Low-Rank-Based Feature Learning with Active Learning for Open-Category Classification

2018.5.24

Related Introduction

Theorem

complete solutions

$$\min_{Z,L} \|Z\|_* + \|L\|_* \quad s.t. X = XZ + LX \longrightarrow Z^* = V_X(I-S)V_X^T \text{ and } L^* = U_X S U_X^T$$

$SVD(X) = U_X \Sigma_X U_X^T$ $\quad S$ is any block-diagonal matrix that satisfies two constraints:

1) its blocks are compatible with Σ_X, i.e., if $(\Sigma_X)_{ii} \neq (\Sigma_X)_{jj}$, then $S_{ij} = 0$

2) both S and $I - S$ are positive semidefinite.

Active Learning

$$\min_{A,B} \sum_{i=1}^{n} (\|x_i - Ba_i\|_2^2 + \gamma \|a_i\|_2^2) \quad s.t. \quad A = [a_1, \ldots, a_n] \in R^{m \times n}, B \subset X, |B| = m$$

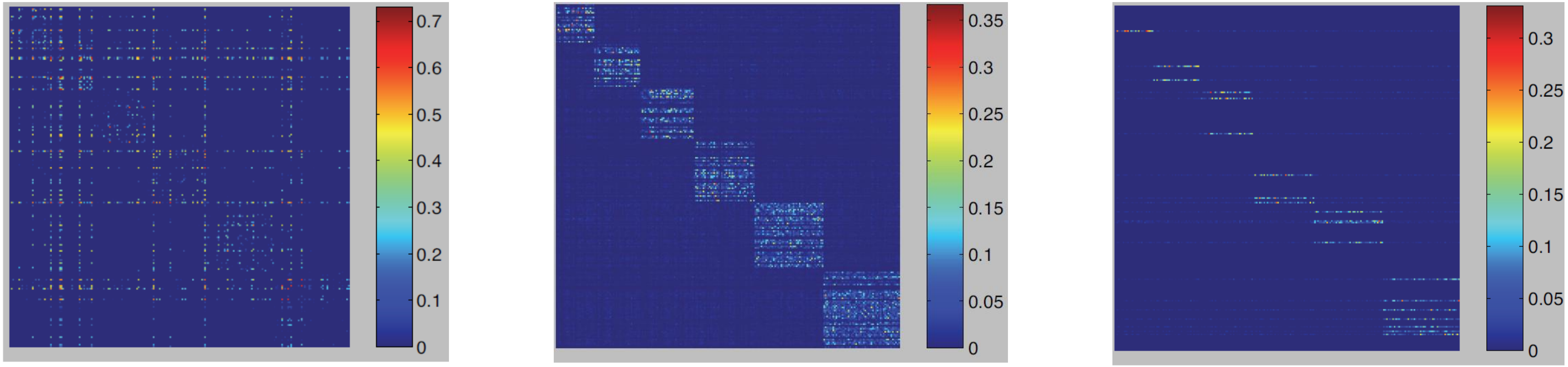
$$\min_{A} \sum_{i=1}^{n} (\|x_i - Xa_i\|_2^2 + \gamma \|a_i\|_2^2) \quad s.t. \quad A = [a_1, \ldots, a_n] \in R^{n \times n}, \quad \|A\|_{2,0} = m$$

