Привет !!

Q. ¿Какая ETS практичнее?

A. И1. здоровый смысл: (1) y_t – годовые данные

~~ETS(AAA)~~

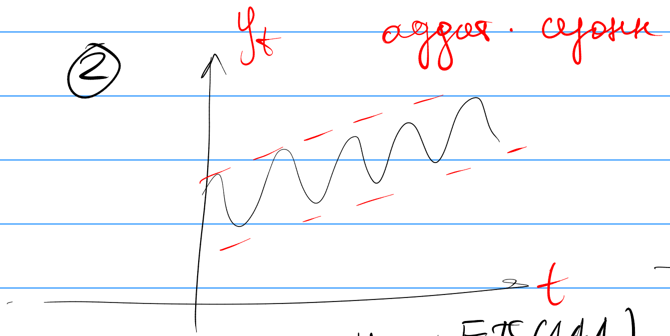
ETS(AMN)

ETS(AAN)

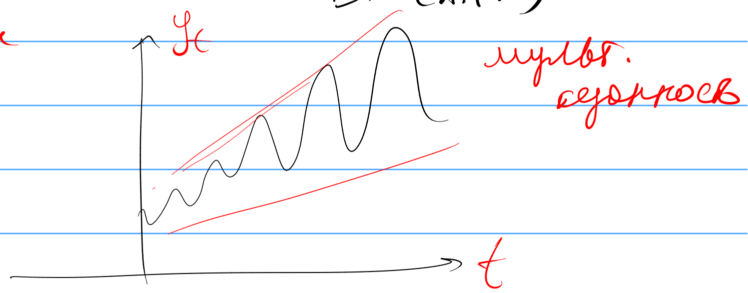
И2. автом. выбор

CV или AIC

(2)



$y_t \sim ETS(AAA)$



$y_t \sim ETS(MAM)$?

$y_t \sim ETS(MA_dM)$?

$\ln y_t \sim ETS(AAA)$?

$bc_\lambda(y_t) \sim ETS(AAA)$?

Тренд: A или Ad

И3. помните о цели

И4. можно оставить несколько конкур-х моделей.

y_t [руб]

идея аддит. ошибки $(\ldots) + u_t$ $\quad u_t$ [руб]

идея мульт. ошибки $(\ldots)\cdot(1+u_t)$ $\quad u_t$ [ ]

часть мультипл. моделей может быть численно неуст-вой

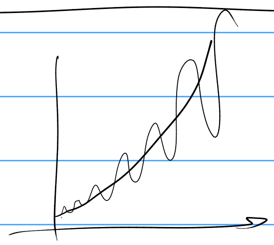
ETS(MMM) ← неуст.

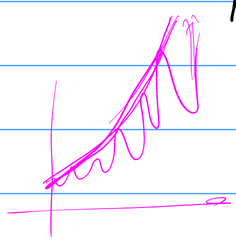
$$
\begin{cases}
u_t \sim N(0;\sigma^2), \text{ незав} & \text{} \} \text{ случ-сть} \\
s_t = s_{t-12}\cdot(1+\gamma u_t) & \} \text{ сезонный множитель} \\
b_t = b_{t-1}\cdot(1+\beta u_t) & \} \text{ скорость роста} \\
l_t = l_{t-1}\cdot b_{t-1}\cdot(1+\alpha u_t) & \} \text{ сглаж., очищ. от сезонности } y_t \\
y_t = l_{t-1}\cdot b_{t-1}\cdot s_{t-12}\cdot(1+u_t) & \} \text{ исход. ряд}
\end{cases}
$$

