

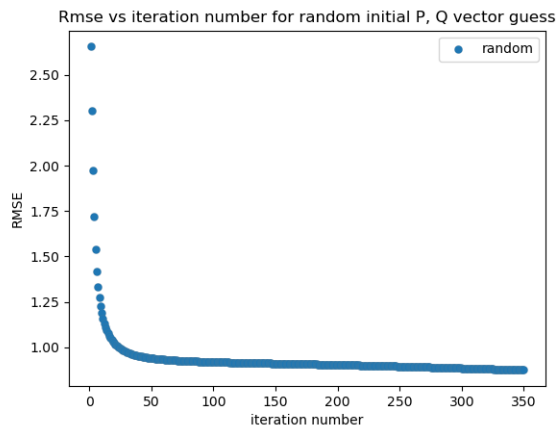
Computer Simulation II- PYU44C01 Linear algebra assignment 2

Matrix factorisation - recommender systems

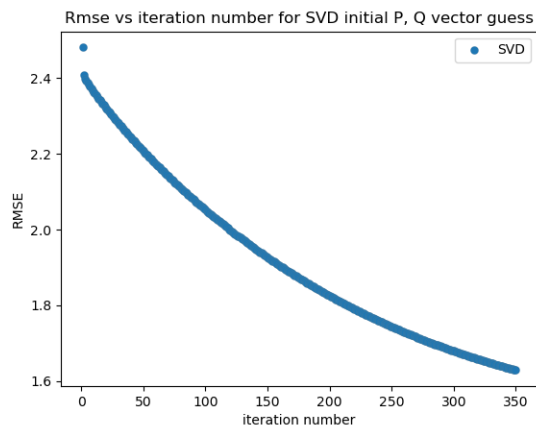
Name: Benjamin Stott, ID: 18336161

April 11, 2022

0.1 RMSE vs iterations for random initial matrix P and Q $k = 4$ (decreases at each step as required)



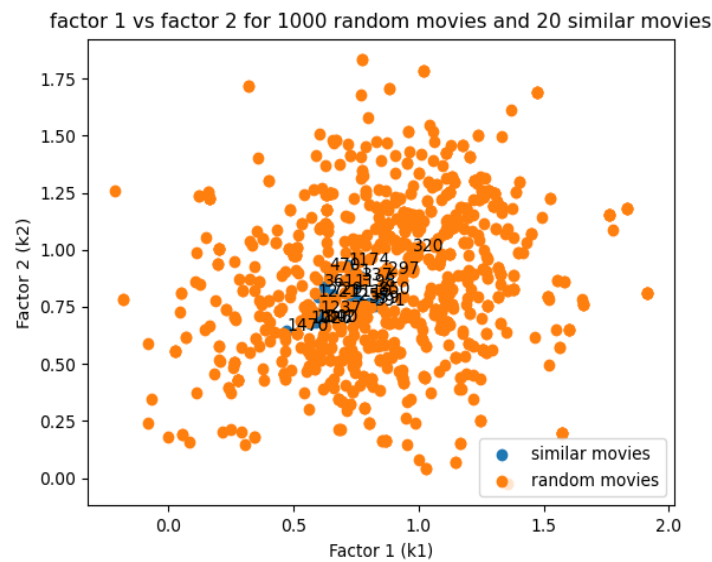
0.2 RMSE vs iterations for SVD initial matrix P and Q $k = 4$



As the iteration number increases, so too the rmse decreases for this sample as expected. This takes longer to converge though which could be due to the fact that I had to replace NaN values in the R matrix with 0s so that the SVD process could run without error(i.e the matrix is now sufficiently different to the starting matrix that it behaves differently). Since SVD is reliant on having a full matrix, an educated guess at the NaN values would likely serve better(e.g an average rating for a particular movie to fill a blank spot as opposed to a zero).

0.3 Factor 1 vs Factor 2, 20 similar movies(dense collection of labelled points and 1000 random movies from the u1.test set)

0.4



(a) factor 1 vs factor 2

0.5 what does factor 1 vs factor 2 tell us?

The similar movies are arranged as a cluster of points while dissimilar movies randomly occupy the surrounding space according to their particular characteristics. Movies that are similar can be traced out in a straight line, correlation would then decreases as one goes to either side of this vector. Note this would be a vector with maximum variance and not say one traced through the middle densest points only. This is an effective method of recommending movies as latent features correlated with the items and users tastes can be used to create a vector representing their specific needs. The data is also easily interpreted as it can be observed visually how it functions by the naked eye.