CS727 HW6

Zhili Feng

December 5, 2018

1. (8.2.5)

   (a) Let $f(x) = |x|$. Then $f : \mathbb{R} \to \bar{\mathbb{R}}$ is a proper closed convex function. Let $x_0 = 0$, then it is clear that x_0 is a unique global minimizer of f, and we have $\partial f(x_0) = [-1, 1]$. Let's pick any $x_0^* \in \partial f(x_0)$ that is non-zero, say $x_0^* = -1/2$. Then according to the statement, $-x_0^* = 1/2$ should be a descent direction, i.e:

   $$\exists \mu < 0, \text{ such that for all small } t > 0, f(t/2) \leq f(0) + \frac{\mu t}{2} = \frac{\mu t}{2}$$

   But OTOH, $f(t/2) = t/2 > \frac{\mu t}{2}$ for any $\mu < 0$. Therefore we reach to a contradiction.

   (b) We first show that the min norm solution is attained (notice that closest to the origin in Euclidean metric means this point has the smallest Euclidean norm)
   We are trying to pick x_0^* such that:

   $$\begin{aligned} \min_x \quad & \|x\| \\ \text{s.t.} \quad & x \in \partial f(x_0) \end{aligned}$$

   According to proposition 8.1.2, $\partial f(x_0)$ is closed and convex. We claim that the above minimum is actually attained. Since $x_0 \in \text{ridom} f$, by proposition 8.1.3, we know that $\text{ridom} f \subset \text{dom}\partial f$, so $\partial f(x_0)$ is non-empty and we can pick any $x \in \partial f(x_0)$, let $\|x\| = t < \infty$. Consider the following auxiliary optimization problem:

   $$\begin{aligned} \min_x \quad & \|x\| \\ \text{s.t.} \quad & x \in \partial f(x_0), \\ & \|x\| \leq t \end{aligned}$$

   The feasible set is compact because it's closed and bounded. Since $\|\cdot\|$ is a continuous function, we know it attains its minimum on the feasible set at some point x_0^*. Note that x_0^* also minimizes the original optimization problem. Since x_0 does not minimize f (globally), we know that $0 \notin \partial f(x_0)$. So we know that $x_0^* \neq 0$. Therefore the question's statement "one chooses the element x_0^* closest to the origin in the Euclidean norm" makes sense.

   Claim 1. The directional derivative of convex f can be written as:

   $$f'(x; v) = \sup_{g \in \partial f(x)} \langle g, v \rangle$$

   Proof. By proposition 8.2.2, since f is convex and $x \in \text{ridom} f$, $f'(x; v)$ is closed and proper, specifically we have $f'(x; v) = \text{cl} f'(x; v)$. Then by theorem 8.2.3, we have that $\text{cl} f'(x; v)$ is a support function of $\partial f(x)$. By definition, that means:

   $$f'(x; v) = \text{cl} f'(x; v) = \sup_{g \in \partial f(x)} \langle g, v \rangle$$

   $\square$

Substitute $x = x_0$ and $v = -x_0^*$ we get:

$$f'(x_0; -x_0^*) = \inf_{t>0} \frac{1}{t}[f(x_0 - tx_0^*) - f(x_0)] = \sup_{g \in \partial f(x_0)} \langle g, -x_0^* \rangle = - \inf_{g \in \partial f(x_0)} \langle g, x_0^* \rangle$$

Now, we show that $\inf_{g \in \partial f(x_0)} \langle g, x_0^* \rangle$ is attained. In fact, the infimum is attained at $g = x_0^*$. Assume otherwise that the infimum is attained at $y \in \partial f(x_0)$ and $\langle y, x_0^* \rangle < \|x_0^*\|^2$. By proposition 8.1.2, $\partial f(x_0)$ is convex. Hence for any $\lambda \in [0, 1]$, we have $h = x_0^* + \lambda(y - x_0^*) \in \partial f(x_0)$. Then:

$$\|h\|^2 = \|x_0^*\|^2 + \lambda^2 \|y - x_0^*\|^2 + 2\lambda \langle x_0^*, y - x_0^* \rangle$$

Based on how we define y, we have $\langle x_0^*, y \rangle < \langle x_0^*, x_0^* \rangle$, so $\langle x_0^*, y - x_0^* \rangle < 0$. Then for small enough λ, we would have $\lambda \|y - x_0^*\|^2 + 2 \langle x_0^*, y - x_0^* \rangle < 0$ therefore $\lambda^2 \|y - x_0^*\|^2 + 2\lambda \langle x_0^*, y - x_0^* \rangle < 0$. This leads to $\|h\|^2 < \|x_0^*\|^2$, which contradicts the fact that $x_0^* \in \partial f(x_0)$ has the smallest norm.
Hence we conclude

$$f'(x_0; -x_0^*) = \inf_{t>0} \frac{1}{t}[f(x_0 - tx_0^*) - f(x_0)] = -\|x_0^*\|^2$$

Note that $\frac{1}{t}[f(x_0 - tx_0^*) - f(x_0)]$ is non-decreasing in t (lemma 7.2.3 and definition 8.2.1). Pick $\epsilon = \frac{1}{2}\|x_0^*\|^2$. Then by definition of infimum, $\exists t' > 0$ such that

$$t'^{-1}[f(x_0 - t'x_0^*) - f(x_0)] \leq -\|x_0^*\|^2 + \frac{1}{2}\|x_0^*\|^2 \Longrightarrow f(x_0 - t'x_0^*) \leq f(x_0) - \frac{t}{2}\|x_0^*\|^2$$

and due to the monotonicity of the difference quotient in t, we have $\forall t \leq t'$:

$$t^{-1}[f(x_0 - tx_0^*) - f(x_0)] \leq -\|x_0^*\|^2 + \frac{1}{2}\|x_0^*\|^2 \Longrightarrow f(x_0 - tx_0^*) \leq f(x_0) - \frac{t}{2}\|x_0^*\|^2$$

Let $\mu = -\frac{\|x_0^*\|}{2}$, the above inequality then proves the statement.