问题与实验

实验目的

- 尝试利用fold-in绕开LatLRR的训练实现快速测试
- 比较不同特征（raw data、XZ、LX）的性状表现，期望LX较好

2018.6.7

$$\min_{Z,L}\|Z\|_* + \|L\|_* + \|E\|_1 \quad s.t.\, X = XZ + LX + E \xrightarrow[\text{Yale Face}]{2016\times 160} \text{17~20 minutes}$$

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

ECML 2013

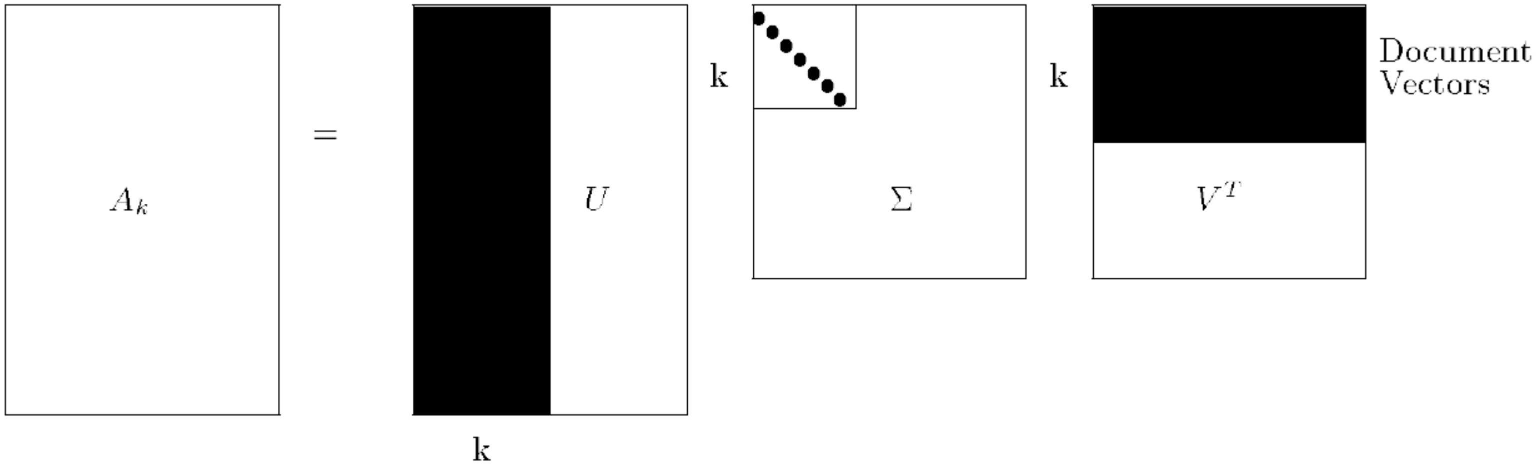
$SVD(X) = U_X\Sigma_X V_X^T$ S is any block-diagonal matrix that satisfies two constraints:

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.

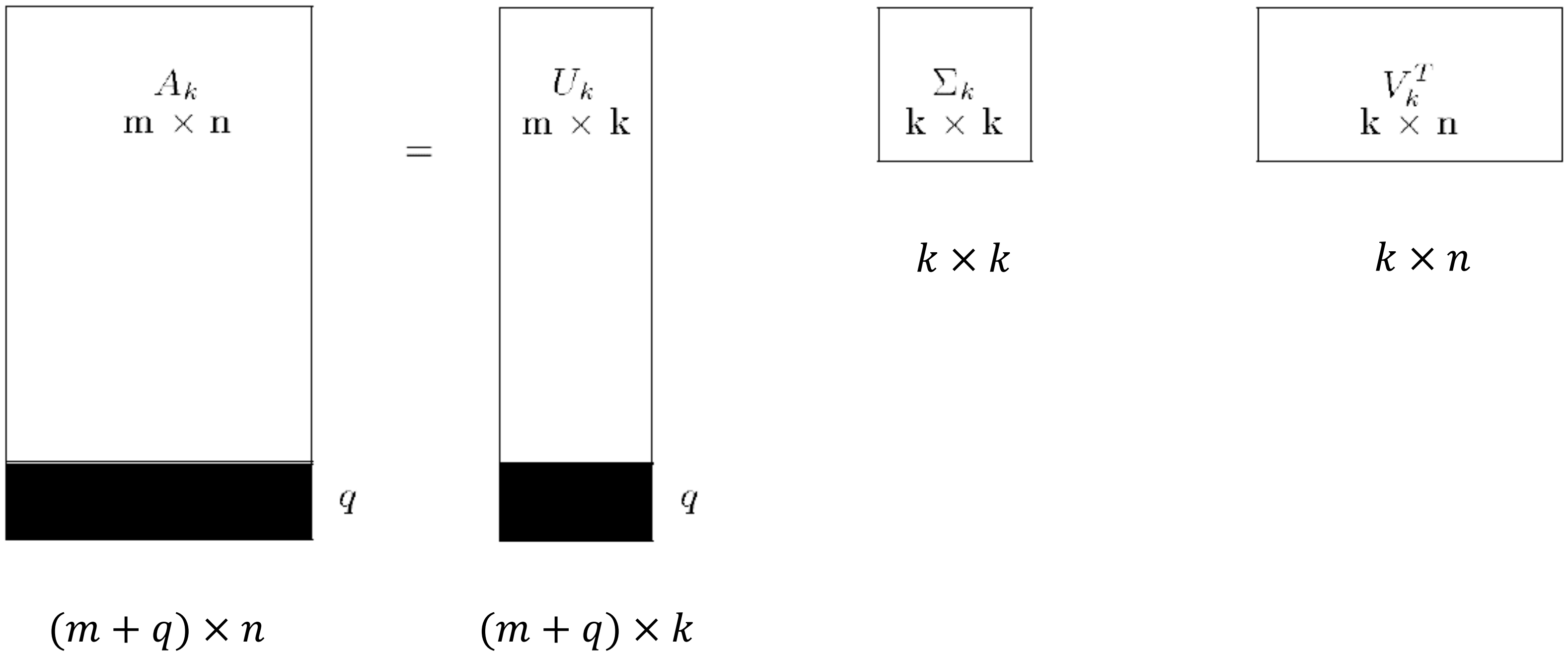
V_X是通过对AL已经选出的点进行SVD分解所得,提前计算V_X、S，新加入一个测试点利用SVD快速计算出新的V_{X_new}，再利用$Z^* = V_X(I-S)V_X^T$直接获得新的关联度矩阵Z_{new}^*

