Robust PCA via Weighted Low Rank

Yao Zhang

1 Original Model

$$\min_{L,S} \|L\|_* + \lambda \|S\|_1, \quad s.t.\ W \circ F = W \circ L + S \tag{1}$$

2 Lagrangian function $\mathcal{L}$ for the Original Model

$$\mathcal{L}(L, S, Y, \mu) = \|L\|_* + \lambda \|S\|_1 + \langle Y, W \circ (F - L) - S \rangle + \frac{\mu}{2} \|W \circ (F - L) - S\|_F^2 \tag{2}$$

3 Derivative Trace Hadamard Production

1. The trace is equivalent to the inner product:
$$X \cdot Y = tr(X^T Y) \tag{3}$$
where $\cdot$ is inner product operator, $X, Y \in \mathbb{R}^{m \times n}$.

2. The Hadamard and inner products commute:
$$X \circ Y \cdot Z = X \cdot Y \circ Z \tag{4}$$
where $\circ$ is Hadamard product operator, $X, Y \in \mathbb{R}^{m \times n}$.

Proof. It is easy to verify that Hadamard product operator is supper to inner product operator.

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \ddots & \vdots & \vdots \\ x_{m-1,1} & x_{m-1,2} & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \circ \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \ddots & \vdots & \vdots \\ y_{m-1,1} & y_{m-1,2} & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} \cdot \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \ddots & \vdots & \vdots \\ z_{m-1,1} & z_{m-1,2} & \ddots & \vdots \\ z_{m1} & z_{m2} & \cdots & z_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} x_{11}y_{11} & x_{12}y_{12} & \cdots & x_{1n}y_{1n} \\ x_{21}y_{21} & x_{22}y_{22} & \cdots & x_{2n}y_{2n} \\ \vdots & \ddots & \vdots & \vdots \\ x_{m-1,1}y_{m-1,1} & x_{m-1,2}y_{m-1,2} & \ddots & \vdots \\ x_{m1}y_{m1} & x_{m2}y_{m2} & \cdots & x_{mn}y_{mn} \end{bmatrix} \cdot \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \ddots & \vdots & \vdots \\ z_{m-1,1} & z_{m-1,2} & \ddots & \vdots \\ z_{m1} & z_{m2} & \cdots & z_{mn} \end{bmatrix}$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{m} x_{ij} y_{ij} z_{ij}$$

$$
\begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \ddots & \vdots & \vdots \\ z_{m-1,1} & z_{m-1,2} & \ddots & \vdots \\ z_{m1} & z_{m2} & \cdots & z_{mn} \end{bmatrix} \cdot \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \ddots & \vdots & \vdots \\ y_{m-1,1} & y_{m-1,2} & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} \circ \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \ddots & \vdots & \vdots \\ z_{m-1,1} & z_{m-1,2} & \ddots & \vdots \\ z_{m1} & z_{m2} & \cdots & z_{mn} \end{bmatrix}
$$

$$
= \sum_{j=1}^{n} \sum_{i=1}^{m} x_{ij} y_{ij} z_{ij}
$$

□

$$
\because r = A \cdot C \circ B = A \circ C \cdot B \tag{5}
$$

where r is a function which dependents on A, B, C and $r \in \mathbb{R}$.

$$
\therefore dr = A \circ C \cdot dB \Rightarrow \frac{\partial r}{\partial B} = A \circ C \tag{6}
$$

4 Derivative $\frac{\partial (A \cdot A \circ B)}{\partial A}$

$$
A \cdot A \circ B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{m-1,1} & a_{m-1,2} & \ddots & \ddots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{m-1,1} & a_{m-1,2} & \ddots & \ddots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \circ \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ b_{m-1,1} & b_{m-1,2} & \ddots & \ddots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}
$$

$$
= \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij}^2 b_{ij}
$$

so, we can get that

$$
\frac{\partial A \cdot A \circ B}{\partial A} = 2A \circ B \tag{7}
$$

5 Computing for the Original Model

1. update L

$$
L^* = \arg\min_{L} \|L\|_* - \langle Y, W \circ L \rangle + \frac{\mu}{2} \|W \circ (F - L) - S\|_F^2 \tag{8}
$$

