Lecture 7 01/10/2019

proof: $f^*(y)$ is convex, where $f^*(y) = \max_{x\in\mathbb{R}^n} y^T x - f(x)$

Let's denote $g_x(y) := y^T x - f(x)$. Note that function $g_x(y)$ is linear w.r.t y. Thus $g_x(y)$ is convex w.r.t y. By convexity of $g_x(y)$ we have

$$g_x(tz+(1-t)s) \le t\, g_x(z) + (1-t)\, g_x(s) \quad \forall x\in\mathbb{R}^n$$

Let's take the maximum

$$\max_{x\in\mathbb{R}^n} g_x(tz+(1-t)s) \le \max_{x\in\mathbb{R}^n} t\, g_x(z) + (1-t)\, g_x(s)$$

$$\le t \max_{x\in\mathbb{R}^n} g_x(z) + (1-t) \max_{x\in\mathbb{R}^n} g_x(s)$$

By definition we have that

$$f^*(y) = \max_{x\in\mathbb{R}^n} g_x(y)$$

Thus using the above inequality we get

$$f^*(tz+(1-t)s) \le t f^*(z) + (1-t) f^*(s)$$

$\Rightarrow$ $f^*(y)$ is convex.

Convex conjugate of ℓ_1-norm

$$f^*(y) = \max_x \quad x^T y - \|x\|_1$$

$$= \max_x \sum_{i=1}^n x_i y_i - \sum_{i=1}^n |x_i|$$

$$= \max_x \sum_{i=1}^n x_i y_i - |x_i|$$

$$= \sum_{i=1}^n \max_{x_i} x_i y_i - |x_i|$$

using separability of the function w.r.t x.)

Let's work with $\max_{x_i} \; x_i y_i - |x_i| = \max_{x_i} \; x_i y_i - x_i \, \mathrm{sign}(x_i)$

$$= \max_{x_i} \quad x_i\left(y_i - \mathrm{sign}(x_i)\right)$$

Say $y_i > 1$ then if $x_i = +\infty$ the maximum is $+\infty$

Say $y_i < -1$ then if $x_i = -\infty$ the maximum is $+\infty$.

say $-1 \leq y_i \leq 1$ then maximum is attained for $x_i = 0$ which gives maximum value equal to zero. Thus

$$\max_{x_i} x_i y_i - |x_i| = \begin{cases} 0 & \text{if } |y_i| \leq 1 \\ +\infty & \text{otherwise} \end{cases}$$

Thus

$$f^*(y) = \sum_{i=1}^{n} \max_{x_i} x_i y_i - |x_i| = \begin{cases} 0 & \text{if } \|y\|_\infty \leq 1 \\ +\infty & \text{otherwise.} \end{cases}$$

Proof that : $f(x) - \mu D \leq f_\mu(x) \leq f(x)$

Let's prove the upper bound.

$$f_\mu(x) = \max_{y \in \operatorname{dom} f^*} x^T y - f^*(y) - \mu d(y)$$

$$\leq \max_{y \in \operatorname{dom} f^*} x^T y - f^*(y) + \max_{y \in \operatorname{dom} f^*} -\mu d(y)$$

$$= f(x) + \max_{y \in \operatorname{dom} f^*} -(-\mu d(y))$$

$$= f(x) - \min_{y \in \operatorname{dom} f^*} \mu d(y)$$

But we assumed that $\min_{y \in \operatorname{dom} f^*} d(y) = 0$, thus

we get

$$f_\mu(x) \leq f(x).$$

proof $f(x) - \mu D \le f_\mu(x) \le f(x)$

where $D = \sup_{y \in \mathrm{dom} f^*} d(y)$

We assumed that $\mathrm{dom} f^* \subseteq \mathbb{R}^n$ is closed and bounded. This means that $\mathrm{dom} f^*$ is compact. We also assumed that $d(y)$ is continuous. These assumptions imply that

$$\max_{y \in \mathrm{dom} f^*} d(y) < +\infty \implies D < +\infty.$$ [weierstrauss extreme value theorem]

Let's prove $f(x) - \mu D \le f_\mu(x)$.

$$f_\mu(x) = \max_{y \in \mathrm{dom} f^*} x^T y - f^*(y) - \mu d(y)$$

$$\ge \max_{y \in \mathrm{dom} f^*} x^T y - f^*(y) - \mu \cdot D$$

$$= f(x) - \mu D$$