Using Linear Algebra Intelligent Information Retrieval

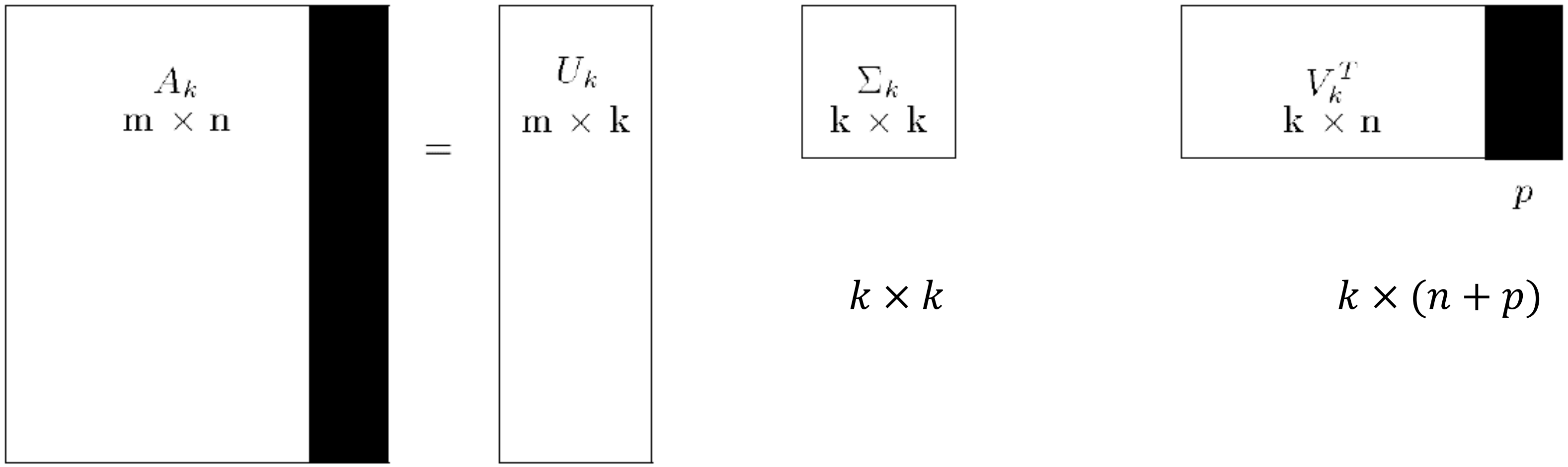
1995 SIAM



$m \times n$ $m \times k$ $k \times k$ $k \times n$

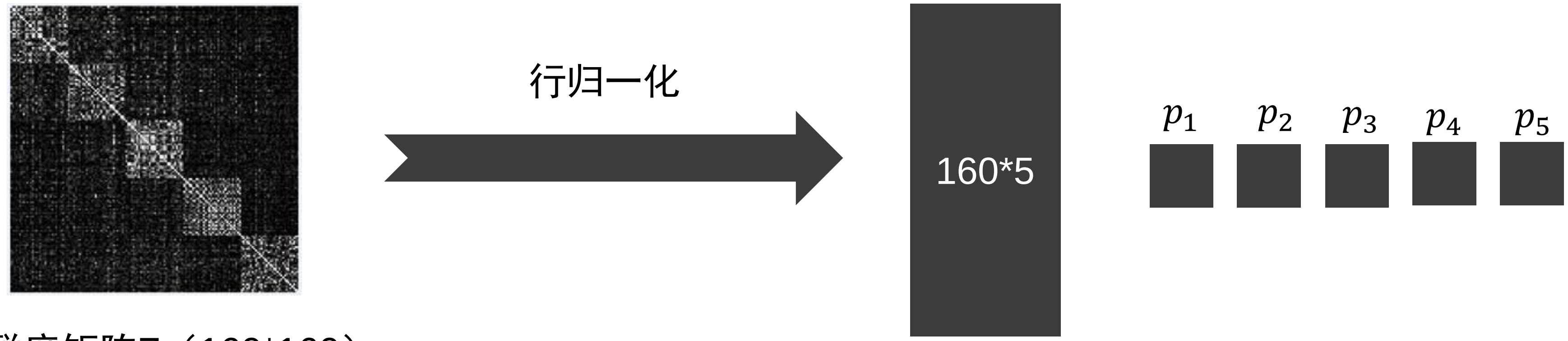


$$\hat{t} = t^T V_k \Sigma_k^{-1} \qquad (q \times k) = (q \times n)\ (n \times k)\ (k \times k)$$

