

Методы матричной факторизации и их применение в рекомендательных системах

Сергей Александрович Попов, группа 422

Кафедра статистического моделирования
Математико-механический факультет
Санкт-Петербургский Государственный Университет

Научный руководитель — к. ф.-м. н., доцент П. В. Шпилёв

Рецензент — к. ф.-м. н., доцент А. Г. Берлинков

8 июня 2022

Задача рекомендательной системы: отбор наиболее привлекательных для пользователя объектов.

Коллаборативная фильтрация — набор методов определения пользовательских предпочтений на основе данных о предпочтениях большого количества пользователей.

Методы матричной факторизации — подкласс методов коллаборативной фильтрации, основывающийся на разложении матрицы оценок, оставляемых пользователями объектам рекомендаций.

Пользовательская обратная связь:

- Явная (например, пользовательские оценки)
- Неявная (история покупок или факт просмотра фильма)

Функция предпочтений (scoring function)

- $\hat{y}_\theta : U \times I \rightarrow \mathbb{R}$ — функция предпочтений
- $i \in I$ — индекс объект рекомендаций
- $u \in U$ — индекс пользователя
- $\theta \in \Theta$ — параметризация функции предпочтений

Векторное представление

- $\hat{y}_\theta(u, i) = \langle \phi_\theta(u), \psi_\theta(i) \rangle$ — функция предпочтений, заданная скалярным произведением
- $\phi_\theta(u) : U \rightarrow \mathbb{R}^d, \psi_\theta(i) : I \rightarrow \mathbb{R}^d$
- $\phi_\theta(u)$ — d -мерное представление пользовательского профиля
- $\psi_\theta(i)$ — d -мерное представление объекта рекомендаций

Матрица предпочтений — содержит однородную обратную связь по объектам рекомендаций.

$$R_{(u,i)} = \begin{cases} y(u,i), & (u,i) \in S \\ 0, & (u,i) \notin S \end{cases}$$

Матрицы представлений

- $\hat{y}_\theta(u,i) = \langle \mathbf{w}_u, \mathbf{h}_i \rangle$, где \mathbf{w}_u и \mathbf{h}_i — строки матриц W и H соответственно.
- $W \in \mathbb{R}^{U \times d}$ — матрица представлений пользователей.
- $H \in \mathbb{R}^{I \times d}$ — матрица представлений объектов.
- $\theta = [W, H]$ — параметризация функции предпочтений.
- $R \approx WH^T$ — приближение матрицы предпочтений.

Цель: приближение R матрицами меньшей размерности.

Задача: поиск матриц W и H , наилучшим образом приближающих матрицу предпочтений R с учетом регуляризации параметров:

$$\arg \min_{W, H} \sum_{u, i: r_{ui} \neq 0} (r_{ui} - \langle \mathbf{w}_u \mathbf{h}_i \rangle)^2 + \lambda \sum_{i=1}^m \sum_{j=1}^d w_{uj}^2 + \lambda \sum_{i=1}^n \sum_{j=1}^d h_{ij}^2,$$

где m — число пользователей, а n — число объектов.

Проблема «холодного старта»: Если для заданного объекта $i \in I$ отсутствует обратная связь, то такой объект невозможно рекомендовать.

$$i \in I : r_{ui} = 0 \quad \forall u \in U$$

- 1 Изучить строение рекомендательных систем и распространенные методы матричной факторизации, используемые в них.
- 2 Реализовать рекомендательную систему на основе набора данных MovieLens 1M (Harper и др. 2006).
- 3 Провести сравнительный анализ нескольких алгоритмов матричной факторизации, используемых в рекомендательных системах.
- 4 Предложить модификацию одного из распространенных методов матричной факторизации, демонстрирующую лучшие результаты на наборе данных.
- 5 Разработать и реализовать подход для решения проблемы «холодного старта».

Используемый набор данных: MovieLens 1M (Harper и др. 2006)

- 1 млн. оценок на 3953 фильмов от 6141 пользователей.
- Только явная обратная связь: **оценки от 1 до 5.**

Доступные данные:

- ID пользователя,
- ID фильма,
- Пользовательская оценка фильма,
- Название фильма,
- IMDB ID фильма.

