低秩稀疏表示算法求解

1. 鲁棒性主成分分析（Robust Principal Component Analysis, RPCA）

$min_{L,E}\, rank(L)+\lambda\|E\|_0 \quad s.t. X=L+E$ (RPCA)

$min_{L,E}\|L\|_*+\lambda\|E\|_1 \quad s.t. X=L+E$ (RPCA 最终求解模型)

假设背景存在单一的子空间，没有考虑到高光谱影像中复杂的背景地物信息

2. 低秩表示算法（Low Rank Representation）

$min_Z\, rank(Z) \quad s.t. X=DZ$ (LRR 初始模型)

$min_{Z,E}\|Z\|_*+\lambda\|E\|_{2,1} \quad s.t. X=DZ+E$ （LRR 最终求解模型）

3. 低秩和稀疏表示（Low Rank and Sparse Representation）

$min_{Z,E}\|Z\|_*+\|Z\|_1+\lambda\|E\|_{2,1} \quad s.t. X=DZ+E$ （LRASR 最终求解模型）

4. 字典学习的方法

上面方法采用整个矩阵作为过完备字典，实际运算计算代价较大，因此采用了学习字典方法，用少量的原子构成字典

$x=D\alpha+v$ (系数α是稀疏的)

$\hat{\alpha}=argmin\,\|x-D\alpha\|_2+\gamma\|\alpha\|_1$ (采用稀疏编码方法来求稀疏系数)

$D^{(n+1)}=D^n-\mu\sum_{i=1}^{M}(D^n\alpha_i-x_i)\alpha_i^T$ (字典学习过程)

5. 单/多局部窗口低秩表示模型（SLW_LRRSTO/MLW_LRRSTO）

$$min_{Z,E}\|Z\|_*+\eta\left\|Z-\tilde{Z}\right\|_F^2+\lambda\|E\|_{2,1} \quad s.t. X=DZ+E,\ \mathbf{1}_d^T\cdot Z=\ \mathbf{1}_n^T$$

矩阵的各种范数含义：

$\|\cdot\|_*$ 核范数，矩阵的奇异值之和，用于$rank(\cdot)$的松弛求解

$\|\cdot\|_0$ 0 范数，矩阵中非零元素的个数，通常松弛到$\|\cdot\|_1$之后进行求解

$\|\cdot\|_1$ 1 范数，矩阵元素的绝对值之和（稀疏规则算子，L1 范数是 L0 范数的最优凸近似，而且它比 L0 范数要容易优化求解）

$\|E\|_{2,1}$ $\ell_{1,2}$范数，矩阵中各列的 2 范数之和

$\|\cdot\|_F$　　F 范数，矩阵元素的绝对值的平方和，再开方

参数λ称为正则化参数，用于控制稀疏解的稀疏度，λ取值越大，解α就越稀疏。当正则化参数λ足够大时，解α为零向量；随着λ的逐渐减小，解向量α的稀疏度也逐渐减小；当λ逐渐减小到 0 时，解向量α便变成了$\|y-X_s\alpha\|_2^2$最小化的向量。

Part I RPCA 求解的具体过程：

$$min_{Z,E}\|A\|_*+\lambda\|E\|_1 \quad s.t. X=A+E$$

利用增广拉格朗日法求解：

$$L(A,E,Y,\mu)=\|A\|_*+\lambda\|E\|_1+\boxed{\langle Y,X-A-E\rangle+\frac{\mu}{2}\|X-A-E\|_F^2}$$

$$=\|A\|_*+\lambda\|E\|_1+\boxed{\frac{\mu}{2}\left\|X-A-E+\frac{Y}{\mu}\right\|_F^2-\frac{1}{2\mu}\|Y\|_F^2}$$

利用配方法化简：

$$\langle Y,X-A-E\rangle+\frac{\mu}{2}\|X-A-E\|_F^2$$

$$=\frac{\mu}{2}\left[\|X-A-E\|_F^2+\frac{2}{\mu}\langle Y,X-A-E\rangle\right]$$

$$=\frac{\mu}{2}\left[\|X-A-E\|_F^2+2\langle\frac{Y}{\mu},X-A-E\rangle\right]$$

$$=\frac{\mu}{2}\left[\|X-A-E\|_F^2+2\langle\frac{Y}{\mu},X-A-E\rangle+\left\|\frac{Y}{\mu}\right\|_F^2-\left\|\frac{Y}{\mu}\right\|_F^2\right]$$

$$=\frac{\mu}{2}\left[\left\|X-A-E+\frac{Y}{\mu}\right\|_F^2-\left\|\frac{Y}{\mu}\right\|_F^2\right]$$

$$=\frac{\mu}{2}\left\|X-A-E+\frac{Y}{\mu}\right\|_F^2-\frac{\mu}{2}\left\|\frac{Y}{\mu}\right\|_F^2$$

$$=\boxed{\frac{\mu}{2}\left\|X-A-E+\frac{Y}{\mu}\right\|_F^2-\frac{1}{2\mu}\|Y\|_F^2}$$

