International Journal of Control and Automation
Vol.7, No.8 (2014), pp. 79-92
http://dx.doi.org/10.14257/ijca.2014.7.8.08

Application and Research of Improved Probability Matrix Factorization Techniques in Collaborative Filtering

Zhijun Zhang[1,2,3] and Hong Liu[1,2]*

[1]School of Information Science and Engineering, Shandong Normal University, Jinan, Shandong, 250014, China
[2]Shandong Provincial Key Laboratory for Novel Distributed Computer Software Jinan, Shandong, 250014, China
[3]School of Computer Science and Technology, Shandong Jianzhu University, Jinan, Shandong, 250101, China
zhangzj@sdjzu.edu.cn, *lhsdcn@jn-public.sd.cninfo.net

Abstract

The matrix factorization algorithms such as the matrix factorization technique (MF), singular value decomposition (SVD) and the probability matrix factorization (PMF) and so on, are summarized and compared. Based on the above research work, a kind of improved probability matrix factorization algorithm called MPMF is proposed in this paper. MPMF determines the optimal value of dimension D of both the user feature vector and the item feature vector through experiments. The complexity of the algorithm scales linearly with the number of observations, which can be applied to massive data and has very good scalability. Experimental results show that MPMF can not only achieve higher recommendation accuracy, but also improve the efficiency of the algorithm in sparse and unbalanced data sets compared with other related algorithms.

Keywords: *Matrix Factorization, Collaborative Filtering, Recommendation system, SVD, PMF*

1. Introduction

Traditional collaborative filtering approaches can neither handle large data sets, nor solve the problem of data scarcity. In this paper, to solve the problems mentioned above, basic methods of matrix factorization are discussed, and the collaborative filtering (CF) technologies based on matrix factorization algorithms are deeply analyzed. Collaborative filtering is one of the most widely used techniques in recommendation systems, which filters some unrelated information using the degree of information association, and retains the useful parts to users. Tapestry, GroupLens, Ringo and Video Recommender are relatively early recommender systems [1].Tapestry is the first proposed recommender system based on collaborative filtering, in which the target users need to clearly list other users whose behavior are more similar to theirs. GroupLens is a kind of automated collaborative filtering recommender system based on user ratings for recommending movies and news. Recommender systems such as Ringo and Video usually recommend music and movies via e-mail. Breese [2] *et al.*, preceded an in-depth analysis of various collaborative filtering recommendation algorithms and their improvements. Amazon.com, Jester and WWW are some other recommender systems based on collaborative filtering [3].

ISSN: 2005-4297 IJCA

Currently, collaborative filtering algorithm is mainly classified into three categories: memory-based collaborative filtering, model-based collaborative filtering and hybrid recommendation. The most frequently used model in collaborative filtering recommender systems is the k-Nearest Neighbor (kNN), which includes two kinds of technologies of users-based recommendation and items-based recommendation [4]. Examples of model-based approaches include factor model [5], Bayesian classification model [6], clustering model [7] and graph model [8]. In the model-based approaches, training datasets are firstly used to learn a predictive model offline, and then the model is applied to online systems to make recommendations. The key of the algorithm is how to learn an effective predictive model [9].

The problem to be solved in this paper is mainly about data sparsity and scalability, which is ubiquitous in traditional collaborative filtering recommendation algorithms and accordingly a modified probability matrix factorization algorithm (MPMF) is presented. Firstly, basic method of matrix factorization (MF) was discussed deeply and the collaborative filtering technologies based on MF algorithm were analyzed extensively. The MF algorithms such as singular value factorization (SVD) and the probability matrix factorization (PMF) *etc.*, were analyzed from the aspects of their performance, and the existing problems of these algorithms were also stated. Secondly, the impact of user feature vectors dimensions and item feature vectors dimensions on the recommendation accuracy and recommendation efficiency was observed by means of experiments, as well as the consistency of the recommendation results of PMF algorithms both in the training set and data sets. Finally, the optimal parameters of MPMF algorithm were determined by experiments. Experimental results in public dataset Netflix indicated that MPMF algorithm was superior to traditional algorithms of collaborative filtering recommendation in both running efficiency and accuracy.

The remainder of this paper is organized as follows. In Section 2, we provide an overview of several major approaches for recommender systems and related work. Section 3 gives the definition of the fundamental matrix factorization model. Section 4 presents our work on MPMF. Section 5 gives experimental data sets and evaluation metrics. The results of the experiment are presented in Section 6, followed by the conclusion and future work in Section 7.