We denote that

$$\begin{aligned}
q &= -\langle Y, W \circ L\rangle + \frac{\mu}{2}\|W \circ (F-L) - S\|_F^2 \\
&= -Y \cdot W \circ L + \frac{\mu}{2}\left[\langle W \circ (F-L) - S, W \circ (F-L) - S\rangle\right] \\
&= -Y \cdot W \circ L + \frac{\mu}{2}\langle W \circ F, W \circ F - W \circ L - S\rangle - \frac{\mu}{2}\langle W \circ L, W \circ F - W \circ L - S\rangle \\
&\quad - \frac{\mu}{2}\langle S, W \circ F - W \circ L - S\rangle \\
&= -Y \cdot W \circ L + \frac{\mu}{2}(W \circ F \cdot W \circ F) - \frac{\mu}{2}(W \circ F \cdot W \circ L) - \frac{\mu}{2}(W \circ F \cdot S) - \frac{\mu}{2}(W \circ L \cdot W \circ F) \\
&\quad + \frac{\mu}{2}(W \circ L \cdot W \circ L) + \frac{\mu}{2}(W \circ L \cdot S) - \frac{\mu}{2}(S \cdot W \circ F) + \frac{\mu}{2}(S \cdot W \circ L) + \frac{\mu}{2}(S \cdot S)
\end{aligned} \tag{9}$$

then we can get that

$$\frac{\partial q}{\partial L} = \frac{\partial\left(-Y \cdot W \circ L - \frac{\mu}{2}(W \circ F \cdot W \circ L) - \frac{\mu}{2}(W \circ L \cdot W \circ F) + \frac{\mu}{2}(W \circ L \cdot W \circ L) + \mu(W \circ L \cdot S)\right)}{\partial L}$$

And

$$\begin{aligned}
& -Y \cdot W \circ L - \frac{\mu}{2}(W \circ F \cdot W \circ L) - \frac{\mu}{2}(W \circ L \cdot W \circ F) + \frac{\mu}{2}(W \circ L \cdot W \circ L) + \mu(W \circ L \cdot S) \\
=& -Y \circ W \cdot L - \frac{\mu}{2} W \circ F \circ W \cdot L - \frac{\mu}{2} W \circ W \circ F \cdot L + \frac{\mu}{2} L \cdot L \circ W \circ W + \mu W \circ S \cdot L
\end{aligned}$$

so, we can get that

$$\frac{\partial q}{\partial L} = -Y \circ W - \frac{\mu}{2}(W \circ F \circ W) - \frac{\mu}{2}(W \circ W \circ F) + \mu(L \circ W \circ W) + \mu W \circ S \tag{10}$$

According to LADMM, minimizing Equation 8 can be replaced by solving the following:

$$\begin{aligned}
L^{k+1} &= \arg\min_L \|L\|_* + \langle \nabla_L q(L_k), L - L_k\rangle + \frac{\eta}{2}\|L - L_k\|_F^2 \\
&= \arg\min_L \|L\|_* + \frac{\eta}{2}\left\| L - L_k + \frac{\nabla_L q(L_k)}{\eta}\right\| \\
&= \arg\min_L \frac{1}{\eta}\|L\|_* + \frac{1}{2}\left\| L - \left(L_k - \frac{\nabla_L q(L_k)}{\eta}\right)\right\|
\end{aligned} \tag{11}$$

2. update S

$$\begin{aligned}
S^{k+1} &= \arg\min_S \lambda\|S\|_1 + \langle Y, W \circ (F-L) - S\rangle + \frac{\mu}{2}\|W \circ (F-L) - S\|_F^2 \\
&= \arg\min_S \lambda\|S\|_1 + \frac{\mu}{2}\left\| W \circ (F-L) - S + \frac{Y}{\mu}\right\|_F^2 \\
&= \arg\min_S \frac{\lambda}{\mu}\|S\|_1 + \frac{1}{2}\left\|\left(W \circ (F-L) + \frac{Y}{\mu}\right) - S\right\|_F^2
\end{aligned} \tag{12}$$

3. update Y

$$Y^{k+1} = Y^k + \mu(W \circ (F-L) - S) \tag{13}$$

6 Our Model

$$\min_{L,S} \|L\|_{S_{2/3}}^{2/3} + \lambda \|S\|_{l\ 2/3}^{2/3}, \quad s.t.\ W \circ F = W \circ L + S \tag{14}$$

7 Lagrangian function $\mathcal{L}$ for Our Model

$$\mathcal{L}(L, S, Y, \mu) = \|L\|_{S_{2/3}}^{2/3} + \lambda \|S\|_{l_{2/3}}^{2/3} + \langle Y, W \circ F - W \circ L - S \rangle + \frac{\mu}{2} \|W \circ F - W \circ L - S\|_F^2 \tag{15}$$

Jan 19, 2019