Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра теории функций
и стохастического анализа

ЖАДНЫЕ АЛГОРИТМЫ ДЛЯ РЕШЕНИЯ ЗАДАЧ ОПТИМИЗАЦИИ

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студентки 4 курса 412 группы
направления 01.03.02 Прикладная математика и информатика
Кондратьевой Елизаветы Вадимовны

Научный руководитель
д.ф-м.н, доцент ________________________________ С.П. Сидоров

Зав. кафедрой
д.ф-м.н, доцент ________________________________ С.П. Сидоров

Саратов, 2018

ВВЕДЕНИЕ

Актуальность темы. Финансовый рынок – это чрезвычайно сложная структура со множеством участников и финансовых посредников, оперирующих с разнообразными финансовыми инструментами и выполняющих широкий набор функций по обслуживанию и управлению финансовыми процессами. В последние годы в нашей стране в связи с развитием рыночной экономики существенно повысился интерес к постановке и решению задач теории инвестиций. Среди этих задач значительное место занимают задачи оптимизации портфелей активов.

Выбирая различные варианты распределения капитала между объектами, в которые инвестируется капитал, мы получим различные результаты, если под результатом понимать величину дохода, полученного в течение заранее определенного периода. Очевидно, оптимальное распределение инвестируемого капитала должно обеспечивать в некотором смысле наилучший результат. В то же время, решение о структуре распределения капитала принимается часто в условиях неопределенности, когда доходность от вложения капитала в объекты инвестирования носит случайный характер. Тем самым появляется риск вложения капитала и задача оптимизации портфеля инвестиций должна ставиться и решаться в условиях наличия риска.

Целью бакалаврской работы является составление оптимального портфеля ценных бумаг, используя стратегию слежения за индексом (репликация индекса). Это пассивная финансовая стратегия, которая состоит в имитации (репликации) доходности заданного индекса или портфеля. Цель инвестора – найти веса активов в своем портфеле, чтобы получившийся портфель имел минимальную ошибку слежения, в качестве которой обычно используют дисперсию разности между доходностью индекса и доходностью портфеля. В данной работе решение проблемы слежения за индексом рассматривается с ограничением на кардинальность, т.е. с ограничением на максимальное количество активов, удерживаемых в портфеле. В данной работе рассматривается жадный алгоритм для решения данной задачи и изучаются свойства сходимости жадных алгоритмов для решения некоторых задач оптимизации.

Объект исследования — составление оптимального портфеля ценных бумаг с ограничением на максимальное количество активов.

Предмет исследования — финансовый рынок ценных бумаг и свойства сходимости жадных алгоритмов для решения задач оптимизации, в частности задачи слежения за индексом.

Для достижения поставленных целей в работе необходимо решить следующие **задачи**:

— определить основные понятия, необходимые для работы с портфелем ценных бумаг;

— рассмотреть теорию управления портфелем Марковица;

— рассмотреть понятие бинарной классификации и задачи оптимизации;

— рассмотреть неразделимые данные, проблемы оптимизации с экспоненциальной функции потерь и логистической функции потерь;

— рассмотреть нахождение разреженных решений и алгоритмы оптимизации для решения задач бинарной классификации;

— описать алгоритм жадного координатного спуска и его свойства;

— анализ и сложность метода жадного координатного спуска;

— рассмотреть метод Франка-Вулфа для условной оптимизации;

— рассмотреть метод минимизации Франка-Вулфа;

— анализ и сложность метода Франка-Вулфа;

— рассчитать основные практические и теоретические характеристики данной системы.

Практическая значимость проводимого исследования состоит в составление оптимального портфеля ценных бумаг, используя стратегию слежения за индексом, решение проблемы рассматривается с ограничением на кардинальность. По результатам этих вычислений можно делать выводы о рентабельности и доходности заданного индекса или портфеля.

Структура и содержание бакалаврской работы. Работа состоит из введения, четырех разделов, заключения, списка использованных источников и приложения.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обосновывается актуальность темы работы, формулируется цель работы и решаемые задачи, отмечается практическая значимость полученных результатов.

В **первом** разделе приводятся основные понятия портфеля ценных бумаг, теория управления портфелем Марковица.

