

Оптимизация экономического портфеля с использованием методов машинного обучения

Бакалаврский диплом

Михаил Давыдов

ФПМИ МФТИ, кафедра дискретной математики

21 июня 2024 г.

Введение

Инвесторы вкладывают свои деньги в акции, и хотят использовать для распределения денег эффективные алгоритмы.

Введение

Инвесторы вкладывают свои деньги в акции, и хотят использовать для распределения денег эффективные алгоритмы. Если инвестор изначально не обладает информацией о стоимости активов, то в качестве модели можно использовать модель многоруких бандитов

Введение

Инвесторы вкладывают свои деньги в акции, и хотят использовать для распределения денег эффективные алгоритмы. Если инвестор изначально не обладает информацией о стоимости активов, то в качестве модели можно использовать модель многоруких бандитов

- Есть n рычагов, i -ый рычаг соответствует какому-то распределению со средним m_i . Изначально распределения, как и средние, неизвестны.

Введение

Инвесторы вкладывают свои деньги в акции, и хотят использовать для распределения денег эффективные алгоритмы. Если инвестор изначально не обладает информацией о стоимости активов, то в качестве модели можно использовать модель многоруких бандитов

- Есть n рычагов, i -ый рычаг соответствует какому-то распределению со средним m_i . Изначально распределения, как и средние, неизвестны.
- При нажатии на i -ый рычаг выдается награда в соответствии с i -ым распределением.

Введение

Инвесторы вкладывают свои деньги в акции, и хотят использовать для распределения денег эффективные алгоритмы. Если инвестор изначально не обладает информацией о стоимости активов, то в качестве модели можно использовать модель многоруких бандитов

- Есть n рычагов, i -ый рычаг соответствует какому-то распределению со средним m_i . Изначально распределения, как и средние, неизвестны.
- При нажатии на i -ый рычаг выдается награда в соответствии с i -ым распределением.
- Задача – найти $\arg \max_{\mathbf{p} \in \Delta^n} \mathbf{p}^T \cdot \mathbf{m} = \sum_{i=1}^n p_i m_i$, где

$\Delta^n = \{\mathbf{p} = (p_1, \dots, p_n) : (\sum_{i=1}^n p_i = 1) \wedge (\forall i \ p_i \geq 0)\}$, \mathbf{p} отвечает за долю от общих средств, вкладываемых в каждый актив на каждом шаге.

Введение

Инвесторы вкладывают свои деньги в акции, и хотят использовать для распределения денег эффективные алгоритмы. Если инвестор изначально не обладает информацией о стоимости активов, то в качестве модели можно использовать модель многоруких бандитов

- Есть n рычагов, i -ый рычаг соответствует какому-то распределению со средним m_i . Изначально распределения, как и средние, неизвестны.
- При нажатии на i -ый рычаг выдается награда в соответствии с i -ым распределением.
- Задача – найти $\arg \max_{\mathbf{p} \in \Delta^n} \mathbf{p}^T \cdot \mathbf{m} = \sum_{i=1}^n p_i m_i$, где $\Delta^n = \{\mathbf{p} = (p_1, \dots, p_n) : (\sum_{i=1}^n p_i = 1) \wedge (\forall i \ p_i \geq 0)\}$, \mathbf{p} отвечает за долю от общих средств, вкладываемых в каждый актив на каждом шаге.
- Равносильно нахождению рычага с наибольшим средним

Учет рисков

- Появляется дисперсия σ_i^2 .

Учет рисков

- Появляется дисперсия σ_i^2 .
- Задача – найти

$$\arg \max_{\mathbf{p} \in \Delta^n} \left(\mathbf{p}^T \cdot \mathbf{m} - \lambda (\mathbf{p}^T)^2 \cdot \sigma^2 \right) = \sum_{i=1}^n p_i m_i - \lambda \sum_{i=1}^n p_i^2 \sigma_i^2$$

где $\lambda > 0$ – коэффициент отвращения, или неприятия к риску.