2. Related Work

The Netflix dataset is very sparse with its sparse degree of 1%. In order to solve the problem of data sparsity and algorithmic scalability in the collaborative filtering algorithms, the researchers at home and abroad have conducted a series of studies and proposed a variety of solutions. Chen [10] *et al.*, have conducted collaborative filtering recommendation of products that target users may be interested in combined with the revenue of recommendation results. Yang [11] *et al.*, put forward a collaborative filtering method based on inferences. Xue [12] *et al.*, proposed a method based on clustering in 2005, which divided users into k classes and chose all users in the nearest class from target users as neighbors and calculated the similarity between them; however the uncertainty of the number of clusters still remains a problem in this method. Tomoharu [13] *et al.*, recommended goods that users are most likely to purchase through collaborative filtering algorithm based on maximum entropy principle. Park [14] *et al.*, improve the recommendation accuracy of Yahoo! by combining collaborative filtering and search engine.

Recently an important research trend in recommender systems is the application of both latent semantic models and matrix factorization techniques in collaborative filtering systems. What they have in common is the complementing of the scoring matrix through dimensionality reduction methods. Matrix factorization techniques can provide more accurate recommended items to users through the reduction of dimensions of the sparse matrix. At present matrix factorization methods mainly include nonnegative matrix factorization (NMF) [15], singular value decomposition (SVD) [15], probability matrix factorization (PMF) and Bayesian probability matrix factorization (BPMF).

In 1999, both Lee and Seung published nonnegative matrix factorization algorithm in Nature [17], which immediately drew the attention of many researchers in collaborative filtering systems. The main idea of NMF is that high dimensional matrix can be decomposed into the product of two or more low-dimensional matrices in order to study the nature of high-dimensional matrix in low-dimensional space [18].

Sarwar *et al.*, published an article on the application of the singular value decomposition algorithm to the design of the recommender system in 2000 [19]. In SVD, the feature values of data set are denoted by singular value and are ranked according to their importance, which achieves the purpose of dimensionality reduction by means of discarding unimportant feature vectors. SVD algorithm did not get much attention in the field of recommender systems when it was proposed after a few years because of its two disadvantages- over-fitting and lack of precision. In 2006, Simon Funk proposed an improved SVD method in his Blog, which attracted the attention of the matrix factorization method throughout the academic field [20]. The matrix factorization method proposed by Simon Funk was subsequently called latent factor model (LFM) by Netflix Prize winner Koren. Matrix factorization can lead to better recommendation accuracy after many tests of KDDCup [21] and Netflix [22] competition.

3. The Definition of Fundamental Matrix Factorization Model

The notations of this paper as summarized in Table 1.

Table 1. Mathematical Notations

Notation	Description
N	Number of users
M	Number of items
R	User-item rating matrix of N users on M items
$U \in R^{D \times N}$	User’s latent matrix of feature vectors
$V \in R^{D \times M}$	Item’s latent matrix of feature vectors
U_i	Specific user’s latent feature vectors (column vector)
V_j	Item-specific latent feature vectors (column vector)
u	Current user
v	Current item
r_{ij}	Real rating by user i on item j
$\hat{r}_{ij}$	Prediction rating by user i on item j

In this paper, the matrix is denoted in capital letters (*e.g.*, R) and scalar is denoted in lowercase letters (*e.g.*, i, j). R^T represents the transpose matrix of R. R^{-1} represents the inverse matrix of R. The film is a concrete expression of item which can be considered to be concept equivalent.

Matrix factorization model mapped the collection of users and items to the joint latent factor space whose dimension is D, and the interaction between users and items is modeled through the inner product in the space [23]. Accordingly, each user u is associated with a vector p_u $\left(p_u \in R^D\right)$, and each item i is associated with a vector q_i $\left(q_i \in R^D\right)$. For a given user u, the elements of p_u measure the extent of interest the user has on each element of the item. For a given item i, the elements of q_i measure the extent to which the item possesses those factors. Both p_u and q_i can be either positive or negative. The marking $\hat{r}_{ui}$ of user u on item i can be modeled through the inner product of user feature vector p_u and item feature vector q_i, which can be estimated as follows:

$$\hat{r}_{ui} = p_u q_i^T \tag{1}$$

The major challenge is computing the mapping of each item and each user to factor vectors, in which feature vector q_i, p_u can be trained by optimizing the following loss function:

$$C(D) = \sum_{(u,i)\in D_A} (r_{ui} - p_u^T q_i)^2 + \lambda\left(\|p_u\|^2 + \|q_i\|^2\right) \tag{2}$$

Here, D_A is the training set. $(r_{ui} - p_u^T q_i)^2$ is the squared error between the predicted value and the actual score. $\lambda\left(\|p_u\|^2 + \|q_i\|^2\right)$ is used to prevent fitting of overtraining, and parameter λ is used to achieve cross-validation. Equations (2) is optimized through stochastic gradient descent and alternating least squares method.