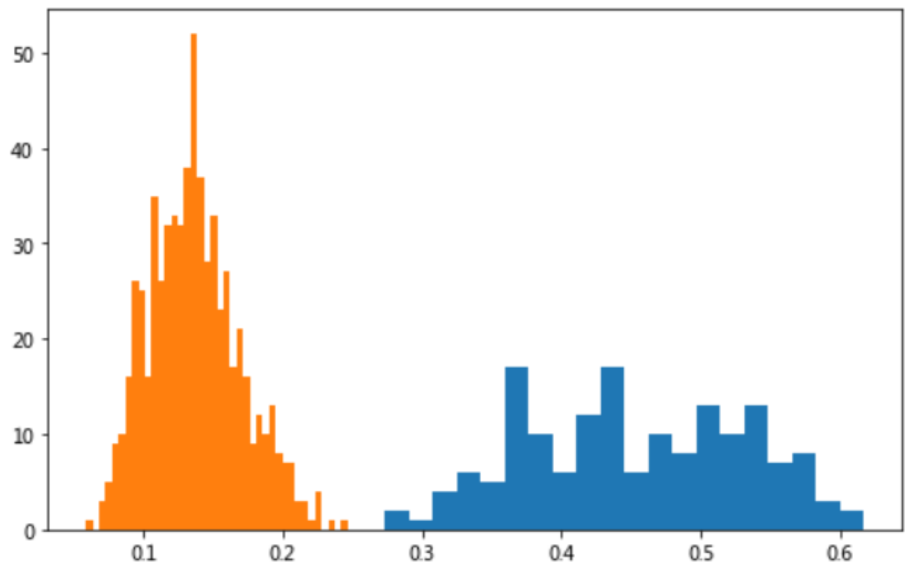


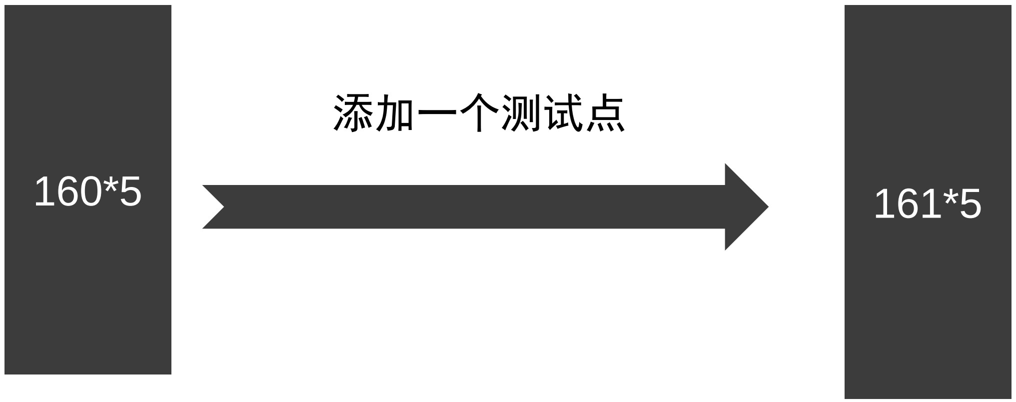
$m \times (n+p)$ $\qquad$ $m \times k$

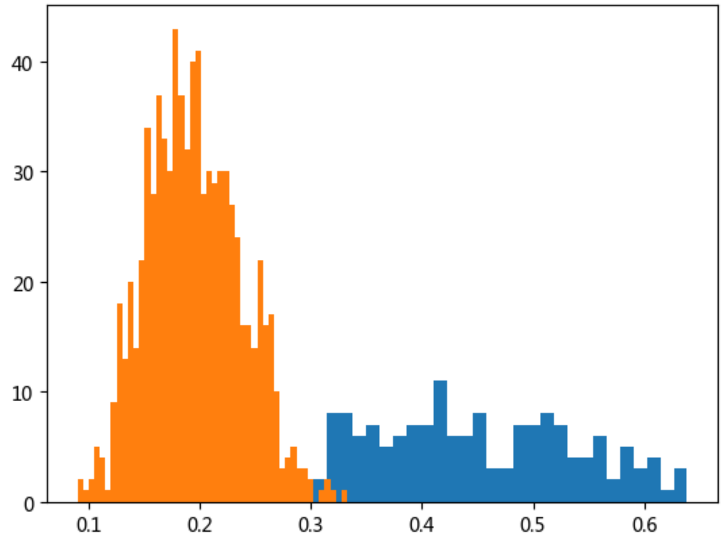
$\hat{d} = d^T U_k \Sigma_k^{-1} \qquad (k \times p) \neq (p \times m)\ (m \times k)\ (k \times k)$

