



## **2.3. Модель авторегрессии – скользящего среднего ARMA(p, q) AutoRegressive Moving Average models**



## Модель авторегрессии – скользящего среднего ARMA(p, q) AutoRegressive Moving Average models

Процесс, описываемый уравнением

$$y_t = \underbrace{\alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p}}_{AR(p)} + \underbrace{\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}}_{MA(q)}$$

называется процессом *авторегрессии – скользящего среднего*,  
 $ARMA(p, q)$ , где  $p$  – порядок авторегрессии,  
 $q$  – порядок скользящего среднего.

$$y_t - \alpha_1 y_{t-1} - \alpha_2 y_{t-2} - \dots - \alpha_p y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

$$\alpha_p(L)y_t = \theta_q(L)\varepsilon_t,$$

$$\alpha_p(L) = (1 - \alpha_1 L - \alpha_2 L^2 - \alpha_3 L^3 - \dots - \alpha_p L^p),$$

$$\theta_q(L) = (1 + \theta_1 L + \theta_2 L^2 + \theta_3 L^3 + \dots + \theta_q L^q),$$

$$ARMA(p, 0) = AR(p)$$

$$ARMA(0, q) = MA(q)$$

## *ARMA(p, q)*

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

- Чтобы процесс был **стационарным**, необходимо и достаточно, чтобы все корни  $z_i$  характеристического уравнения AR(p)-процесса лежали вне единичного круга.

$$1 - \alpha_1 z - \alpha_2 z^2 - \alpha_3 z^3 - \dots - \alpha_p z^p = 0$$

$$\alpha_p(L)y_t = \theta_q(L)\varepsilon_t,$$

$$\exists \alpha_p^{-1}(L)$$

- Если процесс стационарен, то *ARMA(p, q)* можно представить как MA( $\infty$ ).

*ARMA(p, q)*

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

- **Условия обратимости** процесса: определяются условиями обратимости МА-части. Если МА-часть обратима, то весь процесс обратим.

$$\alpha_p(L)y_t = \theta_q(L)\varepsilon_t,$$

$$\exists \theta_q^{-1}(L)$$

- Процесс *ARMA(p, q)* можно представить как  $AR(\infty)$



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

# ARMA(p, q): условие обратимости и стационарности

## Пример

*ARMA(2, 1)*

$$y_t = 2 + 1.3y_{t-1} - 0.4y_{t-2} + \varepsilon_t - 0.5\varepsilon_{t-1}$$

1. Показать, стационарен ли **ARMA** -процесс и представить **ARMA**-процесс в виде **MA**, если это возможно.
2. Показать, обратим ли **ARMA** -процесс и представить **ARMA**-процесс в виде **AR**, если это возможно.



# ARMA(1, 1)

## *ARMA(1,1)*

$$y_t = \alpha y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1} \rightarrow y_t - \alpha y_{t-1} = \varepsilon_t + \theta \varepsilon_{t-1}$$

*В операторной записи:*

$$(1 - \alpha L) y_t = (1 + \theta L) \varepsilon_t$$

**Стационарность:**  $|\alpha| < 1$ ,

**Обратимость:**  $|\theta| < 1$ .

$$E y_t = 0$$

$$V(y_t) = \frac{1 + \theta^2 + 2\alpha\theta}{1 - \alpha^2} \sigma^2$$

$$\rho(1) = \frac{(\alpha + \theta)(1 + \alpha\theta)}{1 + \theta^2 + 2\alpha\theta}$$

