Rating Prediction of Collaborative Filtering using Quasi-Newton Approach

M.Munafur Hussaina[1], Dr.R.Parimala[2]
munafurjalal@gmail.com, rajamohanparimala@gmail.com
Affiliated to Bharathidasan University, Tiruchirappalli
Research Scholar[1], Asst.Professor[2], Department of Computer Science,
Periyar E.V.R.College,
Tiruchirappalli, Tamil Nadu, India.

Abstract

Consumers search the product and its ratings rather than reviews to make purchasing decisions. Rating gives out different intention for different users and product. Collaborative Filtering models developed using Matrix factorization algorithm to predict user's rating of unrated items. There is a need for finding missing ratings for the products. A most popular latent factor Quasi-Newton method is applied to predict the unbounded missing ratings. And also in bounded rating prediction problem, the rating reviews are found using Projected Quasi-Newton algorithm. Experiments conducted on Movie Lens review dataset and real-world dataset. The result shows that the Projected Quasi-Newton algorithm gives better performance.

Keywords: Recommendation System, Newton Method, Hessian Matrix, Quasi-Newton, Rating Prediction Problem (RPP), Collaborative Filtering.

I. Introduction

In this decade, technologies play a major role in every sector, especially in consumer-oriented business. They try to bring any new mechanism on buying and selling environment to the user. World economy increases every year through online shopping. A person who likes shopping may want to look at hundreds of different items and read dozens of reviews before buying the products. E-Commerce Company needs to offer a reliable service based on customer's interest. Product recommendation helps the customer to make better buying decisions. Collaborative filtering is the technique used for recommendation system that constructs an intelligent recommenders' model to give better recommendations as information to the customers. The main categories of product recommendation system are Content-Based Filtering (CBF), Collaborative Filtering (CF), Complementary Filtering and Hybrid recommendation system which use a combination of CBF and CF. Researchers used the Collaborative Filtering as core part of their recommendation systems. The websites offer the recommendation information to the users based on likes and dislikes of similar users. Similarity measures used to determine the set of users having the same behaviour regard to selected items. Star rating is kind of reviewing in numeric scale based on quality and quantity that given by customers or viewers.

The rating prediction is a recommendation task aims to predict user's rating for those items which are similar or dissimilar or not yet rated. Predictions computed automatically from the user's feedback provided in the past [10]. Rating prediction is defined in terms of Root Mean Square Error (RMSE) and it is performed with rating based data. The Root Mean square error on rating prediction problem (RPP) is root mean square of observed and predicted ratings. Item-based approach recommends the items based on user ratings on those items. Item based filtering is faster and more stable than user based filtering. Collaborative filtering faces issues in data sparsity - user gives rating to a limited number of available items, especially when the catalogue is very large. This results in a sparse user-item rating matrix with insufficient data for identifying similar users or items. The issue scalability faces scalability of algorithms with large and real-world dataset. Some algorithms are computationally expensive to run the large datasets. Matrix Factorization (MF) recommendation algorithms emerge locally and produce results for scalability. Without comprising the precision of the system, the recommendation system finding ways in diversity increase. The issues overwhelmed by Latent Factor Model(LFM).The most frequently used algorithm in CF filtering is the K-

nearest neighbours' algorithm and factor analysis. Nearest neighbour method used to identify items with similar users like the inverted neighbourhood which occurs to identify less similar. The LFM measure does not require similarity computation to make inference or prediction. LFM technique projects the users and items by latent vectors in a low dimensional space to evaluate response values with the help of statistical model. In this paper, a proposed work focus on efficient optimization algorithm named Latent Factor Broyden-Fletcher-Goldfarb-Shanno (LF-BFGS) to predict unknown ratings of rating prediction problem. LF-BFGS is an optimization algorithm in the family of quasi-Newton methods which uses a limited amount of computer memory.

II. Literature review

[1] Wei-Sheng et.al developed an Alternating minimization algorithm based on Newton method and carry out factorization for training using MATLAB Package. Novel optimization techniques were used on large-scale training factorization to reduce the running time. The stochastic gradient algorithms are proposed and performed coordinate descent methods in multithreaded environment.

