Name: Uzair Akbar ID: 03697290

Convex Optimization for ML & CV

Weekly Exercises 4 - Convex Conjugate

Exercise 1

Proving the hint is straightforward:

$\because \|Au\|_2 = \sqrt{\|Au\|_2^2} = \sqrt{\langle Au, Au\rangle} = \sqrt{u^T A^T A u}$

$\Rightarrow \quad = \sqrt{u^T u} = \sqrt{\langle u, u\rangle} = \sqrt{\|u\|_2^2}$ [$\because$ A orthonormal]

$\therefore \|Au\|_2 = \|u\|_2$ ——— (i)

Now, we simplify the projection operator Π_C by rewriting it in terms of a simpler projection onto a unit ball $C' := \{x \in \mathbb{R}^n \,|\, \|x\|_\infty \leq 1\}$ on l_∞-norm.

Note that: $x \in C' \iff u = A^T x \in C$ ——— (ii)

Now,

$$\because \Pi_C(v) := \arg\min_{u \in C} \frac{1}{2}\|u - v\|_2^2$$

$$\Rightarrow \quad = \arg\min_{\|Au\|_\infty \leq 1} \frac{1}{2}\|u - v\|_2^2$$

$$\Rightarrow \quad = A^T \arg\min_{\|x\|_\infty \leq 1} \frac{1}{2}\|A^T x - v\|_2^2 \quad \text{[From (ii)]}$$

$$\Rightarrow \Pi_C(v) = A^T \underset{x \in C'}{\arg\min} \frac{1}{2} \| A(A^T x - v) \|_2^2$$

[From (i)]

$$\Rightarrow \qquad = A^T \underset{x \in C'}{\arg\min} \frac{1}{2} \| x - Av \|_2^2 \qquad [\because A \text{ orthonormal}]$$

$$\therefore \Pi_C(v) = A^T \Pi_{C'}(Av)$$

Hence, $\Pi_C(v)$ is only an orthonormal transformation of a projection of Av onto a l_∞ unit ball.

Exercise 2

$$J(u) := \frac{1}{q}\|u\|_q^q = \sum_{i=1}^{n} \frac{1}{q}|u_i|^q, \quad q \in [1, +\infty]$$

$$\therefore J^*(v) := \sup_u \langle u, v\rangle - J(u) = \sup_u \sum_{i=1}^{n} u_i v_i - \sum_{i=1}^{n} \frac{1}{q}|u_i|^q$$

We can take the i^{th} element and maximize for it to maximize for u:

$$\left.\frac{d\left(u_i v_i - \frac{1}{q}|u_i|^q\right)}{du_i}\right|_{u_i=u_i^*} = 0$$

$$\Rightarrow v_i - |u_i^*|^{q-1}\,\mathrm{sign}(u_i^*) = 0$$

$$\Rightarrow v_i = |u_i^*|^{q-1}\,\mathrm{sign}(u_i^*) \qquad \text{——— (i)}$$

$$\Rightarrow |v_i| = |u_i^*|^{q-1}$$

$$\Rightarrow |v_i|^{1/q-1} = |u_i^*|$$

$$\Rightarrow |v_i|^{1/q-1}\,\mathrm{sign}(u_i^*) = |u_i^*|\,\mathrm{sign}(u_i^*) = u_i^*$$

$$\therefore u_i^* = |v_i|^{1/q-1}\,\mathrm{sign}(u_i^*) = |v_i|^{1/q-1}\,\mathrm{sign}(v_i)$$

[From (i), $\Rightarrow \mathrm{sign}(v_i) = \mathrm{sign}(u_i)$]

Hence, $J^*(v)$ becomes:

$$\therefore J^*(v) = \sum_{i=1}^{n} u_i^* v_i - \sum_{i=1}^{n} \frac{1}{q}|u_i^*|^q$$

$$\Rightarrow J^*(v) = \sum_{i=1}^{n} |v_i|^{q/q-1} - \sum_{i=1}^{n} \frac{1}{q} |v_i|^{q/q-1}$$

$$\Rightarrow \quad = \sum_{i=1}^{n} \left(1 - \frac{1}{q}\right) |v_i|^{q/q-1} = \sum_{i=1}^{n} \left(1 - \frac{1}{q}\right) |v_i|^{(1-1/q)}$$

$$\Rightarrow \quad = \sum_{i=1}^{n} \frac{1}{p} |v_i|^{1/p} \qquad \left[\text{Substituting } \frac{1}{p} = 1 - \frac{1}{q}\right]$$

