Convex optimization based low-rank matrix decomposition for image restoration

Alexey Bauman, Nikolay Kozyrskiy, Nikolay Skuratov

December 18th

Introduction

PCA seeks an optimal estimate of A, by the following constrained optimization:

$$\min_{A,E} \|E\|_F \qquad (1)$$
$$\text{s.t. rank } (A) \leq r, D = A + E$$

We can recover low-rank matrix A exactly from $D = A + E$ with gross but sparse errors E, by solving the following convex optimization problem:

$$\min_{A,E} \|A\|_* + \lambda\|E\|_1 \qquad (2)$$
$$\text{s.t. } D = A + E$$

Introduction

For a given vector $y \in R^n$ and the thresholding weight vector $w \in R^n_{++}$, the non-uniform singular value operator $S_w[y]$ satisfies:

$$S_w[y] = \arg\min_x \left(\frac{\mu}{2}\|x-y\|_2^2 + \|w \odot x\|_1\right) \qquad (3)$$

$\mu > 0$ is the penalty factor. As μ approaches 0, any solution to Eq. (3) approaches the solution set of (2). In other words, the non-uniform singular value operator satisfies:

$$S_w[y] = S_{w0}[z] = \arg\min_x \left(\frac{\mu}{2}\|x-z\|_2^2 + w_0\|x\|_1\right) \qquad (4)$$

where $z = \mathrm{sign}(y) \odot (|y| + w_0\mathbf{1} - w)$,
$\|y\| = \mathrm{sign}(y) \odot y$, and vector $1 \in R^n$ with all elements equal to 1.

Introduction

The singular value shrinkage operator: Consider the *SVD* of matrix $X \in R^{n_1 \times n_2}$ of rank r

$$Y = U\Sigma V^T, \quad \Sigma = \text{diag}\left(\{\sigma_i\}_1 \le i \le r\right)$$

For each $w \in R^r_{++}$, we introduce the soft-thresholding operator as follows: $D_w[Y] = US_w[\Sigma]V^T$, where $S_w[\Sigma] = \text{diag}\left(\left\{(\sigma_i - w_i)_+\right\}\right)$

Introduction

For each $w \in R^r_{++}$ and $Y \in R^{m\times n}$, the non-uniform singular value soft-thresholding operator obeys.

$$D_w[y] = D_{w_0}[z] = \arg\min_X \left(\frac{\mu}{2}\|X - Z\|_F^2 + w_0\|X\|_*\right)$$

where $Z = U\left(\Sigma + w_0 I - W\right)V^T, \quad Y = U\Sigma V^T$ denotes the problem defined in (6).$w_0 = \max\left(\{w_i\}\right), I \in R^{r\times r}$ is an unit matrix, and $W = \mathrm{diag}(w)$ is a diagonal matrix.

New Problem

We utilize a similar scheme for assigning weights to the singular value of a matrix. We attach a smaller weight to a larger singular value of the matrix, and vice versa. We obtain the following result for the nuclear norm minimization problem.

$$\begin{aligned} \min_{A,E} \textstyle\sum_{j=1}^{n} w_A \left(\sigma_j^2(A) + \gamma\right)^{p/2} + \lambda \left\| W_E \odot E \right\|_1 \\ \text{s.t.} \quad D = A + E \end{aligned}$$

Algo 1

Algorithm 1. Alternating direction method (ADM)

Step 1: Assuming iteration number $k=0$, initialize weight $w_A{}^{(0)} = \mathbf{1} \in R^n$, $W_E{}^{(0)} = \mathbf{1} \cdot \mathbf{1}^T \in R^{m\times n}$(assuming $m \geq n$).
Step 2: Using Algorithm 2 to solve the reweighted nuclear norm and ℓ_1-norm minimization problem

$$\left(A^{(k+1)}, E^{(k+1)}\right) = \underset{A,E}{\operatorname{argmin}} \left(\sum_{j=1}^{n} w_{A,j}^{(k)} \left(\sigma_j^2(A) + \gamma\right)^{p/2} + \lambda \| W_E^{(k)} \odot E \|_1, \text{s.t.} D = A + E \right), \tag{11}$$

where $w_A{}^{(k)}$ and $W_E{}^{(k)}$ are weight coefficients, we can obtain the estimated value $(A^{(k+1)}, E^{(k+1)})$. We defined $p=1$.
Step 3: For $i=1,\cdots,m$ and $j=1,\cdots,n$ update the weight coefficients as follows:

$$\begin{cases} w_{A,j}^{(k+1)} = \dfrac{1}{\sigma_j^{(k)}(A) + \varepsilon_A} \\ w_{E,ij}^{(k+1)} = \dfrac{1}{\left|E_{ij}^{(k)}\right| + \varepsilon_E} \end{cases}, \tag{12}$$

