Research on the Design and Application of Implicit Semantic Model in Bundle Recommendation Method

Zhitao Li, Shuqing Li*

Nanjing University of Finance and Economics

*Corresponding author e-mail: leeshuqing@163.com

Abstract: The recommendation system has been studied for a long time. Data of more users and commodities has been used to improve precision of the recommendation system. However, few studies are on utilization of bundling information. The bundling information refers to the inherent connection of the commodities, and depended by objective properties of commodities. This paper adds bundling information to the recommendation method based on traditional implicit semantic model and collaborative filtering method. This paper sets the bundling coefficient to show the connection degree between commodities and recommended decisive factors. Experiment shows the precision of recommendation could be improved with bundling information in specific environment through changing bundling coefficient.

1. Introduction

As network emerges, massive information regarding commodities, music, films, animation and news, is exposed. In the beginning when information is little, internet users acquire information through portal. When classified search of web portal is insufficient to contain information amount, the search engine appears and plays an important role. Users could get information concerned by inputting key words. However, many people browse information without distinct purpose, or, they do not know what information they want. Then, suitable information shall be recommended to users. For specific shopping malls, the shopping mall may recommend information from multiple angles in view of commodity characteristics. In usual, they recommend the hottest commodities and commodities with the highest reputation to users. Aforesaid method recommends the same information to different users from all angles of commodities, and the commodities selected are suitable for most people. Nonetheless, users are independent individuals, and every user has his/her unique interests. The recommendation of hot commodities could not meet in-depth demands of users. We wish the recommendation system could recommend sports commodities to sports enthusiasts, maternal and baby products to pregnant women, books in certain field to scholars, i.e., providing targeted and personalized recommendation.

The way of personalized recommendation could recommend different products to different users. Commodities recommended are various due to different interests of two users. Then, what is the way to learn interests of different users for commodities? Scholars respectively proposed collaborative filtering recommendation and model-based recommendation according to behavioral data generated by users when purchasing commodities, including purchasing behavior, commodity scoring and comment, and shopping context information to predict probability of every user to purchase certain commodity, or feedback to certain commodity. Those methods not only improve precision of recommendation, but
also greatly increase recommendation coverage compared to former recommendation methods. Some common commodities may be matching with interest characteristics of certain users probably.

The personalized recommendation has been applied to most business situations, but the actual shopping process is more complicated. Many merchants in the shopping mall sell commodities by promotions in order to attract users. Promotions are provided in various forms, including discount, coupons and two-for-one, etc. Bundle sales is a kind of promotion form. Two commodities are sold together, and users may enjoy certain preference when purchasing two commodities at the same time. Thus, when the user purchases certain commodity, his/her probability to buy another commodity will be changed. The bundle recommendation considers aforesaid method. When certain commodity is not purchased by users, and certain commodity is on given recommendation list, the probability of users to purchase commodities in other recommended list or comment on commodities may be changed. When certain commodity is purchased by users, due to bundling, the probability of users to purchase commodities in recommendation list or their evaluation may be changed. This paper aims to analyze and process such change so that the recommendation acquires favorable effect in commodity system with additional bundle factors.

2. Studies concerned

The studies on recommendation system have been carried out long. In 1992, David Goldberg published “Using collaborative filtering to weave an information Tapestry” [1], and proposed the collaborative filtering algorithm for the first time. Subsequently, the researches on the collaborative filtering algorithm are carried out steadily by scholars to improve precision of recommendation [2][3][4]. In 2006, Netfix held recommendation system competition, and attracted many scholars who put forward many recommendation system algorithms. The latent factor model (LFM) [5], scalable collaborative filtering [6], and Markov Chain [7] have been applied to recommendation successively. The matrix decomposition method is prominent in later business field, and now is still applied extensively to commercial recommendation system. Scholars make further optimization on that basis. Some add influence of neighborhood to the matrix decomposition method, and link user historical behaviors with final recommendation [8]; some add influence of time factor to LFM, i.e., the interest characteristics of shopping may be different due to different time [9][10]. Aforesaid improvements play certain role in different scenarios. Besides, scholars make contribution to other orientations. Professor Adomavicius G carried out in-depth study on recommendation based on context [11]; others studied solutions of cold startup [12][13]; general fusion of search, recommendation and advertising is one of the development trends and research directions of the future recommendation system [14]. At present, LFM keeps being optimized [20][21], and for new recommendation system technology based on in-depth learning, bundle recommendation is one of the research topics. Based on early studies and LFM, this paper makes improvement. The difference is that this paper considers influence of bundle. The factor affecting final recommendation effect is the relationship between commodities, which will not be changed with historical behaviors of users. Scholars have not concerned such relationship in aforesaid research methods.

