

УДК 336.763.4

**РЕЖИМНО-ЗАВИСИМАЯ ПАРАМЕТРИЗАЦИЯ ОПЦИОННОЙ
КОНСТРУКЦИИ ВХОДА В АКТИВ****Бураков Никита Сергеевич,**

Количественный исследователь

IRQ, 115054, город Москва, Дубининская ул., д. 53 стр. 5, этаж 2, помещение 7а

E-mail: nikhbrown35@gmail.com

SPIN-код автора: 3077-1888

Аннотация

В работе рассматривается задача параметризации опционной конструкции входа в базовый актив при ограниченном капитале обеспечения. Инвестор заинтересован в долгосрочном владении активом, однако предпочитает вход по более низкой цене, не теряя участия в росте рынка. Для этого исследуется конструкция из короткого put и длинного call с параметризацией страйков через коэффициенты относительно спота.

Для решения этой задачи предлагается конструкция, состоящая из проданного put опциона (источник премиального дохода и инструмент условного входа по сниженной цене) и купленного call-опциона (инструмент участия в росте). Ключевым вопросом является выбор страйков обоих инструментов. Предлагается двухэтапный подход.

На первом этапе по длинной истории дневных цен выполняется сценарный перебор коэффициентов страйков на дискретной сетке, при этом оценка опционных премий производится посредством синтетического дэска на основе модели Black-Scholes-Merton.

На втором этапе восстанавливается отображение из пространства рыночных признаков (реализованная волатильность, её динамика, моментум, диапазоны) в исторически оптимальные коэффициенты с помощью мультывыходной регрессии CatBoost (MultiRMSE).

Предложенный подход переводит выбор страйков из области экспертной эвристики в воспроизводимую количественную процедуру, адаптирующуюся к текущему волатильностному режиму рынка.

Ключевые слова: опционная конструкция, параметризация страйков, синтетический дэск Блэка-Шоулза-Мертон, реализованная волатильность, волатильностный режим, сценарный бэкстест, градиентный бустинг, CatBoost, мультывыходная регрессия, признаковое пространство OHLCV

**REGIME-DEPENDENT PARAMETERIZATION OF AN OPTION-BASED
ENTRY STRUCTURE****Burakov Nikita Sergeevich,**

Quantitative Researcher

IRQ, 115054, Moscow, Dubininskaya St., 53 bldg. 5, floor 2, room 7a

E-mail: nikhbrown35@gmail.com

Author's SPIN code: 3077-1888

ABSTRACT

This paper studies the parameterization problem for an option-based entry structure into an underlying asset under constrained capital allocation. The investor is interested in longterm ownership of the asset but prefers entering at a lower level without losing upside participation if the market rises without a pullback.

To address this, we analyze a two-leg structure consisting of a short put and a long call, where strikes are parameterized by spot-relative coefficients. A two-stage methodology is proposed. First, a scenario-based grid search over strike coefficients is performed on a long daily history, while option premiums are computed through a synthetic option board based on the Black–Scholes–Merton framework. Second, a mapping from market-state features (realized volatility, volatility acceleration, momentum, and range-related metrics) to historically optimal coefficients is estimated using multi-output CatBoost regression (MultiRMSE).

The resulting framework converts strike selection from discretionary heuristics into a reproducible quantitative procedure adaptive to current volatility regimes.

Keywords: option structure, strike parameterization, Black–Scholes–Merton synthetic desk, realized volatility, volatility regime, scenario backtest, gradient boosting, CatBoost, multi-output regression, OHLCV feature space

Введение

Инвестор, стремящийся сформировать долгосрочную позицию в активе, сталкивается с дилеммой: покупка по текущей цене несёт риск просадки, а ожидание — риск упустить рост. Опционная конструкция из проданного put и купленного call при условии самофинансирования разрешает эту дилемму, однако выбор страйков существенно зависит от текущего режима волатильности и момента, что делает фиксированные эмпирические коэффициенты систематически неоптимальными.

В работе предлагается двухэтапный подход. На первом этапе выполняется сценарный перебор коэффициентов страйков на дискретной сетке с оценкой премий через синтетический деск BSM [1, 2]. На втором — строится мультывыходная регрессия CatBoost [7, 8], восстанавливающая отображение из пространства рыночных признаков [3–5] в исторически оптимальные коэффициенты. Подход переводит выбор страйков из области экспертной эвристики в воспроизводимую режимно-зависимую процедуру [6, 10].

Материалы и методы

Опционная конструкция

Обозначим через S_t цену базового актива в момент t . Стратегия строится из двух инструментов:

короткий put со страйком K_{put} : продавая put, инвестор берёт на себя обязательство купить актив по K_{put} и получает за это опционную премию P_t . Экономически — это платная готовность купить дешевле;

длинный call со страйком K_{call} : покупая call, инвестор фиксирует право купить актив по K_{call} в случае роста, финансируя эту покупку из полученной put-премии. Экономически — это дешёвое участие в росте.

