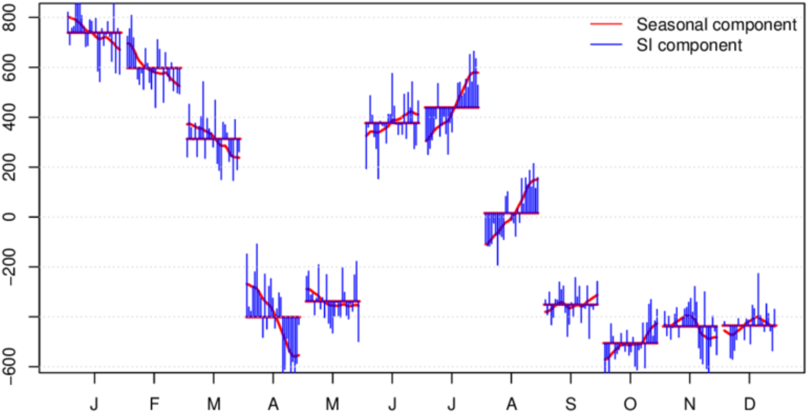
Seasonal Adjustment with Stable Seasonal Filter. Illustration



Seasonal Adjustment with $S_{3\times 3}$ Seasonal Filter. Illustration



Month Plot



The straight red line shows the average seasonal component (stable filtered), red line shows smoothed seasonal component ($S_{n \times m}$ filtered), blue line shows detrended observations

The X12 Seasonal Adjustment Method

The X12 method for time series adjustment were developed by the U.S. Bureau of Census* (1998) and soon it was adopted by many statistical agencies around the world

It has been integrated into X12-ARIMA software and a number of commercially available software packages such as SAS and STATISTICA

The X12 family of methods is based on seasonal filtering to estimate seasonality and seasonally adjust the data

*A principal agency of the U.S. Federal Statistical System, responsible for producing data about the American people and economy

The X12 Method Pipeline

Step 1. Obtain a first estimate of the trend component, $\{\hat{\mu}_t\}$

Moving average or parametric trend estimate

Step 2. Detrend the original series

Calculate $x_t = y_t - \hat{\mu}_t$ for additive decomposition or $x_t = \frac{y_t}{\hat{\mu}_t}$ for multiplicative decomposition

Step 3. Obtain an estimate of the seasonal component, $\{\hat{s}_t\}$, for detrended series $\{x_t\}$

Apply a seasonal filter to $\{x_t\}$ and center the estimate to fluctuate around zero or one, depending on the chosen decomposition. Use a stable seasonal filter or some $S_{n\times m}$ seasonal filter (e.g. $S_{3\times 3}$)

Step 4. Deseasonalize the original series

Calculate $d_t = y_t - \hat{s}_t$ for additive decomposition or $d_t = \frac{y_t}{\hat{s}_t}$ for multiplicative decomposition

The X12 Method Pipeline

Step 5. Obtain a second estimate of the trend component, $\{\hat{\mu}_t\}$, using the deseasonalized series d_t

Moving average or parametric trend estimate

Step 6. Detrend the original series again

Calculate $x_t = y_t - \hat{\mu}_t$ for additive decomposition or $x_t = \frac{y_t}{\hat{\mu}_t}$ for multiplicative decomposition

Step 7. Obtain an estimate of the seasonal component, $\{\hat{s}_t\}$, for detrended series $\{x_t\}$

Apply a seasonal filter to $\{x_t\}$ and center the estimate to fluctuate around zero or one, depending on the chosen decomposition. Use a stable seasonal filter or some $S_{n\times m}$ seasonal filter (e.g. $S_{3\times 5}$)

Step 8. Deseasonalize the original series

Calculate $d_t = y_t - \hat{s}_t$ for additive decomposition or $d_t = \frac{y_t}{\hat{s}_t}$ for multiplicative decomposition. $\{d_t\}$ is the final deseasonalized series