Convex Optimization for ML & CV

Name: Uzair Akbar ID: 03697290

Weekly Exercise 5 - Convex Conjugate

Exercise 1

$$\because \|X\|_{1,2} = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} X_{i,j}^2 \right)^{1/2} = \sum_{i=1}^{m} l_2(X_i)$$

$$\Rightarrow \quad = \sum_{i=1}^{m} l_2\left([e_i^T X]^T\right) = \sum_{i=1}^{m} l_2(e_i^T X)$$

Where $X_i = [X_{i,1}, \dots X_{i,j}, \dots X_{i,n}]^T = [e_i^T X]^T$,

and e_i is the i^{th} standard basis in $\mathbb{R}^m$.

Now,

$$\because \|X\|_{1,2} = \sum_{i=1}^{m} l_2(e_i^T X)$$

$$\left[\begin{array}{l} \because \partial \tilde{J} = \partial J_1 + \partial J_2 \\ \quad \text{for } \tilde{J} = J_1 + J_2, \\ \text{and,} \\ \because \partial \tilde{J} = K^T \partial J(K \cdot x) \\ \quad \text{for } \tilde{J} = J(K \cdot x) \end{array}\right]$$

$$\Rightarrow \partial \|X\|_{1,2} = \sum_{i=1}^{m} e_i \, \partial l_2(e_i^T X)$$

$$\Rightarrow \quad = \sum_{i=1}^{m} \overbrace{\underbrace{e_i}_{\in \mathbb{R}^{m\times 1}} \, \partial l_2\left(\underbrace{[e_i^T X]^T}_{\in \mathbb{R}^{1\times n}}\right)}^{\in \mathbb{R}^{m\times n}}$$

$$[\because l_2(e_i^T X) = l_2([e_i^T X]^T)]$$

$$\Rightarrow \partial \|X\|_{1,2} = \sum_{i=1}^{m} e_i \, \partial l_2(X_i)$$

$$\therefore \partial \|X\|_{1,2} = \begin{bmatrix} \partial l_2(X_1) \\ \vdots \\ \partial l_2(X_i) \\ \vdots \\ \partial l_2(X_m) \end{bmatrix} \in \mathbb{R}^{m \times n}$$

Where the l_2 subdifferential is given by:

$$\partial l_2(x) = \{ p \in \mathbb{R}^n \mid \|p\|_2 \leq 1, \ \langle p, u \rangle = \|x\|_2 \}$$

$$= \left\{ p \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} \|p\|_2 \leq 1 \ , & x = \underline{0} \\ p = x / \|x\|_2 \ , & x \neq \underline{0} \end{array} \right\}$$

[$\underline{0}$ denotes zero vector]

Exercise 2

Part A $\quad g(x,t) = \begin{cases} t f(x/t), & t > 0 \\ \infty, & \text{otherwise} \end{cases}$

$$\because g^*(y,s) = \sup_{\substack{(x,t) \\ \in \mathbb{R}^n \times \mathbb{R}}} \left\langle \begin{bmatrix} x \\ \hdashline t \end{bmatrix}, \begin{bmatrix} y \\ \hdashline s \end{bmatrix} \right\rangle - g(x,t)$$

$$\Rightarrow \quad = \sup_{\substack{(x,t) \in \\ \mathbb{R}^n \times \mathbb{R}}} \left[x^T \,\vdots\, t \right] \begin{bmatrix} y \\ \hdashline s \end{bmatrix} - g(x,t)$$

$$\Rightarrow \quad = \sup_{\substack{(x,t) \in \\ \mathbb{R}^n \times \mathbb{R}}} x^T y + st - g(x,t)$$

$$\Rightarrow \quad = \sup_{\substack{(x,t) \in \mathbb{R}^n \times \mathbb{R} \\ t > 0}} x^T y + st - t f(x/t)$$

[$\because g(\cdot,t) = \infty$ for $t \leq 0$, hence $t > 0$ should be considered]

$$\Rightarrow \quad = \sup_{\substack{(x,t) \in \mathbb{R}^n \times \mathbb{R} \\ t > 0}} t\left(s + \left\{ \left(\frac{x}{t}\right)^T y - f\left(\frac{x}{t}\right) \right\} \right)$$

$$\Rightarrow \quad = \sup_{\substack{(z,t) \in \mathbb{R}^n \times \mathbb{R} \\ t > 0}} t\left(s + \left\{ \langle z, y \rangle - f(z) \right\} \right)$$

[Setting $z := x/t$]

$$\Rightarrow \quad = \sup_{t > 0} t\left(s + f^*(y) \right)$$

$$\therefore g^*(y,s) = \begin{cases} 0, & f^*(y) \leq -s \\ \infty, & \text{otherwise} \end{cases}$$

Part B

$$\because g^{**}(x,t) = \sup_{(y,s)\in\mathbb{R}^n\times\mathbb{R}} \left\langle \begin{bmatrix} y \\ \hdashline s \end{bmatrix}, \begin{bmatrix} x \\ \hdashline t \end{bmatrix} \right\rangle - g^*(y,s)$$

$$= \sup_{\substack{(y,s)\\ \in\mathbb{R}^n\times\mathbb{R}}} y^T x + st - g^*(y,s)$$

$$\Rightarrow g^{**}(x,t) = \sup_{f^*(y)\leq -s} y^T x + st \qquad \left[\because g^*(\cdot,\cdot) = \infty \text{ otherwise}\right]$$

