ЕМ-АЛГОРИТМ ДЛЯ РЕЙТИНГ-СИСТЕМ

Сергей Николенко

СПбГУ — Санкт-Петербург

05 ноября 2020 г.

Random facts:

- 5 ноября в Великобритании — Guy Fawkes' Night; в эту ночь отмечается провал Порохового заговора, когда группа католиков-заговорщиков попыталась взорвать Парламент Великобритании в Лондоне в ночь на 5 ноября 1605 года
- 5 ноября 1935 г. регулярное троллейбусное движение открылось в Киеве, 5 ноября 1952 г. — во Владимире, 5 ноября 1964 г. — в Чернигове
- 5 ноября 1898 г. филиппинские националисты на острове Негрос восстали против испанского господства и основали Республику Негрос, которая просуществовала три года и стала протекторатом США
- 5 ноября 1917 г. Тихон был избран патриархом Московским, а Владимир Ленин инициировал Октябрьскую революцию
- 5 ноября 1940 г. Франклин Рузвельт стал первым и единственным президентом США, избранным на третий срок
- 5 ноября 1996 г. президенту России Борису Ельцину сделали операцию шунтирования на сердце; обязанности Президента РФ в это время исполнял Виктор Черномырдин

Рейтинг-системы, обучающиеся ММ-алгоритмами

Частичные сравнения

- Рейтинг-система — это модель, которая ранжирует участников (игроков) в единый линейный порядок по данным сравнений небольших подмножеств этих игроков (турниров).
- Более того, результаты турниров зашумлены (отчасти случайны).
- Соответственно, и применяются они в таких ситуациях (пример: контекстная реклама в Bing).

Рейтинг Эло

- Первая известная рейтинг-система, основанная на байесовском подходе — это рейтинг Эло.
- Суть модели:
  - сила игры шахматиста в одной партии — случайная величина;
  - *рейтинг* — это ожидание этой величины; мы пытаемся оценить это ожидание;
  - исходная модель Эло — нормальное распределение силы игры вокруг рейтинга.

- Значит, сила игры в конкретной партии распределена как

$$p(x) = N(x; s, \beta) = \frac{1}{\beta\sqrt{2\pi}} e^{-\frac{1}{2\beta^2}(x-s)^2}.$$

- Сила игры задаётся двумя параметрами: средним s (собственно рейтингом) и дисперсией β^2.
- Эло предположил, что дисперсия β^2 постоянна (и даже от игрока не зависит), а среднее — это как раз рейтинг, который мы пытаемся оценить.

Рейтинг Эло

- Значит, математически говоря, мы ищем

$$\arg\max_{s,\beta^2} p(s,\beta^2 \mid D) = \arg\max_{s,\beta^2} \frac{p(D \mid s,\beta^2)p(s,\beta^2)}{p(D)} =$$
$$= \arg\max_{s,\beta^2} p(D \mid s,\beta^2)p(s,\beta^2).$$

- Как мы знаем, нормальное распределение является самосопряжённым, поэтому если сила игры нормально распределена вокруг рейтинга, то логично взять нормальное распределение как априорное для рейтинга.
- Таким образом, рейтинг игрока складывается из двух чисел: его среднего значения μ и дисперсии σ^2.
- Значение μ отображается в таблице рейтингов, а σ^2 показывает, насколько достоверна имеющаяся оценка.

Рейтинг Эло

- Предположим, что встречаются два игрока с некоторыми априорными распределениями на рейтинги $N(s_1; \mu_1, \sigma_1^2)$ и $N(s_2; \mu_2, \sigma_2^2)$.
- Тогда сила игры каждого из них в этой конкретной партии имеет распределение

$$p(x \mid \mu, \sigma) = \int_{-\infty}^{\infty} p(x \mid s) p(s \mid \mu, \sigma) ds =$$

$$= \int_{-\infty}^{\infty} N(x; s, \beta) N(s; \mu, \sigma) ds =$$

$$= \int_{-\infty}^{\infty} \frac{1}{\beta\sqrt{2\pi}} e^{-\frac{1}{2\beta^2}(x-s)^2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(s-\mu)^2} ds = N(x; \mu_x, \sigma_x).$$

то есть мы снова приходим к нормальному распределению, но с другими параметрами.

Рейтинг Эло

- Задача обучения заключается в том, чтобы после новой партии принять во внимание её результат и пересчитать рейтинги.
- Эло разработал специальные аппроксимации и очень простые алгоритмы для этого случая (через «ожидаемые очки в турнире»), чтобы каждый шахматист мог сам на калькуляторе свой рейтинг посчитать, но они нас сейчас не очень интересуют.
- Сейчас есть обобщения рейтинга Эло, мы поговорим о них позже.

Модели Брэдли--Терри

- Другой подход к рейтинг–системам — модели Брэдли–Терри (Bradley–Terry).
- Модель предполагает, что для участников $1, \dots, n$ можно подобрать такие рейтинги γ_i, $i = 1..n$, что вероятность победы участника i над участником j равна

$$p(i \text{ побеждает } j) = \frac{\gamma_i}{\gamma_i + \gamma_j}.$$

- Основная задача заключается в том, чтобы найти $\gamma = (\gamma_1, \dots, \gamma_m)$ максимального правдоподобия из имеющихся данных D.

