Динамическое ценообразование с помощью Томсонского Сэмплирования*

А. А. Харъ, Ю. В. Дорн

В наше время появляется все больше и больше онлайн магазинов, в которых продают товары разных категорий. И у людей возникает проблема: каким образом устанавливать цены? Как не сделать их слишком высокими, чтобы покупатели в принципе приобретали определенный товар, или же не сделать слишком низкими, чтобы не потерять свою прибыль? В этой статье мы изучим алгоритм динамического ценообразования. Делать мы это будем, основываясь на стандартных алгоритмах для многоруких бандитов (игровые автоматы с несколькими ручками, у которых разные распределения на выигрыши). Это непростая проблема, потому что поведение покупателя неизвестно, и, производя, так называемый exploration, мы получаем бандитский фидбек (знание лишь о той ручке, которую мы выбрали).

Ключевые слова: *многорукие бандиты, Томсонское Сэмплирование, прибыль, спрос, эластичность, "пассивные"алгоритмы*

DOI:

1 Введение

Динамическое ценообразование является сложной задачей из-за того, что поведение покупателя неизвестно, а также фидбек бандитский. Стандартный подход к этой проблеме состоит из следующих 4 шагов:

- собирание данных о ценых на товар и спрос на этот товар при различных его стоимостях;
- создание статической модели для спроса как функцию от цены, оценка параметров модели с использованием собранных данных;
- используя обученную функцию спроса, оптимизировать некоторые метрики (например, прибыль) для получения новой оптимальной цены;
- установление полученной цены на товар в ближайшие несколько дней и повторение снова.

Это, так называемый, "пассивный"подход к динамическому ценообразованию. Такие подходы "близоруки"и пытаются оптимизировать метрику кратковременно. В [1] показано, что такие подходы приводят к неполному обучению и не очень хорошему результаты в перспективе.

Мы же попробуем избавиться от этой проблемы таки образом:

- предложим и реализуем простой, "активный"алгоритм для динамического ценообразования, называемый Max-Rev-TS, который рассматривает проблему динамического ценообразования как оптимизационную проблему в условиях неопределенности;
- наша система точно оценит параметры для функции спроса и максимизирует метрику (прибыль) в условиях ограничений;
- используя априорное распределение неизвестных параметров функции спроса, а также используя Байесовское правило, мы получил апостериорное распределение;

*Работа выполнена при финансовой поддержке РФФИ, проекты № №00-00-00000 и 00-00-00001.

- покажем, что Max-Rev-TS дает серьезное улучшение качества по сравнению с "пассивными"алгоритмами.

2 Постановка задачи

Стандартный подход к модели спроса в проблеме ценообразования - это предположение что функция спроса из какого-либо параметрического семейства и затем оценивание параметров, используя статистические техники. Много таких моделей используется при ценообразовании, и они дают действительно неплохую прибыль, среди них линейная модель, лог-линейная модель, логитная модель [2]. Во всех эти моделях считается, что функция спроса постоянна, но в реальности она не постоянна, поэтому мы будет этом учитывать. Функция спроса на время t для вещи i задается уравнением:

$$d_{i,t}(p_i) = f_{i,t}\left(\frac{p_i}{p_{i,t-1}}\right)^{\gamma_{*,i}}, \tag{1}$$

где $f_{i,t}$ - прогноз спроса для товара i на день t если цена на него $p_{i,t-1}$, $\gamma_{*,i}$ - эластичность товара i.
$\gamma_{*,i}$ "показывает как спрос товара i меняется относительно его цены.
Если $p_{i,t}$ близко к $p_{i,t-1}$, то можно произвести следующую аппроксимацию:

$$d_{i,t}(p_i) \approx f_{i,} + (p_i - p_{i,t-1})\frac{f_{i,t}\gamma_{*,i}}{p_{i,t-1}} \tag{2}$$

С такой аппроксимацией, оценка "эластичности"может быть сведена к линейной регрессии, которую можно решать, используя обыкновенный метод наименьших квадратов. Далее вместо того, чтобы смотреть на все данные, мы будем смотреть только на недавние (примерно 1-2 месяца) для оценки "эластичности".
Оценка эластичности гораздо сложнее прогноза спроса, так как товар имеет только k конкретных цен в последние несколько месяцев, поэтому для оценки "эластичности"для этого товара, мы должны использовать все k точек.

