Привет !!

Моделир-ие вр-х рядов.

Борис Демешев
wiki.cs.hse.ru

$$5 \, \text{ДЗ}$$
$$\text{кр} + \text{эку}$$

Какие задачи?

1 ряд:

① прогнозировать
  точечное → точит. интервал
  волатильность

② обнаружить момент разладки

③ заполнить пропуски

④ дезагрегировать
  годовые → месячные

⑤ разложить на составляющие

→ Ссылка на Вики !

$y_t \rightarrow$ тренд
$ \rightarrow$ сезонность
$ \rightarrow$ остаток
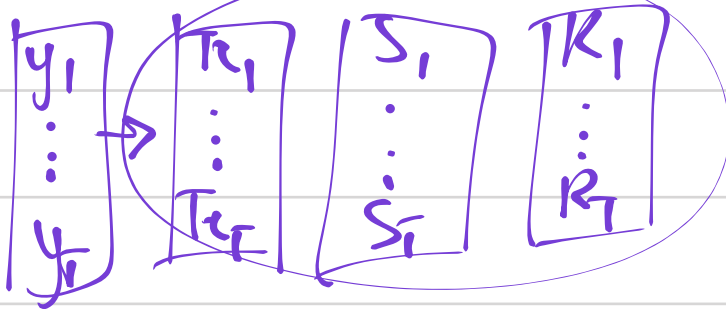
⑥ ...

к рядов:

① старые задачи

② классиф. ряды

③ кластериз. ряды

④ измерять близость рядов

⑤ ...

Заговор: <u>Временные ряды - не ряды!</u>

мат. анализ:
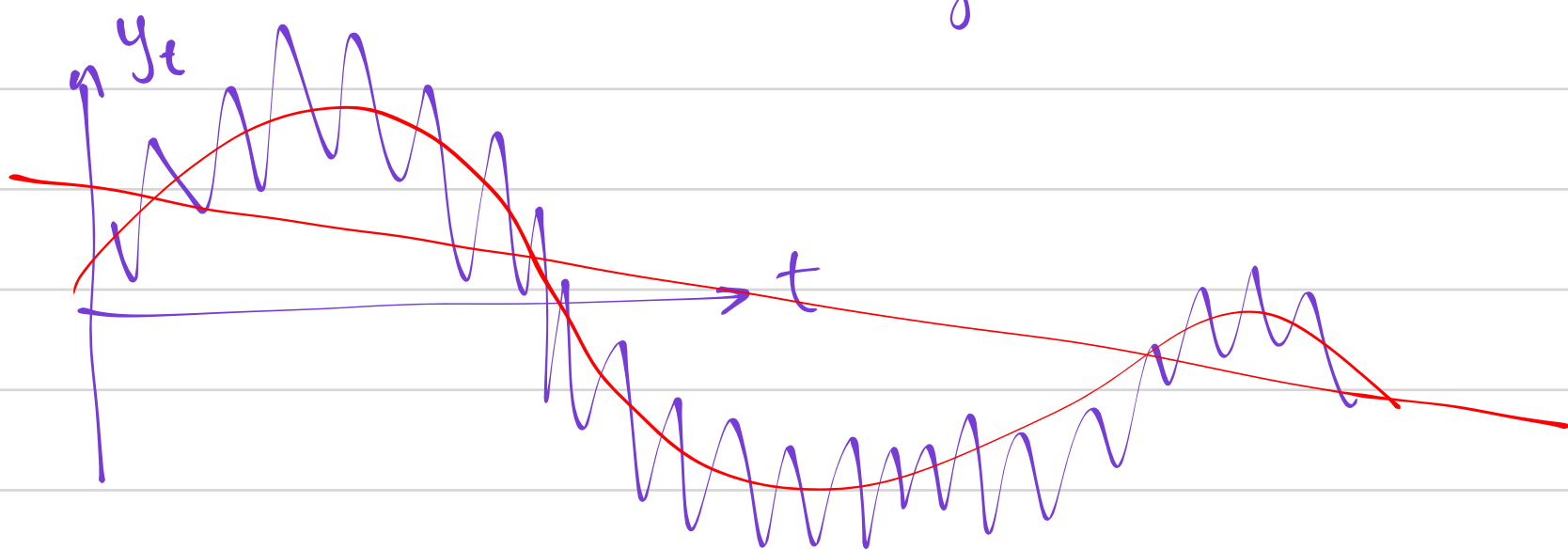
      послед-сть: $a_1, a_2, a_3, \ldots$