$$\min_{A} \sum_{i=1}^{n} (\|x_i - Xa_i\|_2^2) + \gamma \|A\|_{2,1}$$

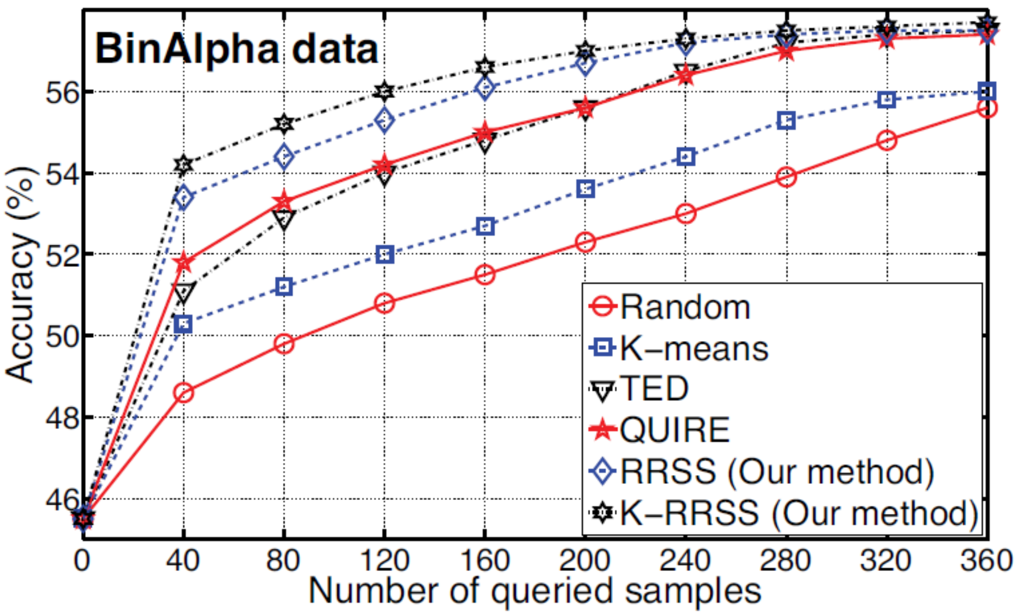
sensitive to data outliers

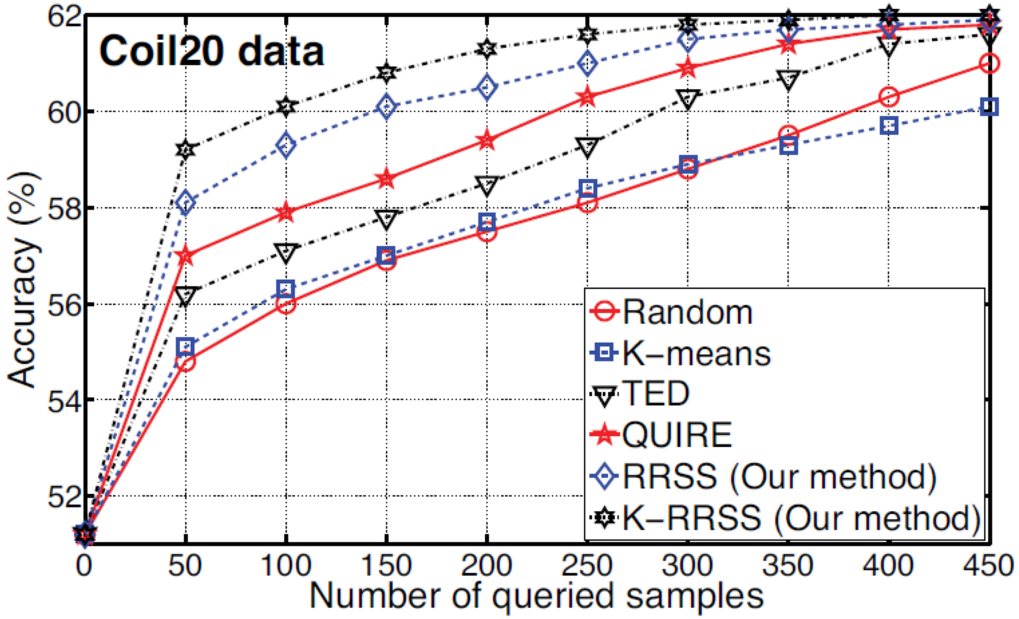
$$J = \min_{A} \|(X - XA)^T\|_{2,1} + \gamma \|A\|_{2,1}$$

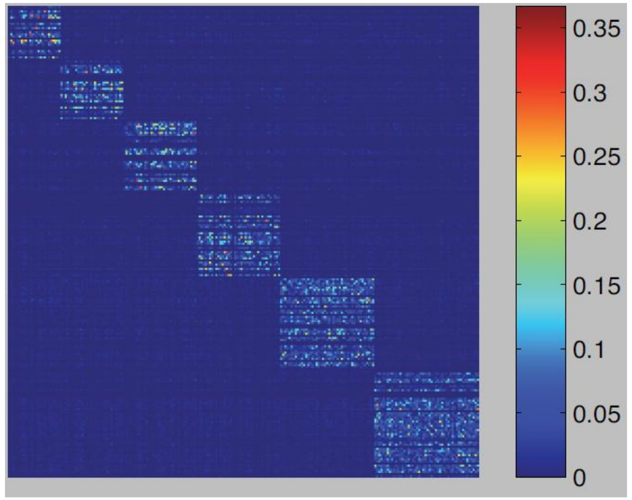
Sort the rows of A by **the row-sum values of the absolute** A in the **decreasing order**. Therefore, the active learning task can be performed **by selecting the m samples** corresponding to the top m rows of A.



