

*Боцман Дмитрий Алексеевич,
студент,
МИРЭА – Российский технологический университет,
г. Москва*

*Шугайлов Никита Александрович,
студент,
МИРЭА – Российский технологический университет,
г. Москва*

**ВЕКТОРНЫЙ АНАЛИЗ СТРУКТУРНОГО СХОДСТВА
ИНВЕСТИЦИОННЫХ АКТИВОВ: ПРИМЕНЕНИЕ
МОДИФИЦИРОВАННОЙ КОСИНУСНОЙ МЕРЫ ДЛЯ ОЦЕНКИ
РИСК-ПРОФИЛЯ**

Аннотация. В статье предложен многомерный подход к оценке структурного сходства инвестиционных активов с использованием модифицированной косинусной меры. Традиционные методы оценки диверсификации уязвимы к рыночному шуму и опираются на ценовую ковариацию. Для решения этой проблемы разработан алгоритм векторной алгебры с применением инверсии риск-метрик, позволяющий перевести показатели доходности, волатильности и максимальной просадки в единое пространство признаков. На основе эмпирических данных 18 глобальных биржевых фондов (ETF) за 2016–2025 годы построена тепловая матрица кросс-секторного сходства и проведено сопоставление активов с экстремальными целевыми профилями инвестора. Результаты подтверждают глобализацию отраслевых рисков и предлагают новый математический инструментарий для формирования устойчивых инвестиционных портфелей, нивелирующий эффект ложной диверсификации.

Ключевые слова: векторный анализ, косинусное сходство, инвестиционный портфель, диверсификация рисков, биржевые фонды (ETF), риск-профиль, инверсия рисков.

Botsman Dmitry Alexeevich,
student,

MIREA – Russian technological university,
Moscow

Shugailov Nikita Alexandrovich,
student,

MIREA – Russian technological university,
Moscow

**VECTOR ANALYSIS OF STRUCTURAL SIMILARITY OF
INVESTMENT ASSETS: APPLICATION OF A MODIFIED COSINE
MEASURE FOR RISK PROFILE ASSESSMENT**

Abstract. *The article proposes a multidimensional approach to assessing the structural similarity of investment assets using a modified cosine measure. Traditional diversification assessment methods are vulnerable to market noise and rely heavily on price covariance. To address this, a vector algebra algorithm employing risk-metric inversion was developed, translating returns, volatility, and maximum drawdown into a unified normalized feature space. Based on empirical data from 18 global Exchange-Traded Funds (ETFs) over 2016–2025, a cross-sector similarity heatmap was constructed, and assets were matched against extreme investor target profiles. The results confirm the globalization of sectoral risks and offer a novel mathematical toolkit for building resilient investment portfolios, effectively mitigating the illusion of diversification.*

Keywords: *vector analysis, cosine similarity, investment portfolio, risk diversification, exchange-traded funds (ETF), risk profile, risk inversion.*

В условиях нарастающей макроэкономической нестабильности фундаментальной проблемой современного инвестиционного анализа становится эффект «ложной диверсификации». Классические подходы к распределению капитала, опирающиеся на историческую ковариацию цен, демонстрируют существенную уязвимость в периоды системных шоков. Исторический опыт показывает, что в моменты острой нехватки ликвидности срабатывает эффект рыночного заражения, при котором корреляция большинства классов активов (включая традиционно защитные) стремится к единице. Географическая и номинальная отраслевая диверсификация перестают обеспечивать защиту портфеля, что требует разработки принципиально новых, многомерных метрик оценки структурного сходства активов.

Объектом исследования выступает глобальный рынок биржевых инвестиционных фондов (ETF), отражающий динамику ключевых секторов мировой экономики.

Предмет исследования — структурное сходство риск-профилей инвестиционных активов и методы его многомерной оценки на основе векторной алгебры.

Цель исследования заключается в разработке и апробации векторной модели кросс-секторного сходства активов с применением модифицированной косинусной меры для выявления скрытой структурной идентичности фондов и оптимизации целевых инвестиционных стратегий.

Задачи исследования:

1. Сформировать эмпирическую базу исторических котировок отраслевых и суверенных ETF за десятилетний период (2016–2025 гг.).
2. Разработать алгоритм трансформации базовых метрик (доходность, волатильность, максимальная просадка, коэффициент Шарпа) в единое пространство признаков с применением метода инверсии рисков.
3. Построить тепловую матрицу попарного косинусного сходства для выявления кластеров ложной диверсификации.

