Denoising with subspace learning
ST790 Milestone Report

Bowen Liu, Renfei Gong, Haoyu Chen

April 7, 2017

1 Introduction

Image noise are produced when being created, saved, transmitted or presented. In order to restore a corrupted image, various optimization methods are at hand (Chi et al., 2013), and most of them reqiure the knowledge of noise position to achieve a better performance. In most cases, however, we have no information about the exact positions of noise pixels, and common optimization methods do no better than blur the corrupted image. Inspired by minimizing nuclear norm method, we found dimensionality reduction technique can lower the noise in a significant way. In fact, random noise points often increase the rank of the original image matrix, and reducing rank can fix this problem. Hence our denoising problem is actually subspace learning with corrupted information.

1.1 The PCA approach

A traditional dimensionality reduction method with full information is Principal Component Analysis (PCA). We will first introduce the subspace learning problem in a PCA setting.

Let $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n)}$ be the column vectors in an image. In this project, we assume that $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n)}$ are iid with some certain distribution. In the expressions below, the vector $\mathbf{x}$ refers to a vector with the same distribution as $\mathbf{x}^{(i)}$. Let Π denote a projection matrix from $\mathbb{R}^p$ to an $r-$dimensional subspace. Our goal is to find such a projection matrix that minimize the mean square error

$$f(\Pi) = \mathbf{E}\|\mathbf{x} - \Pi\mathbf{x}\|_2^2 = \mathbf{E}\left(\|\mathbf{x}\|_2^2 - \mathbf{x}^\top \Pi \mathbf{x}\right) = \mathbf{E}\left(\|\mathbf{x}\|_2^2\right) - \langle \Pi, \mathbf{E}(\mathbf{x}\mathbf{x}^\top)\rangle,$$

where the inner product of two matrices $\langle A, B\rangle$ is defined as $\langle A, B\rangle = \mathrm{tr}(A^\top B)$. Since the first term $\mathbf{E}\left(\|\mathbf{x}\|_2^2\right)$ is free of Π, our goal is to minimize $-\langle \Pi, \mathbf{E}(\mathbf{x}\mathbf{x}^\top)\rangle$, or to maximize $\langle \Pi, \mathbf{E}(\mathbf{x}\mathbf{x}^\top)\rangle$. Denote $\mathbf{E}(\mathbf{x}\mathbf{x}^\top)$ by C, and the PCA approach to maximize $\langle \Pi, C\rangle$ is to take $\Pi = \sum_{i=1}^k \mathbf{v}_i \mathbf{v}_i^\top$, where $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are the k leading eigenvectors of C. Before implementing the PCA approach, we need to find an estimator for C.

2 Our Contribution

2.1 Realizing POPCA

In subspace learning problem where we have no control on which attributes are missing, Gonen et al. (2016) derived Partially Observed PCA (POPCA) method to account for the missing values.

That is, if denote $\hat{\mathbf{x}}$ as the observed values of true $\mathbf{x}$, we assume that

$$\hat{\mathbf{x}}_i = \begin{cases} \mathbf{x}_i, & \text{w.p. } 1-p \\ 0, & \text{w.p. } p \end{cases}$$

to get an unbiased estimator for matrix C, Mitliagkas et al. (2014) gives

$$\hat{C}_{ij} = \begin{cases} \frac{1}{(1-p)^2}\hat{\mathbf{x}}_i\hat{\mathbf{x}}_j, & i \neq j \\ \frac{1}{1-p}\hat{\mathbf{x}}_i^2, & i = j \end{cases}$$

We realized this method and the result is shown in the figure below.

2.2 Deriving the uniform noise case

In the denoising problem, however, we are faced with noised information instead of missing values. We assume each noise pixel has a uniformly distributed color value independently. Thus,

$$\hat{\mathbf{x}}_i = \begin{cases} \mathbf{x}_i, & \text{w.p. } 1-p \\ U_i, & \text{w.p. } p \end{cases},$$

where U_i is a uniformly distributed noise from color spectrum. For simplicity, we assume $U_i \sim \text{Unif}(0, 1)$. From this assumption, an unbiased estimator for C can be derived as

$$\hat{C}_{ij} = \begin{cases} \frac{\hat{\mathbf{x}}_i\hat{\mathbf{x}}_j - \frac{1}{2}\hat{\mathbf{x}}_i - \frac{1}{2}\hat{\mathbf{x}}_j + \frac{1}{4}p^2}{(1-p)^2}, & i \neq j \\ \frac{\hat{\mathbf{x}}_i^2 - \frac{1}{3}p}{1-p}, & i = j \end{cases}$$

The performance of this approach relies on the accuracy of estimating noise rate p. It can perform better than the previous method with an accurate p.