$$\therefore J^*(v) = \frac{1}{p} \|v\|_p^p \quad , \quad \frac{1}{p} = 1 - \frac{1}{q}$$

$$J(u) := \begin{cases} \frac{1}{2} \|u\|_2^2 & , \|u\|_2 \leq \varepsilon \\ +\infty & , \text{otherwise.} \end{cases}$$

$$\because J^*(v) := \sup_u \langle u, v \rangle - J(u)$$

$$\Rightarrow J^*(v) = \sup_{\|u\|_2 \leq \varepsilon} \langle u, v \rangle - \frac{1}{2} \|u\|_2^2$$

Because otherwise $J^*(v) = -\infty$ for $\|u\|_2 \nleq \varepsilon$, which is clearly not the supremum.

We can also write:

$$J^*(v) = \langle u^*, v \rangle - \frac{1}{2} \|u^*\|_2^2$$

Where

$$u^* = \arg \max_{\|u\|_2 \leq \varepsilon} \langle u, v \rangle - \frac{1}{2} \|u\|_2^2$$

$$\Rightarrow u^* = \arg\min_{\|u\|_2 \le \varepsilon} \frac{1}{2}\|u\|_2^2 - \langle u, v\rangle$$

$$\Rightarrow \quad = \arg\min_{\|u\|_2 \le \varepsilon} \frac{1}{2}\|u\|_2^2 - \langle u, v\rangle + \frac{1}{2}\|v\|_2^2$$

[$\because$ we are minimizing over u only]

$$\Rightarrow \quad = \arg\min_{\|u\|_2 \le \varepsilon} \frac{1}{2}\left\{\langle u, v\rangle - 2\langle u, v\rangle + \langle u, v\rangle\right\}$$

$$\Rightarrow \quad = \arg\min_{\|u\|_2 \le \varepsilon} \frac{1}{2}\langle u-v, u-v\rangle$$

$$\Rightarrow u^* = \arg\min_{\|u\|_2 \le \varepsilon} \frac{1}{2}\|u-v\|_2^2 =: \Pi_{\|u\|_2 \le \varepsilon}(v)$$

Therefore,
if $v \in \{u \mid \|u\|_2 \le \varepsilon\} \Rightarrow \Pi_{\|u\|_2 \le \varepsilon}(v) = v$,
hence,

$$u^* = v \quad \text{if} \quad \|v\|_2 \le \varepsilon.$$

Otherwise, $u^* = \Pi_{\|u\|_2 \le \varepsilon}(v) = \varepsilon \dfrac{v}{\|v\|_2}$

Substituting u^* in $J^*(v)$:

$$\Rightarrow J^*(v) = \begin{cases} \frac{1}{2}\|v\|_2^2 , & \|v\|_2 \le \varepsilon \\ \varepsilon\|v\|_2^2 - \frac{\varepsilon^2}{2} , & \text{otherwise.} \end{cases}$$

$$J(u) := \sum_{i=1}^{n} u_i \log u_i + \delta_{\Delta^{n-1}}(u)$$

$$\because J^*(v) = \sup_u \langle u, v \rangle - J(u)$$

$$\Rightarrow \quad = \sup_u \left[\sum_{i=1}^{n} \{ u_i v_i - u_i \log u_i \} - \delta_{\Delta^{n-1}}(u) \right]$$

$$\Rightarrow J^*(v) = \sup_{u \in \Delta^{n-1}} \sum_{i=1}^{n} (u_i v_i - u_i \log u_i)$$

We'll take the i^{th} element and maximize for u_i in order to maximize for u:

$$\left. \frac{\partial \left(\sum_{i=1}^{n} \{ u_i v_i - u_i \log u_i \} - \delta_{\Delta^{n-1}}(u) \right)}{\partial u_i} \right|_{u_i = u_i^*} = 0$$

$$\Rightarrow v_i - \log u_i^* - 1 + \gamma = 0 \quad , \quad \gamma := \left. -\frac{\partial \delta_{\Delta^{n-1}}(u)}{\partial u_i} \right|_{u_i = u_i^*}$$

$$\Rightarrow u_i^* = e^{(-\gamma + v_i - 1)}$$

$$\because \sum_{i=1}^{n} u_i^* = 1 \qquad [\because u^* \in \Delta^{n-1}]$$

$$\Rightarrow \sum_{i=1}^{n} e^{(-\gamma + v_i - 1)} = 1$$

$$\Rightarrow \log \left(\sum_{i=1}^{n} e^{(-\gamma + v_i - 1)} \right) = 0$$

$$\Rightarrow \log \left\{ e^{-(\gamma+1)} \sum_{i=1}^{n} e^{v_i} \right\} = 0$$

$$\Rightarrow \log \left(\sum_{i=1}^{n} e^{v_i} \right) = \gamma + 1 \qquad \text{--- (ii)}$$