Учет рисков

- Появляется дисперсия σ_i^2 .
- Задача – найти

$$\arg \max_{\mathbf{p} \in \Delta^n} \left(\mathbf{p}^T \cdot \mathbf{m} - \lambda (\mathbf{p}^T)^2 \cdot \sigma^2 \right) = \sum_{i=1}^n p_i m_i - \lambda \sum_{i=1}^n p_i^2 \sigma_i^2$$

где $\lambda > 0$ – коэффициент отвращения, или неприятия к риску.

- В этой трактовке задачи вектор вероятностей может не сосредотачиваться в одном рычаге.

Задачи

- 1 Проанализировать известные подходы в классической задаче о многоруких бандитах на предмет применимости для распределений, отличных от нормального.

Задачи

- 1 Проанализировать известные подходы в классической задаче о многоруких бандитах на предмет применимости для распределений, отличных от нормального.
- 2 Придумать алгоритмы и подходы для решения задачи о многоруких бандитах с учетом степени отвращения к риску.

Задачи

- 1 Проанализировать известные подходы в классической задаче о многоруких бандитах на предмет применимости для распределений, отличных от нормального.
- 2 Придумать алгоритмы и подходы для решения задачи о многоруких бандитах с учетом степени отвращения к риску.
- 3 Протестировать созданные подходы на степенных распределениях.

Задачи

- 1 Проанализировать известные подходы в классической задаче о многоруких бандитах на предмет применимости для распределений, отличных от нормального.
- 2 Придумать алгоритмы и подходы для решения задачи о многоруких бандитах с учетом степени отвращения к риску.
- 3 Протестировать созданные подходы на степенных распределениях.
- 4 Оценить степень применимости оценки риска через дисперсию

1 Greedy и ϵ -greedy

$$A_t = \begin{cases} \arg \max_a Q_t(a), & \text{with probability } 1 - \epsilon, \\ \text{a random action,} & \text{with probability } \epsilon. \end{cases}$$

Алгоритмы

1 Greedy и ϵ -greedy

$$A_t = \begin{cases} \arg \max_a Q_t(a), & \text{with probability } 1 - \epsilon, \\ \text{a random action,} & \text{with probability } \epsilon. \end{cases}$$

2 Стратегия с позитивной инициализацией ($\forall a Q_t(a) = d, d > 0$)

1 Greedy и ϵ -greedy

$$A_t = \begin{cases} \arg \max_a Q_t(a), & \text{with probability } 1 - \epsilon, \\ \text{a random action,} & \text{with probability } \epsilon. \end{cases}$$

2 Стратегия с позитивной инициализацией ($\forall a Q_t(a) = d, d > 0$)

3 Upper-Confidence Bound selection

$$A_t = \arg \max_a \left[Q_t(a) + c \sqrt{\frac{\ln t}{N_t(a)}} \right], \quad c > 0$$

1 Greedy и ϵ -greedy

$$A_t = \begin{cases} \arg \max_a Q_t(a), & \text{with probability } 1 - \epsilon, \\ \text{a random action,} & \text{with probability } \epsilon. \end{cases}$$

2 Стратегия с позитивной инициализацией ($\forall a Q_t(a) = d, d > 0$)

3 Upper-Confidence Bound selection

$$A_t = \arg \max_a \left[Q_t(a) + c \sqrt{\frac{\ln t}{N_t(a)}} \right], \quad c > 0$$

4 Gradient bandit

Алгоритмы

1 Greedy и ϵ -greedy

$$A_t = \begin{cases} \arg \max_a Q_t(a), & \text{with probability } 1 - \epsilon, \\ \text{a random action,} & \text{with probability } \epsilon. \end{cases}$$

2 Стратегия с позитивной инициализацией ($\forall a Q_t(a) = d, d > 0$)

3 Upper-Confidence Bound selection

$$A_t = \arg \max_a \left[Q_t(a) + c \sqrt{\frac{\ln t}{N_t(a)}} \right], \quad c > 0$$