分别更新 A、E、Y、μ：

1)　固定 E，Y 更新 A：

$$(U,S,V)=svd(X-E_k+\mu_k^{-1}Y_k)$$
$$A_{k+1}=US_{\mu_k^{-1}}[S]V^T$$

2)　固定 A，Y 更新 E：

$$\mathrm{E}_{k+1}=S_{\lambda\mu_k^{-1}}[X-A_{k+1}+\mu_k^{-1}Y_k]$$

3) 固定 A，E 更新 Y：

$$Y_{k+1} = Y_k + \mu_k(X - A_{k+1} - E_{k+1})$$

4) 更新μ

$$\mu_{k+1} = \rho\mu_k$$

相关的公式：

① 软阈值收缩算子（用于升级 E 的计算）：

$S_\varepsilon[W] = argmin_X\ \varepsilon\|X\|_1 + \frac{1}{2}\|X - W\|_F^2$　　L1 范数求解式（用于升级 E）

$$S_\varepsilon[x] = \begin{cases} x - \varepsilon, & if\ x\ > \varepsilon \\ x + \varepsilon, & if\ x\ < -\varepsilon \\ 0, & otherwise \end{cases}$$

② 奇异值运算（用于升级 A 的计算）：

$US_\varepsilon[S]V^T = argmin_X\ \varepsilon\|X\|_* + \frac{1}{2}\|X - W\|_F^2$　　核范数求解式（用于升级 A）

具体的升级计算过程：

1. 固定 E，Y 更新 A

$L(A, E, Y, \mu) = \|A\|_* + \lambda\|E\|_1 + \frac{\mu}{2}\left\|X - A - E + \frac{Y}{\mu}\right\|_F^2 - \frac{1}{2\mu}\|Y\|_F^2$　　（化简之后的式子）

去掉 E、Y 相关项，只保留与 A 相关的项。

$$L(A, E, Y, \mu) \ \rightarrow \quad \|A\|_* + \frac{\mu}{2}\left\|X - A - E + \frac{Y}{\mu}\right\|_F^2$$

$$L(A, E, Y, \mu) \ \rightarrow \quad \mu\left(\frac{1}{\mu}\|A\|_* + \frac{1}{2}\left\|X - A - E + \frac{Y}{\mu}\right\|_F^2\right)$$

去掉μ，后面提出来一个负号：

$$L(A, E, Y, \mu) \ \rightarrow \quad \frac{1}{\mu}\|A\|_* + \frac{1}{2}\left\|A - \left(X - E + \frac{Y}{\mu}\right)\right\|_F^2$$

形势化简完成，采用奇异值运算（核范数），其中$\varepsilon = \frac{1}{\mu}$，$W = X - E + \frac{Y}{\mu}$

更新式如下：

$$(U, S, V) = svd(X - E_k + \mu_k^{-1}Y_k)$$
$$A_{k+1} = US_{\mu_k^{-1}}[S]V^T$$

2. 固定 A，Y 更新 E

$L(A,E,Y,\mu)=\|A\|_*+\lambda\|E\|_1+\frac{\mu}{2}\left\|X-A-E+\frac{Y}{\mu}\right\|_F^2-\frac{1}{2\mu}\|Y\|_F^2$ （化简之后的式子）

去掉 A、Y 相关项，只保留与 E 相关的项。

$$L(A,E,Y,\mu)\ \rightarrow\quad \lambda\|E\|_1+\frac{\mu}{2}\left\|X-A-E+\frac{Y}{\mu}\right\|_F^2$$

$$L(A,E,Y,\mu)\ \rightarrow\quad \mu\left(\frac{\lambda}{\mu}\|E\|_1+\frac{1}{2}\left\|X-A-E+\frac{Y}{\mu}\right\|_F^2\right)$$

去掉μ，后面提出来一个负号：

$$L(A,E,Y,\mu)\ \rightarrow\quad \frac{\lambda}{\mu}\|E\|_1+\frac{1}{2}\left\|E-\left(X-A+\frac{Y}{\mu}\right)\right\|_F^2$$

