Привет !!

Лекция 1

boris.demeshev@gmail.com
+7 903 - 287 - 34 - 22

Врем. ряд?

Мат. анализ.

посл-ть: $a_1, a_2, a_3, a_4, \ldots$

ряд: $a_1 + a_2 + a_3 + a_4 + \ldots$

Заговор рептилоидов!

Врем ряды - не ряды!

def Послед-ть случ-х величин.

$y_1, y_2, y_3, y_4, y_5, y_6 \ldots$

$y_0, y_1, y_2 \ldots$

$\ldots y_{-3}, y_{-2}, y_{-1}, y_0, y_1, y_2 \ldots$

Задачи?

Один ряд

* прогнозировать
* очистить от шума
  → умение видеть единороги.
* заполнение пропусков.
* характеристики ряда
* обнаружение момента разладки
* разложить на компоненты
  - тренд T_t
  - сезонность S_t
  - остаток R_t

Y_t

...

Много рядов

* схожие задачи
* связаны ли ряды?
* классифицировать ряды
* кластеризовать ряды.
* измерять похожесть рядов
* ...

Осн. фокус курса:
прогнозирование с моделями.

прогнозирование: (ORBIT, PROPHET...)

→ стат. модели для ряда.

→ нет стат. модели (GB, RF..)

Как обойтись без этого курса?

y_t	Ly_t	L^2y_t	Δy_t	$\Delta^2 y_t$	$Ly_t = y_{t-1}$
y_1	-	-	-		$\Delta = 1 - L$
y_2	y_1	-	$y_2 - y_1$		$\Delta y_t = y_t - y_{t-1}$
y_3	y_2	y_1	$y_3 - y_2$		
y_4	y_3	y_2	$y_4 - y_3$		
y_5	y_4	y_3	⋮	⋮	
⋮	⋮	⋮			
y_{2000}	y_{1999}	y_{1998}			

$$\Delta^2 y_t \stackrel{?}{=} \Delta \cdot (\Delta \cdot y_t) = \Delta(y_t - y_{t-1}) = \Delta y_t - \Delta y_{t-1}$$
$$= y_t - 2y_{t-1} + y_{t-2}$$

сезонный лаг / сез. разность

L^{12} $\quad \Delta_{12} = 1 - L^{12}$

$\Delta_{12} y_t = y_t - y_{t-12}$

Идея:

скользящее окно

растущее окно

агрегир. ф-ция

Σ, max, min, med, sd, var, ...

пред. | след.

	пред.	след.
y_1	—	
y_2	—	
y_3	—	
y_4	$\max(y_1, y_2 \dots y_3)$	
y_5	$\max(y_2, y_3, y_4)$	
⋮	⋮	
y_{1000}	$\max(y_{1997}, y_{1998}, y_{1999})$	

Скольз. функц. = 3

max

		след.
y_1	=	
y_2	y_1	
y_3	$\max(y_1, y_2)$	
⋮	⋮	
y_{2000}	$\max(y_{1999}, \dots)$	

расш. окно

max

→ признаки связанные с даты/времени.

→ t, $\sqrt{t}$, $\ln t$

→ $\cos(t)$, $\sin(t)$ $\cos(kt)$ $\sin(kt)$

одуг.

y_t	y_t	t	$\cos t$	$\sin t$	$\cos 2t$	$\sin(2t)$	holid.
y_1		1					0
⋮		2					0
y_t		⋮					0
?	y_0	$T+1$					—

Хочется

→ вероятностные прогнозы

→ предиктивные интервалы.

