Supplementary Material

1 Proofs in Section 2

Before we prove Proposition 1, let us recall the definition of star-convexity and show a lemma.

Definition 1 *(Star-convex functions). A function $f : \mathbf{R}^n \to \mathbf{R}$ is star-convex if there is $x^* \in \operatorname*{argmin}_{x\in\mathcal{X}} f(x)$ such that for all $\alpha \in [0,1]$ and $x \in \mathcal{X}$,*

$$f((1-\alpha)x^* + \alpha x) \leq (1-\alpha)f(x^*) + \alpha f(x). \quad (1)$$

The following lemma characterizes the differentiable star-convex functions.

Lemma 1 *For a differentiable function f, the star convexity condition (1) is equivalent to the following condition*

$$f(x) - f(x^*) \leq \nabla f(x)^\top (x - x^*), \quad (2)$$

where $x^* = \operatorname*{argmin}_{x\in\mathcal{X}} f(x)$.

Proof. Suppose (1) holds. Then we have

$$f(x) - f(x^*) \leq \frac{f(x) - f((1-\alpha)x^* + \alpha x)}{1-\alpha}, \quad (3)$$

for all $\alpha \in [0,1]$. Note that

$$\lim_{\alpha\to 1^-} \frac{f(x) - f((1-\alpha)x^* + \alpha x)}{1-\alpha} = \nabla f(x)^\top (x - x^*),$$

which implies (2). Conversely, suppose that (2) holds. Let us denote

$$d(\alpha) := f((1-\alpha)x^* + \alpha x) - f(x^*).$$

Clearly, (1) is equivalent to

$$d(\alpha) \leq \alpha d(1), \text{ for all } 0 \leq \alpha \leq 1. \quad (4)$$

It remains to show that if f is differentiable then (2) implies (4). In fact, (2) leads to

$$f((1-\alpha)x^* + \alpha x) - f(x^*) \leq \alpha \nabla f((1-\alpha)x^* + \alpha x)^\top (x - x^*),$$

or,

$$d(\alpha) \leq \alpha d'(\alpha).$$

Hence,

$$\left(\frac{d(\alpha)}{\alpha}\right)' = \frac{\alpha d'(\alpha) - d(\alpha)}{\alpha^2} \geq 0,$$

for all $0 < \alpha \leq 1$, implying that $\frac{d(\alpha)}{\alpha}$ is a nondecreasing function for $\alpha \in (0,1]$. Therefore,

$$\frac{d(\alpha)}{\alpha} \leq \frac{d(1)}{1},$$

which proves (4) for $\alpha \in (0,1]$. Since $d(0) = f(x^*) = 0$, (4) in fact holds for all $\alpha \in [0,1]$. □

Proposition 1 *If $f(\cdot)$ is star-convex and smooth with bounded gradient in $\mathcal{X}$, then $f(\cdot)$ is weakly pseudo-convex.*

Proof: From Lemma 1, we have

$$\begin{aligned} f(x) - f(x^*) &\leq \nabla f(x)^\top \|x - x^*\| \\ &\leq M \frac{\nabla f(x)^\top (x - x^*)}{\|\nabla f(x)\|} \end{aligned}$$

where the last inequality is due to the bounded gradient condition $\|\nabla f(x)\| \leq M$ for $x \in \mathcal{X}$. □

Proposition 2 *If $f(\cdot)$ has bounded gradient and satisfies the acute angle condition, then $f(\cdot)$ is weakly pseudo-convex.*

Proof: For all $x \in \mathcal{X}$, we have

$$\begin{aligned} f(x) - f(x^*) &\leq M\|x - x^*\| \\ &\leq \frac{M}{Z} \frac{\nabla f(x)^\top (x - x^*)}{\|\nabla f(x)\|} \end{aligned}$$

where the first inequality follows from the bounded gradient assumption while the second inequality is due to the acute angle condition. □

Proposition 3 *If $f(\cdot)$ has bounded gradient and satisfy the α-homogeneity with respect to its minimum, i.e., there exists $\alpha > 0$ satisfying*

$$f(t(x - x^*) + x^*) - f(x^*) = t^\alpha (f(x) - f(x^*)),$$

for all $x \in \mathcal{X}$ and $t \geq 0$ where $x^* = \arg\min_{x\in\mathcal{X}} f(x)$, then $f(\cdot)$ is weak pseudo-convex.

Proof: By taking the derivative of the equation (3) with respective to t and letting $t = 1$, we have

$$\nabla f(x)^\top (x - x^*) = \alpha(f(x) - f(x^*)).$$

Therefore, we have

$$\begin{aligned} f(x) - f(x^*) &= \frac{1}{\alpha} \nabla f(x)^\top (x - x^*) \\ &\leq \frac{M}{\alpha} \frac{\nabla f(x)^\top (x - x^*)}{\|\nabla f(x)\|}, \end{aligned}$$

which satisfies the weak pseudo-convexity condition with $K = \frac{M}{\alpha}$. □