Аукционы на продажу ключевых слов

Сергей Николенко

Теория экономических механизмов — ИТМО, весна 2008

Outline

Бизнес-модель Google

- Бизнес-модель многих интернет-компаний строится на показах рекламы.
- Overture, Google AdWords...
- Продажа ключевых слов: когда пользователь ищет по тому или иному слову, в поле для рекламы отображаются релевантные платные объявления.
- Как определять «релевантность» — большой вопрос, но другой. Мы сейчас будем продавать показы.

CTR

- Для всякой рекламы i и всякого рекламного места (ключевого слова) j можно оценить *click-through rate* $\mathrm{CTR}_{i,j}$.
- Рекламодатель хочет максимизировать количество кликов или их эффективность.
- Механизм (т.е. сам поисковик) хочет максимизировать свой доход.

Рекламные слоты

- Итак, некоторое количество рекламодателей сражаются за показ их рекламы в рекламной выдаче по тому или иному ключевому слову.
- В рекламной выдаче несколько мест (слотов), причём, разумеется, чем выше реклама в этой выдаче, тем выше будет CTR ($\mathrm{CTR}_{i,j}$ для любого i не возрастает по j).
- Агенты делают ставки, механизм определяет, кому и по сколько продавать рекламные слоты.

Постановка

- Рассмотрим N агентов, которые сражаются за K слотов, $K < N$.
- Задана матрица, составленная из $\mathrm{CTR}_{i,j}$ для разных i от 1 до N и разных j; для удобства сделаем матрицу квадратной, положив $\mathrm{CTR}_{i,j} = 0$ для всех $j > K$.

Постановка

- Каждый агент i имеет свою внутреннюю ценность v_i; в данной модели ценность агенты присваивают тем ситуациям, когда на их рекламу кликает потребитель.
- Если агент i оказывается по итогам аукциона на месте j, то он получит доход, равный $\mathrm{CTR}_{i,j} v_i$.
- С другой стороны, плата (как и в настоящих рекламных моделях Google, Yahoo и др.) взимается за клик, а не за показ.
- В итоге, если агент i, занимая место j, платит p_i, то общая (ожидаемая) прибыль агента будет равна

$$u_i = \mathrm{CTR}_{i,j} \left(v_i - p_i \right).$$

Постановка

- Ограничимся пока ситуацией, в которой перед нами ровно один аукцион на ровно один показ рекламы.
- Обозначим ставку агента i через b_i.
- Механизм должен по вектору $(b_1, b_2, \ldots, b_N)$ определить расстановку агентов, то есть ввести некоторую *функцию ранжирования* (ranking function).
- Нас интересует достаточно узкий класс таких функций: мы предполагаем, что механизм присваивает каждому участнику аукциона вес w_i, а затем ранжирует их по убыванию произведений $w_i b_i$.

Разделимость

Осталось рассмотреть ещё одно важное предположение. Мы не всегда будем им пользоваться, но для результатов об эквивалентности доходности пригодится. Это предположение утверждает, что $\mathrm{CTR}_{i,j}$ зависит и от агента, и от места, но эти факторы можно разделить.

Определение

Набор click-through rates для N агентов и K рекламных мест называется разделимым (separable), если существуют такие $\mu_1, \mu_2, \ldots, \mu_n > 0$ и $\theta_1 \geq \theta_2 \geq \ldots \geq \theta_K > 0$, что

$$\mathrm{CTR}_{i,j} = \mu_i \theta_j.$$

Разделимость

Определение

Набор click-through rates для N агентов и K рекламных мест называется разделимым (separable), если существуют такие $\mu_1, \mu_2, \ldots, \mu_n > 0$ *и* $\theta_1 \geq \theta_2 \geq \ldots \geq \theta_K > 0$, *что*

$$\mathrm{CTR}_{i,j} = \mu_i \theta_j.$$

Иначе говоря, матрицу $\mathbf{CTR} = (\mathrm{CTR}_{i,j})_{i,j}$ можно представить в виде произведения

$$\mathbf{CTR} = \mu\theta^{\perp},$$

где $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 & \mu_2 & \ldots & \mu_N \end{pmatrix}^{\perp}$ и $\boldsymbol{\theta} = \begin{pmatrix} \theta_1 & \theta_2 & \ldots & \theta_K \end{pmatrix}^{\perp}$.

