W02 Важно / сложно?

W01. STL для разделения ряда на компоненты.

$\boxed{\text{W02. – MSTL; – простые стат. модели}}$

MSTL – обобщение STL на случай сезонных эффектов нескольких частот.

$$\begin{pmatrix} y_1 \\ \vdots \\ y_T \end{pmatrix} \to \boxed{MSTL} \to \begin{pmatrix} s_1^{(1)} \\ \vdots \\ s_T^{(1)} \end{pmatrix}, \begin{pmatrix} s_1^{(2)} \\ \vdots \\ s_T^{(2)} \end{pmatrix}, \begin{pmatrix} s_1^{(3)} \\ \vdots \\ s_T^{(3)} \end{pmatrix}, \begin{pmatrix} trend_1 \\ \vdots \\ trend_T \end{pmatrix}, \begin{pmatrix} rem_1 \\ \vdots \\ rem_T \end{pmatrix}$$

единороги?

MSTL – это STL неск-ко раз.

Инт. идея: высокочастотную компоненту можем оценить точнее, чем низкочастотную, будем очищать ряд от сез-сти с высоких частот.

Вход: $(y_t)_{t=1}^{T}$

Шаг 1. С помощью STL выделим высокочаст. сез.-сть

$$(y_t) \to \begin{matrix} (seas_t^{(1)}) \\ (trend_t) \\ (rem_t) \end{matrix} \qquad y_t^{(-1)} = y_t - seas_t^{(1)}.$$

Шаг 2. С – // – STL выделим среднечаст. сезонность

$$(y_t^{(-1)}) \to \begin{matrix} seas_t^{(2)} \\ trend_t \\ rem_t \end{matrix} \qquad y_t^{(-1,2)} = y_t^{(-1)} - seas_t^{(2)}$$

Шаг 3. – // – низкочаст.

$$y_t^{(-1,2,3)} = y_t^{(-1,2)} - seas_t^{(3)}.$$

первичное выделение мн-ва сез-стей.