形势化简完成，采用软阈值收缩算子运算（L1 范数），其中$\varepsilon=\frac{\lambda}{\mu}$， $\mathrm{W}=X-A+\frac{Y}{\mu}$

更新式如下：

$$E_{k+1}=S_{\lambda\mu_k^{-1}}[X-A_{k+1}+\mu_k^{-1}Y_k]$$

3. 固定 A，E 更新 Y

原始式子如下：

$$L(A,E,Y,\mu)=\|A\|_*+\lambda\|E\|_1+\langle Y,X-A-E\rangle+\frac{\mu}{2}\|X-A-E\|_F^2$$

与 Y 相关的式子：

$$L(A,E,Y,\mu)\ \rightarrow\qquad \langle Y,X-A-E\rangle$$

其中 Y 为系数，对其求导，梯度为 $X-A-E$

更新式如下：

$$Y_{k+1}=Y_k+\mu_k(X-A_{k+1}-E_{k+1})$$

Part II 低秩表示（LRR）算法求解：

$$min_{Z,E}\|Z\|_*+\lambda\|E\|_{2,1}\quad s.t.X=DZ+E$$ （LRR 公式）

采用增广拉格朗日法进行求解，不同于 RPCA 算法求解，求解过程中需要引入辅助变量J，使得目标函数具有可分性，求解公式等价于：

$$min_{Z,E}\|J\|_*+\lambda\|E\|_{2,1}\quad s.t.X=DZ+E,\qquad Z=J$$

构造增广拉格朗日函数：

$$L(J,Z,E,Y_1,Y_2,\mu)=\|J\|_*+\lambda\|E\|_{2,1}+\langle Y_1,X-DZ-E\rangle+\langle Y_2,Z-J\rangle$$
$$+\frac{\mu}{2}\|X-DZ-E\|_F^2+\frac{\mu}{2}\|Z-J\|_F^2$$

对上式的两个平方项通过配方法分别进行化简:

$$L(J,Z,E,Y_1,Y_2,\mu)=\|J\|_*+\lambda\|E\|_{2,1}+\langle Y_1,X-DZ-E\rangle+\langle Y_2,Z-J\rangle$$
$$+\frac{\mu}{2}\|X-DZ-E\|_F^2+\frac{\mu}{2}\|Z-J\|_F^2$$
$$=\|J\|_*+\lambda\|E\|_{2,1}+\boxed{\frac{\mu}{2}\|X-DZ-E\|_F^2+\langle Y_1,X-DZ-E\rangle}$$
$$\boxed{+\frac{\mu}{2}\|Z-J\|_F^2+\langle Y_2,Z-J\rangle}$$
$$=\|J\|_*+\lambda\|E\|_{2,1}+\frac{\mu}{2}\left\|X-DZ-E+\frac{Y_1}{\mu}\right\|_F^2+\frac{\mu}{2}\left\|Z-J+\frac{Y_2}{\mu}\right\|_F^2$$
$$-\frac{1}{2\mu}(\|Y_1\|_F^2+\|Y_2\|_F^2)$$

具体化简过程如下:

为了方便，下面分别拿出来对两个平方向进行配方法化简:

对于第一项进行化简，具体过程如下:

$$\frac{\mu}{2}\|X-DZ-E\|_F^2+\langle Y_1,X-DZ-E\rangle$$
$$=\frac{\mu}{2}\left(\|X-DZ-E\|_F^2+\frac{2}{\mu}\langle Y_1,X-DZ-E\rangle\right)$$
$$=\frac{\mu}{2}\left(\|X-DZ-E\|_F^2+2\left\langle\frac{Y_1}{\mu},X-DZ-E\right\rangle\right)$$
$$=\frac{\mu}{2}\left(\|X-DZ-E\|_F^2+2\left\langle\frac{Y_1}{\mu},X-DZ-E\right\rangle+\left\|\frac{Y_1}{\mu}\right\|_F^2-\left\|\frac{Y_1}{\mu}\right\|_F^2\right)$$
$$=\frac{\mu}{2}\left(\left\|X-DZ-E+\frac{Y_1}{\mu}\right\|_F^2-\left\|\frac{Y_1}{\mu}\right\|_F^2\right)$$
$$=\frac{\mu}{2}\left\|X-DZ-E+\frac{Y_1}{\mu}\right\|_F^2-\frac{\mu}{2}\left\|\frac{Y_1}{\mu}\right\|_F^2$$
$$=\frac{\mu}{2}\left\|X-DZ-E+\frac{Y_1}{\mu}\right\|_F^2-\frac{1}{2\mu}\|Y_1\|_F^2$$

