Miniproject 2:
Visualization of Latent Factors from Movies

Ryan Langman, Naoki Eto, and Ryan Casey

Problem 1

We implemented three training algorithms: stochastic gradient descent and alternating least squares for the basic latent factors model described on slide 5 of the mini-project guide, and stochastic gradient descent for the "even more advanced" model on slide 8.

The visualizations were made based off gradient descent of the even more advanced model, with the global average as a baseline rating, parameters a and b as user and movie biases, and dot(u, v) as the influence from latent features. We chose this model because it seemed to have the smallest amount of ambiguity in what the features actually represent, which would hopefully make the result more interpretable.

This model was trained using a learning rate of 0.005, a regularization factor lambda of 0.01, and stopped after 500 epochs. The learning rate and regularization were basically just chosen through trial and error, and the stopping criteria was a result of looking at the in-sample RMSE after every iteration and stopping it when the decrease in error after each iteration was too small to have much effect.

Overall we did not put too much emphasis on trying to optimize parameters or training because such optimizations would probably only have minor effects on the actual visualization.

Problem 2b - Explanation of Plots

In batman_starwars_disney_diehard.png, we can see that all of the Star Wars movies are clustered on the graph together, as well as the Die Hard movies. The Disney movies are rather spread out on the graph, but movies like “The Fox and the Hound” are rather different from movies like “Pocahontas”, as “Pocahontas” is a more romantic and musical movie. But the majority of the Disney movies lie around (-0.4, 0). The Batman movies all lie in the bottom left quadrant, as they negatively trigger both features. Feature 1 (on the X-axis) is triggered in the positive direction for Star Wars movies and the Die Hard movies,

In the filmnoir.png, we see that “Raw Deal (1948)” is significantly far away from the rest of the film noir genre movies. This is one of the three movies that are only classified in “film noir” genre, with two of those movies only having one data point in the data.txt file. So, perhaps this genre is just too obscure that our features don’t really know what to do with it.

We also see that “Akira (1988)” is significantly far away from the rest of the Adventure films, or Animation films, and is not grouped very well with the rest of the Sci-fi or Thriller films. “Akira (1988)” is the only film in the data set that has these four genres, perhaps a reason why it is not really grouped with any of the genres that well. “Akira (1988)” may benefit from being classified as a new genre, Japanese Animation, since American Animation, such as Disney, is quite different from Japanese Animation.

We’ve included a PNG that plots all the movies in each movie genre, and for the most part we can see that each movie genre, and we can see that each genre tends to occupy a particular space in the 2-dimensional representation. Children and Animation movies tend to trigger Feature 1 (on the X-axis) fairly negatively, while Action and Adventure only have a slight bias in this direction. Sci-Fi, Thriller, and War movies have higher activations for Feature 1. So perhaps one interpretation is that Feature 1 activates for movies with higher levels of action, while it deactivates (or more accurately, is activated negatively) for movies for young people that don’t really have a lot going on. It is difficult to say what Feature 2 (on the Y-axis) might represent, since most genres are spread out pretty evenly in the vertical direction and centered pretty close to Y = 0. Action, Horror, War, and Sci-Fi movies have the largest skew in Y in the negative direction. Overall across all movies, we tend to see a cluster slightly to the left of the origin, and the the X-axis. This probably means that that Feature 1 is significantly more important that Feature 2, and we just have more points that negatively activate Feature 1 in the training set.