1) Алгоритм подбора Кхандакар – Hyndman

2) ADF, KPSS

ARIMA(p,d,q)

$d = 1, 2, 3$

ARMA(p,q)

$d = 0$

$y_t \sim$ стац. процесс

$\Delta^k y_t$ – стац

$\Delta^{k-1} y_t$ – нестац

$\Delta^{k-1} y_t \ne \alpha + \beta t +$ стац

Q. Как выбрать наилучшую ARIMA

A_1: Если ряды длинные, много ресурсов $\Rightarrow$
кросс-валидация

A_2: алгоритм KH

<u>Шаг 1.</u> С помощью статистических тестов определить порядок интеграции d
(ADF, KPSS)

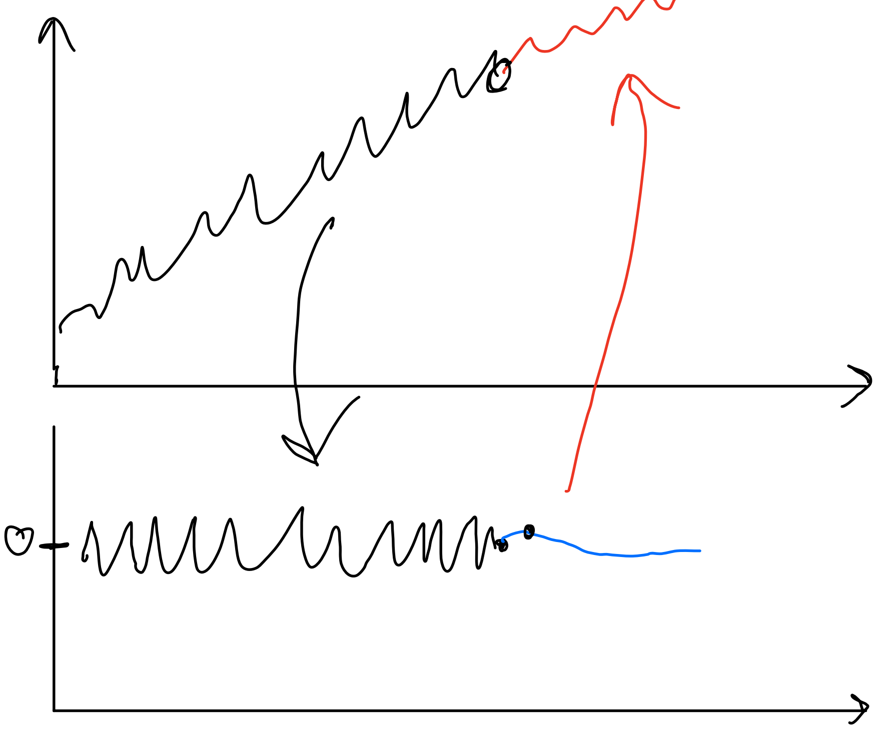
$f(y_1, \dots, y_T) \leftarrow \quad d = 0, \quad y_t \sim ARMA$

$f(\Delta y_2, \dots, \Delta y_T) \leftarrow \quad d = 1, \quad \Delta y_t \sim ARMA$

$f(\Delta^2 y_3, \dots, \Delta y_T) \leftarrow \quad d = 2, \quad \Delta^2 y_t \sim ARMA$

$$y_t \to \Delta^d y_t \xrightarrow{ARMA} \Delta^d \hat{y}_{t+h} \to \hat{y}_{t+h}$$

$$y_t \to \Delta y_t \to \Delta \hat{y}_{t+h} \overset{?}{\to} \hat{y}_{t+h}$$



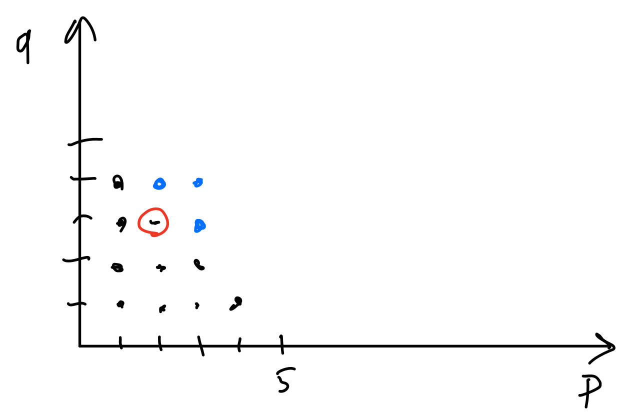
$$\hat{y}_{T+1} = y_T + \Delta \hat{y}_{T+1}$$

$$\hat{y}_{T+h} = y_T + \sum_{k=1}^{h} \Delta \hat{y}_{T+k}$$

Шаг 2. Оценивается пачка ARMA(p, q)

$(p+q \leq 5)$

Выбирается модель с минимальным AICc



1) Почему нельзя брать d с запасом
2) Двойное тестирование ?

parsimonius

KPSS - тест.

