$$f = f_1 + f_2$$

$$< \quad \leq$$

$$<$$

$$f(\alpha x + (1-\alpha)x) = f_1(\ ) + f_2(\ )$$

---

s.g. $\beta^+ = \beta - t \cdot g$

$g \in \partial f(\beta)$

$O(1/\sqrt{k})$ ... any nonsmooth first-order method i.e,

$$\beta^{(k)} \in \beta^{(0)} + \text{span}\{g^{(0)}, \dots g^{(k-1)}\}$$

can't do better than $1/\sqrt{k}$

g.g.d. $O(1/k)$

Backtracking for g.d.

while $f(x - t\nabla f(x)) > f(x) - \frac{1}{2} t \cdot \|\nabla f(x)\|_2^2$

update $t = \beta t$

Y_{ij} observed entries

$$\min_{B} \ \frac{1}{2}\sum_{i,j\in\Omega}(Y_{ij}-B_{ij})^2 + \lambda\|B\|_*$$

$$\|B\|_* = \sum_{i=1}^{r}\sigma_i(B)$$

cvx. surrogate for rank(B)

$$B = U\Sigma V^T \qquad \Sigma = \begin{pmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{pmatrix}$$

$\|B\|_*$ is to rank(B)

as

$\|\beta\|_1$ is to $\|\beta\|_0 \leftarrow$ # of nzs in β

$$\min \ \frac{1}{2}\|P_\Omega(Y) - P_\Omega(B)\|_F^2 + \lambda\|B\|_*$$

$$[P_\Omega(B)]_{ij} = \begin{cases} B_{ij} & \text{if } i,j\in\Omega \\ 0 & \text{else} \end{cases}$$

$$\nabla g(B) = -(P_\Omega(Y) - P_\Omega(B))$$

Prox function:

$$\text{prox}(B) = \underset{Z}{\text{argmin}} \ \frac{1}{2}\|B-Z\|_F^2 + \lambda t\,\|Z\|_*$$

let $B = U\Sigma V^T$, SVD of B
then

$\text{prox}_t(B) = S_{\lambda t}(B)$ matrix-thresholding

$= U\Sigma_{\lambda t}V^T$

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

~~ggd.~~

$$B^+ = S_{\lambda t}\Big(B + t(P_\Omega(Y) - P_\Omega(B))\Big)$$

$t=1$ gives convergent algorithm because

~~∇g is 1-Lipschitz~~

$$B^+ = S_{\lambda t}\big(P_\Omega(Y) + B - P_\Omega(B)\big)$$

$$= S_{\lambda t}\big(P_\Omega(Y) + P_{\Omega^c}(B)\big)$$

soft-impute

$$\text{prox}_t(x) = \underset{z}{\text{argmin}}\ \frac{1}{2t}\|x-z\|_2^2 + I_C(z)$$

$$= \underset{z\in C}{\text{argmin}}\ \|x-z\|_2$$

$$= P_C(x)$$