4 Gradient bandit

В конце все алгоритмы были сравнены в зависимости от их гиперпараметров

Параметры

- Кол-во рычагов: 10
- Количество тестов – 2000
- Длина каждого теста – 1000 шагов

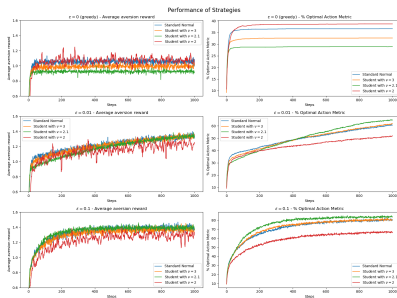
Параметры

- Кол-во рычагов: 10
- Количество тестов – 2000
- Длина каждого теста – 1000 шагов
- Распределения рычагов (все рычаги брались из одного семейства распределений):
 - Стандартное нормальное ($N(0, 1)$ или t_{∞})
 - Распределение Стьюдента с дисперсией 1 и 3-мя степенями свободы t_3 (для единичной дисперсии распределение было домножено на $\sqrt{\frac{1}{3}}$)
 - Распределение Стьюдента нормированное $t_{2.1}$ с 2.1 степенями свободы.
 - Распределение Стьюдента с 2-мя степенями свободы t_2
 - Распределение Коши t_1 (только в задаче без учета риска)

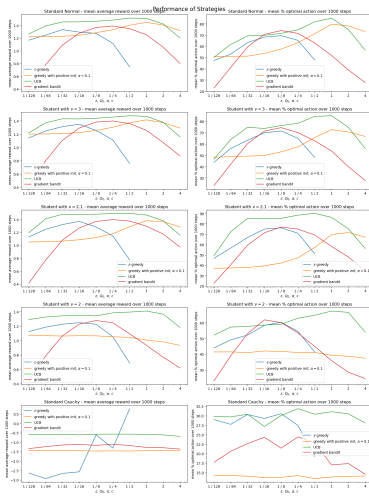
Параметры

- Кол-во рычагов: 10
- Количество тестов – 2000
- Длина каждого теста – 1000 шагов
- Распределения рычагов (все рычаги брались из одного семейства распределений):
 - Стандартное нормальное ($N(0, 1)$ или t_{∞})
 - Распределение Стьюдента с дисперсией 1 и 3-мя степенями свободы t_3 (для единичной дисперсии распределение было домножено на $\sqrt{\frac{1}{3}}$)
 - Распределение Стьюдента нормированное $t_{2.1}$ с 2.1 степенями свободы.
 - Распределение Стьюдента с 2-мя степенями свободы t_2
 - Распределение Коши t_1 (только в задаче без учета риска)

Результаты – ϵ -greedy



Результаты – финальное тестирование



Выводы

Проделанные эксперименты позволяют судить о том, что Gradient bandits, ϵ -greedy и UCB – стратегии показывают высокую эффективность на степенных распределениях. UCB при этом лучшая из стратегий.

Результаты

- 1 Разработаны градиентный жадный и алгоритмический жадный алгоритмы для вычисления

$$\arg \max_{\mathbf{p} \in \Delta^n} = \sum_{i=1}^n p_i m_i - \lambda \sum_{i=1}^n p_i^2 \sigma_i^2$$

, при известных m_i и σ_i^2 , причем последний работает за $O(n \log n)$

Результаты

- 1 Разработаны градиентный жадный и алгоритмический жадный алгоритмы для вычисления

$$\arg \max_{\mathbf{p} \in \Delta^n} = \sum_{i=1}^n p_i m_i - \lambda \sum_{i=1}^n p_i^2 \sigma_i^2$$

, при известных m_i и σ_i^2 , причем последний работает за $O(n \log n)$

- 2 Адаптированы стратегии ϵ -greedy, adaptive ϵ , positive initialization, UCB, gradient bandits.