where ε_A and ε_E are positive preset constants, and the SVT matrix is

$$\Sigma^{(k)} = \operatorname{diag}\left(\left\{ \sigma_i^{(k)} \right\}_{i=1}^{r} \right) \in R^{n\times n}, \tag{13}$$

where $U^{(k)}, \Sigma^{(k)}, V^{(k)}$ are generated by the following SVD:

$$D - A^{(k+1)} + Y^{(k)}/\mu = U^{(k)} \Sigma^{(k)} \left(V^{(k)}\right)^T \tag{14}$$

with $U^{(k)} \in R^{m\times n}, V^{(k)} \in R^{n\times n}$.
Step 4: If the termination conditions are satisfied or k reaches the maximum iteration step k_{max}, output solutions $(\tilde{A}, \tilde{E})$; otherwise, increase k and return to step 2.

Figure 1:

Algo 2

Algorithm 2. ADM for nuclear-norm and ℓ_1-norm minimization problem

INPUT: initial values $(A_0, E_0) \in R^{m\times n} \times R^{m\times n}$, $Y_0 \in R^{m\times n}$, $\mu_0 > 0$, $\rho > 1$, $t=0$

WHILE not converged DO

$$A_{t+1} = D_{\mu_t^{-1} w_A}[D - E_t + \mu_t^{-1} Y_t]$$

$$E_{t+1} = S_{\lambda\mu_t^{-1} w_E}[D - A_{t+1} + \mu_t^{-1} Y_t]$$

$$Y_{t+1} = Y_t + \mu_t(D - A_{t+1} - E_{t+1})$$

$$\mu_{t+1} = \rho\mu_t$$

$$t \leftarrow t+1$$

END WHILE

OUTPUT: solution $(\tilde{A}, \tilde{E})$ to Eq. (10)

Figure 2:

Image restoration results Algo 2

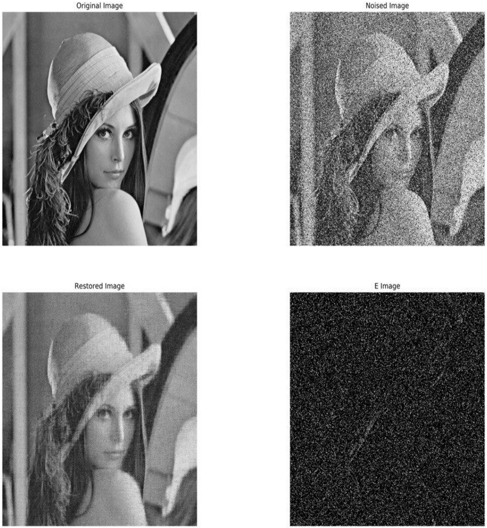

Image restoration results Algo 2

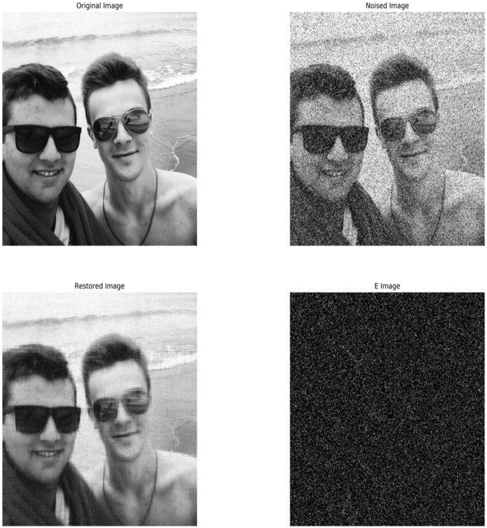

Image restoration results Algo 2

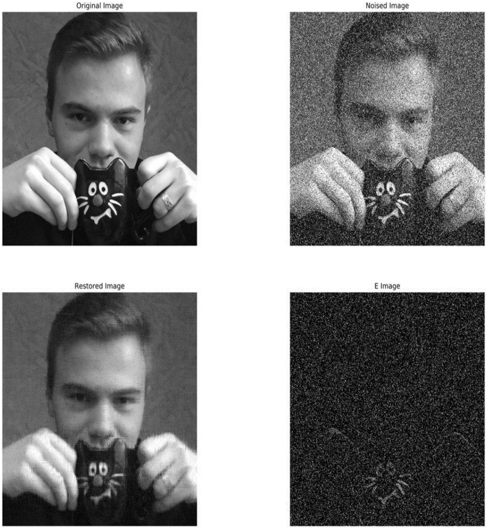

Image restoration results Algo 1

Application

Image denoising tasks (Medical images, Military, etc.).

Numerical data recovery.

Hyperspectral image restoration.

Conclusion

We managed to implement and reconstruct different images with proposed algorithms on a small number of iterations. Considering the results, both of them are working properly.