Модели аукционов

- Разные модели аукционов могут различаться весами w_i, по которым происходит ранжирование, и выплатами, которые делают участники.
- Нас будут интересовать три модели, две из которых относятся к классу *аукционов следующей цены* (next-price auctions).

Модели аукционов

1. Модель *Overture* (он же *Yahoo*). Рекламодатели делают ставки b_i, они ранжируются по убыванию абсолютной ставки b_i, после чего с победителя аукциона берут наименьшую цену из тех, что гарантируют победу при фиксированных ставках остальных (принцип второй цены).
2. Модель *Google AdWords*. Рекламодатели делают ставки, они ранжируются по убыванию показателя $b_i\mathrm{CTR}_{i,j}$, после чего с победителя аукциона берут наименьшую цену из тех, что гарантируют победу при фиксированных ставках остальных (тоже принцип второй цены, но здесь он чуть модифицирован, потому что ранжируются не ставки, а $b_i\mathrm{CTR}_{i,j}$).

Модели аукционов

- Эти две модели можно обобщить.

Определение

Аукцион следующей цены задаётся весами ранжирующей функции $\mathbf{w} = (w_1, w_2, \ldots, w_N)$. *Для данных ставок агентов* $\mathbf{b} = (b_1, b_2, \ldots, b_N)$ *аукцион следующей цены ранжирует их в порядке убывания* $w_i b_i$ *и взимает с агента* i, *попавшего на место* j, *минимальную ставку, которая понадобилась бы ему для того, чтобы остаться на месте* j: *если на* $j+1$ *месте оказался агент* i', *то аукцион взимает с агента* i *сумму* $\frac{w_{i'} b_{i'}}{w_i}$.

Аукционы следующей цены: пример

- Рассмотрим трёх агентов — 1, 2 и 3 — которые сражаются за два рекламных слота. Предположим, что ожидаемый CTR от этих слотов для всех трёх агентов одинаковый и равен

$$\begin{aligned} \mathrm{CTR}_{11} = \mathrm{CTR}_{21} = \mathrm{CTR}_{31} &= 0{,}2, \\ \mathrm{CTR}_{12} = \mathrm{CTR}_{22} = \mathrm{CTR}_{32} &= 0{,}15. \end{aligned}$$

- Кроме того, предположим, что веса аукциона следующей цены для агентов тоже совпадают, а истинные ценности агентов составляют 10, 7 и 2 единицы за клик соответственно.

Аукционы следующей цены: пример

- Тогда, если все агенты ставят свою истинную ценность, агент 1 выигрывает первый слот, платит за клик 7 единиц (ставку второго агента) и получает с показа ожидаемую прибыль
$$(10-7) \times 0,2 = 0,6.$$
- Но если бы агент 1 соврал и поставил 6, он проиграл бы первый слот агенту 2, а сам начал бы платить за клик 2 единицы (ставку третьего агента) и получать с показа ожидаемую прибыль
$$(10-2) \times 0,15 = 1,2.$$
- Следовательно, стратегии «говорить правду» никакого равновесия в этой ситуации не составляют.

VCG

- Без правдивости тяжело. Может быть, нас спасут механизмы Викри-Кларка-Гровса (VCG)? Они всегда правдивы.
- Однако оказывается, что никакой VCG-механизм, даже взвешенный, даже с априорными предпочтениями, не сможет достичь того же результата, что *Overture*-схема работы аукциона.

VCG

- Рассмотрим двух агентов — 1 и 2 — которые сражаются за два рекламных слота, и их CTR таков:

$$\mathrm{CTR}_{1,1} = \mathrm{CTR}_{2,1} = 0,2; \quad \mathrm{CTR}_{1,2} = 0,1; \quad \mathrm{CTR}_{2,2} = 0,2.$$

- Иначе говоря, первому агенту невыгодно проигрывать первый слот, а второй агент к этому вообще нечувствителен.