4. The Improved Probability Matrix Factorization Algorithm

4.1. The Traditional Recommendation Algorithm model-PMF

In 2008 Salakhutdinov and Mnih put forward probability matrix factorization algorithm in NIPS [24] and discussed the deeper probability explanation of MF, predicting user's score from the perspective of probability and resulting in a probability matrix factorization (PMF) model as shown in Figure 1.

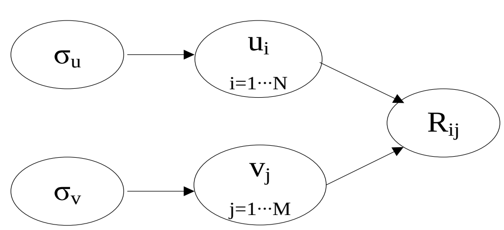


Figure 1. The Graphical Model for PMF

Suppose both user feature vector U_i and movie feature vector V_j conform to Gaussian distributions with mathematical expectation of 0.

$$p(U|\sigma_U^2) = \prod_{i=1}^{N} N(U_i|0, \sigma_U^2 I) \tag{3}$$

$$p(V|\sigma_V^2) = \prod_{j=1}^{M} N\left(V_j|0, \sigma_V^2 I\right) \tag{4}$$

User preference for the movie is the combination of a series of probability and we define the conditional distribution as following:

$$p(R|U, V, \sigma^2) = \prod_{i=1}^{N}\prod_{j=1}^{M}\left[N\left(R_{ij}|U_i^T V_j, \sigma^2\right)\right]^{I_{ij}} \tag{5}$$

Here, $N\left(x \mid \mu, \delta^2\right)$ is the probability density function of the Gaussian distribution with mean μ and variance σ^2, and I_{ij} is the indicator function with the value equal to 1 if user u_i rated movie v_j and equal to 0 otherwise.

The problem existing in PMF algorithm is that the running time of the program will get increasingly long and the recommendation efficiency will accordingly decrease significantly as the dimensions of user and item feature vectors increase. Simultaneously, the issue of solving sparsity of scoring matrix by the matrix factorization technique lies in that its basic idea is to approximate the original score matrix by using a low-rank one, and the approximation goal is to minimize the square error between the predict rated matrix and the original one, which is considerably time-consuming.

4.2. Improved Probability Matrix Factorization Algorithm—MPMF

PMF algorithm is a kind of matrix factorization technique based on probability. Matrix factorization tends to be transformed into an optimization problem with its local optimal solution by iteration. Experiment 1 (as seen in Section 6.2) showed that dimension D has a great impact on execution time. How to reduce the dimension of feature vectors while ensuring the accuracy of recommendation algorithm remains an issue of vital importance, since the running time that the user might accept could not be too long for lots of the recommendation systems. In real systems the key problem affecting the recommendation accuracy is to determine dimensions of both user feature vectors and item feature vectors, the value of which researchers usually give according to their experiences, while how to get an optimal value has been a hard problem. The solution of PMF recommendation algorithm is done by the gradient descent method, during which the rate of error declining becomes increasingly slow, as a result PMF requires more time of iterations and training.

In order to solve these problems in PMF algorithms we propose an improved probability matrix factorization algorithm named MPMF. The algorithm consists of the following four steps: maximizing the log-posterior probability, solving implicit feature vectors, normalization process and limiting the dimension of feature vectors. At last we compare our method MPMF with two popular methods: PMF and traditional collaborative filtering algorithms.