4. Провести сопоставление реальных фондов с экстремальными эталонными векторами целевых стратегий инвестора.

Рабочая гипотеза состоит в том, что переход от анализа ценовой ковариации к расчету косинусного сходства в инвертированном пространстве риск-метрик позволяет доказать глобализацию отраслевых рисков и выявить инструменты, сохраняющие ортогональность к рынку акций в периоды кризисов.

Фундаментальной основой оценки инвестиционных активов является современная портфельная теория, заложенная Г. Марковицем. В своей базовой работе он обосновал правило «ожидаемая доходность — дисперсия» (expected returns — variance of returns), математически доказав, что инвестор при принятии решений должен рассматривать ожидаемую доходность как желаемый фактор, а дисперсию доходности — как нежелательный, стремясь к минимизации последней посредством диверсификации портфеля [1, с. 77]. Дальнейшее развитие данный подход получил в трудах У. Шарпа, который формализовал оценку эффективности через отношение ожидаемой дифференциальной доходности к прогнозируемому стандартному отклонению этой доходности (ex ante Sharpe Ratio). Этот показатель позволил измерять премию за каждую единицу принимаемого инвестором риска [2, с. 50].

Однако, как отмечается в отечественной академической литературе, классические модели оценки риска требуют существенной концептуальной адаптации применительно к развивающимся рынкам капитала, которые подвержены высокой макроэкономической волатильности и инфраструктурным аномалиям [3, с. 119–120]. Исторический опыт макроэкономических шоков доказывает уязвимость традиционной оценки диверсификации на основе дисперсии и ковариации цен. В частности, К. Рейнхарт и К. Рогофф, описывая исторический синдром «в этот раз все иначе» (this-time-is-different syndrome), подчеркивают, что инвесторы систематически игнорируют риски высокой долговой нагрузки, принимая временный бум за

новые структурные реалии, что неизбежно приводит к масштабным кластерным кризисам [4, с. 14].

В моменты таких системных шоков на финансовых рынках срабатывает эффект «заражения» (contagion). По определению Г. Бекаерта и соавторов, рыночное заражение представляет собой избыточную корреляцию между рынками сверх той, что может быть объяснена экономическими фундаментальными факторами [5, с. 1–2]. Под воздействием скрытого влияния общих стресс-факторов корреляция большинства активов стремится к единице, что делает традиционные портфели уязвимыми [5, с. 20].

Практическим подтверждением краха подходов наивной диверсификации стала ситуация на российском рынке коллективных инвестиций в 2022 году. Как отмечают Г. Т. Гафурова и др., реакция рынка на геополитический шок оказалась глубоко неоднородной: хотя 2022 год формально ознаменовался самым большим приростом стоимости чистых активов за счет сегмента закрытых паевых фондов (ЗПИФ), сегмент открытых фондов (ОПИФ) акций понес рекордные убытки [6, с. 500, 506]. Ключевой проблемой, разрушившей инвестиционные портфели, стала заморозка иностранных ценных бумаг, затронувшая 159 паевых фондов (29% от их общего количества) [6, с. 508]. Этот прецедент наглядно продемонстрировал, что номинальная географическая диверсификация перестает защищать капитал инвестора в условиях системных ограничений, что требует перехода к принципиально новым, многомерным метрикам оценки структурного сходства активов.

В контексте поиска оптимальных инструментов для защиты капитала от системных шоков особое внимание исследователей привлекают биржевые инвестиционные фонды (ETF). Традиционно они позиционируются как прозрачные и высоколиквидные инструменты пассивного инвестирования, позволяющие с минимальными издержками реплицировать доходность широких рыночных индексов. Однако глубокий анализ их операционной структуры показывает, что управление реальной стоимостью портфеля ETF

сопряжено с рядом фундаментальных проблем, искажающих заявленный риск-профиль актива.

Как отмечает М. И. Львова, реальная биржевая цена ETF формируется под воздействием рыночного спроса и предложения, а не определяется исключительно чистой стоимостью его базовых активов (Net Asset Value, NAV). В периоды рыночной турбулентности или ограничения ликвидности это приводит к возникновению значительных премий или дисконтов к NAV, что генерирует для инвесторов неконтролируемые риски и отклонения от ожидаемой доходности [7, с. 65–66].

