

基于机器学习的电子商务推荐系统的研究与应用

杨 媛, 彭定涛

贵州大学数学与统计学院, 贵州 贵阳

收稿日期: 2024年9月9日; 录用日期: 2024年9月26日; 发布日期: 2024年11月25日

摘 要

随着互联网的快速发展, 新型的商业运营模式——电子商务使得我们的生活越来越便捷。本文所考虑的推荐系统是电子商务技术中的一种, 其是利用电子商务网站向客户提供商品信息和建议, 帮助用户决定应该购买什么产品, 模拟销售人员帮助客户完成购买过程。协调过滤推荐技术是推荐系统技术中的一种, 其基本思想是用户可以根据兴趣进行分类, 类似的用户有着非常相似的利益, 可以通过协作用对目标用户接收信息, 过滤其他用户使用的建议, 且其算法一般可以分为基于记忆和基于模型两类。本文主要研究推荐系统中协调过滤技术的基于模型的算法, 即矩阵补全问题在推荐系统中的应用。在本文中, 我们构造了矩阵补全问题的非凸连续松弛模型, 并运用加速的迭代阈值算法求解模型, 然后运用真实数据检验模型和算法在推荐系统中的有效性。

关键词

电子商务, 推荐系统, 协调过滤, 矩阵补全问题

Research and Application of the E-Commerce Recommender System Based on Machine Learning

Yuan Yang, Dingtao Peng

School of Mathematics and Statistics, Guizhou University, Guiyang Guizhou

Received: Sep. 9th, 2024; accepted: Sep. 26th, 2024; published: Nov. 25th, 2024

Abstract

With the rapid development of the Internet, the new business operation model—electronic

文章引用: 杨媛, 彭定涛. 基于机器学习的电子商务推荐系统的研究与应用[J]. 电子商务评论, 2024, 13(4): 4701-4707.
DOI: 10.12677/ecl.2024.1341694

commerce makes our life more and more convenient. The recommender system considered in this paper is one of electronic commerce technologies, which uses e-commerce websites to provide customers with commodity information and suggestions, help users decide what products to buy, and simulate sales personnel to help customers complete the purchase process. Collaborative filtering recommended technology is one of the recommender system technologies, its basic idea is that the user can be classified according to interest, similar users have very similar interests, can receive information through collaboration to the target user, filtering other users use suggestions, and the algorithm can be divided into two categories based on memory and based on the model. This paper focuses on the model-based algorithm of collaborative filtering technique in recommender system, namely the application of matrix completion problem in recommender system. In this paper, we construct a non-convex continuous relaxation model for the matrix completion problem, and use the accelerated iterative threshold algorithm to solve the model, and then use the real data to test the effectiveness of the model and the algorithm in the recommender system.

Keywords

E-Commerce, Recommender System, Collaborative Filtering, Matrix Completion Problem

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着互联网的快速发展,电子产品已成我们日常生活中的必需品,我们可以通过它来完成购物和旅游策划等活动。而像这种在互联网开放的网络环境下,基于客户端或者服务端应用方式,买卖双方不谋面地进行各种商贸活动,实现消费者的网上购物、商户之间的网上交易和在线电子支付以及各种商务活动、交易活动、金融活动和相关的综合服务活动的一种新型的商业运营模式称为电子商务[1]。电子商务的存在使得消费者可以通过网络在网上购物、网上支付,节省了客户与企业的时间和空间,大大提高了交易效率。并且在消费者信息多元化的时代,可以通过足不出户的网络渠道,了解世界各地想知道的一切相关信息,然后再享受决策带来的乐趣,已经成为消费者习惯,方便消费者需求。而正是由于电子商务活动不受时间,空间的限制,可以随时随地在网上交易;减少商品流通的中间环节,节省大量的开支,大大降低商品流通和交易的成本,使其具有较高的使用率,所以其技术的研究越来越受到人们的关注。