(1) The posterior probability is regarded as a target function to get its maximum value. The log of prediction equations (5) is given by

$$\ln p\left(U,V|R,\sigma^2,\sigma_V^2,\sigma_U^2\right)=-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{j=1}^{M}I_{ij}\left(R_{ij}-U_i^TV_j\right)^2-\frac{1}{2\sigma_U^2}\sum_{i=1}^{N}U_i^TU_i-\frac{1}{2\sigma_V^2}\sum_{j=1}^{M}V_j^TV_j$$
$$-\frac{1}{2}\left(\left(\sum_{i=1}^{N}\sum_{j=1}^{M}I_{ij}\right)\ln\sigma^2+ND\ln\sigma_u^2+MD\ln\sigma_V^2\right)+C \tag{6}$$

In which C is a constant independent of the parameters. Maximizing the value of equations (6) can be thought of as unconstrained optimization problems, consequently minimizing the value of equations (7) is equivalent to maximizing that of equations (6).

$$E=\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M}I_{ij}\left(R_{ij}-U_i^TV_j\right)^2+\frac{\lambda_U}{2}\sum_{i=1}^{N}\left\|U_i\right\|_{Fro}^2+\frac{\lambda_V}{2}\sum_{j=1}^{M}\left\|V_j\right\|_{Fro}^2 \tag{7}$$

Where, $\lambda_U=\sigma^2/\sigma_U^2$, $\lambda_V=\sigma^2/\sigma_V^2$, $\|\cdot\|_{Fro}^2$ denotes the Frobenius norm.

(2) The implied user feature vector and item feature vector are obtained based on gradient descent.

The local minimum of the objective function given by equations (7) can be gained by means of performing gradient descent in U and V.

Algorithm 1: Gradient Descent Algorithm

1) Selecting number of potential factors K, regularization parameter λ and learning rate η, initialing feature vector p_u and q_i. Typically, the initial value is some random number near the average score.
2) Calculating the partial derivative of C. The partial derivative of p_u and q_i are as follows:

$$\frac{\partial C}{\partial p_u}=-2\cdot q_i e_{ui}+2\lambda p_u \tag{8}$$

$$\frac{\partial C}{\partial q_i}=-2\cdot p_u e_{ui}+2\lambda q_i \tag{9}$$

Here, $e_{ui}\overset{\text{def}}{=}r_{ui}-q_i^Tp_u$.

3) Iterative updating of p_u and q_i:

$$p_u\leftarrow p_u+\alpha(e_{ui}q_i-\lambda\cdot p_u) \tag{10}$$

$$q_i\leftarrow q_i+\alpha(e_{ui}p_u-\lambda\cdot q_i) \tag{11}$$

4) Calculating new Probe RMSE. If the new value is smaller than the previous one, then repeating step (2) and step (3), otherwise the algorithm will be terminated.

(3) Normalization process：in order to map the rating 1,2,3,4,5 to the interval [0-1], we use the following function:

$$g(x) = 1/(1 + exp(-x)) \tag{12}$$

Therefore, the corresponding prediction equation is changed to equations (13).

$$p(R|U,V,\sigma^2) = \prod_{i=1}^{N}\prod_{j=1}^{M}\left[N\left(R_{ij}|g\left(U_i^T V_j\right),\sigma^2\right)\right]^{I_{ij}} \tag{13}$$

(4) Limiting the dimension of the feature vectors: determining the optimal value of dimension D through experiments and determining the relevant parameter values and number of iterations.

Normally we need to set fixed parameters in Algorithm 1 [25], and we can also set an argument $\max\ epoch$ --the maximum number of iterations based on experience. In the experiments of this paper, we set the following initial values: $\lambda = 0.01, \max\ epoch = 50$. The optimum dimension of feature vectors is obtained to be 10 through our experiment 2 (as seen in Section 6.3) and 3 (as seen in section 6.4), as a result we determine D=10 as its optimal value, which has been verified in the literature [26].

From the experiments we can get a conclusion that algorithm MPMF needs less running time than that of PMF, but it can get better prediction accuracy. Compared with other current recommendation algorithm, MPMF is only next to BayesMPF as far as the prediction accuracy is concerned; however it significantly reduces the training time [27].

5. Dataset and Metrics

5.1. Experiment Environment

The PC machine configuration in the experiment: Intel(R) Core(TM) i3-2120 CPU @3.30GHz, 1.96GB RAM. Operating system: Windows XP. Programming environment: Mathworks Matlab R2010a.

5.2. Dataset

Table 2. Netflix Prize

Awards	The winning teams	Test RMSE
Progress Prize 2007	KorBell	0.8723
Progress Prize 2008	BellKor in BigChaos	0.8627
Grand Prize 2009	BellKor’s Pragmatic Chaos	0.8567

The Netflix dataset is one of the most popular data sets in collaborative filtering algorithm. Netflix company established Netflix Prize in 2006. Table 2 summarizes the awards. This dataset was constructed to support participants in the Netflix Prize. The movie rating files contain over 100 million ratings from 480 thousand randomly-chosen, anonymous Netflix customers over 17 thousand movie titles. The ratings are on a scale from 1 to 5 (integral) stars. The date of each rating and the title and year of release for each movie id are also provided.

