

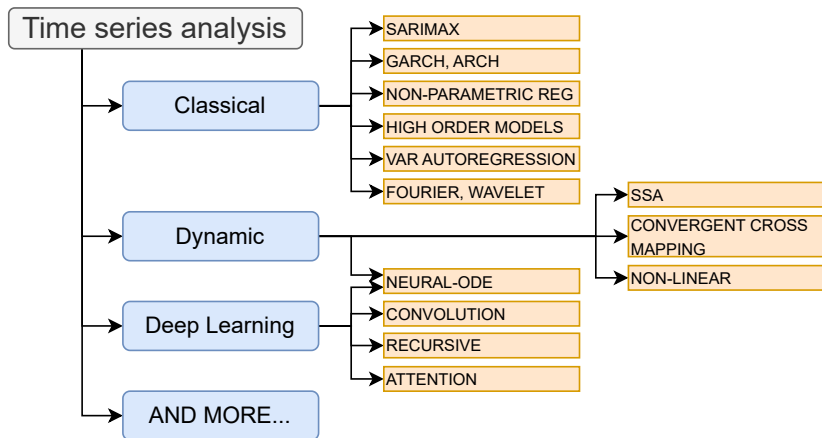
# Mathematical Forecasting Methods

## Лекция 1

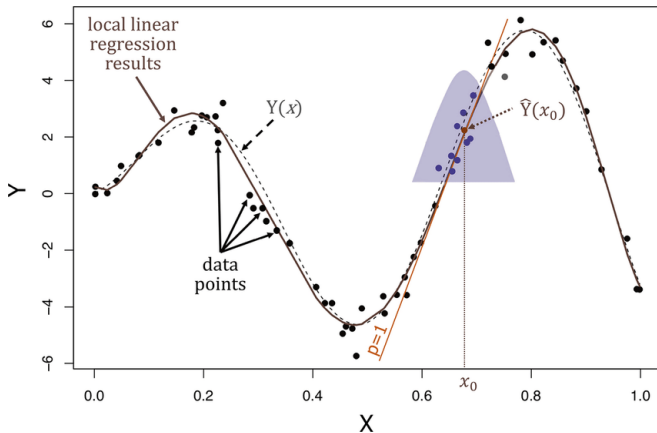
МФТИ

Осень, 2023

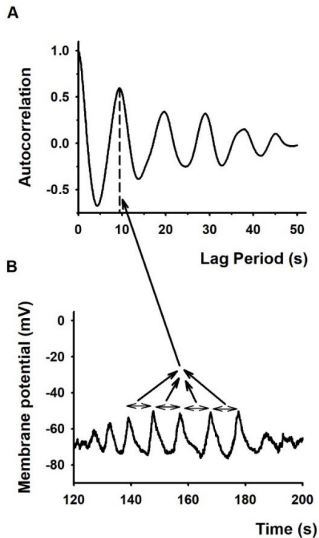
# Forecasting Methods/Models zoo



# Регрессионные модели: непараметрические методы



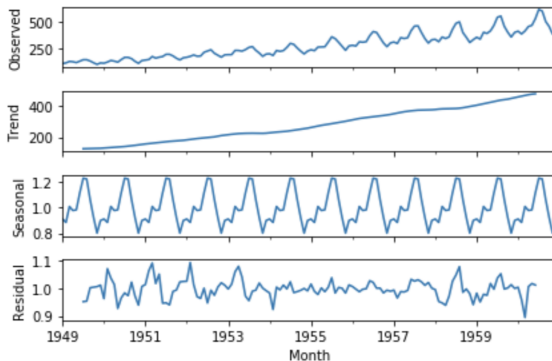
# Регрессионные модели: автокорреляция



# Регрессионные модели: SARIMAX

$$y_t^{(d)} = c + \varepsilon_t + \underbrace{\phi_1 y_{t-1}^{(d)} + \phi_2 y_{t-2}^{(d)} + \dots + \phi_p y_{t-p}^{(d)}}_{\text{Auto-Regressive}} + \underbrace{\theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}}_{\text{Moving Average}}$$