→ умение видеть единорогов

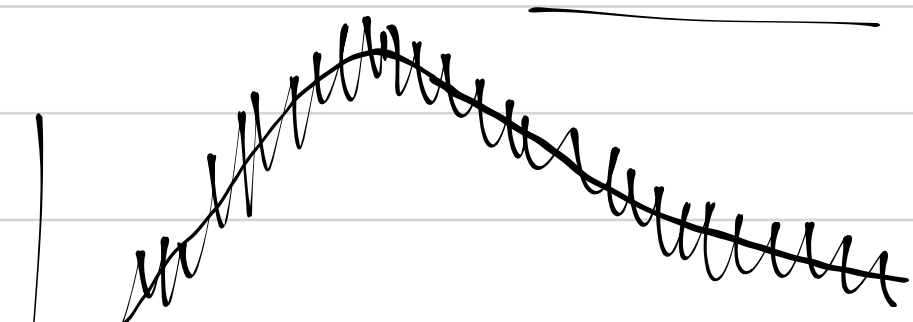
→ можно улучшить прогноз.

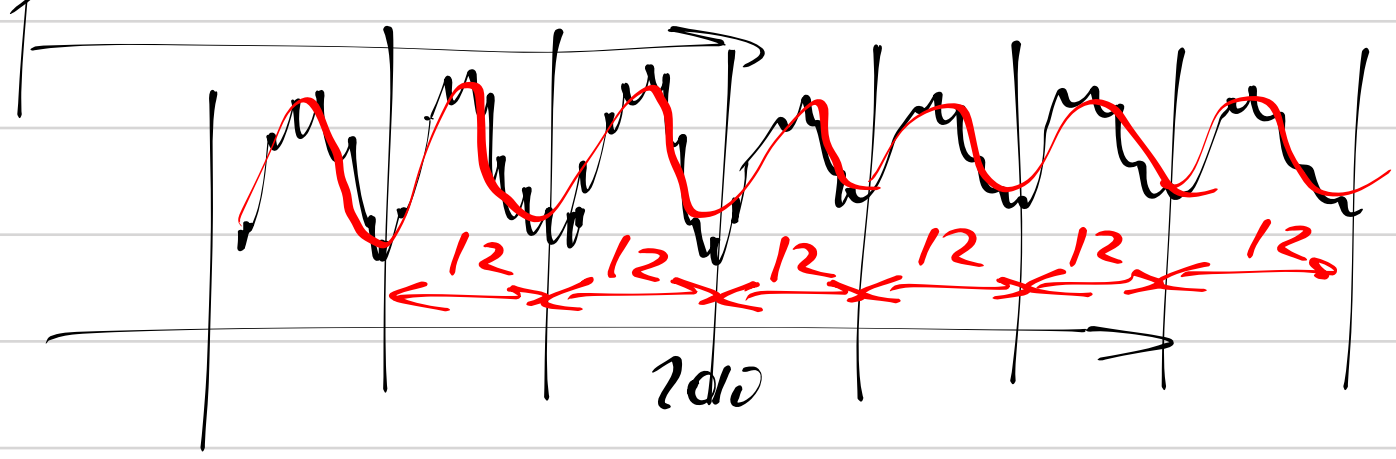
Идея. (неформально)