5.3. Metrics

We use the Root Mean Squared Error (RMSE) metrics to measure the prediction quality of our proposed approach in comparison with other collaborative filtering methods. RMSE is defined as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{|E^P|}\sum_{(u,i)\in E^P}(r_{ui} - \hat{r}_{ui})^2} \tag{14}$$

Where E^P denotes test set, $|E^P|$ denotes the number of tested ratings. The smaller the RMSE value is, the higher the prediction accuracy is.

6. Experimental Analysis

6.1. Experiment Scheme

In this section, we conduct four experiments to compare recommendation quality of our approach with other state-of-the-art collaborative filtering methods. The four experiments are as follows:

Experiment 1. Comparison of algorithm performance: we analyzed the effect of the feature vector dimension D on execution time of PMF by changing its value, which was completed in Section 6.2.

Experiment 2. We analyzed different impacts of dimension D on RMSE both in the training set and in the testing set, which was completed in Section 6.3.

Experiment 3. We analyzed the impact of the feature vector dimension D on RMSE by changing its value , which was completed in Section 6.4.

Experiment 4. Prediction accuracy comparison: we compared the recommendation accuracy of improved MPMF with basic Netflix algorithm, SVD algorithm and PMF algorithm, which was completed in Section 6.5.

The impacts of feature vector dimension D on PMF algorithm is illustrated through experiment 1. The optimal value of the dimension D is determined through experiment 2 and 3. And the recommendation accuracy of the algorithm MPMF is tested through experiment 4.

6.2. The Impacts of Dimension D on Running Time of PMF

We designed Experiment 1 to verify the impacts of dimension D of both the user feature vector and the item feature vector on execution time and the experimental results is shown in Figure 2. As can be seen from the figure, with the value of dimension D increases, the running time of PMF algorithm becomes increasingly long and the operating efficiency really declines substantially.

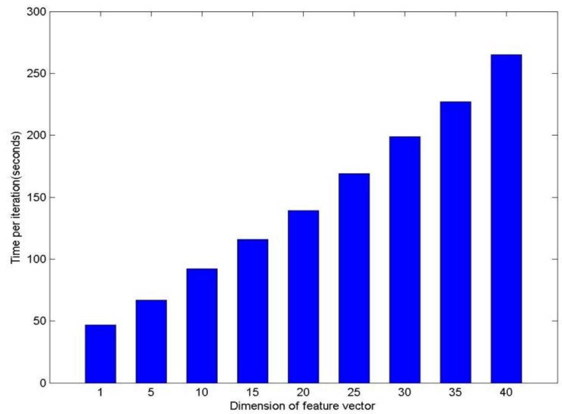
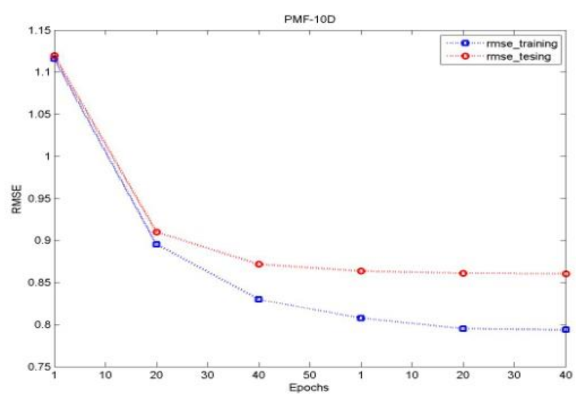


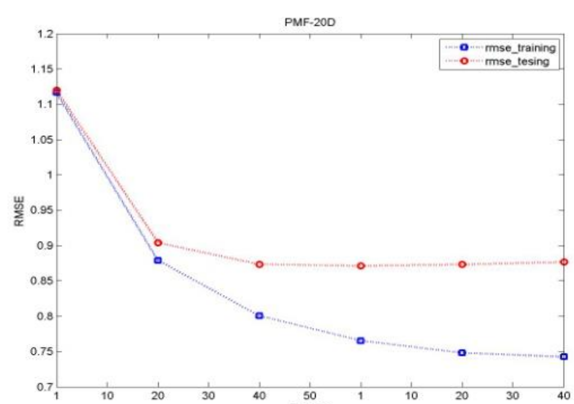
Figure 2. Running Times per Iteration of PMF