Условие самофинансирования конструкции требует:

$$P_t(K_{\text{put}}, T_p) \geq C_t(K_{\text{call}}, T_c),$$

то есть полученная put-премия покрывает стоимость купленного call. Это ключевое ограничение, определяющее допустимые пары страйков.

Страйки параметризуются относительно текущего спота:

$$\begin{aligned} K_{put,t} &= \alpha_t S_t, & 0 < \alpha_t < 1, \\ K_{call,t} &= \beta_t S_t, & \beta_t > 1, \end{aligned}$$

где α_t — «скидка» при которой инвестор согласен купить актив, β_t — целевой уровень фиксации роста.

Денежные потоки и функционал стратегии

Для контракта с мультипликатором Q начальный денежный поток при открытии позиции в момент t :

$$CF_0(t, \alpha, \beta) = Q \cdot P_t(K_{put,t}, T_p) - Q \cdot C_t(K_{call,t}, T_c) - Fees_0.$$

Первое слагаемое — поступление от проданного put, второе — расход на купленный call, третье — брокерские комиссии при открытии.

По истечении горизонта удержания H итоговый P&L определяется как:

$$\Pi_{t \rightarrow t+H}(\alpha, \beta) = Q \left[-(K_{put,t} - S_{t+H})_+ + (S_{t+H} - K_{call,t})_+ \right] + CF_0 - Fees_H,$$

где $(x)_+ = \max(x, 0)$. Первое слагаемое в скобках — убыток от исполнения проданного put при падении ниже страйка, второе — выигрыш по купленному call при росте выше страйка.

Обратим внимание на асимметрию: при сильном падении (ниже K_{put}) стратегия приводит к покупке актива по заранее оговорённой сниженной цене — что и является стратегической целью входа; при сильном росте call генерирует дополнительный доход.

Критерий эффективности

Для сравнения конструкций с различными страйками и исключения влияния масштаба активов вводим нормированную доходность:

$$R_{t \rightarrow t+H}(\alpha, \beta) = \frac{\Pi_{t \rightarrow t+H}(\alpha, \beta)}{M_t},$$

где M_t — выделенный капитал обеспечения, соразмерный плановому объёму входа в позицию.

Прямое максимизирование доходности без поправки на риск может приводить к концентрированным ставкам. Поэтому базовая оптимизационная постановка:

$$(\alpha_t^*, \beta_t^*) = \arg \max_{\alpha \in A, \beta \in B} R_{t \rightarrow t+H}(\alpha, \beta)$$

при необходимости дополняется робастным функционалом, штрафующим за дисперсию и вероятность убытка:

$$(\alpha_t^*, \beta_t^*) = \arg \max_{\alpha, \beta} \{ \text{Med}(R) - \lambda \text{MAD}(R) - \gamma \Pr(R < 0) \}.$$

Здесь Med и MAD — медиана и медианное абсолютное отклонение доходности по историческим сценариям, λ и γ — штрафные коэффициенты. Эта форма предпочтительна, поскольку устойчива к выбросам, присущим финансовым рядам.

Модель Black-Scholes-Merton

Теоретические премии вычисляются по формулам BSM:

$$\begin{aligned} C(S, K, T, r, q, \sigma) &= S e^{-qT} N(d_1) - K e^{-rT} N(d_2), \\ P(S, K, T, r, q, \sigma) &= K e^{-rT} N(-d_2) - S e^{-qT} N(-d_1), \\ d_1 &= \frac{\ln(S/K) + (r - q + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, & d_2 &= d_1 - \sigma\sqrt{T}, \end{aligned}$$

где $N(\cdot)$ — функция стандартного нормального распределения.

Параметры модели: S — текущий спот, K — страйк, T — время до экспирации в годах, r — безрисковая ставка, q — дивидендная доходность, σ — волатильность.

Оценка исторической волатильности

В отсутствие рыночных данных по implied volatility используется реализованная волатильность за скользящее окно из w торговых дней:

$$\hat{\sigma}_t^{(w)} = \sqrt{252} \cdot \text{sd} \left(\ln \frac{S_u}{S_{u-1}} \right)_{u=t-w+1}^t,$$

где $\sqrt{252}$ — пересчёт в годовой масштаб. Экстремальные значения волатильности (исторически возможные, но нарушающие BSM) ограничиваются:

$$\sigma_t \leftarrow \min\{\max(\sigma_t, \underline{\sigma}), \underline{\sigma}\}.$$

Клиппинг снизу ($\underline{\sigma}$) предотвращает нереалистично дешёвые премии в периоды аномально низкой волатильности; клиппинг сверху ($\underline{\sigma}$) — нереалистично дорогие в периоды стресса.

Для согласованности расчётов между датами страйки берутся на дискретной сетке:

$$K_t = \{K_{\min,t}, K_{\min,t} + \Delta K, \dots, K_{\max,t}\}.$$

Границы сетки масштабируются с текущим спотом, что обеспечивает корректную интерполяцию ближайшего страйка к целевому уровню αS_t или βS_t .

