



भारतीय प्रौद्योगिकी संस्थान हैदराबाद  
Indian Institute of Technology Hyderabad

# Using Regularized SVD for Recommender Systems (iterative method)

MA5120 - Numeric Linear Algebra

---

Prajwaldeep and Vishwanath

April 27, 2023

## Team Members

- Mulugu Vishwanath Sharma - MA20BTECH11010
- Prajwaldeep Kamble - MA20BTECH11013

# Introduction

---

## Abstract

This presentation is part of course project for MA5120 - Numeric Linear Algebra. The project is based on the paper "Using Regularized SVD and Application to Recommender Systems" by Shuai Zheng, Chris Ding and Feiping Nie. The paper is about the use of Regularized SVD for Recommender Systems and How we could optimize the RSVD by using the Iterative method. We implemented the algorithm in Python and checked it on actual movie Dataset.

## About Recommender Systems

The Recommender Systems are a really important application of Numeric Linear Algebra in the modern world. they're used extensively by the Streaming platforms (like Netflix and YouTube), movie rating websites (like imdb and my anime list) as well as E-commerce platforms like amazon/flipkart. In our class this was mentioned as an application of the SVD and referred to as 'The Netflix Problem'. Netflix also awards the 'Netflix Prize' for best collaborative filtering algorithm.

## **collaborative Flitering Alogorithms**

The collaborative flitering alogorithms are used to make predictions about the user's interests by collecting information from many users (collaborating). The information collected is used to predict the items (movies, songs, products, etc) that the user might like. Unlike the content-based recommender systems, the collaborative flitering alogorithms do not require the item's metadata. It uses the data from other users who liked the same item to predict the user's interest.

## Regularized SVD

The Regularized SVD is a method to approximate a matrix by using the SVD of the matrix. It is used to reduce the dimensionality of the data. It is also used to remove noise from the data. The main idea behind the RSVD is to add a regularization term to the SVD. The regularization term is used to penalize the large values in the matrix. The RSVD is used in many applications like Recommender Systems, Image Processing, etc. It's primary advantage over the SVD is that it is more robust to noise and overfitting.

## Regularized SVD

Assume there is a matrix  $X \in \mathbb{R}^{m \times n}$ . Regularized SVD (RSVD) tries to find low-rank approximation using regularized factor matrices  $U$  and  $V$ . The objective function is proposed as

$$J_1 = \|X - UV^T\|_F^2 + \lambda \|U\|_F^2 + \lambda \|V\|_F^2 \quad (1)$$

where low-rank regularized factor matrices  $U \in \mathbb{R}^{n \times k}$  and  $V \in \mathbb{R}^{m \times k}$ ,  $k$  is the Number of latent features of the moves that we are considering. Minimizing eq:?? is a multi-variable problem.



## Data Collection

---

## The Dataset-1

- The dataset is 'The movies dataset' from Kaggle. It contains the data of 3,000 movies released on or before July 2017 along with ratings of 1081 users.
- The dataset contains the following information about the movies: Ratings, Budget, Revenue, Genres, Cast, Crew, etc.; metadata about the movies, Userid, Credits etc..
- But We'll be using only the Userid and Ratings from the dataset. As we're using the collaborative filtering algorithm. We don't need the metadata of the movies.

## The Dataset-2

- The dataset is "TMBD 5000 Dataset" The dataset is very Similar to the previous dataset
- It contains all attributes like metadata, cast, crew, budget, revenue, release dates, languages, production companies, countries, TMDB vote counts and vote averages.
- But for collaborative flitering alogorithm we'll be using only the Userid and Ratings from the dataset. We don't need the metadata of the movies.
- It's Glimpse is Similar to the previous dataset.

## Movie Dataset-1

	A	B	C	D
1	user	movie	rating	
2	595	1735	3	
3	534	2940	4	
4	1024	664	2	
5	1076	703	3.5	
6	35	914	4.5	
7	389	2360	4	
8	67	772	5	

## Defining variables

The below representations of all the variables are used in the presentation.

- $m$  - Number of users
- $n$  - Number of movies
- $k$  - Number of latent features
- $X$  -  $m \times n$  matrix of ratings, where  $X_{ij}$  is the rating given by user  $i$  to movie  $j$
- $U$  -  $m \times k$  matrix of user features, where  $U_{ik}$  is the strength of association of user  $i$  with latent feature  $k$
- $V$  -  $n \times k$  matrix of movie features, where  $V_{jk}$  is the strength of association of movie  $j$  with latent feature  $k$

# Algorithms

---

## **Iterative way of finding U and V**

---

## Iterative way of finding U and V

- The objective function is proposed as
$$J_1 = \|X - UV^T\|_F^2 + \lambda \|U\|_F^2 + \lambda \|V\|_F^2$$
- On taking the partial derivative of the above equation with respect to U and equating it to zero, we get

$$U = (XV)(VV^T + \lambda I)^{-1}$$

- On taking the partial derivative of the above equation with respect to V and equating it to zero, we get

$$V = (X^T U)(UU^T + \lambda I)^{-1}$$



## Iterative way of finding $U$ and $V$

**Step 1:** Initialize  $V$  with random values.

**Step 2:** Calculate  $U$  using the equation  $U = (XV)(VV^T + \lambda I)^{-1}$

**Step 3:** Calculate  $V$  using the equation  $V = (X^T U)(UU^T + \lambda I)^{-1}$

**Step 4:** Repeat Step 2 and Step 3 until convergence.

## RSVD in Subspace of SVD

Let  $F\Sigma G^T$  is the SVD of the matrix  $X$ . And the optimal values of RSVD be  $U^*$  and  $V^*$ . Then the following holds true:

$$U^* = F_K * \Omega \quad (2)$$

$$V^* = G_K * \Omega \quad (3)$$

$$(4)$$

where,

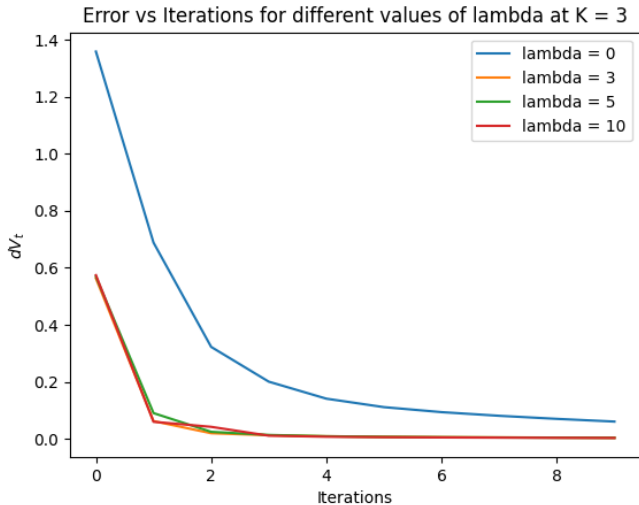
$$F_K = F[:, 1 : K] \quad (5)$$

$$G_K = G[:, 1 : K] \quad (6)$$

$$\Omega = \sqrt{\Sigma - \lambda I_K} \quad (7)$$

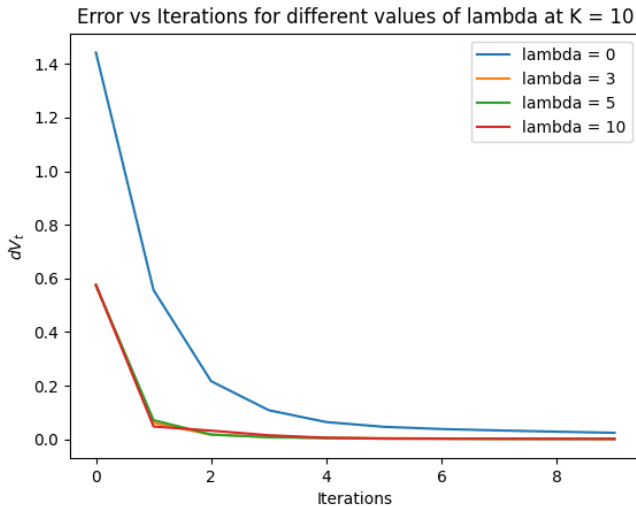
- The computational complexity of the RSVD is  $O(k(m+n)\min(m, n))$ .
- But the iterative way of finding U and V is computationally less expensive than the RSVD because the term  $VV^T + \lambda I$  is much faster to converge than SVD.

# Results



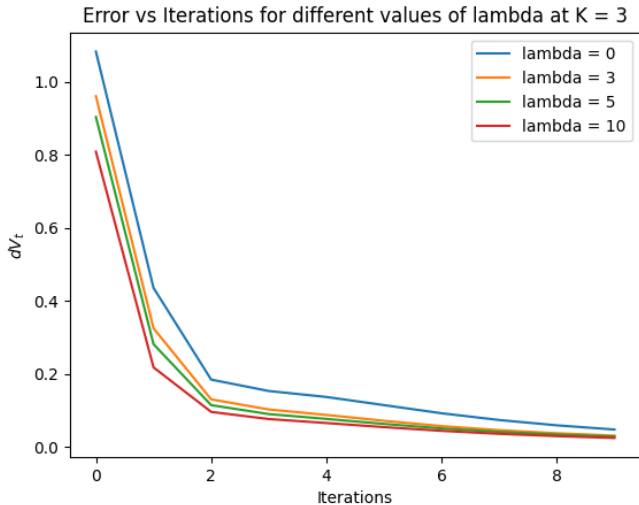
**Figure 2:** Convergence of RSVD on dataset-1 for k=3

# Results



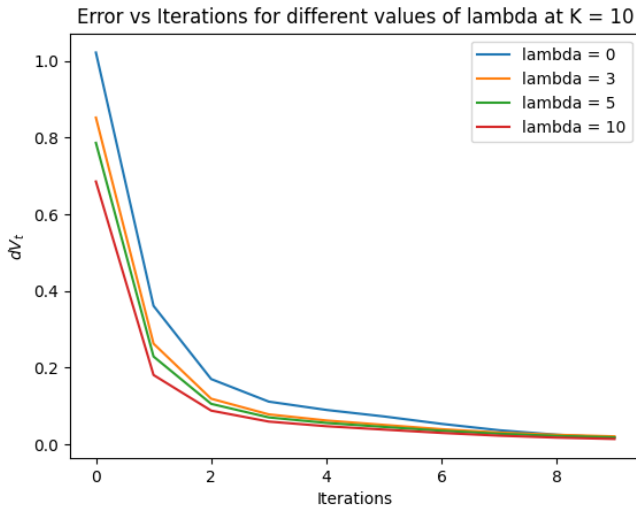
**Figure 3:** Convergence of RSVD on dataset-1 for k=10

# Results



**Figure 4:** Convergence of RSVD on dataset-2 for k=3

# Results



**Figure 5:** Convergence of RSVD on dataset-2 for k=10