8.6 f为闭凸函数. 证明 Moreau 分解

$$x = \mathrm{prox}_f(x) + \mathrm{prox}_{f^*}(x)$$

<u>Proof.</u>

设 $u = \mathrm{prox}_f(x)$, $v = x - u$

由 prox 算子的性质, 我们有 $x - u \in \partial f(u)$

因此 $v \in \partial f(u)$

由定义 $\forall y \in \mathbb{R}^n$, $f(y) \geq f(u) + v^T(y-u)$

$$v^T u - f(u) \geq v^T y - f(y)$$

$$\Rightarrow v^T u - f(u) = \sup_{y \in \mathbb{R}^n} (v^T y - f(y)) = f^*(v)$$

而 $x = \mathrm{prox}_f(x) + \mathrm{prox}_{f^*}(x)$

$\Leftrightarrow v = \mathrm{prox}_{f^*}(x)$

$\Leftrightarrow x - v \in \partial f^*(v)$

$\Leftrightarrow u \in \partial f^*(v)$

由定义, 只需验证 $\forall y \in \mathbb{R}^n$, $f^*(y) \geq f^*(v) + u^T(y-v)$

此即 $f^*(y) \geq v^T u - f(u) + u^T(y-v)$

即 $f^*(y) \geq u^T y - f(u)$

但 $f^*(y) = \sup_{x \in \mathbb{R}^n} x^T y - f(x)$. 证毕.

8.16 写出 robust-PCA

$$\min \|L\|_* + \lambda\|S\|_1$$
$$\text{s.t. } L+S=M$$ 的 ADMM 算法及子问题求解方式.

Solution.

写出增广 Lagrangian，对 $\rho>0$

$$L_\rho(L,S,u) = \|L\|_* + \lambda\|S\|_1 + \langle u, L+S-M\rangle + \frac{\rho}{2}\|L+S-M\|_2^2$$

ADMM 算法为

迭代为对 $k=0,1,2,\cdots$

$$L^{(k+1)} = \arg\min_L L_\rho(L, S^{(k)}, u^{(k)}) \quad (1)$$

$$S^{(k+1)} = \arg\min_S L_\rho(L^{(k+1)}, S, u^{(k)}) \quad (2)$$

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

首先我们考虑 $\text{prox}_{\|\cdot\|_1}$ 的表达式

$$\text{prox}_{t\|\cdot\|_1}(X) = \arg\min_U t\|U\|_1 + \frac{1}{2}\|U-X\|_F^2$$

此问题对所有 U_{ij}, $1\le i,j\le n$ 是可分离的

即只需考虑 $\arg\min_{U_{ij}} |U_{ij}| + \frac{1}{2}|U_{ij}-X_{ij}|^2 = \text{sgn}(X_{ij})\max(|X_{ij}|-t, 0)$

故 $\text{prox}_{t\|\cdot\|_1}(X) = \text{sgn}(X) \otimes \max(|X|-t, 0)$

其中 $\text{sgn}(X)$, $\max(\cdot)$ 作用于每一分量，即 $\text{sgn}(X)_{ij} = \text{sgn}(X_{ij})$;

$$\max(|X|-t, 0)_{ij} = \max(|X_{ij}|-t, 0)$$

$\otimes$ 表示逐项乘积，即 $(A\otimes B)_{ij} = A_{ij}B_{ij}$

我们推导 (1)(2) 的显式表达式.

$$L^{(k+1)} = \arg\min_L \|L\|_* + \frac{\rho}{2}\|L+S^{(k)}-M+\frac{\Upsilon^k}{\rho}\|_F^2$$

$$= \arg\min_L \frac{1}{\rho}\|L\|_* + \frac{1}{2}\|L+S^{(k)}-M+\frac{\Upsilon^k}{\rho}\|_F^2$$

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

$$= U \operatorname{Diag}\left(\operatorname{sgn}(\sigma(A)) \otimes \max\left(\sigma(A) - \frac{1}{\rho} \right) \right) V^T \quad (3)$$

其中 $M - S^{(k)} - \frac{Y^{(k)}}{\rho} = U \operatorname{Diag}\left(\sigma\left(M - S^{(k)} - \frac{Y^{(k)}}{\rho} \right) \right) V^T$ 为约化奇异值分解.

而 $S^{(k+1)} = \arg\min L_\rho \left(L^{(k+1)}, S, Y^{k} \right)$

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

$$= \operatorname{Prox}_{(\lambda/\rho)\|\cdot\|_1} \left(M - L^{(k+1)} - \frac{Y}{\rho} \right)$$

$$= \operatorname{sgn}\left(M - L^{(k+1)} - \frac{Y^{(k+1)}}{\rho} \right) \otimes \max\left(M - L^{(k+1)} - \frac{Y^{(k)}}{\rho} - \frac{\lambda}{\rho}, 0 \right) \quad (4)$$

(3), (4) 给出了显式解. #