机器学习中的优化算法作业4　8.16　8.19 (a)(b)(c)

8.16 先求该问题的 Lagrange 函数：

引入乘子 Λ 作用在约束、$L+S=M$ 上

$$L_\rho(L,S,\Lambda) = \|L\|_* + \lambda\|S\|_1 + \langle \Lambda, L+S-M\rangle + \frac{\rho}{2}\|L+S-M\|_F^2$$

在第 $(k+1)$ 步. 交替方向乘子法分别求解关于 L 和 S 的子问题更新

L^{k+1} 和 S^{k+1}

对于 L 子问题 $L^{k+1} = \arg\min\limits_L L_\rho(L, S^k, \Lambda^k)$

$$= \arg\min_L \left\{ \|L\|_* + \frac{\rho}{2}\| L+S^k-M+\frac{1}{\rho}\Lambda^k\|_F^2 \right\}$$

$$= \arg\min_L \left\{ \frac{1}{\rho}\|L\|_* + \frac{1}{2}\| L+S^k-M+\frac{1}{\rho}\Lambda^k\|_F^2 \right\}$$

$$= U\mathrm{Diag}(\mathrm{prox}_{\frac{1}{\rho}\|\cdot\|_1}(\sigma(A)))V^T$$

其中 $A = M-S^k-\frac{1}{\rho}\Lambda^k$，$\sigma(A)$ 为 A 的所有非奇异值构成的向量

并且 $U\mathrm{Diag}(\sigma(A))V^T$ 是 A 的奇异值分解.

对于 S 子问题 $S^{k+1} = \arg\min\limits_S L_\rho(L^{k+1}, S^k, U^k)$

$$= \arg\min_S \left\{ \lambda\|S\|_1 + \frac{\rho}{2}\| S+L^{k+1}-M+\frac{1}{\rho}\Lambda^k\|_F^2 \right\}$$

$$= \mathrm{prox}_{\frac{\lambda}{\rho}\|\cdot\|_1}(M-L^{k+1}-\frac{1}{\rho}\Lambda^k)$$

此处 $Z=\mathrm{prox}_{\frac{\lambda}{\rho}\|\cdot\|_1}(Y)$ 满足

$$Z_{ij}=\mathrm{sign}(Y_{ij})\max\{|Y_{ij}|-\frac{\lambda}{\rho},0\}$$

对于乘子 Λ，有常规更新

$$\Lambda^{k+1}=\Lambda^k+\tau\rho(L^{k+1}+S^{k+1}-M)$$

因此对于 L 子问题和 S 子问题都有显式解.

8.19 (a)(b)(c) 的算法实现见后文