

INSTRUCTIONS

- **Due:** Thursday, 30 April 2020 at 11:59 PM EDT.
- **Format:** Complete this pdf with your work and answers. Whether you edit the latex source, use a pdf annotator, or hand write / scan, make sure that your answers (tex'ed, typed, or handwritten) are within the dedicated regions for each question/part. If you do not follow this format, we may deduct points.
- **How to submit:** Submit a pdf with your answers on Gradescope. Log in and click on our class 10-315, click on the appropriate *Written* assignment, and upload your pdf containing your answers. Don't forget to submit the associated *Programming* component on Gradescope if there is any programming required.
- **Policy:** See the course website for homework policies and Academic Integrity.

Name	
Andrew ID	
Hours to complete (both written and programming)?	

For staff use only

Q1	Q2	Q3	Total
/10	/6	/12	/ 28

Q1. [10pts] Programming: Recommender Systems

The following questions should be completed after you work through the programming portion of this assignment.

After running `matrix_factorization_alt_min` with `K=20`, `alpha=0.001`, and `num_epoch=200`, which of the first 10 users (user indices 0-9) do you predict would rate The Lightning Thief (book index 6) the highest? Which would rate it the lowest? What are these respective ratings?

- (a) [5pts] Which user would rate it the *highest*? What is the predicted rating?

User index:

Rating:

- (b) [5pts] Which user would rate it the *lowest*? What is the predicted rating?

User index:

Rating:

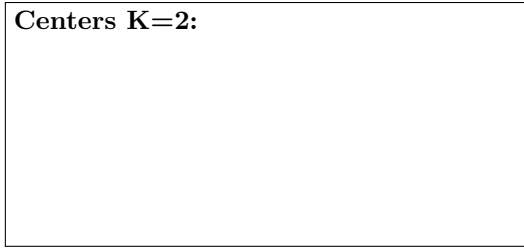
Q2. [6pts] Programming: K-means

The following questions should be completed after you work through the programming portion of this assignment.

(a) [2pts] K=2

Include the images of the cluster centers after running k-means with two clusters.

Centers K=2:



(b) [2pts] K=5

Include the images of the cluster centers after running k-means with five clusters.

Centers K=5:



(c) [2pts] K=10

Include the images of the cluster centers after running k-means with ten clusters.

Centers K=10:



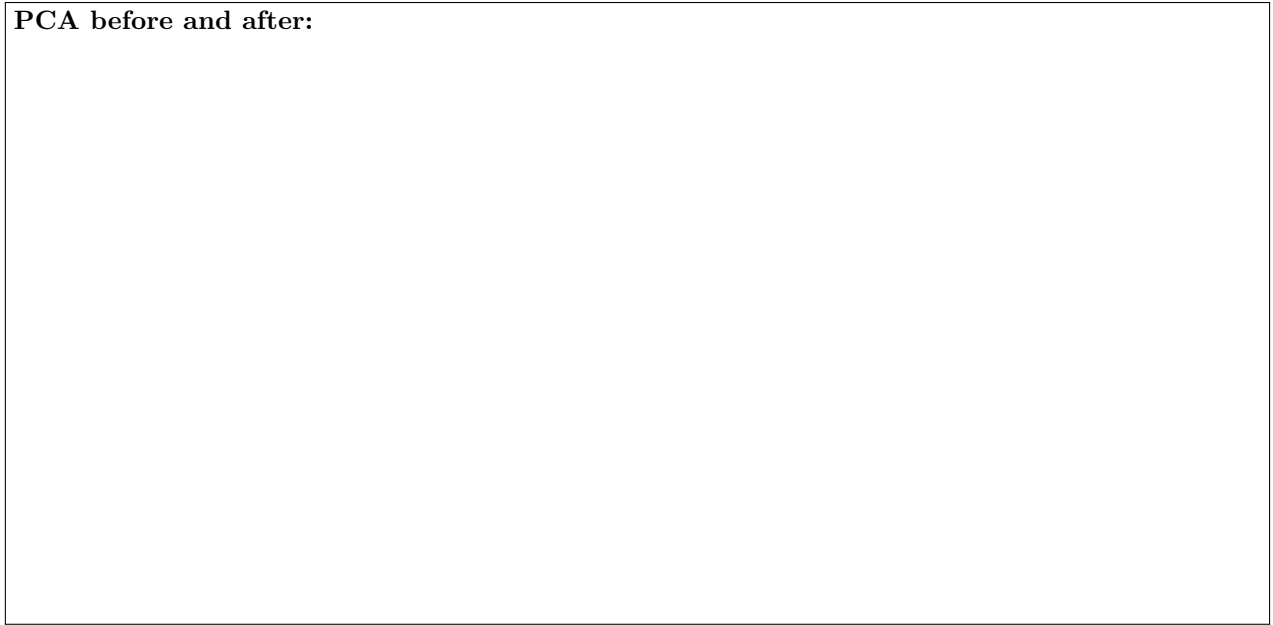
Q3. [12pts] Programming: GMM

The following questions should be completed after you work through the programming portion of this assignment.