      ряд   : $a_1 + a_2 + a_3 + \ldots$

<u>опр</u> <u>врем ряд</u> — послед-сть случ-х величин.

$$(y_t) \qquad (y_t)_{t=1}^{\infty} \qquad \boxed{(y_t)_{t=-\infty}^{t=+\infty}}$$

$\boxed{\text{Сглаживание ряда.}}$

     „Умерие видеть единорогов"

   „идея"   ⟶ модель 1
              ⟶ модель 2
              ⟶ модель 3

$\boxed{\text{форм. идея } 1}$ Взять скользящее среднее (moving average) с окном ширины h.

$$\left\{ \quad \tilde{y}_t = \frac{y_{t-1} + y_t + y_{t+1}}{3} \qquad h = 3 \quad \right\}$$

$$(y_t) \xrightarrow{MA} (\tilde{y}_t) \qquad \text{„} \tilde{y} = MA(y) \text{"}$$

ещё одна формализация идеи сглаживание

LOESS = LOcal regrESSion
LOWESS w = weighted
локальная регрессия.

$\hat{\beta}_1 + \hat{\beta}_2 \cdot t$

сглаж-не проход-от сами

OLS [ordinary least squares]

$\hat{y}_t = \hat{\beta}_1 + \hat{\beta}_2 \cdot t$

цел. ф-ция:
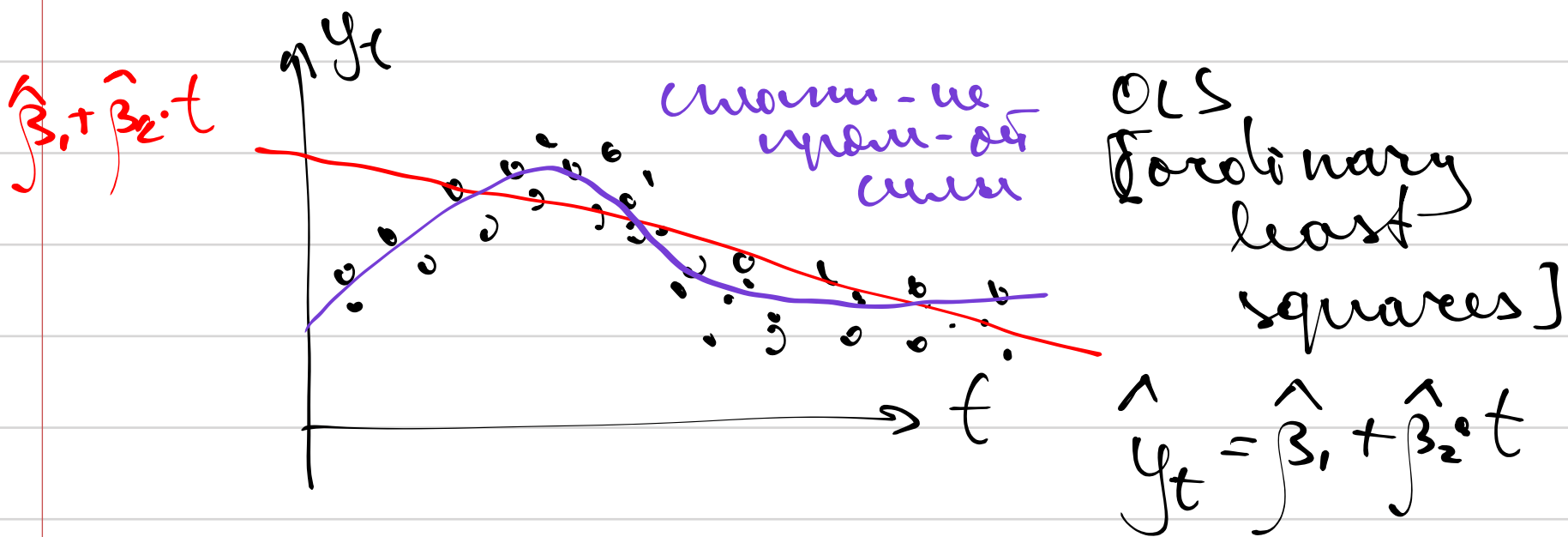
$$\min_{\hat{\beta}_1, \hat{\beta}_2} \sum_{t=1}^{T} \left(y_t - \hat{y}_t \right)^2$$

$$Q(\hat{\beta}_1, \hat{\beta}_2) = \sum_{t=1}^{T} \left(y_t - (\hat{\beta}_1 + \hat{\beta}_2 \cdot t) \right)^2$$

идея LOESS

не вошли для лок-ой регрессии

лок. линия регрессии в т. a.