The study on bundle relationship started earlier in economy [15][16]. The bundle recommendation is the comprehensive measurement of inherent commercial bundle relationship and the recommendation system. Suppose various bundling relationships between commodities are given by merchants, and users’ demands are confirmed, the users pay the minimum cost to acquire commodities acquired with bundle commodity purchase study method [17]. The Bundle Recommendation Problem (BRP) proposed by Tao Zhu, et. al. is similar to this paper. This method establishes BRP framework, and proposes a revenue function including bundle coefficient. The bundle relationship between commodities is given to acquire a commodities list so that the benefits reach the maximum after bundle relationship is satisfied [18]. This problem is solved by the Branch and Bound algorithm [19]. Since the solution is NP-hard problem, in order to reduce complexity, candidate set is set. Commodities most possibly to be recommended will be put into the candidate set, and then Branch and Bound algorithm is used to get results. This paper proposes different revenue function. Meanwhile, the bundle relationship is acquired from commodity information to dig out deeper relationship of commodities so that the revenue function could get favorable result regardless of profit information.
3. Establishment of bundle recommendation system

3.1 Establishment of bundle system

There is inherent information of a commodity to express its characteristics, including category information, price information and image information. Except independent information, commodities have inherent connection. The inherent information of commodities could be used to quantize relationship between commodities, and shown by function \( b(i, j) \), in which, \( i, j \) are commodities and \( b \) is the function.

The information of similar commodity categories and the same merchant will undoubtedly make commodities highly similar; on the contrary, different categories and different merchants will make commodities more different; in the same time, the diversity degree is increased.

Suppose the full information of commodity \( i \) is expressed by vector \( k_i = (i_1, i_2, \ldots, i_n) \), and full information of commodity \( j \) is expressed by vector \( k_j = (j_1, j_2, \ldots, j_n) \); then, the similar bundle coefficient of two commodities \( b(i, j) \) can be shown in:

\[
\frac{k_i \cdot k_j}{\|k_i\| \cdot \|k_j\|}
\]

Formula 1

3.2 Bundle coefficient setting method in LFM

LFM is one of the most classical recommendation algorithms. In this section, LFM for scoring prediction is introduced, followed by integration of bundle commodity information to LFM.

Users’ scoring of commodities is affected by multiple factors. In summary, the factors are, on the one hand, personalized characteristics such as interests and personality of users; on the other hand, goods characteristics of functions and performance (regardless of influence of context such as time and place on scoring). Therefore, we wish to establish a prediction function \( p(u, i) \) to predict score made by user \( u \) to commodity \( i \). The vector \( q(i) \) is used to represent characteristics of commodity \( i \). The vectors of commodities have totally \( F \) dimensions, showing \( F \) pieces of characteristic information. The element \( q(i, k) \) of \( q(i) \) in the \( k \)th dimension means the size of common points between commodity \( i \) and the \( k \)th characteristic. Correspondingly, vector \( p(u) \) is used to express characteristics of user \( u \), and there are totally \( F \) dimensions. The \( p(u, k) \) in the \( k \)th dimension means the preference of user \( u \) for the \( k \)th characteristic. Thus, the scoring prediction formula shall be:

\[
p(u, i) = \sum_{k=1}^{F} p(u, k) \cdot q(i, k)
\]

Formula 2

Aforesaid scoring prediction has acquired favorable effect. In actual scoring, however, more circumstances shall be considered. For instance, in the 5-point scoring system, some users score loosely and give 5 points for acceptable commodities; while some demanding users may give 4 or even 3 points to good commodities with certain defects. In addition, the quality of commodity is important. Apparently, the score of commodities with high quality and low price will be higher than those with high price and worse quality. Furthermore, the environment of the scoring system is an important factor affecting results. Different platforms have their own conventional rules for scoring, which shall be considered in prediction. Make \( b_u(u) \) as the deviation generated by user individuals, \( b_i(i) \) as deviation generated by commodity quality, and average the average point scored by users on the platform to commodities. The scoring prediction formula is changed to:

\[
p(u, i) = \sum_{k=1}^{F} p(u, k) \cdot q(i, k) + b_u(u) + b_i(i) + \text{average}
\]

