$$\lim_{n \to \infty} \frac{\sum x_i}{n} = \bar{x} \quad (1)$$

?

$$\lim_{n \to \infty} x_n = \bar{x} \quad (2)$$

?

$$\{x_n\} = (-1)^n \qquad \lim_{n \to \infty} \left[\frac{\sum x_i}{n} - \bar{x} \right] \stackrel{?}{=} 0$$

$$\left[\frac{\sum_{i=1}^{n} x_i - n\bar{x}}{n} \right] = \left[\frac{\overbrace{(x_1 - \bar{x}) + (x_2 - \bar{x})}^{\text{finite}} + \cdots \overbrace{(x_n - \bar{x})}^{\varepsilon}}{n} \right]$$

$$= 0$$

$$\lim_{n \to \infty} x_n \to \bar{x}$$

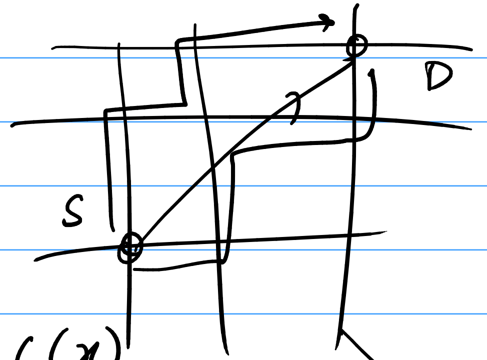
Projected Gradient Method (contd)

$$P_\Omega(x_0) = \frac{1}{2}\min_{x\in\Omega} \|x - x_0\|^2$$

When $\Omega = L_2$ ball, i.e.

$$\Omega : \{x \mid \|x\|_2 \leq 1\}$$

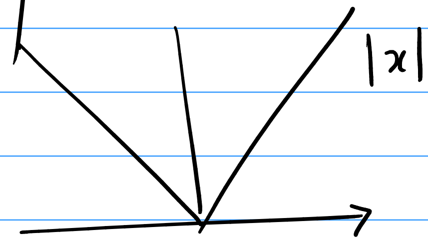
Say $\Omega : \{x \mid \|x\|_1 \leq 1\}$



$$c(x) = 1 - \|x\|_1 \geq 0$$

To solve: $\mathcal{L}(x,\lambda) = f(x) - \lambda c(x)$

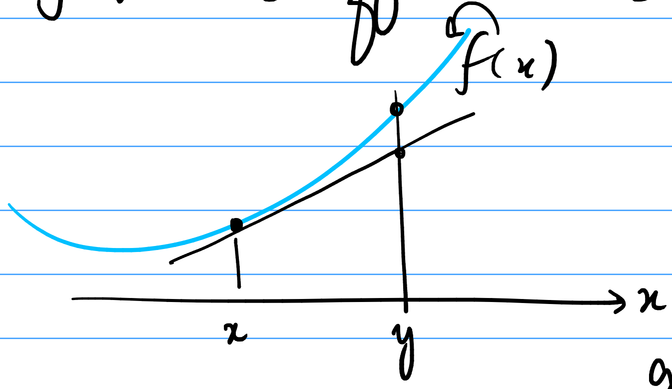
$\nabla_x \mathcal{L}$ $\qquad$ ∇c ?



$$\|x\|_1 = \Sigma |x_i|$$

Subgradients (when gradients don't work)

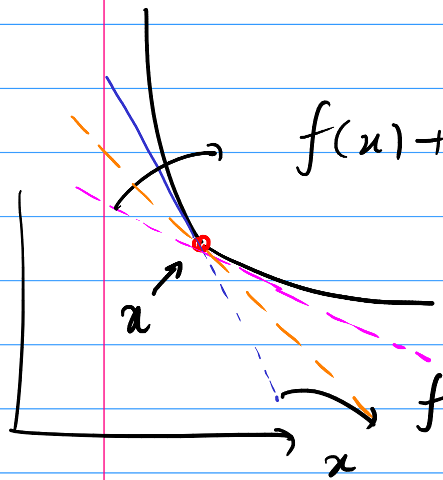
↳ for a differentiable convex fn.