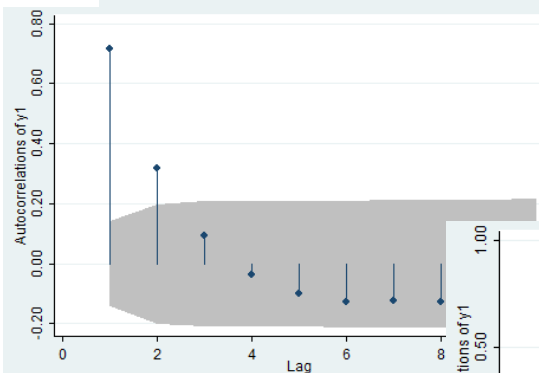
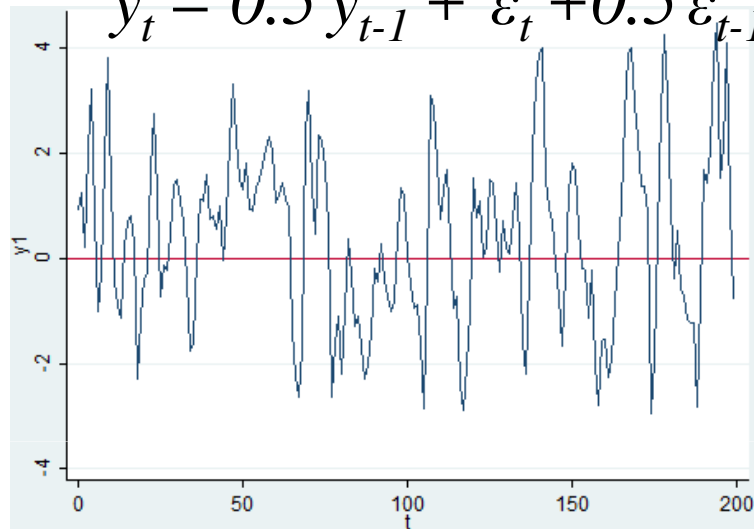
$$\rho(\tau) = \alpha \rho(\tau - 1), \quad \tau > 1$$



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ ЭКОНОМИКИ И СТАТИСТИКИ  
(NRUE)

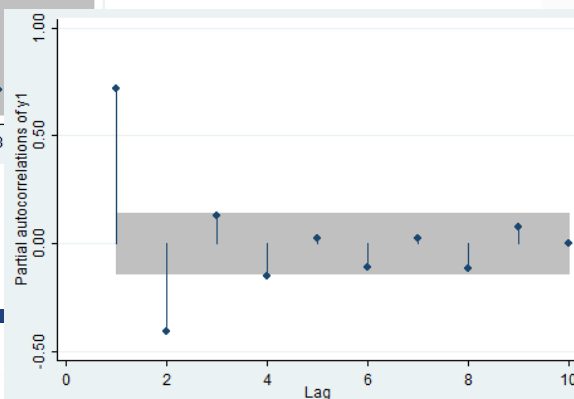
# ARMA(1, 1)

$$y_t = 0.5 y_{t-1} + \varepsilon_t + 0.5 \varepsilon_{t-1}$$

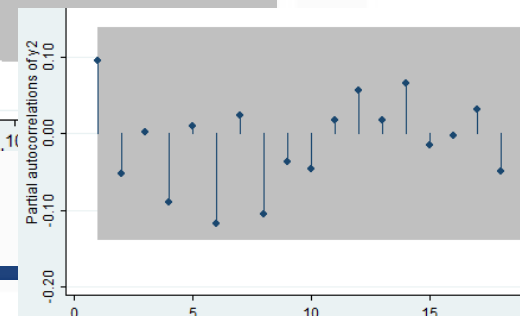
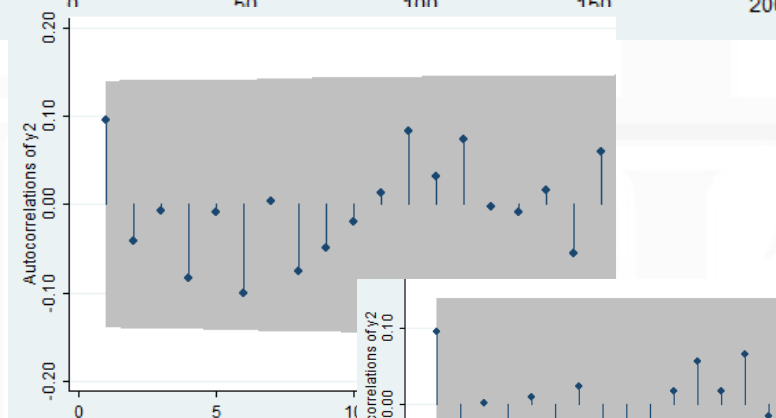
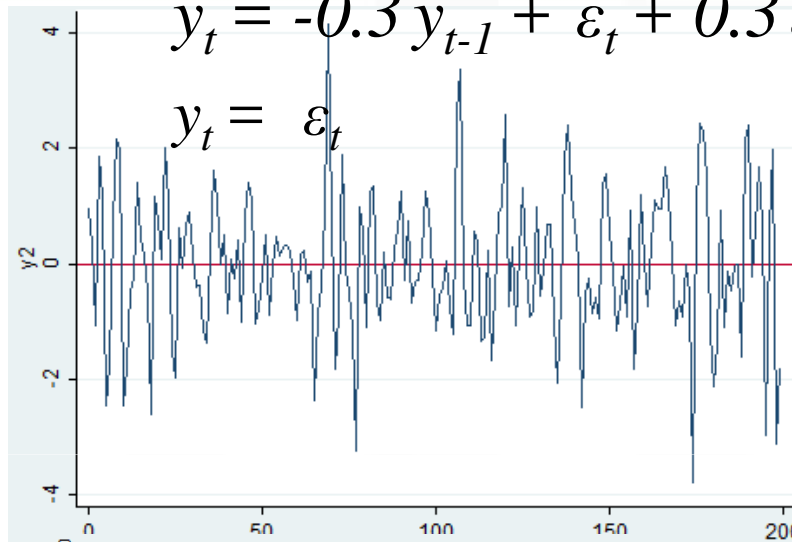


