实验探究

2018.6.28

$$J = \min_{A,c} \|(X - XA)^T\|_{2,1} + \gamma\|AA^T - C\|_F^2 + \lambda\|c\|_1$$

$$\frac{\partial J}{\partial A} = X^T XAU - X^T XU + 4\gamma(AA^T - C)A = 0 \qquad U: u_{ii} = \frac{1}{2\|x_i - Xa_i\|_2}$$

For each i compute:

u_{ii}

For each i compute:

$a_i = A[:, i]$

(update M)

update M

compute :

$\min_c \gamma\|AA^T - C\|_F^2 + \lambda\|c\|_1$

$M = AA^T$

$$u_{ii}X^T Xa_i - u_{ii}X^T x_i + 4\gamma(M - C)a_i = 0$$

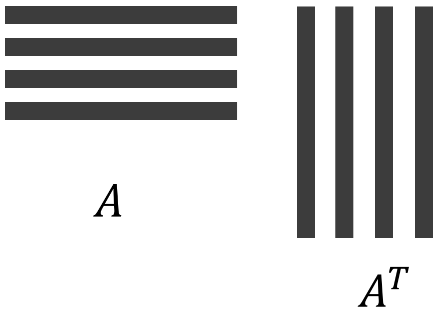
$$a_i = u_{ii}[u_{ii}X^T X + 4\gamma(M - C)]^{-1}X^T x_i$$

$M = AA^T A$

$$u_{ii}X^T Xa_i - u_{ii}X^T x_i + 4\gamma m_i - 4\gamma Ca_i = 0$$

$$a_i = [u_{ii}X^T X - 4\gamma C]^{-1}[u_{ii}X^T x_i - 4\gamma m_i]$$

$$\min_{A,c} \gamma \|AA^T - C\|_F^2 + \lambda \|c\|_1$$



对于式(11.12), 可先计算 $\boldsymbol{z} = \boldsymbol{x}_k - \frac{1}{L}\nabla f(\boldsymbol{x}_k)$, 然后求解

$$\boldsymbol{x}_{k+1} = \arg\min_{\boldsymbol{x}} \frac{L}{2}\|\boldsymbol{x} - \boldsymbol{z}\|_2^2 + \lambda\|\boldsymbol{x}\|_1 \ . \tag{11.13}$$

令 x^i 表示 $\boldsymbol{x}$ 的第 i 个分量, 将式(11.13)按分量展开可看出, 其中不存在 $x^i x^j$ $(i \neq j)$ 这样的项, 即 $\boldsymbol{x}$ 的各分量互不影响, 于是式(11.13)有闭式解

$$x_{k+1}^i = \begin{cases} z^i - \lambda/L, & \lambda/L < z^i \ ; \\ 0, & |z^i| \leqslant \lambda/L \ ; \\ z^i + \lambda/L, & z^i < -\lambda/L \ , \end{cases} \tag{11.14}$$

其中 x_{k+1}^i 与 z^i 分别是 $\boldsymbol{x}_{k+1}$ 与 $\boldsymbol{z}$ 的第 i 个分量. 因此, 通过 PGD 能使 LASSO 和其他基于 L_1 范数最小化的方法得以快速求解.

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 $E = \arg\min_E \lambda/\mu||E||_1 + 0.5||E - (X - XZ - LX + Y_1)/\mu||_F^2$.

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

$$\frac{\partial J}{\partial A} = X^T XAU - X^T XU + 4\gamma(AA^T - C)A = 0$$

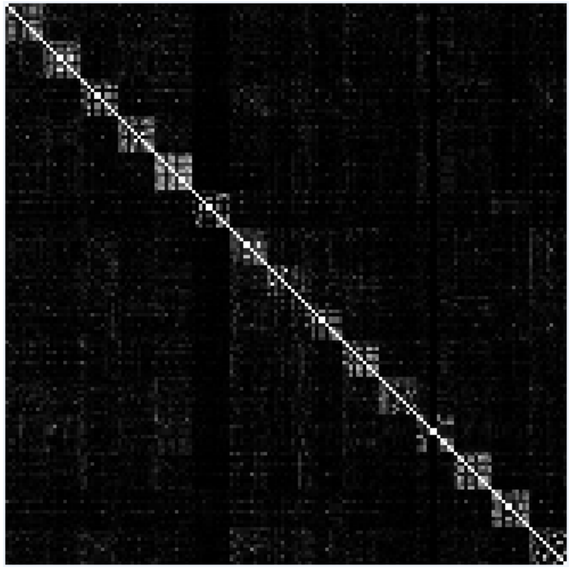
$$\gamma = 0 \rightarrow A = I$$

$$J = \min_{A,c} \ \|(X - XA)^T\|_{2,1} + \gamma\|AA^T - C\|_F^2 + \lambda\|c\|_1 + \beta Tr(A)$$

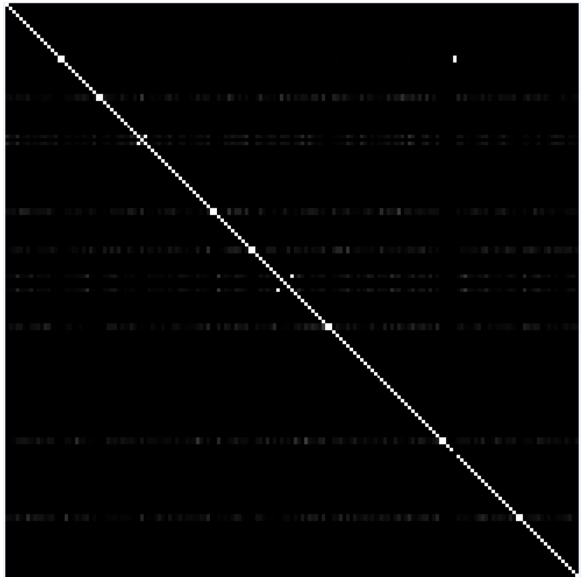
$$J = \min_{A} \ \|(X - XA)^T\|_{2,1} + \gamma\|A\|_{2,1}$$

$$J = \min_{A,c} \ \|(X - XA)^T\|_{2,1} + \gamma\|AA^T - C\|_F^2 + \lambda\|c\|_1$$

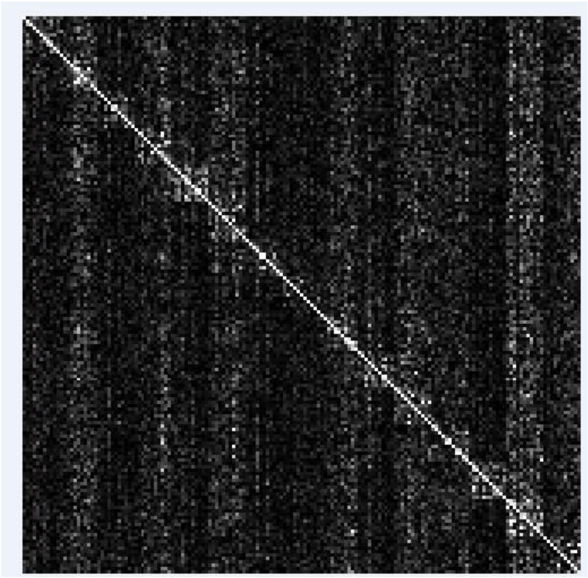
$\gamma = 20$



$\gamma = 0.1$



$\gamma = 25$



$$J = \min_{A,c} \|(X - XA)^T\|_{2,1} + \gamma\|AA^T - C\|_F^2 + \lambda\|c\|_1 + \beta Tr(A)$$

For each i compute:

u_{ii}

For each i compute:

$a_i = A[:,i]$

(update M)

compute :

