МОДЕЛИ ДОХОДНОСТЕЙ АКТИВОВ В СРЕДНЕ-ДИСПЕРСИОННОМ АНАЛИЗЕ МАРКОВИЦА НА КРИПТОВАЛЮТНЫХ РЫНКАХ

Полузёров Т. Д.

БГУ ФПМИ

Структура работы

1. Портфельная теория
2. Методы оценки характеристик доходностей
3. Проверка статегий на реальных данных
4. Результаты

Однопериодная задача инвестирования

Доступны N активов. S_i^0, S_i^1 — цены i-го актива в моменты времени $t = 0$ и $t = 1$ соответсвенно.
Доходность актива за период

$$r_i := \frac{S^1 - S^0}{S^0}, i = \overline{1, N}$$

Необходимо сформировать портфель

$$b = (b_1, \ldots, b_N)$$

где b_i — число преобретаемых активов i-го типа.
Портфель покупается в момент $t = 0$ и продается в момент $t = 1$ по рыночным ценам.

Требования к портфелю

Пусть инвестор имеет капитал x.
Перейдем к долям инсвестирования капитала x в доступные активы.

$$\omega_i := \frac{b_i S_i^0}{x}, i = \overline{1, N}$$

Цена портфеля в момент времени $t = 0$

$$X^0 = x$$

В момент $t = 1$

$$X^1 = (1 + R)X^0 = \sum_{i=1}^{N} \omega_i r_i = \omega^T r$$

где R — доходность порфтеля

Предположим что случайные величины доходностей $r_i, i = \overline{1, N}$ известны, тогда матожидание и дисперсия случайной величины доходности портфеля равны

$$\mu_X := \mathbb{E}\left[R\right] = \sum_{i=1}^{N} \omega_i \mathbb{E}\left[r_i\right] = \omega^T \mu$$

$$\sigma_X^2 := \sum_{i=1, j=1}^{N} \omega_i \omega_j \mathbf{Cov}\left(r_i, r_j\right) = \omega^T \Sigma \omega$$

где $\mu = \mathbb{E}\left[r\right]$, а Σ — ковариационная матрица случайного вектора r.

Оптимизационная задача

Введем параметр $\tau \in [0, +\infty)$ — толерантность к риску и сформурируем критерий оптимальности

$$f(\mu_X, \sigma_X^2) = \tau\omega^T\Sigma\omega - \omega^T\mu \to \min_{\omega}$$

Оптимизационная задача имеет вид

$$\begin{cases} \tau\omega^T\Sigma\omega - \omega^T\mu \to \min_\omega \\ \omega^T e = 1 \\ \omega \geq 0 \end{cases}$$

При $\tau = 0$ имеем портфель максимальной доходности, при $\tau \to +\infty$ — минимального риска.

На практике в момент времени $t = 0$ случайные величины доходностей r неизвестны.
Необходимо оценить характеристики $\hat{\mu}_X$ и $\hat{\Sigma}$.
Известные подходы:

- Модель CAPM (Capital Asset Pricing Model), У. Шарп, Дж Линтер
- Теория APT (Arbitrage Pricing Theroy), С. Росс, Р. Ролл.

Предлагается иной подход — применить методы машинного обучения и модели временных рядов для решения задачи регрессии по историческим данным.

Линейная регрессия

Пусть X — матрица объектов-признаков, где признаки это лаги ряда, y — соотвествующие истинные доходности. $\theta = (\theta_0, \dots, \theta_k)$ — параметры.

$$a(x) = \theta_0 + \theta_1 x_1 + \dots + \theta_k x_k$$

Оптмальные веса θ^* методом наименьших квадратов

$$Q(a) = ||X\theta - y||^2 \to \min_\theta$$

Случайный лес

ARIMA

Входные данные

Общие характеристики данных:

- Криптовалютная биржа OKX.
- 8 наиболее популярных активов
- Временной период с 1 января 2022 по 1 января 2025.
- Дневной таймфрейм
- Период инвестирования — 1 неделя

Проведение тестирования:

- Данные до 1 января 2024 используются для подбора числа лагов и гиперпараметров моделей
- Тестирование проводится на данных за 2024 год

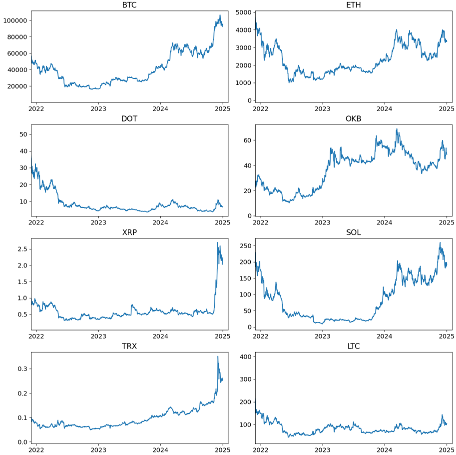


Рис.. Цены активов

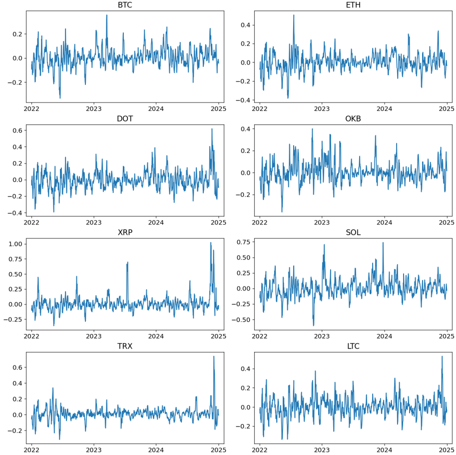
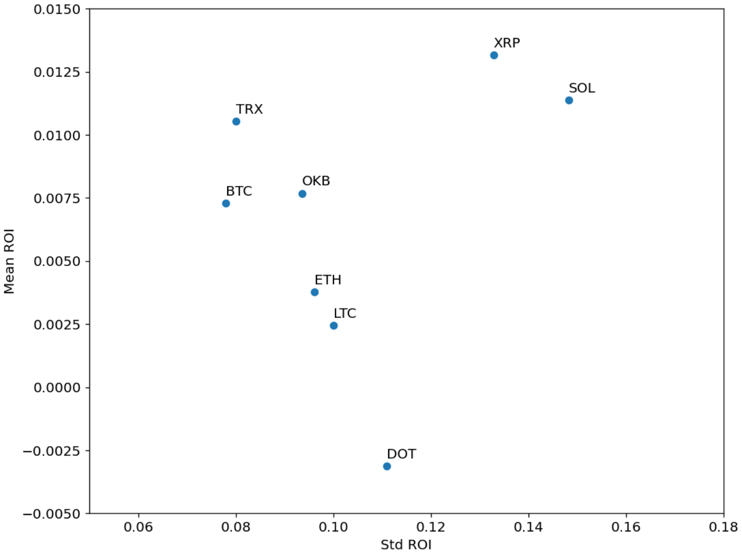
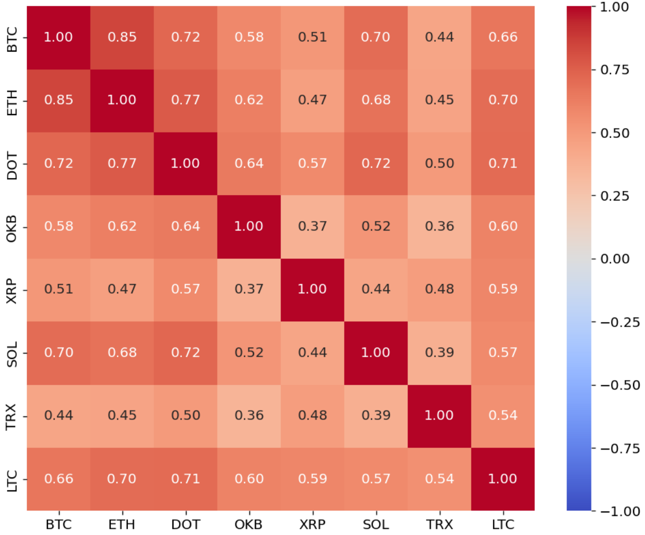