Formula 3

We have acquired user’s historical score data set \( K \), and the points marked by user \( u \) for commodity \( i \) is recorded as \( r(u, i) \). The scoring of users for new commodity may be predicted by machine learning. In the formula, average may be acquired by arithmetic mean of user’s historical points, and \( p(u), q(i), b_u(u) \) and \( b_i(i) \) are unknown. The gradient descent algorithm is used, and the loss function shall be:
In order to prevent over-fitting, the regular item coefficient is added, and the score prediction formula is further optimized to:

\[
\text{loss} = \sum_{(u,i) \in K} (r(u, i) - p(u, i))^2 + \lambda (\| p(u) \|^2 + \| q(i) \|^2 + \| b_u(u) \|^2 + \| b_i(i) \|^2)
\]

Formula 5

Solve loss function derivative of \( p(u,k) \), \( q(i,k) \), \( b_u(u) \) and \( b_i(i) \) respectively, then,

\[
\frac{\partial \text{loss}}{\partial p(u,k)} = \sum_i -2(r(u,i) - p(u,i)) \cdot q(i,k) + 2\lambda p(u,k) \\
\frac{\partial \text{loss}}{\partial q(i,k)} = \sum_u -2(r(u,i) - p(u,i)) \cdot p(u,k) + 2\lambda q(i,k) \\
\frac{\partial \text{loss}}{\partial b_u(u)} = \sum_i -2(r(u,i) - p(u,i)) + 2\lambda b_u(u) \\
\frac{\partial \text{loss}}{\partial b_i(i)} = \sum_u -2(r(u,i) - p(u,i)) + 2\lambda b_i(i)
\]

Formulas 6, 7, 8, 9

Based on aforesaid function, further optimization is made by adding bundle commodity elements to make the prediction results more accurate. Actually, users who score certain commodity may refer to points of other similar commodities. There is a bundle relationship between commodities: when a user likes certain commodity, he/she may possibly imply the preference for other commodities. The preference may be either positive or negative. In other way, the more the user likes certain commodity, the more he/she probably is found of another commodity. Meanwhile, it means he/she may dislike certain commodity greatly. Such commodity relationship is shown by commodity characteristics in original LFM method. However, the commodity characteristics are acquired by machine learning in a concealed way. We have owned commodity property database, and could get commodity bundle data referring to inherent characteristics of commodities. That relationship is more obvious and could amplify influence on relationship between commodities. The \( b(i,j) \) is used to show bundle value between two commodities, and the formula is changed to:

\[
p(u,i) = \sum_{k=1}^{P} p(u,k) \cdot \left( q(i,k) + \mu \sum_{j} q(j,k) b(i,j) \right) + b_u(u) + b_i(i) + \text{average}
\]

Formula 10

In Formula 10, the point of one commodity may be decided by multiple commodities. Every commodity characteristic is a vector and added to the scored commodity. The influence on the scored commodity shall be \( b(i,j) \). The parameter \( \mu \) is added. When \( \mu \) is 0, the LFM method of bundle is not considered. As \( \mu \) becomes larger, the proportion of commodity bundle will be larger.

Various parameters can be acquired through machine learning based on former user scoring data to predict the following scoring. The derivatives of \( p(u,k) \) and \( q(i,k) \) are changed to:

\[
\frac{\partial \text{loss}}{\partial p(u,k)} = \sum_i -2(p(u,i) - r(u,i)) \cdot \sum_{j,k} b(i,j) q(j,k) \\
\frac{\partial \text{loss}}{\partial q(i,k)} = \sum_u -2(p(u,i) - r(u,i)) \cdot p(u,k) \cdot \sum_{j=1} b(i,j)
\]

Formulas 11, 12

With other derivatives remain unchanged, the gradient descent algorithm is used to respectively solve each variant so as to get the scoring prediction model in demand.
3.3 Adding bundle coefficient to top-N recommendation

Top-N recommendation is one important link in commercial recommendation system. The top-N recommendation based on collaborative filtering method is firstly introduced in this section, followed by top-N recommendation added with bundle coefficient.