$\min_{c} \gamma\|AA^T - C\|_F^2 + \lambda\|c\|_1$

$$U: u_{ii} = \frac{1}{2\|x_i - Xa_i\|_2}$$

$$\frac{\partial J}{\partial A} = X^T XAU - X^T XU + 4\gamma(AA^T \; - C)A + \beta I = 0$$

$$M = AA^T$$

$$u_{ii}X^T Xa_i - u_{ii}X^T x_i + 4\gamma(M - C)a_i + \beta I = 0$$

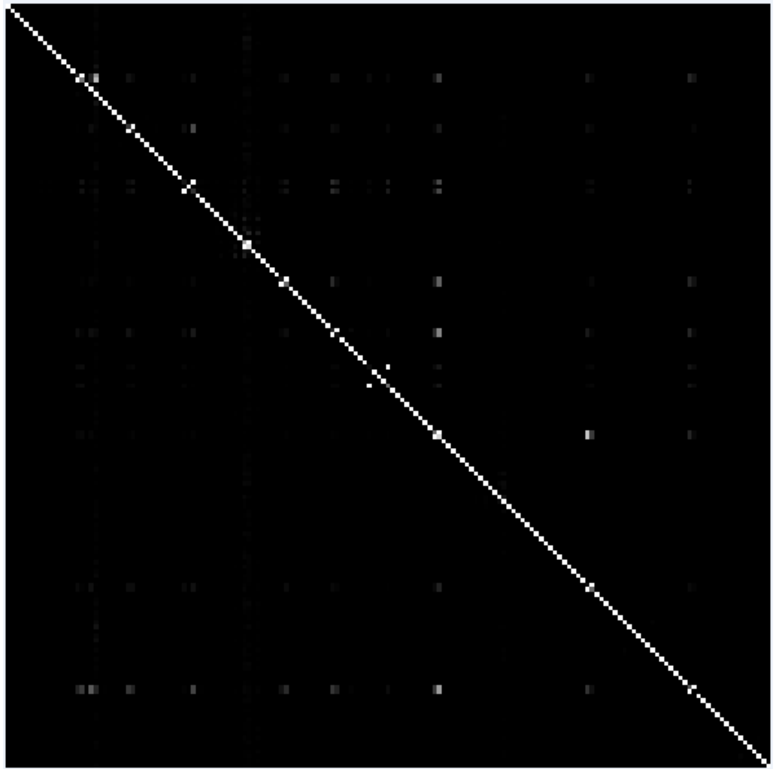
$$a_i = [u_{ii}X^T X + 4\gamma(M - C)]^{-1}[u_{ii}X^T x_i - \beta I_i]$$

$$\text{condition} = \left\|A^{(k+1)} - A^{(k)}\right\|_\infty$$

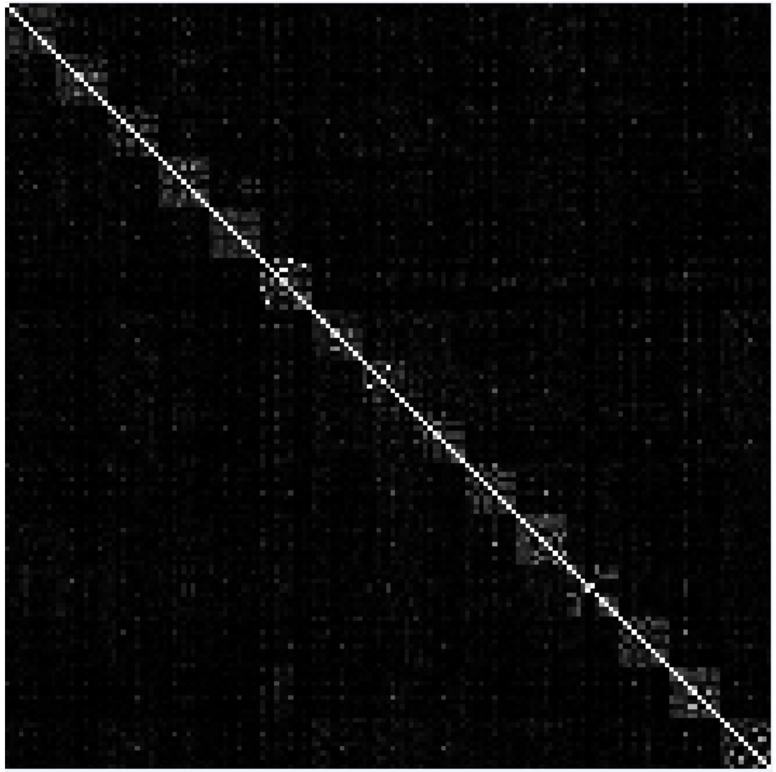
$$J = \min_{A,c} \ \|(X - XA)^T\|_{2,1} + \gamma\|AA^T - C\|_F^2 + \lambda\|c\|_1 + \beta Tr(A)$$

$$\gamma = 0.1 \quad \lambda = 100$$

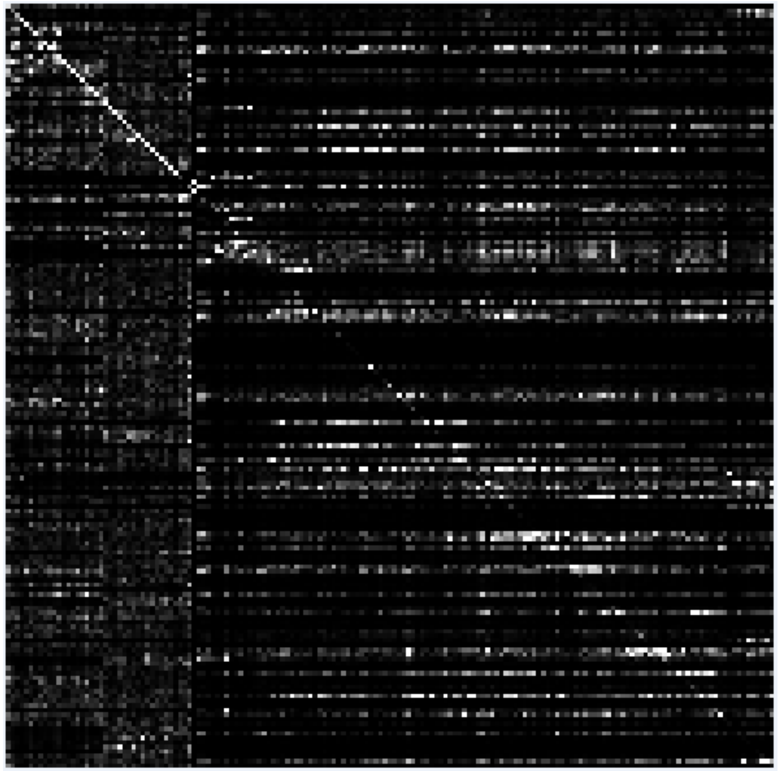
$\beta = 1$



$\beta = 10$



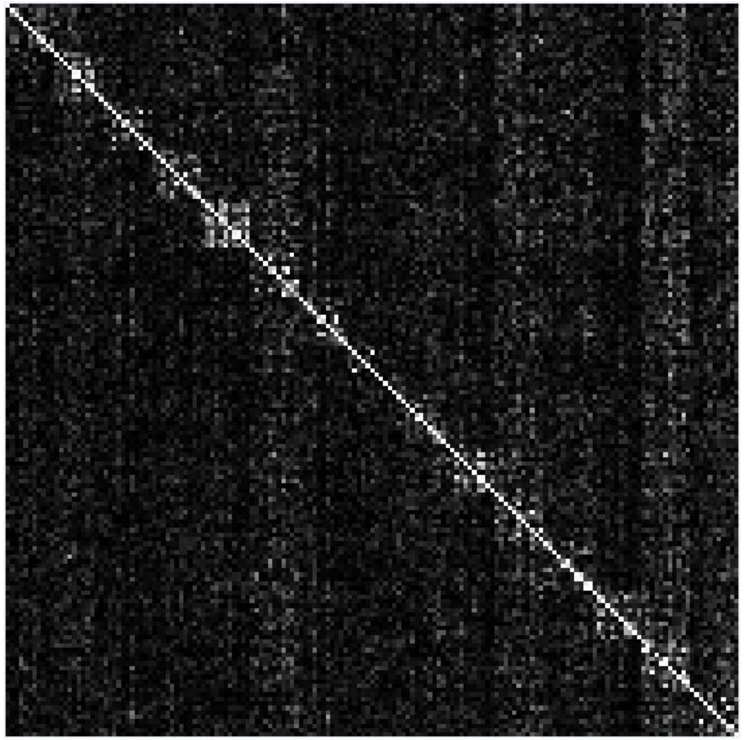
$\beta = 25$



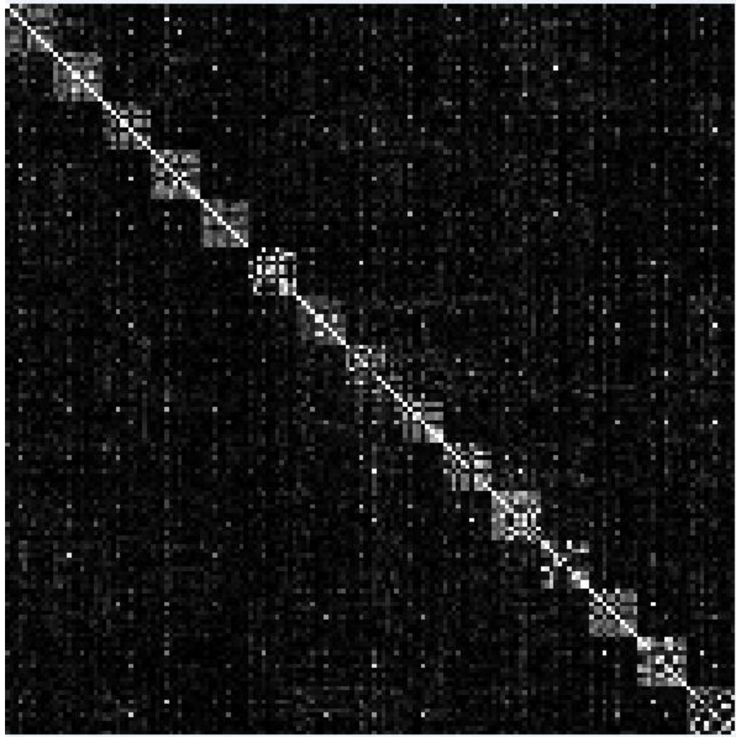
$$J = \min_{A,c} \ \|(X - XA)^T\|_{2,1} + \gamma\|AA^T - C\|_F^2 + \lambda\|c\|_1 + \beta Tr(A)$$