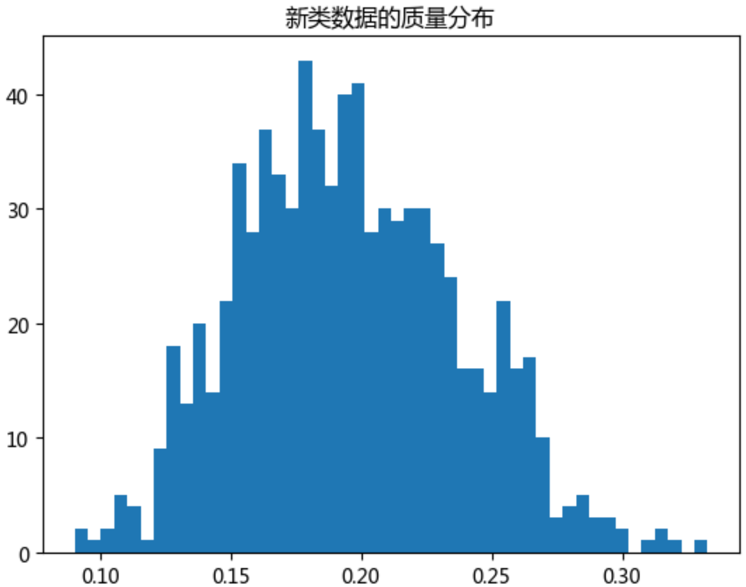


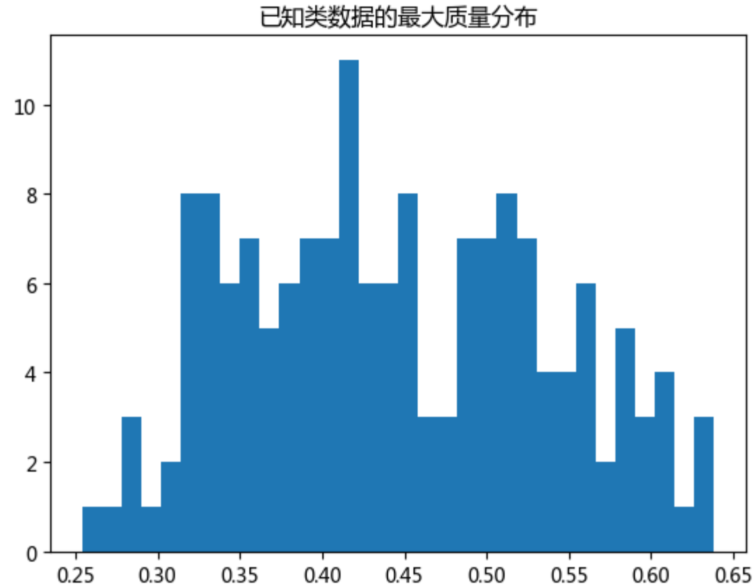
关联度矩阵Z（160*160）

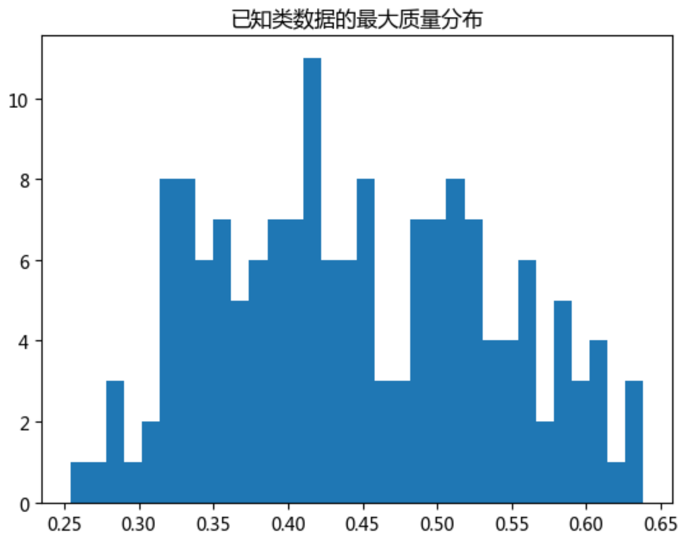


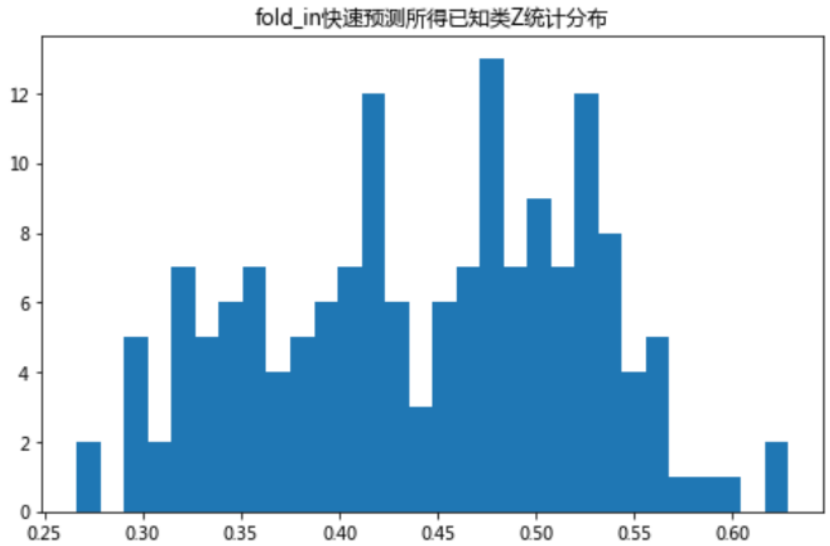

160*5
添加一个测试点
161*5

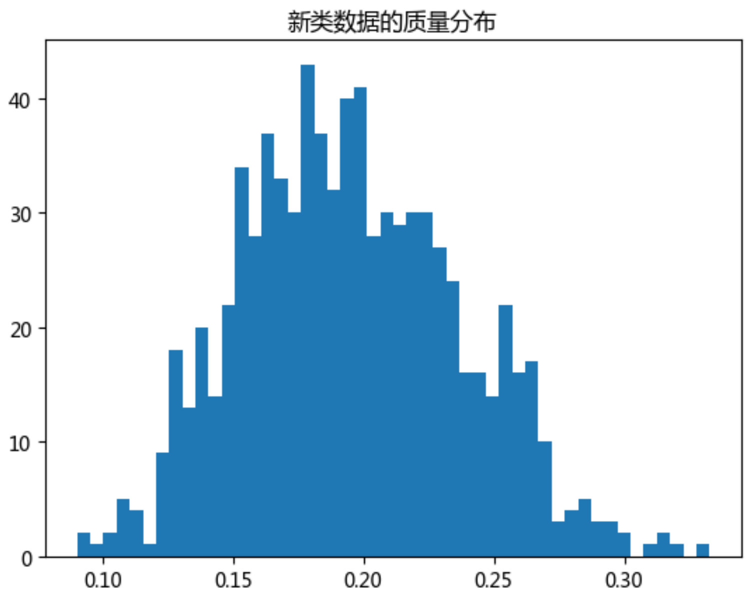


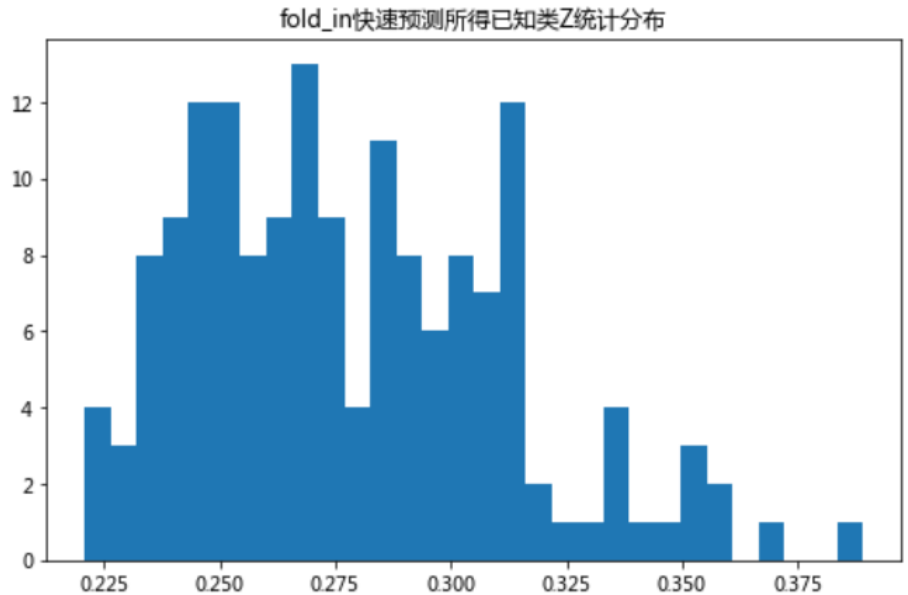

新类数据的质量分布


已知类数据的最大质量分布

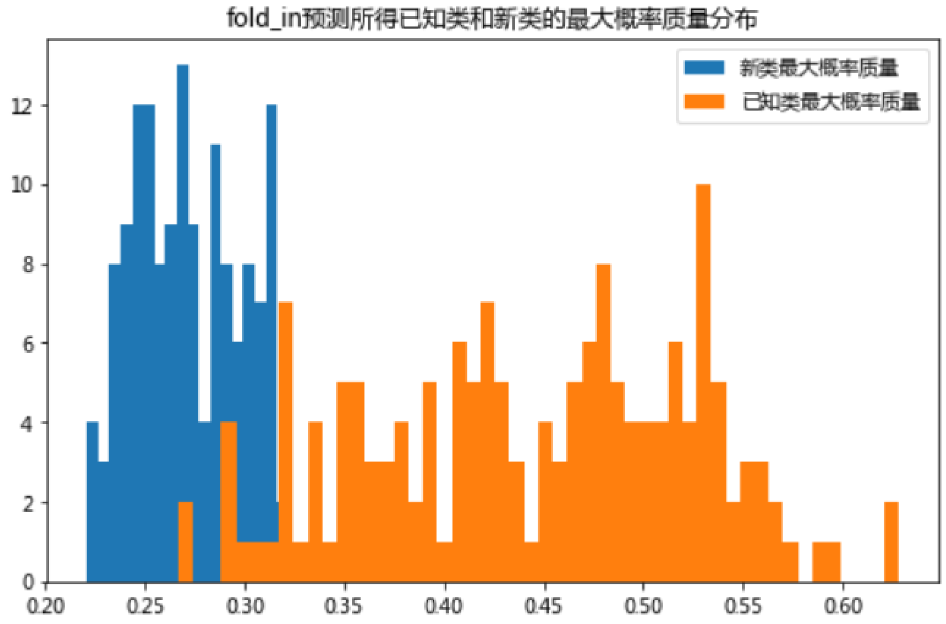

已知类数据的最大质量分布
0
2
4
6
8
10
0.25
0.30
0.35
0.40
0.45
0.50
0.55
0.60
0.65


fold_in快速预测所得已知类Z统计分布
0
2
4
6
8
10
12