не вошли для сгл. в т a

где пров-ти лок-ую линию регрессии?

линия лок. об регрессии в т. в.

$$Q(\hat{\beta}_1, \hat{\beta}_2, x) \rightarrow \min_{\hat{\beta}_1, \hat{\beta}_2} \quad \hat{\beta}_1(x) \quad \hat{\beta}_2(x)$$
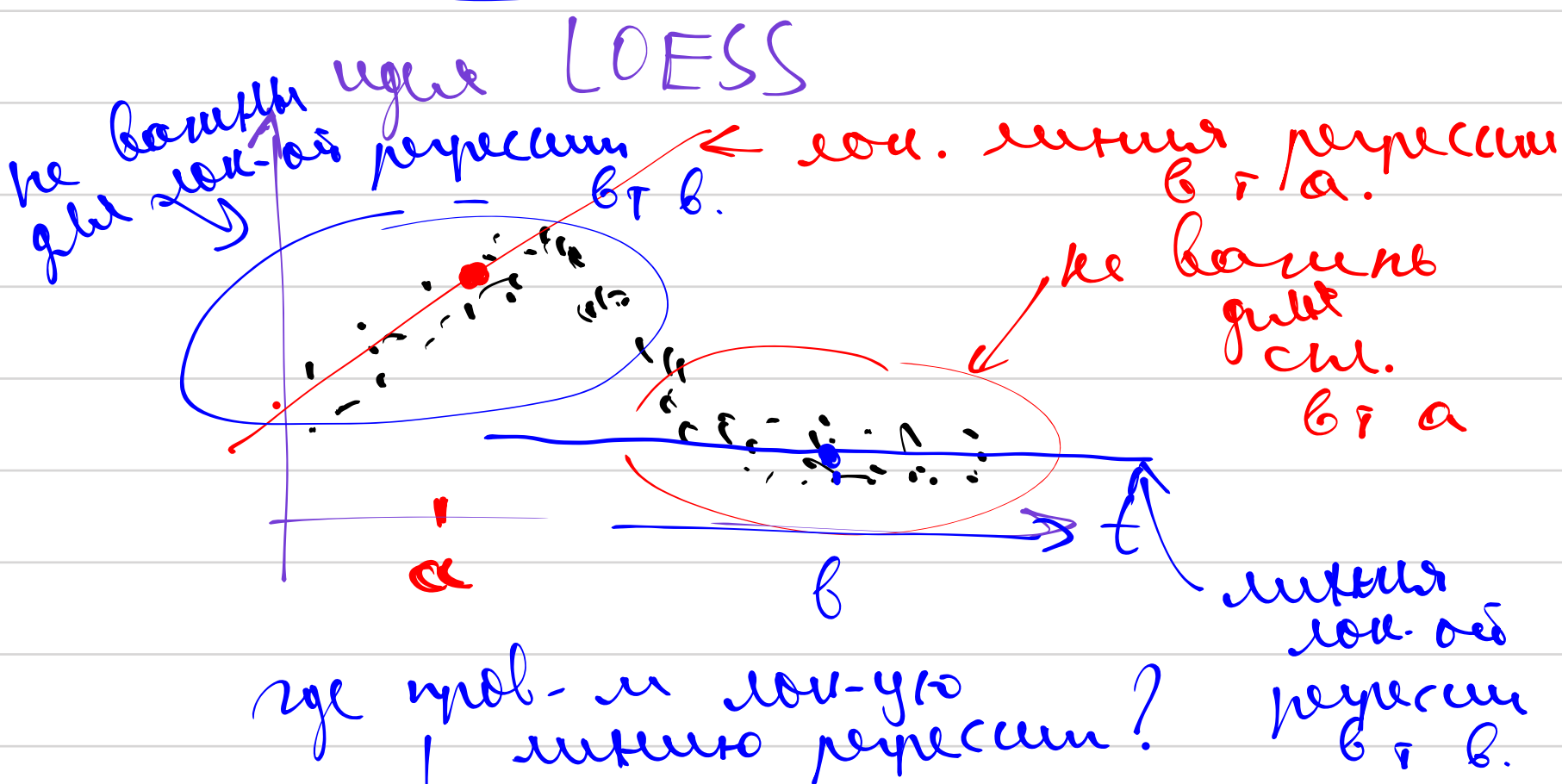
$$Q(\hat{\beta}_1, \hat{\beta}_2, x) = \sum_{t=1}^{T} \underbrace{K(x_t, x)}_{} \cdot (y_t - \hat{y}_t)^2$$