$$f(y) \geq \underbrace{f(x) + \nabla f(x)^T (y-x)}_{\text{supporting hyperplane.}}$$

↳ A subgradient g of a convex fn is s.t.

$$f(y) \geq \underbrace{f(x) + g^T (y-x)} \quad \forall\, x, y \leftarrow$$



$f(x) + g_2^T (y-x)$

$f(x) + g_1^T (y-x)$

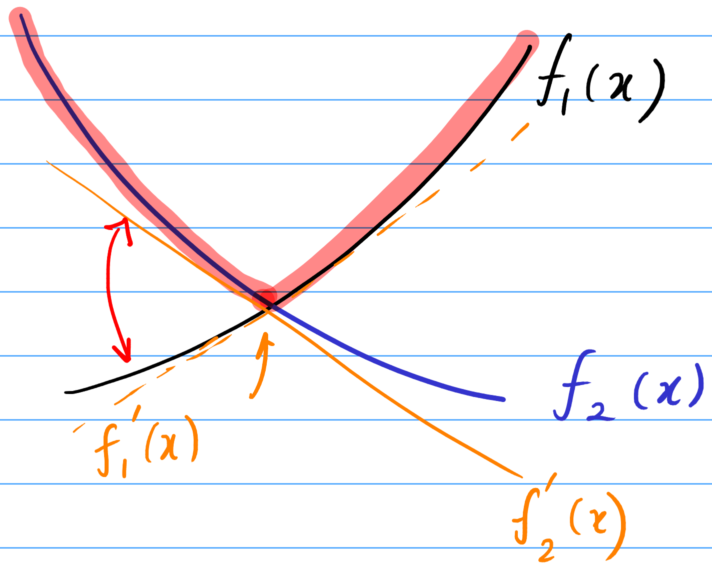
Both supporting hyperplanes, alt, they never cut the fn.

[g_1, g_2 are subgradients of f at x]

↳ Sub differential: $\partial f(x) = \{ g \mid g^T(y-x) \le f(y) - f(x) \}$

↳ closed, convex set.

e.g.



$f(x) = \max(f_1(x), f_2(x))$

1) $f_1(x) > f_2(x) \rightarrow f(x) = f_1(x)$

$\delta f = \{ \nabla f_1(x) \}$

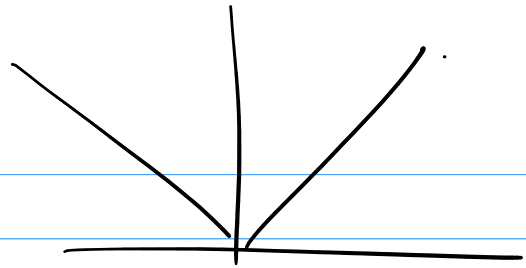
2) $f_1(x) < f_2(x) \rightarrow f(x) = f_2(x)$

$\delta f = \{ \nabla f_2(x) \}$

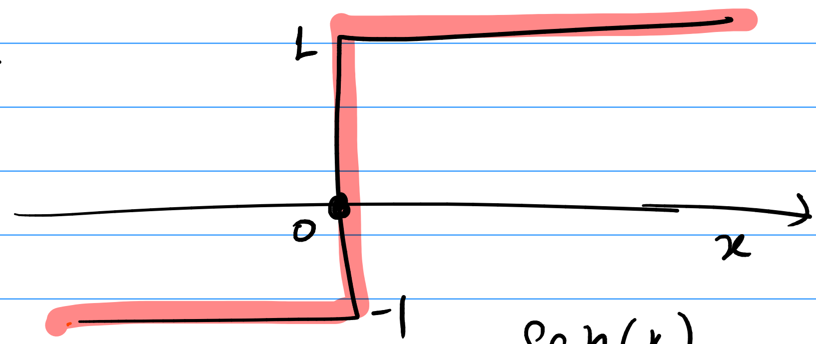
3) $f_1(x) = f_2(x)$, $\delta f = \{ [\nabla f_1(x), \nabla f_2(x)] \}$

i.e. any no in this range

e.g. $f(x) = |x|$



$$\begin{cases} x>0, & \delta f = 1 \\ x<0, & \delta f = -1 \\ x=0, & \delta f = [-1,1] \end{cases}$$



sgn(x)

good candidate for δf.