[2] Mahdi Nasiri et.al proposed matrix factorization with initial latent factor for preprocessing of users and items matrices. The method accuracy using optimization-based matrix factorization technique was improved and also increased the speed of convergence.

[3] Herlan et.al discussed the learning algorithm using BFGS Quasi-Newton method and updates the weight of back propagation (BP) using ANN. The BP with BFGS improves the convergences of learning process with optimal architecture and increased the accuracy for glass dataset.

[4] Jianke et.al offered a approach to social recommendation problem with graph Laplacian regularization to capture social relationship among users. The approach formulated a problem into a low-rank semi-definite problem using Quasi-Newton algorithm. The experiment conducted on Epinions dataset is more efficient and effective result for the social recommendation task.

[5] Liwei wu et.al proposed a Collaborative Ranking (CR) algorithm called Primal-CR and Primal-CR++ with Newton update. This reduces the time complexity for the items rated by the users. The results are compared and showed the same efficiency for both the algorithms with minimum loss in ranking.

[6] Ajit.P.Singh et.al projected a Collective matrix factorization, allows handling larger matrices using a sampled approximation to the gradient and Hessian. This method produces an efficient Newton update for the projection with fast rate of convergence.

[7] Fabio Sigrist presented tuning parameter for tree-based Newton boosting method for predicting accuracy. The results of unified framework of gradient, Newton boosting and hybrid variant algorithms are compared. The empirical results showed Newton boosting converged faster and achieves lower training loss.

[8] A.N.Nikolakopoulos and G Karypis et.al proposed uncoupled random walks based collaborative filtering method called as RecWalk, this method perform a novel random walk rely on customer weightage of past preferences on products via uncoupled Markov chain, which helps to find the past performances of customers.

[9] Agoritsa et.al conducted a random walk based model called as scholars walk, which caught the later relationships between the different courses in a university that recommended the courses to the students based on earlier knowledge and wisdom of the crowd. This framework considers as an important for educational institution to bring out the efficient in learning between students.

III. Proposed Methodology

This work resolves the product cold-start where a new product has been added to the database has not yet received enough rating. Matrix factorization is a common approach to CF represented as sparse user item/product

rating matrix. Gradient descent method follows iterative approaches that start from random feature vectors and gradually improves the solutions till there is no change in feature vectors, and/or convergence is reached. The stochastic gradient descent method is not feasible to find the solution in reasonable time. A Newton's method is used to replace the gradient descent, which has the advantages of fast convergence speed and less effort of tedious learning rate. The Quasi-Newton BFGS method is used for its effective, robustness and self-correcting properties, as well as for the super linear convergence approximation of the objective function. Quasi-Newton method is one of the unconstrained optimization techniques, substituting derivative computation using Hessian Matrix. Figure.1 shows the process of PLF-BFGS method to predict the unknown rating prediction. The size of |nxm| dimension of user-item real rating matrix is taken and converting into gradient vector. The inverse hessian matrix used with gradient vector for updating the rating value.

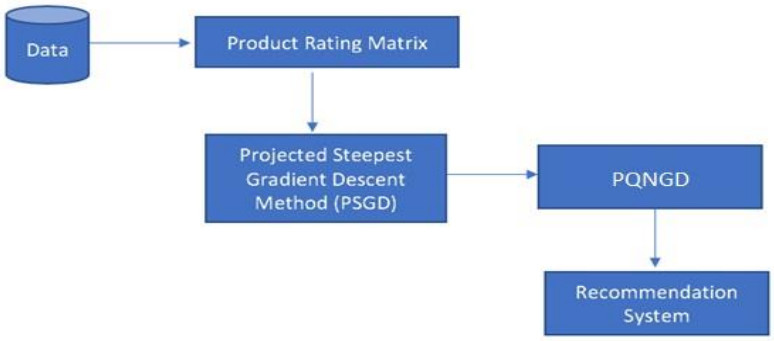

Figure 1. Process of Projected BFGS

The rating prediction is a task of recommender system to predict user ratings for the missing one's (items which are not rated by the user).The user-item rating matrix of a product j with a vector p_j and use i with a vector u_i is denoted as