(a) 10-NN similarity graph (b) Learned matrix A (c) The top 20 rows of A

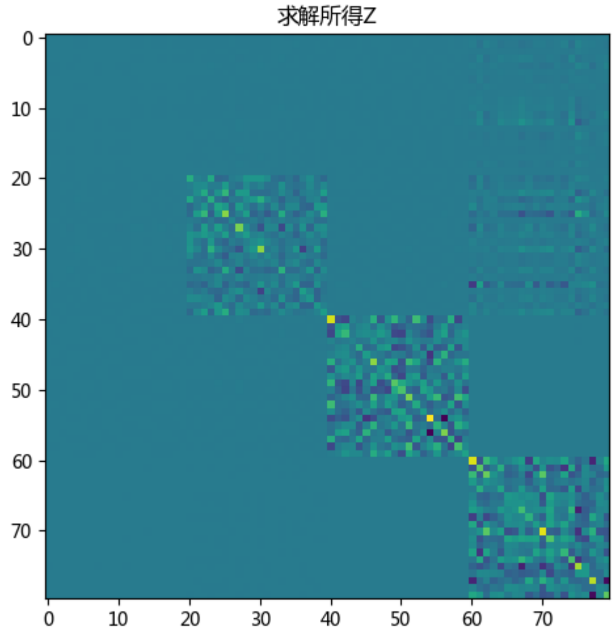






(b) Learned matrix A

$$\min_{Z}\|Z\|_* \quad s.t.\, X = XZ$$



如何利用“关联度矩阵”来区分 new class data？

计算权重质量分布的熵

$$x_i - Xa_i \quad x_i : d\times 1 \;\; X : d\times n \quad a_i : n\times 1$$

test point: $x = Xa$ Total Mass: $M = \sum_{k=1}^{n} a_k$

$$p_x = [p_1, p_2, \dots, p_6]$$

$$H_x = -\sum_j p_j$$

$$\begin{cases} H_x > \delta & new\ class\ data \\ H_x < \delta & known\ class\ data \end{cases}$$