Модели Брэдли--Терри

- Если принять априорное распределение равномерным, можно просто максимизировать правдоподобие

$$p(D|\gamma) = \prod_{i=1}^{m}\prod_{j=1}^{m}\left(\frac{\gamma_i}{\gamma_i+\gamma_j}\right)^{w_{ij}},$$

  где w_{ij} — то, сколько раз x_i обыграл x_j при их попарном сравнении ($w_{ii} = 0$ по определению).
- Взяв логарифм, будем максимизировать

$$l(\gamma) = \sum_{i=1}^{m}\sum_{j=1}^{m}\left(w_{ij}\log\gamma_i - w_{ij}\log(\gamma_i+\gamma_j)\right).$$

- Максимизировать будем классическим MM-алгоритмом (minorization–maximization), фактически вариационным приближением.
- Заметим, что
$$1 + \log \frac{x}{y} - \frac{x}{y} \leq 0.$$

Упражнение. Докажите это.

- Рассмотрим вспомогательную функцию

$$Q(\gamma, \gamma^{(k)}) = \sum_{i,j} w_{ij} \left[\log \gamma_i - \frac{\gamma_i + \gamma_j}{\gamma_i^{(k)} + \gamma_j^{(k)}} - \log\left(\gamma_i^{(k)} + \gamma_j^{(k)}\right) + 1 \right].$$

Упражнение. Используя предыдущее неравенство, докажите, что $Q(\gamma, \gamma^{(k)}) \leq l(\gamma)$.

- Чтобы найти $\max_\gamma Q\left(\gamma, \gamma^{(k)}\right)$, можно просто взять производные $\frac{\partial Q}{\partial \gamma_l}$:

$$\frac{\partial Q}{\partial \gamma_l} = \sum_{i,j} w_{ij} \left[\frac{\delta_{il}}{\gamma_i} - \frac{\delta_{il} + \delta_{jl}}{\gamma_i^{(k)} + \gamma_j^{(k)}} \right] =$$

$$= \frac{1}{\gamma_l} \sum_j w_{lj} - \sum_j \frac{w_{lj}}{\gamma_i^{(k)} + \gamma_j^{(k)}} - \sum_i \frac{w_{il}}{\gamma_i^{(k)} + \gamma_j^{(k)}}.$$

Модели Брэдли--Терри

- Если w_l — общее количество побед игрока l ($w_l = \sum_j w_{lj}$), и N_{ij} — количество встреч между игроками i и j ($N_{ij} = w_{ij} + w_{ji}$), получаем

$$\frac{w_l}{\gamma_l} - \sum_j \frac{N_{lj}}{\gamma_i^{(k)} + \gamma_j^{(k)}} = 0.$$

- В результате правило пересчёта на одной итерации выглядит так:

$$\gamma_l^{(k+1)} := w_l \left[\sum_j \frac{N_{lj}}{\gamma_i^{(k)} + \gamma_j^{(k)}} \right]^{-1}.$$

- Получили алгоритм оценки рейтингов. Правда, он пока работает только для ситуации, когда игроки встречаются один на один и выигрывают или проигрывают.
- Но даже в шахматах бывают ничьи. Как их учесть?

Модели Брэдли--Терри

- Если возможны ничьи, их вероятность можно описать дополнительным параметром $\theta > 1$, и это приводит к модели, в которой

$$\begin{aligned} p(i \text{ побеждает } j) &= \frac{\gamma_i}{\gamma_i + \theta\gamma_j}, \\ p(j \text{ побеждает } i) &= \frac{\gamma_j}{\theta\gamma_i + \gamma_j}, \\ p(i \text{ и } j \text{ играют вничью}) &= \frac{(\theta^2 - 1)\gamma_i\gamma_j}{\left(\gamma_i + \theta\gamma_j\right)\left(\theta\gamma_i + \gamma_j\right)}. \end{aligned}$$

Модели Брэдли--Терри

- Аналогично можно вводить другие обобщения.
- Например, если результат может зависеть от порядка элементов в паре (скажем, команды проводят «домашние» матчи и «гостевые»), можно ввести дополнительный параметр θ, характеризующий, насколько большое преимущество дают «родные стены», и рассмотреть модель с

$$p(i \text{ побеждает } j) = \begin{cases} \frac{\theta\gamma_i}{\theta\gamma_i+\gamma_j}, & \text{если } i \text{ играет дома,} \\ \frac{\gamma_j}{\theta\gamma_i+\gamma_j}, & \text{если } j \text{ играет дома.} \end{cases}$$

- Можно даже обобщить на случай, когда в одном турнире встречаются несколько игроков: пусть перестановка π подмножества игроков $A = \{1, \ldots, k\}$ (результат турнира) имеет вероятность

$$p_A(\pi) = \prod_{i=1}^{k} \frac{\gamma_{\pi(i)}}{\gamma_{\pi(i)} + \gamma_{\pi(i+1)} + \ldots + \gamma_{\pi(k)}}.$$

Модели Брэдли--Терри

- Можно показать, что такая модель эквивалентна весьма естественной «аксиоме Люса»: для любой модели, в которой вероятности игроков обыграть друг друга в любой паре не равны нулю, для любых подмножеств игроков $A \subset B$ и любого игрока $i \in A$

$$p_B(i \text{ побеждает}) = $$
$$= p_A(i \text{ побеждает}) p_B(\text{побеждает кто-то из множества } A).$$

- Но для крупных турниров это перестаёт работать; и совсем трудно что-то осмысленное сделать, если игроки соревнуются не поодиночке, а в командах.

Спасибо!

Спасибо за внимание!