OLX GROUP

ADAM MICKIEWICZ UNIVERSITY POZNAŃ

Part 4: Matrix factorization

Robert Kwieciński

OLX Group and Adam Mickiewicz University

May 22, 2020

MF models

KNN models are either user-based or item-based. Here we will see the method which utilizes both user and item similarities simultanously.

Matrix factorization idea

Let's represent each user and item as a vector in the same space. The items the closest to the target user are items which we should recommend him.

MF models

KNN models are either user-based or item-based. Here we will see the method which utilizes both user and item similarities simultanously.

Matrix factorization idea

Let's represent each user and item as a vector in the same space. The items the closest to the target user are items which we should recommend him.

Matrix factorization

Let k be given (it's a number of latent factors).
Consider $P \in \mathbb{R}^{m \times k}$ and $Q \in \mathbb{R}^{n \times k}$ which are respectively matrices of user and item latent factors. We estimate

$$\hat{R}_{ui} = PQ^T$$

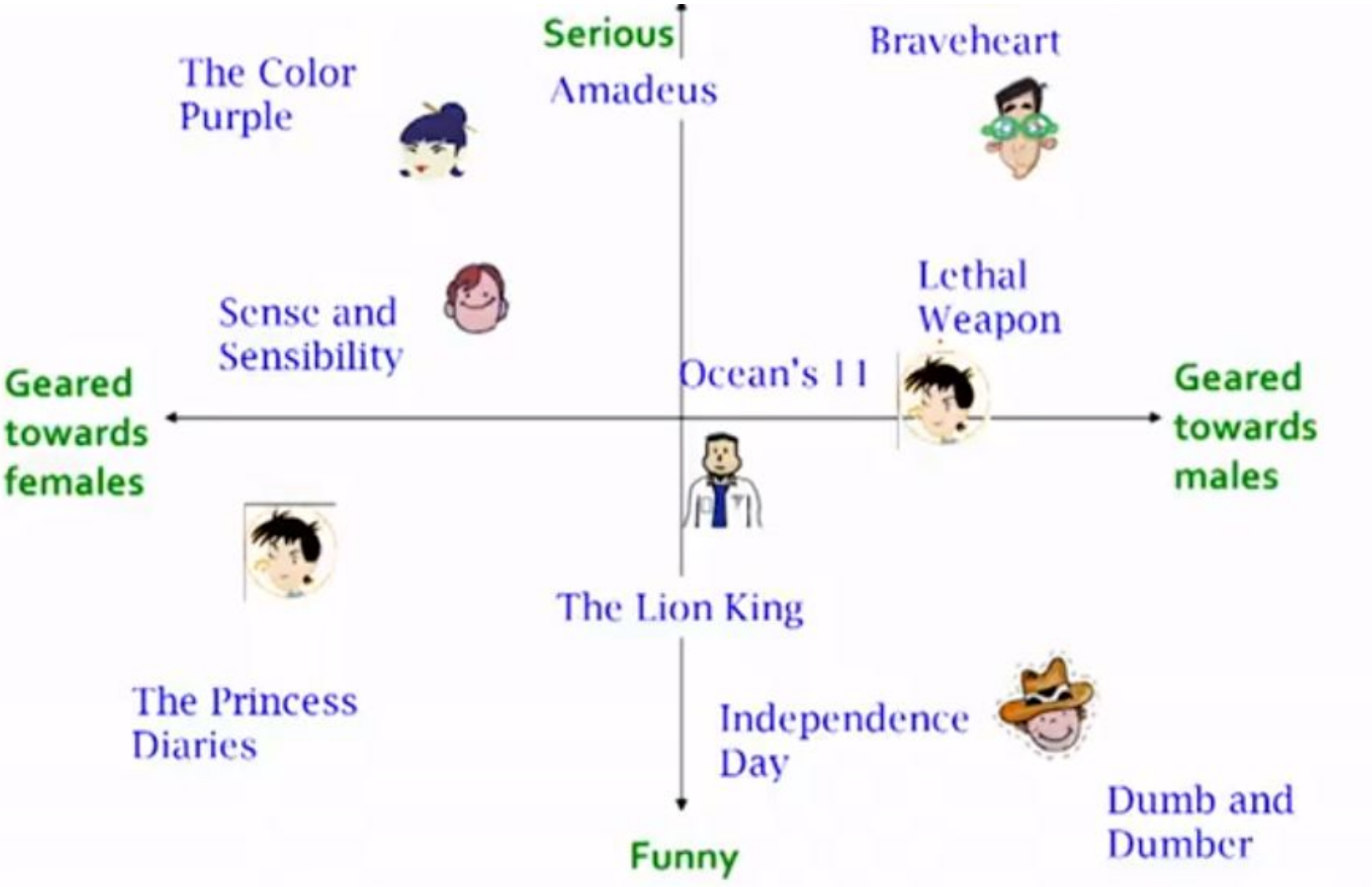
.