The principle of recommendation based on item collaborative filtering method is that the user who purchases certain commodity will also purchase other similar commodities. In this method, the similarity $w(i,j)$ between commodity $i$ and commodity $j$ shall be measured firstly. The user purchasing behavior is used to estimate similarity. The more users purchased the same two commodities, the two commodities will be more similar. We establish item vector $N$, with dimensions equaling to number of users. Among the vectors, the dimension of user purchasing the commodity is 1, and other is 0. Thus, the similarity can be shown as:

$$w(i,j) = \frac{N(i) \cdot N(j)}{|N(i)||N(j)|}$$  
Formula 13

Make $N(u)$ to express commodity set purchased by users, and the probability of user $u$ to purchase commodity $i$ shall be:

$$p(u,i) = \sum_{j \in N(u)} w(i,j)$$  
Formula 14

In former recommendation system, if 10 commodities are recommended, we directly traverse all commodities $i$ of user $u$, and select the 10 commodities that the user most probably purchase. However in actual situations, when the recommendation list has certain commodity that one user probably purchase, the probability of the user to buy other commodities in the list may be changed correspondingly. For instance, if one user wishes to buy a TV, the recommendation system acutely notices user demands and recommends 10 kinds of TV. This recommendation is not friendly because the user only needs one TV. When the recommendation list recommends the most matching TV sets, other products in the list are ineffective; contrarily, some other commodities may not be purchased as possibly as the TV sets according to original algorithm, but after the list shows TV products more complying with expectations of the user, the effect of other commodities recommended may be better. Therefore, the diversity of recommendation is very important. When we select $M$ pieces of commodities that the user most probably purchase $R(u) = p(i1)$, select 10 recommended commodities considering relationship between commodities.

The revenue function $R(u)$ of certain user is established. After considering relationship between two commodities, the probability of purchasing the $i$th commodity may be represented by $p(i) + p(i,i1) + p(i,i2) + ... + p(i,in)$, in which, $p(i,in)$ means the influence of commodity in on $i$. Thus, $p(i,in)$ and $p(in,i)$ are not equivalent. $R(u)$ can be shown as:

$$R(u) = \sum_i p(i) + \mu \sum_{i,j} p(i,j)$$  
Formula 15

The establishment of function $p(i,j)$ may be variant to meet different demands. We select the commodity sequence with maximum $R(u)$ for recommendation.

4. Experiment

In order to verify the effect of the method proposed in this paper, experiment is carried out based on data. The Wal-Mart commodity data downloaded from http://jmcauley.ucsd.edu/data/amazon is processed. The data from 2008 to 2014 is saved, and inactive users (with comment information less than 8 pieces) and unpopular commodities (with comment information less than 8 pieces) are filtered. The data set finally acquired contains 19146 commodities and 44337 pieces of user evaluation information, totally 545346 pieces of information. This paper aims to establish recommendation module of the data by comparing different LFM methods so as to get the best LFM method for the group of data.

Due to restricted conditions, the offline experiment is adopted to divide the data set into training set
and test set instead of providing online actual test effect. The data in training set is used to get model framework and test the effect. The 10-fold cross-validation is carried out, i.e., the data set is divided to 10 groups. One group is selected as test set and the other 9 groups are training sets. The effects of the 10 times are averaged, showing the effect of offline experiment.

4. 1 Experiment of LFM method improvement

Two indexes are used to evaluate effect of the scoring system, mean square error (RMSE) and mean absolute error (MAE).

In the test set, the predicted scoring value \( p(u,i) \) of every user-merchant pair \((u, i)\) in every test set, and the index result is acquired according to the actual average value \( r(u,i) \). The larger the gap between RMSE and MAE, the larger the error will be. When the gap is 0, there is no error.