Дополнительные (внешние) данные:

- Изображение постера фильма.

Функция потерь:

$$L(\theta) := \sum_{(u,i)} c_{ui} (p_{ui} - \langle \mathbf{w}_u, \mathbf{h}_i \rangle)^2 + \lambda \sum_{u=1}^m k_u \sum_{j=1}^d w_{uj}^2 + \lambda \sum_{i=1}^n l_i \sum_{j=1}^d h_{ij}^2,$$

где k_u — число объектов с оценкой пользователя $u \in U$, l_i — число пользователей, оценивших объект $i \in I$, $c_{ui} = \alpha r_{ui}$ — значимость оценки r_{ui} , α — гиперпараметр масштаба.

$$c_{ui} = \alpha r_{ui} \qquad p_{ui} = \text{sign}(r_{ui})$$

Метод покоординатного спуска (Ну и др. 2008):

- 1 Случайная инициализация матриц W и H .
- 2 При фиксированных параметрах W , поиск локального минимума H .
- 3 При фиксированных параметрах H , поиск локального минимума W .
- 4 Повтор шагов 2 и 3 до сходимости.

Функция потерь:

$$L(\theta) := \sum_{(u,i)} c_{ui} (p_{ui} - (\langle \mathbf{w}_u, \mathbf{h}_i \rangle + \mathbf{b}_i^{item} + \mathbf{b}_u^{user} + \mu))^2 \\ + \lambda \sum_{u=1}^m k_u \sum_{j=1}^d w_{uj}^2 + \lambda \sum_{i=1}^n l_i \sum_{j=1}^d h_{ij}^2,$$

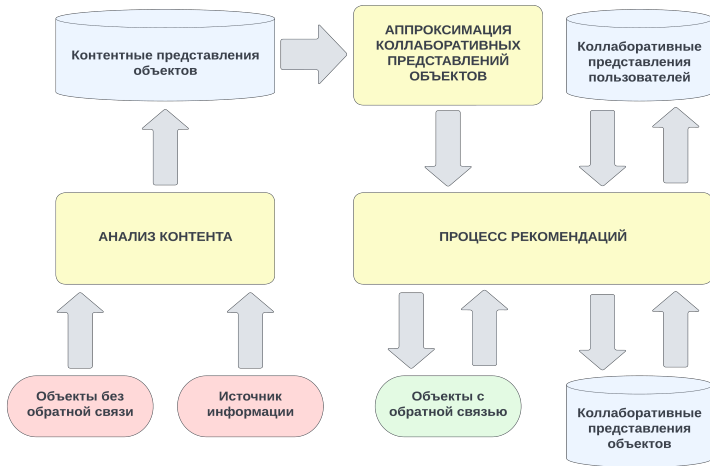
где k_u — число объектов с оценкой пользователя $u \in U$, а l_i — число пользователей, оценивших объект $i \in I$, \mathbf{b}^{user} и \mathbf{b}^{item} — обучаемые вектора пользовательского и объектного отклонения предпочтений, μ — среднее значение p_{ui} .

Проблема: коллаборативная рекомендательная система не может рекомендовать объекты, для которых отсутствует пользовательская обратная связь.

Решение: оценка коллаборативного представления $\mathbf{q}_i \in Q$ объекта $i \in I$ по внешним данным $\mathbf{p}_i \in P$, отражающим свойства объекта, при помощи отображения $\Phi : P \rightarrow Q$. На практике, $P \subseteq \mathbb{R}^k$ и $Q \subseteq \mathbb{R}^d$.

Гибридная рекомендательная система позволяет совмещать представления объектов, полученные на основе пользовательских оценок и при помощи анализа свойств объекта (контента).

Гибридные рекомендательные системы



Блок-схема гибридной рекомендательной системы, использующей внешние данные.

Метрика схожести $\omega : P \times P \rightarrow \mathbb{R}_+$ сопоставляет двум вещественным векторам из пространства P некоторый коэффициент схожести.

