

## Appendix: Proof of Theorem 1

### Theorem Restatement

To solve the standard nuclear norm minus Frobenius norm minimization (TNFM) problem in closed-form, we propose a proximal operator, defined as follows,

$$\mathbf{P}_\lambda(\mathbf{Z}) = \arg \min_{\mathbf{B}} \frac{1}{2} \|\mathbf{B} - \mathbf{Z}\|_F^2 + \lambda \|\mathbf{B}\|_{t,*-F}, \quad (1)$$

where  $\mathbf{Z} \in \mathbb{R}^{m \times n}$  is a given matrix. Assume that  $\lambda > 0$  and  $\mathbf{Z} \in \mathbb{R}^{m \times n}$  admits SVD as  $\mathbf{U}_Z \text{Diag}(\boldsymbol{\sigma}(\mathbf{Z})) \mathbf{V}_Z^\top$ , without loss of generality, let  $m \geq n$ . The closed-form solution of Equation (1) is given by

$$\mathbf{B}^* = \mathbf{P}_{\lambda,t}(\mathbf{Z}) = \mathbf{U}_Z \text{Diag}(\boldsymbol{\varrho}^*) \mathbf{V}_Z^\top, \quad (2)$$

where

$$\varrho_i^* = \begin{cases} \sigma_i(\mathbf{Z}), & \text{if } 0 \leq i < t+1, \\ \left(1 + \frac{\lambda}{\|\mathcal{S}_\lambda(\mathbf{r})\|_2}\right) \mathcal{S}_\lambda(\sigma_i(\mathbf{Z})), & \text{if } t+1 \leq i \leq n, \end{cases} \quad (3)$$

with  $\mathbf{r} = [0, \dots, 0, \sigma_{t+1}(\mathbf{Z}), \dots, \sigma_n(\mathbf{Z})]^\top$  and  $\mathcal{S}_\lambda(\mathbf{r})_i = \max(\mathbf{r}_i - \lambda, 0)$  be the soft shrinkage.

### Proof

Consider  $\mathbf{B} = \mathbf{U}_B \boldsymbol{\Sigma}_B \mathbf{V}_B^\top$  to be its SVD. The first term of problem (1) can be rewritten as

$$\frac{1}{2} \|\mathbf{B} - \mathbf{Z}\|_F^2 = \frac{1}{2} (\|\mathbf{B}\|_F^2 - 2\langle \mathbf{B}, \mathbf{Z} \rangle + \|\mathbf{Z}\|_F^2). \quad (4)$$

According to Von Neumann's trace inequality, the matrix inner product in (4) can be derived as follows:

$$\begin{aligned} \langle \mathbf{B}, \mathbf{Z} \rangle &= \text{tr}(\mathbf{B}^\top \mathbf{Z}) \\ &= \text{tr}(\mathbf{V}_B \boldsymbol{\Sigma}_B \mathbf{U}_B^\top \mathbf{Z}) \\ &= \text{tr}(\boldsymbol{\Sigma}_B \mathbf{U}_B^\top \mathbf{Z} \mathbf{V}_B) \\ &\leq \sum_{i=1}^n \sigma_i(\mathbf{B}) \cdot \sigma_i(\mathbf{Z}) \cdot \sigma_i(\mathbf{U}_B^\top \mathbf{V}_B) \\ &= \sum_{i=1}^n \sigma_i(\mathbf{B}) \cdot \sigma_i(\mathbf{Z}). \end{aligned} \quad (5)$$

The equality of (5) occurs if and only if

$$\mathbf{U}_B = \mathbf{U}_Z, \mathbf{V}_B = \mathbf{V}_Z, \quad (7)$$

since

$$\begin{aligned} &\text{tr}(\boldsymbol{\Sigma}_B \mathbf{U}_B^\top \mathbf{Z} \mathbf{V}_B) \\ &= \text{tr}(\boldsymbol{\Sigma}_B \mathbf{U}_B^\top (\mathbf{U}_Z \boldsymbol{\Sigma}_Z \mathbf{V}_Z^\top) \mathbf{V}_B) \\ &= \text{tr}(\boldsymbol{\Sigma}_B (\mathbf{U}_Z^\top \mathbf{U}_Z) \boldsymbol{\Sigma}_Z (\mathbf{V}_Z^\top \mathbf{V}_Z)) \\ &= \text{tr}(\boldsymbol{\Sigma}_B \boldsymbol{\Sigma}_Z) = \sum_{i=1}^n \sigma_i(\mathbf{B}) \cdot \sigma_i(\mathbf{Z}). \end{aligned} \quad (8)$$

