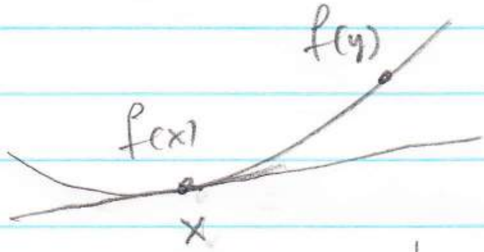
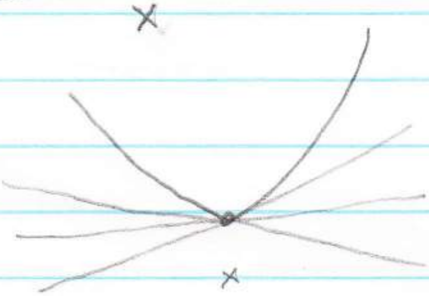
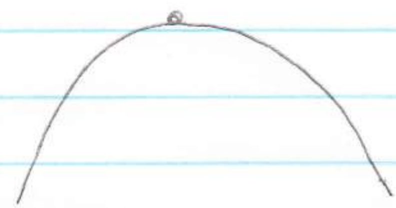
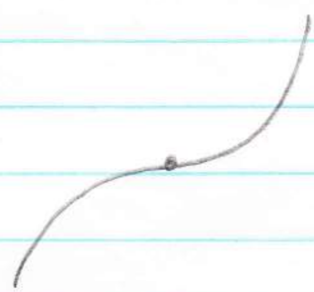








$f(y) \geq f(x) + g^T(y-x)$ all y

$|y| \geq 0 + gy$

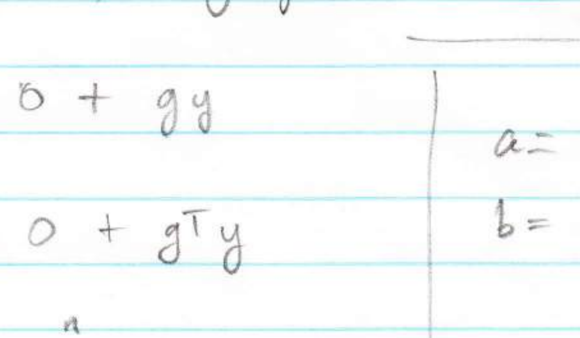
$\|y\|_2 \geq 0 + g^T y$

$\|y\|_1 = \sum_{i=1}^{n} |y_i|$

$a = \nabla f_1(x) \in \mathbb{R}^n$

$b = \nabla_2 f(x) \in \mathbb{R}^n$

$g \in$ line$\{a, b\} \in \mathbb{R}^n$
conv$\{a, b\}$



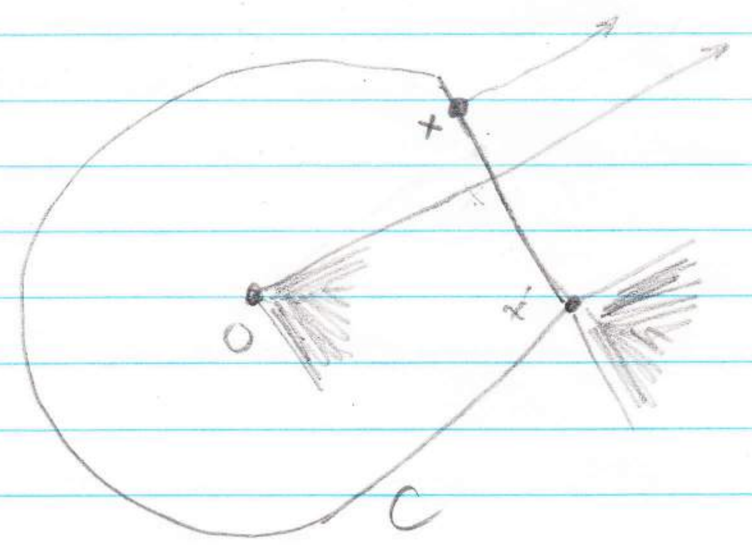
$f(y) \geq f(x) + g_1^T(y-x)$ all y

$f(y) \geq f(x) + g_2^T(y-x)$ all y

$\alpha g_1 + (1-\alpha) g_2$

$$\begin{aligned} & f(x) + (\alpha g_1 + (1-\alpha) g_2)^T(y-x) \\ = \ & \alpha(f(x) + g_1^T(y-x)) + (1-\alpha)(\text{"} \quad g_2^T \quad \text{"}) \\ \leq \ & \alpha f(y) + (1-\alpha) f(y) \\ = \ & f(y) \end{aligned}$$