Кроме того, внутренняя структура таких фондов зачастую страдает от скрытой дисперсии и непрозрачности. Использование методов «коллективного инвестирования», сложности с оценкой неликвидных или неторгуемых активов внутри фонда, а также издержки и запаздывания при регулярной ребалансировке приводят к устойчивому расхождению (tracking error) между эталонным индексом и реальным портфелем ETF [7, с. 67–68].

Таким образом, структурная непрозрачность и подверженность биржевых фондов внутренним дисбалансам доказывают, что инвестор не может опираться исключительно на номинальные (маркетинговые) характеристики ETF или их ценовые котировки в спокойные периоды. Подобная специфика объекта исследования обосновывает необходимость применения многомерного математического аппарата для оценки истинного, скрытого сходства риск-профилей различных фондов.

Классические подходы к скорингу активов, опирающиеся на простые коэффициенты корреляции временных рядов, обладают существенными ограничениями из-за высокого уровня рыночного шума. Как доказано в фундаментальных работах по экономфизике, для выявления истинной иерархической структуры финансового рынка необходимо отказываться от простых корреляционных матриц в пользу обобщенных метрик расстояния (в частности, $d(i, j) = 1 - p_{ij}^2$). Построение на их основе минимальных остовных

деревьев и субдоминантных ультраметрических иерархий позволяет выявлять скрытые экономические факторы, влияющие на ценообразование [8, с. 3, 6].

Этот подход находит активное применение в современной российской практике. Использование алгоритмов кластеризации на фондовом рынке позволяет математически точно выделять группы активов с различным уровнем волатильности и доходности, а также фиксировать паттерны поведения инвесторов в отношении принятия системных рисков в моменты макроэкономических разногласий [9, с. 99, 110]. Более того, применение алгоритмов иерархической кластеризации решает фундаментальную проблему классической портфельной теории Марковица — избыточную размерность данных. Как показывают современные исследования, разделение активов на однородные кластеры позволяет многократно (в пять и более раз) снизить вычислительную сложность алгоритмов при сохранении приемлемого и контролируемого качества оптимального инвестиционного решения [10, с. 63, 70].

Однако прямой перенос математических алгоритмов на сырые исторические ряды котировок несет в себе критические риски. М. Лопес де Прадо подвергает жесткой критике классическую эконометрику за неспособность обучаться и учитывать сложные нелинейные закономерности, указывая на фатальную проблему переобучения моделей (*backtest overfitting*). Для получения объективных результатов необходим переход от зашумленных временных рядов цен к структурированному многомерному пространству признаков (*feature space*), поскольку именно оценка важности признаков (*feature importance*), а не простое историческое симулирование, является истинным инструментом научного исследования [11, с. 15, 17, 114].

Анализ научной литературы показывает, что, несмотря на активное внедрение алгоритмов кластеризации для снижения размерности и группировки рисков [10, с. 63; 5, с. 99], подавляющее большинство исследователей по-прежнему опираются на метрику евклидова расстояния или корреляцию Пирсона применительно к рядам исторической доходности.

В настоящей работе предлагается заполнить данный методологический пробел. Опираясь на строгий математический аппарат информационного поиска и векторной алгебры [12], исследование предлагает применение косинусной меры в инвертированном векторном пространстве риск-метрик. Такой подход позволяет математически оценивать фундаментальную структурную идентичность активов, эффективно нивелируя влияние эффекта ложной диверсификации и ценового шума.

Для проведения структурного анализа была сформирована репрезентативная выборка, включающая 18 глобальных биржевых инвестиционных фондов (ETF), торгуемых на мировых площадках. Выборка охватывает как широкие рыночные индексы, так и узкоспециализированные отраслевые сектора (информационные технологии, энергетика, финансовый сектор, агробизнес) и суверенные долговые инструменты (государственные облигации США, Европы и Китая).

Временной горизонт исследования охватывает десятилетний период с 1 января 2016 года по 31 декабря 2025 года. Данный интервал был выбран целенаправленно, поскольку он включает в себя фазы стабильного макроэкономического роста, период пандемического шока ликвидности (2020 г.) и фазу глобальной инфляции и геополитической турбулентности (2022–2023 гг.). Эмпирическую базу составили ежедневные скорректированные цены закрытия (Adjusted Close), учитывающие реинвестирование дивидендов и сплиты акций. Данные были выгружены с аналитической платформы Yahoo Finance [13]. Общий объем наблюдений для каждого актива составил $n = 2520$ торговых дней.