Обычно покупатели покупают товары не по одному, а несколько сразу, кладя их в корзину, именно такую ситуацию мы и будем в дальнейшем рассматривать. Прибыль от товара i в корзине $\mathbb{B}$ может быть оценена следующим образом:

$$Rev_{i,t}(p_{i,t}) = p_{i,t} \times d_{i,t}(p_{i,t}) \approx p_{i,t}(f_{i,t} + (p_{i,t} - p_{i,t-1})\frac{f_{i,t}\gamma_{*,i}}{p_{i,t-1}}) = \frac{p_{i,t}^2 f_{i,t}\gamma_{*,i}}{p_{i,t-1}} - p_{i,t}f_{i,t}\gamma_{*,i} + p_{i,t}f_{i,t} \tag{3}$$

Пусть $p = [p_1, p_2, \ldots, p_{|\mathbb{B}|}]$ - вектор цен на товары в корзине $\mathbb{B}$. Оптимизируем, используя оценку на величины $\gamma_{*,i},\ f_{i,t}$.

$$p_t = argmax_p \sum_{i \in \mathbb{B}} \frac{p_i^2 f_{i,t}\gamma_{*,i}}{p_{i,t-1}} - p_i f_{i,t}\gamma_{*,i} + p_i f_{i,t} \tag{4}$$

$\Pi_0(\gamma_*) = N(\mu_0, \Sigma_0)$
$l(R_t, Rev_t, \gamma_*) = N(R_t, Rev_t, \sigma^2)$
$Rev_t(p, \gamma_*, f_t) = \sum_{i \in \mathbb{B}} \frac{p_i^2 f_{i,t}\gamma_{*,i}}{p_{i,t-1}} - p_i f_{i,t}\gamma_{*,i} + p_i f_{i,t}$
$\Pi_t(\gamma_*) \propto \Pi_{t-1}(\gamma_*) N(R_t, Rev_t, \sigma^2)$

Алгоритм 1 Max-Rev

0: **input** Корзина $\mathbb{B}$, содержащая B товаров, период времени T, в течении которого хотим максимизировать прибыль
• выбирается вектор $\gamma_* \neq 0$ (игрок его не знает)
For $t = 1, \dots, T$
• игрок "играет"вектор p_t для корзины $\mathbb{B}$
• игрок получает стохастический выигрыш R_t, такой что, $\mathbb{E}[R_t] = Rev_t(p_t, \gamma_*, f_{1,t}, \dots, f_{B,t})$
End for

Алгоритм 2 Max-Rev-TS

0: **input** Корзина $\mathbb{B}$, содержащая B товаров, период времени T, в течении которого хотим максимизировать прибыль
• выбирается вектор $\gamma_* \neq 0$ (игрок его не знает)
• инициализируется априорное $\Pi_0(\gamma_*)$
For $t = 1, \dots, T$
• Сэмплируем γ_t из Π_{t-1}
• Используя прогноз спроса получаем $f_{i,t}$ для всех $i \in \mathbb{B}$
• Решаем Max-Rev оптимизационную задачу с $f_{i,t}$, $p_{i,t-1}$ и γ_t как оценку γ_*, получаем p_t
• Применяем цену p_t, получаем выигрыш R_t
• Изменяем распределение $\Pi_t(\gamma_*)$ по формуле, приведенной выше
End for

Теорема: Пусть γ_* - истинный вектор эластичностей. Тогда $||\mathbb{E}(\gamma_* - \mu_t)|| \to 0$, при $t \to +\infty$.

Доказательство:

$$\Sigma_t^{-1} = \Sigma_{t-1}^{-1} + \frac{\theta_t \theta_t^T}{\sigma^2} = \sum_{i=1}^{t} \frac{\theta_i \theta_i^T}{\sigma^2};$$

Рассмотрим такой раунд τ, что матрица $\Sigma_\tau^{-1} = \sum_{i=1}^{\tau} \frac{\theta_i \theta_i^T}{\sigma^2}$ невырождена.

Тогда $\forall\ t > \tau \hookrightarrow$:

$$\mathbb{E}(\gamma_* - \mu_t) = \gamma_* - \mathbb{E}[\mu_t] = \gamma_* - \mathbb{E}\big[\big(\Sigma_{t-1}^{-1} + M_t^{-1}\big)^{-1}\big(\Sigma_{t-1}^{-1}\mu_{t-1} + \frac{R_t - \bar{R}_t}{\sigma^2}\theta_t\big)\big] =$$

$$= \gamma_* - \big(\Sigma_{t-1}^{-1} + M_t^{-1}\big)^{-1}\big(\Sigma_{t-1}^{-1}\mathbb{E}[\mu_{t-1}] + \frac{\theta_t}{\sigma^2}\mathbb{E}[R_t] - \frac{\bar{R}_t\theta_t}{\sigma^2}\big) =$$

$$= \gamma_* - \big(\Sigma_{t-1}^{-1} + \frac{\theta_t\theta_t^T}{\sigma^2} + \lambda I\big)^{-1}\big(\Sigma_{t-1}^{-1}\mathbb{E}[\mu_{t-1}] + \frac{\theta_t}{\sigma^2}(\theta_t^T\gamma_* + \bar{R}_t - \bar{R}_t)\big) =$$

