A seasonal model

Hans Karlsen

UiB: 2021-02-10 12:57:21

A basic model for seasonal data

$$X_t = m_t + s_t + Y_t, \qquad t \geq 1 \tag{1.1}$$

where

i) For a postive integer d,

$$s_{t+d} \equiv s_t$$

ii) $\sum_{k=1}^{d} s_k = 0$.

iii) For a positive integer p,

$$m_t = \sum_{j=0}^{r} a_j \, t^j$$

iv) The time series $\{Y(t),\, t \geq 1\}$ is (weakly) stationary with zero mean.

Parametric model

Note that

i) The periodic part of $\{X_t,\ t \geq 0\}$ is represented by the d general different levels $\{s_t,\ t = 1, \ldots, d\}$.

ii) The function m_t is a polynominal trend. If $r \geq 1$ and t is large, this will be the dominating term.

iii) If $\{Y_t,\ t \geq 0\}$ is defined by a parametric time series model, for instance an ARMA(p,q) model, then (1.1) could be seen as a parametric model. The number parameters is $r + d + p + q$.

iv) An alternative model for the seasonal part is

$$s_t = \sum_{k=1}^{M_1} \alpha_k \cos((k/d)2\pi) + \sum_{k=1}^{M_2} \beta_k \sin((k/d)2\pi)$$

The number of parameters is $M_1 + M_2$ and is less or equal to d. This model is called harmonic regression in the textbook.

v) $s_t = s_k$ if $t = k \mod (d)$ and $\{s_k,\ k = 1, \ldots, d\} = \{s_{k+h},\ k = 1, \ldots, d\}$ for any integer h. In R notation the mod is `%%` . An example is `17%%12=5`.

Estimation

i) Estimate the trend first with aid of a filter that eliminates the seasonal component.

ii) Estimate the seasonal component from detrended series.

iii) Eventually estimate parameters of the stationary component when both a trend and a seasonal component are adjusted for.

With n observations

$$\widehat{m}_t = \begin{cases} \sum_{\ell=-q}^{q} X_{t+\ell}, & \text{for } d = 2q+1 \\ \sum_{\ell=-q+1}^{q-1} X_{t+\ell} + 2^{-1} X_{t-q} + 2^{-1} X_{t+q}, & \text{for } d = 2q \end{cases}$$

For $d = 2q+1$,

$$\sum_{\ell=-q}^{q} s_{t+\ell} = \sum_{\ell=1}^{2q+1} s_{t+\ell-q} = \sum_{\ell=1}^{d} s_\ell = 0$$

For $d = 2q$, $s_{t-q} = s_{t+d-q} = s_{t+q}$ so that

$$\sum_{\ell=-q+1}^{q-1} s_{t+\ell} + 2^{-1} s_{t-q} + 2^{-1} s_{t+q} = \sum_{\ell=1}^{2q-1} s_{t-q+\ell} + s_{t+q} = \sum_{\ell=1}^{d} s_{t-q+\ell} = 0$$

This means that $\widehat{m}_t$ does not depend on the seasonal component. But there some edge effects, m_t is defined for $t > q$ and $t \leq n - q$.

Estimating the levels

Let

$$U_t = X_t - \widehat{m}_t, \qquad t \in [q+1, n-q]$$

$$A_k = \left\{ t \in [q+1, n-q] \colon t = k \mod (d) \right\}, \qquad k = 1, \ldots, d$$

$$w_k = \frac{\sum_{j \in A_k} U_{jd+k}}{\sum_{j \in A_k} 1}, \qquad k = 1, \ldots, d$$

$$\widehat{s}'_k = w_k, \qquad k = 1, \ldots, d$$

$$\widehat{s}_k = \widehat{s}'_k - d^{-1} \sum_{j=1}^{d} \widehat{s}'_j$$

since we should have $\sum_{k=1}^{d} \widehat{s}_k = 0$.

Estimating the parameters in the trend

Assume that $d = 2q + 1$.
We use a regression model;

$$\begin{aligned} \widehat{m}_t &= d^{-1} \sum_{\ell=-q}^{q} X_{t+\ell} \\ &= d^{-1} \sum_{\ell=0}^{d-1} X_{t-q+\ell} \end{aligned} \tag{1.2}$$

$$
\begin{aligned}
\widehat{m}_{t+q} &= d^{-1} \sum_{\ell=0}^{d-1} X_{t+\ell} \\
&= d^{-1} \sum_{\ell=0}^{d-1} m_{t+\ell} + d^{-1} \sum_{\ell=0}^{d-1} Y_{t+\ell} \\
&= d^{-1} \sum_{\ell=0}^{d-1} \sum_{j=0}^{r} a_j (t+\ell)^j + e_t
\end{aligned}
$$

Expanding

Now,

$$\begin{aligned}
d^{-1}\sum_{\ell=0}^{d-1}\sum_{j=0}^{r} a_j(t+\ell)^j &= d^{-1}\sum_{\ell=0}^{d-1}\sum_{j=0}^{r} a_j(t+\ell)^j \\
&= d^{-1}\sum_{\ell=0}^{d-1}\sum_{j=0}^{r} a_j \sum_{k=0}^{j}\binom{j}{k} t^k \ell^{j-k} \\
&= \sum_{k=0}^{r}\left\{d^{-1}\sum_{j=k}^{r} a_j \sum_{\ell=0}^{d-1}\binom{j}{k}\ell^{j-k}\right\} t^k \\
&= \sum_{k=0}^{r} b_k t^k
\end{aligned}$$

$$b_k$$

with

$$b_k = \sum_{j=k}^{r} \left(d^{-1} \sum_{\ell=0}^{d-1} \binom{j}{k} \ell^{j-k} \right) a_j$$

$$= \sum_{j=0}^{r} c_{kj} a_j$$

with

$$c_{kj} = \begin{cases} 0, & \text{for } j < k; \\ 1, & \text{for } j = k; \\ d^{-1} \sum_{\ell=0}^{d-1} \binom{j}{k} \ell^{j-k}, & \text{for } j \geq k. \end{cases}$$

Least square

Hence

$$\mathbf{b} = \mathbb{C}\,\mathbf{a}$$

and

$$\widehat{m}_{t+q} = \sum_{r=1}^{r} b_j t^j + e_{t+q}, \qquad t = 1 \leq n - 2q$$

Let $V_t = \widehat{m}_{t+q}$, $m = n - 2q$ and $u_t = e_{t+q}$. Then

$$V_t = \sum_{r=1}^{r} b_j t^j + u_t, \qquad t = 1, \ldots, m$$

We estimate the b_js by least square.

$$\begin{aligned} \mathbf{V} &= \mathbb{T}\mathbf{b} + \mathbf{u} \\ \widehat{\mathbf{b}} &= (\mathbb{T}^T\mathbb{T})^{-1}\mathbb{T}^t\mathbf{V}. \end{aligned}$$

Matrix formulation

The solution for $\mathbf{a}$ is

$$\widehat{\mathbf{a}} = \mathbb{C}^{-1}(\mathbb{T}^T\mathbb{T})^{-1}\mathbb{T}^T\mathbf{V}$$

with

$$\mathbf{V} = \begin{bmatrix} \widehat{m}_{q+1} \\ \widehat{m}_{q+1} \\ \vdots \\ \widehat{m}_{n-q} \end{bmatrix},$$

$$t_{ij} = i^j, \qquad i = 1, \ldots, m, \qquad j = 0, \ldots, r.$$

The observations are only present in the vector $\mathbf{V}$ and $\mathbf{V} = \mathbb{A}\mathbf{X}$ for the matrix $\mathbb{A}$ defined by (1.2).