Результаты

- 1 Разработаны градиентный жадный и алгоритмический жадный алгоритмы для вычисления

$$\arg \max_{\mathbf{p} \in \Delta^n} = \sum_{i=1}^n p_i m_i - \lambda \sum_{i=1}^n p_i^2 \sigma_i^2$$

, при известных m_i и σ_i^2 , причем последний работает за $O(n \log n)$

- 2 Адаптированы стратегии ϵ -greedy, adaptive ϵ , positive initialization, UCB, gradient bandits.
- 3 Создана стратегия с коррективкой дисперсии, переработана стратегия VDBE.

Метрики

- Среднее сожаление: если $V_{max} = (\mathbf{p}_{max}^T \cdot \mathbf{m} - \lambda(\mathbf{p}_{max}^2)^T \cdot \sigma^2)$, где \mathbf{p}_{max} – оптимальный вектор вероятностей, то $Reg_t = V_{max} - (\mathbf{p}_t^T \cdot \mathbf{q} - \lambda(\mathbf{p}_t^2)^T \cdot \mathbf{s}^2)$. Если $Reg_t < 0$, то алгоритм “переоценивает” себя.

Метрики

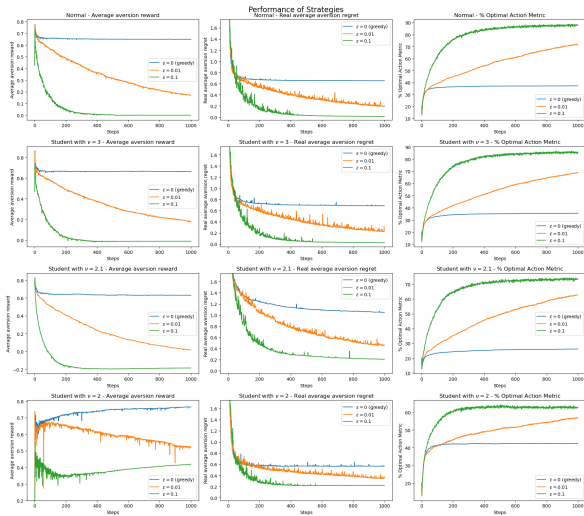
- Среднее сожаление: если $V_{max} = (\mathbf{p}_{max}^T \cdot \mathbf{m} - \lambda(\mathbf{p}_{max}^2)^T \cdot \boldsymbol{\sigma}^2)$, где \mathbf{p}_{max} – оптимальный вектор вероятностей, то $Reg_t = V_{max} - (\mathbf{p}_t^T \cdot \mathbf{q} - \lambda(\mathbf{p}_t^2)^T \cdot \mathbf{s}^2)$. Если $Reg_t < 0$, то алгоритм “переоценивает” себя.
- Среднее реальное сожаление:
 $RReg_t = V_{max} - (\mathbf{p}_t^T \cdot \mathbf{m} - \lambda(\mathbf{p}_t^2)^T \cdot \boldsymbol{\sigma}^2)$. Всегда ≥ 0 .

Метрики

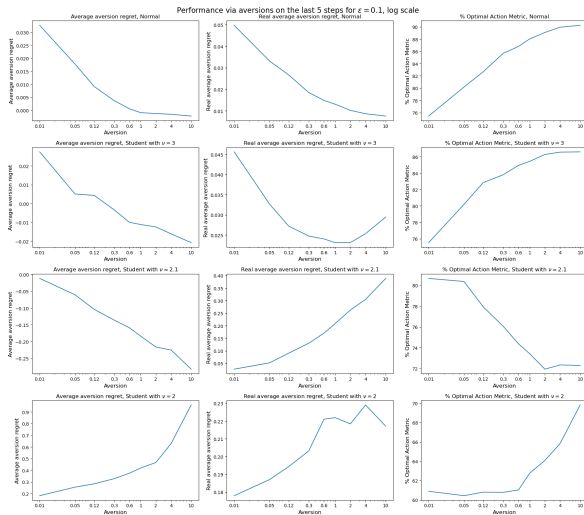
- Среднее сожаление: если $V_{max} = (\mathbf{p}_{max}^T \cdot \mathbf{m} - \lambda(\mathbf{p}_{max}^2)^T \cdot \boldsymbol{\sigma}^2)$, где \mathbf{p}_{max} – оптимальный вектор вероятностей, то $Reg_t = V_{max} - (\mathbf{p}_t^T \cdot \mathbf{q} - \lambda(\mathbf{p}_t^2)^T \cdot \mathbf{s}^2)$. Если $Reg_t < 0$, то алгоритм “переоценивает” себя.
- Среднее реальное сожаление:
 $RReg_t = V_{max} - (\mathbf{p}_t^T \cdot \mathbf{m} - \lambda(\mathbf{p}_t^2)^T \cdot \boldsymbol{\sigma}^2)$. Всегда ≥ 0 .
- Процент оптимальных действий:

$$\delta = 1 - \frac{1}{2} \sum_{i=1}^n |p_i - p_{i,max}|$$

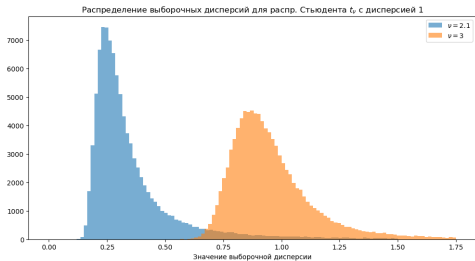
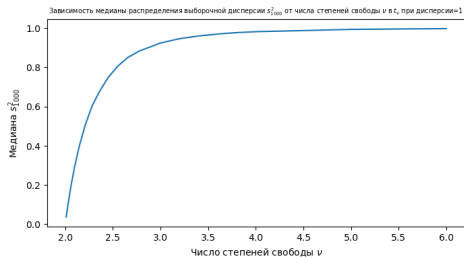
Результаты – ϵ -greedy



Результаты – ϵ -greedy

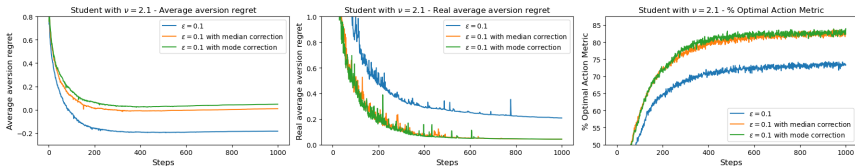


Результаты – Причины плохих метрик для $t_{2.1}$

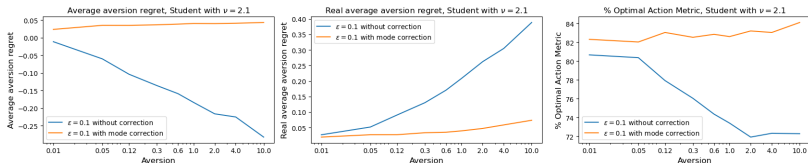


Результаты – Коррекция выборочной дисперсии

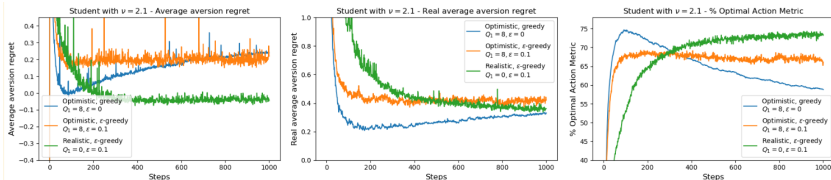
Performance of Strategies



Performance via aversions on the last 5 steps for $\epsilon = 0.1$ with and without correction, log scale

