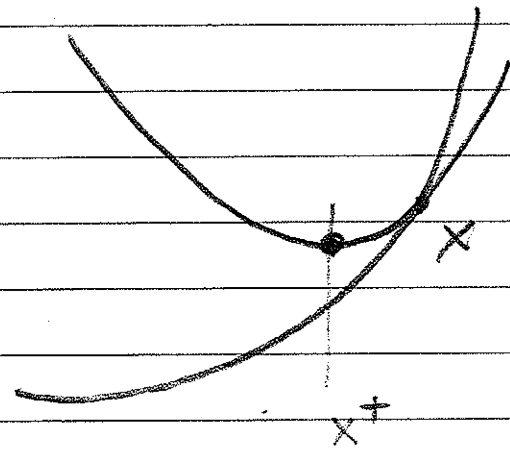


$$\underbrace{\text{prox}_t(x) = \min_z \frac{1}{2t}\|x - z\|^2 + h(z)}_{\lambda\|z\|_1}$$

$$x^+ = \text{prox}_t(x - \underbrace{t\nabla g(x)})$$

$$\downarrow$$

$$x^+ = S_{\lambda t}(\underbrace{x + tA^T(y - Ax)})$$

$$f(x) = \frac{1}{2}\|y - Ax\|^2 + \lambda\|x\|_1$$

$$\partial f(x) \ni \underbrace{-A^T(y - Ax) + \lambda v}$$ where

$v_i = \text{sign}(x_i), \; x_i \neq 0$

$\in [-1, 1], \; x_i = 0.$

Proof.

- ∇g Lipschitz with $L>0$.

$$f(y) = g(y) + h(y)$$

$$\leq g(x) + \nabla g(x)^T(y-x) + \frac{L}{2}\|y-x\|^2 + h(y)$$

$$y = x^+ = x - tG_t(x)$$

$$f(y) \leq g(x) - t\nabla g(x)^T G_t(x) + \frac{Lt^2}{2}\|G_t(x)\|^2 + h(x - tG_t(x))$$

$$x - tG_t(x) = \operatorname{prox}_t(x - t\nabla g(x))$$

$$= \operatorname*{argmin}_z \frac{1}{2t}\|z - (x - t\nabla g(x))\|^2 + h(z)$$

$$= \operatorname*{argmin}_z \nabla g(x)^T(z-x) + \frac{1}{2t}\|z-x\|^2 + h(z)$$

$$\nabla g(x) + \frac{1}{t}(z-x) + v = 0, \quad v \in \partial h(z)$$

$$\nabla g(x) - G_t(x) + v = 0, \quad v \in \partial h(x - tG_t(x))$$

$$G_t(x) - \nabla g(x) \in \partial h(x - tG_t(x))$$

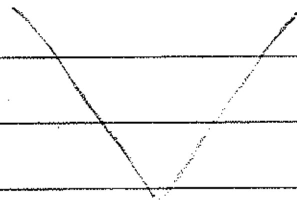
$$h(x) \geq h(x - tG_t(x)) + (G_t(x) - \nabla g(x))^T tG_t(x)$$

$$h(x - tG_t(x)) \leq h(x) - t(G_t(x) - \nabla g(x))^T G_t(x)$$

$$f(x - tG_t(x)) \leq \overbrace{g(x) + h(x)}^{f(x)} - \left(1 - \frac{Lt}{2}\right)t\|G_t(x)\|^2$$

$$\|\nabla g(x) - \nabla g(y)\| \le L\|x-y\|$$

$$\|v_x - v_y\| \le L\|x-y\|.$$



$$\boxed{x^+ = x - t\, G_t(x)}$$

$$x^+ = \text{prox}_t(x - t\nabla g(x)).$$

$$G_t(x) = \frac{x - \text{prox}_t(x - t\nabla g(x))}{t}.$$

$\approx \|X\|_*$

$$\frac{1}{2}\sum (A_{ij} - X_{ij})^2 + \lambda\, \text{rank}(X)$$

$\|x\|_0$

$$\left(P_\Omega(X)\right)_{ij} = \begin{cases} X_{ij} & i,j \in \Omega \\ 0 & \text{otherwise} \end{cases}$$

$$\underbrace{\frac{1}{2}\| P_\Omega(A) - P_\Omega(X) \|_F^2}_{g(X)} + \underbrace{\lambda \|X\|_*}_{h(X)}.$$

- $$\nabla g(X) = -\left(P_\Omega(A) - P_\Omega(X)\right)$$

- Prox

$$\text{prox}_t(X) = \underbrace{\text{argmin}_Z \; \frac{1}{2t}\|X - Z\|_F^2 + \lambda \|Z\|_*}$$

$$= S_{\lambda t}(X)$$

$$X = U \Sigma V^T \qquad U, V \text{ orthogonal columns}$$
$$\Sigma \text{ diagonal}$$

$$(\Sigma_\lambda)_{ii} = \max\{\Sigma_{ii} - \lambda, 0\}.$$

$$\text{prox}_t(X) = U \Sigma_\lambda V^T$$

Fact: $Z = U \Sigma V^T$

$$\partial \|Z\|_* = \{ UV^T + W : \|W\| \le 1, \; U^T W = 0, \; WV = 0 \}$$

$$\mathrm{prox}_t(X) = \operatorname{argmin} \frac{1}{2t}\|X-Z\|_F^2 + \lambda\|Z\|_*$$

$$(Z-X) + \lambda t\partial\|Z\|_* \ni 0.$$

$$X^+ = S_{\lambda t}(X + t(P_\Omega(A) - P_\Omega(X)))$$

$$\nabla g(X) = -(P_\Omega(A) - P_\Omega(X))$$

Lipschitz. $L = 1$

$$\|\nabla g(Y) - \nabla g(X)\|_F = \|P_\Omega(Y) - P_\Omega(X)\|_F$$

$$\leq \|Y - X\|_F$$

$t = 1$

$$X^+ = S_\lambda(X + P_\Omega(A) - P_\Omega(X))$$

$$= S_\lambda(P_\Omega(A) + P_\Omega^\perp(X))$$

where $P_\Omega(X) + P_\Omega^\perp(X) = X$