Suggested Solutions to Assignment 2

1.

2.

3. (a)

   (b)

   (c)

4. Notice the right derivative function $\varphi'(\cdot;1)$ is increasing on the open interval. Define

$$\mathcal{I}_1 = \{x : \varphi \text{ is not differentiable at } x\}, \mathcal{I}_2 = \{x : \varphi'(\cdot;1) \text{ is not continuous at } x\},$$

think of why $\mathcal{I}_1 \subset \mathcal{I}_2$ and $\mathcal{I}_2$ is countable.

5.

6. For $i = 1, 2, \cdots, n$, let e_i denote the n-length column vector whose ith component is one while all the others are zeros. The n-length column vector of all ones is denoted by e. Fix $x \in \mathbb{R}^n$ arbitrarily, notice that there exists $i^* \in \{1, 2, \cdots, n\}$ with

$$x_{i^*} = \max\{x_1, x_2, \cdots, x_n\}.$$

On the one hand, for any $i \in \{1, 2, \cdots, n\}$,

$$x_i = e_i^T x \le \max_{y \in \Delta_n} y^T x \qquad \text{since } e_i \in \Delta_n.$$

Then

$$\max\{x_1, x_2, \cdots, x_n\} \le \max_{y \in \Delta_n} y^T x.$$

On the other hand, for any $y = (y_1, y_2, \cdots, y_n) \in \Delta_n$,

$$y^T x = y^T [x - x_{i^*} e] + y^T [x_{i^*} e] \le x_i^* = \max\{x_1, x_2, \cdots, x_n\},$$

where the inequality above holds by $\cancel{y^T = e}$ $y^T e = 1$ and $y_i \ge 0$, $x_i - x_{i^*} e \le 0$ for any $i = 1, 2, \cdots, n$. Then

$$\max_{y \in \Delta_n} y^T x \le \max\{x_1, x_2, \cdots, x_n\}.$$

7. (a)

   (b) Fix $d \in \mathbb{R}^n$ arbitrarily. Notice that for any $\varepsilon > 0$, $\exists \delta > 0$ such that

$$\left| \frac{f(x + \alpha d') - f(x)}{\alpha} - f_H'(x; d) \right| < \frac{\varepsilon}{2} \text{ if } 0 < \alpha < \delta \text{ and } \|d' - d\| < \delta.$$

For any d^* with $\|d^* - d\| < \frac{\delta}{2}$, notice that for the same ε again, $\exists \delta^* > 0$ such that

$$\left| \frac{f(x + \alpha d') - f(x)}{\alpha} - f_H'(x; d^*) \right| < \frac{\varepsilon}{2} \text{ if } 0 < \alpha < \delta^* \text{ and } \|d' - d^*\| < \delta^*.$$

Let $\delta_{\min} = \min\left\{\delta^*, \frac{\delta}{2}\right\}$, then for any α, d' with $0 < \alpha < \delta_{\min}$ and $\|d' - d^*\| < \delta_{\min}$ (so $\|d' - d\| < \delta$ meanwhile), it holds that

$$\left|\frac{f\left(x + \alpha d'\right) - f(x)}{\alpha} - f'_H\left(x; d^*\right)\right| < \frac{\varepsilon}{2},$$
$$\left|\frac{f\left(x + \alpha d'\right) - f(x)}{\alpha} - f'_H\left(x; d\right)\right| < \frac{\varepsilon}{2}.$$

Then

$$\left|f'_H\left(x; d^*\right) - f'_H\left(x; d\right)\right| \leq \left|f'_H\left(x; d^*\right) - \frac{f\left(x + \alpha d'\right) - f(x)}{\alpha}\right| + \left|\frac{f\left(x + \alpha d'\right) - f(x)}{\alpha} - f'_H\left(x; d\right)\right| < \varepsilon.$$

In conclusion, fix $d \in \mathbb{R}^n$ arbitrarily, for any $\varepsilon > 0$, $\exists \delta > 0$ such

$$\left|f'_H\left(x; d^*\right) - f'_H\left(x; d\right)\right| \leq \varepsilon \text{ if } \|d^* - d\| < \frac{\delta}{2}.$$

(c) Choose $\delta_1 > 0$ such that $\mathbb{B}_{\delta_1}(x) \subset V$.

i. Fix d arbitrarily. For any $\varepsilon > 0$, $\exists \delta_2 > 0$ such that

$$\left|\frac{f(x + \alpha d) - f(x)}{\alpha} - f'(x; d)\right| < \frac{\varepsilon}{2} \text{ if } 0 < \alpha < \delta_2.$$

For the same ε again, $\exists \delta_3 = \frac{\delta_1}{\frac{\varepsilon}{2L_f} + \|d\|}$ such that

$$\left|\frac{f(x + \alpha d') - f(x)}{\alpha} - \frac{f(x + \alpha d) - f(x)}{\alpha}\right| \leq L_f \|d' - d\| < \frac{\varepsilon}{2}$$
$$\text{if } 0 < \alpha < \delta_3, \|d' - d\| \leq \frac{\varepsilon}{2L_f}.$$