Substituting u^* into $J^*(v)$:

$$\therefore J^*(v) = \sum_{i=1}^{n} u_i^* v_i - u_i^* \log u_i^*$$

$$= \sum_{i=1}^{n} e^{(-\gamma-1+v_i)} \cdot v_i - e^{(-\gamma-1+v_i)} \cdot \log e^{(-\gamma-1+v_i)}$$

$$= \sum_{i=1}^{n} e^{(-\gamma-1+v_i)} \cdot v_i - (-\gamma-1+v_i) \cdot e^{(-\gamma-1+v_i)}$$

$$= \sum_{i=1}^{n} e^{(-\gamma-1+v_i)} \cdot (v_i + \gamma + 1 - v_i)$$

$$= (\gamma+1) \sum_{i=1}^{n} e^{(-\gamma-1+v_i)} = (\gamma+1) \sum_{i=1}^{n} u_i^* = \gamma + 1 \quad [\because u^* \in \Delta^{n-1}]$$

$$\therefore J^*(v) = \log\left(\sum_{i=1}^{n} e^{v_i}\right) \quad [\text{from (ii)}]$$

Exercise 3

We have to show that

$$\|\Pi_C(u) - \Pi_C(v)\| \leq \|u-v\| \quad , \quad \forall\, u,v \in \mathbb{E}.$$

Consider the projection of u onto C:

$$\Pi_C(u) := \arg\min_{v \in C} \frac{1}{2} \|v-u\|_2^2$$

Then, equivalently via the variational inequality:

$$\langle u - \Pi_C(u) , x - \Pi_C(u) \rangle \leq 0 \quad , \forall\, x \in C, \forall u \in \mathbb{E}$$

$$\Rightarrow \langle u - \Pi_C(u), \Pi_C(v) - \Pi_C(u) \rangle \leq 0 \quad \text{— (i)}$$

$$[\because \Pi_C(v) \in C \ , \forall v \in \mathbb{E}]$$

Similarly, for v:

$$\Rightarrow \langle v - \Pi_C(v), \Pi_C(u) - \Pi_C(v) \rangle \leq 0 \quad \text{— (ii)}$$

Subtracting (i) from (ii):

$$\Rightarrow \langle v - \Pi_C(v), \Pi_C(u) - \Pi_C(v) \rangle - \langle u - \Pi_C(u), \Pi_C(v) - \Pi_C(u) \rangle \leq 0$$

$$\Rightarrow \langle v - u + \Pi_C(u) - \Pi_C(v) , \Pi_C(u) - \Pi_C(v) \rangle \leq 0$$

$$\Rightarrow \langle \Pi_C(u) - \Pi_C(v), \Pi_C(u) - \Pi_C(v) \rangle \leq \langle u-v , \Pi_C(u) - \Pi_C(v) \rangle$$

$$\cancel{= \|u-v\| \cdot \|\Pi_C(u) - \Pi_C(v)\| \cdot \cos\theta}$$

$$\Rightarrow \cancel{\|\Pi_c(u)-\Pi_c(v)\|_2^2 \leq \|u-v\| \cdot \|\Pi_c(u)-\Pi_c(v)\|_2}$$

$$\Rightarrow \langle \Pi_c(u)-\Pi_c(v), \Pi_c(u)-\Pi_c(v) \rangle \leq \|u-v\| \cdot \|\Pi_c(u)-\Pi_c(v)\|_\star$$ [Cauchy-Schwarz]

$$\Rightarrow \|\Pi_c(u)-\Pi_c(v)\| \cdot \|\Pi_c(u)-\Pi_c(v)\|_\star \leq \|u-v\| \cdot \|\Pi_c(u)-\Pi_c(v)\|_\star$$ [Cauchy-Schwarz with Equality]

$$\therefore \|\Pi_c(u)-\Pi_c(v)\| \leq \|u-v\|$$

Note that for $\Pi_c(u)-\Pi_c(v) = 0$ or $u-v=0$, we do not require Cauchy-Schwarz as this inequality holds with equality in that case.

