Evaluating the Augmented Lagrangian Method for matrix completion used for Collaborative Filtering

Sven Hammann, Stefan Dietiker, Jannick Griner
Department of Computer Science, ETH Zurich, Switzerland

***Abstract*—We evaluate the performance of a collaborative filtering algorithm that uses the Augmented Lagrangian Method (ALM) to minimize the nuclear norm of the reconstructed matrix. We present a slightly simpler version of this algorithm and show that is still performs well on different data sets, including data that contains corrupted data from malicious attackers. In particular, we show that the algorithm performs better than robust principal component analysis on these data sets, which is specifically used to filter out corrupted entries. We also compare the ALM algorithm to a baseline SVD algorithm and show that it performs better in terms of mean squared error, but has a longer runtime.**

I. Introduction

Recommender systems are widely used in practice to present items, e.g. books, movies or music, that are of interest to the user. Collaborative filtering algorithms try to predict ratings of users on items they have not yet rated. They do so by relating the ratings of the user for other items to those of other users.

In our model, we consider an $m \times n$ matrix that consists of the ratings of m users for n items, where many values are missing, denoted by a special value. The goal of the collaborative filtering algorithm is to predict the missing values as accurately as possible. Its performance is measured by comparing a full original data matrix with the matrix that was reconstructed by the algorithm, given a part of the original matrix with missing entries.

The main assumption of many collaborative filtering algorithms including the one used in this paper is that the original matrix has low rank, or at least approximately low rank, i.e., many dimensions that have small singular values. The reason for this lies in the assumed correlations between ratings: If two or more users have similar ratings on some items, it is likely that there is some underlying smaller dimension responsible for these ratings. These dimensions can be interpreted as topics of interest.

A natural procedure is now to find a solution that respects the known entries, i.e., the reconstruction is equal to the original on the entries that are known, and that has minimal rank. A good approximation for minimization of the rank is minimizing the nuclear norm, i.e., the sum of singular values. This can be solved by convex optimization. The algorithm we use was presented in [1] and uses the Augmented Lagrangian Method (ALM).

An additional consideration is that of corrupted entries in the original data matrix. These values are not missing, but they may not reflect a true rating of a user. An important real-world example are shilling attacks, where companies upvote their own products and downgrade the products of competitors. In a scenario where such attacks are expected, it is advisable to use a different method that takes care of these corruptions. We compare the algorithm using ALM with another algorithm that performs Robust Principal Component Analysis (RPCA), as introduced in [2], and can deal with corrupted entries.

We compare the algorithm using ALM from [1] to an algorithm using RPCA (as in [2]) and to another baseline algorithm using one Singular Value Decomposition. We present our own model to generate artificial test data that simulates users rating hotels that belong to different groups, and attackers that may want to benefit one group and hurt the others. We test our algorithms on our own artificial data as well as on a different data set. We then choose the ALM algorithm based on its performance on these data sets.

II. Models and Methods

A. Related Work

The algorithm we use is based on Algorithm 6 in [1]. We use a simplified version of it that does not include the update step for μ and uses the economy-size singular value decomposition directly provided by MATLAB rather than an external library.

We compare these results to those of another algorithm based on [2], Section 1.6, that explains how Robust PCA can be used for matrix completion.

B. Model

We formulate our problem as in Section 3.2 of [1]. Given an original matrix D and a number of observed entries Ω, we wish to reconstruct the complete matrix D. Under certain conditions on the rank of D and the number of samples [TODO more detailed, cite Exact matrix completion via convex optimization], D can be recovered exactly. In the case of collaborative filtering, we may be given a matrix D that is only approximately low-rank, i.e., it has many dimensions with small singular values. In this case, a good approximation is still possible.

We want so solve the following problem:

$$\underset{A}{\text{minimize}}\ ||A||_* \text{ subject to } A_{ij} = D_{ij}\ \forall (i,j) \in \Omega$$

Let D_Ω be the matrix such that $(D_\Omega)_{ij} = D_{ij}\ \forall (i,j) \in \Omega$, and $(D_\Omega)_{ij} = 0\ \forall (i,j) \notin \Omega$. That is, the missing values are set to zero. We can then rewrite the constraint by introducing an auxiliary variable matrix E:

$$\underset{A}{\text{minimize}}\ ||A||_* \text{ subject to}$$

$$A + E = D_\Omega \quad E_{ij} = 0\ \forall (i,j) \in \Omega$$

Since E may be nonzero only on the coordinates of missing values, its only purpose is to compensate for the values set in A on these coordinates. As D_Ω is zero there, if $A_{ij} = x$, we must have $D_{ij} = -x$ to fulfill the constraint.

Introducing E makes it possible to solve this problem using the Augmented Lagrangian Method.

III. Results

IV. Discussion

V. Summary

References

[1] Z. Lin, M. Chen, and Y. Ma, "The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices," *ArXiv e-prints*, Sep. 2010.

[2] Y. M. Emmanuel J. Candès, Xiaodong Li and J. Wright, "Robust principal component analysis?" *Journal of ACM*, vol. 58, no. 1, pp. 1–37, 2011.