VCG

- Обозначим через w_1 и w_2 веса агентов в VCG-аукционе, а через $k_{i,j}$ — предпочтение, оказываемое этим аукционом тому, чтобы расположить агента i выше агента j ($k_{i,j}$ — константа для данного аукциона и фиксированных i и j).
- Отметим, что $w_1 > 0$ и $w_2 > 0$, ведь любой агент должен иметь шанс выиграть аукцион. Тогда первый агент будет выигрывать первый слот в VCG-аукционе тогда и только тогда, когда

$$w_1(0,2b_1) + w_2(0,2b_2) + k_{1,2} > w_1(0,1b_1) + w_2(0,2b_2) + k_{2,1}.$$

VCG

- Получается, что это событие совершенно не зависит от ставки второго агента, и первый агент будет выигрывать, если его ставка превышает $\frac{k_{2,1}-k_{1,2}}{0{,}1 w_2}$.
- А в любой схеме ранжирования с ненулевыми весами ставок у второго агента есть шанс оказаться выше первого, что и приводит нас к противоречию.
- Это эффективно и формально правильно, но нехорошо: возможность победить должна быть, даже если это может привести к неэффективности.
- Кроме того, это, конечно, приводит к неоптимальности: продавцу выгоднее получить цену побольше, а тут получается, что первый агент выигрывает при любой ненулевой ставке (если нет предпочтений).

Поэтапный аукцион

- Рассмотрим третью схему.

Определение

Поэтапный аукцион на K рекламных слотов, как и аукцион следующей цены, определяется вектором весов $(w_1, \ldots, w_N)$, принимает ставки рекламодателей b_i и ранжирует их в порядке убывания $w_i b_i$ (будем предполагать без потери общности, что агент i занял в этом порядке i-е место). Затем поэтапный аукцион взимает с агента i, где $1 \leq i \leq K$, цену

$$p_i = \sum_{j=i}^{K} \left(\frac{\mathrm{CTR}_{i,j} - \mathrm{CTR}_{i,j+1}}{\mathrm{CTR}_{i,i}} \right) \frac{w_{j+1}}{w_i} b_{j+1}.$$

Поэтапный аукцион

- Идея формулы:

$$\mathrm{CTR}_{i,i} p_i = \sum_{j=i}^{K} \left(\mathrm{CTR}_{i,j} - \mathrm{CTR}_{i,j+1}\right) \frac{w_{j+1}}{w_i} b_{j+1}.$$

- За разные клики агент платит по-разному:
  - за те клики, которые агент i получил бы, будь он на позиции $i+1$, он платит столько же, сколько агент, который на ней находится (тут возникает рекурсия — цена для агента на позиции $i+1$ при таком подходе зависит от цены агента $i+2$ и так далее);
  - за дополнительные клики от позиции i агент доплачивает по принципу второй цены: платит минимальную цену, которая позволила бы ему оставаться на позиции i.

Поэтапный аукцион

- Рассмотрим четырёх агентов, которые сражаются за три рекламных слота.
- Предположим, что их CTR задаются следующим образом:

$$\begin{array}{llll} \mathrm{CTR}_{1,1} = 0,5; & \mathrm{CTR}_{2,1} = 0,45; & \mathrm{CTR}_{3,1} = 0,4; & \mathrm{CTR}_{4,1} = 0, \\ \mathrm{CTR}_{1,2} = 0,45; & \mathrm{CTR}_{2,2} = 0,4; & \mathrm{CTR}_{3,2} = 0,2; & \mathrm{CTR}_{4,2} = 0, \\ \mathrm{CTR}_{1,3} = 0,3; & \mathrm{CTR}_{2,3} = 0,25; & \mathrm{CTR}_{3,3} = 0,15; & \mathrm{CTR}_{4,3} = 0, \end{array}$$

Поэтапный аукцион

- Предположим также, что их веса и ставки равны

$$\begin{array}{llll} w_1 = 5; & w_2 = 5; & w_3 = 4; & w_4 = 3. \\ b_1 = 30; & b_2 = 25; & b_3 = 18; & b_4 = 22. \end{array}$$

- Давайте разберёмся, что будет происходить в поэтапном аукционе с такими весами и такими CTR.

Поэтапный аукцион

- Во-первых, агенты будут ранжированы в порядке, соответствующем их номерам:
$$w_1 b_1 = 150 > w_2 b_2 = 125 > w_3 b_3 = 72 > w_4 b_4 = 66.$$
- Теперь нужно определить, кому какая по итогам аукциона достанется цена за клик. Четвёртый агент не занимает никакого рекламного слота, и поэтому его цена равна нулю:
$$p_4 = 0.$$

Поэтапный аукцион

- Третий агент платит, по условиям поэтапного аукциона,

$$p_3 = \sum_{i=3}^{3} \left(\frac{\mathrm{CTR}_{3,i} - \mathrm{CTR}_{3,i+1}}{\mathrm{CTR}_{3,3}} \right) \frac{w_{i+1}}{w_3} b_{i+1} = \\ = \frac{\mathrm{CTR}_{3,3}}{\mathrm{CTR}_{3,3}} \frac{w_4}{w_3} b_4 = 16,5.$$

- Для последнего нетривиального агента аукцион превращается в точности в аукцион второй цены: он платит столько, сколько нужно, чтобы удержаться на своём месте.

