"Two-stage Sparse Representation Clustering for Dynamic Data Streams" —Supplementary Document

Jie Chen, *Member, IEEE,* Zhu Wang, Shengxiang Yang, *Senior Member, IEEE,* Hua Mao

I. Proof of Theorem 1

In this section, we prove Theorem 1 in the paper regarding the optimization program

$$\begin{aligned}&\min_{\mathbf{Z},\mathbf{D}} \|\mathbf{Z}\|_0 + \frac{\alpha}{2}\|\mathbf{X}-\mathbf{D}\mathbf{Z}\|_F^2 + \frac{\beta}{2}\|\mathbf{D}\|_F^2 \\ &s.t. \quad diag(\mathbf{Z}) = \mathbf{0}\end{aligned} \tag{1}$$

Given the fixed $\mathbf{J}_{k+1}$, $\mathbf{Z}_{k+1}$ is updated by the following scheme:

$$\begin{aligned}&\mathbf{Z}_{k+1} = \min_{\mathbf{Z}_{k+1}} \frac{1}{\mu_k}\|\mathbf{Z}_{k+1}\|_0 + \frac{1}{2}\left\|\mathbf{Z}_{k+1} - \left(\mathbf{J}_{k+1} + \frac{\mathbf{Y}_k}{\mu_k}\right)\right\|_F^2, \\ &\mathbf{Z}_{k+1} \leftarrow \mathbf{Z}_{k+1} - d(\mathbf{Z}_{k+1}).\end{aligned} \tag{2}$$

The hard thresholding operator $\mathcal{H}_{\sqrt{\lambda}}(x)$ is defined as follows [1]:

$$\mathcal{H}_{\sqrt{\lambda}}(x) = \begin{cases} 0, & if \quad |x| \le \sqrt{\lambda} \\ x, & if \quad |x| > \sqrt{\lambda} \end{cases} . \tag{3}$$

The closed-form solution of the first part of (2) is obtained using the operator $\mathcal{H}$:

$$\mathbf{Z}_{k+1} = \mathcal{H}_{\sqrt{\frac{1}{\mu_k}}}\left(\mathbf{J}_{k+1} + \frac{\mathbf{Y}_k}{\mu_k}\right). \tag{4}$$

Theorem 1 *The convergence condition $\|\mathbf{Z}_{k+1} - \mathbf{J}_{k+1}\|_{\max} < \varepsilon$ will eventually be satisfied as k increases if ρ and μ satisfy the following conditions:*

$$\rho > 2 \quad and \quad \mu > 0$$

where k represents the number of iterations and ε is a small positive number, e.g., $\varepsilon = 10^{-4}$.

Proof *Given the optimal $\mathbf{Z}_k$, $\mathbf{J}_k$ and $\mathbf{D}_k$ at the k-th iteration, where $k > 1$, we continue to optimize $\mathbf{Z}_{k+1}$ and $\mathbf{J}_{k+1}$ by fixing $\mathbf{D}_k$ and $\mathbf{Y}_k$ at the $(k+1)$-th iteration. According to* (4)*, we know that $\mathbf{Z}_{k+1}$ has a closed-form solution. Thus, we have the following equality:*

$$\|\mathbf{Z}_{k+1} - \mathbf{J}_{k+1}\|_{\max} = \left\|\mathcal{H}_{\sqrt{\frac{1}{\mu_k}}}\left(\mathbf{J}_{k+1} + \frac{\mathbf{Y}_k}{\mu_k}\right) - \mathbf{J}_{k+1}\right\|_{\max}. \tag{5}$$

Suppose $\rho > 2 \quad and \quad \mu > 0$, and we get $\mu_k \to \infty$ when $k \to \infty$ according to $\mu_k = \rho\mu_{k-1}$. This indicates that we will obtain

$$\mathcal{H}_{\sqrt{\frac{1}{\mu_k}}}\left(\mathbf{J}_{k+1} + \frac{\mathbf{Y}_k}{\mu_k}\right) = \mathbf{J}_{k+1} + \frac{\mathbf{Y}_k}{\mu_k}$$

as k steadily increases. According to (5)*, we get*

$$\begin{aligned}\|\mathbf{Z}_{k+1} - \mathbf{J}_{k+1}\|_{\max} &= \left\|\frac{\mathbf{Y}_k}{\mu_k}\right\|_{\max} \\ &= \left\|\frac{\mathbf{Y}_{k-1} + \mu_{k-1}(\mathbf{Z}_k - \mathbf{J}_k)}{\mu_k}\right\|_{\max} \\ &\le \left\|\frac{\mathbf{Y}_{k-1}}{\mu_k}\right\|_{\max} + \left\|\frac{\mu_{k-1}(\mathbf{Z}_k - \mathbf{J}_k)}{\mu_k}\right\|_{\max} \\ &= \left\|\frac{\mathbf{Y}_{k-1}}{\rho\mu_{k-1}}\right\|_{\max} + \left\|\frac{\mathbf{Z}_k - \mathbf{J}_k}{\rho}\right\|_{\max}.\end{aligned}$$