Примеры:

$$\text{dot}(\mathbf{p}_1, \mathbf{p}_2) = \mathbf{p}_1 \cdot \mathbf{p}_2^T,$$

$$\cos(\mathbf{p}_1, \mathbf{p}_2) = \frac{\mathbf{p}_1 \cdot \mathbf{p}_2^T}{\|\mathbf{p}_1\| \cdot \|\mathbf{p}_2\|},$$

$$\text{corr}(\mathbf{p}_1, \mathbf{p}_2) = \frac{\sum_{k=1}^n (p_{1k} - \bar{p}_1) (p_{2k} - \bar{p}_2)}{\sqrt{\sum_{k=1}^n (p_{1k} - \bar{p}_1)^2 \cdot \sum_{k=1}^n (p_{2k} - \bar{p}_2)^2}},$$

Оценка $\hat{\mathbf{q}}_i$ с использованием косинусного расстояния :

$$\hat{\mathbf{q}}_i = \frac{\sum_{j \in B_i^k} \cos(\mathbf{p}_i, \mathbf{p}_j) \cdot \mathbf{q}_j}{\sum_{j \in B_i^k} \cos(\mathbf{p}_i, \mathbf{p}_j)}.$$

ResNet (He и др., 2015) — сверточная нейронная сеть, предварительно обученная классификации на наборе данных ImageNet (Deng и др., 2009).

Адаптация нейронной сети:

- Последние слои нейронной сети были заменены полносвязными.
- До-обучение задаче регрессии на основе коллаборативных представлений фильмов из обучающей выборки.
- При до-обучении нейронной сети использовались параметры:

Количество слоев	Число параметров	Скорость обучения
34	21 852 992	0.004028
50	25 655 360	0.002815
101	44 647 488	0.002481
152	60 291 136	0.001202

iBALS показал средний прирост **Recall@50** в размере **1.5%**:

Размерность Итерации	160			240		
	20	40	80	20	40	80
iALS	0.542	0.546	0.547	0.543	0.547	0.549
iBALS	0.551	0.554	0.556	0.552	0.555	0.557

Recall@50 для метода ближайших соседей (**kNN**) и **ResNet** на объектах без обратной связи в тренировочной выборке:

Число соседей	1	2	3	4	5
kNN	0.0515	0.0549	0.0569	0.0576	0.0582

Число слоев	34	50	101	152
ResNet	0.1247	0.1291	0.1411	0.1458

Сравнение iBALS и iALS

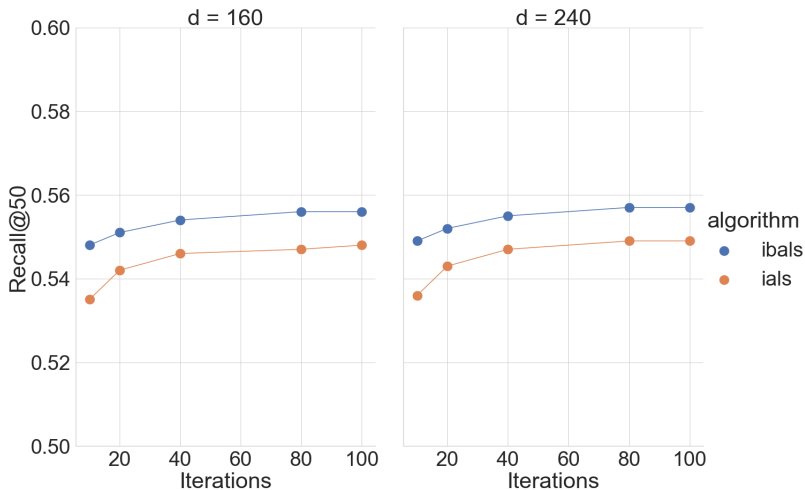


График **Recall@50** для **iALS** и **iBALS** в зависимости от размерности представлений и количества итераций при обучении.

- Изучена теория и архитектура рекомендательных систем.
- Реализована система рекомендаций для сравнения алгоритмов матричной факторизации и методов решения проблемы холодного старта на наборе данных MovieLens 1M.
- В результате сравнения алгоритмов было продемонстрировано превосходство iBALS над iALS.
- При помощи нейронных сетей была реализована гибридная рекомендательная система, продемонстрировавшая $\text{Recall@50} = 0.1458$ и $\text{Precision@50} = 0.16484$ на объектах, для которых отсутствовала обратная связь в обучающей выборке.