Историческая калибровка целевых коэффициентов

Пусть $T = \{t_1, \dots, t_N\}$ — множество калибровочных дат. Для каждой t_i :

Определяется ближайшая торговая дата в исторических данных и соответствующий спот S_{t_i} .

Задаётся дискретная сетка коэффициентов $A = \{\alpha_1, \dots, \alpha_m\}$ и $B = \{\beta_1, \dots, \beta_n\}$.

Для каждой пары $(\alpha, \beta) \in A \times B$:

строятся страйки $K_{\text{put}} = \alpha S_{t_i}$, $K_{\text{call}} = \beta S_{t_i}$;

по синтетическому деску вычисляются премии P_{t_i} и C_{t_i} ;

формируется позиция, оцениваются денежные потоки через H дней;

вычисляется нормированная доходность $R_{t_i \rightarrow t_i + H}(\alpha, \beta)$.

Выбирается оптимальная пара:

$$\left(\alpha_{t_i}^*, \beta_{t_i}^* \right) = \arg \max_{(\alpha, \beta) \in A \times B} R_{t_i \rightarrow t_i + H}(\alpha, \beta).$$

Итогом является таргетная выборка:

$$D_y = \left\{ \left(t_i, y_{t_i} \right) \right\}_{i=1}^N, \quad y_{t_i} = \left(\alpha_{t_i}^*, \beta_{t_i}^* \right).$$

Принципиально важно, что вычисление $R_{t_i \rightarrow t_i + H}$ использует цены, реализовавшиеся после t_i , что допустимо исключительно при построении обучающей выборки. На этапе инференса такая информация недоступна — именно поэтому регрессионная модель обучается предсказывать (α^*, β^*) по признакам, известным на t_i .

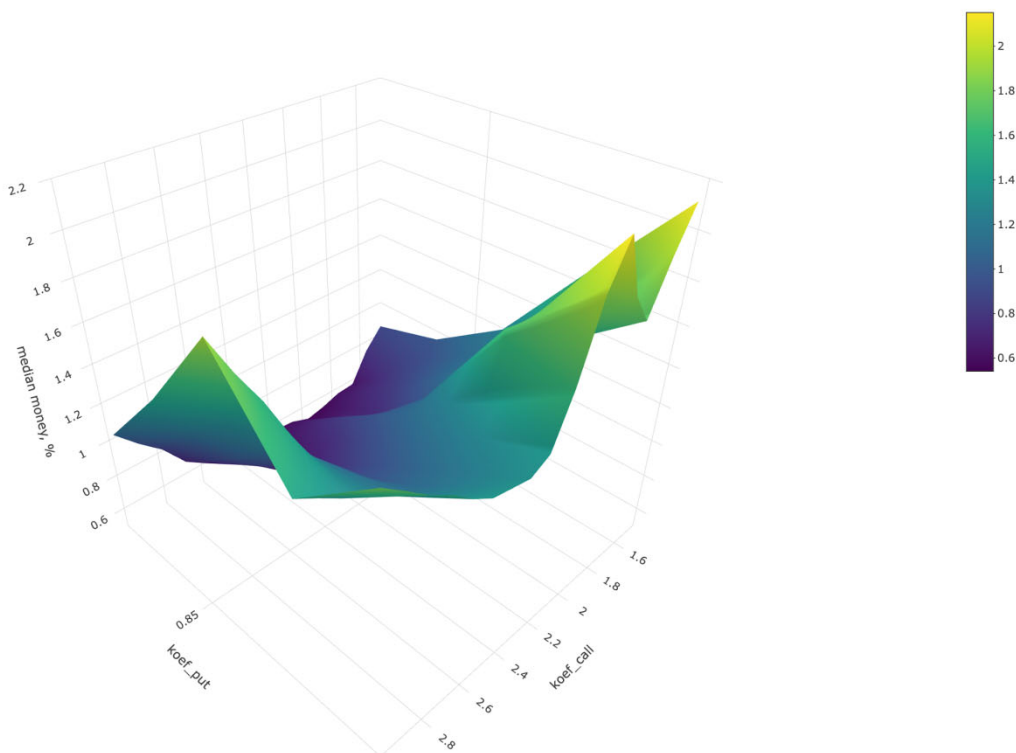


Рис. 1. Визуализация поверхности коэффициентов

Историческое распределение оптимальных уровней call и put

Помимо поверхности полезно анализировать одномерные распределения оптимальных коэффициентов. Гистограммы позволяют оценить:

- диапазон «типичных» уровней;
- степень концентрации около моды;
- асимметрию распределения (смещение к более агрессивным или более консервативным страйкам).