Therefore, problem (1) can be rewritten as follows:

$$\begin{aligned} &\arg \min_{\mathbf{B}} \frac{1}{2} \|\mathbf{B} - \mathbf{Z}\|_F^2 + \lambda \|\mathbf{B}\|_{t,*-F} \\ &= \arg \min_{\mathbf{B}} \frac{1}{2} \|\mathbf{B}\|_F^2 - \sum_{i=1}^n \sigma_i(\mathbf{B}) \sigma_i(\mathbf{Z}) + \frac{1}{2} \|\mathbf{Z}\|_F^2 \\ &\quad + \lambda \left( \sum_{i=t+1}^n \sigma_i(\mathbf{B}) - \left( \sum_{i=t+1}^n \sigma_i(\mathbf{B})^2 \right)^{\frac{1}{2}} \right) \\ &= \arg \min_{\mathbf{B}} \sum_{i=1}^t \left( \frac{1}{2} \sigma_i(\mathbf{B})^2 - \sigma_i(\mathbf{B}) \sigma_i(\mathbf{Z}) \right) + \sum_{i=t+1}^n \left( \frac{1}{2} \sigma_i(\mathbf{B})^2 \right. \\ &\quad \left. - \sigma_i(\mathbf{B}) \sigma_i(\mathbf{Z}) + \lambda \sigma_i(\mathbf{B}) \right) - \lambda \left( \sum_{i=t+1}^n \sigma_i(\mathbf{B})^2 \right)^{\frac{1}{2}}. \end{aligned} \quad (9)$$

Up to now, original problem (1) has been equivalently transformed into the combination of independent quadratic equations for each  $\sigma_i(\mathbf{B})$ . Let  $F(\boldsymbol{\sigma}(\mathbf{B}))$  denote the objective function of (9). The minimum of  $F$ , denoted as  $\sigma_i(\mathbf{B})^*$ , is given by the first-order optimality condition

$$\frac{\partial F}{\partial \sigma_i(\mathbf{B})} = 0. \quad (10)$$

When  $0 \leq i < t+1$ , it is trivial to obtain

$$\sigma_i(\mathbf{B})^* = \sigma_i(\mathbf{Z}). \quad (11)$$

When  $t+1 \leq i < n+1$ , equation (10) is expressed as

$$\left(1 - \frac{\lambda}{\|\boldsymbol{\sigma}(\mathbf{B})\|_2}\right) \boldsymbol{\sigma}(\mathbf{B}) = \boldsymbol{\sigma}(\mathbf{Z}) - \lambda. \quad (12)$$

Let  $\mathbf{r} = [0, \dots, 0, \sigma_{t+1}(\mathbf{Z}), \dots, \sigma_n(\mathbf{Z})]^\top \in \mathbb{R}^n$  and  $\mathcal{S}_\lambda(\mathbf{r}) = \max(\mathbf{r} - \lambda, 0)$ , the solution of (12) is

$$\sigma_i(\mathbf{B})^* = \left(1 + \frac{\lambda}{\|\mathcal{S}_\lambda(\mathbf{r})\|_2}\right) \cdot \mathcal{S}_\lambda(\sigma_i(\mathbf{Z})). \quad (13)$$

Combining (11) and (12), the minimum of  $F$  is obtained at

$$\sigma_i(\mathbf{B})^* = \begin{cases} \sigma_i(\mathbf{Z}), & \text{if } 0 \leq i < t+1, \\ \left(1 + \frac{\lambda}{\|\mathcal{S}_\lambda(\mathbf{r})\|_2}\right) \mathcal{S}_\lambda(\sigma_i(\mathbf{Z})), & \text{if } t+1 \leq i < n+1. \end{cases} \quad (14)$$

Therefore, the optimal solution of problem (1) is

$$\mathbf{B}^* = \mathbf{U}_Z \text{Diag}(\boldsymbol{\sigma}(\mathbf{B})^*) \mathbf{V}_Z^\top. \quad (15)$$

□