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

subgradients at β are

$$\beta - y + \lambda s$$

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

$$s_i \begin{cases} = \text{sign}(\beta_i) & \beta_i \neq 0 \quad i=1,\dots n \\ \in [-1,1] & \beta_i = 0 \end{cases}$$

now set to 0

$$\begin{aligned} y_i - \beta_i &= \lambda\,\text{sign}(\beta_i) & \beta_i \neq 0, \; i=1,\dots n \\ |y_i - \beta_i| &\leq \lambda & \beta_i = 0. \end{aligned}$$

plug in $\beta = S_\lambda(y)$

- $y_i > \lambda$, $\beta_i = y_i - \lambda$, $y_i - \beta_i = \lambda$ ✓
- $y_i < -\lambda$, similar
- $-\lambda \leq y_i \leq \lambda$, $\beta_i = 0$, $|y_i| \leq \lambda$ ✓

$\varepsilon = 1/100$

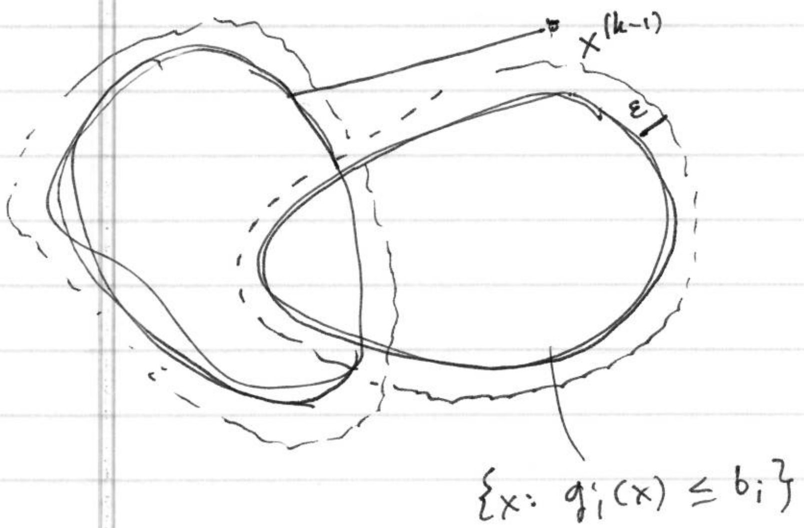
compare $1/\varepsilon$ vs $1/\varepsilon^2$.

$$\beta^+ = \beta - t \cdot \left(\sum (y_i - p_i(\beta)) x_i + \lambda s \right)$$

$$s_i = \begin{cases} \text{sign}(\beta_i) & \beta_i \neq 0 \\ [-1, 1] & \text{else} \end{cases}, \quad i = 1..n$$

$$\text{dist}(x, C_i) = \min_{y \in C_i} \|y - x\|_2$$

$$\underline{f(x^{(k-1)}) = \|x^{(k-1)} - P_{C_i}(x^{(k-1)})\|_2}$$



$$\underbrace{m^2 G^2}_{(\text{Lipschitz bd})^2 \ \text{for} \ f = \sum_{i=1}^{m} f_i}$$