# Supplementary Material:
# Fixed-Rank Representation for Unsupervised Visual Learning

**Risheng Liu**[†‡], **Zhouchen Lin**[♭♯], **Fernando De la Torre**[‡] and **Zhixun Su**[†]
[†]School of Mathematical Sciences, Dalian University of Technology
[♯]Key Lab. of Machine Perception (MOE), Peking University
[‡]Robotics Institute, Carnegie Mellon Univesity
[♭]Microsoft Research Asia

## 1. Proof of Theorem 1

$$\min_{\mathbf{Z},\mathbf{L},\mathbf{R}} \|\mathbf{Z}-\mathbf{L}\mathbf{R}\|_F^2, \ s.t. \ \mathbf{X}=\mathbf{X}\mathbf{Z}. \tag{1}$$

**Theorem 1** *Let $[\mathbf{V}_X]_{1:m} = [[\mathbf{V}_X]_1, [\mathbf{V}_X]_2, \cdots, [\mathbf{V}_X]_m]$. Then for any fixed $m \leq r_X$,*

$$(\mathbf{Z}^*, \mathbf{L}^*, \mathbf{R}^*) := (\mathbf{V}_X\mathbf{V}_X^T, [\mathbf{V}_X]_{1:m}, [\mathbf{V}_X]_{1:m}^T)$$

*is a globally optimal solution to (1) and the minimum objective function value is $(r_X - m)$.*

The proof of this theorem is based on the following lemma.

**Lemma 1** *(Courant-Fischer Minimax Theorem [3]) For any symmetric matrix $\mathbf{A}\in\mathbb{R}^{n\times n}$, we have that*

$$\lambda_i(\mathbf{A}) = \max_{\dim(\mathcal{S})=i} \min_{\mathbf{0}\neq\mathbf{y}\in\mathcal{S}} \mathbf{y}^T\mathbf{A}\mathbf{y}/\mathbf{y}^T\mathbf{y}, \text{ for } i = 1,2,...,n,$$

*where $\mathcal{S}\subset\mathbb{R}^n$ is some subspace and $\lambda_i(\mathbf{A})$ is the $i$-th largest eigenvalue of $\mathbf{A}$.*

**Proof** First, by the well known Eckart-Young theorem [2], given $\mathbf{Z}$, we have

$$\min_{\mathbf{L},\mathbf{R}} \|\mathbf{Z}-\mathbf{L}\mathbf{R}\|_F^2 = \sum_{i=m+1}^{d} \sigma_i^2(\mathbf{Z}), \tag{2}$$

where $\sigma_i(\mathbf{Z})$ is the $i$-th largest singular value of $\mathbf{Z}$. Now we prove that

$$\text{if } \mathbf{X}=\mathbf{X}\mathbf{Z} \text{ then } \sigma_{r_X}(\mathbf{Z}) \geq 1. \tag{3}$$

By $\mathbf{X}=\mathbf{X}\mathbf{Z}$, we have that $\text{rank}(\mathbf{Z}) \geq r_X$. Then (2) and (3) imply that the minimum objective function value is no less than $r_X - m$. Indeed, by the compact SVD of $\mathbf{X}$ and $\mathbf{X}=\mathbf{X}\mathbf{Z}$, we have

$$\mathbf{V}_X^T = \mathbf{V}_X^T\mathbf{Z}, \tag{4}$$

By Lemma 1, $\sigma_i(\mathbf{Z}) = \max_{\dim(\mathcal{S})=i} \min_{\mathbf{0}\neq\mathbf{y}\in\mathcal{S}} \|\mathbf{Z}^T\mathbf{y}\|_2/\|\mathbf{y}\|_2$, where $\|\cdot\|_2$ is the $l_2$ norm of a vector. So by choosing $\mathcal{S}=\mathcal{R}(\mathbf{V}_X)$ and utilizing (4),

$$\begin{aligned} \sigma_{r_X}(\mathbf{Z}) &\geq \min_{\mathbf{0}\neq\mathbf{y}\in\mathcal{R}(\mathbf{V}_X)} \|\mathbf{Z}^T\mathbf{y}\|_2/\|\mathbf{y}\|_2 \\ &= \min_{\mathbf{b}\neq\mathbf{0}} \|\mathbf{Z}^T\mathbf{V}_X\mathbf{b}\|_2/\|\mathbf{V}_X\mathbf{b}\|_2 \\ &= \min_{\mathbf{b}\neq\mathbf{0}} \|\mathbf{V}_X\mathbf{b}\|_2/\|\mathbf{V}_X\mathbf{b}\|_2 = 1. \end{aligned} \tag{5}$$

Next, when $\mathbf{Z}=\mathbf{V}_X\mathbf{V}_X^T$, it can be easily checked that the objective function value is $(r_X - m)$. Again, by Eckart-Young theorem, $\mathbf{L}\mathbf{R} = [\mathbf{V}_X]_{1:m}[\mathbf{V}_X]_{1:m}^T$. Thus we have $(\mathbf{V}_X\mathbf{V}_X^T, [\mathbf{V}_X]_{1:m}, [\mathbf{V}_X]_{1:m}^T)$ is a globally optimal solution to (1), thereby completing the proof of the theorem.

