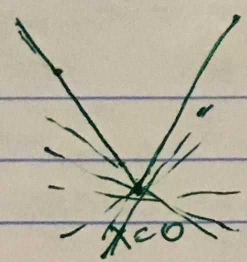
$|x|$



$$y_i - \beta_i = \lambda v_i$$

$$= \begin{cases} \lambda & \beta_i > 0 \\ -\lambda & \beta_i < 0 \\ [-\lambda, \lambda] & \beta_i = 0 \end{cases}$$

$$\begin{bmatrix} y_i - \beta_i = \lambda \,\mathrm{sign}(\beta_i) & \beta_i \neq 0 \\ |y_i - \beta_i| \leq \lambda & \beta_i = 0 \end{bmatrix}$$

$$\Downarrow$$

$$\cancel{y \Rightarrow} \quad |y_i| \leq \lambda$$

$$\beta_i = \begin{cases} y_i - \lambda & y_i > \lambda \\ 0 & |y_i| \leq \lambda \equiv -\lambda \leq y_i \leq \lambda \\ y_i + \lambda & y_i < -\lambda \end{cases}$$

$$\mathrm{prox}(\beta^{(k-1)} - t \nabla \underbrace{g(\beta^{(k-1)})}_{= (y - X\beta)}) \rightarrow \beta$$

$$\min_z \|\beta - z\|^2 + h(\beta)$$

$Y \approx B$ $\quad$ B- low rank

$$\sum_{(i,j) \in \Omega} (Y_{ij} - B_{ij})^2 + \lambda \|B\|_0 \longrightarrow \mathrm{rank}(B)$$

$$\downarrow \qquad = \sum_i 1_{\sigma_i(B) \neq 0}$$

$$\beta = \begin{bmatrix} \sigma_1(B) \\ \vdots \\ \sigma_r(B) \end{bmatrix} \qquad \sum_i \sigma_i(B) = \mathrm{tr}(B)$$

$$\mathrm{Prox}_t(B) = \underset{Z}{\arg\min} \frac{1}{2t}\|B-Z\|_F^2 + \lambda\|Z\|_{tr}$$

$$0 \in Z - B + \lambda t\, \partial\|Z\|_{tr}$$

$U\Sigma_\lambda V^T$
$\shortparallel$
$Z = \tilde{U}\tilde{\Sigma}_\lambda\tilde{V}^T$

Let W st. $0 = \tilde{U}\tilde{\Sigma}_\lambda\tilde{V}^T - U\Sigma V^T + \lambda t(\tilde{U}\tilde{V}^T + W)$

$B = U\Sigma V^T$
$\shortparallel$
$\tilde{U}\tilde{\Sigma}\tilde{V}^T + \tilde{U}_\perp\tilde{\Sigma}_\perp\tilde{V}_\perp^T$

$\Rightarrow 0 = \cancel{\tilde{U}\tilde{\Sigma}\tilde{V}^T} - \cancel{\tilde{U}\tilde{\Sigma}\tilde{V}^T} - \tilde{U}_\perp\tilde{\Sigma}_\perp\tilde{V}_\perp^T + \cancel{\lambda t\tilde{U}\tilde{V}^T} + \lambda t W - \cancel{\lambda t\tilde{U}\tilde{V}^T}$

$\because \Sigma_\lambda = \tilde{\Sigma} - \lambda t I$

$\tilde{U}, \tilde{V}$ - singular vectors corresp to singular values $\Sigma_{ii} > \lambda t$

$(\because \tilde{U}\Sigma_\lambda\tilde{V}^T = \tilde{U}\tilde{\Sigma}\tilde{V}^T - \tilde{U}(\lambda t I)\tilde{V}^T)$

$$\Rightarrow \quad W = \frac{\tilde{U}_\perp\tilde{\Sigma}_\perp\tilde{V}_\perp^T}{\lambda t}$$

$\tilde{U}^T W = 0 \qquad W\tilde{V} = 0$

$$\|W\|_{op} = \max_i \frac{(\tilde{\Sigma}_\perp)_{ii}}{\lambda t} \leq 1$$