$f(x) = g(x) + \sum h_i(x_i)$

assume x is coord. wise min.

$$
\begin{aligned}
f(y) - f(x) &= g(y) - g(x) + \sum h_i(y_i) - h_i(x_i) \\
&\geq \nabla g(x)^T (y - x) + \sum h_i(y_i) - h_i(x_i) \\
&= \sum \underbrace{\nabla_i g(x)(y_i - x_i) + h_i(y_i) - h_i(x_i)}_{\geq 0}
\end{aligned}
$$

$0 \in \nabla_i g(x) + \partial h_i(x_i)$ sub grad. optimality applied along i^{th} coord axis

$-\nabla_i g(x) \in \partial h_i(x_i)$

$h_i(y_i) \geq h_i(x_i) - \nabla_i g_i(x)(y_i - x_i)$ all y_i

$\nabla_i g(x)(y_i - x_i) + h_i(y_i) - h_i(x_i) \geq 0$ ✓

P: $\min_\beta \frac{1}{2}\|y - X\beta\|_2^2 + \lambda\|\beta\|_1$

D: $\max_u \frac{1}{2}\|y\|_2^2 - \frac{1}{2}\|y - u\|_2^2$ s.t. $\|X^T u\|_\infty \leq \lambda$

dual feas point u_0. $\left(\text{eg. } u_0 = y \cdot \frac{\lambda}{\lambda_{max}} \right)$

then $\underset{\overset{\parallel}{\gamma}}{g(u_0)}$ lower bounds g^*.

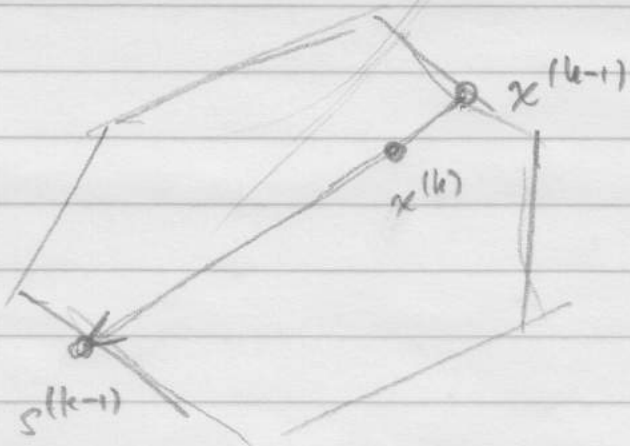
$$
\hat{f}^{quad}(y) = f(x^{(k-1)}) + \nabla f(x^{(k-1)})^T (y - x^{(k-1)}) + \frac{1}{2t_k}\|y - x^{(k-1)}\|_2^2
$$

$$\hat{f}^{lin}(y) = f(x^{(k-1)}) + \nabla f(x^{(k-1)})^T (y - x^{(k-1)})$$

$$s^{(k-1)} = \underset{y \in C}{\operatorname{argmin}}\ \hat{f}^{lin}(y)$$

$$x^{(k)} = x^{(k-1)} + \gamma_k (s^{(k-1)} - x^{(k-1)})$$

$$\gamma_k = \frac{2}{k+1}, \quad k = 1, 2, \ldots$$



$$\max_{\|s\|_1 \leq 1} a^T s = \|a\|_*$$

l_1 norm : subgradients of $\|a\|_\infty = \max\{|a_1|, \ldots |a_p|\}$

if $|a_i| = \|a\|_\infty$

$\text{sign}(a_i) \cdot e_i \in \partial \|a\|_\infty$

$$\max_{\|s\|_q \le 1} a^T s$$

$$= \max_{\sum s_i^q \le 1} a^T s$$

$$\|A\|_{op}$$

$uv^T \in \partial \|A\|_{op}$ where u, v are top left/right sing vec.