本文所考虑的推荐系统是电子商务技术中的一种,其是利用电子商务网站向客户提供商品信息和建议,帮助用户决定应该购买什么产品,模拟销售人员帮助客户完成购买过程。个性化推荐就是根据用户的兴趣爱好,购买行为等,向用户推荐其所感兴趣的信息与商品的过程,其主要包含以下步骤:收集用户相关数据,如购买历史,浏览记录,用户喜好兴趣等;利用推荐技术对数据进行处理分析,确定推荐的内容;将确定好的结果推送用户,利于用户决策。常用的推荐技术[2]:基于内容推荐技术,其是建立在项目的内容信息上作出推荐的,而不需要依据用户对项目的评价意见,更多地需要用机器学习的方法从关于内容的特征描述的事例中找到用户的兴趣资料;协同过滤推荐技术,其利用用户的历史喜好信息计算用户之间的距离,然后利用目标用户的最近邻居用户对商品评价的加权评价来预测目标用户对特定商品的喜好程度,系统从而根据这一喜好程度来对目标用户进行推荐;基于关联规则推荐技术,其是以关联规则为基础,把已购商品作为规则头,规则体为推荐对象;基于效用推荐技术,其是在用户需要和可选集之间匹配的评估之上,通过计算商品对用户的效用来作出推荐;基于知识推荐技术,其不依赖

于用户对商品的评分数据量, 而是通过推断用户的需求与偏好来作出推荐; 最后一种较为常见的推荐技术就是将上述推荐技术组合在一起。而正是由于推荐系统的存在, 我们可以在海量的信息和产品中, 提取所需数据, 节省了大量的时间与精力的同时也提高了效率。

本文主要考虑的是推荐系统中的协调过滤推荐技术。其基本思想是用户可以根据兴趣进行分类, 类似的用户有着非常相似的利益, 可以通过协作用对目标用户接收信息, 过滤其他用户使用的建议[3]。协调过滤推荐技术的算法一般可以分为基于记忆和基于模型两类。基于记忆的算法是根据系统中所有被打过分的产品信息给出预测, 基于模型的算法是收集打分的相关数据进行学习并判断用户行为模型, 从而对某个产品进行预测打分。上述两种方法的不同主要在于基于模型的算法不是基于一些启发规则进行预测计算, 而是基于对已有数据应用统计和机器学习得到的模型进行预测[4]。

本文主要研究推荐系统中协调过滤技术的基于模型的算法的一种, 即矩阵补全问题在推荐系统中的应用。矩阵补全问题[5]是指将不完整的数据表示为一个具有低秩性质(数据本身结构的高度线性相关性)或者近似低秩的矩阵, 再根据矩阵中已知元素将低秩矩阵中缺失元素补全, 从而得到一个完整矩阵的过程。而其在推荐系统中常用于解决协调过滤中数据稀疏(数据缺失)问题, 即矩阵补全的元素代表每个用户对每个产品或选项的评分, 而由于用户相关数据过多, 容易造成评分的缺失, 这时我们就可以利用矩阵补全问题及相关算法对缺失数据进行补全。一般的矩阵补全问题模型如下所示:

$$\min_{X \in R^{m \times n}} \text{rank}(X) \quad \text{s.t. } X_{i,j} = M_{i,j}, (i,j) \in \Omega \quad (1)$$

其中 $\text{rank}(X) = \|\sigma(X)\|_0 = \sum_{i=1}^r \sigma_i^0(X)$ 是矩阵 $X \in R^{m \times n}$ 的秩函数, $X, M \in R^{m \times n}$ 是一个具有一些可用抽样项的未知矩阵, Ω 是已知元素的位置集合。然后由于秩函数是非凸不连续函数, 故求解问题(1)是困难的, 所以我们采取运用 Capped-L1 范数(非凸但连续)松弛上述模型(1)得到如下模型:

$$\min_{X \in R^{m \times n}} F(X) := f(X) + \lambda \Phi(X) \quad (2)$$

其中 $\Phi(X) := \sum_{i=1}^r \phi(\sigma_i(X))$, $\phi(t) = \min\left\{1, \frac{t}{v}\right\}$, $v > 0$ 是矩阵 $X \in R^{m \times n}$ 的 Capped-L1 范数, $\lambda > 0$ 是正则参数