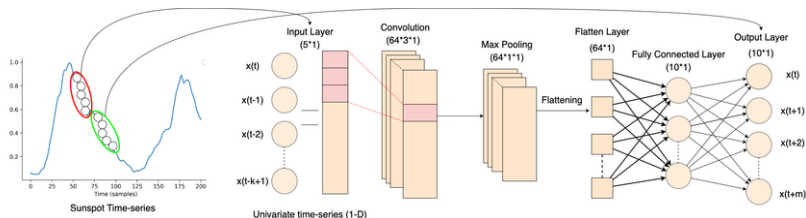
Integrated



credit:

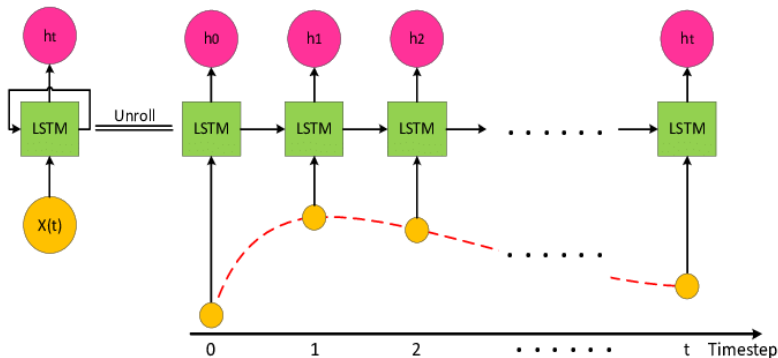
<https://www.geeksforgeeks.org/python-arima-model-for-time-series-forecasting/>

# Регрессионные модели: сверточные нейронные сети

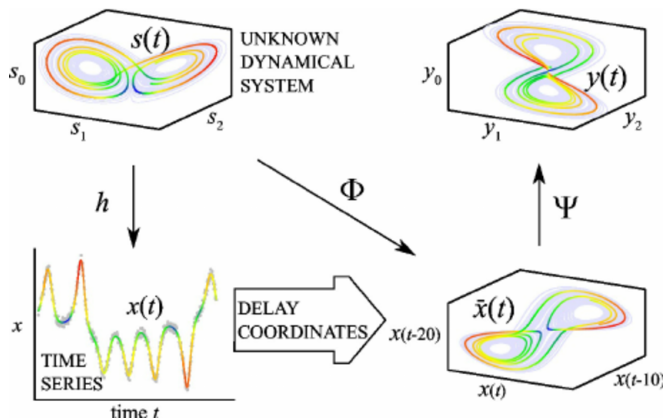


credit: <https://www.researchgate.net/figure/One-dimensional-Convolutional-Neural-Network-for-time-series>