$\boxed{\text{нач. усл – тоже пар-ры!}}$: $l_0, b_0, s_0, s_1, \ldots, s_{11}$

норм: $s_0\cdot s_1\cdot\ldots\cdot s_{11}=1$

$s_0=1$

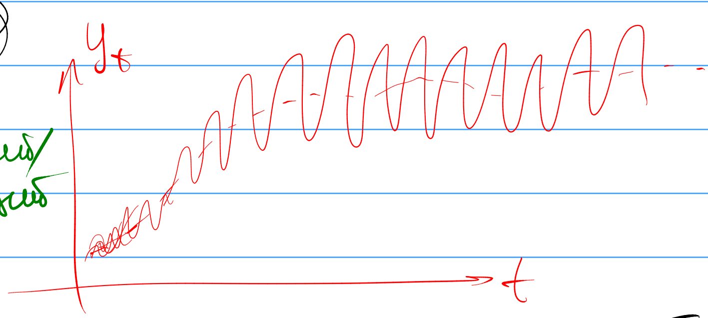




ETS(MA_dM)

u_t[ ]

демпированный/затухающий



$\ln y_t \sim \cancel{ETS(AAA)}$

(корр)
$$
\begin{cases}
u_t \sim N(0;\sigma^2), \text{ незав} \\
s_t = s_{t-12}\cdot(1+\gamma u_t) \\
\underbrace{b_t}_{\text{руб}} = \phi\cdot b_{t-1} + (l_{t-1}+\phi b_{t-1})\cdot \underbrace{\beta\cdot u_t}_{\text{безр.}} \\
l_t = (l_{t-1}+\phi b_{t-1})\cdot(1+\alpha u_t) \\
y_t = (l_{t-1}+\phi b_{t-1})\cdot s_{t-12}\cdot(1+u_t)
\end{cases}
$$

оценка сгл. l_t по прошлому

нач. условия:

$l_0, b_0, s_0, s_1, \ldots, s_{11}$

$s_0\cdot s_1\cdot\ldots\cdot s_{11}=1$

Автоматический выбор.

→ CV. (растущее и скользящее окно).

! дешёвинка ! [не от к ETS]

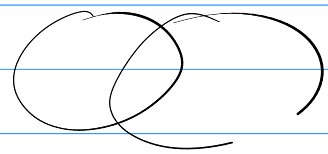
пожарищи предикторов.

y_t	$L y_t$	$\max\{L y_t, L^2 y_t\}$	$\cos(\frac{\pi}{2} t)$	t
y_1	NA	NA		1
y_2	y_1	NA		2
		$\max(y_1, y_2)$		3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_T	y_{T-1}			
?	y_T	$\max(y_T, y_{T-1})$	$\cos(..)$	$T+1$

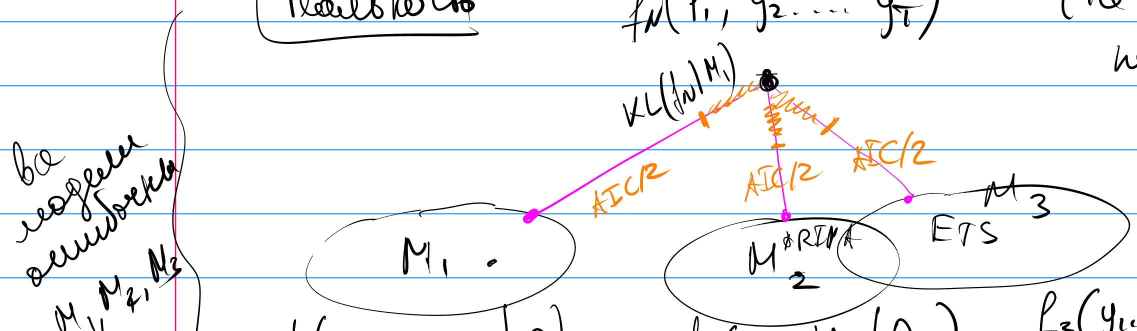
град. бустинг, случ. лес. (это алгоритмы для нерегр. данных)

обычная CV (не под вр ряд) может хорошо работать.

AIC инф. критерий Акаике.