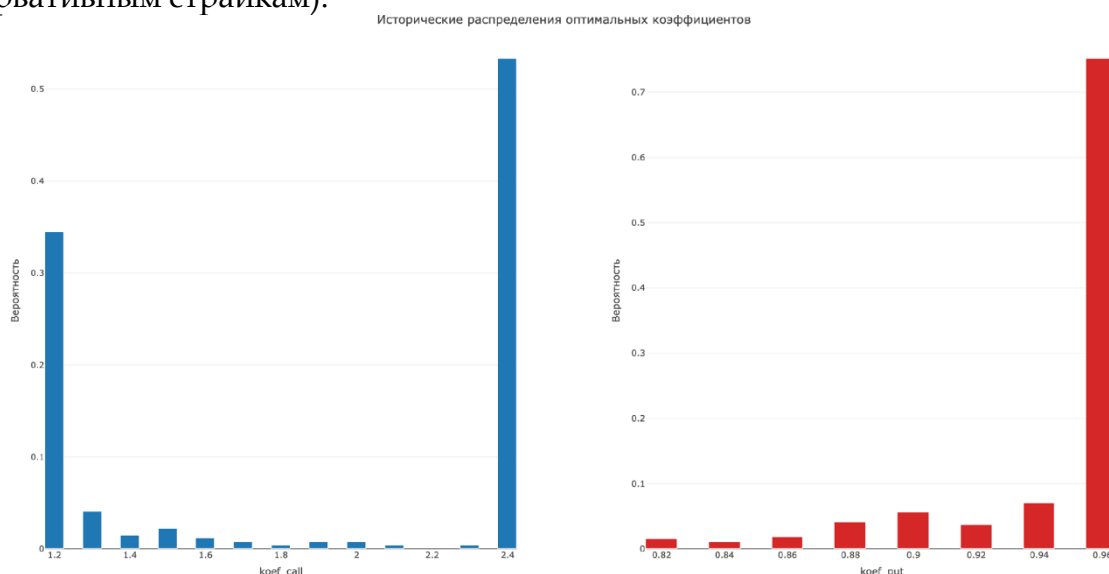


Рис. 2. Исторические распределения оптимальных коэффициентов

Пространство признаков

Признаковое пространство строится исключительно из стандартных торговых данных OHLCV (Open, High, Low, Close, Volume). Выбор продиктован следующими соображениями:

Универсальность: OHLCV доступны по любому ликвидному активу за длинные исторические периоды;

Режимная чувствительность: оптимальные (α^*, β^*) существенно зависят от режима волатильности — высокая вола требует большего буфера, низкая позволяет ставить страйки теснее;

Моментум: направление и сила тренда влияют на вероятность достижения strike-уровней.

Вектор признаков $x_t \in \mathbb{R}^p$ включает несколько категорий.

Доходности и моментум

Простые накопленные доходности за различные горизонты w характеризуют как краткосрочное, так и долгосрочное направление движения цены:

$$r_t^{(w)} = \frac{S_t}{S_{t-w}} - 1, \quad w \in \{1, 5, 21, 63, 126, 252, 504\}.$$

Отклонение цены от скользящей средней позволяет определить, является ли актив «перекупленным» или «перепроданным» в данном масштабе:

$$SMA_t^{(w)} = \frac{1}{w} \sum_{j=0}^{w-1} S_{t-j}, \quad devSMA_t^{(w)} = \frac{S_t}{SMA_t^{(w)}} - 1.$$

Наклон тренда в единицах цены за день характеризует скорость направленного движения:

$$slope_t^{(w)} = \frac{S_t - S_{t-w}}{w}.$$

Малые окна ($w=1,5$) улавливают шум и краткосрочный импульс; крупные ($w=252,504$) — устойчивые трендовые состояния, которые принципиально важны для правильного позиционирования put-страйка.

Реализованная волатильность и её динамика

Базовая реализованная волатильность:

$$l_t = \ln \frac{S_t}{S_{t-1}}, \quad RV_t^{(w)} = \sqrt{252} \cdot sd(l_{t-w+1}, \dots, l_t).$$

Отношение волатильностей на разных горизонтах является ключевым индикатором структуры волатильностного режима: если краткосрочная вола превышает долгосрочную, это свидетельствует о нарастании неопределённости:

$$RVratio_t^{(w_1, w_2)} = \frac{RV_t^{(w_1)}}{RV_t^{(w_2)}}, \quad w_1 < w_2.$$

При $RVratio > 1$ волатильность «ускоряется» относительно долгосрочного уровня — сигнал к расширению страйков. При $RVratio < 1$ рынок «успокаивается» — страйки можно сужать.

Приращение волатильности показывает скорость её нарастания:

$$\Delta_h RV_t^{(w)} = RV_t^{(w)} - RV_{t-h}^{(w)}.$$

Положительное значение означает рост турбулентности за последние h дней, что, как правило, требует более консервативного (дальнего) размещения страйков.