The process of data set training is the process to solve each parameter of the formula and make the loss function minimum. The gradient descent method is used to solve parameters, in which alpha is the learning speed, \( N \) is the iteration frequency of gradient descent, \( F \) is number of characteristics, and \( \lambda \) is the regularization term coefficient. Make alpha=0.02 and \( F=10 \), and change \( N \) value to get the experiment results, as shown in Table 1:

Table 1. Error acquired by predicted scoring under different gradient descent iteration frequency \( N \) according to LFM method

| N  | F  | alpha | \( \lambda \) | RMSE  | MAE  |
|----|----|-------|------------|-------|------|
| 2  | 10 | 0.02  | 0.1        | 1.1216| 0.9109|
| 4  | 10 | 0.02  | 0.1        | 1.0812| 0.8127|
| 8  | 10 | 0.02  | 0.1        | 1.0786| 0.8102|
| 12 | 10 | 0.02  | 0.1        | 1.0782| 1.8105|

The larger the \( N \) iteration frequency is, the better the effect will be, and the slower the iteration will be. The experimental data shows when \( N \) iteration frequency reaches more than 4 times, the precision of prediction has reached a relatively high value. Make \( N=4 \) and \( F=10 \), and change the value of alpha to get the experiment results, as shown in Table 2:
Table 2. Error acquired by predicted scoring under different learning speeds alpha according to LFM method

| N | F  | alpha | lambd | RMSE  | MAE  |
|---|----|-------|-------|-------|------|
| 4 | 10 | 0.01  | 0.1   | 1.1584| 1.1332|
| 4 | 10 | 0.015 | 0.1   | 1.0915| 1.1050|
| 4 | 10 | 0.02  | 0.1   | 1.0812| 0.8127|
| 4 | 10 | 0.03  | 0.1   | 1.1008| 1.1384|

Apparently, neither too fast nor too slow learning speed is suitable. If the alpha is too small, the change is not obvious, while alpha is too large, it is easy to skip the optimal solution. A relatively favorable level can be reached when alpha is at 0.02. Next, make N=4 and alpha=0.02, and change F value to get the experiment results, as shown in Table 3:

Table 3. Error acquired by predicted scoring under different characteristics numbers F according to LFM method

| N | F  | alpha | lambd | RMSE  | MAE  |
|---|----|-------|-------|-------|------|
| 4 | 5  | 0.02  | 0.1   | 1.0859| 0.8204|
| 4 | 10 | 0.02  | 0.1   | 1.0812| 0.8127|
| 4 | 20 | 0.02  | 0.1   | 1.0939| 0.8229|
| 4 | 50 | 0.02  | 0.1   | 1.2901| 1.0211|

Apparently, too low or high characteristics numbers will lead to lower prediction effect. If F is too small, the characteristics number is small, and fitting is not enough due to insufficient data utilization; when F is too large, many unnecessary characteristics will be trained, resulting in over-fitting. When F value is around 10, the effect acquired is favorable. At that time, RMSE and MAE are both small, and the predicted effect is ideal. After adjusting parameters ceaselessly, the most ideal parameter value is about F=10, alpha=0.02 and lambd=0.1. The parameter μ is added on that basis; when μ is 0, it shows the original circumstances. Solve according to the gradient descent method again, and get the experiment results, as shown in Table 4:

Table 4. RMSE and MAE value acquired by predicted scoring under different μ values according to improved LFM method

| N | F  | lambd | μ  | RMSE  | MAE  |
|---|----|-------|----|-------|------|
| 4 | 10 | 0.1   | 0  | 1.0557| 0.7666|
| 4 | 10 | 0.1   | 0.01| 1.0586| 0.7561|
| 4 | 5  | 0.1   | 0.005| 1.0492| 0.7538|
| 4 | 5  | 0.1   | 0.008| 1.0513| 0.7470|

The experiment shows after μ value is added, the value of RMSE and MAE is increased slightly. However, the μ value must be controlled within a reasonable scope, i.e., to control proportion of relationship between commodities to scoring prediction. When the μ value is too small, the inherent relationship between commodities are not considered; while when the μ value is too large, the characteristics of commodities will be neglected. Through adjustment of μ value, we could coordinate proportion of relationship between commodities and characteristics of commodities so as to get more ideal recommendation system. When μ is 0.005, an ideal effect could be acquired.