## 2. Proof of Corollary 2

**Corollary 2** *Under the assumption that subspaces are independent and data* $\mathbf{X}$ *is clean, there exits a globally optimal solution* $(\mathbf{Z}^*, \mathbf{L}^*, \mathbf{R}^*)$ *to problem (1) with the following structure:*

$$\mathbf{Z}^* = diag(\mathbf{Z}_1, \mathbf{Z}_2, ..., \mathbf{Z}_k), \tag{6}$$

*where* $\mathbf{Z}_i$ *is an* $n_i \times n_i$ *matrix with* $rank(\mathbf{Z}_i) = d_{C_i}$ *and*

$$\mathbf{L}^*\mathbf{R}^* \in \mathcal{R}(\mathbf{Z}^*) = \mathcal{R}(\mathbf{X}^T). \tag{7}$$

The proof of this corollary is based on the following lemma.

**Lemma 2** *[1] Let* $\mathbf{X} = \mathbf{U}_X \Sigma_X \mathbf{V}_X^T$ *be the compact SVD. Under the same assumption in Corollary 2,* $\mathbf{V}_X \mathbf{V}_X^T$ *is a block diagonal matrix that has exactly* $k$ *blocks. Moreover, the* $i$*-th block on its diagonal is an* $n_i \times n_i$ *matrix with rank* $d_{C_i}$.

**Proof** By the proof of Theorem 1, we have that $\mathbf{Z}^* = \mathbf{V}_X \mathbf{V}_X^T$ is a globally optimal solution to (1) and any globally optimal $\mathbf{L}^*$ and $\mathbf{R}^*$ are in the range space $\mathcal{R}(\mathbf{Z}^*)$. So we have that $\mathbf{L}^*\mathbf{R}^* \in \mathcal{R}(\mathbf{Z}^*) = \mathcal{R}(\mathbf{X}^T)$. By Lemma 2, we achieve the block diagonal structure (6) for $\mathbf{Z}^*$, which concludes the proof.

## 3. Proof of Corollary 3

**Corollary 3** *Assume that the columns of* $\mathbf{Z}^*$ *are normalized (i.e.* $\mathbf{1}_n^T \mathbf{Z}^* = \mathbf{1}_n^T$*) and fix* $m = k$*, then there exists globally optimal* $\mathbf{L}^*$ *and* $\mathbf{R}^*$ *to problem (1) such that*

$$\mathbf{L}^*\mathbf{R}^* = diag(n_1 \mathbf{1}_{n_1} \mathbf{1}_{n_1}^T, n_2 \mathbf{1}_{n_2} \mathbf{1}_{n_2}^T, ..., n_k \mathbf{1}_{n_k} \mathbf{1}_{n_k}^T). \tag{8}$$

**Proof** By Corollary 2 and the normalization assumption, $\mathbf{Z}^* = \text{diag}(\mathbf{Z}_1^*, \mathbf{Z}_2^*, ..., \mathbf{Z}_k^*)$, where $\mathbf{Z}_i^*$ is an $n_i \times n_i$ for subspace $\mathcal{C}_i$ and $\mathbf{1}_{n_i}$ is an eigenvector of $\mathbf{Z}_i^*$ with eigenvalue 1. Thus there exists a basis $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_k]$, each vector of which with the form $\mathbf{h}_i = [\mathbf{0}, \mathbf{1}_{n_i}^T, \mathbf{0}]^T$ is eigenvector of $\mathbf{Z}$ with eigenvalue 1. By the Eckart-Young theorem (similar to the proof of Theorem 1), we have that $\mathbf{L}^* = \mathbf{H}$ and $\mathbf{R}^* = \mathbf{H}^T$ are globally optimal solutions to (1), which directly leads (8).

## 4. Proof of Corollary 4

$$\min_{\mathbf{Z},\mathbf{L},\mathbf{R}} \|\mathbf{Z} - \mathbf{L}\mathbf{R}\|_F^2, \ s.t. \ \mathbf{X} = \mathbf{Z}\mathbf{X}, \tag{9}$$

Let $[\mathbf{U}_X]_{1:m} = [[\mathbf{U}_X]_1, [\mathbf{U}_X]_2, \cdots, [\mathbf{U}_X]_m]$. Then we have the following corollary

**Corollary 4** *For any fixed* $m \leq r_X$,

$$(\mathbf{Z}^*, \mathbf{L}^*, \mathbf{R}^*) := (\mathbf{U}_X \mathbf{U}_X^T, [\mathbf{U}_X]_{1:m}, [\mathbf{U}_X]_{1:m}^T)$$

*is a globally optimal solution to (9) and the minimum objective function value is* $(r_X - m)$.

**Proof** The proof of Theorem 1 directly leads to the above corollary.

## References


[1] J. Costeira and T. Kanade. A multibody factorization method for independently moving objcets. *IJCV*, 29(3):159–179, 1998. 2

[2] C. Eckart and G. Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1936. 1

[3] G. Golub and C. Van Loan. *Matrix Computations*. Johns Hopkins University Press, 3rd edition, 1996. 1