В соответствии с гипотезой исследования о необходимости перехода от зашумленных временных рядов к структурной оценке активов, на первом этапе методологии был произведен расчет четырех базовых макропризнаков. Каждый ETF был представлен в виде вектора в четырехмерном пространстве $V = [x_1, x_2, x_3, x_4]$, где координаты отражают фундаментальный риск-профиль актива:

- x_1 — Ожидаемая доходность (Return). Рассчитывалась как среднегодовой темп роста (CAGR) на основе ежедневных логарифмических доходностей за весь анализируемый период. Данная координата отражает потенциал актива к преумножению капитала.

- x_2 — Волатильность (Volatility). Вычислялась как аннуализированное стандартное отклонение ежедневных доходностей фонда. Данная метрика выступает классической мерой рыночного риска и неопределенности в парадигме Г. Марковица [1, с. 77].

- x_3 — Максимальная просадка (Maximum Drawdown). Рассчитывалась как максимальный исторический убыток от локального ценового максимума до последующего минимума. Включение данного непараметрического показателя критически важно для компенсации эффекта «толстых хвостов» (fat tails) нормального распределения, так как он фиксирует реальную стрессоустойчивость актива в моменты обвалов.

- x_4 — Коэффициент Шарпа (Sharpe Ratio). Вычислялся как отношение избыточной доходности актива к его волатильности [2, с. 50]. В качестве безрисковой ставки (Risk-Free Rate) для расчета избыточной доходности была принята доходность 10-летних казначейских облигаций США (US 10-Year Treasury Yield), что соответствует глобальному характеру выборки и десятилетнему горизонту исследования. Координата выступает интегральной мерой эффективности, показывая премию, которую рынок выплачивает инвестору за каждую единицу принятого риска.

Таким образом, сырые исторические данные котировок были трансформированы в матрицу признаков размерностью 18×4 . Однако прямое вычисление метрик расстояния на данном этапе математически некорректно из-за различного масштаба и разнонаправленности векторов (например, рост доходности — это позитивный фактор, а рост волатильности — негативный).

Стоит отметить, что включение коэффициента Шарпа (x_4) наряду с сырыми показателями доходности (x_1) и волатильности (x_2) формирует

проблему частичной мультиколлинеарности признаков. Однако в контексте построения многомерного векторного пространства данный шаг является математически обоснованным. Коэффициент Шарпа выступает в роли нелинейного весового индикатора (utility weight): он позволяет алгоритму расчета расстояний дифференцировать активы с «пустой» (некомпенсированной) волатильностью от активов, где принятие риска сопровождается соразмерной премией. Таким образом, предложенный набор метрик формирует эвристическое признаковое пространство, отражающее многомерную природу инвестиционного риска.

Поскольку рассчитанные макропризнаки имеют различную размерность (например, доходность измеряется в процентах, а коэффициент Шарпа является абсолютной величиной), для корректного применения метрик расстояния необходимо было привести все значения к единому масштабу. На данном этапе применялась линейная минимаксная нормализация (Min-Max Scaling), переводящая каждую координату в диапазон [0; 1]:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Критически важным методологическим шагом стала адаптация пространства под расчет косинусного сходства. В классической финансовой математике векторы часто разнонаправлены: рост координат x_1 (доходность) и x_4 (Шарп) является позитивным для инвестора, в то время как рост координат x_2 (волатильность) и x_3 (просадка) — негативным. Расчет углового расстояния в таком неоднородном пространстве приводит к математическим искажениям.

Для решения этой проблемы в исследовании был применен авторский метод «инверсии риск-метрик». Отрицательные по своей экономической сути координаты были инвертированы таким образом, чтобы значение 1.0 всегда отражало «идеальное» (безрисковое) состояние актива, а 0.0 — наихудшее.