用来权衡拟合项与正则项, f 是损失函数。对于模型(2)我们将考虑运用加速的迭代阈值算法进行求解, 并用真实数据检验我们的模型与算法对推荐系统的效用。

2. 模型与算法

2.1. 符号说明

对任意矩阵 $X \in R^{m \times n}$, $\sigma_i(X)$ 表示矩阵 $X \in R^{m \times n}$ 的第 i 个大的奇异值, 且 $\sigma_1(X) > \sigma_2(X) > \dots > \sigma_r(X) > 0$, r 表示矩阵 $X \in R^{m \times n}$ 的秩。矩阵 $X \in R^{m \times n}$ 的富比尼范数为 $\|X\|_F = \left(\sum_{i,j} X_{i,j}^2\right)^{1/2} = \left(\sum_{i=1}^r \sigma_i^2(X)\right)^{1/2}$ 。 L 表示函数 $f(X)$ 梯度 Lipschitz 连续的 Lipschitz 常数。

2.2. 模型

本文考虑模型(2)中损失函数为最小二乘函数($f(X) = \frac{1}{2} \|P_\Omega(X - M)\|_F^2$)的矩阵补全问题模型, 即:

$$\min_{X \in R^{m \times n}} F(X) := f(X) + \lambda \Phi(X) = \frac{1}{2} \|P_\Omega(X - M)\|_F^2 + \lambda \Phi(X),$$

其中 $M \in R^{m \times n}$ 是观测矩阵, Ω 是指标 (i, j) 的集合, P_Ω 是在稀疏矩阵子空间上的正交投影。

2.3. 算法

本文考虑加速的迭代阈值算法, 即将 Nesterov's 加速[6]与奇异值收缩阈值算法[7]结合, 并对外推系数采取自适应更新[8] [9], 具体如下:

算法 1: 加速的迭代阈值算法(ASVST)

初始化: 选择参数 $\lambda > 0$, $tol > 0$, $\beta_0 \in (0, \beta_{\max}]$, $\beta_{\max} < 1$, $\mu > L$, $Y_0 = X_0 \in R^{m \times n}$, $t > 1$ 。并令 $k = 0$ 。

步骤 1: 通过 $G^k = Y^k - \frac{1}{\mu} \nabla f(Y^k)$ 计算 G^k 及其奇异值分解(SVD):

$$G^k = U^k \text{Diag}(\sigma(G^k)) V^{k^T}.$$

步骤 2: 计算 $X^k \in S_{\mu}^{\lambda \Phi}(G^k)$ 。

步骤 3: 计算 $T^k = X^k + \beta_k(X^k - X^{k-1})$ 。

步骤 4: 如果 $F(T^k) \leq F(X^k)$, 计算

$$Y^{k+1} = T^k;$$

$$\beta^{k+1} = \min\{t\beta^k, \beta_{\max}\}.$$

否则

$$Y^{k+1} = X^k;$$

$$\beta^{k+1} = \frac{\beta^k}{t}.$$

步骤 5: 若 $\frac{\|X^k - X^{k-1}\|_F}{\max\{\|X^k\|_F, 1\}} < tol$ 则令 $Y^* := Y^{k+1}$; 其他设 $k := k + 1$ 且返回步骤 1。

3. 应用

下面运用真实数据实验来检验运用加速的迭代阈值算法的矩阵补全问题在推荐系统的效果。即用户在日常生活娱乐或在为旅游做攻略时的选择, 我们通过在电商平台上对用户历史数据的分析, 推送用户其感兴趣的相关信息, 为用户提升效率, 在此之中我们规定对于每个用户只要选项中任意一个的评分超过该用户的所有评分平均数就可以推送给用户。具体步骤如下:

首先我们从数据库(下载网址: <https://archive.ics.uci.edu/>)中提取 200 个用户关于教堂, 度假村, 海滩, 公园, 博物馆, 商场, 动物园, 观景点, 古迹, 园林的 10 个选项的评分得到一个 200×10 的矩阵 M , 随机选择其中 10% 的元素给其赋值为 0 生成需要补全的矩阵 N 。然后利用加速的迭代阈值算法补全矩阵得到矩阵 Y 。最后通过比较需要补全的哪一项前后推送是否一致检验算法效果(允许 10% 的误差)。

在实验中, 我们的目的是基于 N 中一些已知元素, 恢复秩为 r 的目标矩阵 N 。为了观测 N 中已知元素的位置, 我们通过采样率 ($sr = |\Omega|/mn$) 对 N 进行采样, 其中 m, n 是 N 的维度。令自由度 $df := r(m + n - r)$, 则 $os := |\Omega|/df$ 是过采样率。下面给出实验中所用到的参数数据: $X_0 = P_{\Omega}(M)$, $\lambda_0 = \|P_{\Omega}(M)\|$, $\lambda_{final} = 1e^{-6}$, $v = 6$, $\mu = 2$, $\alpha = 5$, $\eta = 0.9$, $tol = 1e^{-4}$, $maxiter = 1e^3$, 其中 η 表示 λ_k 折射率, 即 $\lambda_{k+1} = \eta\lambda_k$ 。

由于实验数据过多, 我们仅截取其中一部分对我们实验进行说明。对从数据库中提取关于用户 - 选项 - 评分的 20×10 矩阵即已知矩阵如表 1 所示, 然后对已知矩阵进行随机选择 10% 赋值为 0 生成需要

矩阵补全的矩阵如表 2 所示,最后利用我们所提出的模型与算法对矩阵进行补全得到的结果如表 3 所示。

Table 1. A 20×10 matrix about the user-option-score extracted from the database
表 1. 从数据库中提取关于用户 - 选项 - 评分的 20×10 矩阵

用户	教堂	度假村	海滩	公园	博物馆	商场	动物园	观景点	古迹	园林
1	1.26	1.3	1.31	5	5	5	2.73	0.84	5	1.25
2	1.25	1.28	1.3	1.34	5	5	2.73	0.79	5	1.24
3	1.25	1.27	1.3	1.34	5	5	2.73	0.78	5	1.24
4	1.25	1.27	1.3	1.34	5	5	2.73	0.79	5	1.24
5	1.25	1.27	1.3	1.34	5	5	2.73	0.79	5	1.24
6	1.25	1.27	1.3	1.33	5	5	2.73	0.77	5	1.24
7	1.25	1.27	1.32	1.33	4.77	5	2.73	0.77	5	1.24
8	1.25	1.27	1.32	1.34	4.76	5	2.73	0.78	5	1.24
9	1.25	1.27	1.32	1.33	4.76	5	2.74	0.78	4.77	1.24
10	1.25	1.27	1.32	1.33	4.77	5	2.74	0.78	4.77	1.24
11	1.33	1.27	1.32	1.33	4.76	5	2.74	0.7	0.72	0.75
12	0.77	5	1.32	1.33	4.09	5	2.75	0.69	0.72	0.74
13	0.76	1.34	1.32	1.33	4.1	5	2.83	0.68	0.71	0.74
14	0.76	1.33	1.32	1.33	4.77	5	2.83	0.68	0.88	0.73
15	1.25	1.33	1.32	1.33	4.11	5	2.83	0.68	0.83	0.73
16	0.75	1.28	1.32	1.33	4.12	5	2.83	0.67	0.81	5
17	1.24	1.28	1.32	1.33	4.11	5	2.85	0.68	0.7	0.72
18	1.24	1.35	1.32	1.36	4.11	4.79	2.76	5	0.69	5
19	1.24	1.25	1.33	1.37	4.11	4.77	5	0.7	0.77	4.78
20	1.24	5	1.33	1.33	4.1	4.77	5	0.66	0.74	4.77