6.3. Comparison of RMSE in Training Set and Testing Set

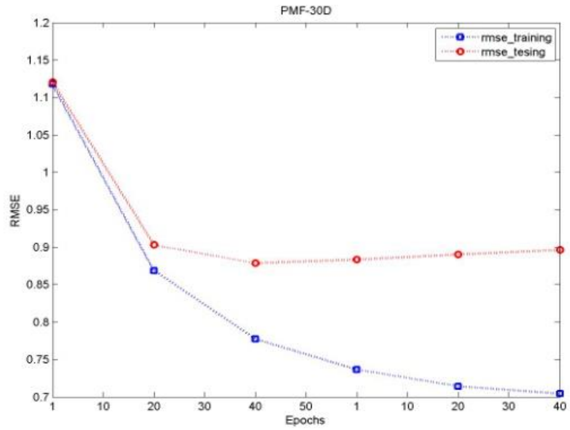
By means of changing the value of the feature vector dimension D, different effect about dimension D on RMSE in the training set and testing set is analyzed. The experiment result is shown in Figure 3 (a) to (d).



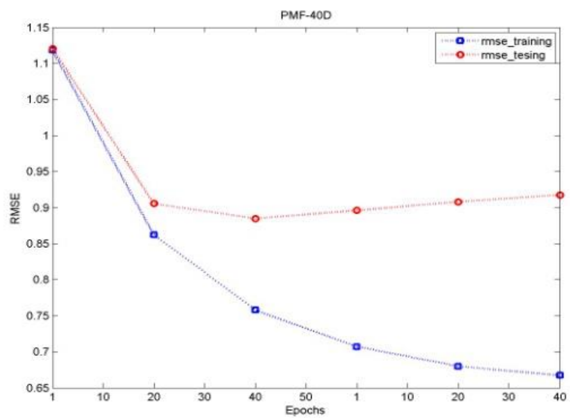
(a) Dimensionality=10



(b) Dimensionality=20



(c) Dimensionality=30



(d) Dimensionality=40

Figure 3. Deviation of RMSE in the Training Set and Testing Set

Figure 3 shows that adding the dimension of a feature vector can improve the recommendation accuracy of PMF algorithms. However, from the 4 figures, we could also come to a conclusion that when the dimensions of matrix feature vector increased to some extent (D>10) the deviation of RMSE between training set and testing set became increasingly large.

6.4. Impacts of Dimension D on Recommendation Precision

By changing the value of feature vector dimension D, the impact of dimension D on RMSE was analyzed and the experimental results were shown in Figure 4. As can be seen from Figure 4, with the increase of feature vectors dimension the values of RMSE grow bigger especially when iterations Epochs are larger than 10. Therefore, as the dimension of feature vectors raises to a certain extent the recommendation accuracy has actually lowered due to the increase of noise.

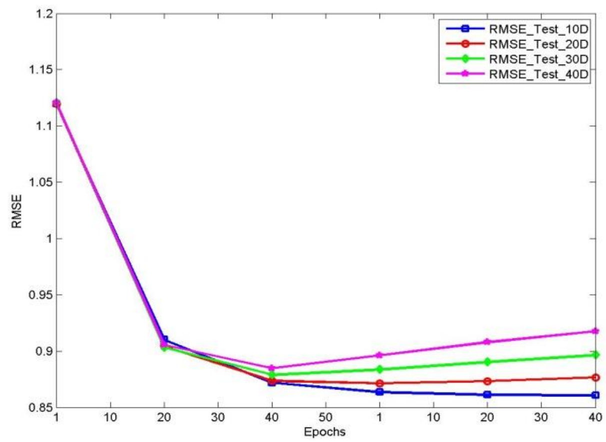


Figure 4. The Impact of Dimensionality D on the RMSE

6.5. Comparison of Recommendation Accuracy

We compared the recommendation accuracy of improved MPMF with Baseline of Netflix algorithm, SVD algorithm and PMF algorithm. Evaluation metrics is the RMSE and the value of the PMF algorithm is the average of the RMSE when D=5, 10, 20, 30, 40. Comparison results are shown in Figure 5 with abscissa indicating the iteration number of Epochs and ordinate representing RMSE.