- Координата волатильности была трансформирована в метрику «Безопасности цены»: $x_{2(safe)} = 1.0 - x_{2(norm)}$

- Координата максимальной просадки была трансформирована в метрику «Стрессоустойчивости»: $x_{3(safe)} = 1.0 - x_{3(norm)}$

В результате было сформировано единое векторное пространство признаков $V' = [x_{1(norm)}, x_{2(safe)}, x_{3(safe)}, x_{4(norm)}]$

На финальном этапе методологии для выявления скрытого структурного сходства между активами применялась модифицированная косинусная мера (Cosine Similarity). В отличие от евклидова расстояния, косинусная мера оценивает не абсолютную разницу между точками, а угол между векторами, что позволяет эффективно выявлять активы с идентичным «поведением» риск-профиля независимо от их абсолютной доходности. Формула расчета сходства между векторами A и B имеет вид:

$$\sin(A, B) = \cos(\theta) = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

Для классификации активов относительно целей инвестора в созданное 4-мерное пространство были искусственно внедрены два эталонных экстремальных (ортогональных) вектора:

1. Консервативный профиль (Вектор сохранения капитала): $V_{cons} = [0.0, 1.0, 1.0, 0.0]$. Данный вектор математически описывает инвестора, для которого важна исключительно 100-процентная безопасность от рыночных скачков и просадок, а премия за риск и доходность не имеют значения.

2. Агрессивный профиль (Вектор преумножения): $V_{aggr} = [1.0, 0.0, 0.0, 1.0]$. Этот вектор описывает чистый поиск сверхдоходности и эффективности по Шарпу при полном игнорировании рисков обвала капитала.

Расчет косинусного сходства между реальными рыночными ETF и данными эталонными векторами позволяет построить объективную, математически обоснованную иерархию активов, нивелируя влияние маркетингового шума и номинальной отраслевой принадлежности.

Сбор сырых исторических котировок, расчет многомерного признакового пространства и векторные вычисления производились в среде программирования Python с использованием библиотек Pandas и NumPy.

В рамках апробации предложенной методологии была проанализирована выборка из 18 глобальных биржевых инвестиционных фондов. Для обеспечения репрезентативности в выборку были включены инструменты, отражающие различные региональные рынки (США, Европа, Китай) и сектора экономики (технологии, финансы, энергетика, агробизнес), а также защитный актив в виде суверенного долга США. Структурный состав анализируемой выборки представлен в Таблице 1.

Таблица 1. Эмпирическая выборка исследования (2016-2025 гг.)

Тикер	Название фонда (ETF)	Базовый сектор / Регион
CBON	VanEck China Bond ETF	Государственные облигации (Китай)
CHIE	Global X MSCI China Energy	Энергетический сектор (Китай)
CHIQ	Global X MSCI China Consumer	Потребительский сектор (Китай)
CHIX	Global X MSCI China Financials	Финансовый сектор (Китай)
CQQQ	Invesco China Technology ETF	Технологический сектор (Китай)
EUFN	iShares MSCI Europe Financials	Финансовый сектор (Европа)
EUNH.DE	iShares STOXX Europe 600 Health	Здравоохранение (Европа)
EXV1.DE	iShares STOXX Europe 600 Banks	Банковский сектор (Европа)
EXV3.DE	iShares STOXX Europe 600 Tech	Технологический сектор (Европа)
EXV4.DE	iShares STOXX Europe 600 Telecom	Телекоммуникации (Европа)
GOVT	iShares U.S. Treasury Bond ETF	Государственные облигации (США)
KSTR	KraneShares SSE STAR Market 50	Инновации и биотехнологии (Китай)
MOO	VanEck Agribusiness ETF	Агропромышленный комплекс (Глобально)
SMH	VanEck Semiconductor ETF	Полупроводники (США/Глобально)
SOXX	iShares Semiconductor ETF	Полупроводники (США)

XLE	Energy Select Sector SPDR	Энергетический сектор (США)
XLF	Financial Select Sector SPDR	Финансовый сектор (США)
XLK	Technology Select Sector SPDR	Технологический сектор (США)

На основе инвертированного 4-мерного признакового пространства была построена матрица косинусного сходства (Рисунок 1). Математической особенностью полученной матрицы является смещение базового уровня сходства: большинство значений превышает порог 0.50. Данный феномен выступает следствием применения метода «инверсии рисков». Поскольку все координаты были переведены в строго положительную логику максимизации выгоды, векторы фондов локализовались в едином положительном ортанте признакового пространства.