$$\gamma = 20 \quad \lambda = 1000$$

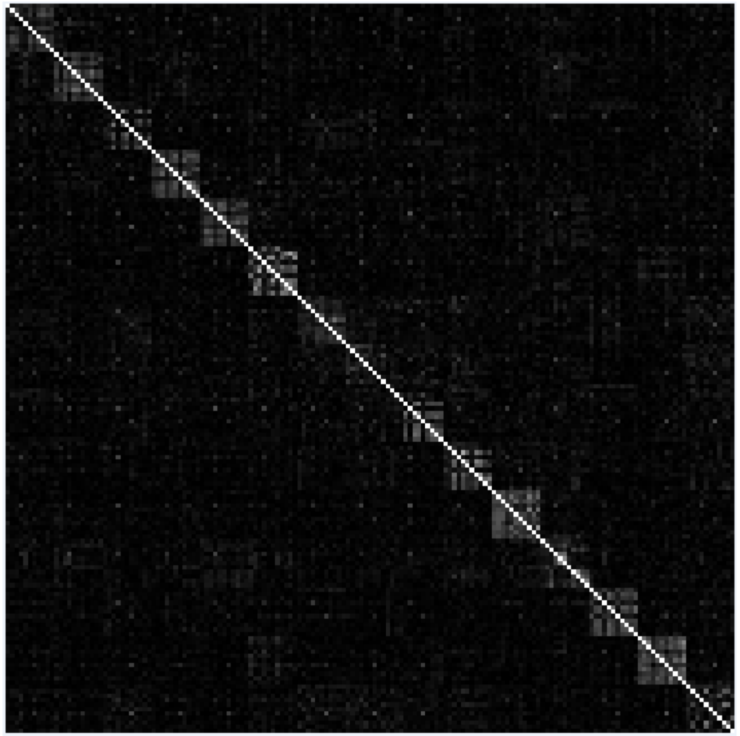
$\beta = 1$

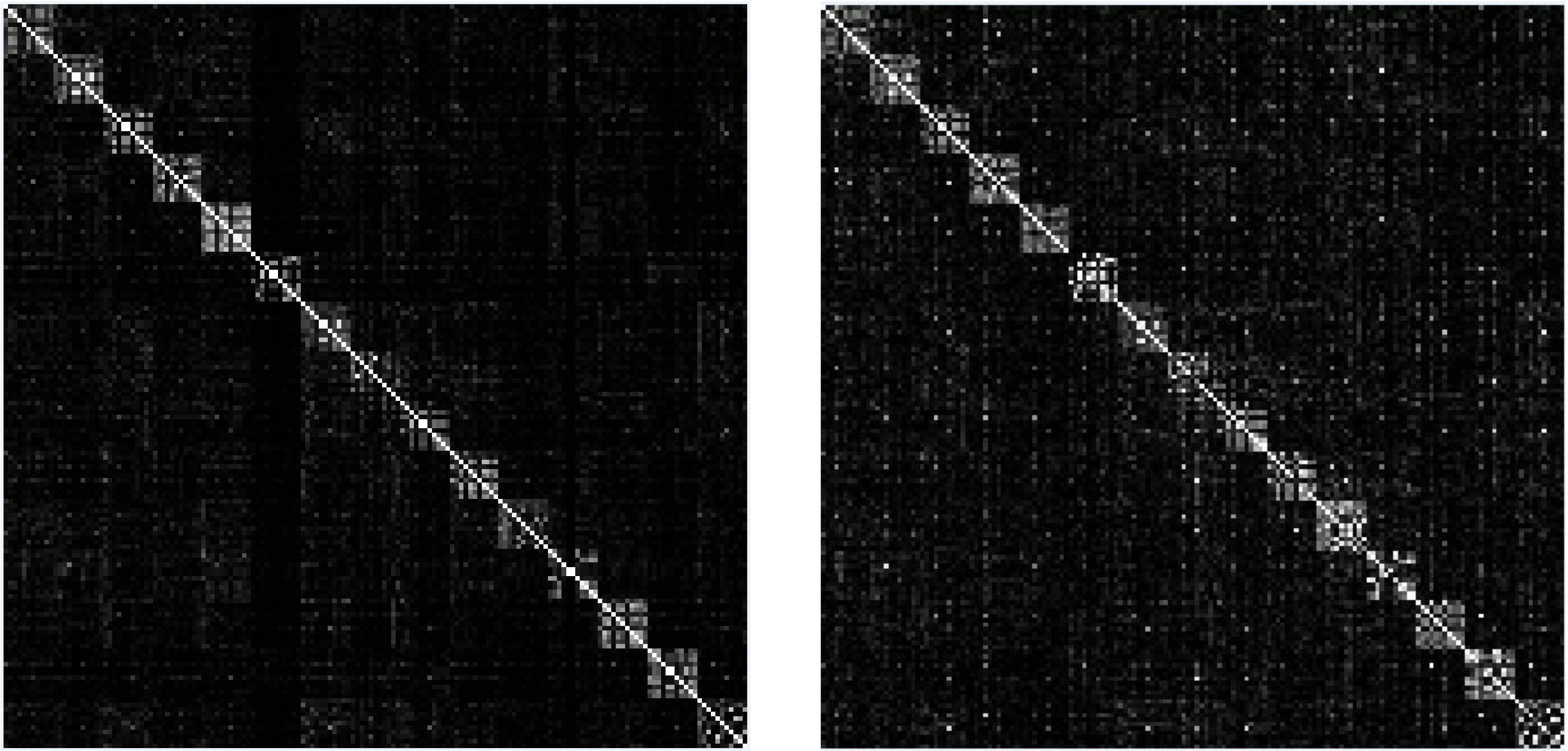


$\beta = 10$



$\beta = 25$





$$J = \min_{A,c} \ \|(X - XA)^T\|_{2,1} + \gamma\|AA^T - C\|_F^2 + \lambda\|c\|_1 + \beta Tr(A)$$

$$\min_{c} \gamma\|AA^T - C\|_F^2 + \lambda\|c\|_1$$

$$\boldsymbol{J = \min_{A,c} \ \left\|(X - XA)^T\right\|_{2,1} + \gamma\left\|\alpha I \circ AA^T - C\right\|_F^2 + \lambda\|c\|_1 + \beta Tr(A)}$$

$$\boldsymbol{\min_{c} \alpha\gamma\left\|AA^T - C\right\|_F^2 + \lambda\|c\|_1}$$

$$M = AA^T$$

$$u_{ii}X^T X a_i - u_{ii}X^T x_i + 4\gamma(\alpha I \odot M - C)a_i + \beta I = 0$$

$$a_i = [u_{ii}X^T X + 4\gamma(\alpha I \odot M - C)]^{-1}[u_{ii}X^T x_i - \beta I_i]$$

发散

0	132.754615871
1	250.022028446
2	246.974418147
3	67.1678533437
4	3895.17266484
5	2.60201592225e+13
6	3.38989496042e+15
7	6.27534517879e+15
8	3.40474522527e+15
9	5.69296228877e+14
10	7.08299711849e+13
11	5.61912608559e+13
12	1.10123962491e+14
13	1.2769774467e+14
14	1.25850866227e+14

不稳定收敛

0	298.306167738
1	249.396290217
2	63.2561886152
3	28.0695339225
4	13.4056337035
5	8.03934969875
6	7.16633535709
7	8.01880042924
8	10.5997939241
9	9.68991452295
10	11.3607668235
11	10.3851267168
12	8.32750860286
13	9.57717915154
14	17.3764564148