Поэтапный аукцион

- Цена второго агента — это цена третьего за те клики, которые второй получил бы на третьем месте, плюс вторая цена за «добавочную стоимость», которую он получает на своём втором месте:

$$p_2 = \sum_{i=2}^{3} \left(\frac{\mathrm{CTR}_{2,i} - \mathrm{CTR}_{2,i+1}}{\mathrm{CTR}_{2,2}} \right) \frac{w_{i+1}}{w_2} b_{i+1} =$$
$$= \frac{\mathrm{CTR}_{2,2} - \mathrm{CTR}_{2,3}}{\mathrm{CTR}_{2,2}} \frac{w_3}{w_2} b_3 + \frac{\mathrm{CTR}_{2,3}}{\mathrm{CTR}_{2,2}} \frac{w_4}{w_2} b_4 = 13,65.$$

- Заметим, что второй агент в итоге платит меньше, чем третий; и вес у него больше, и CTR тоже больше.

Поэтапный аукцион

- И, наконец, подсчитаем цену за клик для первого агента:

$$p_1 = \sum_{i=1}^{3} \left(\frac{\mathrm{CTR}_{1,i} - \mathrm{CTR}_{1,i+1}}{\mathrm{CTR}_{1,1}} \right) \frac{w_{i+1}}{w_1} b_{i+1} =$$

$$= \frac{\mathrm{CTR}_{1,1} - \mathrm{CTR}_{1,2}}{\mathrm{CTR}_{1,1}} \frac{w_2}{w_1} b_2 +$$

$$+ \frac{\mathrm{CTR}_{1,2} - \mathrm{CTR}_{1,3}}{\mathrm{CTR}_{1,1}} \frac{w_3}{w_1} b_3 + \frac{\mathrm{CTR}_{1,3}}{\mathrm{CTR}_{1,1}} \frac{w_4}{w_1} b_4 = 14,74.$$

Правдивость поэтапного аукциона

Теорема

Для любых фиксированных весов $w_1, \ldots, w_n$ поэтапный аукцион является правдивым. Более того, это единственный правдивый аукцион, который ранжирует агентов по убыванию $w_i b_i$.

Правдивость поэтапного аукциона

- Без потери общности перенумеруем агентов так, чтобы

$$w_1 b_1 \geq w_2 b_2 \geq \ldots \geq w_n b_n.$$

- Чтобы доказать правдивость, предположим противное: пусть агенту k выгодно изменить свою ставку. Поскольку цена зависит только от ранга агента, «соврать» означает изменить свой ранг.

Правдивость поэтапного аукциона

- Давайте среди всех рангов, дающих агенту k максимальную прибыль на один показ, выберем ранг l, который ближе всех находится к истинному его рангу k.
- Мы хотели бы доказать, что $l = k$; предположим, что это не так.

Правдивость поэтапного аукциона

- Тогда, если $l > k$, то агенту оказывается выгодно сменить l на $l-1$: прибыль его изменится на
$$(\mathrm{CTR}_{k,l-1} - \mathrm{CTR}_{k,l})\left(v_k - \frac{w_l}{w_k}b_l\right) \geq 0,$$
что даёт противоречие с минимальностью l.
- А если $l < k$, то противоречие с максимальностью аналогично получается при переходе от l к $l+1$: прибыль изменится на
$$(\mathrm{CTR}_{k,l+1} - \mathrm{CTR}_{k,l})\left(v_k - \frac{w_l}{w_k}b_l\right) \geq 0$$
(раньше оба сомножителя были неотрицательными, теперь они оба неположительные).