$$\hat{r}_{ij} = u_i \, . p_j{}^T \tag{1}$$

where u_i is of size n and p_j is of size m. A sample of the user-item rating matrix is shown in Figure.2.The squared prediction error MSE is defined as the difference between the actual and predicted ratings then

$$e_{ij}{}^2 = \sum_{ij}(r_{ij} - \hat{r}_{ij})^2 + \lambda[\|u\|_F^2 + \|p\|_F^2] \tag{2}$$

r_{ij} is the actual rating by user i for product j and $\hat{r}_{ij}$ is the ratings to predict. The objective function eij^2 is calculated with bias value λ=1 for rating prediction. The power of CF method projects users and items into latent vector space of reduced dimensionality to predict missing rating. The gradient descent is an optimization technique used to reduce the error rate. In rating prediction model, the objective is to minimize the root mean square error value.

		Items					
		1	*2*	...	*i*	...	*m*
Users	*1*	5	3		1	2	
	2		2				4
	:			5			
	u	3	4		2	1	
	:					4	
	n			3	2		

Figure 2: User-Item Rating Matrix

BFGS uses quadratic Taylor approximation of the objective function f(x) in x.

$$f(x+d) = f(x) + d^T g(x) + \frac{1}{2} d^T H(x) d \quad (3)$$

where g(x) is the gradient vector and H(x) is the Hessian matrix. The necessary condition for a local minimum of f(x) is g(x)+H(x)d=0. If the function is quadratic, the optimum solution arises in a single step. The Newton direction d for line search is $d=H(x)^{-1}g(x)$. In quasi-Newton method, the hessian matrix is approximated in each iteration. The BFGS method uses iterative Hessian as

$$H_{k+1} = H_k - \frac{H_k S_k S_k{}^T H_k}{S_k{}^T H_k S_k} + \frac{Y_k Y_k{}^T}{Y_k{}^T S_k} \quad (4)$$

Where

$$s_k = x_{k+1} - x_k \quad (5)$$

$$y_k \leftarrow g_{k+1} - g_k \quad (6)$$

where initial Hessian H_0 is the identity matrix. Iterative Hessian saves a large amount of computation time when input data is in sparse. The loop is repeated with changes of gradient until the tolerance is met. The missing rating will be predicted and the root mean square error (RMSE) is recorded. The bound constrained rating prediction problem as in equation (2) $\min f(x), x \in R^n$ states that $l \leq x \leq u$ where l and u are bounded rating. The rating solution projects on to the set $\{l, u\}$. The PLF-BFGS algorithm predicts the unknown ratings or missing ratings with bounded constraints $l \leq r_{ij} \leq u$.

3.1 Algorithm

LF-BFGS

Input: Real Rating Matrix
Output: Predicting Missing ratings with fast convergence
Variables
$(c^{(0)})$ - as Gradient vector
$(H^{(0)})$ - as Hessian matrix
$z^{(0)}$ - as matrix
k – Number of iteration
Step 1: Read the real rating matrix
Step 2: Initialize the gradient vector($c^{(0)}$), tolerance, β, α and k.
Step 3: Compute the gradient matrix,
$g^{(k)} = - (H^{(k)})^{-1} c^{(k)}$
Step 4: Calculate the step length,
$d^{(k)} = f (z^{(k)} + dg^{(k)})$
Step 5: Find the rate of gradient $y^{(k)}=c^{(k+1)}-c^{(k)}.y^{(k)}$.
Step 6: Compute $s^{(k)}=dg^{(k)}$
Step 7: Update Hessian Matrix as in (4)
Step 8:IF$\| g^{(k)} \|$ < tolerance then
Print "Converged"
Loop
Calculate the overall squared error,
$e_{ij}^2 = (r_{ij} - \hat{r}_{ij})^2$
End loop
Else // $\| c^{(k)} \|$ > error tolerance
go to Step 3.
Step 9: Stop.

a. Dataset

Bench mark Movie Lens is a web based recommender system dataset publically available from Group Lens Research, recommends movies for the users to watch [12]. The size of 1M (100k) has 943 users and 1682 movies with 100000 ratings in the dataset. Randomly selected 100 records from the database are taken for implementation.