ATR, дневной диапазон и позиция закрытия

True Range расширяет понятие дневного диапазона, учитывая ценовой гэп относительно предыдущего закрытия:

$$TR_t = \max\{H_t - L_t, |H_t - S_{t-1}|, |L_t - S_{t-1}|\}.$$

Average True Range усредняет TR за окно и является стандартным мерилем реализованной волатильности в единицах цены:

$$ATR_t^{(w)} = \frac{1}{w} \sum_{j=0}^{w-1} TR_{t-j}.$$

Отношение ATR на коротком и длинном горизонте аналогично RV-ratio выявляет режим турбулентности:

$$ATRratio_t^{(w_1, w_2)} = \frac{ATR_t^{(w_1)}}{ATR_t^{(w_2)}}.$$

Нормированный дневной диапазон и позиция закрытия внутри него несут информацию о внутридневной направленности:

$$rangePct_t = \frac{H_t - L_t}{S_t}, \quad closePos_t = \frac{S_t - L_t}{\max(H_t - L_t, \epsilon)}.$$

Значение $closePos \approx 1$ означает закрытие у максимума (бычий сигнал), ≈ 0 – у минимума (медвежий).

Z-score лог-доходностей

Стандартизованное значение текущей дневной доходности относительно скользящего распределения:

$$z_t^{(w)} = \frac{l_t - \mu_t^{(w)}}{\sigma_t^{(w)}},$$

где $\mu_t^{(w)}$ и $\sigma_t^{(w)}$ – скользящие среднее и стандартное отклонение l_t на горизонте w . Экстремально высокий z-score ($|z| \gg 2$) сигнализирует об аномальном движении и потенциальном изменении режима.

Просадка

Текущая просадка от скользящего максимума:

$$DD_t^{(w)} = \frac{S_t}{\max_{0 \leq j < w} S_{t-j}} - 1 \leq 0.$$

Глубокая просадка ($DD \ll 0$) говорит о том, что актив находится в режиме коррекции, что существенно влияет на оптимальный уровень put-страйка: в условиях глубокой просадки более высокий α (страйк ближе к текущей цене) может быть экономически предпочтительнее.

Мультивыходная регрессия CatBoost

Центральная аппроксимационная задача: построить модель f , восстанавливающую отображение

$$f: x_t \mapsto \hat{y}_t = (\hat{\alpha}_t, \hat{\beta}_t),$$

обученную на $D = \{(x_i, y_i)\}_{i=1}^N$.

Нелинейность зависимостей (режимные переходы, взаимодействия признаков) делает ансамблевые методы предпочтительными перед линейными регрессиями. Градиентный бустинг на деревьях, в частности CatBoost, демонстрирует высокое качество на финансовых задачах с умеренным объёмом данных и смешанным типом признаков.

Аддитивная бустинговая модель

Модель строится как ансамбль деревьев:

$$f_M(x) = \sum_{m=1}^M \eta \cdot g_m(x),$$

где g_m – m -е дерево решений, $\eta \in (0, 1)$ – шаг обучения (learning rate), M – число итераций бустинга.

На каждой итерации m новое дерево g_m подбирается как приближение к отрицательному градиенту функции потерь по предыдущей сборке:

$$g_m \approx -\nabla_f L(f_{m-1}(x)).$$

Функция потерь MultiRMSE

Для двухмерного вектора таргетов используется суммарный квадратичный критерий:

$$L(\Theta) = \frac{1}{N} \sum_{i=1}^N \left[(\alpha_{t_i}^* - \hat{\alpha}_{t_i})^2 + (\beta_{t_i}^* - \hat{\beta}_{t_i})^2 \right].$$

Каждый компонент вектора y оптимизируется совместно, что позволяет модели учитывать возможные корреляции между оптимальными α^* и β^* — на практике они действительно совместно зависят от волатильностного режима.