对于第二项进行化简，具体过程如下:

$$
\begin{aligned}
&\frac{\mu}{2}\|Z-J\|_F^2+\langle Y_2,Z-J\rangle \\
&=\frac{\mu}{2}\left(\|Z-J\|_F^2+\frac{2}{\mu}\langle Y_2,Z-J\rangle\right) \\
&=\frac{\mu}{2}\left(\|Z-J\|_F^2+2\left\langle\frac{Y_2}{\mu},Z-J\right\rangle+\left\|\frac{Y_2}{\mu}\right\|_F^2-\left\|\frac{Y_2}{\mu}\right\|_F^2\right) \\
&=\frac{\mu}{2}\left(\left\|Z-J+\frac{Y_2}{\mu}\right\|_F^2-\left\|\frac{Y_2}{\mu}\right\|_F^2\right) \\
&=\frac{\mu}{2}\left\|Z-J+\frac{Y_2}{\mu}\right\|_F^2-\frac{\mu}{2}\left\|\frac{Y_2}{\mu}\right\|_F^2 \\
&=\frac{\mu}{2}\left\|Z-J+\frac{Y_2}{\mu}\right\|_F^2-\frac{1}{2\mu}\|Y_2\|_F^2
\end{aligned}
$$

具体的升级计算过程：

化简之后的 LRR 式子为：

$$
L(J,Z,E,Y_1,Y_2,\mu)=\|J\|_*+\lambda\|E\|_{2,1}+\frac{\mu}{2}\left\|X-DZ-E+\frac{Y_1}{\mu}\right\|_F^2+\frac{\mu}{2}\left\|Z-J+\frac{Y_2}{\mu}\right\|_F^2
$$

$$
-\frac{1}{2\mu}(\|Y_1\|_F^2+\|Y_1\|_F^2)
$$

1. **固定 Z，E，Y_1，Y_2 更新 J**

式子中去掉其他项，只保存与 J 有关的项，则有

$$
L(\boldsymbol{J},Z,E,Y_1,Y_2,\mu)\ \rightarrow\|J\|_*+\frac{\mu}{2}\left\|Z-J+\frac{Y_2}{\mu}\right\|_F^2
$$

对上述式子进行化简：

$$
L(\boldsymbol{J},Z,E,Y_1,Y_2,\mu)\ \rightarrow\mu\left(\frac{1}{\mu}\|J\|_*+\frac{1}{2}\left\|Z-J+\frac{Y_2}{\mu}\right\|_F^2\right)
$$

去掉μ，然后对后面提出一个负号，进行化简后为：

$$
L(\boldsymbol{J},Z,E,Y_1,Y_2,\mu)\ \rightarrow\frac{1}{\mu}\|J\|_*+\frac{1}{2}\left\|J-\left(Z+\frac{Y_2}{\mu}\right)\right\|_F^2
$$

根据奇异值运算形式，可得　$\varepsilon=\mu^{-1}$，　$W=Z+\frac{Y_2}{\mu}$

$$
\boxed{\mathrm{U}S_\varepsilon[S]V^T=argmin_X\ \varepsilon\|X\|_*+\frac{1}{2}\|X-W\|_F^2}\qquad\text{(奇异值运算)}
$$

因此，**最后更新式子为**:

$$(U,S,V)=svd\left(Z_k+\frac{Y_2}{\mu}\right)$$
$$J_{k+1}=US_{\mu_k^{-1}}[S]V^T$$

2. 固定 J，E，Y₁，Y₂ 更新 Z

去掉其他项，保留与 Z 有关的项

$$L(J,\boldsymbol{Z},E,Y_1,Y_2,\mu)\ \rightarrow\ \frac{\mu}{2}\left\|X-DZ-E+\frac{Y_1}{\mu}\right\|_F^2+\frac{\mu}{2}\left\|Z-J+\frac{Y_2}{\mu}\right\|_F^2$$