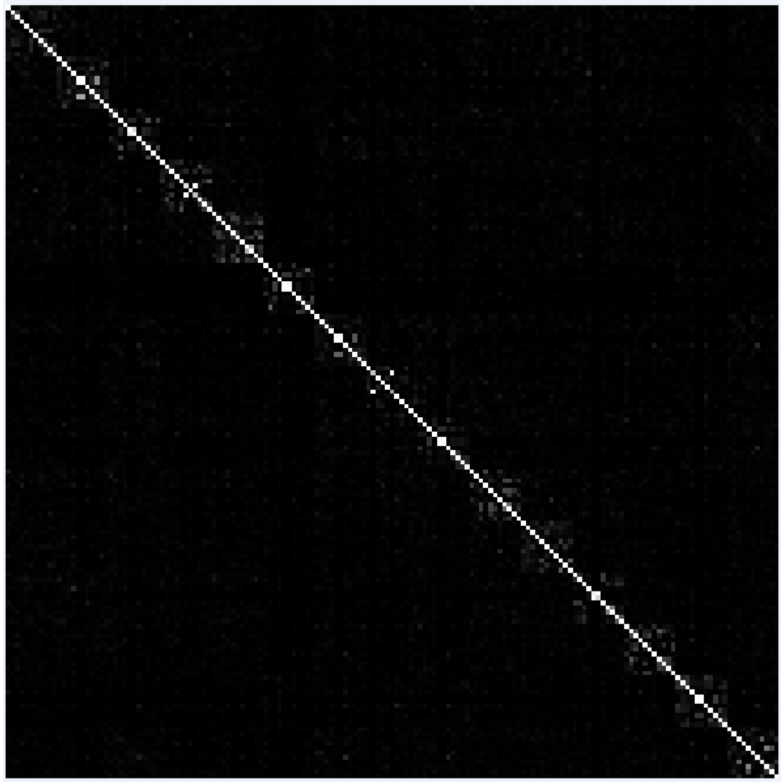
稳定收敛

0	91.6665835061
1	2.18943673032
2	1.71088645991
3	1.63700719974
4	0.806797026423
5	0.5676192167
6	0.410936619691
7	0.164639672102
8	0.0958034126188
9	0.0370229018337
10	0.0146193499335
11	0.00710508556323
12	0.0030408995484
13	0.00141397072025
14	0.000642247208244
15	0.000282608451969
16	0.000136134408896
17	5.66143035831e-05
18	2.88865602959e-05
19	1.13433723171e-05

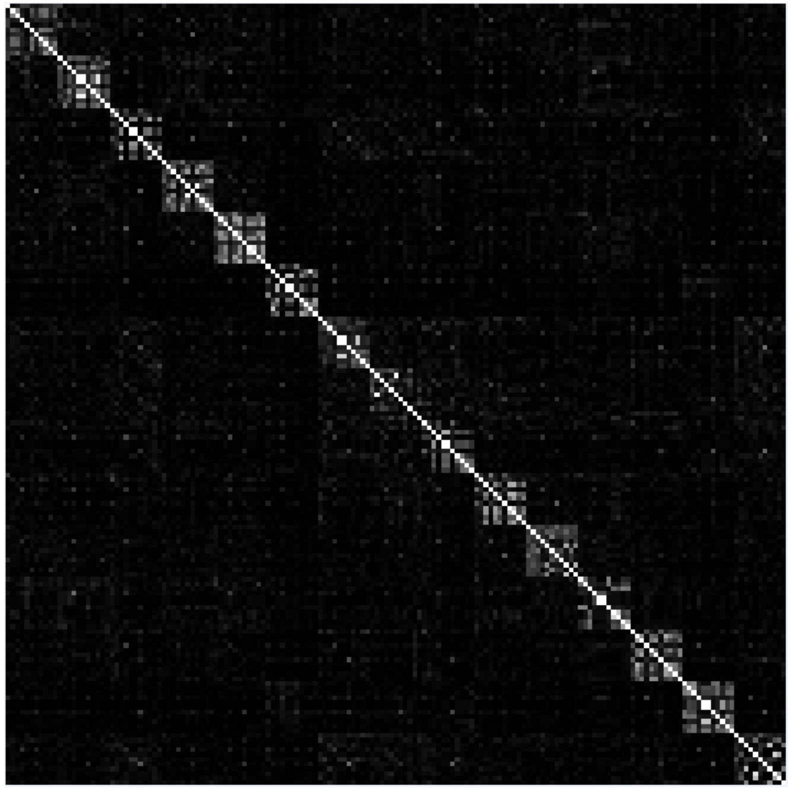
$$J = \min_{A,c} \left\|(X - XA)^T\right\|_{2,1} + \gamma\left\|\alpha I \circ AA^T - C\right\|_F^2 + \lambda\|c\|_1 + \beta Tr(A)$$