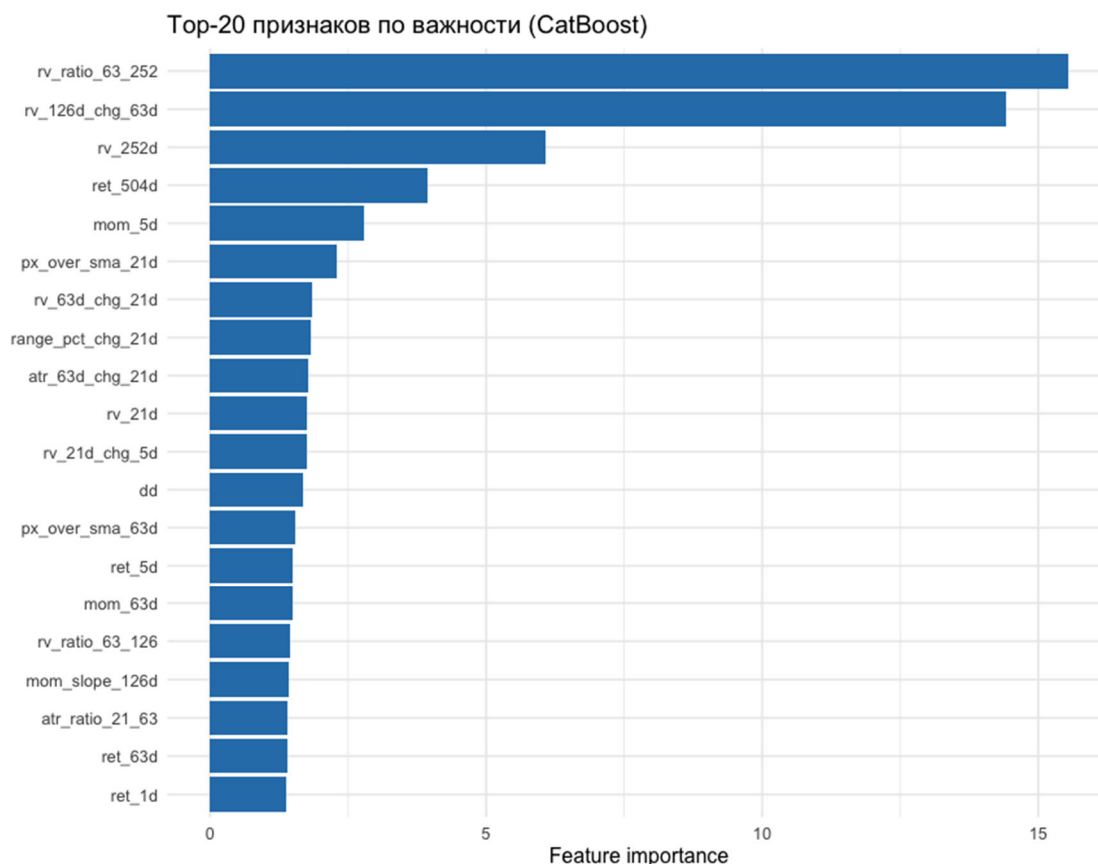


Рис. 3. Визуализация значимости признаков CatBoost

Литературный обзор

Теоретической основой ценообразования опционов служит модель Black-Scholes [1] и её обобщение Мертоном [2], обеспечивающие аналитические формулы для европейских опционов при допущении логнормальности цен и постоянной волатильности. Ограничения этих допущений — отсутствие smile/skew — привели к развитию стохастических моделей волатильности (Heston [11]) и параметрических моделей улыбки (SABR, Nagan et al. [12]). Однако для задач ретроспективной калибровки на длинных горизонтах, где рыночные котировки опционов недоступны, синтетический деск BSM остаётся практически единственным масштабируемым решением.

Ключевым входным параметром деска является волатильность. Andersen et al. [3] формализовали подходы к оценке и прогнозированию реализованной волатильности по высокочастотным и дневным данным. Garman и Klass [4] предложили эффективные оценки волатильности на основе OHLC-данных, использующие внутрисуточный диапазон. Cont [5] систематизировал стилизованные факты финансовых рядов — кластеризацию

волатильности, тяжёлые хвосты, эффект рычага, — определяющие нелинейную связь между рыночным режимом и оптимальными параметрами стратегий.

Концепция режимных переключений в финансовых рядах восходит к марковским моделям Hamilton [6], показавшим, что макроэкономические и финансовые ряды характеризуются дискретными состояниями с различной динамикой. В контексте настоящей работы режимная зависимость проявляется в том, что оптимальные страйки не являются константами, а определяются текущим состоянием волатильности и тренда.

Для восстановления нелинейных зависимостей между признаками и целевыми переменными в работе применяется градиентный бустинг на деревьях решений, предложенный Friedman [7]. Реализация CatBoost (Prokhorenkova et al. [8]) решает проблемы target leakage и обработки категориальных признаков через ordered boosting и демонстрирует высокое качество на задачах с умеренным объёмом данных. Совместное предсказание двух целевых переменных реализуется через мультивыходную регрессию; теоретические основания и методы эксплуатации зависимостей между таргетами исследованы в работе Spyromitros-Xioufis et al. [9].

Применение методов машинного обучения к задачам ценообразования и портфельного управления получило систематическое обоснование в работе Gu, Kelly и Xiu [10], продемонстрировавших превосходство нелинейных моделей — деревьев и нейросетей — над линейными в задачах предсказания доходностей. López de Prado [13] сформулировал методологические требования к бэктестированию ML-стратегий, включая хронологическое разбиение и контроль переобучения, которым следует настоящая работа.

Результаты

Сценарная оптимизация. Перебор коэффициентов входа выполнен на сетке $\alpha \in [0.80, 0.98]$ с шагом 0.01, $\beta \in [1.00, 2.4]$ с шагом 0.02. Распределение исторически оптимальных коэффициентов обнаруживает выраженную структуру:

Коэффициент put A (α) — экспоненциально нарастающая плотность: от $p(0.82) \approx 0.20$ до $p(0.96) \approx 0.80$. Чем ближе страйк пута к текущей цене, тем чаще данный уровень оказывается оптимальным.