Thus,

$$\|\mathbf{Z}_k - \mathbf{J}_k\|_{\max} \ge \frac{\rho}{2}\|\mathbf{Z}_{k+1} - \mathbf{J}_{k+1}\|_{\max}.$$

Then,

$$\begin{aligned}&\|\mathbf{Z}_k - \mathbf{J}_k\|_{\max} - \|\mathbf{Z}_{k+1} - \mathbf{J}_{k+1}\|_{\max} \\ &\ge \left(\frac{\rho}{2} - 1\right)\|\mathbf{Z}_{k+1} - \mathbf{J}_{k+1}\|_{\max}\end{aligned}$$

According to $\|\mathbf{Z}_{k+1} - \mathbf{J}_{k+1}\|_{\max} > 0$ and $\rho > 2$, we get

$$\|\mathbf{Z}_k - \mathbf{J}_k\|_{\max} - \|\mathbf{Z}_{k+1} - \mathbf{J}_{k+1}\|_{\max} > 0$$

when $\mathbf{Z}_k - \mathbf{J}_k \neq 0$. This means there exists a certain k with two conditions, i.e., $\rho > 2$ and $\mu_1 > 0$, such that the following inequality holds:

$$\|\mathbf{Z}_{k+1} - \mathbf{J}_{k+1}\|_{\max} \le \varepsilon,$$

where $\mu = \mu_1$. Hence, convergence will eventually be achieved as k gradually increases if $\rho > 2$ and $\mu > 0$. □

II. Proof of Theorem 2

In this section, we prove Theorem 2 in the paper.

Theorem 2 *Suppose that convergence is achieved after the k-th iteration in Algorithm 1. The sparsity ratio (SR) of a matrix Z is defined as $SR(\mathbf{Z}_k) = \frac{\|\mathbf{Z}_k\|_0}{num(\mathbf{Z}_k)}$, where num($\mathbf{Z}_k$) is the number of elements in $\mathbf{Z}_k$. The SR of $\mathbf{Z}$ will always remain stable, i.e., $|SR(\mathbf{Z}_{k+1}) - SR(\mathbf{Z}_k)| < \varepsilon$, after k iterative computations, if*

$$\mu_{k-1} > 1 \quad and \quad \rho > 1$$

J. Chen is with the College of Computer Science, Sichuan University, Chengdu 610065, China (E-mail: chenjie2010@scu.edu.cn).
Z. Wang is with the Law School, Sichuan University, Chengdu 610065, China (E-mail: wangzhu@scu.edu.cn).
S. Yang is with the School of Computer Science and Informatics, De Montfort University, Leicester LE1 9BH, U. K. (e-mail: syang@dmu.ac.uk).
H. Mao is with the Department of Computer and Information Sciences, Northumbria University, Newcastle, NE1 8ST, U. K. (e-mail: hua.mao@northumbria.ac.uk).

where $\|\mathbf{Z}_k\|_0$ *counts the number of nonzero entries in the matrix* $\mathbf{Z}_k$, $\varepsilon = 1e^{-6}$ *and* $k > 1$.

Proof *Let* $\mathbf{Z}_{min}^{k+1}$ *be the minimum absolute value among all elements except zeros in the matrix* $\mathbf{Z}_{k+1}$. *According to* (3), *we have:*

$$\mathbf{Z}_{min}^{k+1} > \sqrt{\frac{1}{\mu_k}},$$

where $\mathbf{Z}_{k+1} = \mathcal{H}_{\sqrt{\frac{1}{\mu_k}}}\left(\mathbf{J}_{k+1} + \frac{\mathbf{Y}_k}{\mu_k}\right)$. *Suppose Algorithm 1 converges after the* k*-th iteration, and so*

$$\mathbf{Z}_{k+1} \approx \mathbf{J}_{k+1}.$$

Because $\mathbf{Z}_{min}^{k} > \sqrt{\frac{1}{\mu_{k-1}}}$ *and* $\mu_k > \mu_{k-1} > 1$, *the number of nonzero elements in* $\mathbf{Z}_{k+1}$ *remains unchanged after the* k*-th iteration. This indicates that the SR of* $\mathbf{Z}$ *remains stable, i.e.,* $|SR(\mathbf{Z}_{k+1}) - SR(\mathbf{Z}_k)| < \varepsilon$ *at least before the* k*-th iteration.* □

REFERENCES

[1] T. Blumensath and M. E. Davies, “Iterative thresholding for sparse approximations,” *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 629–654, Sept. 2008.