$$\rho(1) = \frac{(\alpha + \theta)(1 + \alpha\theta)}{1 + \theta^2 + 2\alpha\theta}$$

$$\rho(\tau) = \alpha\rho(\tau-1), \quad \tau > 1$$



$$y_t = -0.3 y_{t-1} + \varepsilon_t + 0.3 \varepsilon_{t-1}$$





## Примеры описания поведения ACF

$$y_t = \alpha_1 y_{t-1} + \varepsilon_t + b_1 \varepsilon_{t-1}$$

Модель	ACF	PACF
ARMA(1, 1) $a_1 > 0$	Экспоненциальное убывание с лага 1; знак $\rho(1)$ совпадает со знаком $(a_1 + b_1)$	Осциллирующее убывание с лага 1; $\rho_{part}(1) = \rho(1)$
ARMA(1, 1) $a_1 < 0$	Осциллирующее убывание с лага 1; знак $\rho(1)$ совпадает со знаком $(a_1 + b_1)$	Экспоненциальное убывание с лага 1; $\rho_{part}(1) = \rho(1)$ ; знак $\rho_{part}(k)$ совпадает со знаком $\rho(1)$ , $k > 1$
ARMA( $p, q$ )	Осциллирующее или прямое убывание, начинающееся с лага $q$	Осциллирующее или прямое убывание, начинающееся с лага $p$

**$ACF\ ARMA(p, q) \sim ACF\ AR(p)$**

Носко

**$PACF\ ARMA(p, q) \sim PACF\ MA(q)$**





# ARMA: ACF и PACF

## Примеры описания поведения ACF

**Table 2.1** Properties of the ACF and PACF

Process	ACF	PACF
White noise	All $\rho_s = 0$ ( $s \neq 0$ )	All $\phi_{ss} = 0$
AR(1): $a_1 > 0$	Direct geometric decay: $\rho_s = a_1^s$	$\phi_{11} = \rho_1$ ; $\phi_{ss} = 0$ for $s \geq 2$
AR(1): $a_1 < 0$	Oscillating decay: $\rho_s = a_1^s$	$\phi_{11} = \rho_1$ ; $\phi_{ss} = 0$ for $s \geq 2$
AR( $p$ )	Decays toward zero. Coefficients may oscillate.	Spikes through lag $p$ . All $\phi_{ss} = 0$ for $s > p$ .
MA(1): $\beta > 0$	Positive spike at lag 1. $\rho_s = 0$ for $s \geq 2$	Oscillating decay: $\phi_{11} > 0$ .
MA(1): $\beta < 0$	Negative spike at lag 1. $\rho_s = 0$ for $s \geq 2$	Geometric decay: $\phi_{11} < 0$ .
ARMA(1, 1) $a_1 > 0$	Geometric decay beginning after lag 1. Sign $\rho_1 = \text{sign}(a_1 + \beta)$	Oscillating decay after lag 1. $\phi_{11} = \rho_1$
ARMA(1, 1) $a_1 < 0$	Oscillating decay beginning after lag 1. Sign $\rho_1 = \text{sign}(a_1 + \beta)$	Geometric decay beginning after lag 1. $\phi_{11} = \rho_1$ and $\text{sign}(\phi_{ss}) = \text{sign}(\phi_{11})$ .
ARMA( $p, q$ )	Decay (either direct or oscillatory) beginning after lag $q$ .	Decay (either direct or oscillatory) beginning after lag $p$ .