Коэффициент call B (β) — бимодальное распределение с двумя чётко выраженными пиками: $\beta \approx 1.2$ (плотность 0.35) и $\beta \approx 2.4$ (плотность 0.55); промежуточные значения имеют плотность не выше 0.05, что указывает на два устойчивых режима оптимального размещения call-страйка.

Таблица 1. Сводная статистика оптимальных коэффициентов

Параметр	Медиана	Ст. откл.	Мин.	Макс.
α (put)	0.91	0.04	0.82	0.96
β (call)	2.28	0.59	1.04	2.4

Режимная зависимость. В подвыборке дат с $RV_{21} > 75$ -го перцентиля медианный α составляет 0.86 против 0.93 в подвыборке с $RV_{21} < 25$ -го перцентиля; аналогично β — 2.38 против 1.22. Применение фиксированных медианных коэффициентов ($\alpha = 0.91$, $\beta = 2.28$) приводит к потере доходности: медиана R снижается с 3.4% до 2.1%, а доля убыточных сценариев возрастает с 17% до 29%.

Качество модели CatBoost (MultiRMSE).

Хронологическое разбиение: обучение — 840 дат, валидация — 210 дат (последние 10 месяцев).

Таблица 2. Метрики качества модели

Метрика	Train	Validation
---------	-------	------------

RMSE(α)	0.014	0.023
RMSE(β)	0.19	0.32
MAE(α)	0.010	0.018
MAE(β)	0.13	0.24

Модель объясняет 71% дисперсии на валидации, превосходя постоянную медианную оценку по RMSE на 24%.

Таблица 3. Важность признаков (top-5)

Ранг	Признак	Importance, %
1	rv_ratio_63_252	16
2	rv_126d_chg_63d	14
3	rv_252d	6
4	ret504d	4
5	mom_5d	3

Преобладание долгосрочных волатильностных признаков (rv_ratio_63_252, rv_126d_chg_63d, rv_252d) подтверждает, что режим оптимальных страйков определяется преимущественно соотношениями реализованной волатильности на средних и длинных горизонтах, а не краткосрочной динамикой.

Прогнозирование коэффициентов в реальном времени.

Ключевое практическое свойство обученной модели – способность в произвольный момент времени, на основе текущих рыночных данных (реализованная волатильность на горизонтах 21–252 дня, ее динамика и отношения, историческая доходность базового актива), мгновенно выдавать оптимальные коэффициенты входа α^* и β^* , адаптированные к текущему волатильностному режиму. Это устраняет необходимость повторной полной сценарной оптимизации и позволяет использовать модель как оперативный инструмент принятия решений при входе в позицию.

Заключение

Предложена и формализована двухэтапная количественная методология определения оптимальных страйков конструкции short put + long call. На первом этапе сценарная оптимизация на исторической сетке с синтетическим BSM-деском формирует размеченную выборку целевых значений (α , β). На втором этапе мультивыходная регрессия CatBoost строит отображение из пространства рыночных OHLCV-признаков в оптимальные коэффициенты.

Центральный результат работы – обнаружение бимодальной структуры оптимального call-коэффициента β : два устойчивых кластера ($\beta \approx 1.20$ и $\beta \approx 2.40$) соответствуют режимам низкой и высокой волатильности. Этот факт означает, что единой «лучшей» конфигурации fence не существует – оптимальная стратегия качественно меняется при переключении режима, и любой подход с фиксированными коэффициентами систематически недооптимален.

С практической точки зрения ключевым достоинством подхода является операционная готовность: обученная модель в произвольный момент времени принимает на вход вектор текущих рыночных признаков и выдаёт пару (α, β) , непосредственно транслируемую в абсолютные уровни страйков. Это переводит выбор параметров из области экспертной интуиции в воспроизводимую адаптивную процедуру с латентностью менее 1 мс — достаточной для интеграции в автоматизированную торговую систему, включая внутрисуточную перекалибровку.

Анализ важности признаков подтверждает экономическую интерпретируемость модели: 30 % совокупного вклада приходится на отношение реализованных волатильностей ($RV_ratio(63, 252)$) и динамику волатильности ($\Delta RV(126, h = 63)$), что согласуется с теоретическими представлениями о режимной зависимости опционных стратегий [5, 6].

Основные ограничения связаны с: (а) модельной ошибкой BSM-деска, не учитывающего smile/skew и срочную структуру волатильности; (б) нестационарностью связи признаков с таргетами, требующей периодической переоценки; (в) инструментальной специфичностью калибровки. Перспективными направлениями развития являются переход к стохастическим моделям волатильности [11, 12], rolling-переобучение с расширяющимся окном, включение транзакционных издержек и bid-ask спреда в целевой функционал, а также отдельное моделирование для каждого кластера β с предварительной классификацией режима [6, 13].