$g^T y \leq g^T x$ for all $y \in C$



$$A + B = \{a + b : a \in A,\ b \in B\}$$

$$\partial f(x) = \operatorname{conv}\left(\bigcup_{i:\ f_i \text{ active at } x} \partial f_i(x) \right)$$

$$\|x\|_p = \max_{z:\ \|z\|_q \le 1} x^T z \qquad \text{where } \frac{1}{p} + \frac{1}{q} = 1$$

$$\partial \|x\|_p = \operatorname{conv}\left(\bigcup_{z:\ x^T z = \|x\|_p} \{z\} \right)$$

$$= \operatorname*{argmax}_{\|z\|_q \le 1} x^T z$$

$$f(y) \ge f(x^*) + g^T (y - x^*) \quad \text{all } y$$

$$f(y) \ge f(x^*) + 0 \quad \text{all } y$$

For convex f this means :

$$0 \in \{\nabla f(x^*)\} \iff 0 = \nabla f(x^*)$$

$$\boxed{\min \quad f(x) + I_C(x)}$$

$$0 \in \partial f(x) + \partial I_C(x)$$

$$0 \in \{\nabla f(x)\} + N_C(x)$$

$$\iff$$

$$-\nabla f(x) \in N_C(x)$$

$$\iff -\nabla f(x)^T (x-y) \geq 0 \quad \text{all } y \in C$$

$$\iff \nabla f(x)^T (y-x) \geq 0 \quad \text{all } y \in C$$

$$g^T (y-x) \geq 0 \quad \text{all } y \in C$$

for some $g \in \partial f(x)$

KKT conditions

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

$$\boxed{\nabla f(\beta) = -X^T(y - X\beta)}$$