## 2.4. Прогнозирование процессов $ARMA(p, q)$



# Сравнение моделей прогнозирования

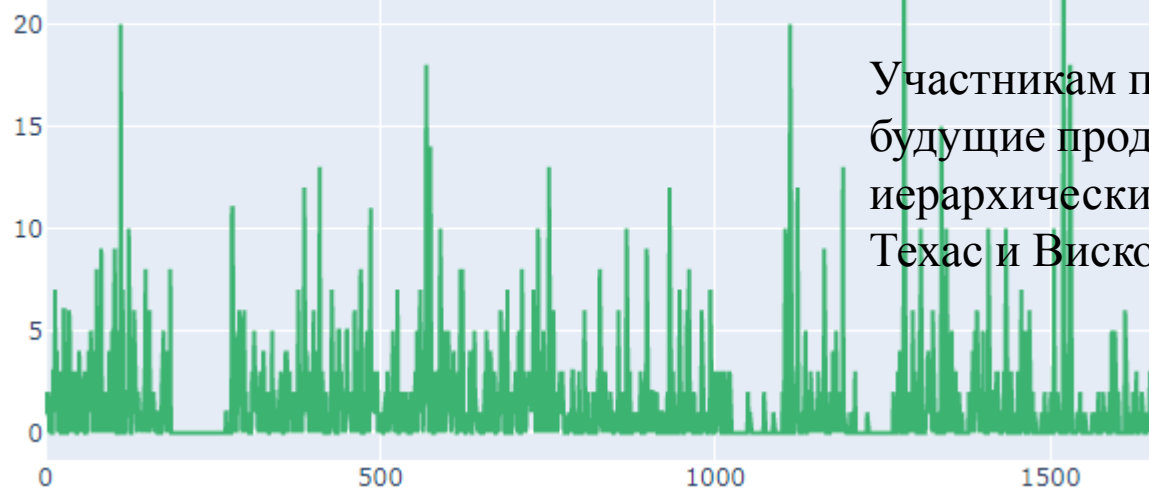
## M5 Competition

### Прогнозы в бизнесе:

<https://www.kaggle.com/tarunpapuraju/m5-competition-eda-models#Modeling->

- Полностью автоматизированные прогнозы сложно настроить
- Аналитики, ответственные за продукт, обладают глубокими знаниями в своей сфере, но не являются специалистами в области прогнозирования ВР
- Спрос на квалифицированное прогнозирование не покрыт

### M5 Competition



Участникам предлагается спрогнозировать будущие продажи в Walmart на основе иерархических продаж в штатах Калифорния, Техас и Висконсин.