$$\underbrace{\hat{y}_t = \hat{\beta}_1 + \hat{\beta}_2 \cdot x_t}_{\text{[каже t]}}$$

$\underline{k(x_t, x)}$: чем дальше x_t от x,
тем меньше k.

$$k(x_t, x) \geq 0$$

Все модели ложные, но некотор-ые полезны.

яро.   заполняну-ое:

$$k(x_t, x) = \exp\left(-\frac{(x_t - x)^2}{h^2}\right)$$

если $h >> 0$, то $k(x_t, x) \approx 1$ и $\quad \hat{\beta}_1(x) \approx \hat{\beta}_1^{OLS}$
$\hat{\beta}_2(x) \approx \hat{\beta}_2^{OLS}$

$-h \approx 0$

$h >> 0$
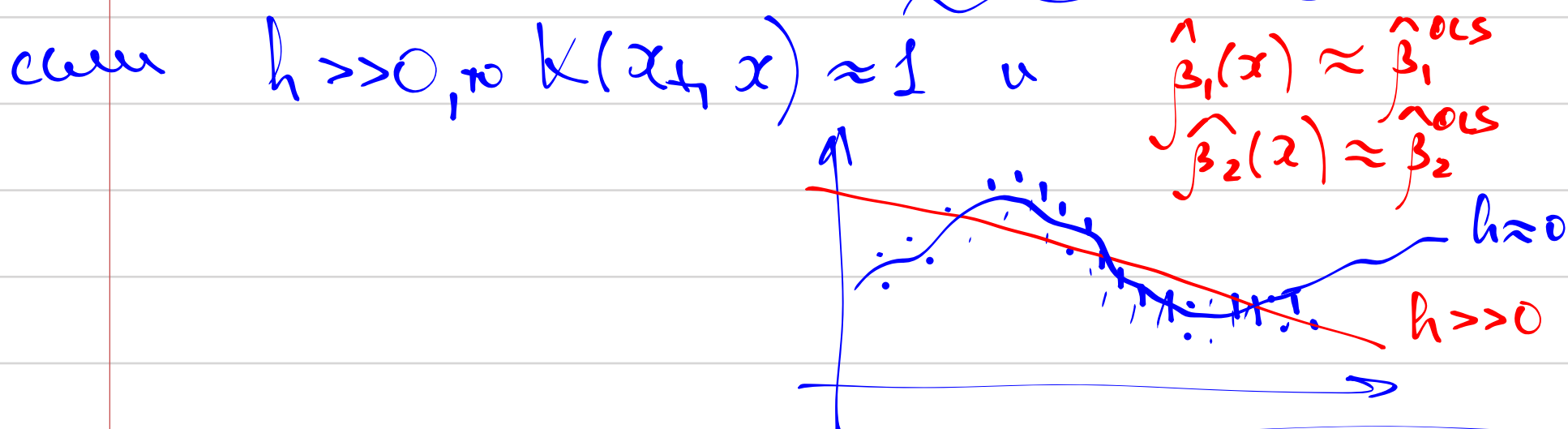
LOESS

$$k(x_t, x) = \begin{cases} 1, & \text{если } x_t - \text{это один} \\ 0, & \text{иначе} \end{cases} \text{ из 5-ти ближ-х соседей } x$$

идея сглаживания $\rightarrow$ MA $\rightarrow$ LOESS

Сглаш. ряд:
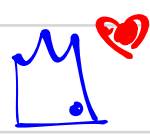
$$\hat{y}_t^{LOESS} = \hat{\beta}_1(t) + \hat{\beta}_2(t) \cdot t$$

