The final predictions given by the recommendation system are the mean of two different predictions.

1. SVD prediction
2. Baseline prediction

A more comprehensive reading is available at http://sifter.org/~simon/journal/20061211.html. Also search keywords "latent factor model" on Youtube.

3.1.1 SVD prediction:

SVD prediction is a matrix completion method. It predicts ratings by decomposing the `urm` which is the user-rating matrix. `urm` is a `ixj` matrix where `i` and `j` are, respectively, number of users and number of films. Every `(i,j)`th element is the rating user `i` gave for the movie `j`. For a movie `j` that user `i` hasn't rated, the `(i,j)`th element is 0.

However, for implementing proper SVD, we cannot have empty, or zero, elements in the user-rating matrix. We fix this problem by mean normalization. For every movie the user hasn't rated, we consider its rating to be the average of the user's ratings for the movies he/she has rated.

We have two matrices by this time: `original_matrix` and `urm`. The 0 elements in `urm` have been normalized and the `original_matrix` still has all of its zeros. We do this so that we can later classify which ratings are real and which are predicted by our program.

SVD is performed on the normalized `urm`, which gives us `U`, `S` and `V` matrices.

$$urm = USV$$

The matrices `U` and `V` give us user-taste and film-genre ideas respectively. The matrix `S` gives us the strength of how much a film is a particular genre and how much a user likes a particular genre. To recompose the `urm` from matrices `U`, `S` and `V`, we multiply `S` and `V`'s transpose. The resulting matrix `P` connects the films to the strength of their genres.

$$P = S * V^T$$

Then, for every `(i,j)`th we want to find predicted rating for, we multiply (dot product) `i`th row from matrix `U` and `j`th column from matrix `P`. Prediction for rating given by user 1 for film 12 is given by:

$$F = U[1] * P[12]$$

We can construct a new predictions matrix that has all the guessed ratings.

Therefore, we start out with a sparse matrix of user ratings and in the end have the same matrix filled with predicted values based on existing values.

3.1.2 Baseline Prediction:

To predict the rating given by user `i` for film `j`, we first find `ratings_mean`, which is the mean of all ratings in `original_matrix`. Then we find the mean `film_ratings_mean` of the ratings given by different users for the film we have to predict rating for. We subtract `ratings_mean` from `film_ratings_mean` to see how well-rated this film is compared to all other movies in the database.

Similarly, we get the intuition on the user's taste based on how critical they are with their ratings. We find the mean `user_ratings_mean` of ratings the user has so far given. We subtract `ratings_mean` from `user_ratings_mean` to see how different the user's taste is from the rest of the database. Then we add all three:

$$ratings\_mean + (film\_ratings\_mean - ratings\_mean) + (user\_ratings\_mean - ratings\_mean)$$

This gives us another naïve rating.

The final prediction is the mean between this rating and the rating generated from SVD prediction.