Hochschule Offenburg
offenburg.university

Block 2

Recommender Systems II

Janis Keuper

Basics of Recommender Systems

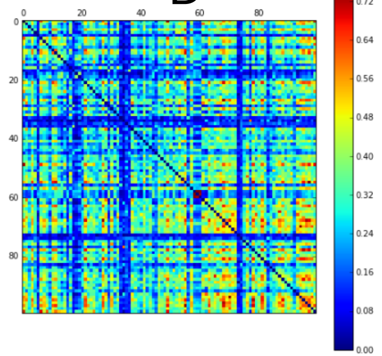
Recall: Item-Item Collaborative Filter

IV Problem: scalability and complexity in the number of products

D of size $n \times n$ with distance d_{ij}, for $i \in \{0, \dots, n-1\}$ and $j \in \{0, \dots, n-1\}$

How many different products does Amazon offer?

D



~ 300 million !

→ **huge matrix (~100 billion entries) with almost only zeros!**

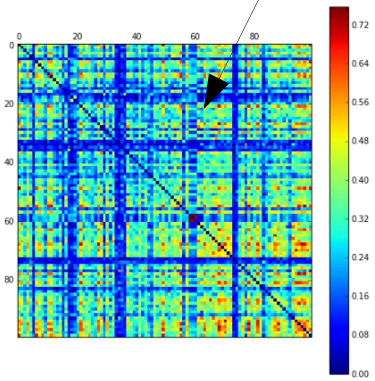
→ **a standard Solution: Singular Value Decomposition (need to do some math first)**

Hochschule Offenburg
offenburg.university

Standard Approach for Collaborative Filters: rSVD

Use regularized Singular Value Decomposition (rSVG) to handle huge recommendation matrix

expected distance $\tilde{d}_{ij} := q_i^T p_j$



Approximate entry in distance matrix by vector factorization.

Saves space, because dimension of p and q are chosen to be much smaller than n:

$$\dim(q) \ll n$$

→ p and q in implicit new compact basis for D

Netflix Prize and SVD

Stephen Gower

April 18th 2014

Abstract

Singular Value Decompositions (SVD) have become very popular in the field of Collaborative Filtering. The winning entry for the famed Netflix Prize had a number of SVD models including SVD++ blended with Restricted Boltzmann Machines. Using these methods they achieved a 10 percent increase in accuracy over Netflix's existing algorithm.

In this paper I explore the different facets of a successful recommender model. I also will explore a few of the more prominent SVD based models such as Iterative SVD, SVD++ and Regularized SVD. This paper is designed for a person with basic knowledge of decompositions and linear algebra and attempts to explain the workings of these algorithms in a way that most can understand.

Introduction

On October 2nd, 2006 Netflix began their contest to find a more accurate movie recommendation system to replace their current system, Cinematch. They promised a prize of one million dollars to anyone who could improve over the Cinematch system by at least 10% Root Mean Squared error (RSME). After three years of the contest, in 2009 the grand prize was awarded to team "Bellkor's Pragmatic Chaos", much of this paper draws from papers written by one of the members of the team, Yehuda Koren.

But before one can examine the algorithm developed to win the prize, first it must be known what must be addressed. Recommender systems are important to services such as Amazon, Netflix and other such online providers that aim to attract users. For Amazon, they want to present their users with products they want so they are more likely to spend their money through their site. On the part of Netflix, their motivation is to help users find movies and shows that fit their tastes because they are more likely to maintain a subscription.

Basics of Recommender Systems

Standard Approach for Collaborative Filters: rSVD

Use regularized Singular Value Decomposition (rSVG) to handle huge recommendation matrix

$$\text{expected distance } \tilde{d}_{ij} := q_i^T p_j$$

Approximate entry in distance matrix by vector factorization.

Short cut: compute set of p and q vectors directly, via optimization:

$$minimum(q,p) \sum_{i,j \in n} \| d_{i,j} - q_i^T p_j \|^2$$


Netflix Prize and SVD

Stephen Gower

April 18th 2014

Abstract

Singular Value Decompositions (SVD) have become very popular in the field of Collaborative Filtering. The winning entry for the famed Netflix Prize had a number of SVD models including SVD++ blended with Restricted Boltzmann Machines. Using these methods they achieved a 10 percent increase in accuracy over Netflix's existing algorithm.

In this paper I explore the different facets of a successful recommender model. I also will explore a few of the more prominent SVD based models such as Iterative SVD, SVD++ and Regularized SVD. This paper is designed for a person with basic knowledge of decompositions and linear algebra and attempts to explain the workings of these algorithms in a way that most can understand.

Introduction

On October 2nd, 2006 Netflix began their contest to find a more accurate movie recommendation system to replace their current system, Cinematch. They promised a prize of one million dollars to anyone who could improve over the Cinematch system by at least 10% Root Mean Squared error (RSME). After three years of the contest, in 2009 the grand prize was awarded to team "Bellkor's Pragmatic Chaos", much of this paper draws from papers written by one of the members of the team, Yehuda Koren.