Exercise 4

Let $C := \bigcap_{1 \le i \le n} C_i \ni u^*$, $C' := \bigcap_{\substack{i \in I \\ 1 \le i \le n}} C_i$

$J(u) := \sum_{\substack{j \notin I \\ 1 \le j \le n}} d^2(u, C_j)$, $\delta_{C'}(u) := \sum_{\substack{i \in I \\ 1 \le i \le n}} \delta_{C_i}(u)$

$$\tilde{J}(u) := J(u) + \delta_{C'}(u)$$

Given the above definitions, we have to show that

$$u^* \in C \iff 0 \in \partial \tilde{J}(u^*) = \partial J(u^*) + N_{C'}(u^*)$$

Proof $\Rightarrow$

Assume $u^* \in C$. Then

$\because N_{C'}(u^*) := \{ p \in E \mid \langle p, v - u^* \rangle \le 0, \ \forall \ v \in C' \}$

$\Rightarrow 0 \in N_{C'}(u^*)$ [By definition, setting $p = 0$]

Now,

$\because \partial J(u^*) := \{ p \in E \mid J(v) \ge J(u^*) + \langle p, v - u^* \rangle, \ \forall v \in E \}$

$\Rightarrow J(v) \ge J(u^*) + \langle p, v - u^* \rangle$, for $\forall \ p \in \partial J(u^*)$

$\Rightarrow J(v) \ge 0 + \langle p, v - u^* \rangle$ [$\because J(u^*) = 0$ by definition]

$\Rightarrow J(v) \ge 0$ [assume $p = 0$]

Which is consistent with the definition of $J(\cdot)$, hence, $p = 0 \in \partial J(u^*)$.

$$\Rightarrow 0 \in \partial \tilde{J}(u^*) \qquad [\text{Minkowski sum of } \partial J(u^*) \text{ and } N_{C'}(u^*)]$$

$$\Rightarrow u^* \in \arg\min_{u \in \mathbb{E}} \tilde{J}(u)$$
$$= \arg\min_{u \in C'} J(u).$$

Proof $\Leftarrow$

Assume $0 \in \partial \tilde{J}(u^*)$, then from Minkowski sum:

$$\Rightarrow 0 \in \partial J(u^*) \quad \text{and} \quad 0 \in \partial \delta_{C'}(u^*) := N_{C'}(u^*)$$

For $0 \in \partial J(u^*)$, it follows that

$$\because J(v) \geq J(u^*) + \langle p, v - u^* \rangle \qquad [\text{Definition of } \partial J(u^*)]$$

$$\Rightarrow J(v) \geq J(u^*), \forall v \in \mathbb{E} \qquad [\because p = 0 \in \partial J(u^*)]$$

$$J(u^*) \geq 0 \qquad [\text{By definition}]$$

$$\Rightarrow 0 \geq J(u^*) \geq 0 \qquad [\text{for } v \in C,\ J(v) = 0 \text{ by definition}]$$

$$\therefore J(u^*) = 0 \qquad [\text{Sandwich theorem}]$$

$$\Rightarrow u^* \in C \setminus C' \qquad [\text{By definition of } J(u)]$$

$\underline{\hspace{2cm}}$ (i)

Also, for $0 \in \partial \delta_{C'}(u^*) := N_{C'}(u^*)$,
it follows that

$$\because \delta_{C'}(v) \geq \delta_{C'}(u^*) + \langle p, v - u^* \rangle$$ [Definition of subdifferential]

$$\Rightarrow \delta_{C'}(v) \geq \delta_{C'}(u^*), \forall v \in E$$ [$\because p = 0 \in N_{C'}(u^*)$]

$$\Rightarrow \qquad \delta_{C'}(u^*) \geq 0$$ [By definition]

$$\Rightarrow \quad 0 \geq \delta_{C'}(u^*) \geq 0$$ [for $v \in C$, $\delta_C(v) = 0$ by definition]

$$\therefore \delta_{C'}(u^*) = 0$$ [Sandwich theorem]

$$\Rightarrow u^* \in C' \quad \text{---(ii)}$$ [By definition of $\delta_{C'}(u)$]

Hence, from (i) and (ii),

$$\Rightarrow u^* \in (C') \cup (C \setminus C')$$

$$\therefore u^* \in C$$