# Сравнение моделей прогнозирования

## M5 Competition

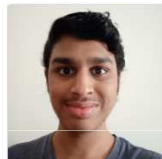
«Наивный метод»  
naïve method

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

Простая СС

$$\hat{y}_{T+h|T} = \frac{1}{T} \sum_{i=1}^T y_i$$

M5 Competition

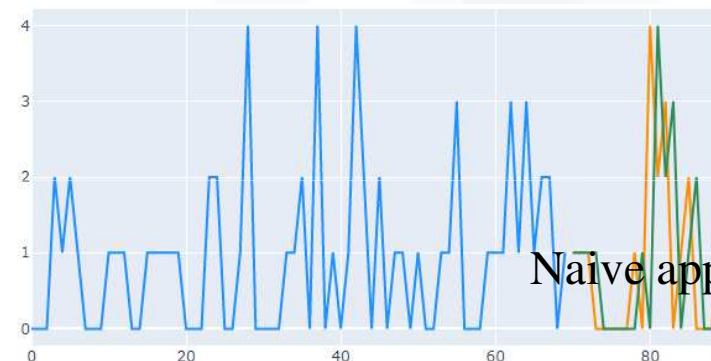


Tarun Paparaju

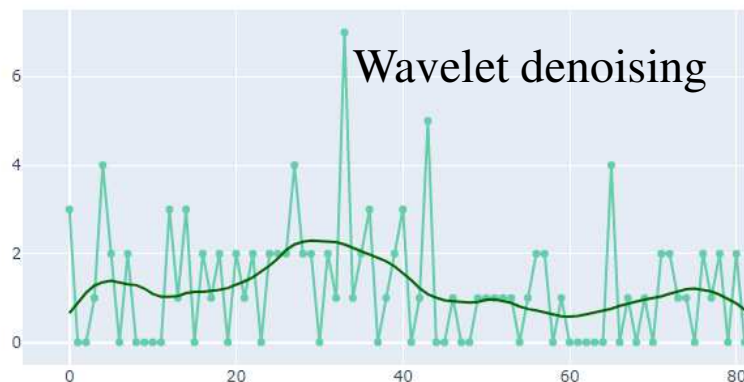
CS Undergrad at UMass  
Bengaluru, Karnataka, India  
Joined 3 years ago · last seen  
https://arirangi

Последнее наблюдение важно в прогнозе

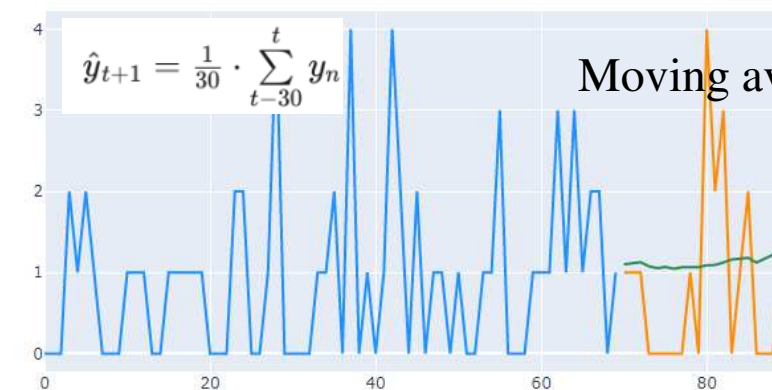
все наблюдения имеют одинаковый вес  
в прогнозе



Naive approach



Wavelet denoising



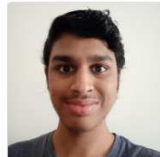
Moving average



# Сравнение моделей прогнозирования

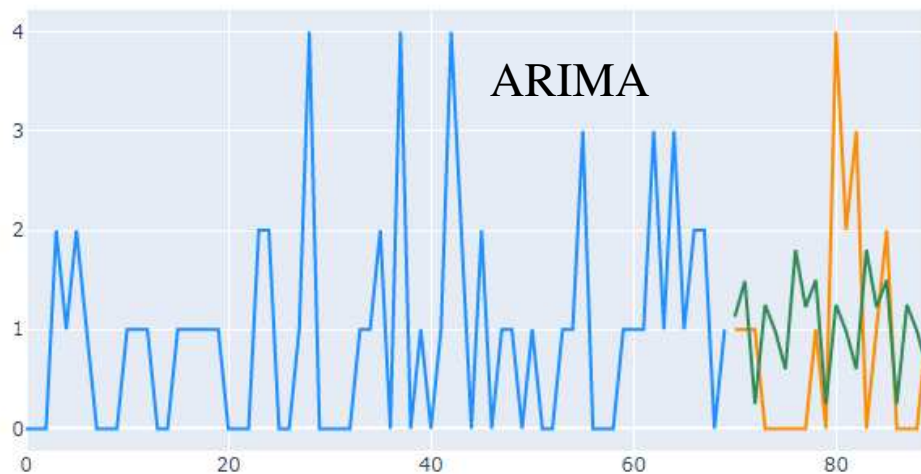
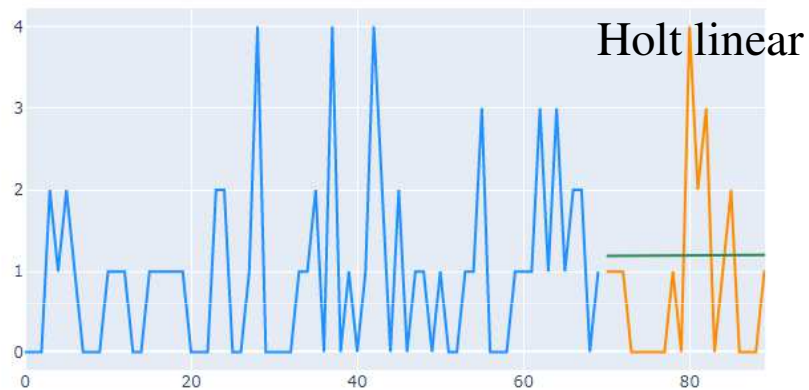
## M5 Competition