Во **втором** разделе рассмотрены основные понятия бинарной классификации и задачи оптимизации, понятия о неразделимых данных, проблемы оптимизации с экспоненциальной функции потерь и логистическая функция потерь. Предположим, что данные задачи неразделимы, как это бывает в большинстве практических случаев. Рассмотрим следующую задачу оптимизации:

$$LELP: \quad L^*_{exp} := mimize_{\lambda} L_{exp}(\lambda) := ln(\frac{1}{m}) \sum_{i=1}^{m} exp((-A\lambda)_i), \quad \lambda \in R^n.$$

$L_{exp}(\lambda)$ есть целевая функция задачи LELP и называется функцией потерь. Эта функция является убывающей и монотонной функцией $(A\lambda)_i$ при $i = 1, \dots, m$, поэтому при сведении к минимуму эта функция будет давать большие значения $(A\lambda)_i$ для $i = 1, \dots, m$.

Еще одна функция потерь, которая часто используется – логистическая функция потерь $L_{logit}(\cdot) R^n \to R$, которая определяется по формуле:

$$L_{logit}(\lambda) := \frac{1}{m} \sum_{i=1}^{m} \ln(1 + exp((-A\lambda)_i)).$$

Она имеет очень привлекательные статистические свойства в дополнение к некоторым свойствам оптимизации. Эта функция также является монотонно убывающей функцией $(A\lambda)_i$ для $i = 1, \dots m$.

Логистической задачей оптимизации потерь является минимизация логистической функции потерь $L_{logit}(\dots)$, которая записывается следующим образом:

$$L^*_{logit} := min_{\lambda} L_{logit}(\lambda) := \frac{1}{m} \sum_{i=1}^{m} \ln(1 + exp((-A\lambda)_i)), \quad \lambda \in R^n.$$

Вышеуказанные задачи оптимизации являются иллюстративными задачами оптимизации, которые возникают достаточно естественно в машинном обучении. Пусть λ^* является оптимальным (или близким к оптимальному) решением одной из проблем оптимизации MMP, LELP. Пусть $\| \lambda^* \|_0$ обозначает число ненулевых компонент в векторе λ^*. Во многих случаях важно, чтобы $\| \lambda^* \|_0$ было небольшим, то есть λ^* являлось бы разреженным вектором (т.е. λ^* имеет очень мало ненулевых компонент). Это полезно при управлении вычислениями и памятью, особенно при $n >> 0$, при задачах большой размерности. Но в большинстве случаев разреженность λ^* возникает из того, что решение зависит от очень небольшого количества основных функций (компонентов) и т.д. В присутствии потенциального переобученного решения на основе данных, опыт работы во многих контекстах показал, что разреженные решения, как правило, являются более точными при прогнозировании на вневыборочных данных. Поэтому в дополнение к оптимизации функций потерь в MMP, LELP, мы также ищем решение, которое является разреженным в том смысле, что $\|\lambda^*\|_0$ относительно мало. Мы могли бы попытаться обеспечить соблюдение требований разреженности, добавив ограничение типа "$\|\lambda^*\|_0 \leq r$"к задаче оптимизации, для некоторых малых значений. Тем не менее, этот прямой подход к решению вопроса разреженности делает задачу очень сложной – на самом деле такая задача является NP-полной в теоретической информатике. Вместо этого, мы видим, что некоторые алгоритмы, которые обычно называют «жадными», по своей конструкции способствуют или гарантируют получение разреженных решений.

В этом разделе рассмотриваются два основных типа «жадных» алгоритмов выпуклой оптимизации, а именно:

1. Жадный координатный спуск. Методы координатного спуска являются методами решения задач выпуклой оптимизации, которые используют одну координату на каждой итерации. Эти методы были впервые разработаны на заре развития теории оптимизации в качестве одного из способов проведения вычислений, но с малой нагрузкой на память. Сегодня такие методы испытывают огромный всплеск интереса из-за их способности находить разреженных решения, то есть решения, которые используют очень мало координат;

2. Метод Франка-Вульфа. Метод Франка-Вульфа создан Маргаритой Франк и Филиппом Вульф, также известен как метод «условного градиента».

Третий раздел посвящен принципам построения жадного координатного спуска, его алгоритму и свойствам.

