MA 323: Project 1

Due date: Oct 29, 2025

Problem 1. The MNIST (Modified National Institute of Standards and Technology) dataset is a classic benchmark in machine learning and computer vision, widely used for testing image classification algorithms. It contains 70,000 grayscale images of handwritten digits from 0 to 9, each normalized and centered in a 28×28 pixel frame.

1. Randomly select 5,000 images from the dataset and fix them for use in the subsequent tasks.

2. View each image as a vector in $\mathbb{R}^{28\times 28=784}$. Compute the pairwise distance matrix under the ℓ_1, ℓ_2, and ℓ_∞ norms, and visualize your results using a heatmap. To better reveal the structure of the data, you may arrange the rows based on the label information of the images (i.e., group all 0's together, 1's together, etc.).

3. Next, view each image as a 2D distribution. One way to do this is by assuming each pixel "exists" if its value is above a certain threshold (e.g., 80). Then compute the pairwise W_1-distance between these images and repeat the procedure from the previous step. For this step, you may use the `ot` library (assuming you are using Python) to compute the W_1 distance. Note that this step may be time-consuming since computing W_1 is relatively expensive; one way to mitigate this is to parallelize your computation. Briefly summarize these results and compare them with those from the previous step.

4. Select all images whose labels correspond to your favorite number. Vectorize each image as an element of $\mathbb{R}^{28\times 28=784}$, and stack the resulting vectors into a matrix (with each column representing one image). Use SVD to compute a low-rank approximation of this matrix, and determine the smallest rank that still provides a reasonable approximation of the original matrix. This rank should reflect some intrinsic complexity of the class you chose.