Collaborative Filtering Based on Orthogonal Non-negative Matrix Factorization

Meng Huang¹, *, JiHong OuYang², Chen Wu¹, a, Liu Bo¹, b

¹ School of Software, Jilin University Software Engineering, Changchun, 130001, China.
² School of Computer Science, Jilin University, Changchun, 130001, China.
*Corresponding author email: huangmeng17@mails.jlu.edu.cn,
a wuchen17@mails.jlu.edu.cn, b boliu17@mails.jlu.edu.cn

Abstract. In order to study the application of orthogonal nonnegative matrix factorization (ONMF) in collaborative filtering, so as to improve the recommendation accuracy of collaborative filtering, firstly, the theoretical knowledge of the existing matrix factorization model was analyzed and discussed, and then linearization correction was added to the matrix factorization model. Secondly, orthogonal constraints were added to the traditional weighted non-negative matrix factorization model (WNMF) to make the algorithm decompose the original data into non-negative matrices. Finally, the NMF collaborative filtering algorithm based on unit factorization and graph regularization correction (RTGNMF) was proposed. Three models, RTGNMF, positive ONMF and WNMF, were compared in NMAE/RMSE on real simulated data sets. The results show that once Tikhonov is used to correct the parameters in NMF model in single graph, the RMSE value will continue to decline based on the parameter adjustment of D1 dataset. RTGNMF, ONMF and WNMF all change the recommendation performance of high-dimensional data to a certain extent. ONMF has higher recommendation accuracy than WNMF. The robustness and adaptability of RTGNMF, ONMF and WNMF decrease in turn. RTGNMF and ONME make up for the shortcomings of current collaborative filtering algorithms to a great extent, and have obvious advantages over traditional algorithms. Adding linear correction in the iteration process of matrix factorization can converge and oscillate progressively, and adding orthogonalization constraint can significantly improve the redundancy of data and effectively improve the recommendation accuracy of collaborative filtering.

1. Introduction
Collaborative filtering algorithm, based on the basic model of evaluation moments [1], combines the preferences of groups with similar interests and common experience to predict and recommend information of interest or valuable personal preferences [2]. Of course, this algorithm does not necessarily only filter information of interest to users, history information that users are not particularly interested in is also particularly important for detailed algorithms [3]. Therefore, collaborative filtering algorithm is to use this way to effectively evaluate the scoring data, so as to achieve accurate recommendation. Generally, collaborative filtering algorithm is divided into evaluation collaborative filtering and group filtering collaborative filtering. In specific applications, because of its excellent computing speed and robustness, collaborative filtering algorithm has become a major algorithm in e-commerce recommendation system [4]. Based on the research results of many scholars, three algorithms, RTGNMF, ONMF and WNMF, are proposed to detect the effectiveness of the algorithm. In fact, the
collaborative filtering algorithm has a higher recommendation accuracy. Its essence can be summarized as the filling problem of matrix default values [5].

Collaborative filtering algorithm is the most important algorithm of recommendation system algorithm. Its advantage lies in filtering the information which is disorderly, low information quality and low value. By using similar similarity, clustering, classification and other methods, all kinds of similar information of data can be found, and then the potential historical information and evaluation are used to predict and assess things (including users and items) with high accuracy [6]. Thus, for collaborative filtering recommendation algorithm, compared with content-based filtering algorithm, because it can utilize potential hidden information (including data information and evaluation shared by Alibaba, No. 1 Store, Tencent and other platforms), the accuracy of its recommendation is beyond doubt [7]. With the development of recommendation system, collaborative filtering algorithm has always been one of the most important algorithms in the field of e-commerce. Since the discovery of collaborative filtering algorithm, it has been applied in various fields, and its popularity is self-evident. In recent years, with the development of e-commerce, people pay more and more attention to collaborative algorithms, and its development types are changing with each passing day. Essentially, recommendation algorithms based on collaborative filtering are usually based on machine learning and artificial intelligence [8]. Collaborative filtering algorithm based on these theories has higher recommendation accuracy, and can also get more diversified and innovative recommendation technology, which makes collaborative filtering recommendation algorithm succeed in recommendation system [9].

Based on the above background, the merits and demerits of various computational models were studied on the basis of non-negative matrix. To solve the problems of data redundancy, sparse matrix features and frequent operation of identification matrix in other models, a collaborative filtering algorithm based on orthogonal non-negative matrix factorization (ONMF) was proposed. From the experimental results, compared with other algorithm models, the effectiveness of the algorithm is improved significantly, which provides a new theoretical basis for future recommendation systems.