The secondly Samsung Mobile real review (SMRR) datasets are crawled and collected from different websites like flipkart.com, amazon.com, mysmartprice.com, snapdeal.com and gadgets.com [11]. Dataset has information about

Product Name, ratings, price and text review information. The 5-star scale rating is deployed in movie lens and SMRR data set. A subset of review rating of size 100 is select for the study. The experimentation is conducted out in R language.

IV. Results and Findings

The proposed algorithms are used to minimize the RMSE (Root Mean Square Error) to find the missing rating of the rating prediction problem (RPP). The relationship between user u_i and item p_j of dot product of rating matrix converted into real rating matrix of size nxm. This conversion of rating matrix used to make prediction to rate from user u to item p. The learning factor k=2 is initialized in the algorithm to find the gradient descent. The epoch is used for 1000 iterations. Hessian matrix $H^{(k)}$ is initialized as identity matrix and multiplied with gradient $g^{(k)}$ until step length replaces the old gradient value. Update of new gradient with quasi-Newton formula as in (4). The Hessian Matrix as $H^{(k+1)}$ is calculated for every iteration until the tolerance met. The result reduces the error rate and predicts the new ratings for LF-BFGS method. A LF-BFGS algorithm extended to bound constraints. The proposed model uses the SMRR dataset that consist of variables namely UID, ITEM, PRICE, RATINGS as shown in Table.1. The implementation was written in R programming Language. Results are depicted in Figure.3. Table.2 shows the results of RMSE value for proposed method.

Table.1 A Sample of SMRR Dataset

UID	Name of the Product	Price	Ratings
1	Samsung On7 Pro (Gold)	7,990	3.8
1	Samsung On5 Pro (Gold)	7,490	4
1	Samsung Guru 1200 GT-E1200ZKYINS - Black	1,100	4
2	Samsung Guru Music 2 (Gold)	1,625	4
2	Samsung On7 Pro (Black)	7,990	3.8
3	Samsung Guru Music 2 (SM-B310E, Blue)	1,625	4
3	Samsung Galaxy J8 (Black, 64GB) with Offers	17,990	4
4	Samsung Guru Music 2 (SM-B310E, Black)	1,625	4
4	Samsung On5 Pro (Black)	7,490	4
4	Samsung Guru Music 2 SM-B310E (White)	1,625	4

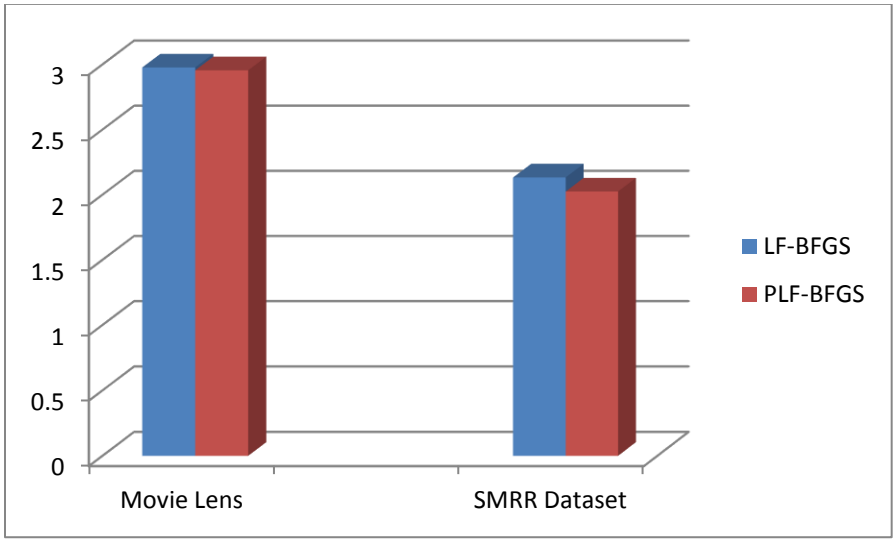

Figure 3. Comparison of Root Mean Square Error Value

Table 2.Error of RPP using LFBFGS and LFBFGS algorithms for Movie Lens & SMRR Dataset