Правдивость поэтапного аукциона

- Теперь о единственности. Рассмотрим аукцион $\mathcal{A}$, который ранжирует по убыванию $w_i b_i$.
- Рассмотрим одного агента x, зафиксируем ставки остальных и рассмотрим функцию $p_{\mathcal{A}}(j)$ — цену, которую аукцион берёт с агента в зависимости от того, куда он попадает по рангу ($p_{\mathcal{A}}(K+1) = 0$).
- Предположим wlog, что остальные проранжированы так, что агент i занимает место i, если M ставит $+\infty$.
- Чтобы доказать единственность, достаточно показать, что

$$p_{\mathcal{A}}(j) - p_{\mathcal{A}}(j+1) = \left(\mathrm{CTR}_{x,j} - \mathrm{CTR}_{x,j+1}\right) \frac{w_{j+1}}{w_x} b_{j+1}.$$

Правдивость поэтапного аукциона

- Предположим сначала, что у x истинная ценность $v_x = \frac{w_{j+1}}{w_x} b_{j+1} + \epsilon$.
- Тогда при правдивой ставке он получает ранг j.
- Дополнительный доход за показ от его j-го места составит $\left(\mathrm{CTR}_{x,j} - \mathrm{CTR}_{x,j+1}\right) v_x$.
- Значит, правдивый аукцион не может брать больше, иначе агенту x будет выгодно занизить ставку и занять $(j+1)$-е место.
- Но ϵ сколь угодно мало, значит, $\leq$ доказали.

Правдивость поэтапного аукциона

- С другой стороны, рассмотрим истинную ценность $v_x = \frac{w_{j+1}}{w_x} b_{j+1} - \epsilon$.
- Совершенно аналогичный рассуждения покажут, что

$$p_{\mathcal{A}}(j) - p_{\mathcal{A}}(j+1) \geq \left(\mathrm{CTR}_{x,j} - \mathrm{CTR}_{x,j+1}\right) \frac{w_{j+1}}{w_x} b_{j+1}.$$

- В итоге из правдивости мы вывели в точности цену поэтапного аукциона, что и требовалось доказать.

Outline

Введение

- Мы теперь знаем, как проводить один аукцион на ключевые слова или рекламные слоты.
- Описали как минимум три разных модели, каждая из которых уже доказала свою эффективность на практике.
- Иначе говоря, мы научились продавать *один* рекламный слот.

Введение

- Но ведь на самом деле слотов много, а у агентов, то есть компаний, размещающих рекламу, денег ограниченное количество.
- Если каждый агент будет участвовать в каждом аукционе, то, скорее всего, многим агентам просто не хватит денег.
- Поэтому нужно решать и вопрос о том, как *покупать* рекламу, а не только продавать.

Введение

- Пусть у каждого рекламодателя есть некий фиксированный бюджет, который он готов потратить на рекламу.
- Бюджет ограничивается на какой-то период времени («день»).
- Каждый раз, когда возникает повод показать рекламу, проходит аукцион между теми рекламодателями, которые выразили интерес к этому рекламному слоту и *у которых ещё остались деньги в сегодняшнем бюджете*.
- Проблема для рекламодателя здесь заключается в том, что не все аукционы одинаково полезны, а бюджет ограничен. Возникает задача оптимизации.

Аукционы на ключевые слова

- В аукционе на ключевые слова рекламодатели делают ставку на то или иное слово, которое пользователь использует в поисковом запросе.
- Но в запросе при этом может оказаться сразу несколько слов!
- А некоторые ключевые слова всё равно стоит учитывать, даже если в самом запросе они не фигурируют.

Аукционы на ключевые слова

- Формально говоря, получается двудольный граф G, множество вершин которого делится на множество ключевых слов K и множество запросов Q.
- Между ключевым словом и запросом есть ребро, если запрос «соответствует» ключевому слову, то есть при таком запросе будет показана реклама по этому ключевому слову.
- Рекламодатели делают ставки на элементы K, но аукционы проводятся, когда поступает тот или иной элемент Q.

Аукционы на ключевые слова

- Формально это можно записать так: рекламодатель определяет ставку b_k для каждого $k \in K$.
- Для каждого запроса $q \in Q$ механизм определяет функцию $b_q(b_{k_1}, \ldots, b_{k_l})$, где $b_{k_1}, \ldots, b_{k_l}$ — соседи q в графе G, а также две функции $\mathrm{Clicks}_q(b_q)$ и $\mathrm{Cost}_q(b_q)$, которые определяют, сколько кликов и по какой цене рекламодатель получит, если подаст ставку b_q на запрос q.