$$c = \arg\max_k\, p_k$$

Active Learning with LRR for feature Learning

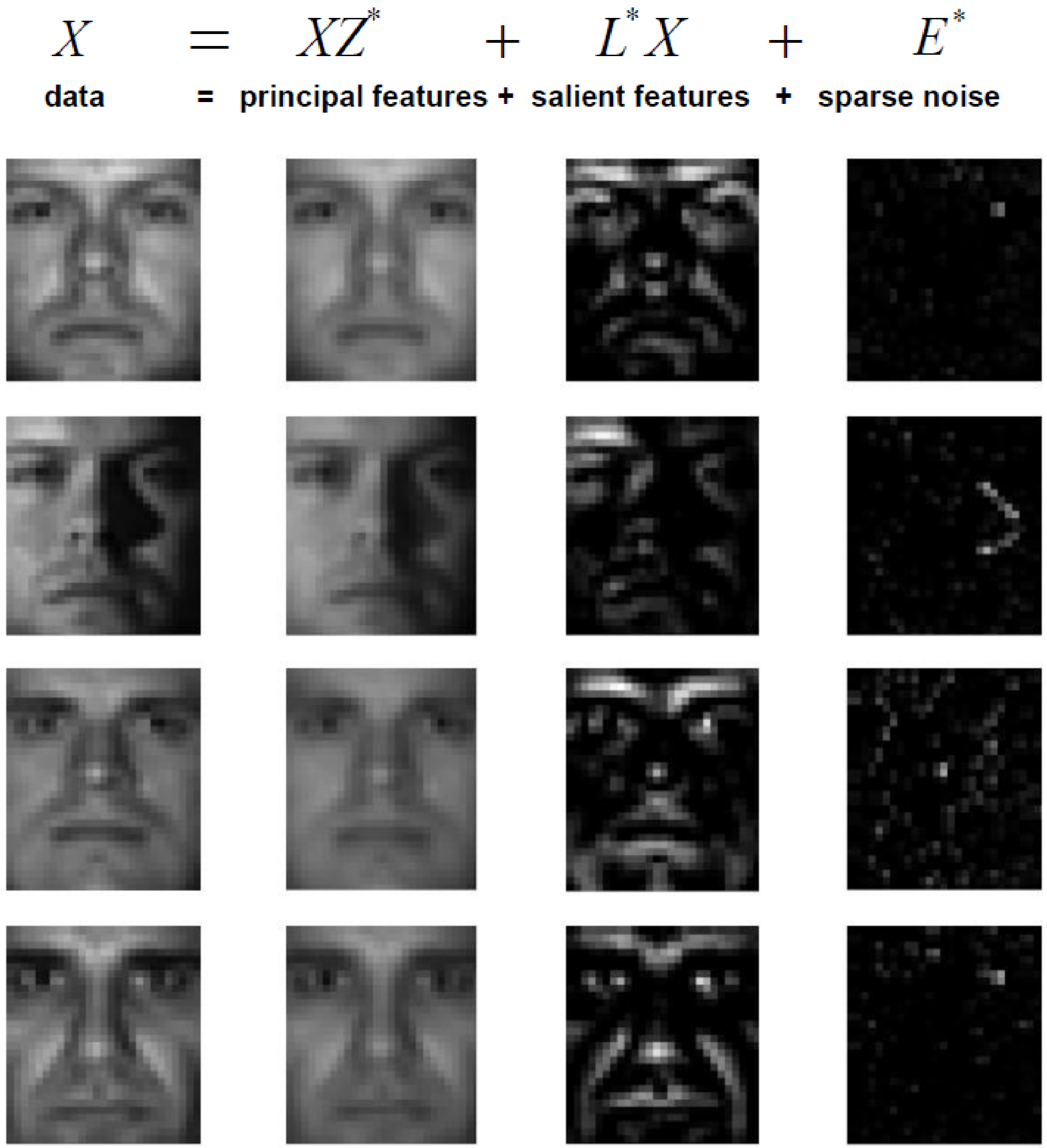
The salient features correspond to the key object parts (e.g., the eyes), which are usually discriminative for recognition

$$J = \min_{A} \ \|(X_o - X_o A)^T\|_{2,1} + \gamma \|A\|_{2,1}$$

$$X_o \in R^{d\times n} \quad A \in R^{n\times n}$$

$$X_o = LX \qquad\qquad L = USU^T = U\Lambda U^T$$

Active learning select $r(1+\epsilon)(\epsilon > 0, r = rank(X_o))$ data points to perform LRR for recognition



$$\min_{A} \ \|(LX - LXA)^T\|_{2,1} + \gamma\|A\|_{2,1} = \min_{A,S} \ \|(US\Sigma V^T - US\Sigma V^T A)^T\|_{2,1} + \gamma\|A\|_{2,1}$$

Optimize

fix A optimize S

fix S optimize A

fix S optimize A

$$J = \min_{A} \ \|(X - XA)^T\|_{2,1} + \gamma\|A\|_{2,1} = \min_{A} \sum_{i=1}^{n} (\|x_i - Xa_i\|_2 + \gamma\|a_i\|_2)$$

$$\frac{\partial J}{\partial a_i} = \frac{-X^T(x_i - Xa_i)}{\|x_i - Xa_i\|_2} + \gamma\frac{a_i}{\|a_i\|_2}$$

$$U: u_{ii} = \frac{1}{\|x_i - Xa_i\|_2}, \qquad V: v_{ii} = \frac{1}{\|a_i\|_2}$$

$$\frac{\partial J}{\partial a_i} = u_{ii}X^TXa_i - u_{ii}X^Tx_i + \gamma v_{ii}a_i = 0 \qquad a_i = (u_{ii}X^TX + \gamma v_{ii}I)^{-1}X^Tx_i$$

$$\frac{\partial J}{\partial A} = X^TXAU - X^TXU + \gamma VA = 0$$

fix A update U V ⟷ fix U V update A

fix A optimize S

$$\min_{A} \|(LX - LXA)^T\|_{2,1} = \min_{A,S} \|(US\Sigma V^T - US\Sigma V^T A)^T\|_{2,1}$$

$$US\Sigma V^T = \sum_{i=1}^{r} s_i\sigma_i u_i v_i^T$$

TIP 2017

$$P = \underset{P}{\operatorname{argmin}} \sum_{i=1}^{N} \|\theta_i P\|_{2,1} + \alpha \|P\|_{2,1} \qquad s.t. P^T B P = I$$