3 Prospect

3.1 Detecting Noises

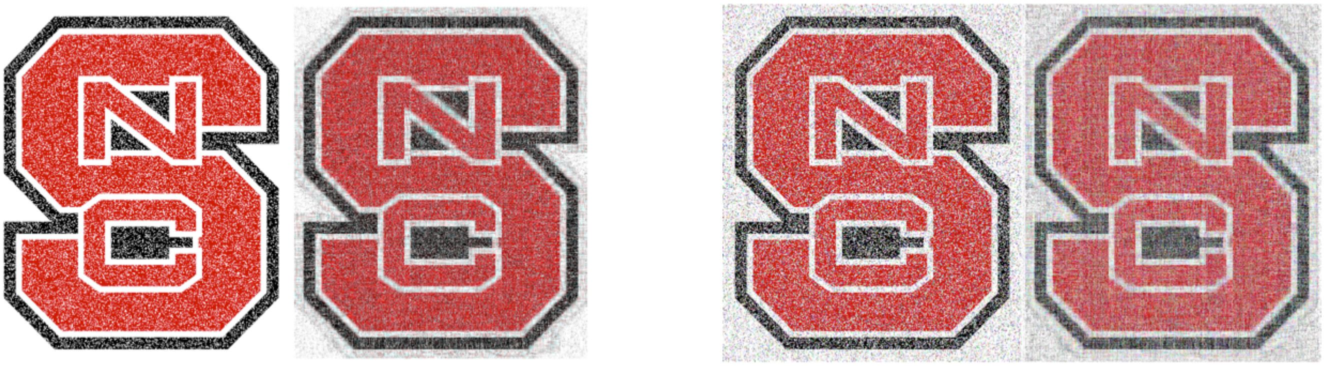
If we can detect the noised pixels, we can replace those pixels with some properly selected ones to make the image as continuous and smooth as possible. However, detecting the noises can be a tough task. An simple idea is, by assuming the original image to be generally smooth, we can label those pixels significantly different from other surrounding pixels as noised ones. Details need further discussion.

3.2 Inferring Noises

If the noises have a uniformly distributed color, it is entirely possible that a noised pixel doesn't have a significantly different color from the surrounding pixels. If we have to make a binary inference whether or not such a pixel is a noised one, there will be a high rate of either Type I or Type II error. Considering such a dilemma, we may estimate the probability that each pixel is a noised one. By doing so, we can estimate the matrix C with some statistical methods.

4 Appendix

4.1 Denoising on a single image



(a) An image with $p = 0.2$ missing pixels and POPCA solution with $rank = 40$

(b) An image with $p = 0.2$ random color noise and POPCA solution with $rank = 40$

Figure 1: A figure with two subfigures

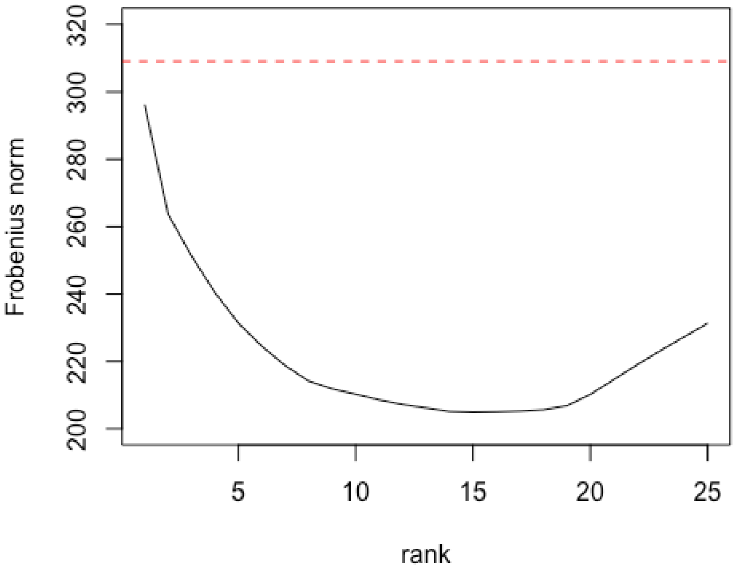


Figure 2: The Frobenius norm of the difference between original and restored image against restored image rank. The red dash line marks Frobenius norm of the difference between original and corrupted image