

# GANs for Trading Strategies Backtesting

Николай Аверьянов, Ярослав Лисняк

ШАД

25 марта 2022 г.

## Reminder: постановка задачи

$S_t = (S_t^1, \dots, S_t^n)$  – векторозначный случайный процесс.

Координаты могут быть созависимы произвольно сложным образом.

В реальной жизни имеем лишь дискретное множество реализаций одного из вероятностных исходов.

Хотим научиться генерировать систему рядов  $S_{\text{gen}}^{(1)}, \dots, S_{\text{gen}}^{(n)}$ , отражающих как можно больше ценных характеристик исходного случайного процесса.

**Мотивация задачи:** получить новые методы для бэктеста торговых стратегий.

# Решение: что мы пробовали

GAN и WGAN.

Разные архитектуры генератора и дискриминатора. Среди них: обычный мlp, свертки (бейзлайны авторов), трансформеры.

Написали кастомный лосс, учитывающий мета-признаки, добавили его к лоссу генератора.

Gradient penalty и Spectral normalization для регуляризации градиентов.

# Подробнее о мета-признаках

1. Моменты ряда:

$$M_k = \int_{t \in T} \|\mathbb{E} S_{\text{real}}^k(t) - \mathbb{E} S_{\text{fake}}^k(t)\| dt$$

2. Стационарные ковариационные функции:

$$C = \int_{t \in T} \|\Sigma(S_{\text{real}}(t)) - \Sigma(S_{\text{fake}}(t))\| dt$$

3. Авторегрессионные функции.

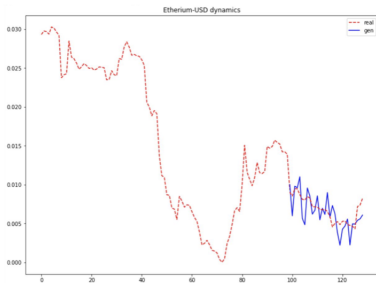
$$A(\tau) = \int_{t \in T} \|\mathbb{E} S_{\text{real}}(t) S_{\text{real}}(t - \tau) - \mathbb{E} S_{\text{fake}}(t) S_{\text{fake}}(t - \tau)\| dt$$

4. Частота колебаний.

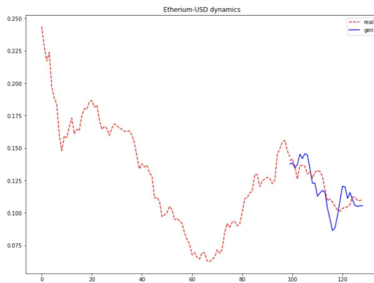
$$Osc(\delta) = \int_{t \in T} \left( S(t + \delta) - (S(t) + S'(t)) \right)^2 dt$$

# Результаты: решаем проблему гладкости

Ни одна из опробованных архитектур дискриминатора не наказывает за повышенную частоту колебаний.



Сверточная сеть без мета-признаков.

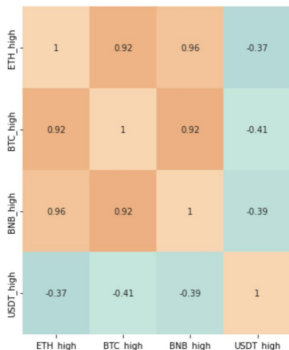


Та же архитектура, обученная с мета-признаками.

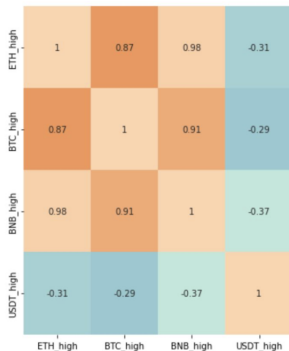
Но с этим справляется придуманный мета-признак!

# Результаты: созависимость рядов

На отложенных данных модель косвенно хорошо воспроизводит корреляционные зависимости между рядами.

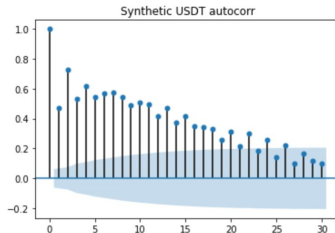
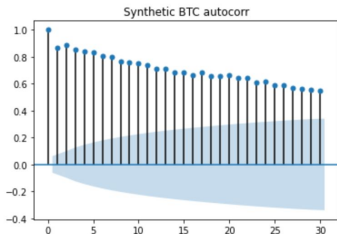
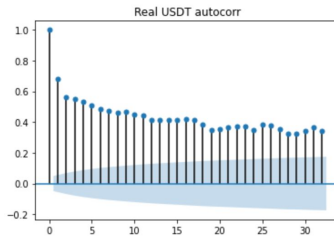
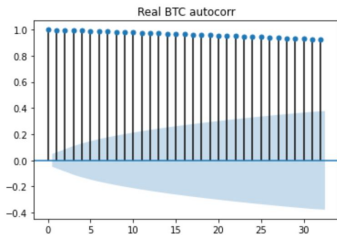


Корреляционная матрица исторических данных.



Корреляционная матрица сгенерированных данных.

# Результаты: автокорреляции



# Результаты: примеры сэмплов