(a) [4pts] PCA

Include the plots of the toy dataset before and after running PCA with $K=2$.

PCA before and after:



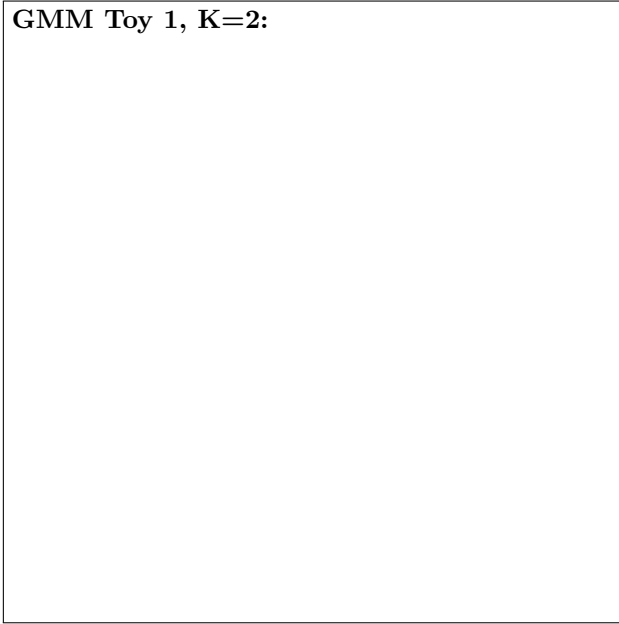
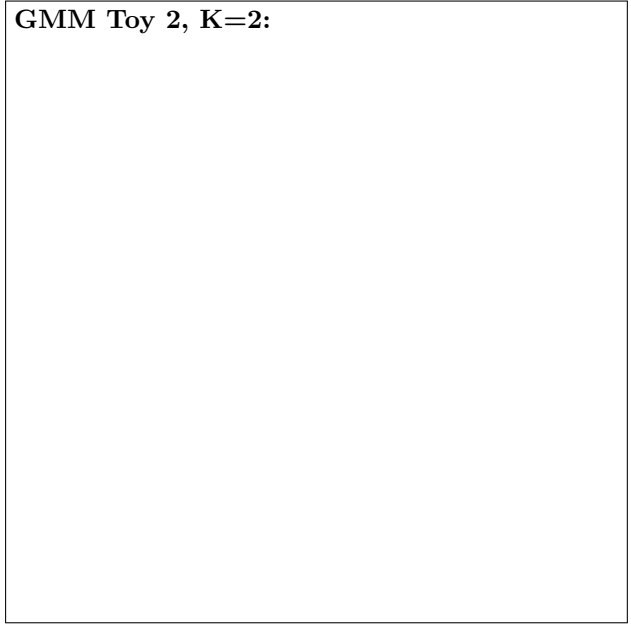
Include the plots of the MNIST zeros and ones dataset after running PCA with $K=2$.

PCA MNIST:

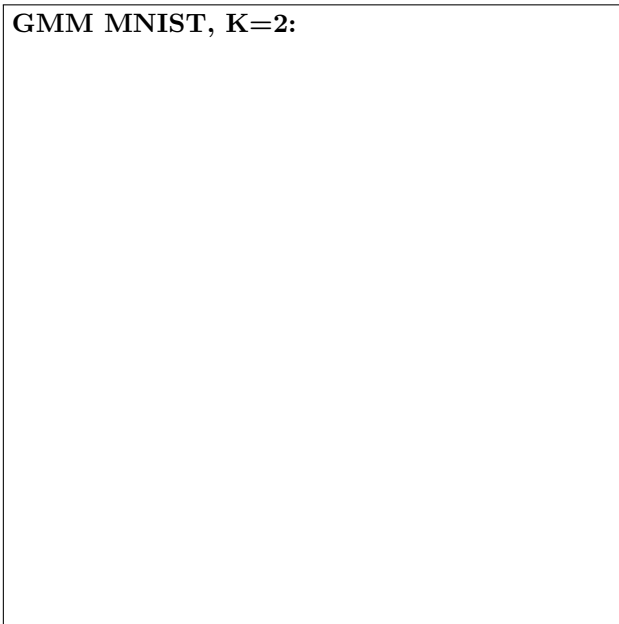


(b) [4pts] GMM Toy Datasets

Include the plots of after learning the GMM parameters for $K=2$ on toy dataset one and two.

GMM Toy 1, $K=2$:**GMM Toy 2, $K=2$:****(c)** [4pts] GMM MNIST Zeros and Ones

Include the plots of after learning the GMM parameters for $K=2$ and $K=5$ on the MNIST zeros and ones dataset.

GMM MNIST, $K=2$:**GMM MNIST, $K=5$:**