上述式子因为是 F 范数的平方，所以不能使用软阈值收缩、奇异值运算等算子进行运算，此处对该式子的 Z 进行求导:

$$\frac{\partial L(J,\boldsymbol{Z},E,Y_1,Y_2,\mu)}{\partial \boldsymbol{Z}}=-\mu D^T\left(X-DZ-E+\frac{Y_1}{\mu}\right)+\mu\left(Z-J+\frac{Y_2}{\mu}\right)$$
$$=-\mu\left(D^TX-D^TDZ-D^TE+\frac{D^TY_1}{\mu}-Z+J-\frac{Y_2}{\mu}\right)$$
$$=-\mu\left(D^TX-D^TE-(D^TD+I)Z+J+\mu^{-1}(D^TY_1-Y_2)\right)$$

令式子等于 0，可得

$$D^TX-D^TE-(D^TD+I)Z+J+\mu^{-1}(D^TY_1-Y_2)=0$$

即，

$$(D^TD+I)Z=D^TX-D^TE+J+\mu^{-1}(D^TY_1-Y_2)$$

因此，可得最后的关于 Z 的更新式子为:

$$\mathrm{Z}_{k+1}=(D^TD+I)^{-1}\left(D^TX-D^TE_k+J_{k+1}+{\mu_k}^{-1}\left(D^TY_{1_k}-Y_{2_k}\right)\right)$$

3. 固定 J，Z，Y₁，Y₂ 更新 E

去掉其他项，保留与 E 有关的项

$$L(J,Z,\boldsymbol{E},Y_1,Y_2,\mu)\ \rightarrow\ \lambda\|E\|_{2,1}+\frac{\mu}{2}\left\|X-DZ-E+\frac{Y_1}{\mu}\right\|_F^2$$

对以上式子进行格式化简

$$L(J,Z,\boldsymbol{E},Y_1,Y_2,\mu)\ \rightarrow\mu\left(\frac{\lambda}{\mu}\|E\|_{2,1}+\frac{1}{2}\left\|X-DZ-E+\frac{Y_1}{\mu}\right\|_F^2\right)$$

$$L(J,Z,\boldsymbol{E},Y_1,Y_2,\mu)\ \rightarrow\mu\left(\frac{\lambda}{\mu}\|E\|_{2,1}+\frac{1}{2}\left\|E-\left(X-DZ+\frac{Y_1}{\mu}\right)\right\|_F^2\right)$$

去掉括号最外面的μ，则有如下形势

$$E = argmin_E \ \frac{\lambda}{\mu}\|E\|_{2,1} + \frac{1}{2}\left\|E - \left(X - DZ + \frac{Y_1}{\mu}\right)\right\|_F^2$$

该公式同软阈值收缩算子形式，其中$\varepsilon = \lambda/\mu$，$W = X - DZ + \frac{Y_1}{\mu}$，故升级 E 的公式为：

$$E_{k+1} = S_{\lambda/\mu}\left[X - DZ + \frac{Y_1}{\mu}\right]$$

4. **固定 J，Z，E 更新 Y_1，Y_2**

同样，去掉其他项，分别保留与 Y_1，Y_2有关的项：

$$L(J, Z, E, \boldsymbol{Y_1}, Y_2, \mu) \rightarrow \ \frac{\mu}{2}\left\|X - DZ - E + \frac{Y_1}{\mu}\right\|_F^2 - \frac{\|Y_1\|_F^2}{2\mu}$$

$$L(J, Z, E, Y_1, \boldsymbol{Y_2}, \mu) \rightarrow \ \frac{\mu}{2}\left\|Z - J + \frac{Y_2}{\mu}\right\|_F^2 - \frac{\|Y_2\|_F^2}{2\mu}$$

两边对 Y_1、Y_2分别求导：

$$\frac{\partial L(J, Z, E, \boldsymbol{Y_1}, Y_2, \mu)}{\partial \boldsymbol{Y_1}} = \ X - DZ - E + \frac{Y_1}{\mu} - \frac{Y_1}{\mu}$$

$$\frac{\partial L(J, Z, E, Y_1, \boldsymbol{Y_2}, \mu)}{\partial \boldsymbol{Y_2}} = \ Z - J + \frac{Y_2}{\mu} - \frac{Y_2}{\mu}$$

此时需要注意，最后的两项不能直接消去，可令其中一个为$Y_{1_{k+1}}$,另一个为Y_{1_k}，同时令求导式等于 0，可得 Y_1的更新式为：

$$Y_{1_{k+1}} = Y_{1_k} + \mu_k(X - DZ_{k+1} - E_{k+1})$$

同理可得 Y_2的更新式为：

$$Y_{2_{k+1}} = Y_{2_k} + \mu_k(Z_{k+1} - J_{k+1})$$

5. **升级μ的式子为**

$$\mu = min(\rho\mu, \mu_{max})$$