

# Mathematical Forecasting Methods

## Лекция 4

МФТИ

Осень, 2023

## Краткое повторение

Модель **SARIMAX**( $p, q, d; P, Q, D$ ) приводит временной ряд к стационарному виду и прогнозирует с наперед заданной точностью (по Теореме Вольда).

$$\begin{aligned}x_t = & \mu + \underbrace{\phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots}_{p \text{ слагаемых AR}} + \underbrace{u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots}_{q \text{ слагаемых MA}} \\& + \underbrace{\alpha_1 x_{t-S} + \alpha_2 x_{t-2S} + \dots}_{P \text{ сезонных слагаемых AR}} + \underbrace{\eta_1 u_{t-S} + \eta_2 u_{t-2S} + \dots}_{Q \text{ сезонных слагаемых MA}} \\& + \underbrace{\beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \dots}_{\text{внешние переменные, не зависящие от } x_t},\end{aligned}$$

где  $\phi, \theta, \alpha, \eta, \beta$  — настраиваемые параметры модели,  
 $x_t$  — значения временного ряда (продифференцированные  $d$  и  $D$  сезонных раз),  
 $u_t$  — значения шумов,  
 $z_i$  — внешние (экзогенные) переменные.

## Краткое повторение

- ▶ Если ранее полагалось, что  $u_t \sim WN(0, 1)$ , то теперь возмущения будут моделироваться с меняющейся во времени дисперсией  $\sigma_t^2$ :

$$u_t = z_t \sqrt{\sigma_t^2},$$

где  $z_t \sim WN(0, 1)$ .

- ▶ Моделирование дисперсии как процесс  $AR(q)$ , используя квадраты оцененных остатков.
- ▶ **ARCH**( $p$ ):  $\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_p u_{t-p}^2$ .
- ▶ Обобщение модели с предположением, что дисперсия зависит от своих прошлых значений.
- ▶ **GARCH**( $p, q$ ):

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i u_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2$$

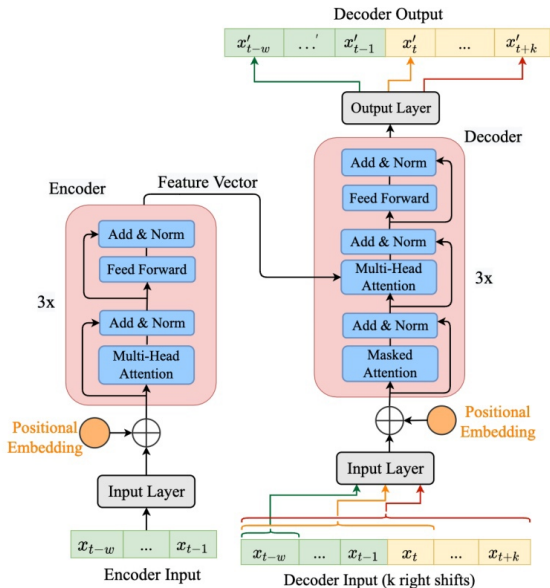
Здесь  $\omega > 0$ ,  $\alpha_i \geq 0 \ \forall i = \overline{1, p}$ ,  $\beta_j \geq 0 \ \forall j = \overline{1, q}$ .

# Введение в нейронные сети для временных рядов

Задача прогнозирования временного ряда

$$x_t \approx f_{t,N}(x_{t-1}, \dots, x_{t-N}; w) =: \hat{x}_t$$

# Механизм Attention и модель Transformer



# Резюме

- ▶ Линейные модели сильно органичены.
- ▶ Нейронные сети моделируют нелинейные зависимости.
- ▶ Модели сверточных нейронных сетей CNN применяются к предыстории и позволяют извлекать сложные закономерности.
- ▶ Рекуррентные нейронные сети RNN хранят внутреннее состояние и эффективнее работают с последовательностями.
- ▶ Модификации классической RNN (например, LSTM и GRU) лучше справляются с затуханием и взрывом градиента.