$$\gamma = 0.5 \quad \lambda = 100 \quad \beta = 0$$

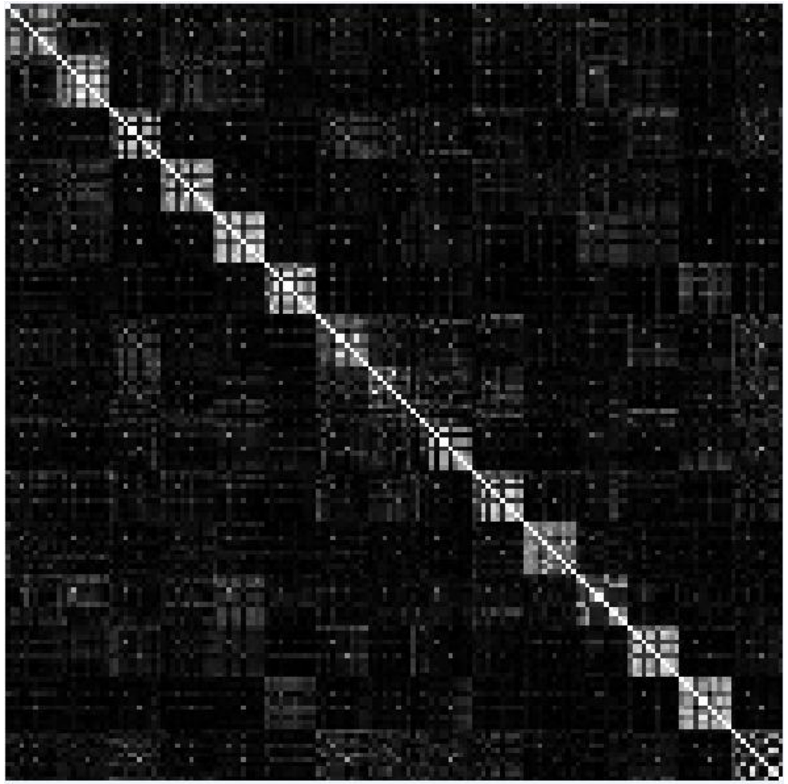
$\alpha = 10$



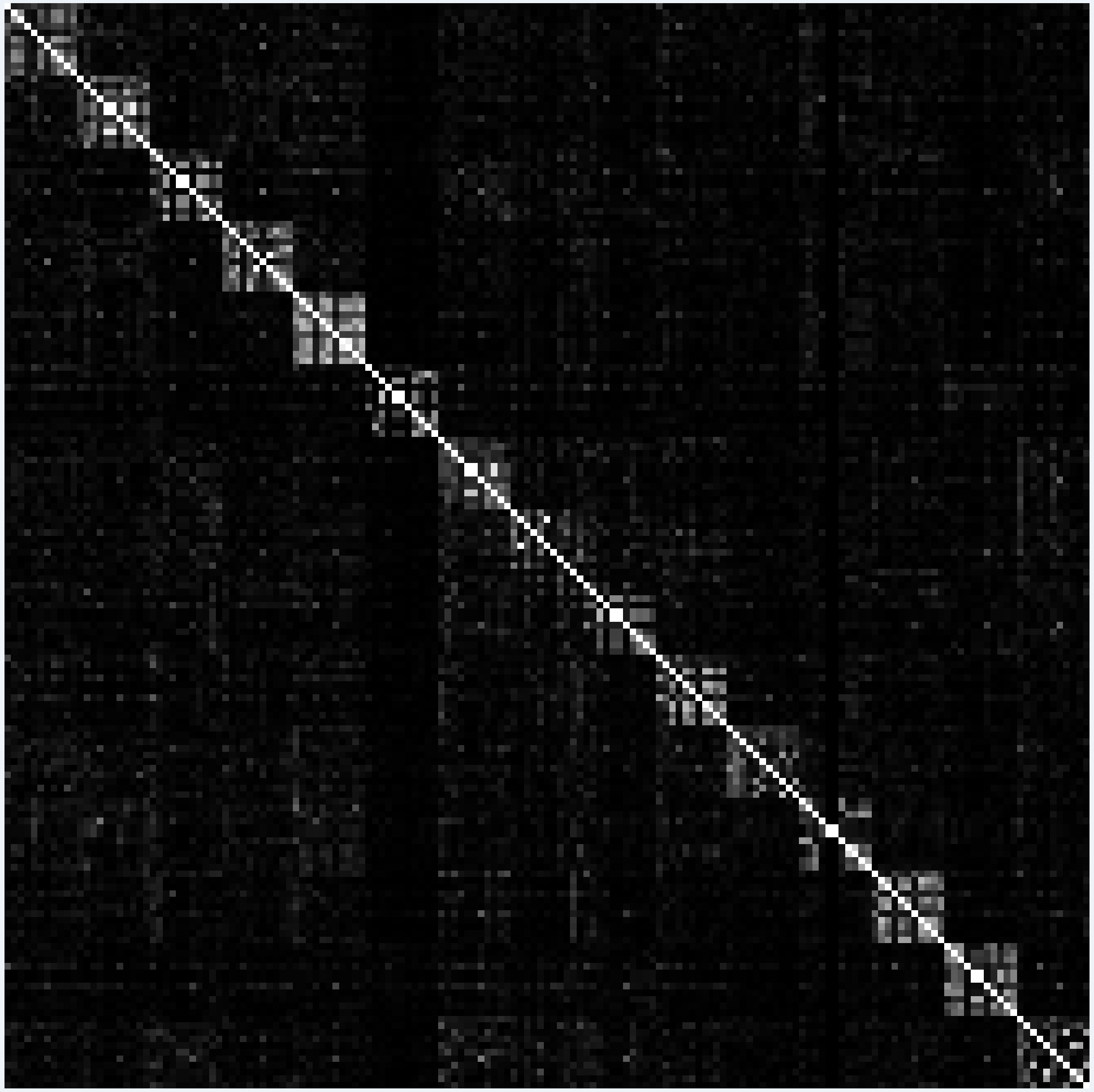
$\alpha = 100$



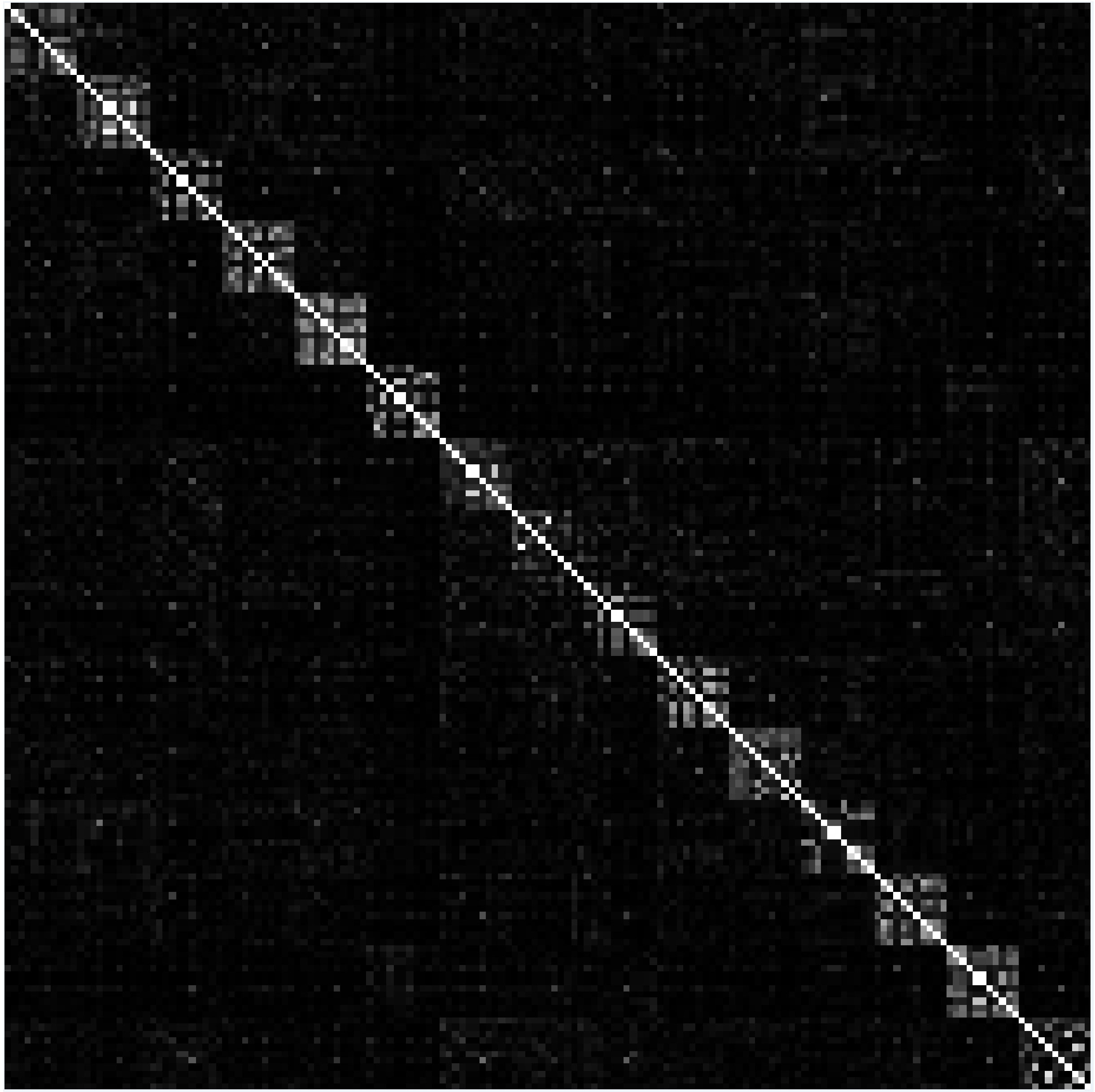
$\alpha = 1000$



IJCAI 2013



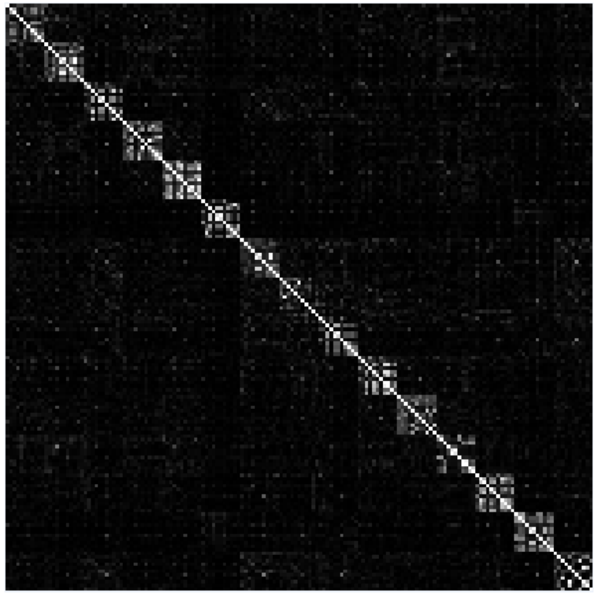
$\alpha = 100$



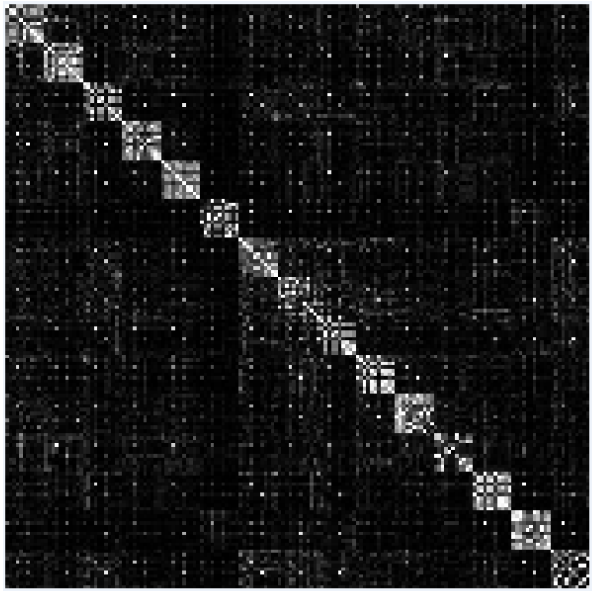
$$\boldsymbol{J} = \min_{\boldsymbol{A},\boldsymbol{c}} \ \left\|(\boldsymbol{X}-\boldsymbol{X}\boldsymbol{A})^{\boldsymbol{T}}\right\|_{2,1} + \boldsymbol{\gamma}\left\|\boldsymbol{\alpha}\boldsymbol{I}\circ\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}}-\boldsymbol{C}\right\|_{\boldsymbol{F}}^{2} + \boldsymbol{\lambda}\|\boldsymbol{c}\|_{1} + \boldsymbol{\beta}\boldsymbol{Tr}(\boldsymbol{A})$$