0.25
0.30
0.35
0.40
0.45
0.50
0.55
0.60


新类数据的质量分布
0
10
20
30
40
0.10
0.15
0.20
0.25
0.30


fold_in快速预测所得已知类Z统计分布
0
2
4
6
8
10
12
0.225
0.250
0.275
0.300
0.325
0.350
0.375

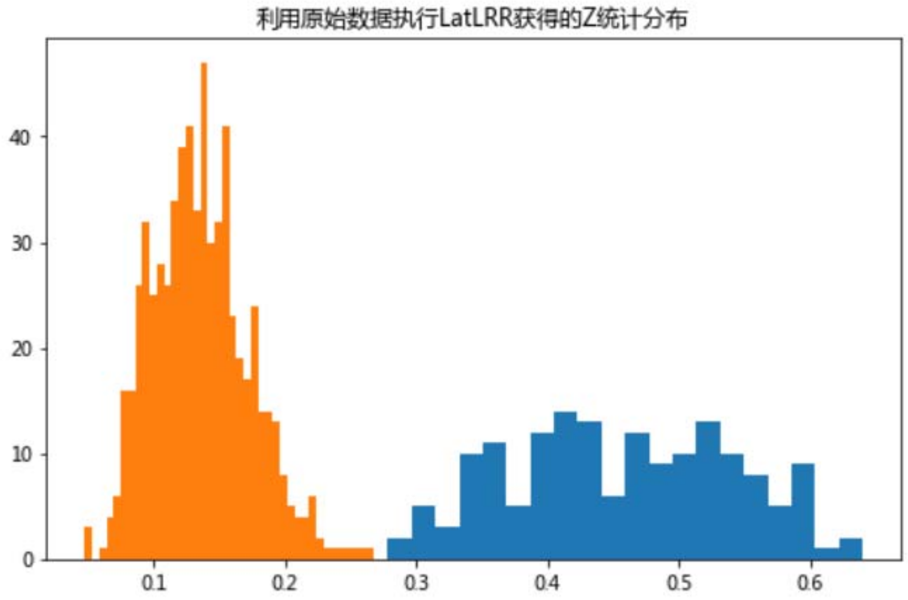
delta = 0.32 acc_known = 0.925 acc_new = 0.89375



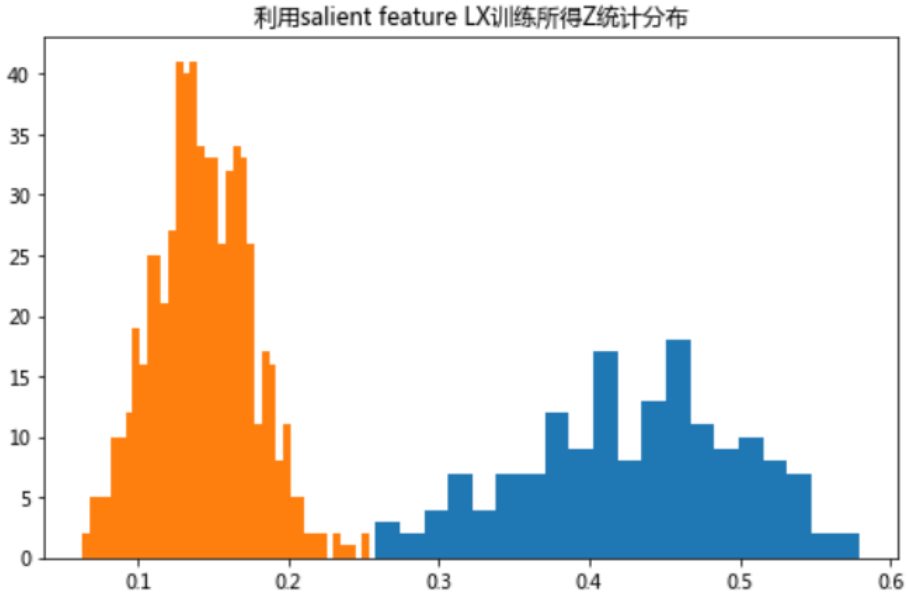
不同特征的比较

p_1 p_2 p_3 p_4 p_5

raw data



LX



$$Accuracy = \frac{A_{new}+A_{known}}{m}$$

AAAI 2017

Table 1: Comparisons of different methods on simulated streams.