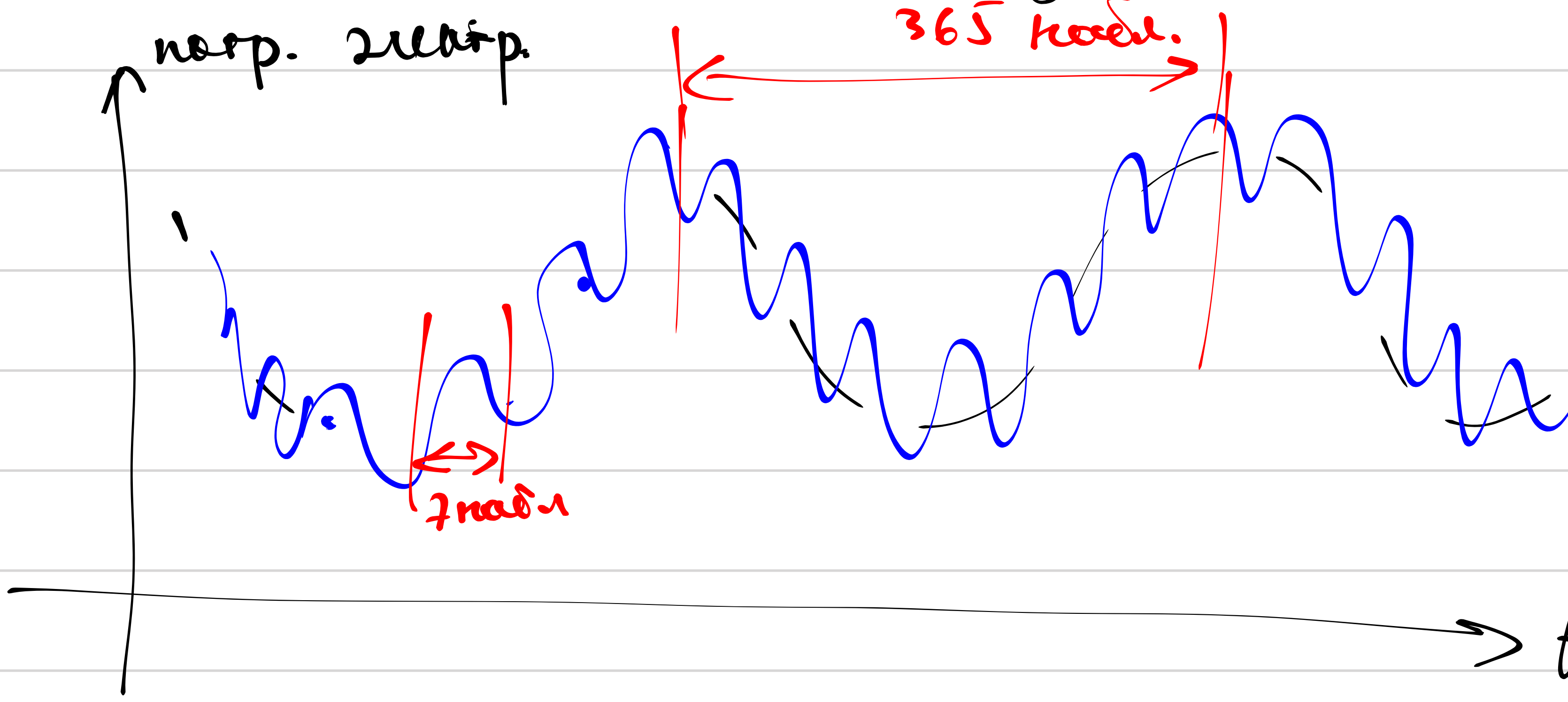
корректировка

Шаг 4. Временно возвращаем $(seas'_t)$ в исх. ряд. Заново извлекаем сезонность высокой частоты.

! обновится и (чуть скорректируется) $(seas_t^{(1)})$ и $(y_t^{(1,2,3)})$

Шаг 5. — // — со средней частотой

Шаг 6. — // — с нулевой частотой.

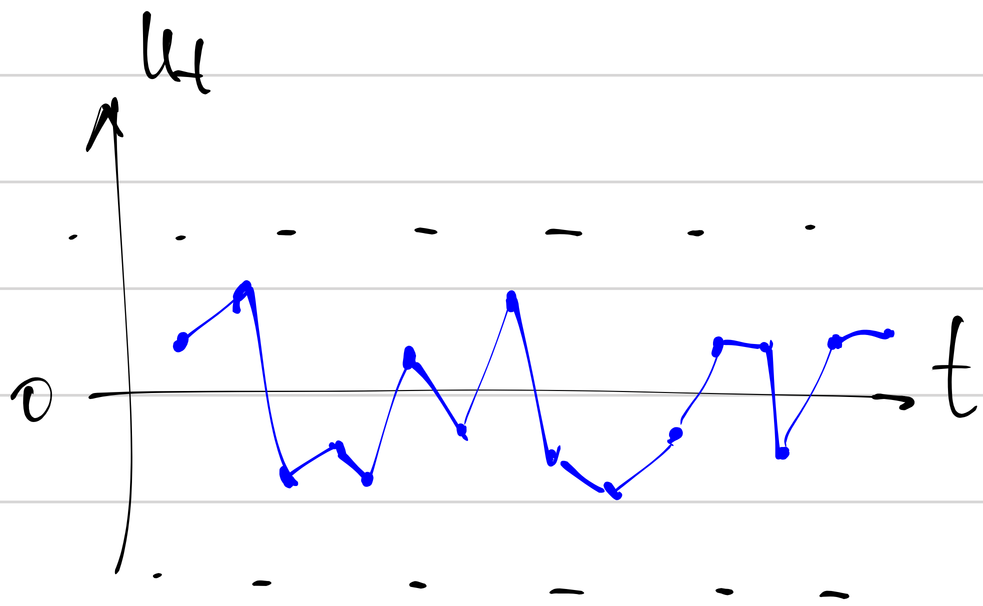


Простые (наивные) модели.