5. Conclusion

This paper puts forward the bundle recommendation model. Not only information of users to commodities, but also characteristic information of commodities is applied to the bundle recommendation model. Through machine learning of user-commodity pair and commodity information, the bundle recommendation model has a better effect; while the prediction becomes more...
accurate, the diversity of commodities is further realized. The only drawback is the cold startup problem of the recommendation system is not solved. For new user and new commodity with data quantity of 0, the recommendation model is unable to guarantee favorable effect. Meanwhile, compared to traditional methods, the speed of the model is not dominant due to expansion of the parameter amount, which shall be further optimized. Due to restriction of data, the bundle information used here is only regarding commodity category and evaluation. A favorable effect has been acquired. With sufficient data, the bundle information could be further expanded by means of promotion and bundle sales so as to get better effect from designated angles such as diversity and novelty.

6. Acknowledgement

This study was funded by Major Basic Research Project of the Natural Science Foundation of the Jiangsu Higher Education Institutions (grant number 19KJA510011).

References

[1] Goldberg D, Nichols D, Oki B M, et al. Using collaborative filtering to weave an information tapestry [J]. Communications of the ACM, 1992, 35(12): 61-71.

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

[3] Linden G, Smith B, York J. Amazon.com recommendations: Item-to-item collaborative filtering [J]. IEEE Internet computing, 2003 (1): 76-80.

[4] Breese J S, Heckerman D, Kadie C. Empirical analysis of predictive algorithms for collaborative filtering [C]. Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., 1998: 43-52.

[5] Koren Y, Bell R, Volinsky C. Matrix factorization techniques for recommender systems [J]. Computer, 2009 (8): 30-37.

[6] Takács G, Pilászy I, Németh B, et al. Scalable collaborative filtering approaches for large recommender systems [J]. Journal of machine learning research, 2009, 10(Mar): 623-656.

[7] Rendle S, Freudenthaler C, Schmidt-Thieme L. Factorizing personalized markov chains for next-basket recommendation [C]. Proceedings of the 19th international conference on World wide web. ACM, 2010: 811-820.

[8] Koren Y. Factor in the neighbors: Scalable and accurate collaborative filtering [J]. ACM Transactions on Knowledge Discovery from Data (TKDD), 2010, 4(1): 1.

[9] Xiang L, Yang Q. Time-dependent models in collaborative filtering based recommender system [C]. Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 01. IEEE Computer Society, 2009: 450-457.

[10] Koren Y. Collaborative filtering with temporal dynamics [C]. Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2009: 447-456.

[11] Adomavicius G, Tuzhilin A. Context-aware recommender systems [M]. Recommender systems handbook. Springer, Boston, MA, 2011: 217-253.

[12] Park S T, Chu W. Pairwise preference regression for cold-start recommendation [C]. Proceedings of the third ACM conference on Recommender systems. ACM, 2009: 21-28.

[13] Schein A I, Popescul A, Ungar L H, et al. Methods and metrics for cold-start recommendations [C]. Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2002: 253-260.

[14] Garcia-Molina H, Koutrika G, Parameswaran A. Information seeking: convergence of search, recommendations and advertising [J]. Communications of the ACM (Accepted), 2011.

[15] Gaeth G J, Levin I P, Chakraborty G, et al. Consumer evaluation of multi-product bundles: An information integration analysis [J]. Marketing letters, 1991, 2(1): 47-57.

[16] Bakos Y, Brynjolfsson E. Bundling information goods: Pricing, profits, and efficiency [J]. Management science, 1999, 45(12): 1613-1630.
[17] Garfinkel R, Gopal R, Tripathi A, et al. Design of a shopbot and recommender system for bundle purchases[J]. Decision Support Systems, 2006, 42(3): 1974-1986.
[18] Zhu T, Harrington P, Li J, et al. Bundle recommendation in ecommerce[C]. Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval. ACM, 2014: 657-666.
[19] Caprara A, Pisinger D, Toth P. Exact solution of the quadratic knapsack problem[J]. INFORMS Journal on Computing, 1999, 11(2): 125-137.
[20] Yang Hojin, Zhu Hongtu, Ibrahim Joseph G. MILFM: Multiple index latent factor model based on high-dimensional features.[J]. Biometrics, 2018, 148(12).
[21] Hui Francis K C, Tanaka Emi, Warton David I. Order selection and sparsity in latent variable models via the ordered factor LASSO.[J]. Biometrics, 2018, 74(4).