2. Method

2.1. Basic classification of recommendation algorithms

In order to solve the situation that the information exceeds the system handling ability, a recommendation system is created. The original intention of the recommendation system design is to solve the problem of information overload. Recommendation system refers to recommending personalized information to users according to their own information needs, browsing history and access traces combined with customers' interests [10]. The core of recommendation system is recommendation algorithm. The speed of recommendation algorithm determines the quality of this recommendation system. At present, there are many kinds of recommendation algorithms, and there is no uniform classification standard. Generally speaking, the commonly used recommendation algorithms are: content-based recommendation algorithm, collaborative filtering-based recommendation algorithm, association rule-based recommendation, utility-based recommendation, knowledge-based recommendation and various combination recommendation algorithms [11]. The specific classification of the recommended algorithm is shown in the following figure. The collaborative filtering algorithm is introduced and studied in detail, and other algorithms are also described in general, and compared with collaborative filtering algorithm.
2.2. Matrix decomposition mathematical model

Assuming that the given user set is $U$ and the given item set is $I$, the item matrix $R = [r_{1,i}, r_{2,i}, \ldots, r_{|U|,i}] \in \mathbb{R}^{I \times |U|}$ can be calculated according to the relevant formula. In this user's item matrix formula, $r_{u,i} \in R$ represents the evaluation of item $i$ by user $u$. For users, there are limited items in real life, which indicates that the characteristics of rating matrix $R$ are sparse \cite{12}. The scoring matrix $R$ is calculated by the cooperative filtering model based on the decomposition of ONMF. Therefore, on this basis, the prediction matrix is $\hat{R} = PQ$. Obviously, it can be seen that the scoring matrix $R$ is composed of two low rank factors, $P$ and $Q$, and $f$ is the characteristic dimension \cite{13}. Its mathematical model is to minimize the objective function and minimize the variance between the observation matrix and the prediction matrix. The general expression of its objective function, that is, the mathematical model, is as follows.

$$L(P,Q) = \|R - PQ\|^2$$  \hspace{5cm} (1)

The above information suggests that the feature of score matrix $R$ is sparse. Tikhonov regularization should be introduced in matrix decomposition to correct the feature sparse score matrix $R$. Through effective correction, the over-fitting phenomenon in the process of calculation can be well solved, thus improving the accuracy of calculation and improving the recommended efficiency. Based on the above reasons, the characteristic matrix is corrected. The method adopted in this paper is the two-norm method. The minimization objective function, i.e. the mathematical model, is recorded as $\arg\min L(P,Q)$, and its general expression is as follows.

$$L_{P,Q}(P,Q) = \|R - PQ\|^2 + \mu_p\|P\|_F^2 + \mu_q\|Q\|_F^2$$  \hspace{5cm} (2)

In the above formula, $\mu_p$ indicates the correction parameters of user characteristic matrix $P$, and $\mu_q$ refers to the correction parameters of item characteristic matrix $Q$. In specific calculation and application, the model can be simplified and easy to be calculated by making $\mu_p = \mu_q = \mu$. 

Figure 1. Classification of recommendation algorithms
2.3. NMF
ONMF was first used in computer. Its main feature is to represent the similarity of specific observation matrices. Essentially, NMF can compensate the negative part of the external value matrix by adjusting the learning rate, so as to find the best parameters and retain only the non-negative part of the eigenvalue matrix. In NMF with penalty term, parameter adjustment is very important. Under the condition of keeping iteration convergence, the two most important factors in designing NMF are the rationality of initialization parameters and the non-negative feature of matrix to be optimized.

By adding non-negative constraints to the general expression of the objective function, the loss function of the most basic collaborative filtering algorithm based on NMF can be obtained as follows:

$$\text{argmin}_{P, Q} \| R - PQ \|^2$$  \hspace{1cm} (3)

Without constraining the nonnegativity of the upper formula, the gradient descent method is used to solve the upper equation, and the following expression can be obtained:

$$p_{uk} \leftarrow p_{uk} \left( R k u \right) \left( PQ^T u, k \right)$$  \hspace{1cm} \hspace{1cm} (4)

$$q_{kj} \leftarrow q_{kj} \left( R P k \right) \left( PQ^T P k, i \right)$$

The gradient descent method is used to transform the loss function, eliminating the negative part of the function, so as to maintain the non-negativity of P and Q. NMF has a very important link, which is the non-negative training link. In this link, the collaborative filtering algorithm based on NMF has a high convergence. However, as previously explained, the R feature of scoring matrix is sparse, so the basic model mentioned above is not suitable for recommendation algorithm. It has been proposed that weighted non-negative matrix factorization (WNMF) be applied to collaborative filtering algorithm to solve the problem of sparse feature of score matrix R. The main methods to solve the problem of sparse feature of score matrix R by WNMF are point residual matrix multiplication and identification matrix.