def (u_t) – б. шум, если (white noise)

$$\begin{cases} E(u_t) = 0 \\ Var(u_t) = \sigma^2 \\ Cov(u_t, u_s) = 0 \text{ при } t \neq s \end{cases}$$

не исключает нелинейные зависимости!



def (y_t) - [слабо] стац-ый процесс
[weakly] stationary

$E(y_t) = \mu$

$Var(y_t) = \gamma_0$

$Cov(y_t, y_{t+k}) = \gamma_k$ ← не зависит от t (но может зависеть от k)

th б. шум - это стац-ый процесс.

def. (y_t) - случайное блуждание, если
y_0 - конст
$y_t = y_0 + u_1 + u_2 + \ldots + u_t$, где (u_t) - б. шум.

→ $y_t = y_{t-1} + u_t$ [накапливаем некоррел-ые слагаемые].

Упр. 1) стац-но ли случ-ое блуждание?
2) $y_t = y_{t-1} + u_t$, $u_t \sim$ б. шум,
доп: $u_t \sim N(0; \sigma^2)$ и независимы

σ^2 известно. посср-ть предиктивный интервал (PI) для $y_{T+1}, y_{T+2}, y_{T+h}$ при обуч-ей 95% выб-ке $[y_1 \ldots y_T]$

① $E(y_t) = E(y_0 + u_1 + \ldots + u_t) = y_0$ [конст]

$Var(y_t) = Var(y_0 + u_1 \ldots + u_t) = 0 + t \cdot \sigma^2 \neq const$

y_t - нестац. процесс

② $(y_{T+1} | y_1 \ldots y_T) \sim ?\ N(y_T; \sigma^2)$

$y_{T+1} = y_T + u_{T+1}$

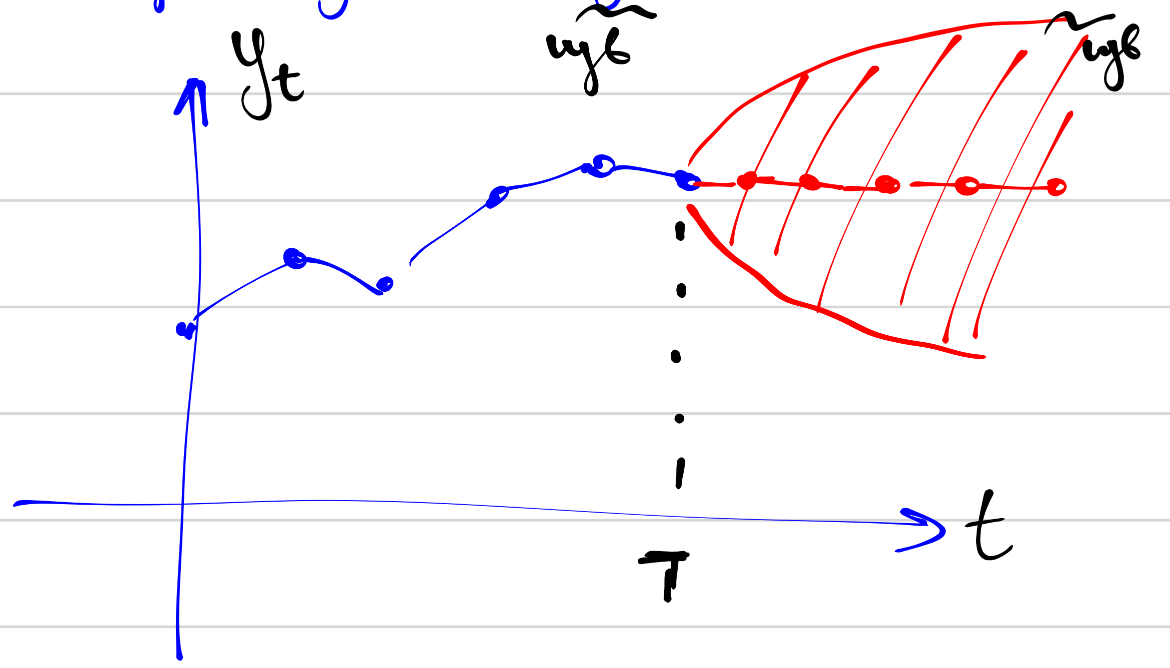
95% PI $[y_T - 1.96 \sqrt{\sigma^2}; y_T + 1.96 \sqrt{\sigma^2}]$