# Регрессионные модели: рекуррентные нейронные сети

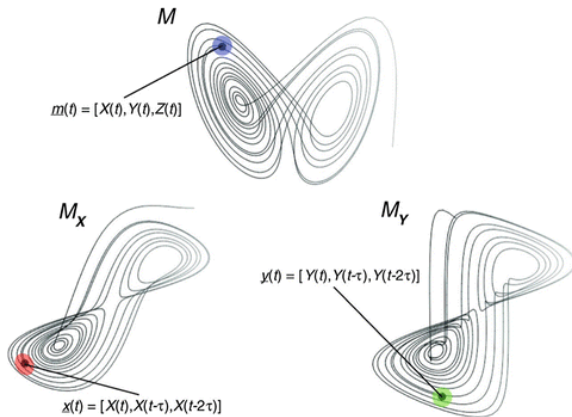


# От статистических к динамическим





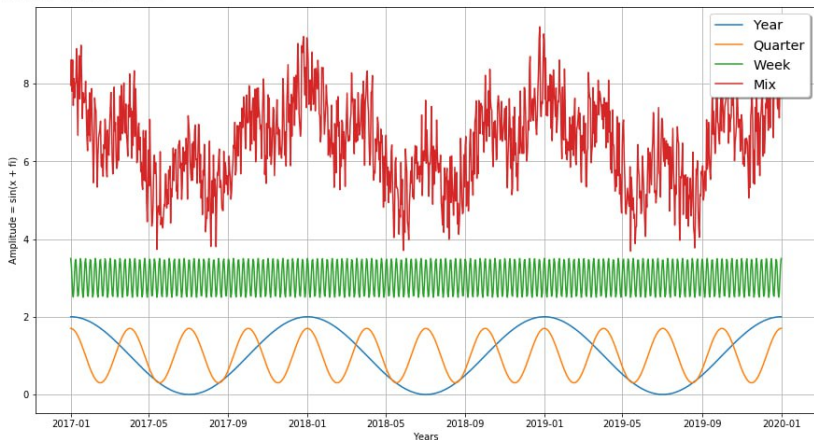
# От статистических к динамическим: перекрёстное отображение



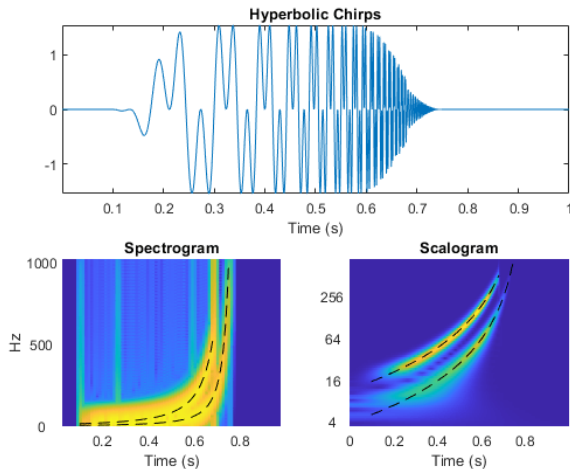
# От статистических к динамическим: анализ Фурье

Частота квартальная: 12

Частота недельная: 144

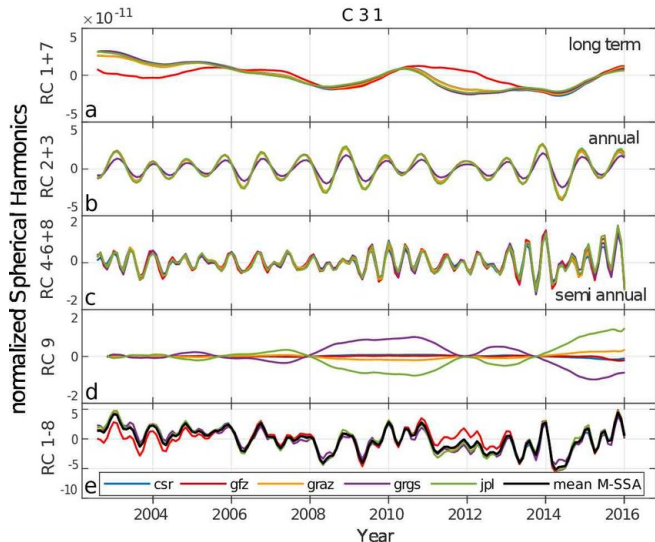


# От статистических к динамическим: вейвлет преобразование



credit: <https://www.mathworks.com/help/wavelet/gs/choose-a-wavelet.html>

# От статистических к динамическим: сингулярный спектральный анализ



# Тензорные методы

$$W \in \mathbb{R}^{M \times N} = \mathcal{G}_1 \in \mathbb{R}^{m_1 n_1 \times r_1} \times_1 \mathcal{G}_2 \in \mathbb{R}^{r_1 \times m_2 n_2 \times r_2} \times_1 \dots \times_1 \mathcal{G}_{d-1} \in \mathbb{R}^{r_{d-2} \times m_{d-1} n_{d-1} \times r_{d-1}} \times_1 \mathcal{G}_d \in \mathbb{R}^{r_{d-1} \times m_d n_d}$$

