COMS 4721 Assignment 4
Matrix Factorization

John Fulgoni

Columbia University in the City of New York

14 April 2015

List of Figures

1   Problem 1

1.1 Part 1

Figure 1 shows the objective function for K means over 20 iterations, and Figure 2 shows the Scatter Plot for K = 2, 3, 4, and 5 for the 500 data pointsm each with a different color/symbol combination.

2   Problem 2

2.1 Part 1

Figure 3 shows the Root Mean Squared Error over 100 iterations of running the MAP Inference Algorithm. Figure 4 shows the log likelihood for the algorithm over 100 iterations.

2.2 Part 2

For my three movies, I chose Pulp Fiction, Jurassic Park, and Angels in the Outfield. Figure 5 shows movies that are the closest euclidean distance to Pulp Fiction, which include some
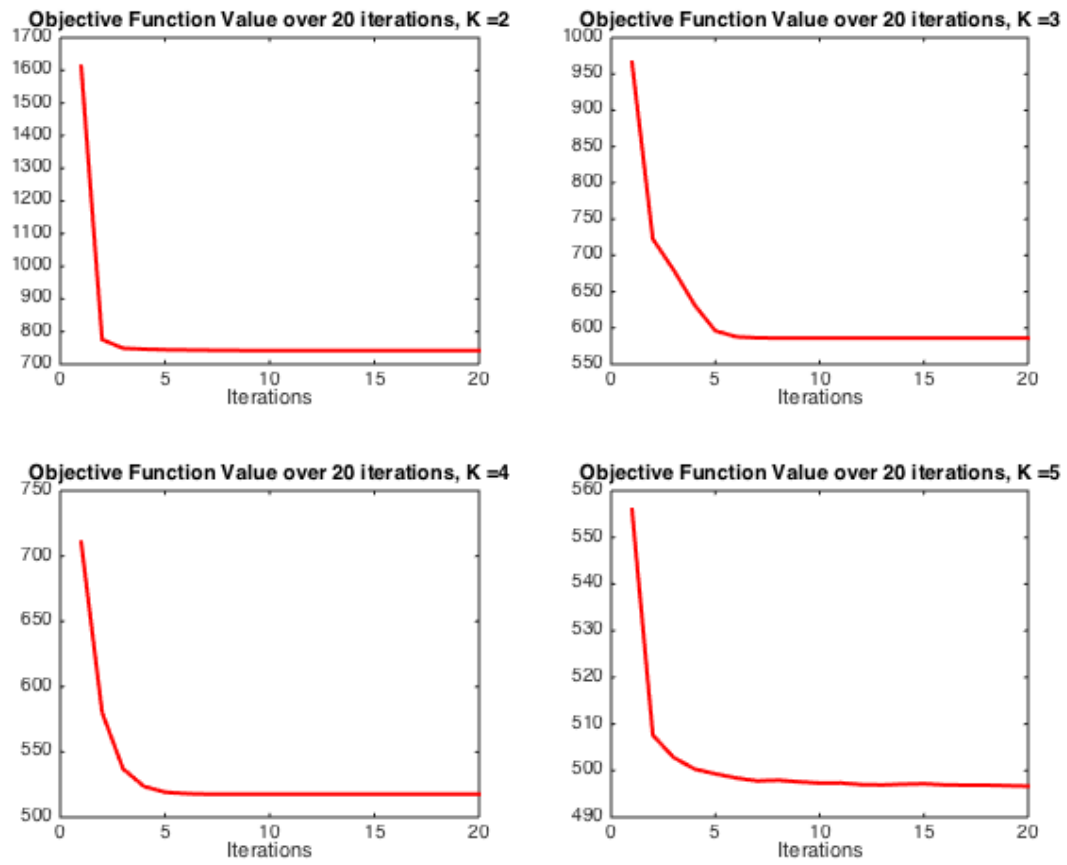
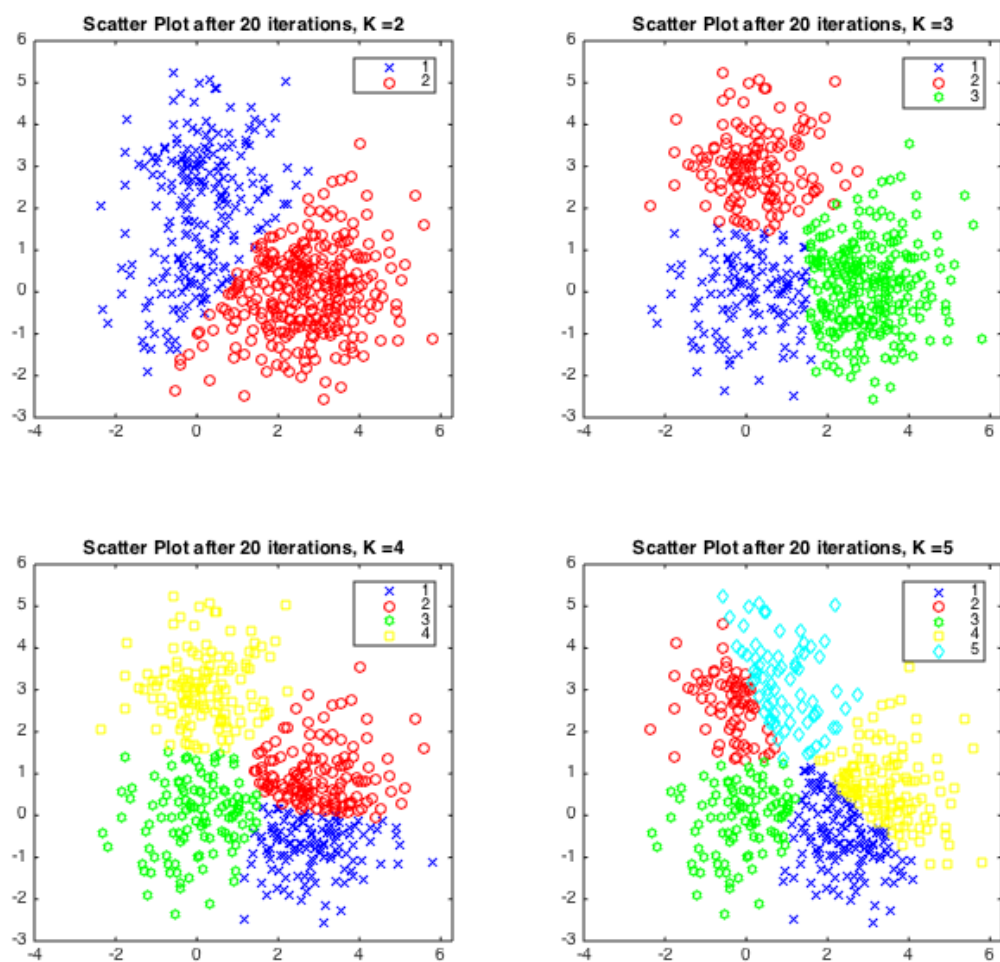Figure 1: Objective Function Plots for K=2, 3, 4, 5
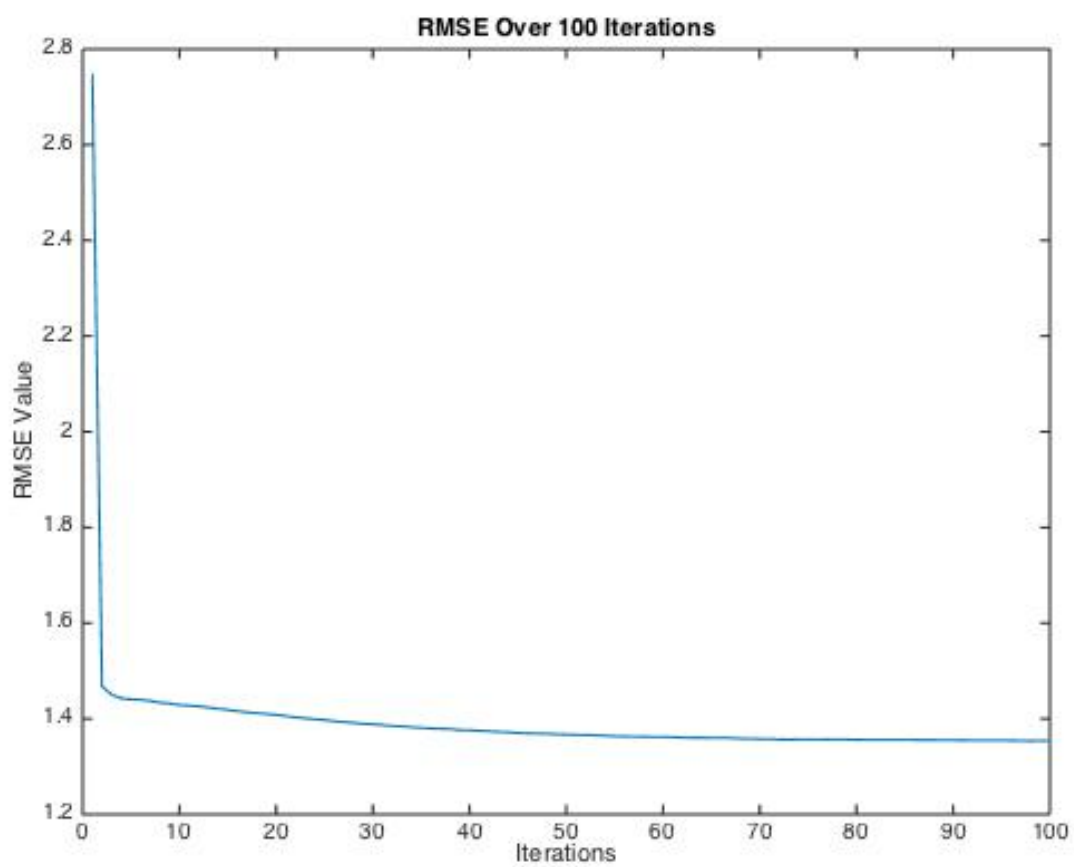
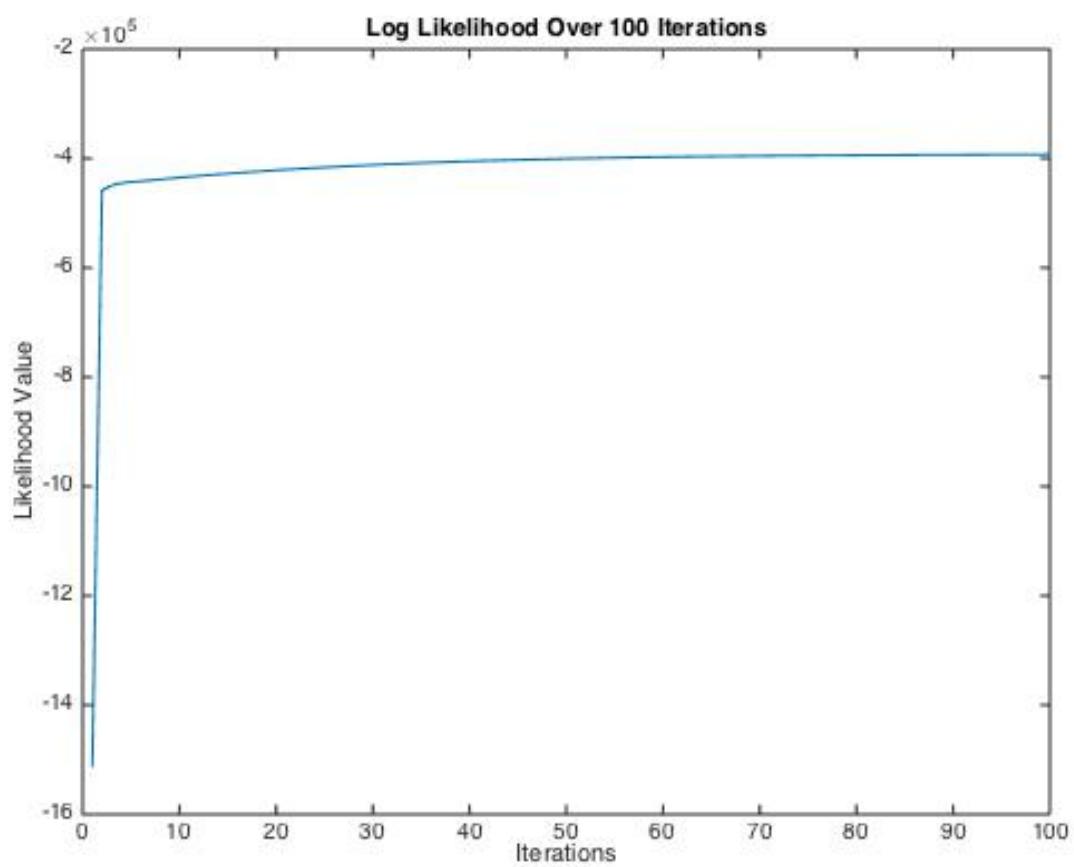Figure 2: Scatter Plots for K=2, 3, 4, 5