$$= \gamma_* - \big(\Sigma_{t-1}^{-1} + \frac{\theta_t\theta_t^T}{\sigma^2}\big)^{-1}\big(\big(\Sigma_{t-1}^{-1} + \frac{\theta_t\theta_t^T}{\sigma^2}\big)\gamma_* + \Sigma_{t-1}^{-1}\big(\mathbb{E}[\mu_{t-1}] - \gamma_*\big)\big) =$$

$$= \gamma_* - \gamma_* - \big(\Sigma_{t-1}^{-1} + \frac{\theta_t\theta_t^T}{\sigma^2}\big)^{-1}\Sigma_{t-1}^{-1}\big(\mathbb{E}[\mu_{t-1}] - \gamma_*\big)\big) =$$

$$= -\left(\Sigma_{t-1}^{-1} + \frac{\theta_t\theta_t^T}{\sigma^2}\right)^{-1}\Sigma_{t-1}^{-1}\left(\mathbb{E}[\mu_{t-1}] - \gamma_*\right)\big) = \left(\Sigma_{t-1}^{-1} + \frac{\theta_t\theta_t^T}{\sigma^2}\right)^{-1}\Sigma_{t-1}^{-1}\left(\gamma_* - \mathbb{E}[\mu_{t-1}]\right)\big).$$

$$\left(\Sigma_{t-1}^{-1} + \frac{\theta_t\theta_t^T}{\sigma^2}\right)^{-1}\Sigma_{t-1}^{-1} = \Big(\sum_{i=1}^{t}\frac{\theta_i\theta_i^T}{\sigma^2}\Big)^{-1}\sum_{i=1}^{t-1}\frac{\theta_i\theta_i^T}{\sigma^2} = \Big(\sum_{i=1}^{t}\theta_i\theta_i^T\Big)^{-1}\sum_{i=1}^{t-1}\theta_i\theta_i^T.$$

Обозначим $x_t = \gamma_* - \mathbb{E}[\mu_t],\ x_{t-1} = \gamma_* - \mathbb{E}[\mu_{t-1}],\ A_t = \sum\limits_{i=1}^{t}\theta_i\theta_i^T,\ A_{t-1} = \sum\limits_{i=1}^{t-1}\theta_i\theta_i^T.$

Тогда $x_t = A_t^{-1}A_{t-1}x_{t-1} \Rightarrow A_t x_t = A_{t-1}x_{t-1}$

$c = ||A_t x_t|| \geqslant \lambda_{min}^{(t)}||x_t||$, где $\lambda_{min}^{(t)}$ - минимальное собственное число матрицы A_t, c - константа $\Rightarrow$

$$||x_t|| \leqslant \frac{c}{\lambda_{min}^{(t)}}. \tag{5}$$

$A_t = A_{t-1} + \dfrac{\theta_t\theta_t^T}{\sigma^2};$

$\forall\ x:\ ||x|| = 1 \hookrightarrow\ A_t x = A_{t-1}x + \dfrac{\theta_t\theta_t^T}{\sigma^2}x.$

Рассмотрим скалярное произведение $x^T A_t x$:

$x^T A_t x = x^T A_{t-1}x + x^T\dfrac{\theta_t\theta_t^T}{\sigma^2}x \geqslant \lambda_{min}^{(t-1)} + \dfrac{(\theta_t^T x)^2}{\sigma^2}\ \Rightarrow$

$||A_t x|| \geqslant x^T A_t x \geqslant \lambda_{min}^{(t-1)}\quad \forall\ x:\ ||x|| = 1\ \Rightarrow$

$$\lambda_{min}^{(t)} \geqslant \lambda_{min}^{(t-1)}, \tag{6}$$

причем равенство достигается только в случае, когда вектор θ_t перпендикулярен $x_{min}^{(t-1)}$ - собственному вектору матрицы A_{t-1}, соответствующему собственному числу $\lambda_{min}^{(t-1)}$.

Тогда в силу (4), (5) получаем, что $\lambda_{min}^{(t)} \to +\infty$ при $t \to +\infty\ \Rightarrow ||x_t|| = ||\gamma_* - \mathbb{E}[\mu_t]|| \to 0$ при $t \to +\infty$.

$\square$.

3 Заключение

В этой статье мы описали алгоритм, который максимизирует функцию прибыли. Пока еще не продемонстрировали, но продемонстрируем, что он работает лучше "пассивных"алгоритмов, которые используются в качестве стандартного алгоритма ценообразования, так как "пассивные"алгоритмы показывают хорошие результаты лишь кратковременно, у нашего же алгоритма такого явления нет.

Литература

[1] Keskin N Bora Harrison, J Michael and Assaf Zeevi. Bayesian dynamic pricing poli- cies: Learning and earning under a binary prior distribution. *Management Science*, 58(3):570–586, 2012.

[2] Kalyan T Talluri and Garrett J Van Ryzin. The theory and practice of revenue management. *Springer Science Business Media*, 68, 2006.

Поступила в редакцию