Реальность $f_N(y_1, y_2 \dots y_T)$ [не знаем и никогда не узнаем]



$f_1(y_1 \dots y_T | \theta_1)$ $\quad$ $f_2(y_1 \dots y_T | \theta_2)$ $\quad$ $f_3(y_1 \dots y_T | \theta_3)$

$D_{KL}(f_N \| f_j)$ $\quad$ $KL(f_N \| f_j)$

$$\frac{AIC_j - AIC_k}{2} \approx KL(f_N|M_j) - KL(f_N|M_k)$$

$$\frac{AIC_j}{2} \approx KL(f_N|M_j) + \underbrace{\text{const}}_{\text{зависит от } f_N}$$

(заучите)

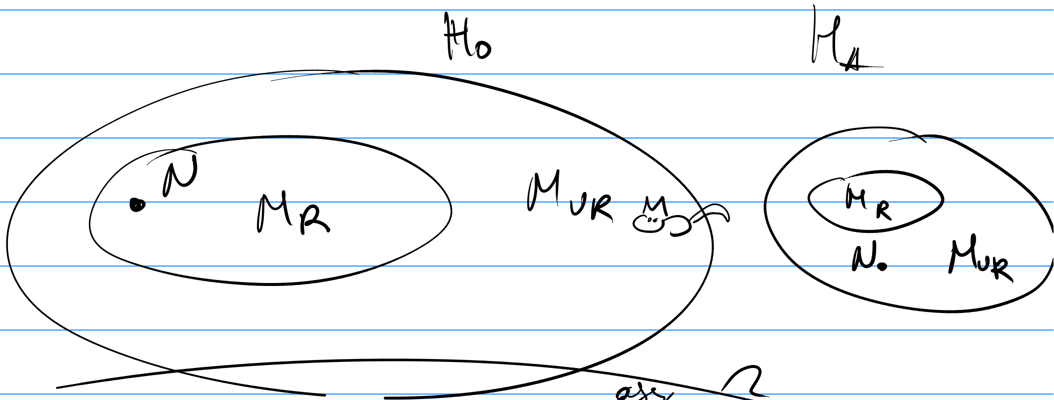
$$AIC = 2\cdot\left[-\ln f(y_1\dots y_T) + k + \frac{k(k+1)}{T-(k+1)}\right]$$

T – число наблюдений

k – число пар-в модели (считая все своб-ые пар-ры)

Q. Зачем в AIC ×2, чтобы потом поделить на 2?

A. В ситуации вложенных моделей, содержащих истину



$$LR = \Delta AIC - \Delta k = AIC_R - AIC_{UR} - \Delta k \overset{asy}{\sim} \chi^2_{\Delta k}$$

Что такое дивергенция KL?

интуиция?

вероятн.	1бит	2бит	3бит	3бит
код	1	01	001	000
y_t	-3	2	5	16
	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

y_t ~ незав.

Истина

α-Centauri

$y_1, y_2, \dots, y_{10^6}$

Земля

$p(u) = P(y_t = u)$

$H(y_t) = E(\log_{\frac{1}{2}} p(y_t))$

опр. ожид-ое число бит на передачу одного y_t (в опт-ой кодировке) — энтропия $H(y_t)$ [бит]

$H(y_1, \dots y_T)$ — " — на передачу $y_1, \dots, y_T$ (в опт. ой кодировке)

Теор. Если $y_1 \dots y_T$ незав. и один. распр. то $H(y_1, \dots y_T) = T \cdot H(y_1)$

Опр. — $CE(f_N \| f_M)$ — сколько бит (в среднем) уходит на передачу сообщ-ия $(y_1 \dots y_T)$, если истинные вер-сти опис-ся f_N, а кодировка оптим-на под f_M (не под f_N).

$$CE(p_N \| p_M) = E_{p_N}\left(\log_{\frac{1}{2}} p_M(y_1 \dots y_T)\right)$$

def $KL(f_N \| f_M) = \underline{CE(f_N \| f_M) - H(f_N)}$ (потери от неопт. кодировки)