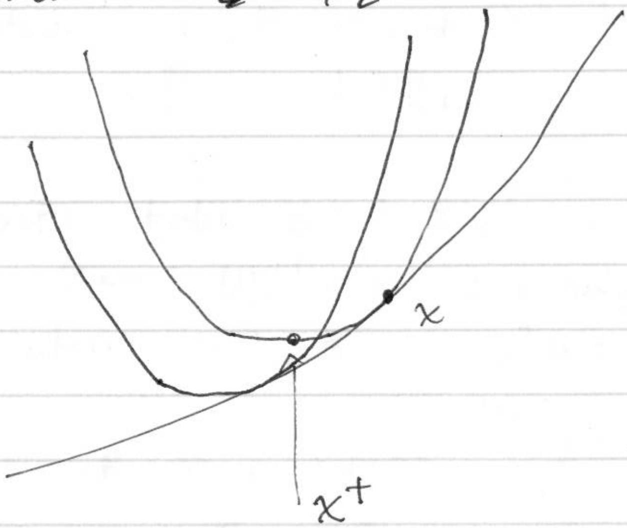
A lower bound on minimizing nonsmooth convex functions: $1/\varepsilon^2$



$$h(\beta) = \lambda \|\beta\|_1$$

$$\text{prox}_t(\beta) = \underset{z}{\text{argmin}} \ \frac{1}{2}\|z-\beta\|_2^2 + \lambda t \|z\|_1$$

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

$$\beta^+ = S_{\lambda t}(\beta - t\nabla g(\beta))$$ prox grad

$$f(\beta) = \frac{1}{2}\|y - X\beta\|_2^2 + \lambda\|\beta\|_1$$

$$\beta^+ = \beta + tX^T(y - X\beta) - \lambda t s$$ sub.g method

$$s \in \partial\|\beta\|_1$$

$$\min_{Z} \quad \frac{1}{2}\|B-Z\|_F^2 + \lambda t \|Z\|_{tr}$$

sub. g. $\quad B-Z + \lambda t \cdot \Gamma$

$\Gamma \in \partial \|Z\|_{tr}$.

Fact: take $Z = U\Sigma V^T$ s.v.d

then $UV^T + W$ is a subg of $\|Z\|_{tr}$

where $\|W\|_{op} \le 1$, $U^T W = 0$, $WV = 0$.

want: $-B + Z + \lambda t \Gamma = 0$

for $Z = S_{\lambda t}(B)$

take $B = U\Sigma V^T$. then

$$U(-\Sigma + \Sigma_{\lambda t})V^T + \lambda t (UV^T + W) = 0$$

$$U(-\Sigma + \Sigma_{\lambda t} + \lambda t I)V^T + \frac{W}{\lambda t} = 0.$$

$$-W = U\left(\frac{-\Sigma + \Sigma_{\lambda t} + \lambda t \cdot I}{\lambda t}\right) \cdot V^T$$

for $\Sigma_{ii} > \lambda t$

$$\frac{-\Sigma_{ii} + (\Sigma_{ii} - \lambda t) + \lambda t}{\lambda t} = 0.$$

$\Sigma_{ii} \le \lambda t$

$$\frac{-\Sigma_{ii} + 0 + \lambda t}{\lambda t} \le 1.$$

checked that $\|W\|_{op} \le 1$ ✓

(also checked $U^T W = WV = 0$ ✓, at home)