$$-\nabla f(\beta) \in \partial\|\beta\|_1 \cdot \lambda$$

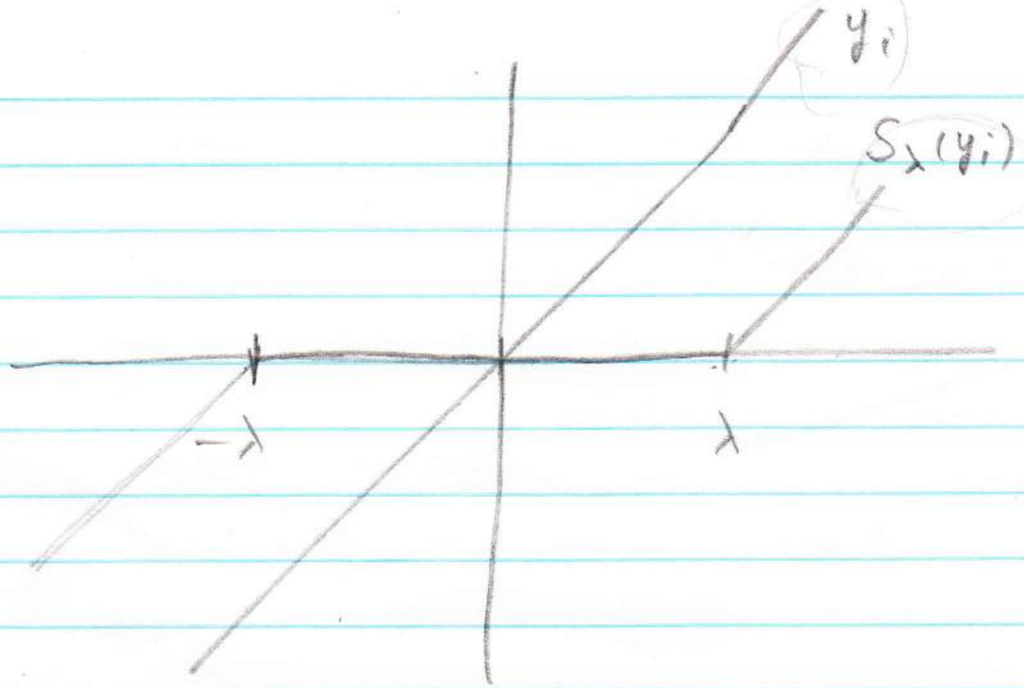
$$\in \lambda v$$

for some $v \in \partial\|\beta\|_1$

$$v_i = \begin{cases} +1 & \beta_i > 0 \\ -1 & < 0 \\ \in [-1, 1] & = 0 \end{cases}$$

$$-\nabla_i f(\beta) = \lambda \operatorname{sign}(\beta_i) \quad \text{when } \beta_i \neq 0$$

$$|\nabla_i f(\beta)| \leq \lambda \quad \text{when } \beta_i = 0$$



$$\begin{cases} y_i-\beta_i = \lambda\, \text{sign}(\beta_i) & \text{when } \beta_i \neq 0 \\ |y_i-\beta_i| \le \lambda & \text{when } \beta_i = 0 \end{cases}$$

suppose $\beta_i = S_\lambda(y_i)$

- if $y_i > \lambda$, $\beta_i = y_i - \lambda > 0$

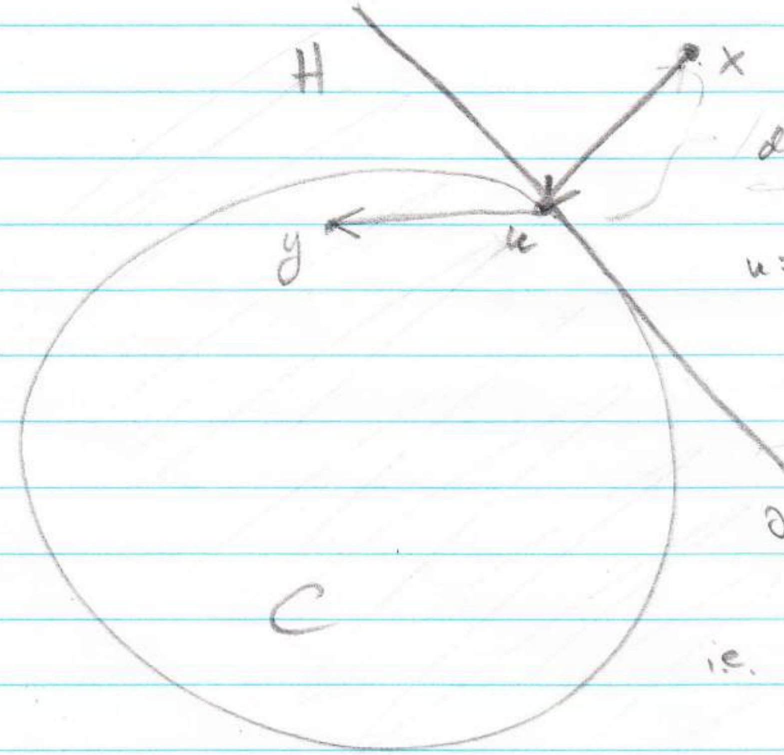
  $y_i - \beta_i = \lambda = \lambda \cdot \text{sign}(\beta_i)$ ✓

- if $y_i < -\lambda$,

  similar

- if $-\lambda \le y_i \le \lambda$, $\beta_i = 0$

  $|y_i - \beta_i| = |y_i| \le \lambda$ ✓



$\mathrm{dist}(x, C)$

$u = P_C(x)$

$$\mathrm{dist}(x, C) = \|x - P_C(x)\|_2$$

$$\partial\, \mathrm{dist}(x, C) = \left\{ \frac{x - P_C(x)}{\|x - P_C(x)\|_2} \right\}$$

i.e.

gradient

$\min\limits_{u \in C} \|y - u\|_2^2$, solution $u = P_C(x)$