But before one can examine the algorithm developed to win the prize, first it must be known what must be addressed. Recommender systems are important to services such as Amazon, Netflix and other such online providers that aim to attract users. For Amazon, they want to present their users with products they want so they are more likely to spend their money through their site. On the part of Netflix, their motivation is to help users find movies and shows that fit their tastes because they are more likely to maintain a subscription.

Basics of Recommender Systems

Standard Approach for Collaborative Filters: rSVD

Use regularized Singular Value Decomposition (rSVG) to handle huge recommendation matrix

expected distance $\tilde{d}_{ij} := q_i^T p_j$

Approximate entry in distance matrix by vector factorization.

Short cut: compute set of p and q vectors directly, via optimization:

$$minimum(q,p) \sum_{i,j \in n} \|d_{i,j} - q_i^T p_j\|^2$$

In practice: regularized optimization via
Stochastic Gradient Descent

Netflix Prize and SVD

Stephen Gower

April 18th 2014

Abstract

Singular Value Decompositions (SVD) have become very popular in the field of Collaborative Filtering. The winning entry for the famed Netflix Prize had a number of SVD models including SVD++ blended with Restricted Boltzmann Machines. Using these methods they achieved a 10 percent increase in accuracy over Netflix's existing algorithm.

In this paper I explore the different facets of a successful recommender model. I also will explore a few of the more prominent SVD based models such as Iterative SVD, SVD++ and Regularized SVD. This paper is designed for a person with basic knowledge of decompositions and linear algebra and attempts to explain the workings of these algorithms in a way that most can understand.

Introduction

On October 2nd, 2006 Netflix began their contest to find a more accurate movie recommendation system to replace their current system, Cinematch. They promised a prize of one million dollars to anyone who could improve over the Cinematch system by at least 10% Root Mean Squared error (RSME). After three years of the contest, in 2009 the grand prize was awarded to team "Bellkor's Pragmatic Chaos", much of this paper draws from papers written by one of the members of the team, Yehuda Koren.

But before one can examine the algorithm developed to win the prize, first it must be known what must be addressed. Recommender systems are important to services such as Amazon, Netflix and other such online providers that aim to attract users. For Amazon, they want to present their users with products they want so they are more likely to spend their money through their site. On the part of Netflix, their motivation is to help users find movies and shows that fit their tastes because they are more likely to maintain a subscription.

Basics of Recommender Systems

Standard Approach for Collaborative Filters: rSVD

Use regularized Singular Value Decomposition (rSVG) to handle huge recommendation matrix

$$expected\ rating = \hat{r}_{ui} = q_i^T p_u$$

Approximate entry in rating matrix by vector factorization

$$minimum(p,q) \sum_{(u,i)\epsilon K} (r_{ui} - q_i^T . p_u)^2 + \lambda(||q_i||^2 + \ ||p_u||^2)$$

In practice: regularized optimization via
Stochastic Gradient Descent

$$minimum(q,p) \sum_{i,j\in n} ||d_{i,j} - q_i^T p_j||^2 + \lambda(||q||^2 + ||p||^2)$$

Netflix Prize and SVD

Stephen Gower

April 18th 2014

Abstract

Singular Value Decompositions (SVD) have become very popular in the field of Collaborative Filtering. The winning entry for the famed Netflix Prize had a number of SVD models including SVD++ blended with Restricted Boltzmann Machines. Using these methods they achieved a 10 percent increase in accuracy over Netflix's existing algorithm.

In this paper I explore the different facets of a successful recommender model. I also will explore a few of the more prominent SVD based models such as Iterative SVD, SVD++ and Regularized SVD. This paper is designed for a person with basic knowledge of decompositions and linear algebra and attempts to explain the workings of these algorithms in a way that most can understand.

Introduction

On October 2nd, 2006 Netflix began their contest to find a more accurate movie recommendation system to replace their current system, Cinematch. They promised a prize of one million dollars to anyone who could improve over the Cinematch system by at least 10% Root Mean Squared error (RSME). After three years of the contest, in 2009 the grand prize was awarded to team "Bellkor's Pragmatic Chaos", much of this paper draws from papers written by one of the members of the team, Yehuda Koren.

But before one can examine the algorithm developed to win the prize, first it must be known what must be addressed. Recommender systems are important to services such as Amazon, Netflix and other such online providers that aim to attract users. For Amazon, they want to present their users with products they want so they are more likely to spend their money through their site. On the part of Netflix, their motivation is to help users find movies and shows that fit their tastes because they are more likely to maintain a subscription.

Literature

[0] free icons taken from https://www.flaticon.com

[1] www.amazon.de

[2] www.netflix.de

[3] www.spotify.com

[4] www.zalando.de

[5] https://blogs.gartner.com/martin-kihn/how-to-build-a-recommender-system-in-python/

[6] http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.895.3477&rep=rep1&type=pdf

[7] https://www.designmantic.com/blog/movie-moods-in-typography/

[8] https://arxiv.org/abs/1205.3193