Рис.. Доходности активов

Таблица. Доходности активов

	BTC	ETH	DOT	OKB	XRP	SOL	TRX	LTC
mean	0.0073	0.0038	-0.0031	0.0077	0.0132	0.0114	0.0105	0.0025
std	0.0779	0.0961	0.1110	0.0936	0.1328	0.1482	0.0800	0.1001
min	-0.3328	-0.3830	-0.3925	-0.3591	-0.3596	-0.6018	-0.3162	-0.3392
25%	-0.0358	-0.0477	-0.0724	-0.0401	-0.0506	-0.0765	-0.0236	-0.0513
50%	0.0026	-0.0013	-0.0084	-0.0016	-0.0003	-0.0034	0.0106	0.0001
75%	0.0446	0.0555	0.0573	0.0494	0.0430	0.0906	0.0373	0.0546
max	0.3566	0.5056	0.6188	0.4003	1.0235	0.7409	0.7407	0.5294


XRP
SOL
TRX
OKB
BTC
ETH
LTC
DOT
Mean ROI
Std ROI
0.0150
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
−0.0025
−0.0050
0.06
0.08
0.10
0.12
0.14
0.16
0.18


BTC 1.00 0.85 0.72 0.58 0.51 0.70 0.44 0.66
ETH 0.85 1.00 0.77 0.62 0.47 0.68 0.45 0.70
DOT 0.72 0.77 1.00 0.64 0.57 0.72 0.50 0.71
OKB 0.58 0.62 0.64 1.00 0.37 0.52 0.36 0.60
XRP 0.51 0.47 0.57 0.37 1.00 0.44 0.48 0.59
SOL 0.70 0.68 0.72 0.52 0.44 1.00 0.39 0.57
TRX 0.44 0.45 0.50 0.36 0.48 0.39 1.00 0.54
LTC 0.66 0.70 0.71 0.60 0.59 0.57 0.54 1.00
BTC ETH DOT OKB XRP SOL TRX LTC
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Модели оценки средней доходности:

1. NAIVE — среднее выборочное
2. MARTINGAL — последнее наблюдаемое
3. ARIMA — модель ARIMA
4. LR — линейная регрессия
5. RF — случайный лес

Качество прогонозирования оценивается по метрике

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(r_i - \hat{r}_i)^2$$

где r_i - истинное значение доходности, а $\hat{r}_i$ - прогнозное значение модели на i-м объекте тестовой выборки.

Таблица. Качество прогнозирования, MSE$\cdot 10^4$

	NAIVE	MARTINGAL	LR	ARIMA	RF
BTC	5.63	1.20	1.58	1.62	2.07
ETH	8.00	1.99	4.47	3.70	5.05
DOT	16.51	3.89	4.09	3.98	5.28
OKB	6.06	1.56	1.77	1.95	2.05
XRP	24.33	5.04	6.98	5.71	6.53
SOL	21.19	4.47	11.86	6.16	5.31
TRX	7.96	1.85	4.19	5.30	6.04
LTC	8.53	2.66	2.24	3.22	5.11

Имея r_t - вектор-столбец доходностей в момент времени t, по истории наблюдений $r_1, \cdots r_n$ выборочная ковариация $\hat{\Sigma}$ рассчитывается как

$$\hat{\Sigma} = \frac{1}{n} \sum_{t=1}^{n} (r_t - \overline{r}) \cdot (r_t - \overline{r})^T \tag{1}$$

где $\overline{r} = \frac{1}{n} \sum_{t=1}^{n} r_t$.