Figure 3: RMSE Over 100 Iterations

Figure 4: Log Likelihood Over 100 Iterations

Bye Bye, Love (1995)	0.8956
Blink (1994)	0.96092
Hackers (1995)	0.98197
Of Love and Shadows (1994)	0.98565
Bad Moon (1996)	0.9907

Figure 5: Closest Movies to Pulp Fiction

Indiana Jones and the Last Crusade (1989)	0.9127
Back to the Future (1985)	1.1056
Fugitive, The (1993)	1.1975
Wild America (1997)	1.2481
Speed (1994)	1.2524

Figure 6: Closest Movies to Jurassic Park

other crime dramas. Figure 6 shows movies close to Jurassic Park, which include other action and thriller movies. Figure 7 shows movies similar to Angels in the Outfield, which include other kid's based sports movies.

2.3   Part 3

For part three I decided to use the built in Matlab k-means function instead of my own implementation. Cluster 1 seems to be a lot of Humphrey Bogart - Old time movies. Cluster 2 I believe is a combination of action/easy going movies. Cluster 3 seems to be a lot of comedies, but also outdoors movies? Cluster 4 seems to be political dramas, with the exception of the Wedding Singer. Cluster 5 seems to be all thriller/action based movies.

Cluster 1   Meet John Doe (1941) U.S. Marshalls (1998) Schindler's List (1993) Dial M for Murder (1954) Casablanca (1942) Mary Poppins (1964) House of Yes, The (1997) Pretty Woman (1990) Cyrano de Bergerac (1990) Young Frankenstein (1974)

Cluster 2   Roman Holiday (1953) Rebel Without a Cause (1955) Extreme Measures (1996) Graduate, The (1967) Sunset Blvd. (1950) Daytrippers, The (1996) 3 Ninjas: High Noon At Mega Mountain (1998) Batman & Robin (1997) Body Snatchers (1993) Money Talks (1997)

Big Green, The (1995)	0.93053
For Richer or Poorer (1997)	0.93769
Operation Dumbo Drop (1995)	0.97804
Tom and Huck (1995)	1.0074
Preacher's Wife, The (1996)	1.0434

Figure 7: Closest Movies to Angels in the Outfield

Cluster 3   Quiz Show (1994) unknown Whole Wide World, The (1996) North by Northwest (1959) Hearts and Minds (1996) Davy Crockett, King of the Wild Frontier (1955) Immortal Beloved (1994) Fled (1996) Billy Madison (1995) Killing Fields, The (1984)

Cluster 4   Year of the Horse (1997) Man of the Year (1995) Citizen Ruth (1996) Great Dictator, The (1940) Money Train (1995) Dangerous Minds (1995) Doom Generation, The (1995) Omen, The (1976) Wedding Singer, The (1998) Moll Flanders (1996)

Cluster 5   Body Snatcher, The (1945) Apollo 13 (1995) Across the Sea of Time (1995) Copycat (1995) Crooklyn (1994) Blob, The (1958) Timecop (1994) Only You (1994) Jaws 3-D (1983) His Girl Friday (1940)