— x —

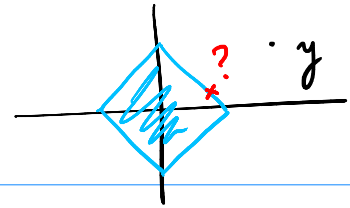
traditionally

$$\mathrm{sgn}(x) = \begin{cases} 1 & x>0 \\ 0 & x=0 \\ -1 & x<0 \end{cases}$$

Here

$$\mathrm{sgn}(x) = \begin{cases} 1 & x>0 \\ [-1,1] & x=0 \\ -1 & x<0 \end{cases}$$

Projection on the L1 ball.



$$P_\Omega(y) = \frac{1}{2} \underset{x\in\Omega}{\operatorname{argmin}} \|y - x\|_2^2 \quad , \quad \Omega : C(x) = 1 - \|x\|_1 \geq 0.$$

y fixed → given to us

$$\mathcal{L}(x,\lambda) = \frac{1}{2}\|x-y\|_2^2 - \lambda(1-\|x\|_1) ,$$

KKT conds $\boxed{\nabla_x \mathcal{L} = 0 = (x^* - y) + \lambda \nabla(\|x\|_1) \quad , \quad \lambda \geq 0.}$

$\lambda C(x) = 0$

① $\lambda = 0 \Rightarrow \nabla_x\mathcal{L} = (x^* - y) = 0 \Rightarrow \underline{x^* = y}$, i.e y was $\in \Omega$.

② $\lambda \neq 0 \Rightarrow C(x) = 0$. $\underline{\nabla(\|x\|_1)\ ?}$

→ $\|x\|_1 = \sum |x_i|$, $\nabla_x\mathcal{L} \to \frac{\partial \mathcal{L}}{\partial x_i} = x_i - y_i + \lambda\, \partial|x_i| = 0$

$x_i = 0 \Rightarrow \quad y_i + \lambda\,\delta|x_i| = 0$

$\Rightarrow \quad \delta|x_i| \neq 0$ in general.

As per KKT: $\boxed{(x_i - y_i) + \lambda\,\partial|x|_i = 0}$

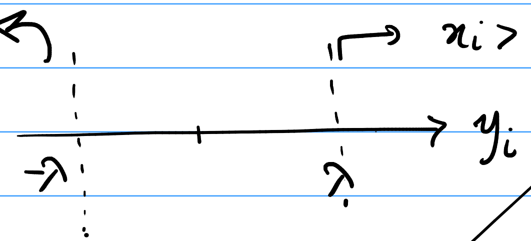
$$\begin{array}{lcl} x_i > 0 & \rightarrow & x_i - y_i + \lambda \times 1 = 0 \\ x_i < 0 & \rightarrow & x_i - y_i + \lambda \times (-1) = 0 \\ x_i = 0 & \rightarrow & x_i - y_i + \lambda\,\delta|x_i| = 0 \end{array} \quad \begin{array}{l} \rightarrow \underline{x_i = y_i - \lambda} \\ \rightarrow \underline{x_i = y_i + \lambda} \\ \rightarrow x_i = y_i + \lambda\,\delta|x_i|. \ \text{e.g. } \delta|x_i| = 0 \end{array}$$