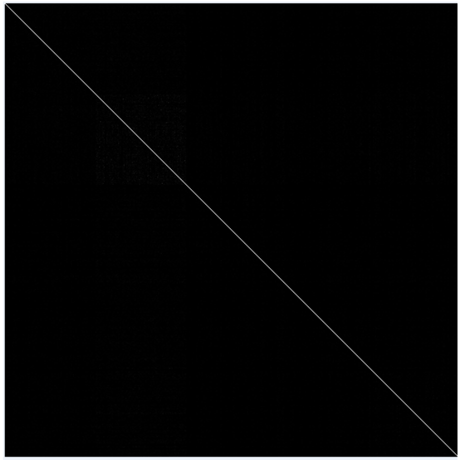
	KDD Cup 99		Forest Cover		MNIST	
Algorithm	Accuracy	F-measure	Accuracy	F-measure	Accuracy	F-measure
iForest+KNN	0.867±0.01	0.882±0.04	0.707± 0.02	0.742± 0.08	0.720 ± 0.06	0.593 ± 0.05
LACU-SVM	0.853±0.05	0.827± 0.02	0.702± 0.07	0.752± 0.11	0.752 ± 0.01	0.697 ± 0.04
SAND-F	0.880±0.03	0.892± 0.03	0.799± 0.02	0.771± 0.03	0.725 ± 0.04	0.622 ± 0.03
ECSMiner	0.857± 0.03	0.852±0.06	**0.823± 0.05**	**0.794± 0.02**	0.745 ± 0.07	0.641 ± 0.03
SENC-non-MaS	0.792± 0.08	0.722±0.13	0.633± 0.02	0.651± 0.05	0.701 ± 0.07	0.652 ± 0.06
SENC-MaS	**0.897± 0.03**	**0.902±0.02**	0.792± 0.04	0.752± 0.07	**0.817 ± 0.05**	**0.710 ± 0.09**

Table 1: F1, Precesion and Recall for seen classes 2,3,5 in MNIST dataset (NIPC means number of instances per class)

	Measure	NIPC	OC-SVM	MOC-SVM	1-vs-Set	OVR-SVM	NNO	ASG-SVM
Seen Class	F1	100	.398±.033	.446±.028	.488±.046	.552±.021	.424±.013	**.570±.029**
		1000	.389±.026	.437±.027	.542±.043	.612±.013	.421±.010	**.624±.024**
	Precision	100	.330±.037	.414±.033	.336±.051	.398±.019	.367±.012	**.539±.027**
		1000	.314±.027	.374±.031	.388±.048	.534±.017	.379±.011	**.566±.022**
	Recall	100	.502±.031	.484±.030	.882±.046	**.898±.024**	.501±.014	.605±.031
		1000	.511±.029	.525±.027	**.897±.044**	.716±.014	.474±.012	.697±.025
Unseen Class	F1	100	.645±.035	.918±.031	.882±.048	.895±.023	.702±.012	**.933±.030**
		1000	.617±.031	.905±.029	.897±.044	.930±.016	.722±.011	**.935±.026**
	Precision	100	.736±.036	.941±.034	.983±.049	**.987±.020**	.763±.014	.957±.028
		1000	.726±.027	.944±.033	**.986±.043**	.968±.018	.763±.012	.966±.024
	Recall	100	.577±.032	.897±.032	.752±.053	.818±.025	.650±.013	**.911±.034**
		1000	.538±.030	.870±.028	.789±.048	.896±.013	.685±0.01	**.907±.026**

IJCAI 2017

compare with mnist



$\lambda \nearrow$

$\lambda \searrow$

Algorithm 1 Solving Problem (6) by Inexact ALM

Initialize: $Z = J = 0, L = S = 0, E = 0, Y_1 = 0, Y_2 = 0, Y_3 = 0, \mu = 10^{-6}, max_u = 10^6, \rho = 1.1$, and $\varepsilon = 10^{-6}$.

while not converged **do**

1. Fix the others and update J by setting $J = \arg\min_J \frac{1}{\mu}||J||_* + \frac{1}{2}||J - (Z + Y_2/\mu)||_F^2$.

2. Fix the others and update S by setting $S = \arg\min_S \frac{1}{\mu}||S||_* + \frac{1}{2}||S - (L + Y_3/\mu)||_F^2$.

3. Fix the others and update Z by setting $Z = (\mathtt{I} + X^T X)^{-1}(X^T(X - LX - E) + J + (X^T Y_1 - Y_2)/\mu)$.

4. Fix the others and update L by setting $L = ((X - XZ - E)X^T + S + (Y_1 X^T - Y_3)/\mu)(\mathtt{I} + XX^T)^{-1}$.

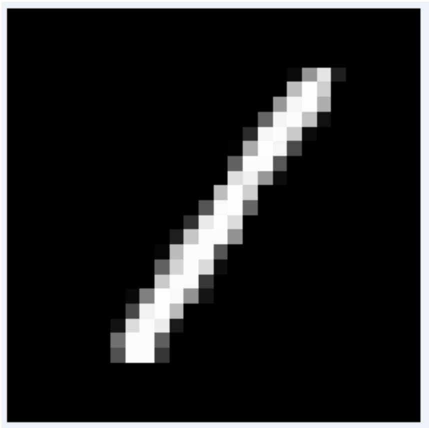
5. Fix the others and update E by setting $E = \arg\min_E \boxed{\lambda/\mu}||E||_1 + 0.5||E - (X - XZ - LX + \boxed{Y_1)/\mu||_F^2}$

6. Update the multipliers by $Y_1 = Y_1 + \mu(X - XZ - LX - E), Y_2 = Y_2 + \mu(Z - J), Y_3 = Y_3 + \mu(L - S)$.

7. Update the parameter μ by $\mu = \min(\rho\mu, max_u)$.

8. Check the convergence conditions: $||X - XZ - LX - E||_\infty < \varepsilon$, $||Z - J||_\infty < \varepsilon$, and $||L - S||_\infty < \varepsilon$.