Наша задача выглядит следующим образом:

$$P : minimize_x f(x), x \in R^n, \tag{1}$$

где $f(\cdot)$ является дифференцируемой выпуклой функцией, определенной на R^n. Обозначим f^* – оптимальное значение целевой функции задачи P, и пусть x^* обозначает оптимальное решение P, если такое решение существует. Пусть $\nabla f(x)$ обозначает градиент $f(\cdot)$ в точке x, и напомним, что неравенство для градиента имеет вид:

$$f(y) \geq f(x) + \nabla f(x)^T (y - x), (x, y \in R^n). \tag{2}$$

Мы измеряем расстояния между точками, используя норму l_1:

$$\|x\|_1 := \sum_{j=1}^{n} \mid x_j \mid .$$

Мы будем измерять величину градиентов в норме l_∞, где норма $\|g\|_\infty = max_j| \ g_j \ |$.

Будем считать, что градиент $f(\cdot)$ Липшицевым. Учитывая наше применение для вышеуказанных норм, это означает, что существует число L, для которого

$$\|\nabla f(y) - \nabla f(x)\|_\infty \leq L\|y - x\|_1, \ x, y \in R^n. \tag{3}$$

Пусть $B_1(c, R)$ обозначает шар по норме l_1 с центром в точке и радиусом R, а именно:

$$B_1(c, R) := x \in R^n : \|x - c\|_1 \leq R.$$

Пусть e^i обозначает $i^t h$ единичный вектор в R^n, а именно $e^i = (0 \ldots 0, 1, 0 \ldots 0)^T$, где появляется 1 в i-ом коэффициенте.

Алгоритм 1 представляет собой жадный метод координатного спуска для решения задачи (P). Предположим, что алгоритм 1 начинает работу в точке $x^0 \leftarrow 0$. Тогда из-за правила обновления, он будет удовлетворять

$$\|x^{k+1}\|_0 \leq \|x^k\|_0 + 1.$$

Эта гарантия разреженности полезна, если решается проблема большой размерности, где $n >> 0$ может иметь порядок 10^6 или даже больше.

Для решения задачи мы используем жадные алгоритмы. Выбор жадного алгоритма для нашего анализа основан на том факте, что жадные алгоритмы показали высокую эффективность при решении прикладных задач. Мы можем предположить, что они являются многообещающими и для решения задачи слежения за индексом. С другой стороны, жадные алгоритмы не обязательно дают оптимальное решение. В данной случае мы используем такие методы для решения задачи репликации индекса.

Алгоритм на каждом шаге своей работы добавляет в портфель актив, который еще не входит в портфель и который наиболее близок к индексу. Процесс включения новых активов в портфель продолжается до тех пор, пока в портфеле не окажется ровно K активов.

Жадный алгоритм. Определение параметров. Инициализировать в точке $x^0, i \leftarrow 0$ На i шаге :

1. Вычислить градиент. Вычислить$\nabla f(x^i)$.

2. Вычисление координат. Рассчитать:

   $j_i \in argmax_{j \in [1,n]} | \nabla f(x^k)_j |$;

   $s_i \leftarrow sgn(\nabla f(x^i)_j)$.

3. Подсчитать размер шага и новой точки:

   $a_i \leftarrow \frac{|\nabla f(x^i)_j|}{L}$;

   $x^{i+1} \leftarrow x^i \leftarrow a_i s_i e^j$.

Также в дипломной работе рассмотрены анализ и сложность метода жадного координатного спуска, а также изучен метод минимизации Франка-Вулфа.

Этот метод основан на том, что множество S удобно для линейной оптимизации. Это означает, что либо S является системой линейных неравенств $S = \{x | Ax \leq b\}$ или, что проблема:

$$LO_c : minimize_x C^T x, X \in S,$$

легко решается для любого заданного целевого вектора c. Предположим, что алгоритм находится на некоторой итерации в точке $x_k \in S$. Функция $f(x)$ в точке $x = x^k$ равна

$$z_1(x) := f(x^k) + \nabla f(x^k)^T(x - x^k).$$

Это разложение в ряд Тейлора первого порядка функции $f(\cdot)$ в точке x^k. Поскольку мы можем легко сделать линейную оптимизацию на S, решаем задачу