Matrix multiplication example

Matrix multiplication example from developers.google.com [1]

Here we assign value 1 to all positive examples which shows us that MF model can be easily applied to implicit feedback datasets.
In our case we will use ratings (scale 1-5) instead.

Latent space visualization



Visualisation of a latent space [2]

Meaning of the latent factors

Latent factors

In the example above we can see tags like funny, serious, geared towards females or males.
In a basic matrix factorization model **we do not know the meaning of any specific coordinate of latent vectors**.
However we can train our model and then realize that a specific coordinate is related with a specific feature.

Meaning of the latent factors

Latent factors

In the example above we can see tags like funny, serious, geared towards females or males.
In a basic matrix factorization model **we do not know the meaning of any specific coordinate of latent vectors**.
However we can train our model and then realize that a specific coordinate is related with a specific feature.

Note that in a basic MF it is not possible to create an embedding for user or item without any interaction.

Latent factors as a function of observed attributes

There exists many ways of utilizing user and item attributes in MF models. One of the most popular and widely used in practise is LightFM https://making.lyst.com/lightfm/docs/home.

Singular Value Decomposition

Singular Value Decomposition theorem

For any $m \times n$ real matrix M there exist orthogonal real matrices $U \in \mathbb{R}^{m \times m}, V \in \mathbb{R}^{n \times n}$ and non-negative diagonal matrix $\Sigma \in \mathbb{R}^{m \times n}$ such that:

$$M = U \Sigma V^T.$$

In explicit we can build U from orthonormal eigenvectors of MM^t, V from orthonormal eigenvectors of $M^t M$ and Σ from singular values of M. It is useful to sort singular value decreasing on the diagonal of the matrix Σ.

Truncated matrices

We can look for smaller matrices $U \in \mathbb{R}^{m \times k}, V \in \mathbb{R}^{n \times k}$ to approximate matrix M with respect to Frobenius norm.
It follows from Eckart-Young-Mirsky Theorem that to achieve it, it is enough to truncate matrices received in SVD.

SVD in recommender systems

Simple approach to use SVD in recommendation systems is the following:

1. Replace missing entries in user-item ratings matrix R by users' means.
2. Decompose received matrix with SVD.
3. Restrict number of latent features to k and truncate received matrices.
4. Use gradient descent to decrease RMSE on known ratings.

Large SVD models on Netflix dataset resulted in RMSE up to 0.905 on the probe set.

If the idea of matrix factorization of gradient descent is not clear for you, you can watch really simple explanations "How does Netflix recommend movies? Matrix Factorization" here.

Python implementation

To do (especially for absent students):

- Go through - *P4. Matrix Factorization* notebook to:
  - check implementation of SVD algorithm
  - observe how learning process influence RMSE on train and test sets
  - observe evaluation measures of SVD
  - learn about embeddings and a way to generate similar items in MF model
  - go to *P4. Appendix - embeddings in high dimensional spaces.ipynb* to see the properties of embeddings in high dimensional spaces
  - **project task 5: implement SVD on top baseline (as it is in Surprise library)** - making changes to our implementation by considering additional parameters in the gradient descent procedure seems to be the fastest option

OLX GROUP

ADAM MICKIEWICZ UNIVERSITY POZNAŃ

References I

[1] Google, "Matrix factorization", https://developers.google.com/machine-learning/recommendation/collaborative/matrix.

[2] S. University, "Lecture 55 — latent factor recommender system", https://www.youtube.com/watch?v=E8aMcwmqsTg&list=PLLssT5z_DsK9JDLcT8T62VtzwyW9LNepV&index=55.