WNMF model is seldom used in practical computer field, and its practicality is not very good even used. There are two reasons: on the one hand, the sparse feature of the score matrix R directly leads to the sparse problem of the identification matrix Y. The difficulty and complexity of the calculation are greatly enhanced by using the residual matrix multiplication and the identification matrix to solve the sparse problem; on the other hand, when iteratively updating, the entrance $r_{ui}$ must check whether its value is 0 every time. In other words, the sparsity of the score matrix R should be calculated before calculating the user u and item i.

2.4. ONMF
The sum of the score matrix R and the low rank evaluation matrix PQ is the Euclidean distance, and its expression formula is as follows.

$$e = \| R - PQ \|^2 = \sum_{(u,i) \in T} (r_{ui} - p_{u,i}q_{i})^2$$  \hspace{1cm} (5)

In the above formula, $p_{u,i}$ suggests the row vector, and $q_{i}$ indicates the column vector. In the final function expression, the inner integral of $p_{u,i}$ and $q_{i}$ can be separately expressed, and the following equation can be obtained.

$$\hat{r}_{ui} = p_{u,i} = \sum_{k=1}^{f} p_{u,k}q_{k,f}$$  \hspace{1cm} (6)

In the process of calculation, the orthogonal decomposition no longer calculates the identification matrix Y, so the calculation becomes much simpler. Moreover, the orthogonal decomposition only calculates and scores the user/item characteristics.
2.5. RTGNMF
The objective function of RTGNMF is shown below:

$$\arg\min_{P,Q} L(P,Q) = \|Q \cdot P\|^2 + \mu \|P\|^2 + \lambda \text{Tr}(Q L_v Q^T)$$

(7)

In the above formula, the regularization of the user characteristic matrix is corrected. \(\mu\) and \(\lambda\) are the parameters of the penalty items of the user characteristic matrix and the item graph, respectively. Only by adjusting the parameters of the penalty items reasonably can RTGNMF converge to the optimum point in the calculation process.

Lagrange multiplier method and KKT condition are used to expand the formulas and constrained optimization of the formulas is carried out. The derivatives of \(P\) and \(Q\) are calculated in time, and the following equations are obtained.

$$\frac{\partial L}{\partial P} = -(R - PQ)^T + \mu R$$

$$\frac{\partial L}{\partial Q} = -P^T (R - PQ) + \lambda L_v Q$$

(8)

Similarly, in the process of experiment, the equation is expanded by using unit strategy to calculate the known scoring elements, which can effectively solve the sparse feature problem of the scoring matrix \(R\).

3. Results and discussion

3.1. Parameter adjustment and data comparison
An important part of RTGNMF calculation is to adjust parameters. Only by adjusting reasonable parameters can RMSE/NMAE curve have no any overload and morbidity, so that RMSE/NMAE curve converges to the best point. Adjusting parameters is a very complex and computational work. The adjustment of parameters is completed under off-line conditions. Once the adjustment parameters are completed, it can consider using empirical values as test sets.

According to the parameter adjustment of D1 data set, once Tikhonov is used to correct in the single graph NMF model, the RMSE value will continue to decline. As shown in the table below, the parameters of RTGNMF and ONMF are summarized in the process of parameter adjustment. It is noteworthy that the parameters are recorded values when all the indexes reach the optimal conditions.

| Dataset | ONMF | RTGNMF |
|---------|------|--------|
| \(\lambda\) | \(\mu\) | \(\lambda\) |
| D1      | 0.48 | 0.04   | 0.03   |
| D2      | 5    | 0.09   | 0.09   |
| D3      | 52   | 0.02   | 0.04   |

Table 1. Parameter values of RTGNMF and ONMF
As shown in the figure above, abscissa 1, 2 and 3 represent D1, D2 and D3 data sets, respectively. From the figure, it is seen that NMAE/RMSE of RTGNMF, ONMF and WNMF is increasing in turn in the three data sets. This is because ONMF introduces graph correction. Therefore, the NMAE/RMSE value of ONMF is significantly lower than that of WNMF, and the effect is also significantly higher than that of WNMF. On the basis of NMF model, RTGNMF not only adopts memory-based recommendation algorithm, but also introduces Tikhonov correction to solve the over-fitting and ill-conditioned phenomena of RMSE/NMAE curve, which makes the effect of RTGNMF better than that of the other two models.