$$LP : minimize_x f(x^k) + \nabla f(x^k)^T(x - x^k), X \in S,$$

которая эквивалентна задаче

$$LP : minimize_x \nabla f(x^k)^T x, X \in S.$$

Пусть x^k обозначает оптимальное решение этой задачи. Продолжаем метод Франка-Вулфа, выбирая следующую итерацию как $x^{k+1} \leftarrow x^k + a(\tilde{x^k} - x^k)$ для некоторого шага. Так как S – выпуклое множество и оба x^k и $\tilde{x^k}$ содержатся в S, то $x^k + a(\tilde{x^k} - x^k) \in S$ для всех $a \in [0, 1]$. Пусть $\tilde{a_k} \in [0, 1]$ обозначает размер шага на итерации k-го метода. Имеем

$$x^{k+1} \leftarrow x^k + a(\tilde{x^k} - x^k), \tilde{a_k} \in [0, 1].$$

Значение $\tilde{a_k}$ можно определить несколькими способами. Один из способов состоит в том, чтобы установить размер шага $\tilde{a_k}$ так, чтобы он соответствовал некоторому заранее определенному правилу, например следующему:

$$\tilde{a_k} = \frac{2}{k + 2}.$$

Еще один способ – это выбрать $\tilde{a_k}$, выполнив поиск $f(\cdot)$ по интервалу $a \in [0, 1]$. То есть мы могли бы определить $\tilde{a_k}$ как оптимальное решение следующей задачи:

$$\tilde{a_k} \leftarrow argmin_a f(x^k + a(\tilde{x^k} - x^k)), 0 \leq a \leq 1.$$

Алгоритм 2 представляет собой формальное описание только что описанного метода Фрэнка-Вулфа.

Посмотрим теперь, как этот метод может быть использован для создания разреженных решений в контексте определенных структурированных задач. Предположим, что наша задача представляет собой проблему условной минимизации:

$$R_r : minimize_x f(x), \|x\|_1 \leq r,$$

где $f(\cdot)$ – дифференцируемая функция потерь, связанная с проблемой двоичной классификации машинного обучения, например:

$$L_{exp}(x) := ln(\frac{1}{m}\sum_{i=1}^{m} exp((-Ax)_i)).$$

Алгоритм метода минимизации Франка-Вулфа.

1. Вычислить градиент. Вычислить $\nabla f(x^k)$, затем решить задачу линейной оптимизации: $\tilde{x^k} \leftarrow argmin_{x \in S} f(x^k) + \nabla f(x^k)^T (x - x^k)$.
2. Находим $x^{k+1} \leftarrow x^k + \tilde{a_k}(\tilde{x^k} - x^k), \tilde{a_k} \in [0, 1]$.

Также представлены два важных предварительных результата. Первый результат гласит о том, что на каждой k-ой итерации метода Франка-Вулфа решение задачи линейной оптимизации приводит к k действительной нижней оценке оптимального значения f^* задачи оптимизации P.

Утверждение. На k итерации метода Франк-Вольфа выполняется

$$f^* \geq f(x^k) + \nabla f(x^k)^T (\tilde{x^k} - x^k).$$

Утверждение. Пусть существует константа $C > 0$, для которой неотрицательный ряд a_i такой что:

$$a_{i+1} \leq a_i(1 - \frac{2}{i+2}) + \frac{C}{(i+2)^2}, i \geq 0.$$

Тогда оно утверждает, что

$$a_k \leq \frac{C}{k+2}, k \geq 1.$$

В **четвертом** разделе написана программа на языке C++ для нахождения оптимального ивестиционного портфеля.

Также были собраны данные по бирже ценных бумаг, в работе использованы открытые данные, которые расположены в Huge Stock Market Dataset.

Подготовлена матрица цен, которая в поседствии была преобразованна в матрицу доходности.

В **заключении** приведены результаты бакалаврской работы.

Основные результаты

1. Определены основные понятия, необходимые для описания портфеля ценных бумаг. Изучены частные случаи управления портфелем Марковица.

2. Определены основные понятия, связанные с бинарной классификацией и задачей оптимизации. Рассмотрены проблемы оптимизации с экспоненциальной функции потерь и логистической функции потерь.

3. Изучены методы жадного координатного спуска и его свойства, а также метод Франка-Вулфа для условной оптимизации.

4. Изучены принципы и алгоритмы построения жадного координатного спуска и метода минимизации Франка-Вулфа.

5. Разработана программа, для нахождения оптимального инвестиционного портфеля.