ряд y_t → тренд $Trend_t$ ← плавно меняется

→ сезонная $Season_t$ ← сост. с сильно-выр-ой период-стью стабильной частоты

→ остаток $Remainder_t$

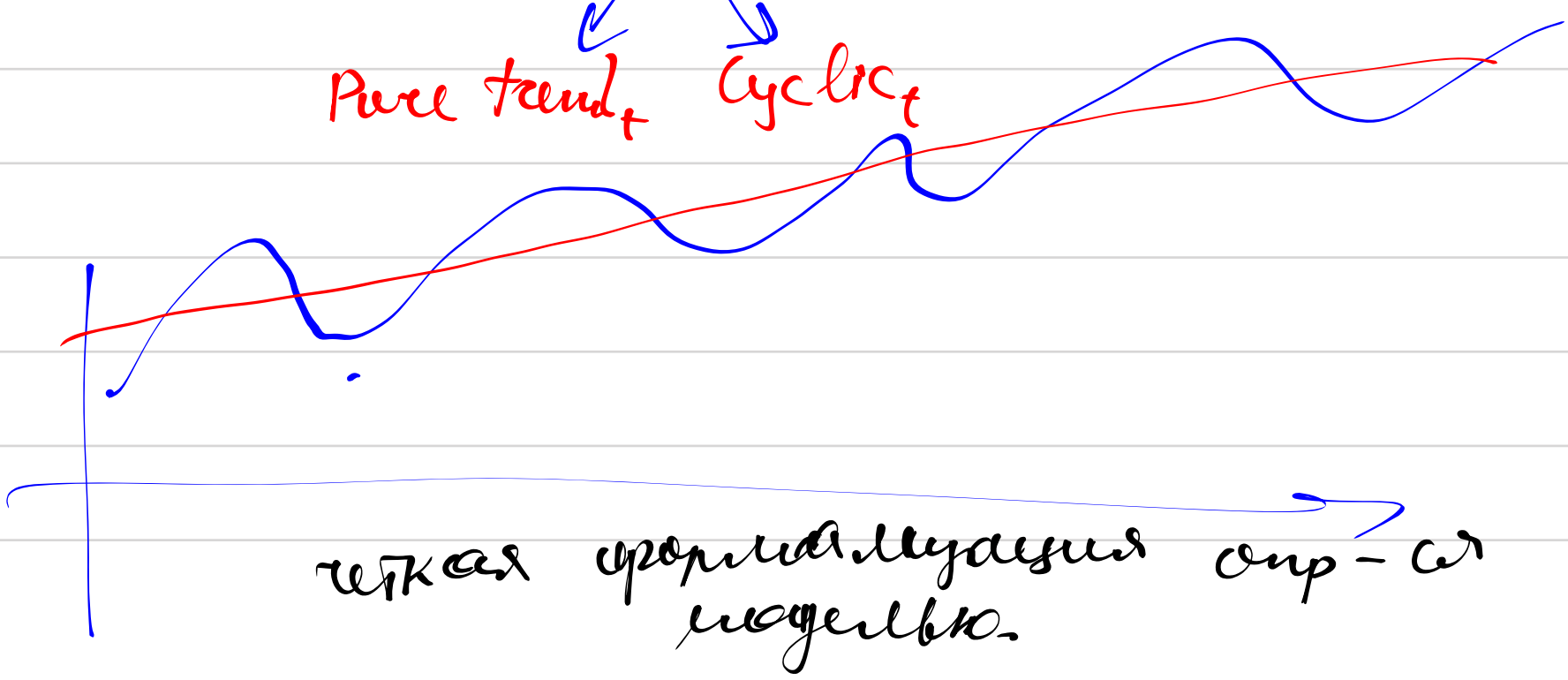




сезонная сост | циклич. сост. частота может плавать

$$y_t = Trend_t + Season_t + Remainder_t$$

Pure $trend_t$ $Cyclic_t$



четкая формализация опр-ся моделью.

Наивная модель.

Случ. блуждание (Random walk).

- → вывод модели
- → оценка модели (Matlib)
- → прогнозирование с уже готовой моделью.

$$\begin{cases} \forall t \\ \forall t \\ \forall t \\ \forall t \end{cases} \begin{cases} y_t = y_{t-1} + (u_t + 2) \\ u_t \sim N(0; 9) \\ F_t = z(y_t, y_{t-1}, y_{t-2}, y_{t-3} \dots) \\ u_t \text{ не зав от } F_{t-1} \end{cases}$$

y_1	5
y_2	9
y_3	12

a) $E(y_4 | F_3)$

$Var(y_4 | F_3)$

95% PI для y_4 (predictive interval)

б) $E(y_5 | \dots)$ .... 95% PI для y_5

$$E(y_4 | F_3) = E(y_3 + u_4 + 2 | F_3) = y_3 + 2 + E(u_4) | F_3) =$$

$$= y_3 + 2 + E(u_4) = y_3 + 2 = 14.$$

$$Var(y_4 | F_3) = Var(y_3 + u_4 + 2 | F_3) = Var(u_4 | F_3) =$$

$$= Var(u_4) = 9.$$

E - expected value {M}
Var - variance {D}

$$(y_4 | F_3) \sim N(14; 9)$$

PI 95% для y_4: $[14 - 1.96 \cdot \sqrt{9}; 14 + 1.96 \cdot \sqrt{9}]$