инт. идея : разбить ряд на (тренд, сезонную сост-ую и остаток.)
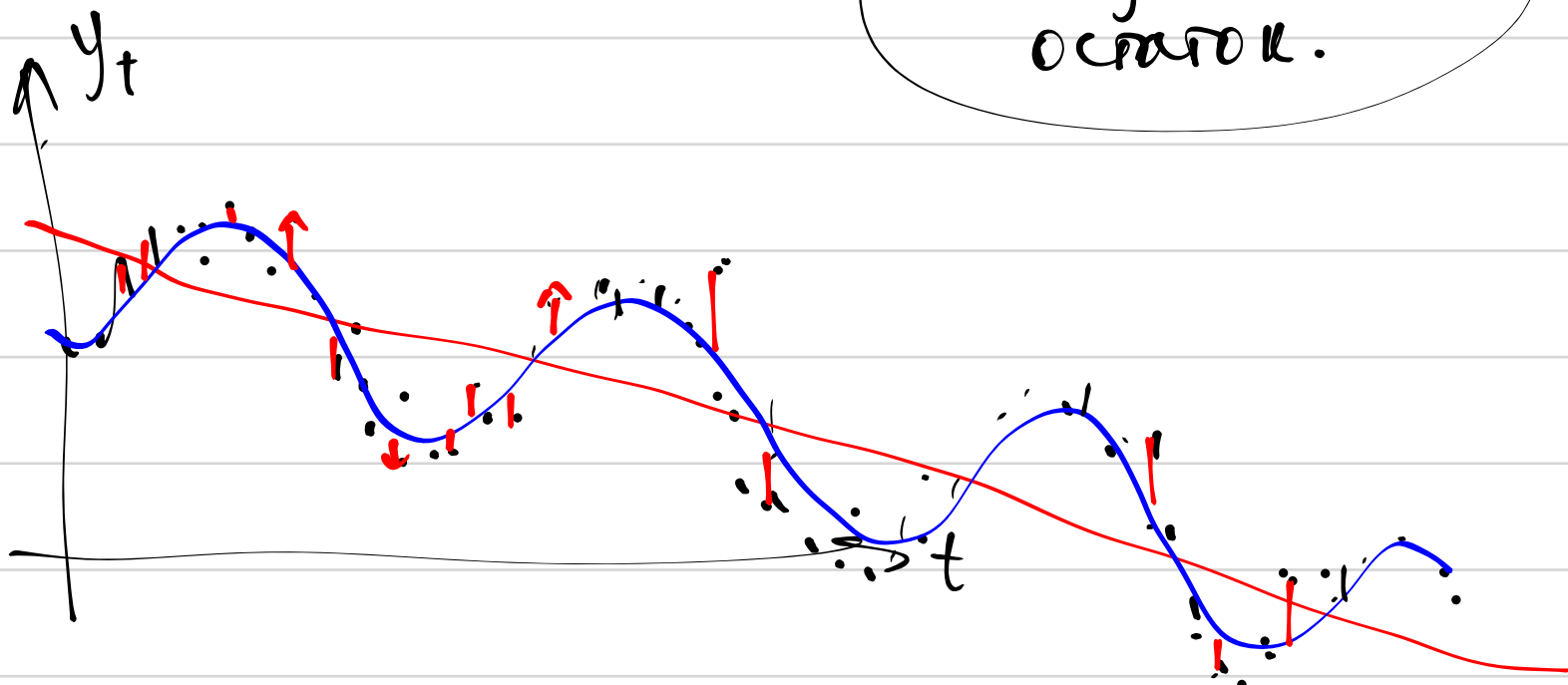
интуитивные ожидания :

остаток: непредсказуемая часть.

тренд : ое. плавно меняющееся.
$$Tr_t \approx Tr_{t-1} + \Delta$$

сезонная сост-ая : $\quad S_t \approx S_{t-12}$

[месяц]

идея разложить ряд на сост-ие $\longrightarrow$ покер-ый алгоритм
(STL), ETS...
PROPHET, DLT...

(STL) - нестатистич. древний быстрый
и робастный {и к выбросам, и пропускам}

цель:
$$y_t = trend_t + season_t + rem_t$$

[STL - это шаманство!]

[да!]

Шаг 1. $Trend_t := 0$

$sez_t := 0$ [Временно зат-ли весь ряд в сезонную сост-ю]

(*)→ **Шаг 2.** Удалили текущий тренд.

$$y_t^{det} = y_t - trend_t$$

Шаг 3. разрезали ряд (y_t) на 12 подрядов.