Аукционы на ключевые слова

- Рекламодателю требуется оптимизировать суммарный трафик при условии, что бюджет ограничен некоторой величиной U, то есть решить задачу

$$\max_{k \in K} \sum_q \mathrm{Clicks}_q(b_q(b_{k_1}, \ldots, b_{k_l}))$$

$$\text{при} \quad \sum_q \mathrm{Cost}_q(b_q(b_{k_1}, \ldots, b_{k_l})) \leq U,$$

где сумма берётся по запросам, подаваемым пользователями (т.е. ожидание).

Аукционы на ключевые слова

- Как и в предыдущем разделе, у каждого запроса есть K слотов, за которые сражаются агенты-рекламодатели.
- Мы сейчас рассматриваем одного агента, поэтому вместо матрицы $\mathbf{CTR}$ нас интересует только вектор $(\mathrm{CTR}_1, \ldots, \mathrm{CTR}_K)$, состоящий из наших click-through rates. ($\mathrm{CTR}_i \geq \mathrm{CTR}_j$ при $i \leq j$).
- Другие агенты подают ставки $(b_1, \ldots, b_{N-1})$; без потери общности можно рассматривать ровно K наивысших ставок других агентов $(b_1, \ldots, b_K)$, а также предполагать, что $b_K = \mathrm{CTR}_K = 0$ и $b_0 = \infty$.

Модель аукциона

- Мы в этом разделе будем рассматривать обобщённый аукцион второй цены: будем платить за каждый клик столько, сколько поставил следующий за нами участник.
- Место в результатах аукциона будет определяться ставкой агента:
$$\mathrm{Pos}(b) = \mathrm{argmax}_{i=1..K}\{b_i :\ b_i \le b\}.$$
- В этих обозначениях платить мы будем, собственно, ровно $b_{\mathrm{Pos}(b)}$ за клик, и, следовательно,

$$\begin{aligned} \mathrm{Cost}_q(b) &= \mathrm{CTR}_{\mathrm{Pos}(b)} b_{\mathrm{Pos}(b)}, \\ \mathrm{Clicks}_q(b) &= \mathrm{CTR}_{\mathrm{Pos}(b)}. \end{aligned}$$

Ландшафты

- Заметим, что пара $(\mathrm{Cost}_q(b), \mathrm{Clicks}_q(b))$ может принимать лишь конечное число допустимых значений.
- Кроме того, $\mathrm{Cost}_q(b)$ и $\mathrm{Clicks}_q(b)$ являются неубывающими функциями от b.
- Более того, их отношение $\frac{\mathrm{Cost}_q(b)}{\mathrm{Clicks}_q(b)}$ тоже не убывает, а численно оно равно ставке следующего участника:

$$\frac{\mathrm{Cost}_q(b)}{\mathrm{Clicks}_q(b)} = b_{\mathrm{Pos}(b)} \leq b.$$

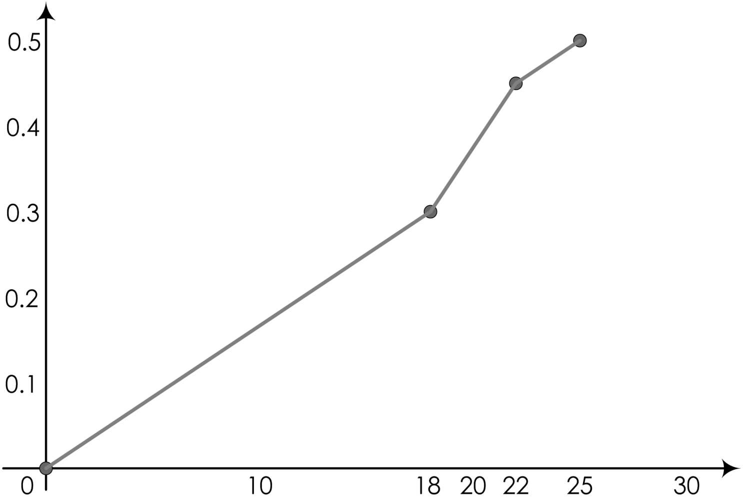
- Таким образом, для каждого запроса q у рекламодателя образуется так называемый *ландшафт* (landscape) возможных значений $(\mathrm{Cost}_q(b), \mathrm{Clicks}_q(b))$.