Walmart



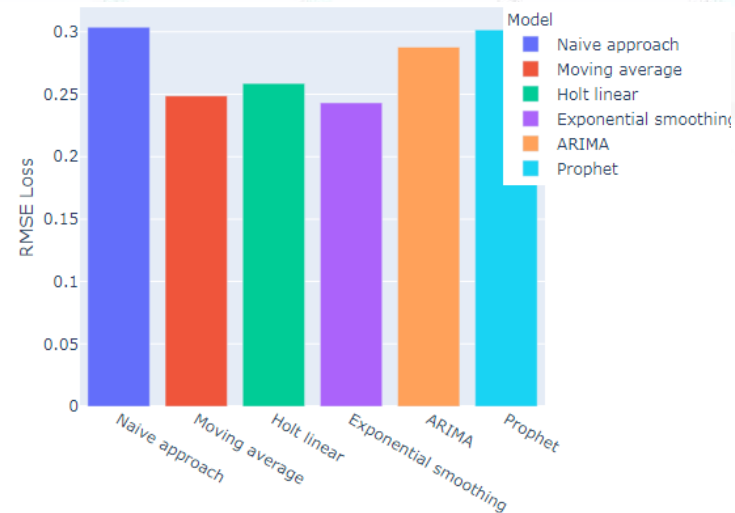
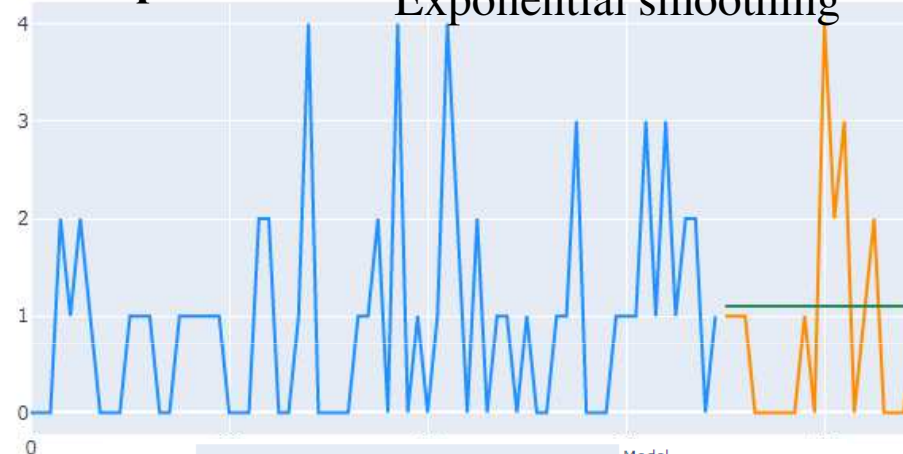
Tarun Paparaju

CS Undergrad at UMass  
Bengaluru, Karnataka, India  
Joined 3 years ago · last seen  
https://srirangi



