CS727 HW4

Zhili Feng

November 12, 2018

1. (6.1.19)

   By definition, we want to show that $\text{epi}(h) = \{(x, \mu) \in \mathbb{R}^{n+1} : \inf_{y \in \mathbb{R}^m} g(x, y) \leq \mu\}$ is convex. It suffices to show that $\forall (x_1, \xi_1), (x_2, \xi_2) \in \text{epi}(h), \forall \lambda \in [0, 1]$, we have $(\lambda x_1 + (1-\lambda)x_2, \lambda\xi_1 + (1-\lambda)\xi_2) \in \text{epi}(h)$. Equivalently, this is saying that $h(\lambda x_1 + (1-\lambda)x_2) = \inf_{y \in \mathbb{R}^m} g(\lambda x_1 + (1-\lambda)x_2, y) \leq \lambda\xi_1 + (1-\lambda)\xi_2$.

   By definition of infimum, for any $\epsilon > 0$, there exists $y_1, y_2 \in \mathbb{R}^m$ such that $g(x_1, y_1) \leq \xi_1 + \epsilon$ and $g(x_2, y_2) \leq \xi_2 + \epsilon$. That is $((x_1, y_1), \xi_1 + \epsilon), ((x_2, y_2), \xi_2 + \epsilon) \in \text{epi}(g)$. Since g is a convex function, then epi(g) is a convex set. So for all $\lambda \in [0, 1]$, $\Big(\big(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2\big), \lambda\xi_1 + (1-\lambda)\xi_2 + \epsilon\Big) \in \text{epi}(g) \iff g(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2) \leq \lambda\xi_1 + (1-\lambda)\xi_2 + \epsilon$.

   OTOH, $h(\lambda x_1 + (1-\lambda)x_2) = \inf_{y \in \mathbb{R}^m} g(\lambda x_1 + (1-\lambda)x_2, y) \leq g(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2) \leq \lambda\xi_1 + (1-\lambda)\xi_2 + \epsilon$. Let $\epsilon \to 0$, we derive $h(\lambda x_1 + (1-\lambda)x_2) \leq \lambda\xi_1 + (1-\lambda)\xi_2$, which finishes the proof.

2. (6.1.22)

   Pick arbitrary $x, y \in \text{dom}\phi$, and pick arbitrary ξ, μ be such that $\phi(x) < \xi$ and $\phi(y) < \mu$. Use proposition 6.1.3, it suffice to show that $\forall \lambda \in (0, 1)$, we have:

   $$\phi(\lambda x + (1-\lambda)y) < \lambda\xi + (1-\lambda)\mu$$

   Pick $(x, \alpha), (y, \beta) \in C$ in the following way: if $(x, \phi(x)) \in C$, then let $\alpha = \phi(x)$. Otherwise pick $\alpha = \phi(x) + \epsilon < \xi$ for small ϵ, such that $(x, \alpha) \in C$. This is doable by definition of ϕ and $\phi(x)$ is strictly less than ξ. Pick (y, β) in a similar fashion.

   Due to the convexity of C, we have $(\lambda x + (1-\lambda)y, \lambda\alpha + (1-\lambda)\beta) \in C$. By the definition of ϕ, we also have $\phi(\lambda x + (1-\lambda)y) \leq \lambda\alpha + (1-\lambda)\beta$. Simple algebra also shows that $\lambda\alpha + (1-\lambda)\beta < \lambda\xi + (1-\lambda)\mu$. These two inequalities give us $\phi(\lambda x + (1-\lambda)y) < \lambda\xi + (1-\lambda)\mu$, which finishes the proof.