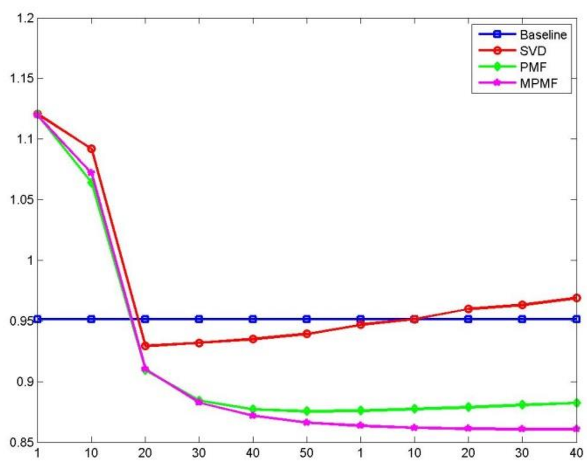


Figure 5. Performance Comparisons with other Approaches

From Figure 5, we can draw the following conclusions:

(1) The problem of over-fitting of SVD is quite serious which starts to over fit when epoch exceeds 10.

(2) The recommendation precision of MPMF algorithm (RMSE=0.8606) is about 7% higher than SVD algorithm (RMSE=0.9293).

(3) The recommendation precision of MPMF algorithm increases about 9.5% than baseline recommendation algorithms of Netflix system (RMSE=0.9293).

(4) The recommendation precision of MPMF algorithm (RMSE=0.8606) raises about 1.7% than PMF algorithm（RMSE=0.8753）.

As a result, MPMF algorithm is much better than the others regardless of the operating efficiency or the prediction accuracy in sparse matrix and unbalanced data sets.

6.6. Analysis of Time Complexity of MPMF

The time complexity of matrix factorization algorithm derives mainly from gradient descent, and the computational cost mainly comes from the objective function C and the corresponding gradient descent formula. Matrices U and V both are sparse. The time complexity of objective functions from equations (6) is $O(D \cdot n_U + D \cdot n_V)$, in which n_U and n_V represent the number of non-zero elements in matrix U and V respectively. Hence, the time complexity is $O(D \cdot n_U + D \cdot n_V)$ at each iteration. Algorithmic time complexity increases lineally with the growth of the number of observations in the sparse matrix, so the algorithm is suited to be used in recommender systems with mass data because of its better extensibility.

7. Conclusion

In this paper, aiming at solving defects of traditional recommender systems, firstly we gave a deep discussion of basic methods of matrix factorization, and then we put forward a kind of improved probability matrix factorization algorithm called MPMF, which has been turned out to be a better way of addressing data sparseness problem with the characteristics of easier programming and lower time complexity. Finally, we verified through experiments that the prediction accuracy of algorithm MPMF is quite high with good scalability on real data sets, and analyzed the impacts of feature vector dimension on recommendation precision and efficiency.

We are ready to deploy and implement an actual system and hope to introduce social relationship information among users in the following work. We are to propose a collaborative filtering method which integrates user ratings and information of user social relations in order to further enhance the recommendation accuracy. Besides, how to determine adaptively the optimal parameters still remains a job of further exploration and research.

Acknowledgements

The authors wish to acknowledge the financial support of the National Natural Science Foundation of China (No. 61272094), Natural Science Foundation of Shandong Province (ZR2010QL01, ZR2012GQ010), A Project of Shandong Province Higher Educational Science and Technology Program (J12LN31, J13LN11), Jinan Higher Educational Innovation Plan (201401214，201303001) and Shandong Provincial Key Laboratory Project.

References

[1] L. Lü, M. Medo and C. H. Yeung, “Recommender systems”, Physics Reports, vol. 519, no. 1, **(2012)**, pp. 1-49.

[2] J. S. Breese, D. Heckerman and C. Kadie “Empirical analysis of predictive algorithms for collaborative filtering”, Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence, Morgan Kaufmann Publishers Inc., **(1998)**, pp. 43-52.

