

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА
Факультет вычислительной математики и кибернетики

**Исследование торговых стратегий на основе
синтетических облигаций**

Выполнил:
студент ВМК МГУ 211 группы
Акрамов Роман Рустамович

Научный руководитель:
Гуров С.И.

Москва, 2025 г.

Содержание

1	Введение	2
2	Анализ литературы	3
3	Методология	4
3.1	Данные и признаки	4
3.2	Целевая переменная	5
3.3	Модель и обучение	5
4	Эксперимент	6
4.1	Обучение модели и метрики	6
4.2	Интерпретация коэффициентов модели	6
4.3	Симуляция торговли	7
5	Графики и визуализация	8
5.1	Временные ряды Bond	8
5.2	Оценка порогового параметра θ	8
6	Выводы	9

1. Введение

Финансовые рынки характеризуются сложностью, нелинейностью и стохастичностью. Для краткосрочной торговли критичны микроуровневые рыночные сигналы, отражённые в заявках (order book). В работе рассматривается построение синтетической облигации через соотношение фьючерсной и спот-цен, что позволяет получать арбитражные сигналы.

Пусть:

$$P^{fwd} = P^{spot} \left(1 + r \frac{t}{365} \right),$$

где P^{fwd} — фьючерсная цена, P^{spot} — спот-цена, r — годовая процентная ставка, t — число дней до экспирации.

С учетом микроизменений в стакане предлагается прогнозировать доходность синтетической облигации, определяемой как отклонение цены от тренда.

2. Анализ литературы

Классическая теория оценки производных инструментов (см. [1]) использует формулу:

$$P^{fwd} = P^{spot} \left(1 + r \frac{t}{365} \right),$$

где P^{fwd} — фьючерсная цена, P^{spot} — спот-цена, r — годовая процентная ставка, t — время до экспирации. Однако при анализе высокочастотных данных (order book) возникают краткосрочные статистические закономерности, для моделирования которых требуются как математические подходы, так и современные вычислительные технологии.

Помимо методов традиционной теории, в исследовании используются инструменты машинного обучения для генерации торговых сигналов. Ключевой подход основан на прогнозировании будущих значений синтетической облигации с использованием линейной регрессии.

Технологический стек

Для реализации эксперимента выбран стек технологий на базе языка Python, что позволяет интегрировать обработку данных, обучение модели и визуализацию результатов в единой экосистеме. Основные компоненты стека:

- **Python.** Язык программирования для обработки данных, реализации алгоритмов машинного обучения и симуляции торговых стратегий.
- **scikit-learn.** Библиотека для машинного обучения, использованная для реализации простых моделей (например, `LinearRegression`) и расчёта регрессионных метрик (MAE, MSE, R^2).
- **joblib.** Инструмент для сериализации и сохранения обученных моделей, что обеспечивает повторяемость экспериментов.
- **Pandas и NumPy.** Библиотеки для структурирования данных, выполнения матричных вычислений и работы со временными рядами.
- **Matplotlib.** Основной инструмент для создания графиков и 3D-визуализации зависимости ключевых параметров (например, итогового баланса от параметров `future_window` и `rolling_window`).

3. Методология

В данном разделе описаны используемые данные, признаки и постановка целевой переменной, что позволяет построить модель прогнозирования синтетической облигации.

3.1. Данные и признаки

Анализируются высокочастотные данные стаканов с интервалом $\Delta t \approx 5$ сек. Каждая запись включает:

- **Временные метки:** `time`, `time_dt`, `server_timefutures`, `server_timestocks`.
- **Идентификаторы инструментов:** тикеры `futures` и `stocks`.
- **Структуры стаканов:** данные по уровням `OFFER` и `BID` для фьючерсов и акций (до 3 уровней).
- **Качественные характеристики:** извлечённые значения цены и объёма для уровней $i = 0, 1, 2$, обозначаемые как $price_i$ и $quantity_i$ соответственно.

Синтетическая облигация определяется по следующим формулам:

$$\text{BondBUY} = \left(\frac{\text{Bid}_{\text{futures}}}{\text{Offer}_{\text{stocks}}} - 1 \right) \cdot 100,$$

$$\text{BondSELL} = \left(\frac{\text{Offer}_{\text{futures}}}{\text{Bid}_{\text{stocks}}} - 1 \right) \cdot 100.$$

Для более детального анализа вводятся дополнительные признаки:

$$\begin{aligned} \text{Trend} &= \text{скользящее среднее}, \\ \text{Momentum} &= \Delta(\text{Bond}) = \text{Bond}(t) - \text{Bond}(t - d), \\ \text{RollingSTD} &= \sqrt{\frac{1}{N} \sum_{i=1}^N (\text{Bond}_i - \text{Trend})^2}, \\ \text{DiffFromTrend} &= \text{Bond} - \text{Trend}, \\ \text{Volume} &= \min(q_{\text{fut}}, q_{\text{stock}}). \end{aligned}$$

Здесь:

- N — количество наблюдений в окне;
- d — шаг для расчёта изменения;
- q_{fut} и q_{stock} — объёмы для фьючерса и акции соответственно.