$$\gamma = 0.5 \quad \lambda = 100 \quad \alpha = 100$$

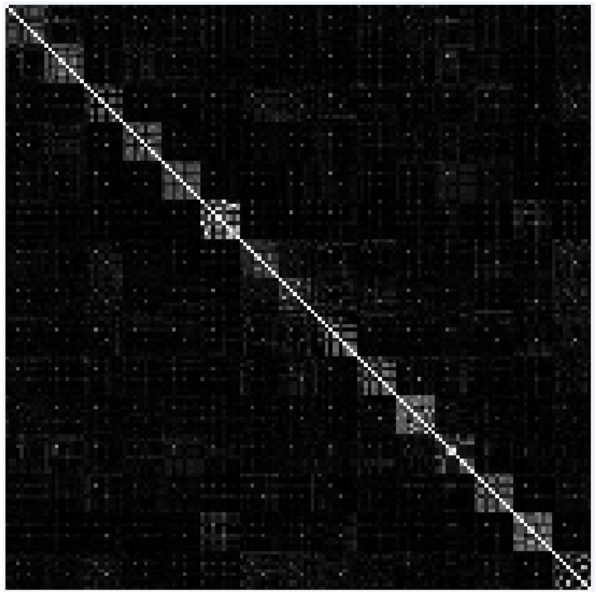
$\beta = 1$



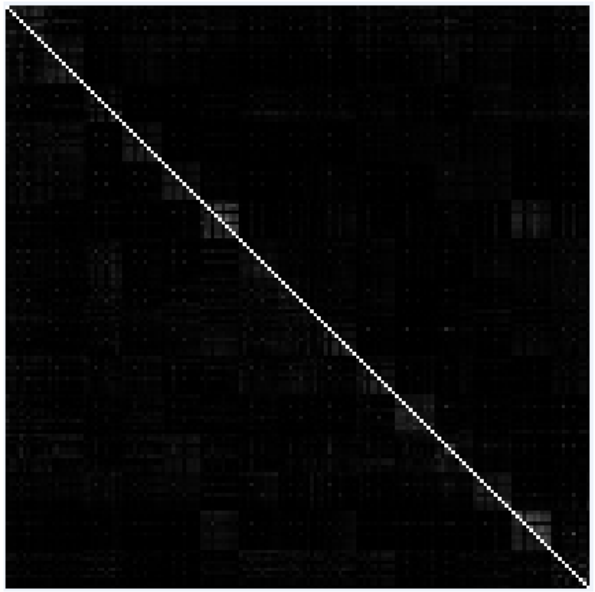
$\beta = 10$



$\beta = 20$



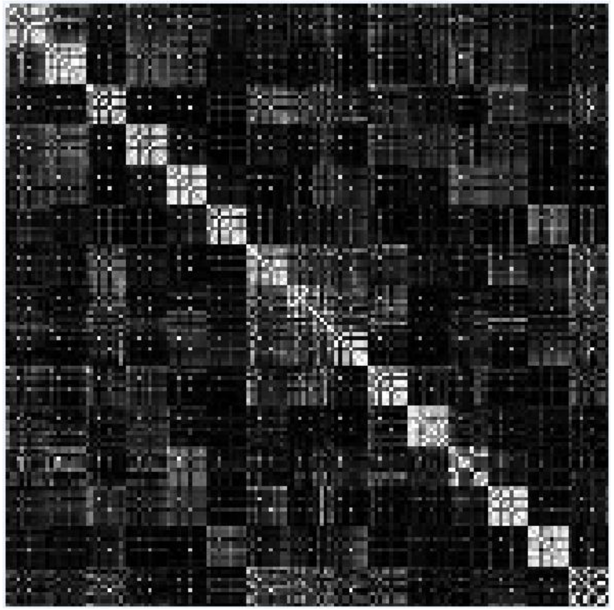
$\beta = 100$



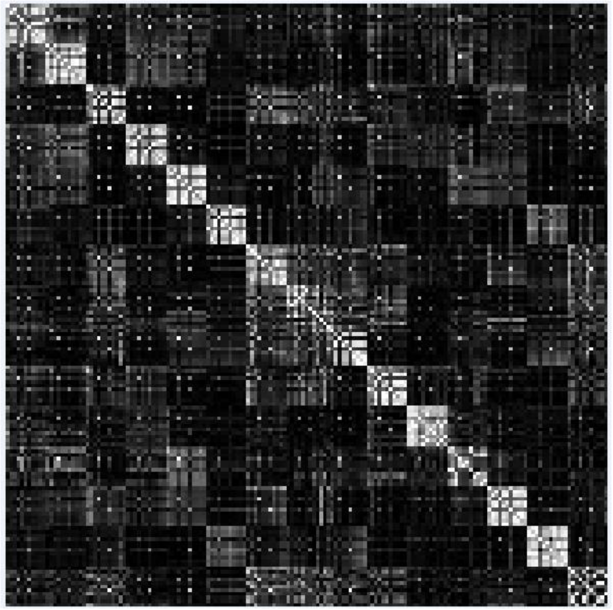
$$J = \min_{A,c} \; \left\|(X - XA)^T\right\|_{2,1} + \gamma\left\|\alpha I \circ AA^T - C\right\|_F^2 + \lambda\|c\|_1 + \beta Tr(A)$$

$$\alpha = 100$$

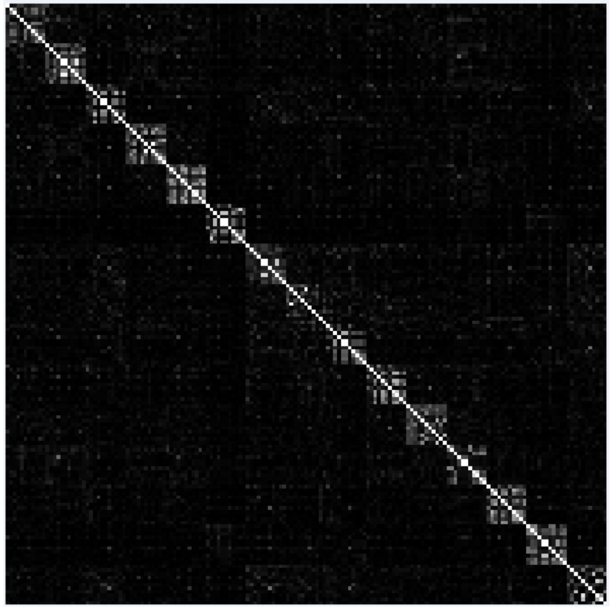
$\gamma = 20\ \lambda = 100\ \beta = 4$



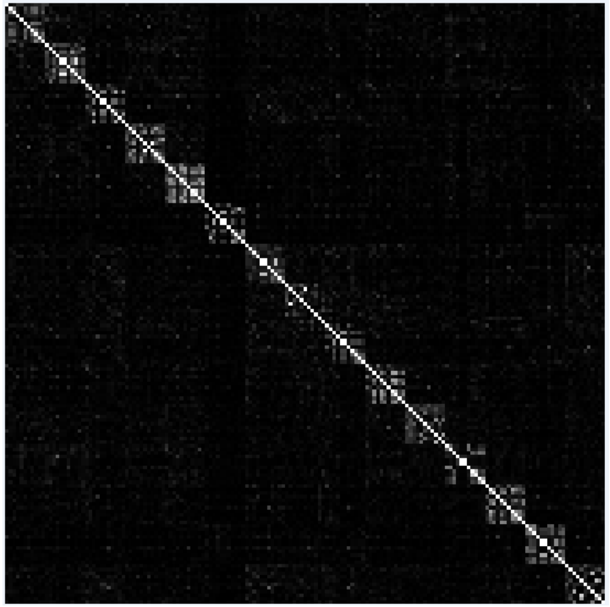
$\gamma = 20\ \lambda = 1000\ \beta = 4$