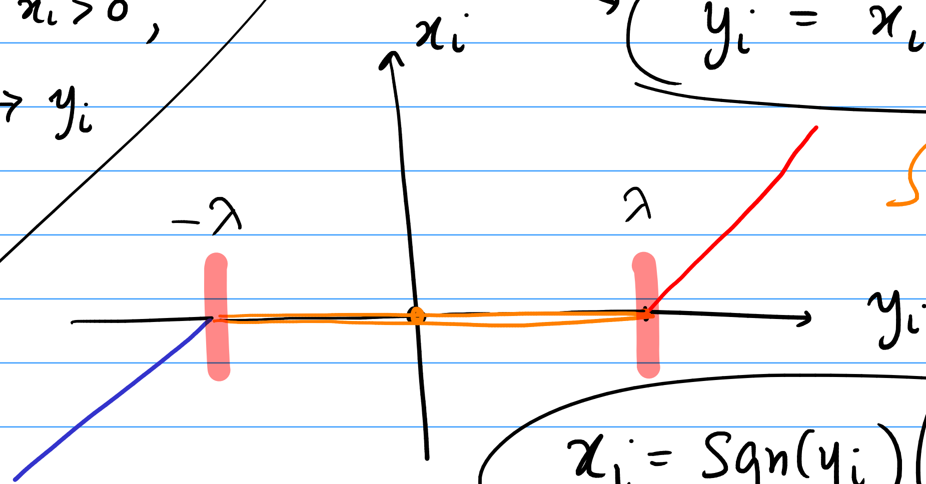
$\delta|x_i|$: any value $\in [-1, 1]$

$$\boxed{x_i = y_i - \lambda\,\mathrm{sgn}(x_i)} \Rightarrow$$

we only know y_i

$$y_i = x_i + \lambda\,\mathrm{sgn}(x_i)$$





(A)

Soft thresholding operator.

Compactly.

$$x_i = \mathrm{Sgn}(y_i)\left(|y_i| - \lambda\right)_+$$

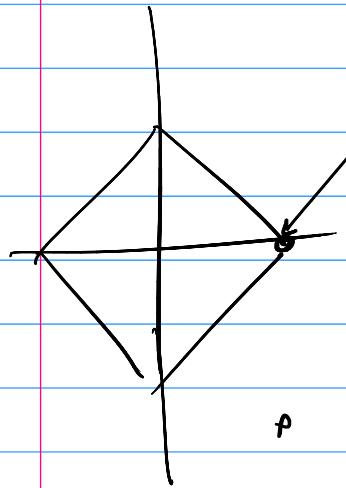
(B)

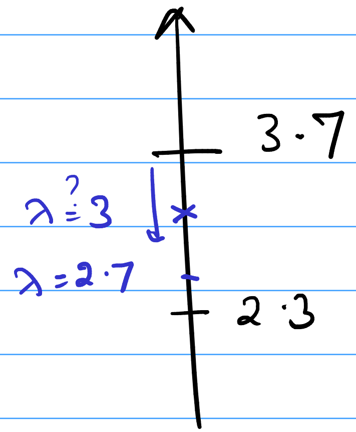
$$(a)_+ = \mathrm{Max}(0, a)$$

$$c(x) = 0 \implies \sum |x_i| = 1 \implies \sum_{i=1}^{n} \left| \operatorname{sgn}(y_i)\left(|y_i| - \lambda\right)_+ \right| = 1.$$

$$\sum_{i=1}^{n} \left(|y_i| - \lambda\right)_+ = 1$$

e.g. $y = \begin{pmatrix} 3.7 \\ 2.3 \end{pmatrix}$





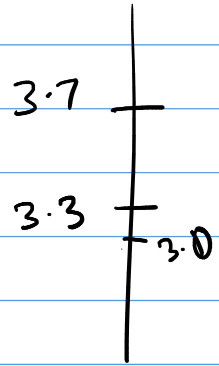
$\implies \lambda = 2.7$ ✓

$$x_i = \operatorname{sgn}(y_i)\left(|y_i| - \lambda\right)_+$$

$x_1 = 1 \times (1) = 1, \quad x_2 = 1 \times (0) = 0$

Soln: $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

e.g. $\begin{pmatrix} 3.7 \\ -3.3 \end{pmatrix}$



$\lambda = 3. \longrightarrow \quad x_1 = 1(0.7)$

$x_2 = -1 \times (0.3)$

Soln $= \begin{pmatrix} 0.7 \\ -0.3 \end{pmatrix}$