Стратегия — принцип по которому в каждый момент времени формируется портфель.
Стратегии Марковица определяются моделью, лежащей в основе оценок $\hat{\mu}$ и $\hat{\Sigma}$, а так же зависят от параметра τ.
Тривиальные стратегии:

1. UNIFORM — равномерное инвестирование во все активы
2. MOST RISKY — наиболее рискованный
3. LESS RISKY — наименее рискованный
4. BEST RETURN — с наибольшей доходностью
5. WORST RETURN — с наименьшей доходностью

Доходность стратегии **ROI** (Return On Investment) определяется аналогично доходности портфеля.

title

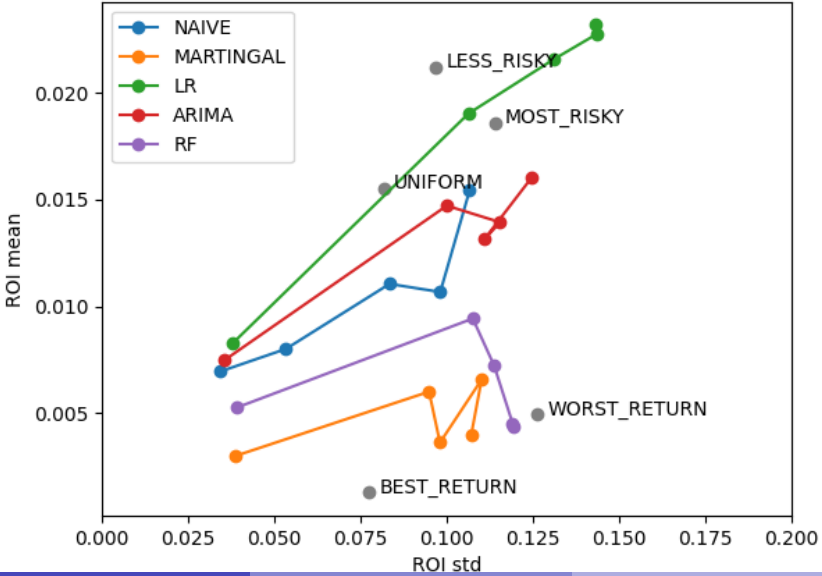

NAIVE
MARTINGAL
LR
ARIMA
RF
LESS_RISKY
MOST_RISKY
UNIFORM
WORST_RETURN
BEST_RETURN
0.020
0.015
0.010
0.005
ROI mean
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175
0.200
ROI std

Таблица. Тривиальные портфели

	mean ROI $\cdot 10^3$	std ROI $\cdot 10^2$
UNIFORM	15.5372	8.1775
MOST RISKY	18.5904	11.3970
LESS RISKY	21.1635	9.6881
BEST RETURN	1.2790	7.7329
WORST RETURN	4.9217	12.6324

Таблица. Средние ROI $\cdot 10^3$

	0.01	0.25	0.50	0.75	1.00
NAIVE	6.9451	8.0025	11.0462	10.6657	15.4227
MARTINGAL	2.9859	6.0019	3.6178	6.5656	3.9942
LR	8.2600	19.0185	21.5537	22.7442	23.1668
ARIMA	7.4648	14.7066	13.9296	13.1520	15.9971
RF	5.2633	9.4236	7.2398	4.3777	4.5050

Таблица. Стандартное отклонение ROI $\cdot 10^2$

	0.01	0.25	0.50	0.75	1.00
NAIVE	3.4328	5.3375	8.3427	9.8131	10.6650
MARTINGAL	3.8577	9.4886	9.8051	11.0111	10.7128
LR	3.7892	10.6314	13.1000	14.3570	14.3071
ARIMA	3.5454	10.0037	11.5506	11.1001	12.4518
RF	3.9215	10.7564	11.3710	11.9397	11.9118

Выводы:

- активы имеют сильную положительную корреляцию
- стремление сформировать портфель с большей доходностью влечет большие риски
- диверсификация действительно позволяет снижать риск портфеля
- как правило, формирование портфеля доминирует над инвестированием в отдельные активы
- линейная модель авторегрессии показала лучшее качество для оценки средней доходности

Дальнейшие шаги по исследованию данной темы:

- рассмотреть другие классы методов прогнозирования временных рядов
- помимо авторегрессионных признаков, учесть влияние внешних факторов на формирование цен
- расширить рассматриваемый набор активов
- исследовать другие таймфреймы и периоды инвестирования

Спасибо за внимание.