$\gamma = 0.5\ \ \lambda = 100(1000)\ \beta = 0$

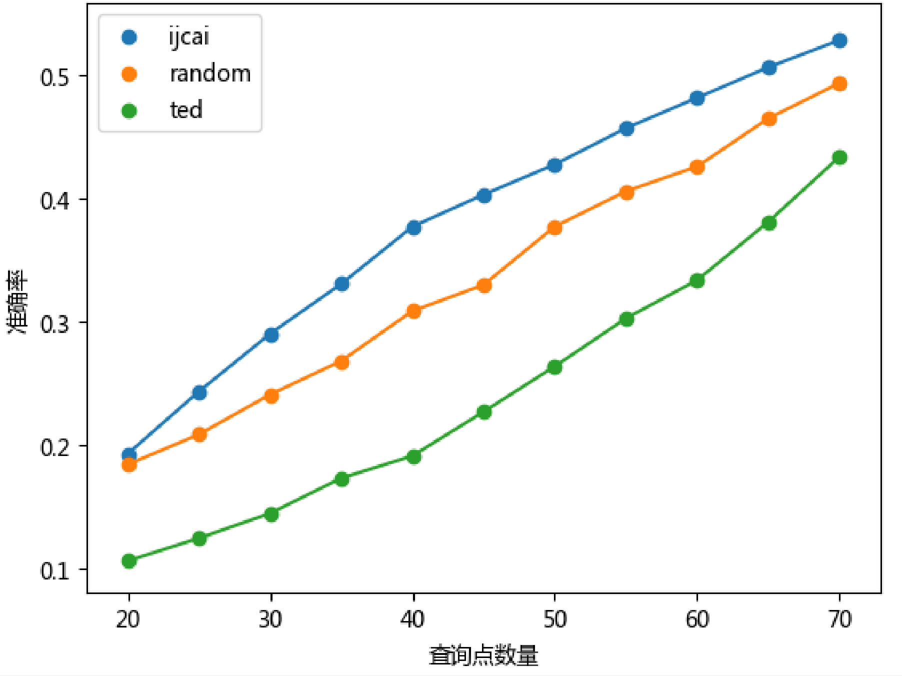


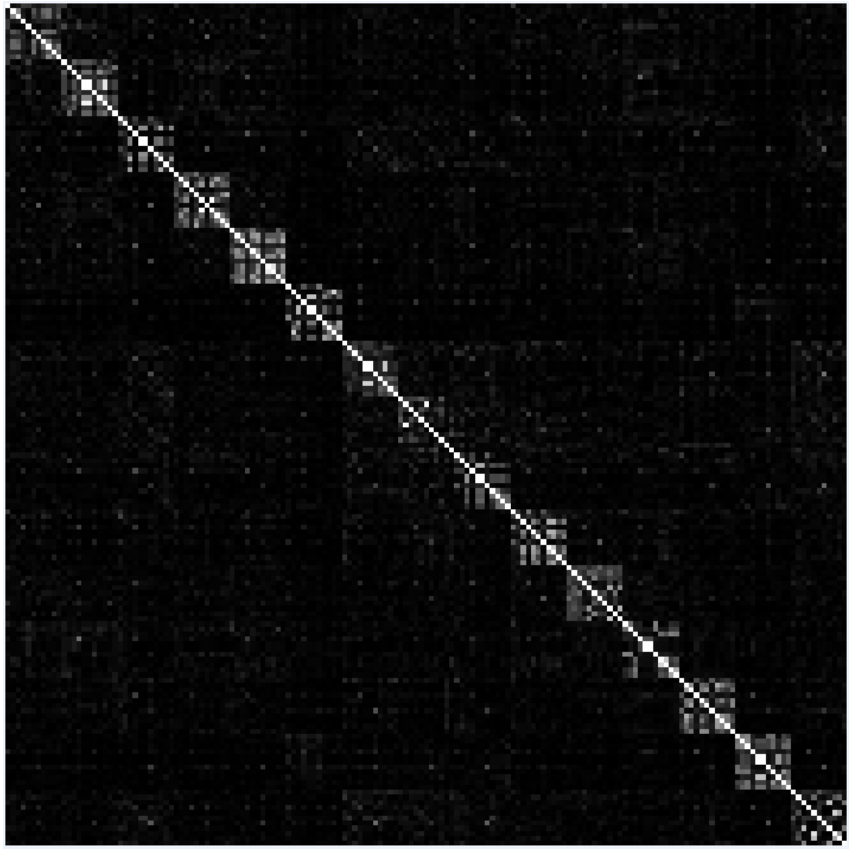
$\gamma = 0.5\ \ \lambda = 10\ \beta = 0$



$$J=\min_{A,c}\ \left\|(X-XA)^T\right\|_{2,1}+\gamma\left\|\alpha I\circ AA^T-C\right\|_F^2+\lambda\|c\|_1+\beta Tr(A)$$

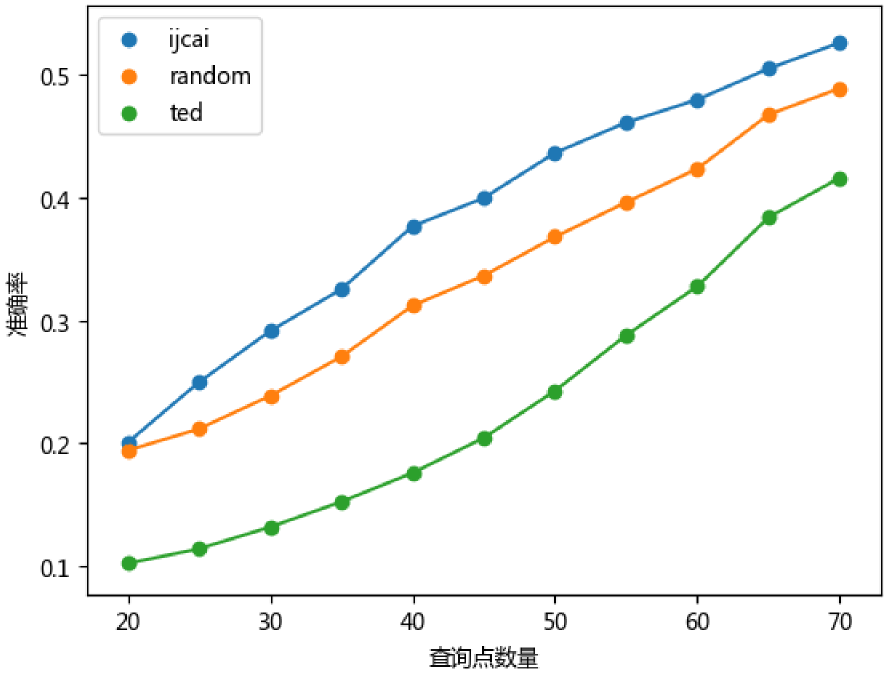
$$\gamma=0.5\quad \lambda=100\quad \alpha=100\quad \beta=0$$

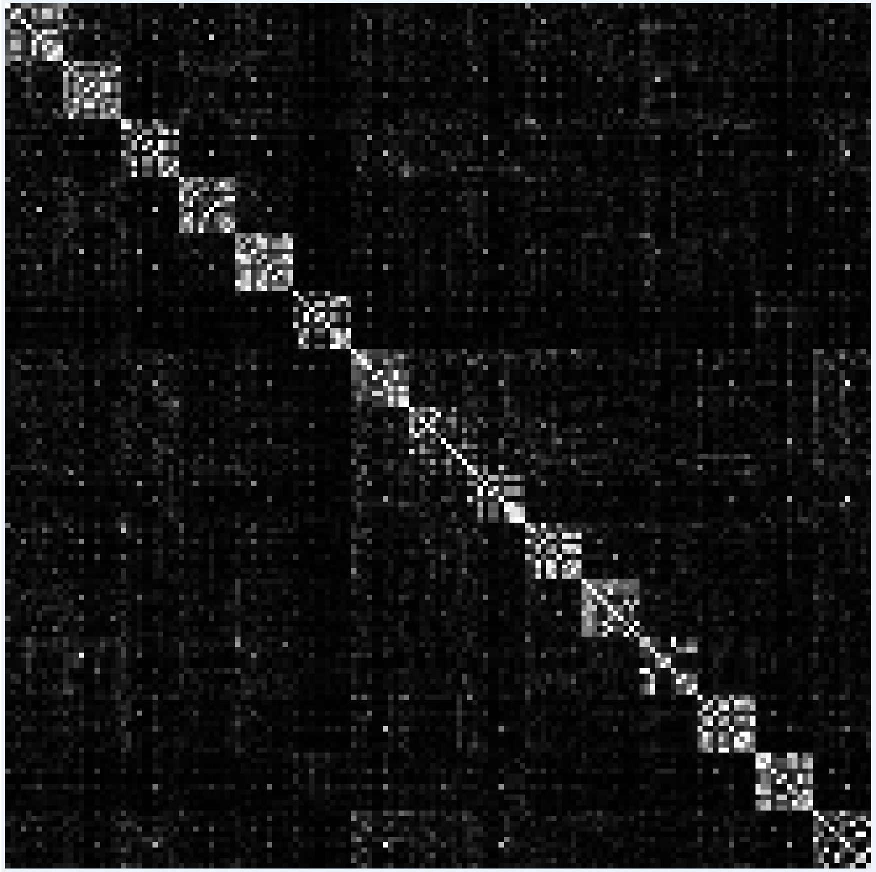




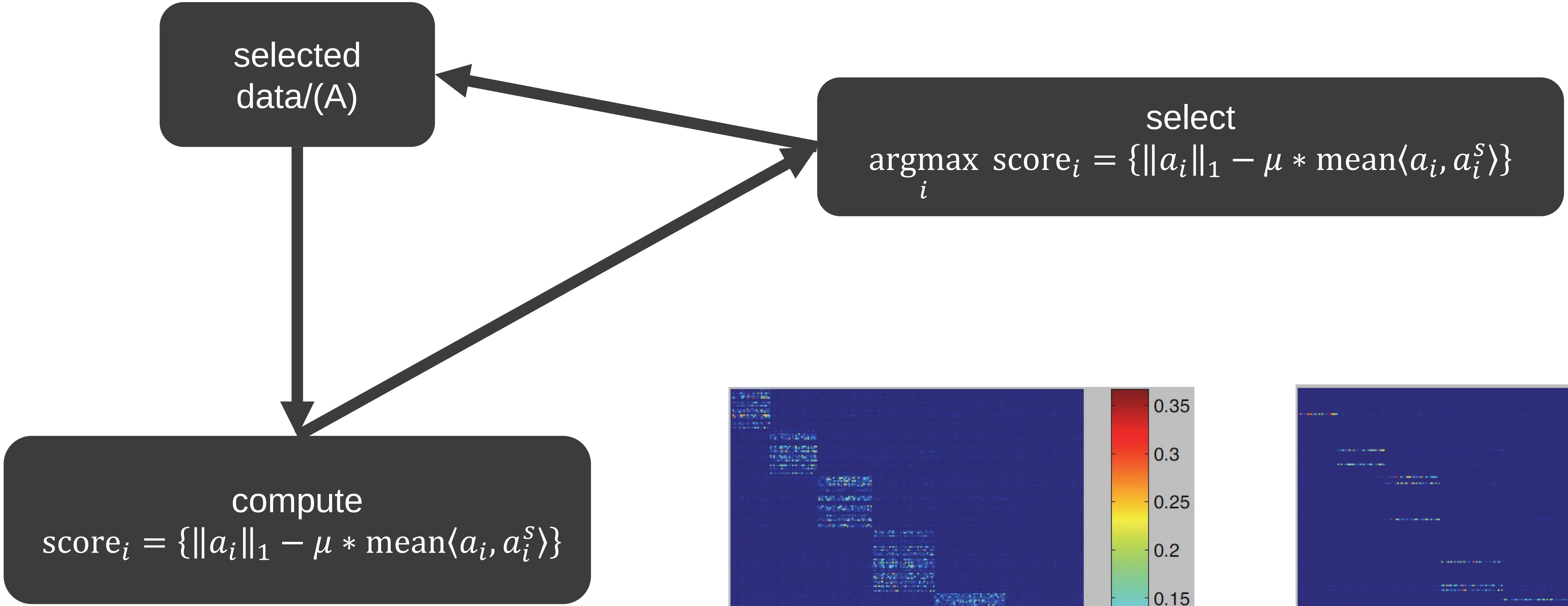
$$J=\min_{A,c}\ \left\|(X-XA)^T\right\|_{2,1}+\gamma\left\|\alpha I\circ AA^T-C\right\|_F^2+\lambda\|c\|_1+\beta Tr(A)$$