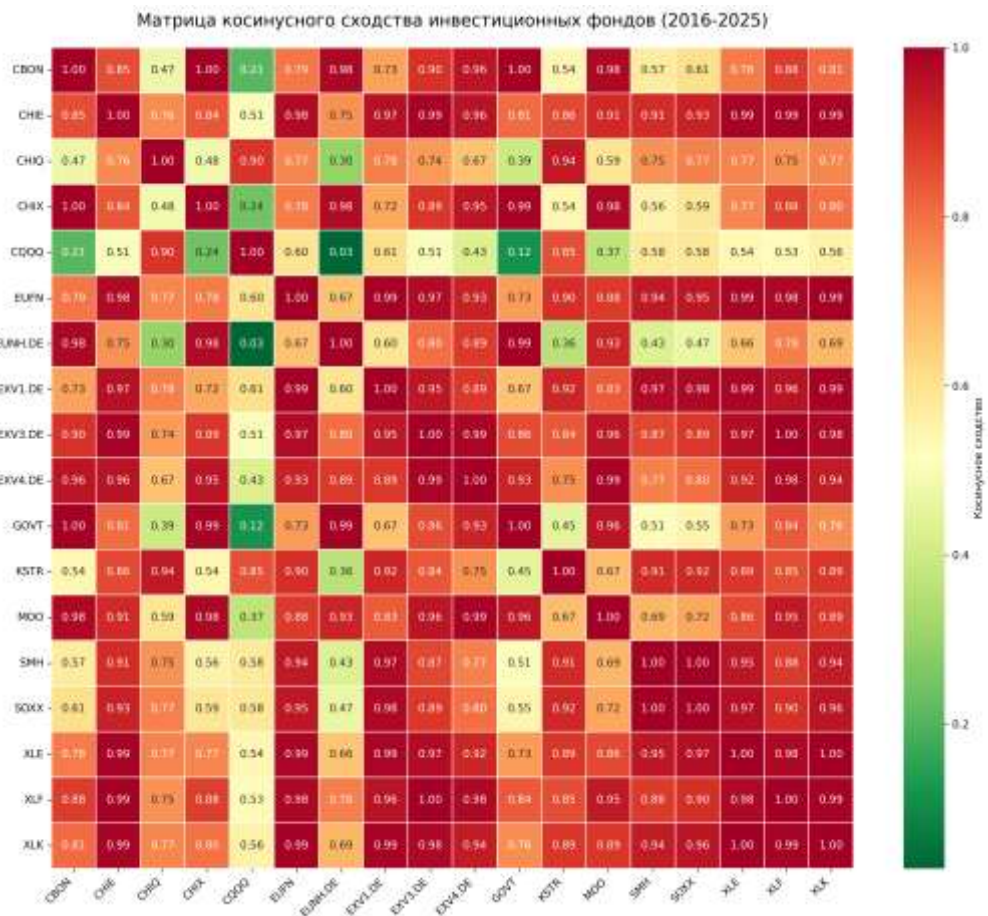


Рисунок 1. Матрица косинусного сходства структурных риск-профилей ETF

В этих условиях наличие у активов глобальной рыночной беты (синхронной реакции на системные макроэкономические шоки) не позволяет им демонстрировать абсолютную ортогональность. В рамках данного исследования пороговые значения для кластеризации были определены эвристически: истинной структурной диверсификацией признаются значения в диапазоне 0.50–0.74, в то время как значения свыше 0.90 (верхний дециль распределения) ассоциируются с высоким уровнем дублирования рисков.

Анализ тепловой матрицы указывает на наличие эффекта, который можно охарактеризовать как «ложную диверсификацию». В частности, фонды технологического сектора США (XLK) и полупроводников (SMH, SOXX) образуют тесный кластер (сходство > 0.95). Аналогичная тенденция прослеживается при кросс-региональном анализе: европейский финансовый сектор (EUFN, EXV1.DE) демонстрирует высокое структурное сходство с американским финансовым сектором (XLF). Это эмпирически демонстрирует, что на исследуемой выборке географическая диверсификация в рамках одной отрасли обладает сниженной эффективностью из-за глобального эффекта рыночного заражения (contagion).