Примеры ландшафтов

- Вспомним пример из прошлого раздела и предположим, что там использовался обобщённый аукцион второй цены.
- Рассмотрим его с позиций первого агента:

$$\begin{aligned} \mathrm{CTR}_1 &= 0.5, & b_1 &= 25, \\ \mathrm{CTR}_2 &= 0.45, & b_2 &= 22, \\ \mathrm{CTR}_3 &= 0.3, & b_3 &= 18, \\ \mathrm{CTR}_4 &= 0, & b_4 &= 0. \end{aligned}$$

Примеры ландшафтов

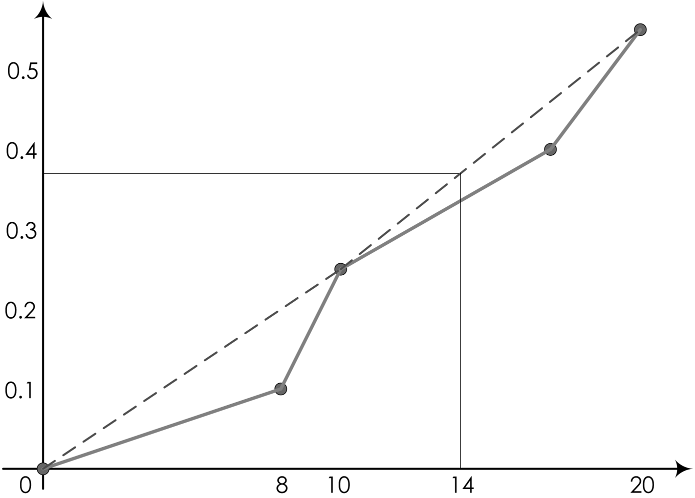


Примеры ландшафтов

- А более сложный ландшафт можно получить, если рассмотреть, к примеру, запрос с пятью слотами и следующими CTR и ставками:

$$\begin{aligned} \mathrm{CTR}_1 &= 0.55, & b_1 &= 20, \\ \mathrm{CTR}_2 &= 0.4, & b_2 &= 17, \\ \mathrm{CTR}_3 &= 0.25, & b_3 &= 10, \\ \mathrm{CTR}_4 &= 0.1, & b_4 &= 8, \\ \mathrm{CTR}_5 &= 0, & b_5 &= 0. \end{aligned}$$

Примеры ландшафтов



Свойства ландшафтов

- Уже ясно, что ландшафт не обязан быть выпуклым.
- Но нечто похожее на выпуклость ему всё же присуще: это уже упоминавшееся свойство неубывания $\frac{\mathrm{Cost}_q(b)}{\mathrm{Clicks}_q(b)}$.

Оптимизация для одного запроса

- Если решать задачу оптимизации для одного запроса, то получится, что нужно просто подавать максимальную ставку, которая не превышает имеющийся бюджет.
- Проблема: множество значений функции $\mathrm{Cost}_q(b)$ конечно, и бюджет может превысить $\mathrm{Cost}_q(b_i)$, но не дотянуть до $\mathrm{Cost}_q(b_{i-1})$.
- Чтобы избавиться от этой неэффективности, мы разрешаем *вероятностные стратегии*: разрешаем агентам подавать в качестве ставок распределения вероятностей над возможными ставками.

Вероятностные стратегии

- Рассмотрим простейшую вероятностную стратегию: выбирать с некоторой вероятностью между двумя вариантами.
- С математической точки зрения такая стратегия — это по сути просто точка на прямой, соединяющей эти варианты на графике ландшафта.
- А оптимальной стратегией будет некоторая точка на выпуклой оболочке точек ландшафта.

Вероятностные стратегии

Лемма

Если рекламодатель делает ставку на одно ключевое слово с бюджетом U, то оптимальной стратегией будет выпуклая комбинация (не более чем) двух ставок, являющихся концами отрезка выпуклой оболочки, пересекающегося с ценой U в наивысшей точке.

(доказательство — упражнение)

Вероятностные стратегии

- Здесь есть один тонкий момент: с некоторой вероятностью рекламодателю придётся превысить свой бюджет.
- Поэтому если бюджет жёсткий, то такая стратегия может и не оказаться оптимальной.
- Однако мы будем предполагать, что бюджет мягкий, к чему есть много разумных аргументов.

Постановка

- Однако ведь на самом деле ключевых слов много, и они взаимодействуют друг с другом, порой довольно хитрым образом.
- Что же делать, как рекламодателю оптимизировать свою прибыль в общей ситуации?

