

Problem 1

1. The code for calculate the Markov matrix, w_t and the top teams is **markov.m**, which needs an input of step t, the data files scores and legends.

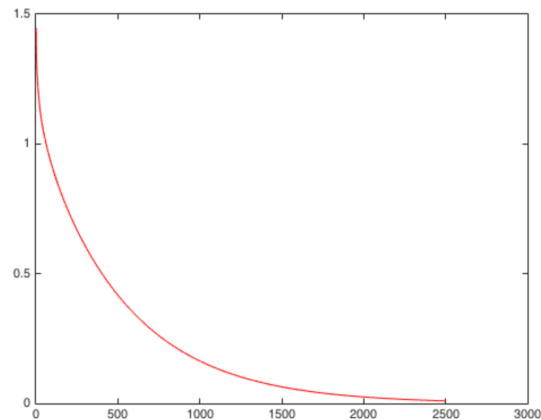
Top 25 teams when t = 10 are: Alabama, Clemson, Mississippi, Ohio State, Oklahoma, Stanford, Michigan St, Houston, Michigan, Iowa, TCU, Florida, LSU, Arkansas, Notre Dame, Florida St, Tennessee, Baylor, Utah, Navy, Oklahoma St, Northwestern, North Carolina, Wisconsin and Oregon.

Top 25 when t = 100 are: Mount Union, Alabama, NW Missouri St, St Thomas, Clemson, Mississippi, Ohio State, Oklahoma, Marian IN, Stanford, Michigan St, Houston, St Francis IN, Linfield, Michigan, Iowa, TCU, Wesley, Emporia St, Florida, LSU, Arkansas, Notre Dame, UW-Whitewater, Florida St.

Top 25 when t = 1000 are: Alabama, Clemson, Mississippi, Ohio State, Oklahoma, Stanford, Michigan St, Houston, Michigan, TCU, Iowa, Florida, LSU, Arkansas, Notre Dame, NW Missouri St, Florida St, Tennessee, Mount Union, Baylor, Utah, Navy, Oklahoma St, Northwestern, North Carolina.

Top 25 when t = 2500 are: Alabama, Clemson, Mississippi, Ohio State, Oklahoma, Stanford, Michigan St, Houston, Michigan, Iowa, TCU, Florida, LSU, Arkansas, Notre Dame, Florida St, Tennessee, Baylor, Utah, Navy, Oklahoma St, Northwestern, North Carolina, Wisconsin, Oregon.

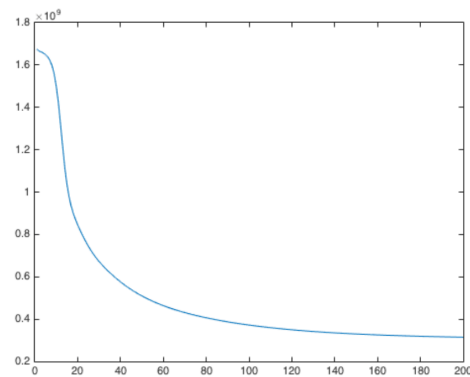
2. Use **computew.m** to compute w_∞ and plot the image of $\|w_t - w_\infty\|_1$ as a function of t. The image is as follow.



Problem 2

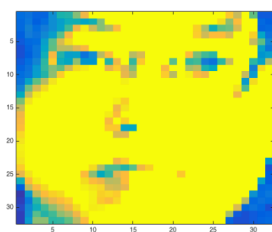
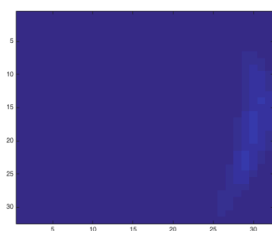
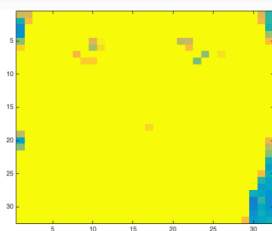
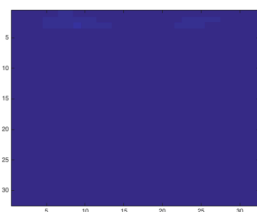
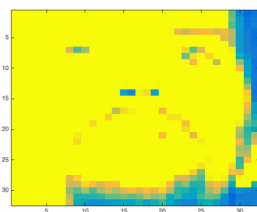
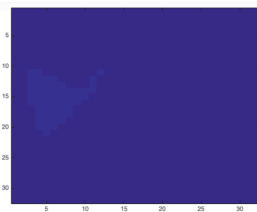
Part 1