$(y_{T+2} | y_1 \dots y_T) \sim ? \ N(y_T ; 2 \cdot \sigma^2)$

$y_{T+2} = y_T + u_{T+1} + u_{T+2}$

95% PI для y_{T+2} $[y_T - 1.96\sqrt{2\sigma^2} ; y_T + 1.96\sqrt{2\sigma^2}]$

95% PI для y_{T+h} $[\underbrace{y_T - 1.96\sqrt{h \cdot \sigma^2}}_{\text{нуб}} ; \underbrace{y_T + 1.96\sqrt{h \cdot \sigma^2}}_{\text{вуб}}]$.



③ $\boxed{\text{как оценить } \sigma^2 \text{ для модели случайного блуждания?}}$

„Используем метод макс-го правдоподобия"

Δy_t $\quad \Delta y_2 = y_2 - y_1 \sim N(0; \sigma^2)$

$\vdots$

$\Delta y_T = y_T - y_{T-1} \sim N(0; \sigma^2)$

$(T-1)$ штука.

ML: $\quad \hat{\sigma}^2 = \dfrac{\sum_{t=2}^{T} (\Delta y_t)^2}{T-1}$ $\quad$ [упр-т]

вывод. $\begin{pmatrix} \Delta y_2 \\ \vdots \\ \Delta y_T \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} ; \begin{pmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{pmatrix} \right)$

$\max\limits_{\sigma^2} f(\Delta y_2 \dots \Delta y_T)$

$$\max_{\sigma^2} f(\Delta y_2)\cdot f(\Delta y_3)\cdot \ldots \cdot f(\Delta y_T)$$

$$\max_{\sigma^2} \ln f(\Delta y_2) + \ldots + \ln f(\Delta y_T)$$

$$\ln f(\Delta y_2) = \ln\left[\frac{1}{\sqrt{2\pi\sigma^2}}\cdot \exp\left(-\frac{(\Delta y_2 - 0)^2}{2\sigma^2}\right)\right]$$

$$= -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{1}{2}\cdot\frac{(\Delta y_2)^2}{\sigma^2}$$

$$\max_{\sigma^2}\left[-\frac{T-1}{2}\ln(2\pi) - \frac{T-1}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\cdot\sum_{t=2}^{T}(\Delta y_t)^2\right]$$

$$\hat{\sigma}^2_{ML} = \frac{\sum_{t=2}^{T}(\Delta y_t)^2}{T-1}$$

Дост-ый игрок [для месячных/кварт/полугод/годовых]

«Тета-метод»

Шаг 1. [Применяем STL к (y_t), получаем: $seas_t$, $trend_t$, rem_t.]

Шаг 2.

$\widehat{rem}_{T+h} = 0$ [ост-к не прогн.]

[для месяч] $\widehat{seas}_{T+h} = seas_{T+h-12}$ [$h = 1, \ldots, 12$]

$\boxed{\text{для } trend_t \text{ исп-се модель:}}$

[x_t вместо $trend_t$]

$$\begin{cases} x_t = \ell_t + (1-\alpha)\cdot u_t \\ \ell_t = \ell_{t-1} + b + \alpha\cdot u_t \\ u_t \sim N(0; \sigma^2) \text{ незав} \end{cases}$$

пар-ры: $\boxed{\alpha, b, \sigma^2, \ell_0}$

MaxLik: $\hat{\alpha}, \hat{b}, \hat{\sigma}^2, \hat{\ell}_0$

$$\boxed{x_t = l_{t-1} + b + u_t}$$

$$\boxed{x_t = l_t + (1-\alpha)\cdot u_t}$$ усл. независ. на посл. инд.