[на прим? и месячных данных]

короче исх-го в ≈ 12 раз

Шаг ④. сглаживали подряды с помощью LOESS

$$C^{jan} = \underline{LOESS}\left(y_{jan}^{det} \right)$$
$$\dots$$
$$C^{dec} = \underline{LOESS}\left(y_{dec}^{det} \right)$$

Шаг ⑤ собираем $C^{jan} \dots \dots C^{dec}$ в один ряд C

[Шаг 6.] сильное-сильно сглаж-ние для C.

$$L = LOESS(\underline{MA}(\underline{MA}(C))$$
[куча настроек]

Шаг 7. выд-и сезонность.
$$seas_t := C_t - L_t$$

Шаг 8. выд-и тренд
$$trend_t := LOESS(y_t - seas_t)$$

к(шаг?12)
[неск-ко раз].

$$(y_t) \longrightarrow \boxed{STL} \begin{array}{l} \rightarrow trend_t \\ \rightarrow seas_t. \\ \rightarrow remainder_t \end{array}$$

↑ пар-ры
* частота ряда
* силы сглаживания.

Как исп-ть STL (разложение на сост-ие)?
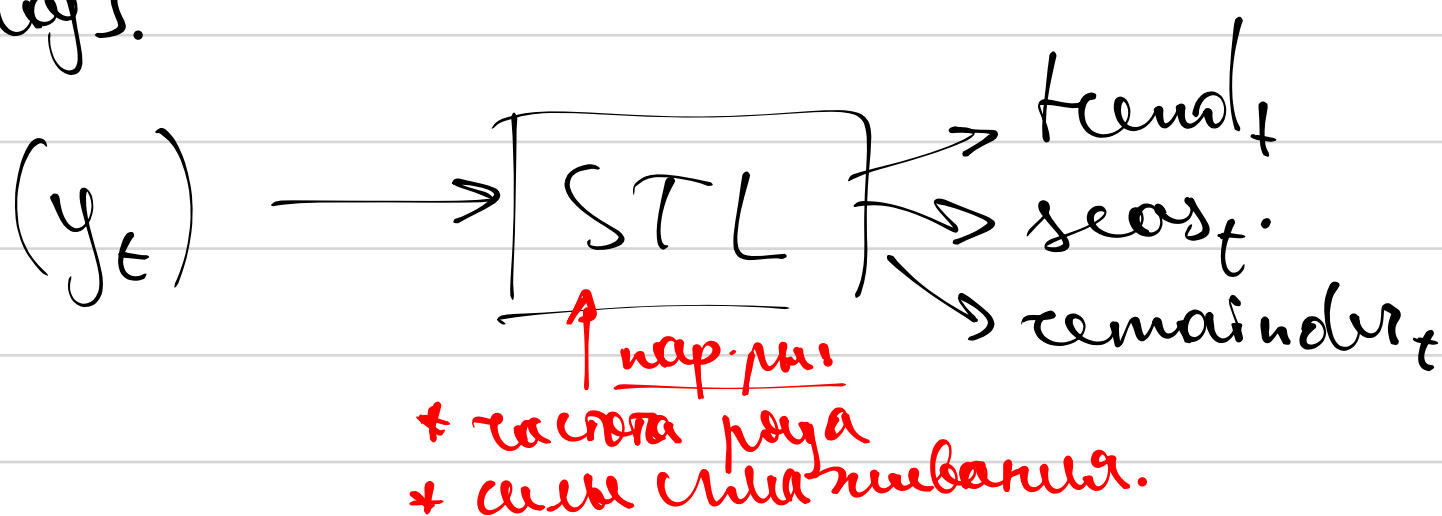
→ для прогнозирования.
наивнейшая: $\widehat{y}_{t+1} = trend_t + seas_{t+1-12} + 0$

→ для получения (хар-к ряда.)

sVar = sample variance

TI - trend intensity

$$TI = \max\left\{0, 1 - \frac{sVar(Rem)}{sVar(Rem + Trend)}\right\}$$

SI - seasonal intensity

$$SI = \max\left\{0, 1 - \frac{sVar(Rem)}{sVar(Rem + Seas)}\right\}$$

Rob Hyndman. ~~Forecos-~~ Forecasting principles & practice