	Algorithm	Movie Lens	SMRR Dataset
RMSE value	LFBFGS	2.97643	2.132463
	PLFBFGS	2.9546	2.024486

V. Conclusion

Algorithmic framework for solving unconstrained RPP and box constrained RPP using Quasi-Newton method. Two methods LF-BFGS and PLF-BFGS were presented. The empirical results of proposed method applied to movie lens rating prediction and SMRR rating prediction problem. The numerical results show that PLF-BFGS method performs substantially better than LF-BFGS method. An algorithmic improvement of rating prediction will be focused in future.

Acknowledgement

We first thank the Almighty and thanks all those who have helped us to make this paper a successful one and giving several openings in research. I also thank the R Community for providing this open source software towards successful implementation of this research work.

Reference

[1] Wei-Sheng Chin, Bo-Wen Yuan, Meng-YuanYang and Chin-Jen Lin, (2016). "An Efficient Alternating Newton Method for Learning Factorization Machines", ACM Transactions on Intelligent Systems and Technology (TIST), 9(6), pp:1-31.

[2] Mahdi Nasiri and Behrouz Minaei (2016, May), "Increasing prediction accuracy in collaborative filtering with initialized factor matrices", Journal of Super computing, Volume 72, pp: 2157-2169.

[3] Herlan Silaban, Muhammad Zarlis and Sawaluddin (2017, December), "Analysis of Accuracy and Epoch on Back-propagation BFGS Quasi-Newton", In Journal of Physics: Conference Series 930,012006, IOP Publishing, pp: 1-6.

[4] Jianke Zhu, Hao Ma, Chun Chen and Jiajun Bu (2011, August), "Social Recommendation Using Low-Rank Semidefinite Program", Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence, pp: 158-163.

[5] Liwei Wu, Cho-Jui Hsieh and James Sharpnack (2017, August), "Large-scale Collaborative Ranking in Near-Linear Time", In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'17), pp: 515-524, ACM.

[6] Ajit P.Singh and Geoffrey J.Gordon (2008, August), "Relational Learning via Collective Matrix Factorization", In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, pp:650-658, ACM.

[7] Fabio Sigrist (2019), "Gradient and Newton Boosting for Classification and Regression", arXiv preprint arXiv:1808.03064v3, pp:1-41.

[8] Athanasios N.Nikolakopoulos and George Karypis (2019), "Boosting Item-based Collaborative Filtering via Nearly Uncoupled Random Walks", arXiv preprint arXiv:1909.03579v1 [cs.IR], pp:1-24.

[9] Agoritsa Polyzou, Athanasios N.Nikolakopoulos and George Karypis (2019), "Scholars Walk: A Markov Chain Framework for Course Recommendation", In: The 12th International Conference on Educational Data Mining, pp. 396 – 401.

[10] Stefan Pero, Tomas Horvath (2013), "Opinion-Driven Matrix Factorization for Rating Prediction", User Modeling, Adaptation and Personalization: 21st International Conference, UMAP, Rome, Italy, pp:1-13.

[11] M.Munafur Hussaina and R.Parimala (February, 2019), "Item Recommendation System Using Matrix Factorization Technique", International Journal for Research in Engineering Application & Management (IJREAM), pp: 503-508.

[12] Badrul M. Sarwar, George Karypis, Joseph A. Konstan and John T. Riedl (July 2000), "Application of Dimensionality Reduction in Recommender System -- A Case Study", Group Lens Research Group / Army HPC Research Center, Department of Computer Science and Engineering University of Minnesota, pp:1-15.