Постановка

- Формально говоря, у нас есть двудольный граф $G = (K \cup Q, E)$ ключевых слов и поисковых запросов.
- Рекламодатель подаёт вектор ставок $\boldsymbol{a} \in \mathbb{R}^{|K|}$ или задаёт распределение вероятностей над $\mathbb{R}^{|K|}$ (вероятностная стратегия).

Постановка

- Для некоторого набора поисковых запросов мы по $\boldsymbol{a}$ определяем $\boldsymbol{b}_q(\boldsymbol{a})$ достаточно простым образом:

$$b_q(\boldsymbol{a}) = \max_{k:(k,q)\in E} a_k.$$

- Суммарно рекламодатель тратит

$$\mathrm{Spend}(\boldsymbol{a}) = \sum_{q\in Q} \mathrm{Cost}_q(b_q(\boldsymbol{a}))$$

и получает траффик

$$\mathrm{Traffic}(\boldsymbol{a}) = \sum_{q\in Q} \mathrm{Clicks}_q(b_q(\boldsymbol{a})).$$

Постановка

- Задача рекламодателя — максимизировать

$$\boldsymbol{E}_{\boldsymbol{a}\in\mathcal{A}}\left[\mathrm{Traffic}(\boldsymbol{a})\right]$$

по распределениям $\mathcal{A}$, соблюдая условие

$$\boldsymbol{E}_{\boldsymbol{a}\in\mathcal{A}}\left[\mathrm{Traffic}(\boldsymbol{a})\right] \le U.$$

- А на вход рекламодателю подаются те самые ландшафты, только теперь возможных запросов много, и у каждого свой ландшафт.

Постановка

- Поиск оптимальной стратегии может оказаться очень сложным, и оптимальный вектор ставок (или оптимальное распределение) может иметь очень сложную структуру.
- Однако, к счастью, можно получить хорошую аппроксимацию, используя крайне ограниченный набор стратегий.

Равномерные стратегии

Определение

Равномерная стратегия рекламодателя — это распределение $\mathcal{A}$ *над векторами* $\boldsymbol{a} \in \mathbb{R}^{|K|}$, *которая присваивает ненулевые вероятности только векторам вида* $(b, b, \ldots, b)$ *для некоторых* $b \in \mathbb{R}$.

Равномерные стратегии

- Проще говоря, равномерная стратегия вместо распределения над $\mathbb{R}^{|K|}$ задаёт распределение над $\mathbb{R}$, а затем просто ставит одинаковые суммы на каждое ключевое слово.
- Прежде всего заметим, что оптимальную равномерную стратегию легко найти.

Поиск оптимальной равномерной стратегии

Теорема

Для вышеописанной задачи оптимизации траффика, в которой ландшафты поисковых запросов суммарно содержат N точек, можно найти оптимальную равномерную стратегию за время $O(N \log N)$. Более того, эта стратегия всегда будет присваивать ненулевые вероятности не более чем двум возможным ставкам.

Поиск оптимальной равномерной стратегии

- Если мы можем делать только одинаковые ставки на все ключевые слова (и, как следствие, на все поисковые запросы), то все ландшафты можно просто свести в один, просуммировав клики и траффик по всем ландшафтам.
- А затем можно воспользоваться стратегией, которую мы уже выработали выше для одного ландшафта: взять наивысшее пересечение с выпуклой оболочкой точек ландшафта и рассмотреть его как равномерную стратегию, варьирующуюся между двумя возможными ставками.
- Сложность этого алгоритма определяется сложностью построения выпуклой оболочки; это можно сделать за $O(N \log N)$.

Оптимальность равномерных стратегий

Теорема

В задаче оптимизации траффика всегда существует равномерная стратегия, которая является $\left(1-\frac{1}{e}\right)$-оптимальной, то есть даёт суммарный траффик, составляющий не менее чем $\left(1-\frac{1}{e}\right)$ долю траффика оптимальной стратегии. С другой стороны, эта оценка точна: для любой равномерной стратегии и любого $\epsilon > 0$ найдётся такой граф и такой набор ставок других агентов, что эта стратегия будет не более чем $\left(1-\frac{1}{e}+\epsilon\right)$-оптимальной.

Спасибо за внимание!

- Lecture notes и слайды будут появляться на моей homepage:
  http://logic.pdmi.ras.ru/~sergey/index.php?page=teaching
- Присылайте любые замечания, решения упражнений, новые численные примеры и прочее по адресам: sergey@logic.pdmi.ras.ru, snikolenko@gmail.com
- Заходите в ЖЖ smartnik.