Kwiatkowski - Phillips - Schmidt - Shin

$KPSS_c$ " с константой "

$KPSS_t$ " с трендом

Долгосрочная дисперсия ряда

λ^2 - долг. дисп. y_t если

$$Var(\bar{y}) = \frac{\lambda^2}{T} + o\left(\frac{1}{T}\right)$$

$$\lim_{T\to\infty} (T\, Var(\bar{y})) = \lambda^2$$

Если бы $y_1, \ldots, y_T$ были независимы и одинаково распр.

$$Var(\bar{y}) = \frac{\sigma^2}{T}$$

$$\bar{y} = \frac{y_1 + \ldots + y_T}{T}$$

Упр. $y_t = u_t + u_{t-1}$, $u_t - WN(0, \sigma^2_u)$

а) $Var(y_t)$

б) $\lim_{T\to\infty} (T \cdot Var(\bar{y}))$

$$T\,Var(\bar y) = T\,\frac{T\,\overset{\text{"}\gamma_0\text{"}}{Var(y_t)} + 2(T-1)\,cov(y_{t-1}, y_t) + 2(T-2)\,cov(y_{t-2}, y_t) + \dots}{T^2}$$

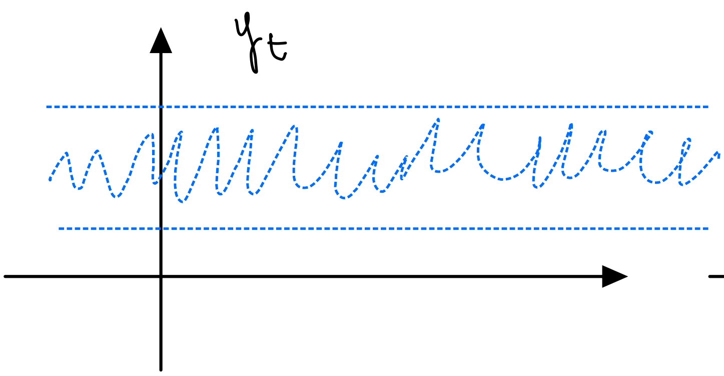
$$\lim(T\,Var(\bar y)) = \gamma_0 + 2\gamma_1 + 2\gamma_2 + \dots = \lambda^2$$

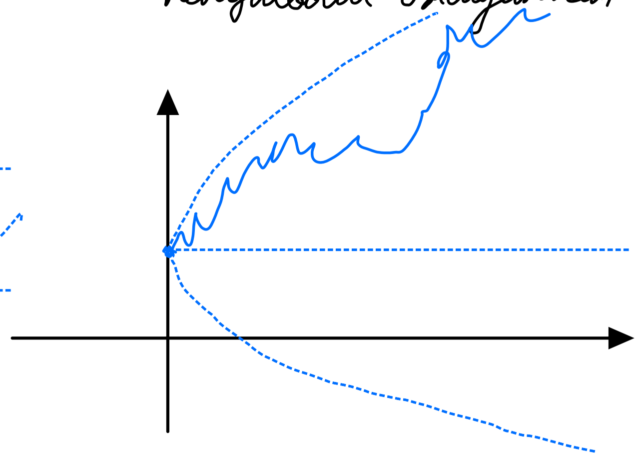
KPSS с

H_0: y_t – стац с ненулевым ожиданием

H_A: $y_t = RW +$ стац. процесс с ненулевым ожиданием

$r_t = r_{t-1} + u_t$





$$y_t = c + r_t + x_t$$

$c \in \mathbb{R}$

x_t – стац. с $E(x_t) = 0$

H_0: $r_t = 0$	H_a: $r_t = r_{t-1} + u_t$
	$r_0 = 0$
	u_t – WN, незав от x_t

Техника: Шаг 1. Регрессия на константу

$$\hat c = \bar y \Rightarrow \hat u_t = y_t - \hat c$$

Шаг 2.