3.2. Целевая переменная

Цель прогнозирования — определить значение синтетической облигации через n шагов. В регрессионной постановке целевые переменные задаются как:

$$\text{target_sell}(t) = \text{BondSELL}(t + n), \quad \text{target_buy}(t) = \text{BondBUY}(t + n).$$

Также возможна бинарная классификация сигналов, при которой целевая переменная определяется следующим образом:

$$\text{class_label} = \begin{cases} 1, & \text{если } \text{BondSELL}(t + n) < \text{TrendSELL}(t) - k \cdot \text{комиссия}, \\ 0, & \text{иначе,} \end{cases}$$

где k — гиперпараметр, задающий чувствительность модели с учетом издержек.

Таким образом, выбранная формулировка задачи обеспечивает математическую строгость при построении модели и позволяет интегрировать практические аспекты торговли (комиссии, задержки, ликвидность) в симуляцию стратегии.

3.3. Модель и обучение

В качестве базового алгоритма выбрана линейная регрессия (например, `LinearRegression`). Обоснование:

1. Интерпретируемость коэффициентов.
2. Отсутствие сильного переобучения при слабом сигнале.
3. Быстрое обучение и возможность интеграции в симуляцию.

Стратегия обучения:

- Разбиение данных по неделям (обучение — первые 2–4 недели, тест — оставшиеся).
- Обучение модели для направления SELL.
- Сохранение модели (например, `joblib`) для повторяемости экспериментов.

4. Эксперимент

В данном разделе проводится оценка эффективности модели на тестовых данных и симуляция торговой стратегии с учётом практических ограничений.

4.1. Обучение модели и метрики

Модель обучалась на следующих признаках:

- BondSELL
- TrendSELL
- RollingSTD_SELL
- DiffFromTrendSELL
- MomentumSELL
- VolumeSELL

при этом целевая переменная определяется как:

$$\text{target_sell}(t) = \text{BondSELL}(t + n).$$

Полученные показатели качества модели следующие:

$$\text{MAE} = 0.268, \quad \text{MSE} = 0.098, \quad R^2 = 0.912.$$

Высокий коэффициент детерминации R^2 свидетельствует об адекватном согласовании модели с тестовыми данными.

4.2. Интерпретация коэффициентов модели

Коэффициенты линейной регрессии, отсортированные по убыванию абсолютного значения, представлены в таблице:

Признак	Коэффициент
TrendSELL	+0.5126
BondSELL	+0.3716
RollingSTD_SELL	+0.3080
DiffFromTrendSELL	−0.1410
MomentumSELL	+0.0814
VolumeSELL	−0.0001

Положительные коэффициенты усиливают прогнозируемое значение целевой переменной, тогда как отрицательные уменьшают его. Например, значение -0.1410 для DiffFromTrendSELL указывает на наличие эффекта обратного отклонения от тренда (mean reversion).

4.3. Симуляция торговли

Для генерации торговых сигналов используется регрессионный прогноз, переходящий в дискретное решение следующим образом:

$$\text{Signal}(t) = \begin{cases} \text{Открыть шорт,} & \text{если } \text{BondBUY}(t) - \text{PredSELL}(t) > \theta, \\ \text{Нет действий,} & \text{иначе,} \end{cases}$$

где порог $\theta = 0.1$ подобран эмпирически.

Расчёт объёма входа:

$$\text{volume} = \min \left(\left\lfloor \frac{\text{balance}}{\text{OFFER_S_P0}} \right\rfloor, \text{VolumeBUY} \right).$$

Цена входа рассчитывается по формуле:

$$\text{bond_buy_rub} = \frac{\text{BondBUY} \cdot \text{OFFER_S_P0}}{100}.$$

Выход из позиции осуществляется через n шагов с учётом комиссий и исполнения по аналогичной схеме.

Итог симуляции (на тестовом наборе):

$$\text{Количество сделок} = 10,$$

$$\text{Итоговый баланс} = 103359.80 \text{ руб.},$$

$$\Delta \text{баланс} = +3.36\%.$$

Таким образом, симуляция демонстрирует положительное изменение баланса, что свидетельствует о потенциале разработанной торговой стратегии.

5. Графики и визуализация

5.1. Временные ряды Bond

Построение графиков $\text{BondSELL}(t)$ и $\text{TrendSELL}(t)$ позволяет визуально оценить схождение к нулю.

5.2. Оценка порогового параметра θ

Анализ зависимости метрик стратегии (количество сделок, прибыль) от порога θ .

6. Выводы

Построенная модель демонстрирует высокую интерпретируемость и адекватное согласование с данными. Линейная регрессия оказывается достаточно эффективной для прогнозирования краткосрочной динамики синтетической облигации, что подтверждается статистическими метриками и симуляцией торговли. Повышение математической строгости позволит в дальнейшем адаптировать модель для нелинейного прогнозирования, учитывая более сложные зависимости и внешние факторы.

Список литературы

- [1] Халл Дж. К. *Опционы, фьючерсы и другие производные финансовые инструменты*. М.: Вильямс, 2011.
- [2] Гудфеллоу И., Бенжио Й., Курвилл А. *Глубокое обучение*. М.: Диалектика, 2018.