ISYE 8803 HDDA

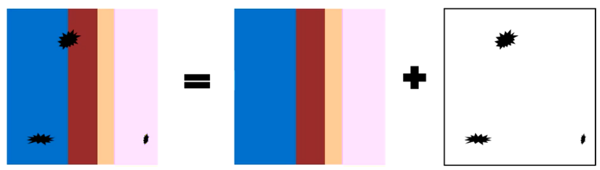
Exam 2

June 27th, 2021

a)

Provided in the lecture notes we can derive the following variables used in PCA:

Robust PCA



$$M = L + S$$

- M is the observed data matrix
- L is a low-rank matrix (to be estimated)
- S is a matrix of sparse outliers (to be estimated)

By finding L and S, we can retrieve low-dimensional linear structure from non-ideal observations.

The formulation for the minimization problem before convex relaxation can be defined as follows:

$$\min\{\text{rank}(L) + \lambda \|S\|_0\}$$

subject to: $\quad M = L + S$

$$\text{rank}(L) = \#\,\{\sigma(L) \neq 0\} \qquad \|S\|_0 = \#\{S_{ij} \neq 0\}$$

Neither the rank(L) or L0 norm is convex, so we need to relax these attributes to create a valid convex optimization problem:

$$\mathrm{rank}(L) = \#\{\sigma(L) \neq 0\} \qquad\qquad \|S\|_0 = \#\{S_{ij} \neq 0\}$$

$$\Big\downarrow \qquad \text{Convex Relaxation} \qquad \Big\downarrow$$

$$\|L\|_* = \sum_i \sigma(L) \qquad\qquad \|S\|_1 = \sum_{i,j} |S_{ij}|$$

Nuclear norm: sum of singular values

L_1 norm: sum of absolute values

b)

Step 1) Given S and Y, and minimizing for L, we can see that the optimization problem gets simplied and we can cancel out terms, then we can derive the frobenius norm of $X - L$. Then with Single value thresholding we can derive L for the next step.

Step 2) Then S get's updated based on the soft-thresholding algorithm.

We then repeat this process until convergence.

- Augmented Langrangian Multiplier form:

$$l(L, S, Y) = \|L\|_* + \lambda\|S\|_1 + \langle Y, M - L - S\rangle + \frac{\mu}{2}\|M - L - S\|_F^2.$$

$$l(L, S, Y; \mu) = \|L\|_* + \lambda\|S\|_1 + \frac{\mu}{2}\left\|M - L - S + \frac{Y}{\mu}\right\|_F^2 + \frac{\mu}{2}\left\|\frac{Y}{\mu}\right\|_F^2$$

Main Idea:

- Given S and Y, Update L

$$\arg\min_L \|L\|_* + \frac{\mu}{2}\left\|M - L - S + \frac{Y}{\mu}\right\|_F^2 \xrightarrow{X = L + S - \frac{Y}{\mu}} L = D_{1/\mu}(X) = US_{1/\mu}(\Sigma)V^T$$

- Given L and Y, Update S

$$\arg\min_S \lambda\|S\|_1 + \frac{\mu}{2}\left\|M - L - S + \frac{Y}{\mu}\right\|_F^2 \Rightarrow S_{ij} = S_{\frac{\lambda}{\mu}}(X) = \mathrm{sgn}(X)\max(|X| - \frac{\lambda}{\mu}, 0)$$

- Given L and S, Update Y

$$Y_{k+1} = Y_k + \mu(M - L - S)$$

c)

.

Figure 1: rank(L) Image output

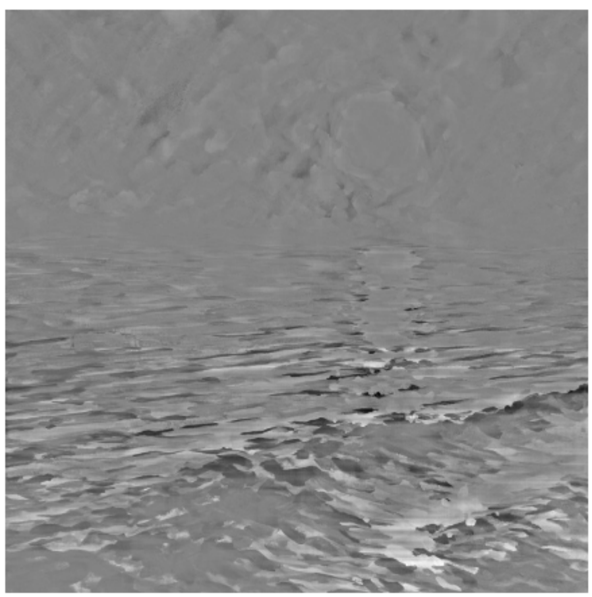


Figure 2: S matrix image for outliers

Figure 3: Original Image

d)
We can see when comparing the outputs, that the low rank L matrix output seemed to be focusing on features related to the following:

-The sun and it's corresponding reflection
-The sky and it's corresponding cloud cover
-The sea and anything realted to the tides

We can see from the Sparse matrix output that a lot of empahsis is places on the darker portion of the images representing the tide. When comparing to previous modules, it seems to have a very similar output compared to edge detection!!

We can see that the L image output seems to create a fairly accurate representation when comparing back to the original image, and accounts for outliers much more effectively than a conventional PCA implementation would yield.