Table 2. The matrix that requires matrix completion
表 2. 需要矩阵补全的矩阵

用户	教堂	度假村	海滩	公园	博物馆	商场	动物园	观景点	古迹	园林
1	1.26	1.3	1.31	5	5	0	2.73	0	5	1.25
2	1.25	1.28	1.3	1.34	5	5	2.73	0.79	5	1.24
3	1.25	1.27	1.3	1.34	5	5	2.73	0.78	5	1.24
4	1.25	1.27	1.3	1.34	0	5	2.73	0.79	5	1.24
5	0	1.27	1.3	1.34	5	0	2.73	0.79	5	1.24
6	1.25	1.27	1.3	1.33	5	5	2.73	0	5	1.24
7	1.25	1.27	1.32	1.33	4.77	5	2.73	0.77	5	1.24
8	1.25	1.27	1.32	1.34	4.76	5	2.73	0.78	5	1.24
9	1.25	1.27	1.32	1.33	4.76	5	2.74	0.78	4.77	1.24

续表

10	1.25	1.27	1.32	1.33	4.77	5	0	0.78	4.77	1.24
11	1.33	1.27	1.32	0	4.76	5	0	0.7	0.72	0.75
12	0.77	5	1.32	1.33	4.09	5	2.75	0	0.72	0.74
13	0.76	1.34	1.32	1.33	4.1	5	2.83	0.68	0.71	0
14	0.76	1.33	0	0	4.77	0	2.83	0.68	0.88	0.73
15	1.25	0	0	1.33	4.11	5	2.83	0.68	0.83	0.73
16	0.75	1.28	1.32	1.33	4.12	5	0	0	0.81	5
17	1.24	1.28	1.32	1.33	4.11	5	2.85	0.68	0.7	0.72
18	1.24	1.35	1.32	1.36	4.11	4.79	2.76	5	0.69	5
19	1.24	1.25	1.33	1.37	4.11	4.77	5	0.7	0.77	4.78
20	1.24	0	1.33	1.33	4.1	4.77	5	0	0.74	4.77

Table 3. The matrix obtained after matrix completion

表 3. 矩阵补全后得到的矩阵

用户	教堂	度假村	海滩	公园	博物馆	商场	动物园	观景点	古迹	园林
1	1.0957	1.1739	0.9480	1.7618	2.1524	3.6272	3.4836	0.1836	2.4994	1.3192
2	1.5144	1.1081	2.3031	1.7522	5.9959	5.7154	2.6852	0.6849	5.1332	1.6086
3	0.9546	1.0427	1.3073	1.2926	2.7813	3.4258	2.7691	0.5726	2.4807	1.1355
4	1.2921	0.3358	1.5276	1.7394	5.1457	4.7712	1.4371	0.6957	5.0798	1.4109
5	1.3432	1.4725	1.9002	1.3500	4.8883	4.7226	2.5342	0.2806	3.8477	1.1494
6	1.3797	0.4051	1.1993	2.3423	4.1705	5.0604	2.8288	0.6980	4.9961	1.8832
7	1.5841	1.1566	0.8288	2.1551	4.2217	4.8088	2.5495	0.8415	4.7664	1.3575
8	1.1376	1.2764	1.2887	1.6233	2.8051	3.7455	3.1995	0.6312	2.6396	1.2282
9	1.7454	1.1800	1.1192	3.1790	3.4259	5.7629	4.9027	0.7061	4.7562	2.2933
10	1.0068	-0.0531	1.7706	1.1231	5.3434	3.7892	-0.3163	0.5996	4.4445	0.8488
11	0.2158	2.2239	1.3246	-0.2245	-0.2689	0.9700	4.2102	0.9720	-1.8191	0.3966
12	0.8698	2.3481	2.5512	0.1894	3.6675	2.9500	2.6760	0.9930	0.7487	0.4021
13	0.1150	1.3346	0.4016	0.0053	-1.0213	-0.0053	2.2361	0.1530	-1.6884	-0.0053
14	-0.2062	-0.6409	0.1495	-0.2516	0.6308	0.1454	-0.2530	1.3733	0.7919	0.3491
15	0.9421	1.2832	1.2373	1.2294	2.3450	3.0381	2.8372	0.8285	1.8994	0.9511
16	2.4942	6.1992	4.1465	2.5812	3.9787	7.3468	11.3160	1.6931	0.7967	2.0341
17	0.8378	1.1494	0.4929	1.8747	-0.1689	2.0767	3.7236	0.7198	0.5145	1.0630
18	2.3621	1.4770	0.5402	5.2363	2.1703	7.9207	9.1618	0.8181	6.0765	3.9207
19	4.5561	10.9623	6.4893	6.0140	4.4659	13.1914	22.6434	2.9207	0.8382	4.5176
20	3.8611	6.4884	3.9014	6.1273	4.6251	12.2153	16.2467	-1.3198	4.6412	4.4662