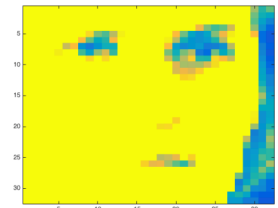
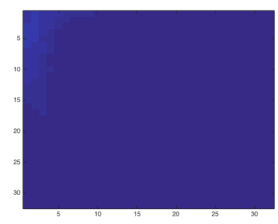
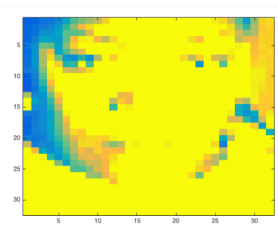
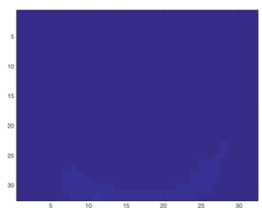
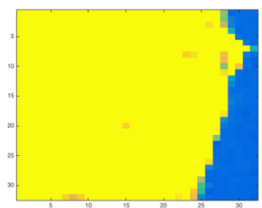
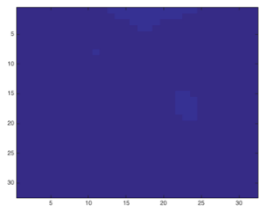
1. Use code **nmf.m** to do the NMF algorithm. And the plot of objective function is:

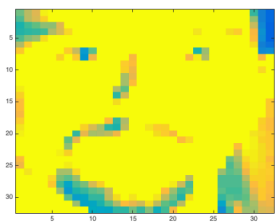
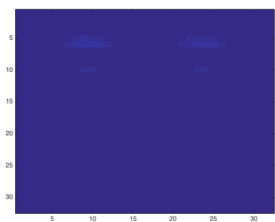
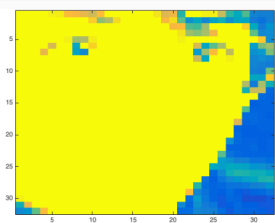
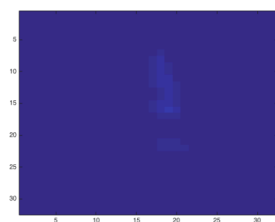
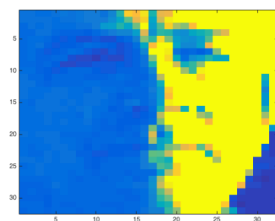
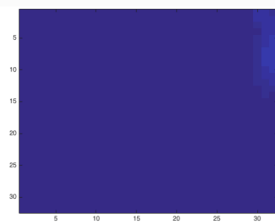


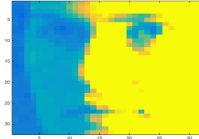
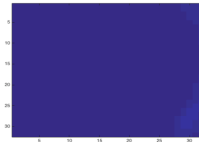
2. I pick the first 10 columns of W, and the corresponding weights in H are: 47.1278928129790; 49.8920393350298; 50.9263833210611; 55.7452571825652; 74.3954740346915; 47.4217666990100; 34.6478783788585; 46.2314956607690; 37.8135967594162; 67.7623021359180.

The images of W and the corresponding X are (first is W1, then is X1, and W2, X2,...):



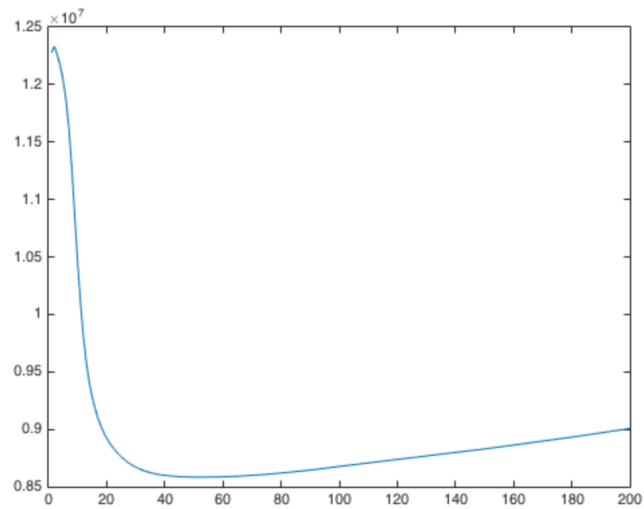






Part 2

1. Use code **nyt.m** to do the NMF algorithm. And the objective function is:



2. Still, pick the first 10 columns of W . Use **topic.m** to get the 10 topics for each document. Topics are:

- (1) system computer technology service company number program information allow available
- (2) political campaign vote party election republican candidate democratic voter leader
- (3) official report agency member spokesman public investigation office yesterday tell
- (4) talk plan leader meeting agreement meet agree deal peace negotiation
- (5) art artist exhibition painting museum collection gallery design photograph color
- (6) president executive vice director chief name chairman senior group television
- (7) states american country government world war foreign military international nation
- (8) father graduate son mrs mother daughter marry receive yesterday degree
- (9) write book fact life history point great read question world
- (10) thing tell ask feel little keep lot put talk start