[3] F. Ricci, L. Rokach and B. Shapira , "Introduction to recommender systems handbook", Recommender Systems Handbook, Springer US, **(2011)**, pp. 1-35.
[4] M. Jahrer, A. Töscher and R. Legenstein , "Combining predictions for accurate recommender systems", Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, **(2010)**, pp. 693-702.
[5] R. Bell, Y. Koren and C. Volinsky, "Modeling relationships at multiple scales to improve accuracy of large recommender systems", Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, **(2007)**, pp. 95-104.
[6] L. M. De Campos, J. M. Fernández-Luna and J. F. Huete, "Combining content-based and collaborative recommendations: A hybrid approach based on Bayesian networks", International journal of approximate reasoning, vol. 51, no. 7, **(2010)**, pp. 785-799.
[7] H. Wu, Y. J. Wang and Z. Wang, "Two-Phase collaborative filtering algorithm based on co-clustering", Journal of Software, vol. 21, no. 5, **(2010)**, pp. 1042-1054.
[8] C. C. Aggarwal, J. L. Wolf and K. L. Wu, "Horting hatches an egg: A new graph-theoretic approach to collaborative filtering", Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, **(1999)**, pp. 201-212.
[9] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions", Knowledge and Data Engineering, IEEE Transactions on, vol. 17, no. 6, **(2005)**, pp. 734-749.
[10] M. C. Chen, L. S. Chen and F. H. Hsu, "HPRS: A profitability based recommender system", Industrial Engineering and Engineering Management, **(2007)**, pp. 219-223.
[11] J. M. Yang and K. F. Li, "An inference-based collaborative filtering approach", Dependable, Autonomic and Secure Computing, (**2007)**, pp. 84-94.
[12] G. R. Xue, C. Lin and Q. Yang, "Scalable collaborative filtering using cluster-based smoothing", Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval, ACM, **(2005)**, pp. 114-121.
[13] T. Iwata, K. Saito and T. Yamada, "Modeling user behavior in recommender systems based on maximum entropy", Proceedings of the 16th international conference on World Wide Web. ACM, **(2007)**, pp. 1281-1282.
[14] S. T. Park and D. M. Pennock, "Applying collaborative filtering techniques to movie search for better ranking and browsing", Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, **(2007)**, pp. 550-559.
[15] Q. Gu, J. Zhou and C. H. Q. Ding, "Collaborative Filtering: Weighted Nonnegative Matrix Factorization Incorporating User and Item Graphs", SDM, **(2010)**, pp. 199-210.
[16] S. Zhang, W. Wang and J. Ford, "Using singular value decomposition approximation for collaborative filtering", E-Commerce Technology, **(2005)**, pp. 257-264.
[17] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization", Nature, vol. 401, no. 6755, **(1999)**, pp. 788-791.
[18] D. D. Tu, C. C. Shu and H. Y. Yu, "Using unified probabilistic matrix factorization for contextual advertisement recommendation", Journal of Software, vol. 24, no. 3, **(2013)**, pp. 454-464.
[19] B. Sarwar, G. Karypis and J. Konstan, "Application of dimensionality reduction in recommender system-a case study", MINNESOTA UNIV MINNEAPOLIS DEPT OF COMPUTER SCIENCE, **(2000)**.
[20] S. Funk, "Netflix update: Try this at home", http://sifter.org/~simon/journal/20061211.html, **(2006)**.
[21] T. Chen, W. Zhang and Q. Lu, "SVDFeature: A Toolkit for Feature-based Collaborative Filtering", Journal of Machine Learning Research, vol. 13, **(2012)**, pp. 3619-3622.
[22] R. M. Bell and Y. Koren, "Lessons from the Netflix prize challenge", ACM SIGKDD Explorations Newsletter, vol. 9, no. 2, **(2007)**, pp. 75-79.
[23] Y. Koren, R. Bell and C. Volinsky, "Matrix factorization techniques for recommender systems", Computer, vol. 42, no. 8, **(2009)**, pp. 30-37.
[24] A. Mnih and R. Salakhutdinov, "Probabilistic matrix factorization", Advances in neural information processing systems, **(2007)**, pp. 1257-1264.
[25] Y. N. Fang, Y. F. Guo and X. T. Ding, "An improved singular value decomposition recommender algorithm based on local structures", Journal of Electronics & Information Technology, vol. 35, no. 6, **(2013)**, pp. 1284-1289.
[26] N. D. Lawrence and R. Urtasun, "Non-linear matrix factorization with Gaussian processes", Proceedings of the 26th Annual International Conference on Machine Learning, ACM, **(2009)**, pp. 601-608.
[27] R. Salakhutdinov and A. Mnih, "Bayesian probabilistic matrix factorization using Markov chain Monte Carlo", Proceedings of the 25th international conference on Machine learning. ACM, **(2008)**, pp. 880-887.

Authors

Zhijun Zhang, he was born in Shandong, China, in 1973. He received his M. S. degree in School of Computer Science and Technology, Shandong University in 2006. Currently, he is a Ph.D. candidate in School of Information Science and Engineering, Shandong Normal University. His recent research interests include information retrieval and recommender systems.

Hong Liu, she is currently the dean of School of Information Science and Engineering, Shandong Normal University, Jinan, China. She is also the director of Shandong Provincial Key Laboratory for Novel Distributed Computer Software Technology. She has contributed over 100 articles to professional journals. Her current research interests include Distributed Artificial Intelligence and Computer Aided Design.