下面进行具体说明为何我们的模型与算法适用于推荐系统：我们从表 2 中可以看出第五位用户关于

教堂与商场的的数据未知, 我们通过模型与算法将表 2 表示的矩阵进行补全得到第五位用户关于 10 个选项的数据如下所示: 1.3432, 1.4725, 1.9002, 1.3500, 4.8883, 4.7226, 2.5342, 0.2806, 3.8477, 1.1494。根据以上数据可以得到他们的平均分为 2.34887, 从而可以得到对第五位用户需要推送商场不用推送教堂的相关信息。然后我们可以从表 1 中得到第五位用户关于 10 个选项的数据如下所示: 1.25, 1.27, 1.3, 1.34, 5, 5, 2.73, 0.79, 5, 1.24。根据以上数据可以得到他们的平均分为 2.492, 从而可以得到对第五位用户需要推送商场不用推送教堂的相关信息。根据以上结果我们知道补全前后结果一致。因此可以得出我们提出的矩阵补全问题的模型可以用于推荐系统, 并且我们的算法是有效的。

4. 总结

在本文中我们构造了损失函数为最小二乘函数和罚函数为 Capped-L1 范数的矩阵补全问题的非凸连续松弛模型, 并通过 Nesterov's 加速与奇异值收缩阈值算法结合建立加速的迭代阈值算法, 然后运用在数据库中所得的真实数据检验了我们的模型和算法在推荐系统中的有效性。

基金项目

国家自然科学基金(12261020), 贵州省科技计划(黔科合基础 ZK[2021]一般 009)和贵州省高层次留学人才创新创业择优资助重点项目([2018]03)。

参考文献

- [1] 黄京华. 电子商务教程[M]. 北京: 清华大学出版社有限公司, 1999.
- [2] 刘平峰, 聂规划, 陈冬林. 电子商务推荐系统研究综述[J]. 情报杂志, 2007, 26(9): 46-50.
- [3] 曾春, 刑春晓, 周立柱. 个性化服务技术综述[J]. 软件学报, 2002, 13(10): 1952-1961.
- [4] 刘建国, 周涛, 汪秉宏. 个性化推荐系统的研究进展[J]. 自然科学进展, 2009, 19(1): 1-15.
- [5] 刘浩洋. 最优化: 建模, 算法与理论[M]. 北京: 高等教育出版社, 2020.
- [6] Nesterov, Y. (2018) Lectures on Convex Optimization. Springer. <https://doi.org/10.1007/s10898-023-01322-8>
- [7] Zhang, X., Peng, D. and Su, Y. (2023) A Singular Value Shrinkage Thresholding Algorithm for Folded Concave Penalized Low-Rank Matrix Optimization Problems. *Journal of Global Optimization*, **88**, 485-508.
- [8] Lin, Z., Li, H. and Fang, C. (2020) Accelerated Optimization for Machine Learning. Springer.
- [9] Yang, X. and Xu, L. (2022) Some Accelerated Alternating Proximal Gradient Algorithms for a Class of Nonconvex Nonsmooth Problems. *Journal of Global Optimization*, **87**, 939-964. <https://doi.org/10.1007/s10898-022-01214-3>