$$\|P\|_{2,1} = 2tr(P^T V P) \quad p^i \neq 0 \qquad V = diag\left(\frac{1}{2}\|p^1\|_2, \frac{1}{2}\|p^2\|_2, \dots, \frac{1}{2}\|p^n\|_2\right) \in R^{n \times n}$$

$$\sum_{i=1}^{N} \|\theta_i P\|_{2,1} = \sum_{i=1}^{r} 2tr\left[(\theta_i P)^T Z_i \theta_i P\right] = tr\left[P^T \left(\sum_{i=1}^{N} \theta_i^T Z_i \theta_i\right) P\right]$$

$$\underset{P, Z_i, V}{\operatorname{argmin}} \; tr\left(P^T \left(\sum_{i=1}^{N} \theta_i^T Z_i \theta_i + \alpha V\right) P\right)$$

update P_{k+1}

$$\left(\sum_{i=1}^{N} \theta_i^T (Z_i)_k \theta_i + \alpha V_k\right) \varphi_j = \lambda_j B \varphi_j$$

first d smallest eigenvalues that are sorted in ascending order

$$\min_{L} \ \|(LX - LXA)^T\|_{2,1} = \min_{S} \ \|(US\Sigma V^T - US\Sigma V^T A)^T\|_{2,1}$$

$$\min_{S} \ \|(US\Sigma V^T - US\Sigma V^T A)^T\|_{2,1} = \min_{S} \ \|(US\Sigma V^T (I - A))^T\|_{2,1}$$

$$= \min_{s_i} \ \left\| \left(\sum_{i=1}^{r} s_i \sigma_i u_i v_i^T \, (I - A) \right)^T \right\|_{2,1}$$

$$\min_{\alpha_i} \ \left\| \sum_{i=1}^{r} \alpha_i \, M_i \right\|_{2,1} \quad s.t. \ \sum_{i} \alpha_i = Const \,, \qquad \alpha_i = s_i \sigma_i, \qquad M_i = (I - A^T) v_i u_i^T$$

$$\|P\|_{2,1} = 2tr(P^T V P) \quad p^i \neq 0 \qquad V = diag\left(\frac{1}{2}\|p^1\|_2, \frac{1}{2}\|p^2\|_2, \dots, \frac{1}{2}\|p^n\|_2\right) \in R^{n\times n}$$

$$\min_S \ \|(US\Sigma V^T - US\Sigma V^T A)^T\|_{2,1} = \min_M \|M\|_{2,1} \quad M = \left(US\Sigma V^T(I-A)\right)^T$$

$$M^T V M = US\Sigma V^T(I-A)V(I-A)^T V\Sigma S U^T = US\Sigma V^T E V \Sigma S U^T$$

$$\|M\|_{2,1} = 2tr(M^T V M) = 2tr(US\Sigma V^T E V\Sigma S U^T) = 2\|S\Sigma V^T E(S\Sigma V^T)^T\|_F^2$$

$$S\Sigma V^T E(S\Sigma V^T)^T = \sum_{i=1}^{r}\sum_{j=1}^{r} s_i s_j \sigma_i \sigma_j v_i^T E v_i$$

$$\min_M \|M\|_{2,1} = \min_{\alpha_i} \sum_{i=1}^{r}\sum_{j=1}^{r} \alpha_i \alpha_j \|E\|_F^2 \quad s.t. \sum_i \alpha_i = Const \quad \alpha_i = s_i \sigma_j$$

$$\underset{P,Z_i,V}{\operatorname{argmin}}\ tr\left(P^T\left(\sum_{i=1}^{N}\theta_i^T Z_i \theta_i + \alpha V\right)P\right)$$

$$s.t. P^T B P = I$$

update P_{k+1}

$$\left(\sum_{i=1}^{N}\theta_i^T (Z_i)_k \theta_i + \alpha V_k\right)\varphi_j = \lambda_j B P$$

$$\underset{P,V}{\operatorname{argmin}}\ tr(P^T V P)$$

$$V_k \varphi_j = \lambda_j B \varphi_j$$

First compute P

$$\underset{S}{\operatorname{argmin}}\ \left\| P - \left(US\Sigma V^T (I - A)\right)^T \right\|_2^2$$