Адаптивные модели будут рассмотрены позже

Exponential smoothing





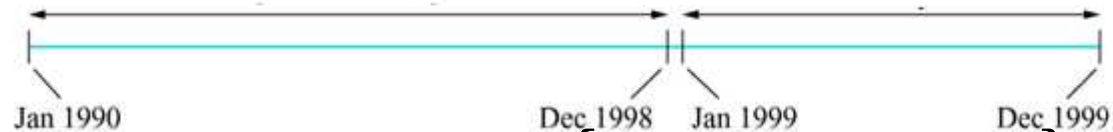
## Прогнозирование процессов ARMA: теоретические предпосылки

Существует два источника ошибок:

- Неопределенность будущих значений случайной величины  $\varepsilon_t$
- Отсутствие точных значений коэффициентов модели
- Значение переменной прогнозируется для некоторого будущего момента времени, при этом лаговые значения переменной можно рассматривать *фиксированными* или *случайными*.
- Первая возможность приводит к условному *прогнозу* (как в модели множественной регрессии), а вторая к *безусловному*.

Из ТВиМС: условная дисперсия СВ не превышает ее безусловную дисперсию  $\rightarrow$  *точность* условного прогноза *выше*.





Прогноз  $y_{T+h}$

$$I_T = \{y_{-\infty}, y_1, \dots, y_{T-1}, y_T\}$$

$I_T$  – информационное множество (наблюдаемые значения)

$h$  – горизонт прогнозирования.

Критерий выбора прогнозирующей функции:

$$E((y_{T+h} - \hat{y}_{T+h})^2 | I_T) \rightarrow \min$$

Наилучший прогноз – при минимальной среднеквадратической ошибке:

$$E\{y_{T+h} | y_1, \dots, y_T\}$$

Точность прогнозирования:

- ошибка прогноза
- дисперсия ошибки

$$MA(1) \quad y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

Пусть коэффициенты модели известны и имеются значения  $y_t$ ,  $t=1, \dots, T$

- **Безусловный** прогноз— это математическое ожидание процесса,

$$\text{т.е. } \hat{y}_{T+1} = E\{\varepsilon_{T+1} + \theta_1 \varepsilon_T\} = 0$$

- **Условный прогноз:**  $\hat{y}_{T+1|I_T} = E\{\varepsilon_{T+1} + \theta_1 \varepsilon_T \mid y_1 \dots y_T\} = \theta_1 \varepsilon_T$

- Ошибка прогноза  $e_{T+1} = y_{T+1} - \hat{y}_{T+1|I_T}$

- Условная дисперсия ошибки прогноза  $\hat{\sigma}_h^2$

- 95% Доверительный интервал прогноза (предположение, что остатки подчиняются **нормальному** распределению)

$$\hat{y}_{T+h|I_T} \pm 1.96 \hat{\sigma}_h$$





НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ  
ТРАНСПОРТА

## Нормальный закон

$$\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz$$

Целые и де-сятичные доли t	Сотые доли t							
	0	1	2	3	4	5	6	7
0,0	0,0000	0,0080	0,0160	0,0239	0,0319	0,0399	0,0478	0,0558
0,1	0797	0876	0955	1034	1113	1192	1271	1350
0,2	1585	1663	1741	1819	1897	1974	2051	2128
0,3	2358	2434	2510	2586	2661	2737	2812	2886
0,4	3108	3182	3255	3328	3401	3473	3545	3616
0,5	3829	3899	3969	4039	4108	4177	4245	4313
0,6	4515	4581	4647	4713	4778	4843	4907	4971
0,7	5161	5223	5285	5346	5407	5467	5527	5587
0,8	5763	5821	5878	5935	5991	6047	6102	6157
0,9	6319	6372	6424	6476	6528	6579	6629	6679
1,0	0,6827	0,6875	0,6923	0,6970	0,7017	0,7063	0,7109	0,7154
1,1	7287	7330	7373	7415	7457	7499	7540	7580
1,2	7699	7737	7775	7813	7850	7887	7923	7959
1,3	8064	8098	8132	8165	8198	8230	8262	8293
1,4	8385	8415	8444	8473	8501	8529	8557	8584
1,5	8664	8690	8715	8740	8764	8789	8812	8836
1,6	8904	8926	8948	8969	8990	9011	9031	9051
1,7	9109	9127	9146	9164	9181	9199	9216	9233
1,8	9281	9297	9312	9327	9342	9357	9371	9385
1,9	9426	9439	9451	9464	9476	9488	9500	9512