(a) Tensorizing for weight matrix in TT format

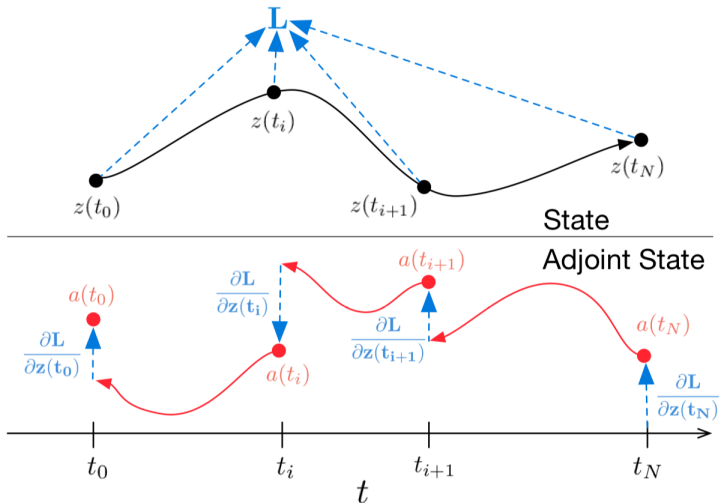
$$K \in \mathbb{R}^{I \times I \times C \times S} = \mathcal{G}_0 \in \mathbb{R}^{I^2 \times r_1} \times_1 \mathcal{G}_1 \in \mathbb{R}^{r_1 \times C_1 S_1 \times r_2} \times_1 \dots \times_1 \mathcal{G}_{d-1} \in \mathbb{R}^{r_{d-1} \times C_{d-1} S_{d-1} \times r_d} \times_1 \mathcal{G}_d \in \mathbb{R}^{r_d \times C_d S_d}$$

(b) Tensorizing for convolutional kernel in TT format

$$K_{3D} \in \mathbb{R}^{t \times h \times w \times C \times S} = \mathcal{G}_0 \in \mathbb{R}^{t \times h \times w \times r_1} \times_1 \mathcal{G}_1 \in \mathbb{R}^{r_1 \times C_1 S_1 \times r_2} \times_1 \dots \times_1 \mathcal{G}_{d-1} \in \mathbb{R}^{r_{d-1} \times C_{d-1} S_{d-1} \times r_d} \times_1 \mathcal{G}_d \in \mathbb{R}^{r_d \times C_d S_d}$$

(c) Tensorizing for 3D convolutional kernel in TT format

# Дифференциальные модели: Neural-ODE



# Временной ряд

- ▶ Временной ряд — это последовательность значений некоторой переменной (или переменных), регистрируемых непрерывно или через некоторые промежутки времени.
- ▶ Скалярным временным рядом  $\{x_i\}_{i=1}^N$  называется массив из  $N$  чисел, представляющих собой значения некоторой измеренной (наблюдаемой) динамической переменной  $x(t)$  с некоторым постоянным шагом  $\tau$  по времени.
- ▶ Основная задача — это задача прогноза (предсказать будущие значения измеряемых характеристик изучаемого объекта на некоторый отрезок времени вперед).

## Временной ряд

Будем называть  $\{x_i\}_{i=1}^N$  временным рядом,  
 $\hat{x}_{i+d} = f_{i,d}(x_1, \dots, x_N; w)$  — моделью временного ряда, где  
 $d = 1, \dots, D$  — горизонт прогнозирования.

Метод наименьших квадратов:

$$Q_t(w) = \sum_{i=1}^N (\hat{x}_i(w) - x_i)^2 \rightarrow \min_w$$

Проблемы:

- ▶ у временных рядов может быть сложная структура,
- ▶ неквадратичная функция потерь подходит не для всех задач,
- ▶ может быть потребность в учете физической природы временного ряда.



- ▶ Этот курс будет посвящен математическим методам вне классических подходов (но и классические тоже будут затронуты).
- ▶ Существует множество различных статистических и динамических методов прогнозирования временных рядов.
- ▶ Существуют методы и модели прогнозирования данных высокой размерности (таких как видео).
- ▶ Принципиальное отличие от других задач машинного обучения — это зависимость наблюдений от предыстории.
- ▶ Классическая постановка сильно ограничена.