end while

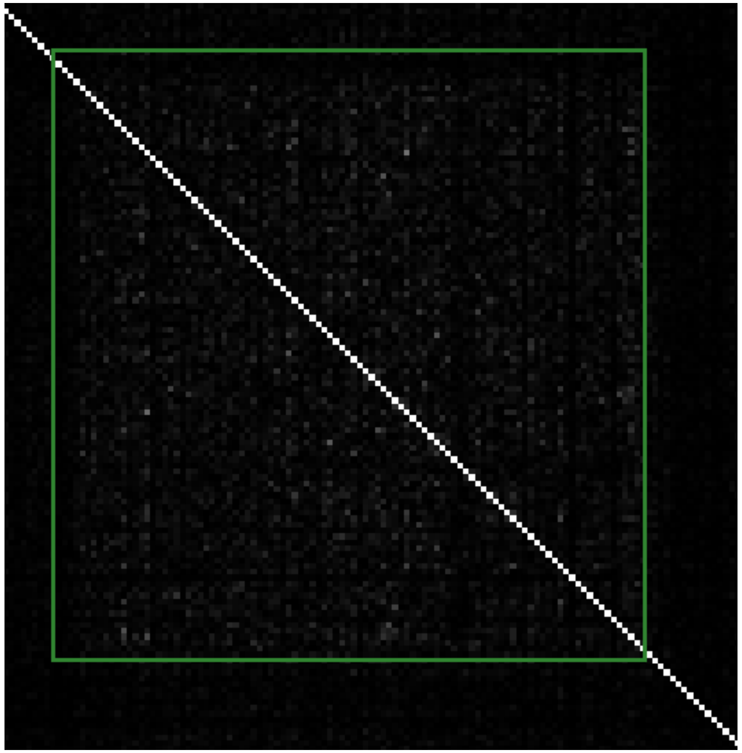


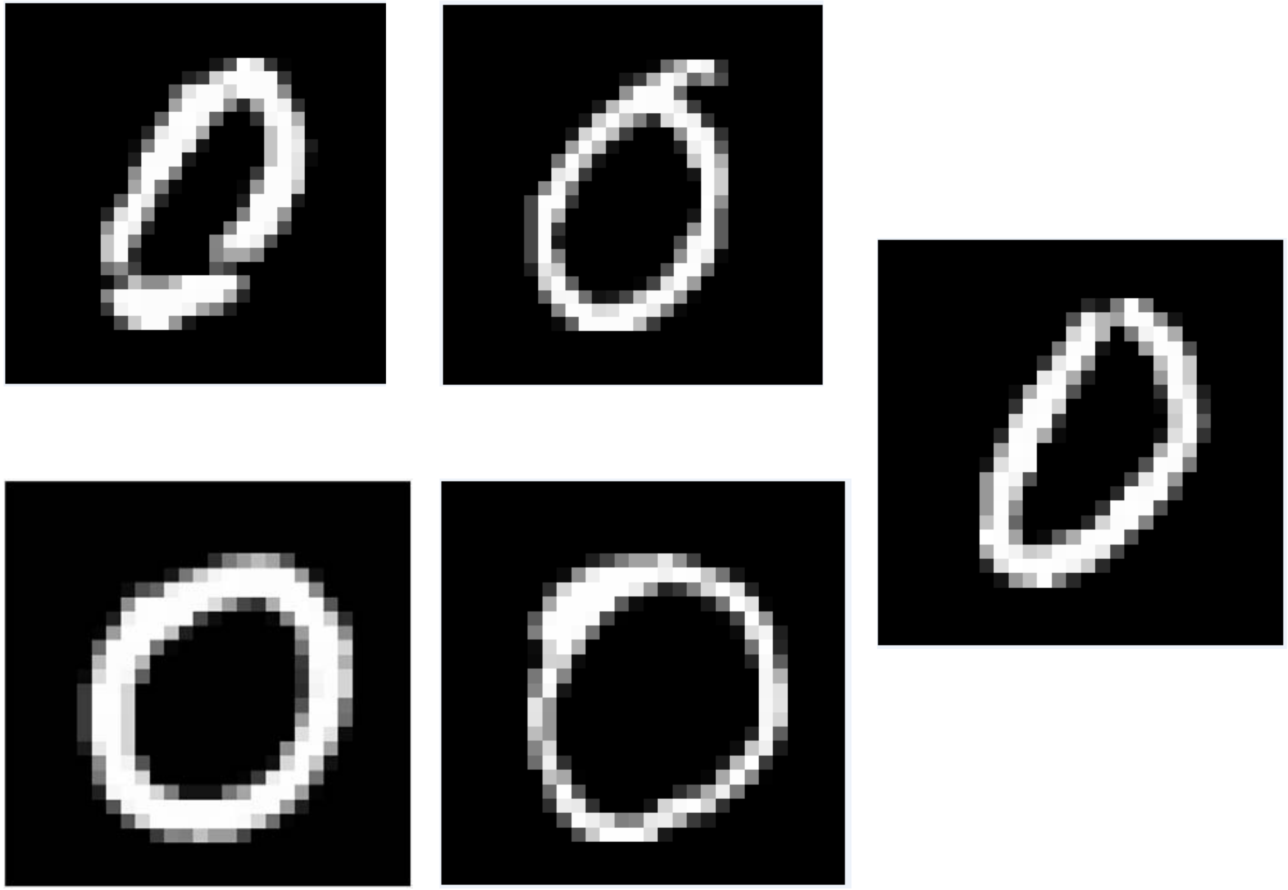
$$\min_{Z,L} \|Z\|_* + \|L\|_* + \|E\|_1 \quad s.t. X = XZ + LX + E$$