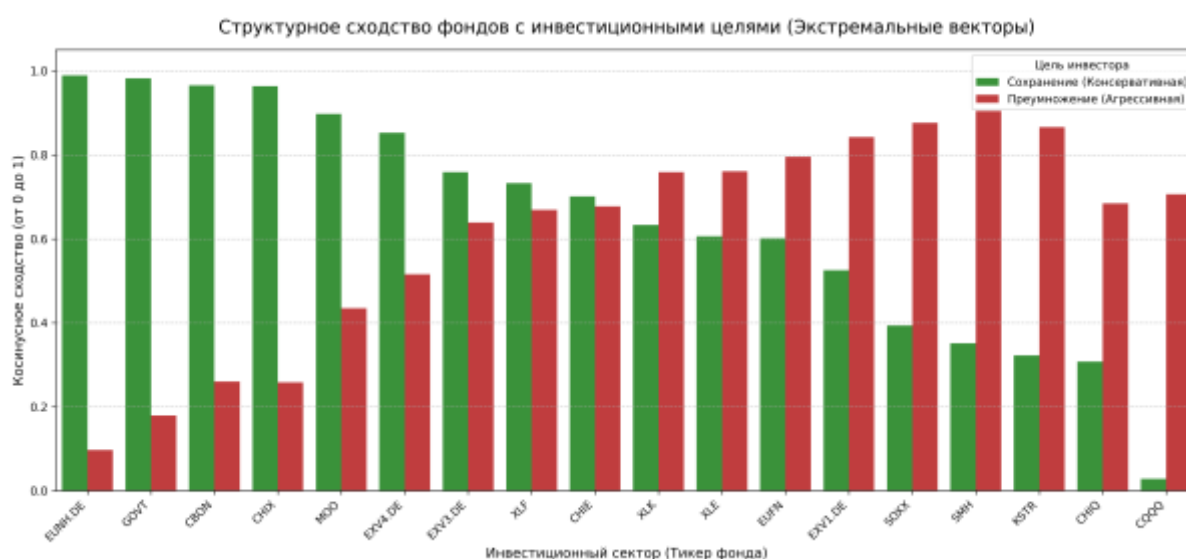


Рисунок 2. Сходство фондов с целевыми инвестиционными стратегиями

Для иллюстрации прикладных возможностей модели было оценено косинусное сходство реальных активов с двумя искусственными эталонными векторами — экстремально Консервативным и экстремально Агрессивным (Рисунок 2). Результаты расчетов показывают структурную дифференциацию фондов:

1. Консервативный профиль. Высокую степень сходства с вектором безопасности, помимо казначейских облигаций США (GOVT), продемонстрировал фонд глобального агропромышленного комплекса (MOO), что указывает на его потенциальную функцию как нетривиального защитного актива на анализируемом отрезке.

2. Агрессивный профиль. Наивысшую степень сходства с вектором преумножения капитала показали технологические фонды (SMH, SOXX), исторически конвертировавшие свою волатильность в соразмерную премию по Шарпу.

3. Зона неэффективности. Фонды китайского рынка (CHIQ, CHIX) на рассматриваемом горизонте продемонстрировали параметры, ортогональные обеим целевым стратегиям, что связано с реализацией специфических региональных рисков и глубокими историческими просадками.

Таким образом, применение многомерной векторной оценки позволяет инвесторам алгоритмически избегать ложной диверсификации и формировать портфели, строго соответствующие их индивидуальным целевым метрикам толерантности к риску.

Несмотря на высокую объясняющую способность предложенной векторной модели, необходимо выделить ряд объективных методологических ограничений, определяющих границы применимости полученных результатов:

1. Статичность анализируемого горизонта. В рамках текущего исследования признаковое пространство формировалось на основе агрегированных данных за весь 10-летний период (in-sample analysis). Данный подход эффективен для выявления фундаментального структурного сходства,

однако он математически сглаживает локальные сдвиги рыночных режимов (regime shifts). Перспективным развитием модели является переход к динамической кластеризации с использованием метода «скользящего окна» (rolling window) для оценки эволюции косинусного сходства в моменты зарождения кризисов.

2. Чувствительность алгоритмов нормализации. Примененный метод линейной минимаксной нормализации (Min-Max Scaling) строго необходим для приведения метрик к масштабированному базису $[0; 1]$, однако он обладает математической уязвимостью к экстремальным рыночным выбросам (outliers). Наличие в выборке актива с аномальной доходностью или беспрецедентным крахом способно искусственно сжать шкалу распределения для остальных фондов, что требует аккуратного подхода к предварительной фильтрации данных.

3. Ограниченность размерности пространства. Текущая 4D-модель опирается исключительно на ценовые макро-метрики. Однако, опираясь на выявленную специфику биржевых фондов [7], реальный риск-профиль ETF уязвим к операционной непрозрачности. Расширение векторного пространства дополнительными координатами (ошибка следования индексу — tracking error, спред ликвидности, комиссии за управление) позволит сделать оценку еще более прецизионной.

