Supplementary Material: Efficient Stochastic Optimization for Low-Rank Distance Metric Learning

Jie Zhang and **Lijun Zhang**
National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China
zhangj@lamda.nju.edu.cn, zhanglj@lamda.nju.edu.cn

Proof of Theorem 2

Note that $_{\mathbb{S}_+\cap\{W|\|W\|_2\leq\tau\}}[X]$ is the optimal solution to the following problem

$$\begin{aligned} \min_{A\in\mathbb{R}^{d\times d}} \quad & \frac{1}{2}\|A - X\|_F^2 \\ \text{s.t.} \quad & A \in \mathbb{S}_+, \quad I - A \in \mathbb{S}_+ . \end{aligned} \tag{6}$$

The Lagrangian function associated with (6) is

$$L(A, Y_1, Y_2) = \frac{\|A - X\|_F^2}{2} - \text{tr}(AY_1) + \text{tr}((A - I)Y_2)$$

where $Y_1 \in \mathbb{R}^{d\times d}$ and $Y_2 \in \mathbb{R}^{d\times d}$ are dual variables for constraints $A \in \mathbb{S}_+$ and $I - A \in \mathbb{S}_+$. Let A^*,Y_1^*,Y_2^* be the optimal primal and dual solutions. The KKT conditions are

$$\begin{aligned} A^* &\in \mathbb{S}_+; \\ I - A^* &\in \mathbb{S}_+; \\ A^* &= X + Y_1^* - Y_2^*; \end{aligned}$$

$\text{tr}(A^*Y^* X + Y^*$ 75)4 Td_[(2)]TJ/F38 9.9626 Tf 9.409 0 Td [(S) where $Y_1 \in \mathbb{R}^{d\times d}$ and $\in \mathbb{R}$ are dual variables for constraints $A \in \mathbb{S}_+$ and $\|A\|_F$. Let $A^*, Y_1^*, {}^*$ be the optimal primal and dual solutions. The KKT conditions are

$$
\begin{aligned}
A^* &\in \mathbb{S}_+; \\
\|A^*\|_F & \\
A^* &= \frac{1}{1+{}^*}(X + Y_1^*); \\
\mathrm{tr}(A^* Y_1^*) &= 0; \\
{}^*(\|A^*\|_F^2 \quad {}^2) &= 0; \\
Y_1^* \in \mathbb{S}_+; \quad {}^* \quad & 0.
\end{aligned}
$$

We complete the proof by noticing that

$$
\begin{aligned}
Y_1^* &= \sum_{i:\lambda_i<0} {}_i \mathbf{u}_i \mathbf{u}_i^\top; \\
\hat{A} = X + Y_1^* &= \sum_{i:\lambda_i>0} {}_i \mathbf{u}_i \mathbf{u}_i^\top; \\
{}^* &= \begin{cases} 0 & \|\hat{A}\|_F \\ \frac{\|\hat{A}\|_F}{} \quad 1 & \|\hat{A}\|_F > \end{cases}; \\
A^* &= \begin{cases} \hat{A} & \|\hat{A}\|_F \\ \frac{}{\|\hat{A}\|_F}\hat{A} & \|\hat{A}\|_F > \end{cases};
\end{aligned}
$$

satisfy these KKT conditions.