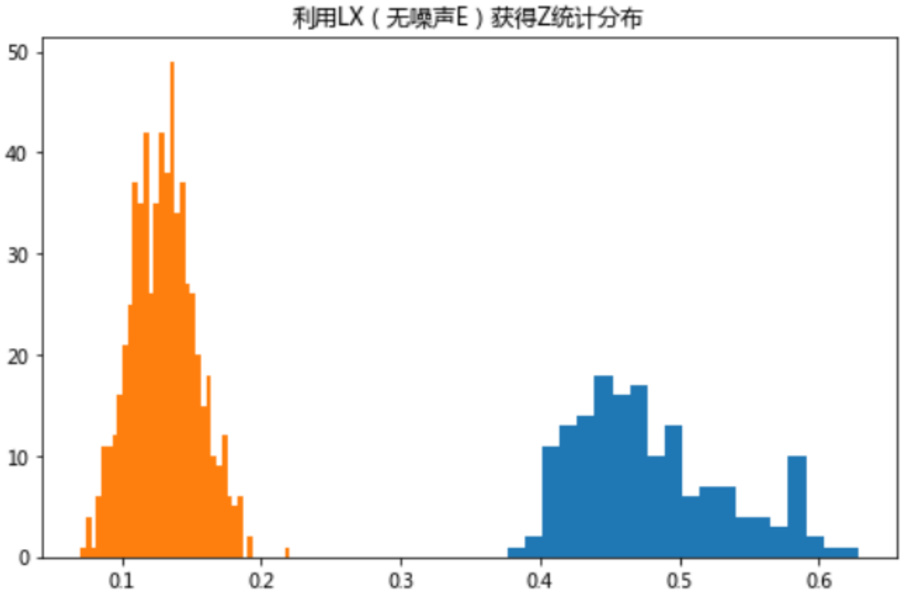
$$\min_{Z,L} \|Z\|_* + \|L\|_* \quad s.t. X = XZ + LX$$

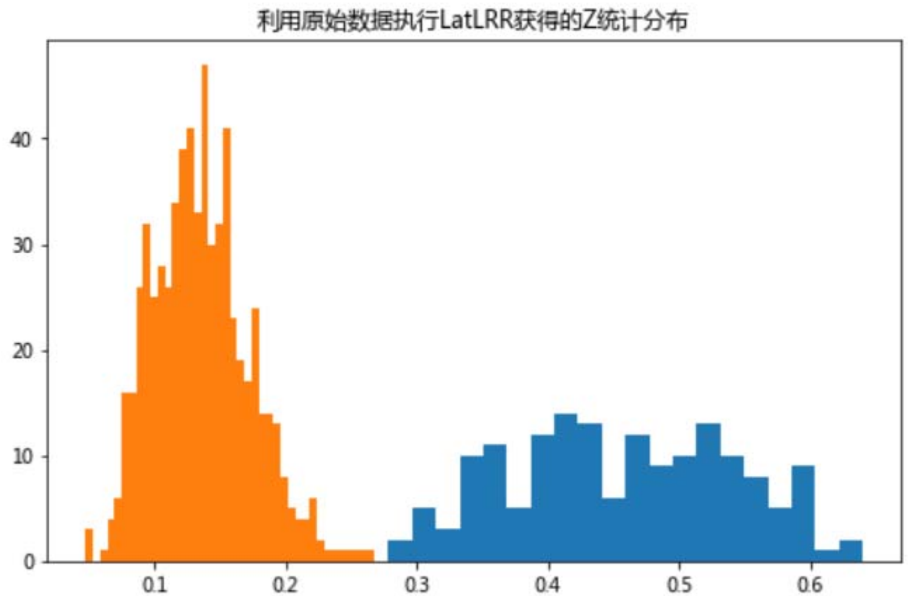
```
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.
    0.    0.    0.    0.    0.    0.    0.  189.  253.  206.    0.    0.
    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.
    0.    0.    0.    0.    0.    0.  151.  246.  253.  174.    0.    0.
    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.
    0.    0.    0.    0.    0.  100.  244.  252.  192.   17.    0.    0.
    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.
    0.    0.    0.    0.   43.  246.  252.  200.   11.    0.    0.    0.
    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.
    0.    0.    0.    3.  170.  253.  245.   73.    0.    0.    0.    0.
    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.
    0.    0.    0.   97.  252.  252.   87.    0.    0.    0.    0.    0.
    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.
    0.    0.    0.  222.  240.  151.   13.    0.    0.    0.    0.    0.
    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.
    0.    0.  189.  253.  206.    0.    0.    0.    0.    0.    0.    0.
    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.
    0.   85.  250.  253.  102.    0.    0.    0.    0.    0.    0.    0.
    0.    0.    0.    0.]
```

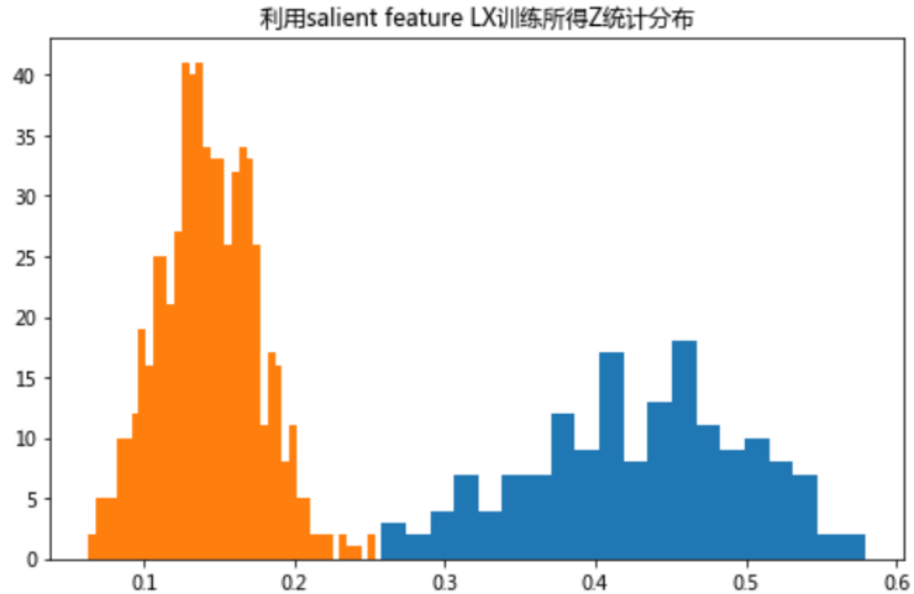
$Z_{ii} = 0/1e-5$






利用LX（无噪声E）获得Z统计分布
0
10
20
30
40
50
0.1
0.2
0.3
0.4
0.5
0.6


利用原始数据执行LatLRR获得的Z统计分布
0
10
20
30
40
0.1
0.2
0.3
0.4
0.5
0.6


利用salient feature LX训练所得Z统计分布
0
5
10
15
20
25
30
35
40
0.1
0.2
0.3
0.4
0.5
0.6