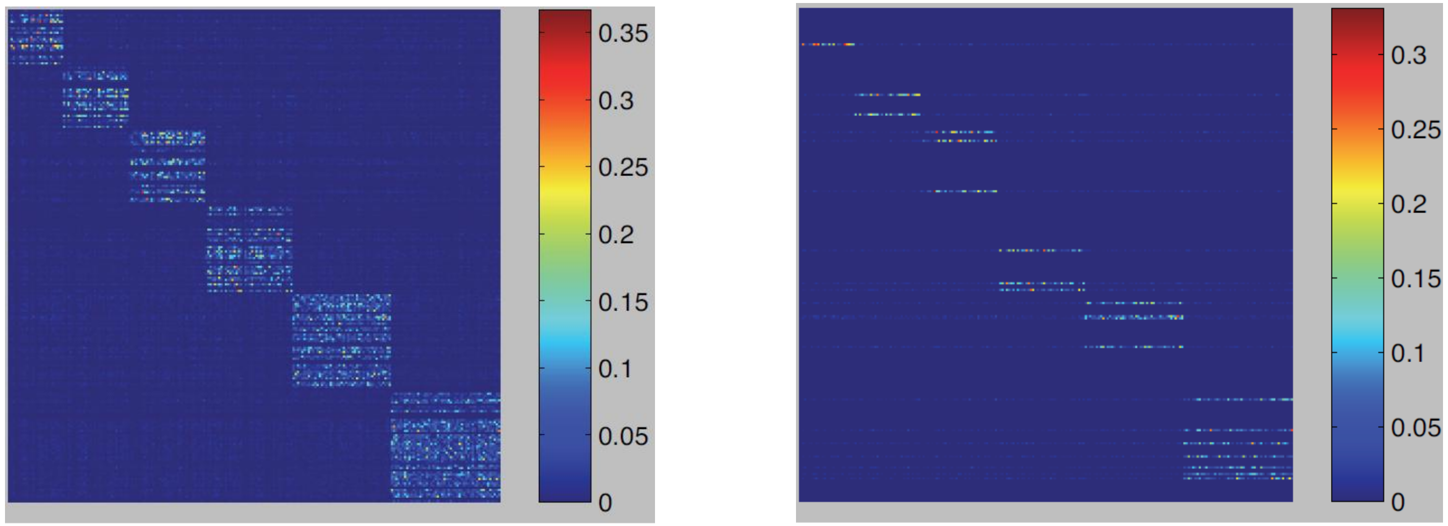
$$\gamma=0.5\quad \lambda=100\quad \alpha=100\quad \beta=4$$





a_i^s:已经选中的数据点对应的A的所在行

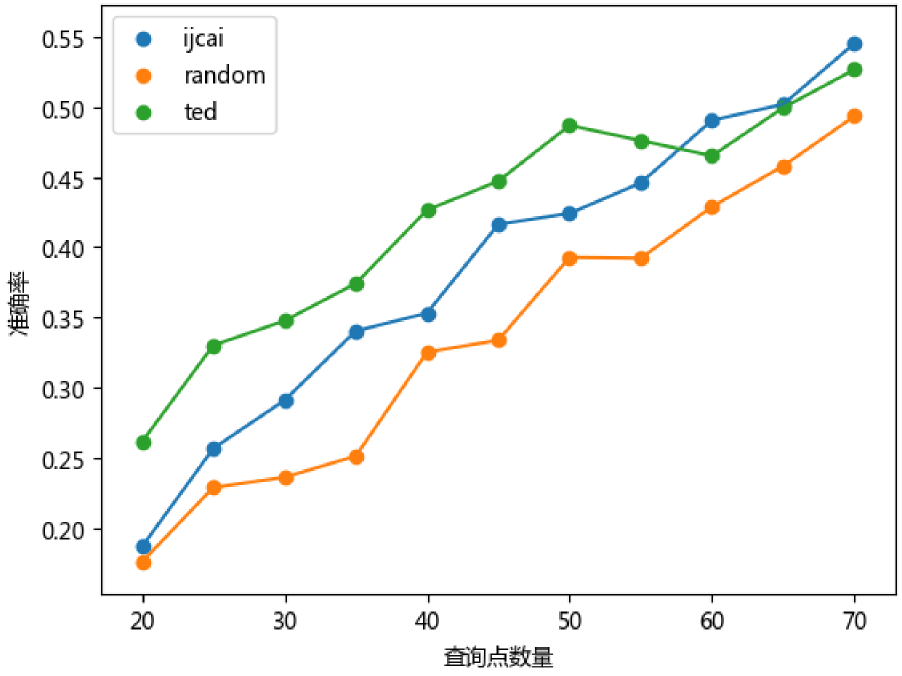


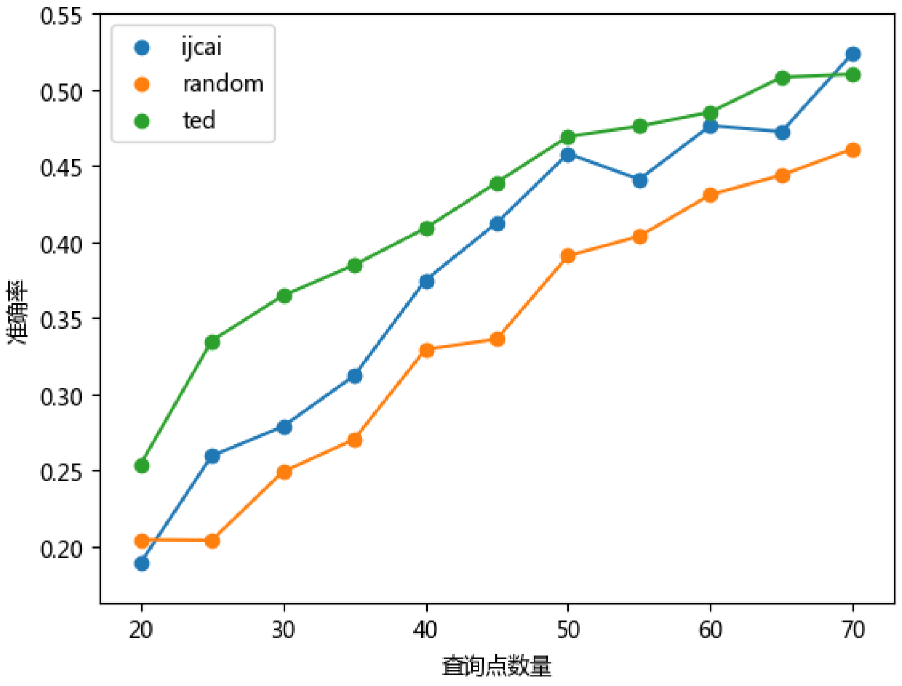


(b) Learned matrix A (c) The top 20 rows of A

$$J = \min_{A,c} \left\| (X - XA)^T \right\|_{2,1} + \gamma \left\| \alpha I \circ AA^T - C \right\|_F^2 + \lambda \|c\|_1 + \beta Tr(A)$$

yaleface 11*15=155 traindata: 77 testdata: 78 $\gamma = 0.5$ $\lambda = 100$ $\alpha = 100$ $\beta = 4$





$$J=\min_{A,c}\ \left\|(X-XA)^T\right\|_{2,1}+\gamma\left\|\alpha I\odot AA^T-C\right\|_F^2+\lambda\|c\|_1+\beta Tr(A)$$

$$\gamma=0.5\quad\lambda=100\quad\alpha=100\ \beta=4$$