$$KPSS = \frac{\sum_{t=1}^{T} S_t^2}{\hat{\lambda}^2 \cdot T^2} \xrightarrow[T \to \infty]{H_0} KPSS_c$$

$$S_t = (\underbrace{y_1 - \hat{c}}_{\hat{u}_1}) + (y_2 - \hat{c}) + \ldots + (\underbrace{y_t - \hat{c}}_{\hat{u}_t})$$

$\hat{\lambda}^2$ — оценка для λ^2

$$\lambda^2 = \gamma_0 + 2\gamma_1 + 2\gamma_2 + \ldots$$

$$\hat{\lambda}^2 = \hat{\gamma}_0 + w_1 2\hat{\gamma}_1 + w_2 2\hat{\gamma}_2 + \ldots$$

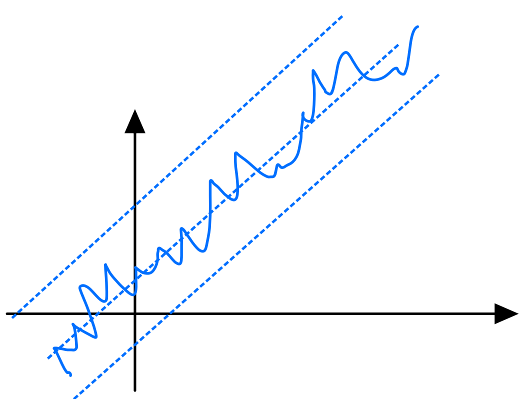
$KPSS_t$ $\quad y_t = c + dt + r_t + x_t$

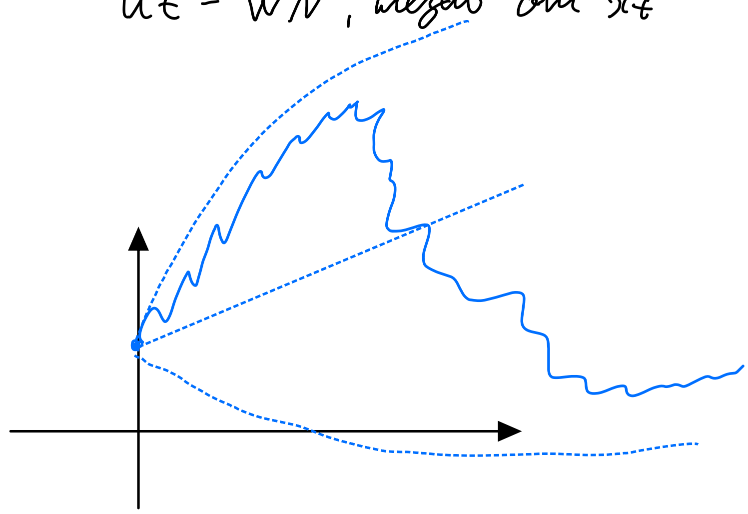
$c, d \in \mathbb{R}$ $\quad x_t$ — стац, $E(x_t) = 0$

$H_0: r_t = 0$ $\qquad H_A: r_t = r_{t-1} + u_t$

$r_0 = 0$

u_t — WN, незав от x_t





Шаг 1. Регрессия $\hat{y}_t = \hat{c} + \hat{d} t$, $\hat{u}_t = y_t - \hat{y}_t$

Шаг 2. $KPSS = \dfrac{\sum_{t=1}^{T} s_t^2}{\hat{\lambda}^2 \cdot T^2} \xrightarrow[T \to \infty]{H_0} KPSS_t$ $\qquad s_t = \hat{u}_1 + \ldots \hat{u}_t$

KPSS для y_t

- H_0 ($y_t \sim$ стац) не отверг.
  $\downarrow$ $d = 0$
  Шаг 2.
  $y_t \sim ARMA(p, q)$

- H_0 отв. в пользу H_A
  H_A (y_t = rw + stat)
  $\downarrow$ $y_t \to \Delta y_t$
  тест KPSS для Δy_t

Идейно: повторное тестирование — это плохо

Для прогноза на 1 шаг вперёд не очень страшно перепутать стац. ряд и нестац.

A. $y_t = y_{t-1} + u_t$

$\hat{y}_{T+1} = y_T$

$\hat{y}_{T+h} = y_T$

B. AR(1)

$y_t = 0{,}99\, y_{t-1} + u_t$

$\hat{y}_{T+1} = 0{,}99\, y_T$

$\hat{y}_{T+h} = (0{,}99)^h\, y_T$