4. Историческая ретроспективность (Ex-post vs Ex-ante). Следуя строгим критериям оценки алгоритмических моделей [11], следует признать, что зафиксированные угловые расстояния до эталонных целевых стратегий отражают исторически реализованную парадигму. Обоснованная экстраполяция выявленных кластеров ложной диверсификации на будущие периоды потребует внедрения методов комбинаторной кросс-валидации (CPCV) для защиты портфелей от эффекта переобучения на исторических данных.

В настоящем исследовании предпринята попытка адаптации методов векторного анализа признакового пространства для решения проблемы оценки

диверсификации инвестиционных портфелей в условиях макроэкономической нестабильности. Авторами предложен эвристический алгоритм, основанный на инверсии классических риск-метрик и расчете косинусного сходства фондов.

Эмпирическая апробация метода на выборке из 18 глобальных ETF за период 2016–2025 гг. позволила сделать следующие выводы:

1. Использование косинусного сходства в нормированном признаковом пространстве позволяет наглядно идентифицировать кластеры с высоким уровнем дублирования рисков. Анализ показал, что номинальная географическая диверсификация в рамках смежных отраслей часто сопровождается значительным структурным сходством активов.

2. Внедрение искусственных ортогональных векторов («Консервативного» и «Агрессивного») предоставляет исследователям и аналитикам дополнительный инструментарий для ранжирования активов сообразно целевым профилям.

3. Продемонстрировано, что на исследуемом горизонте секторальные активы (например, агропромышленный комплекс) могут выступать более эффективными диверсификаторами, чем традиционные международные фонды акций, подверженные эффекту рыночного заражения.

Предложенная методология не претендует на замену классической портфельной теории, однако выступает полезной комплементарной (вспомогательной) эвристической метрикой. Практическая ценность подхода заключается в возможности его интеграции в скрининговые системы риск-менеджмента для первичной фильтрации активов и предотвращения избыточной концентрации скрытых структурных рисков в портфеле.

Список литературы:

1. Markowitz, H. Portfolio Selection // The Journal of Finance. — 1952. — Vol. 7, N 1. — P. 77–91.
2. Sharpe W. F. The Sharpe Ratio // The Journal of Portfolio Management. — 1994. — Vol. 21, N 1. — P. 49–58.
3. Теплова Т. В. Инвестиции: учебник для бакалавров / Т. В. Теплова. — Москва: Издательство Юрайт, 2011. — 724 с.
4. Reinhart, C. M. This Time Is Different: Eight Centuries of Financial Folly —2009. — 27 p.
5. Bekaert G., Harvey C. R., Ng A. Market Integration and Contagion // Journal of Business. — 2005. — Vol. 78, N 1. — P. 39–69.
6. Гафурова Г. Т., Нотфуллина Г. Н., Ковалева Э. Р. Тренды развития паевых инвестиционных фондов в условиях санкций / // Региональная и отраслевая экономика. — 2023. — № 3. — С. 490–514.
7. Львова М. И. Проблемы управления реальной стоимостью портфеля ETF / // Финансы и общество: новые реалии, новые форматы, новые финансовые инструменты. — С. 65–68.
8. Mantegna R. N. Hierarchical structure in financial markets // The European Physical Journal B. — 1999. — Vol. 11, N 1. — 18 p.
9. Файзулин М. С. Сентимент и стадное поведение частных инвесторов: кластерный анализ российского фондового рынка // Финансовый журнал. — 2024. — Т. 16, № 4. — С. 95–113.
10. Полетаев А. Ю., Спиридонова Е. М. Иерархическая кластеризация как метод снижения размерности в задаче оптимизации инвестиционного портфеля Марковица // Моделирование и анализ информационных систем. — 2020. — Т. 27, № 1. — С. 62–71.
11. Lopez de Prado M. Advances in Financial Machine Learning // Hoboken: John Wiley & Sons, 2018. — 366 p.
12. Manning C. D., Raghavan P., Schütze H. Introduction to Information Retrieval – 2008. — 544 p.

13. Yahoo Finance API. Historical Market Data [Electronic resource]. —
URL: <https://finance.yahoo.com> (дата обращения: 25.05.2026).