Now we explore the following three conditions on t:

For $t<0$,

$$\Rightarrow g^{**}(x,t) = \infty \qquad \left[\because st\to\infty \text{ for } s\to-\infty, \text{ and therefore the supremum is } \infty.\right]$$

For $t=0$,

$$\Rightarrow g^{**}(x,t) = \sup_{f^*(y)\leq -s} y^T x = \sup_{y\in \mathrm{dom}(f^*)} (y^T x)$$

$$\therefore g^{**}(x,t) = \sigma_{\mathrm{dom}(f^*)}(x) \qquad \left[\text{Definition of support function}\right]$$

For $t>0$,

$$\Rightarrow g^{**}(x,t) = \sup_{s\leq -f^*(y)} y^T x + st$$

$$= \sup_{y} y^T x - t f^*(y)$$

[By setting highest possible value of s for $\sup(\cdot)$ as $s = -f^*(y)$. $s\leq -f^*(y)$]

$$\Rightarrow g^{**}(x,t) = \sup_y y^T x - t f^*(y)$$

$$= t \sup_y y^T(x/t) - f^*(y)$$

$$= t\, f^{**}(x/t)$$

$$\therefore g^{**}(x,t) = t f(x/t)$$ [For convex, proper and l.s.c. $f(\cdot)$]

Hence,

$$g^{**}(x,t) = \begin{cases} t \cdot f(x/t) & , t > 0 \\ \sigma_{dom(f^*)}(x) & , t = 0 \\ \infty & , t < 0 . \end{cases}$$

Exercise 3

Part 1.

$$\because (Af)^*(u) = \sup_v \langle u, v \rangle - (Af)(u)$$

$$\Rightarrow \quad = \sup_v \langle u, v \rangle - \inf_{Aw=v} f(w)$$

[$\because$ otherwise $(Af)(u) = \infty$]

$$\Rightarrow \quad = \sup_{v=Aw} \langle u, v \rangle - f(w)$$

$$\Rightarrow \quad = \sup_w \langle u, Aw \rangle - f(w)$$

$$\Rightarrow \quad = \sup_w \langle A^T u, w \rangle - f(w)$$

$$\therefore (Af)^*(u) = f^*(A^T u)$$

Part 2.

Assume

$$A^T f^* = (A^T f^*)^{**} \quad \text{— (i)}$$

$$\Rightarrow A^T f^* = (f^{**} \circ A)^* = (f \circ A)^* \quad \text{— (ii)}$$

For (i) to hold, $A^T f^*$ needs to be proper, convex and l.s.c.
For (ii) to hold, f needs to be proper, convex and l.s.c.
Both properties follow from the Biconjugate theorem in the lecture.

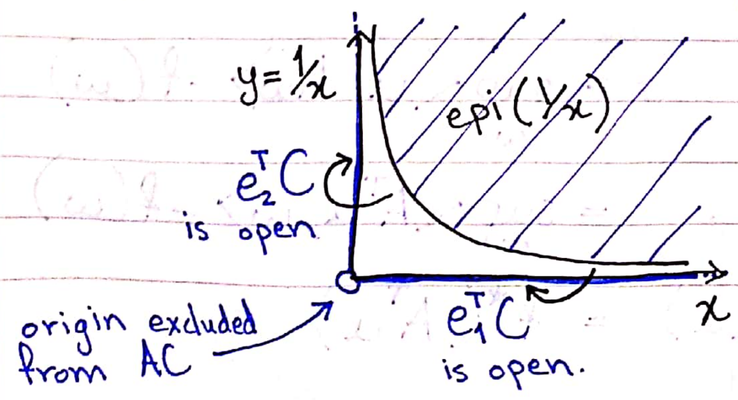
Part 3.

Consider $C := \text{epi}(1/x)$. Then C is closed, convex and non-empty for $x \in \text{dom}(1/x)$.

Now consider $A \in \{e_1^T, e_2^T\}$, with e_i as the i^{th} standard basis in $\mathbb{R}^2$.

For the above scenario, $AC = (0, \infty)$, which is not closed.

That is, the projection of the closed set $C := \text{epi}(1/x)$ onto the x or y-axis is open.



Part 4.

Continuing from the previous example of $C := \text{epi}(1/x)$ and $A \in \{e_1^T, e_2^T\}$, we take a simple convex, closed and proper function $f := \delta_C$ over set C.

Then,

$$\because Af(u) = \inf_{v \in \mathbb{R}^2,\ Av=u} f(v)$$

$$= \inf_{v \in \mathbb{R}^2,\ Av=u} \delta_C(v)$$

$$= \begin{cases} 0, & u > 0 \\ \infty, & \text{otherwise} \end{cases} \quad \left[\begin{array}{l}\text{This holds} \\ \text{for } A \in \{e_1^T, e_2^T\}\end{array}\right]$$

$$\therefore Af(u) = \delta_{AC}(u)$$

But, as we have already seen, $AC = (0, \infty)$, and therefore $Af(u) = \delta_{AC}(u)$ is <u>not</u> closed.