случ блужд со сносом

$x_t = l_t + (1-\alpha) u_t$

$$\boxed{l_t = l_{t-1} + b + \alpha\cdot u_t}$$

(l_t) - случ. блуждание со сносом

$l_t = z_t + b\cdot t$

сл. блужд.

$$x_{T+1} = l_T + b + u_{T+1}$$

$$(x_{T+1} \mid x_1, \dots, x_T) \sim N(l_T + b; \sigma^2)$$

$$\boxed{\hat{x}_{T+1} = \hat{l}_T + \hat{b}}$$

$$x_{T+2} = l_{T+1} + b + u_{T+2} =$$

$$= l_T + b + \alpha u_{T+1} + b + u_{T+2} =$$

$$= l_T + 2b + \alpha u_{T+1} + u_{T+2}$$

$$(x_{T+2} \mid x_1 \dots x_T) \sim N(l_T + 2b; \sigma^2(\alpha^2 + 1))$$

$$\hat{x}_{T+2} = \hat{l}_T + 2\hat{b}$$

для несезонных рядов: ⊕ предиктивные интервалы.

для сезонных рядов: без предиктивных интервалов.

---

def для стац-х рядов определена автоковариац-нная ф-ция

$$\gamma_k = Cov(y_t, y_{t+k})$$

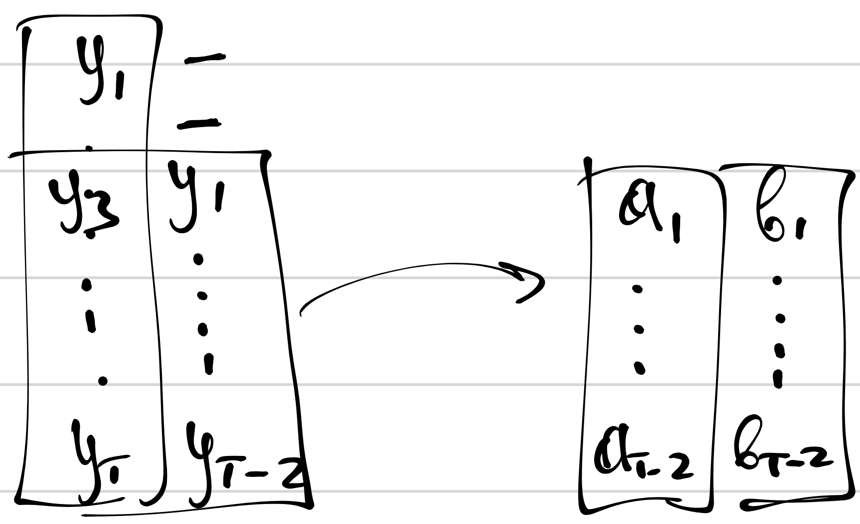
и автокорреляционная ф-ция.

$$\rho_k = Corr(y_t, y_{t+k})$$

для нестац $Cov(y_t, y_{t+k})$ зависит и от t, и от k в общем случае df выборочная ACF (sACF) (s = sample)

$\hat{\rho}_k = \widehat{corr}(y_t, y_{t+k})$, сделанная в предп-ии стац-сти ряд.

$\hat{\rho}_2$



$$\hat{\rho}_2 = s\,corr(a, b) = \frac{\sum_{t=1}^{T-2}(a_t - \bar{a})\cdot(b_t - \bar{b})}{\sqrt{\sum(a_t - \bar{a})^2 \sum(b_t - \bar{b})^2}}$$