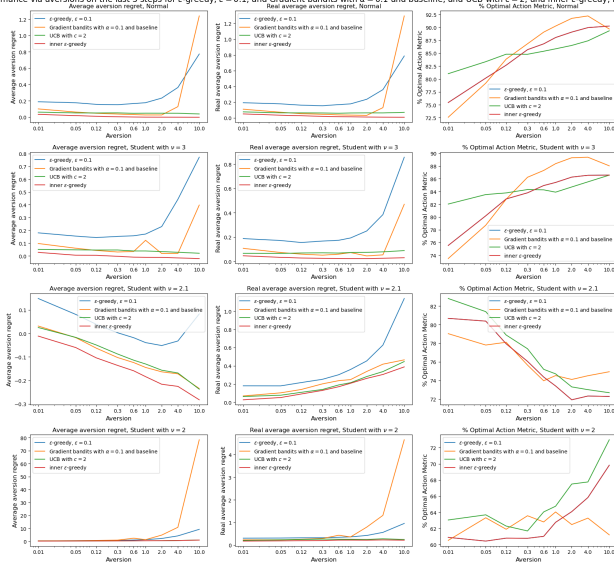


Жадная позитивная инициализация



Результаты – UCB и gradient bandits

Performance via aversions on the last 5 steps for ϵ -greedy, $\epsilon = 0.1$, and Gradient bandits with $\alpha = 0.1$ and baseline, and UCB with $c = 2$, and inner ϵ -greedy, log scale



Выводы

- ϵ -greedy дает высокую эффективность для t_3 и t_∞

Выводы

- ϵ -greedy дает высокую эффективность для t_3 и t_∞
- Для улучшения метрик для $t_{2.1}$ можно использовать коррекцию дисперсии.

Выводы

- ϵ -greedy дает высокую эффективность для t_3 и t_∞
- Для улучшения метрик для $t_{2.1}$ можно использовать коррекцию дисперсии.
- UCS даже без замены дисперсии устраняет разрыв между внутренней и внешней вероятностями в ϵ -greedy

Выводы

- ϵ -greedy дает высокую эффективность для t_3 и t_∞
- Для улучшения метрик для $t_{2.1}$ можно использовать коррекцию дисперсии.
- UCB даже без замены дисперсии устраняет разрыв между внутренней и внешней вероятностями в ϵ -greedy
- Gradient bandits показывают хорошие результаты и могут быть использованы при больших λ

Выводы

- ϵ -greedy дает высокую эффективность для t_3 и t_∞
- Для улучшения метрик для $t_{2.1}$ можно использовать коррекцию дисперсии.
- UCB даже без замены дисперсии устраняет разрыв между внутренней и внешней вероятностями в ϵ -greedy
- Gradient bandits показывают хорошие результаты и могут быть использованы при больших λ
- Оценка рисков через дисперсию при ν , близких к 2, слабоприменима

Заключение

В результате написания работы:

- Проанализированы известные подходы в классической задаче о многоруких бандитах на предмет применимости для распределений, отличных от нормального. По результатам UCSB показывает себя лучше других алгоритмов на всех распределениях.

Заключение

В результате написания работы:

- Проанализированы известные подходы в классической задаче о многоруких бандитах на предмет применимости для распределений, отличных от нормального. По результатам UCB показывает себя лучше других алгоритмов на всех распределениях.
- Придуманы и адаптированы алгоритмы и подходы для решения задачи о многоруких бандитах с учетом степени отвращения к риску. Среди них: жадный алгоритм за $O(n \log n)$, жадный алгоритм через градиентный подъем, ϵ -greedy, UCB, gradient bandits.

Заключение

В результате написания работы:

- Проанализированы известные подходы в классической задаче о многоруких бандитах на предмет применимости для распределений, отличных от нормального. По результатам UCB показывает себя лучше других алгоритмов на всех распределениях.
- Придуманы и адаптированы алгоритмы и подходы для решения задачи о многоруких бандитах с учетом степени отвращения к риску. Среди них: жадный алгоритм за $O(n \log n)$, жадный алгоритм через градиентный подъем, ϵ -greedy, UCB, gradient bandits.
- Протестированы созданные подходы на различных распределениях, получена высокая эффективность стратегий на распределениях Стюдента с $\nu \geq 3$. Сделан вывод о слабой применимости оценки риска через дисперсию при ν , близких к 2.