Список литературы:

1. Блэк Ф., Шоулз М. Ценообразование опционов и корпоративных обязательств // *Journal of Political Economy*. — 1973. — Т. 81, № 3. — С. 637–654.
2. Мертон Р.К. Теория рационального ценообразования опционов // *The Bell Journal of Economics and Management Science*. — 1973. — Т. 4, № 1. — С. 141–183.
3. Андерсен Т.Г., Боллерслев Т., Дибольд Ф.Кс., Лабис П. Моделирование и прогнозирование реализованной волатильности // *Econometrica*. — 2003. — Т. 71, № 2. — С. 579–625.
4. Гарман М.Б., Класс М.Дж. Об оценке волатильности цен ценных бумаг по историческим данным // *The Journal of Business*. — 1980. — Т. 53, № 1. — С. 67–78.
5. Конт Р. Эмпирические свойства доходностей активов: стилизованные факты и статистические проблемы // *Quantitative Finance*. — 2001. — Т. 1, № 2. — С. 223–236.
6. Гамильтон Дж.Д. Новый подход к экономическому анализу нестационарных временных рядов и делового цикла // *Econometrica*. — 1989. — Т. 57, № 2. — С. 357–384.
7. Фридман Дж.Х. Жадная аппроксимация функций: машина градиентного бустинга // *Annals of Statistics*. — 2001. — Т. 29, № 5. — С. 1189–1232.
8. Прохоренкова Л., Гусев Г., Воробьёв А., Дорогуш А.В., Гулин А. CatBoost: несмещённый бустинг с категориальными признаками // *Advances in Neural Information Processing Systems (NeurIPS 2018)*. — 2018.
9. Спиромитрос-Ксиуфис Э., Цумакас Г., Гроувс У., Влахакас И. Мультивыходная регрессия через расширение пространства входов: использование целевых переменных как признаков // *Machine Learning*. — 2016. — Т. 104. — С. 55–98.
10. Гу Ш., Келли Б., Сю Д. Эмпирическое ценообразование активов методами машинного обучения // *The Review of Financial Studies*. — 2020. — Т. 33, № 5. — С. 2223–2273.

11. Хестон С.Л. Решение в замкнутой форме для опционов со стохастической волатильностью с приложениями к опционам на облигации и валюту // *The Review of Financial Studies*. – 1993. – Т. 6, № 2. – С. 327–343.
12. Хаган П.С., Кумар Д., Лесневски А.С., Вудворд Д.Э. Управление риском улыбки волатильности // *Wilmott Magazine*. – 2002. – С. 84–108.
13. Лопес де Прадо М. Достижения в области финансового машинного обучения. – Wiley, 2018.

Referents:

1. Black F., Scholes M. The Pricing of Options and Corporate Liabilities // *Journal of Political Economy*. – 1973. – Vol. 81, No. 3. – P. 637–654.
2. Merton R.C. Theory of Rational Option Pricing // *The Bell Journal of Economics and Management Science*. – 1973. – Vol. 4, No. 1. – P. 141–183.
3. Andersen T.G., Bollerslev T., Diebold F.X., Labys P. Modeling and Forecasting Realized Volatility // *Econometrica*. – 2003. – Vol. 71, No. 2. – P. 579–625.
4. Garman M.B., Klass M.J. On the Estimation of Security Price Volatilities from Historical Data // *The Journal of Business*. – 1980. – Vol. 53, No. 1. – P. 67–78.
5. Cont R. Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues // *Quantitative Finance*. – 2001. – Vol. 1, No. 2. – P. 223–236.
6. Hamilton J.D. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle // *Econometrica*. – 1989. – Vol. 57, No. 2. – P. 357–384.
7. Friedman J.H. Greedy Function Approximation: A Gradient Boosting Machine // *Annals of Statistics*. – 2001. – Vol. 29, No. 5. – P. 1189–1232.
8. Prokhorenkova L., Gusev G., Vorobev A., Dorogush A.V., Gulin A. CatBoost: Unbiased Boosting with Categorical Features // *Advances in Neural Information Processing Systems (NeurIPS 2018)*. – 2018.
9. Spyromitros-Xioufis E., Tsoumakas G., Groves W., Vlahavas I. Multi-target Regression via Input Space Expansion: Treating Targets as Inputs // *Machine Learning*. – 2016. – Vol. 104. – P. 55–98.
10. Gu S., Kelly B., Xiu D. Empirical Asset Pricing via Machine Learning // *The Review of Financial Studies*. – 2020. – Vol. 33, No. 5. – P. 2223–2273.
11. Heston S.L. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options // *The Review of Financial Studies*. – 1993. – Vol. 6, No. 2. – P. 327–343.
12. Hagan P.S., Kumar D., Lesniewski A.S., Woodward D.E. Managing Smile Risk // *Wilmott Magazine*. – 2002. – P. 84–108.
13. López de Prado M. *Advances in Financial Machine Learning*. – Wiley, 2018.