4. Conclusion

On the basis of nonnegative matrix, the advantages and disadvantages of various computing models are studied. In order to solve the problems of data redundancy, sparse matrix features and frequent operation of identification matrix in other models, a collaborative filtering algorithm based on ONMF is presented. At the same time, NMAE/RMSE of three models of RTGNMF, ONMF and WNMF is compared in real simulated data sets. The results show that these three algorithms change the recommendation performance in high-dimensional data to some extent. ONMF has higher recommendation accuracy than WNMF, and the robustness and adaptability of RTGNMF, ONMF and WNMF are reduced in turn. Although some experimental results have been achieved in this paper, there are still many shortcomings. In collaborative filtering algorithm, new problems and challenges are constantly emerging. For example,
in the market economy, the existing Internet industry and recommendation areas of commodity procurement will be extensively used to develop; in the market economy model, the optimal optimization model is used to produce the best recommendation results; in future research, the alternation of different new Internet technologies will bring new development areas. For the traditional recommendation system, even if the mathematical model is established, when the incentive information is too little, there will be similar over-fitting, overload and morbidity phenomena as mentioned in the paper. At the same time, how to ensure the robustness, computational ability, and adaptability of recommendation system and recommendation algorithm will become more and more important.

In conclusion, the collaborative filtering algorithm based on ONMF has great theoretical significance and broad application prospects. At the same time, due to the rapid development of collaborative filtering algorithm, the research on Internet and big data recommendation algorithm is also improving, which has a great space for development.

References
[1] Pompili F, Gillis, Nicolas, Absil, P.-A, et al. Two algorithms for orthogonal nonnegative matrix factorization with application to clustering[J]. Neurocomputing, 2014, 141(4), pp. 15-25.
[2] Kim S, Sael L, Yu H. A Mutation Profile for Top-k Patient Search Exploiting Gene-Ontology and Orthogonal Non-negative Matrix Factorization[J]. Bioinformatics, 2015, 31(22), pp. 3653-3659.
[3] Zhang F, Lu Y, Chen J, et al. Robust collaborative filtering based on non-negative matrix factorization and R1-norm[J]. Knowledge-Based Systems, 2016, 118(C), pp. 177-190.
[4] Bo Li, Guoxu Zhou, Cichocki, Andrzej. Two Efficient Algorithms for Approximately Orthogonal Nonnegative Matrix Factorization[J]. IEEE Signal Processing Letters, 2014, 22(7), pp. 843-846.
[5] Xiang C, Yuan Yuan, Cao Shuai, et al. A SEMG-Force Estimation Framework Based on a FastOrthogonal Search Method Coupled with Factorization Algorithms[J]. Sensors, 2018, 18(7), pp. 2238.
[6] XiuJun Zhang, Chang Liu. Generalized Discriminant Orthogonal Nonnegative Tensor Factorization for Facial Expression Recognition[J]. The Scientific World Journal, 2014, 2014(1), pp. 608158.
[7] Chen W S, Li Y, Pan B, et al. Fast Non-Negative Matrix Factorizations for Face Recognition[J]. International Journal of Pattern Recognition & Artificial Intelligence, 2017, 32(4), pp. 7.
[8] Guterman A, Herrero, Alicia, Thome, Néstor. New matrix partial order based on spectrally orthogonal matrix decomposition[J]. Linear & Multilinear Algebra, 2016, 64(3), pp. 362-374.
[9] Pan, Jinfeng. Measurement matrix optimization method based on matrix orthogonal similarity transformation[J]. Journal of Electronic Imaging, 2016, 25(3), pp. 033025.
[10] Meng Y, Shang, Ronghua, Jiao, Licheng, et al. Dual-graph regularized non-negative matrix factorization with sparse and orthogonal constraints[J]. Engineering Applications of Artificial Intelligence, 2018, 69, pp. 24-35.
[11] Nicolas S, Marjan A, Halandur N. B, et al. The successive projection algorithm as an initialization method for brain tumor segmentation using non-negative matrix factorization[J]. Plos One, 2017, 12(8), pp. e0180268.
[12] Zhang Z, Liao Shouyi, Zhang Hexin, et al. Bilateral Filter Regularized L2 Sparse Nonnegative Matrix Factorization for Hyperspectral Unmixing[J]. Remote Sensing, 2018, 10(6), pp. 816-.
[13] Huang R, Li Xiaorun, Zhao Liaoying. Nonnegative Matrix Factorization With Data-Guided Constraints For Hyperspectral Unmixing[J]. Remote Sensing, 2017, 9(10), pp. 1074.