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

## Построение доверительного интервала для ARMA

$$\hat{y}_{T+h|I_T} \pm c \hat{\sigma}_h$$

$$80\% : \hat{y}_{T+h|I_T} \pm 1.28 \hat{\sigma}_h$$

$$95\% : \hat{y}_{T+h|I_T} \pm 1.96 \hat{\sigma}_h$$

$$99\% : \hat{y}_{T+h|I_T} \pm 2.58 \hat{\sigma}_h$$

%	Значение c
50	0.67
55	0.76
60	0.84
65	0.93
70	1.04
75	1.15
80	1.28
85	1.44
90	1.64
95	1.96
96	2.05
97	2.17
98	2.33
99	2.58



## Прогнозирование процессов ARMA: AR(1)

$$AR(1) \quad y_t = \alpha y_{t-1} + \varepsilon_t$$

- Прогноз на 1 шаг вперед:  $\hat{y}_{T+1} = E\{y_{T+1} \mid y_1 \dots y_T\} = \alpha y_T$

- Прогноз на 2 шага вперед:  $\hat{y}_{T+2} = E\{y_{T+2} \mid y_1 \dots y_T\} = \alpha^2 y_T$

$$\hat{y}_{T+h} = E\{y_{T+h} \mid y_1 \dots y_T\} = \alpha^h y_T$$



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

## Прогнозирование процессов ARMA(1,1)

$$ARMA(1,1) \quad y_t = \alpha_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

Для построения прогноза необходимо объединить прогноз по  
МА и AR.



## Пример

дан случайный процесс  $y_t = 0.5 y_{t-1} + \varepsilon_t$   
 $y_{100} = -1.6$

Выведите формулы прогнозов на 1, 2, 3 шага вперед и вычислите соответствующие прогнозные значения.



## Построение доверительного интервала для ARMA: альтернативный подход (бутстрапирование)

В случае, когда остатки модели **не** подчиняются нормальному распределению, одной из альтернатив является использование метода **бутстрапирования**, который предполагает, что ошибки прогноза некоррелированы.

Ошибка прогноза  $e_t = y_t - \hat{y}_{t|t-1} \rightarrow y_t = \hat{y}_{t|t-1} + e_t$

«Симулируем» следующее наблюдение:  $y_{T+1} = \hat{y}_{T+1|T} + e_{T+1}$   
прогноз                      не известно

Предполагая, что будущие ошибки будут похожи на прошлые ошибки, заменяем  $e_{T+1}$  путем выборки из набора ошибок, которые наблюдались в прошлом (остатки).

Далее:  $y_{T+2} = \hat{y}_{T+2|T+1} + e_{T+2}$



## Построение доверительного интервала для ARMA: альтернативный подход (бутстрапирование)

- Повторяя процедуру бутстрапирования неоднократно, получаем множество возможных значений ошибок.
- рассчитываются процентиля для каждого горизонта прогнозирования.
- Строится бутстрапированный доверительный интервал (**bootstrapped prediction interval**).
- Бутстрапами («bootstrap») называют ремешки на обуви, за которые вытащил сам себя барон Мюнхаузен из болота.
- Бутстрапирование – приближение истинного распределения эмпирическим на основе наблюдаемых данных.



Анатолийев С. (2007) Основы бутстрапирования <http://www.quantile.ru/03/03-SA.pdf>

## Интерпретация параметров ARMA моделей

**Функция импульсного отклика** (impulse responses - IRF) – последовательность  $\psi_{ij}$  в разложении  $MA(\infty)$  стационарного случайного процесса:

$$ARMA \rightarrow MA(\infty): \quad y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots + \psi_K \varepsilon_{t-K} + \dots$$

IRF показывает среднее изменение  $y_{t+\tau}$ , вызванное единичным шоком (единичной инновацией)  $\varepsilon_t$ .

Суммарный эффект от инноваций:

$$\sum_{i=0}^{\infty} IRF(i) = \sum_{i=0}^{\infty} \psi_i, \quad \psi_0 = IRF(0) = 1$$