Hence for the same ε, let $\delta_4 = \min\left\{\delta_2, \delta_3, \frac{\varepsilon}{2L_f}\right\}$, we have for any α, d' with $0 < \alpha < \delta_4$ and $\|d' - d\|$, it holds that

$$\begin{aligned}
&\left|\frac{f(x + \alpha d') - f(x)}{\alpha} - f'(x; d)\right| \\
\leq &\left|\frac{f(x + \alpha d') - f(x)}{\alpha} - \frac{f(x + \alpha d) - f(x)}{\alpha}\right| + \left|\frac{f(x + \alpha d) - f(x)}{\alpha} - f'(x; d)\right| \\
< &\varepsilon.
\end{aligned}$$

In conclusion, $f'_H\left(x; d\right)$ exists. Actually, $f'_H\left(x; d\right) = f'\left(x; d\right)$.

ii. For any $d_1, d_2 \in \mathbb{R}^n$,

$$\begin{aligned}
&\left|f'_H(x; d_1) - f'_H(x; d_2)\right| \\
= &\left|f'(x; d_1) - f'(x; d_2)\right| \\
= &\left|\lim_{\alpha \to 0^+} \frac{f(x + \alpha d_1) - f(x)}{\alpha} - \lim_{\alpha \to 0^+} \frac{f(x + \alpha d_2) - f(x)}{\alpha}\right| \\
\leq &\lim_{\alpha \to 0^+} \left|\frac{f(x + \alpha d_1) - f(x)}{\alpha} - \frac{f(x + \alpha d_2) - f(x)}{\alpha}\right| \\
\leq &L_f \|d_1 - d_2\|.
\end{aligned}$$

8. (a)

(b)

$$\begin{aligned}
&F \text{ is convex} \\
\Leftrightarrow &F(\cdot) + \Omega \text{ is convex} \\
\Leftrightarrow &\forall x, z \text{ and } 0 \leq \lambda \leq 1, \lambda\left[F(x) + \Omega\right] + (1 - \lambda)\left[F(z) + \Omega\right] \subset F\left(\lambda x + (1 - \lambda) z\right) + \Omega \\
\Leftrightarrow &\forall x, z \text{ and } 0 \leq \lambda \leq 1, \lambda F(x) + (1 - \lambda) F(z) + \Omega \subset F\left(\lambda x + (1 - \lambda) z\right) + \Omega \\
\Leftrightarrow &\forall x, z \text{ and } 0 \leq \lambda \leq 1, \lambda F(x) + (1 - \lambda) F(z) - F\left(\lambda x + (1 - \lambda) z\right) + \Omega \subset \Omega \\
\overset{(\star)}{\Leftrightarrow} &\lambda F(x) + (1 - \lambda) F(z) - F\left(\lambda x + (1 - \lambda) z\right) \in R\left(\Omega\right),
\end{aligned}$$

where $(\star)$ holds by T1(b), Midterm Exam.

9. For any $g_x \in \partial f(x)$ and $g_z \in \partial f(z)$,

$$\begin{aligned} f(x) &\geq f(z) + \langle g_z, x - z \rangle; \\ f(z) &\geq f(x) + \langle g_x, z - x \rangle. \end{aligned}$$

Summing the two inequality,

$$f(x) + f(z) \geq f(z) + f(x) + \langle g_z, x - z \rangle + \langle g_x, z - x \rangle,$$

i.e.,

$$0 \geq \langle g_z, x - z \rangle - \langle g_x, x - z \rangle,$$

i.e,

$$0 \leq -\langle g_z, x - z \rangle + \langle g_x, x - z \rangle = \langle g_x - g_z, x - z \rangle.$$

10.

11. Let $z = \mathcal{A}(x)$, consider the problem

$$\begin{aligned} \min_{x \in \mathbb{R}^m} \quad & f(x) + g(z), \\ \text{s.t.} \quad & \mathcal{A}(x) = z. \end{aligned}$$

The Lagrangian function:

$$\begin{aligned} L(x, z; y) &= f(x) + g(z) + \langle y, z - \mathcal{A}(x) \rangle \\ &= \langle y, -\mathcal{A}(x) \rangle + f(x) + \langle y, z \rangle + g(z) \\ &= -\big(\langle -\mathcal{A}^T y, x \rangle - f(x)\big) - (\langle -y, z \rangle - g(z)) \end{aligned}$$

The dual function is

$$\begin{aligned} h(y) &= \inf_{x \in \mathbb{R}^m, z \in \mathbb{R}^n} L(x, z; y) \\ &= -\sup_{x \in \mathbb{R}^m, z \in \mathbb{R}^n} \big(\langle -\mathcal{A}^T y, x \rangle - f(x)\big) - (\langle -y, z \rangle - g(z)) \\ &= -\left(f^* \circ \mathcal{A}^T\right)(y) - g^*(-y). \end{aligned}$$

Hence the dual problem is

$$\max_{y \in \mathbb{R}^n} h(y) = -\left(f^* \circ \mathcal{A}^T\right)(y) - g^*(-y).$$

12. (a) $(\Rightarrow)$ If $\varphi(\cdot, u)$ is proper, then $\exists x^*$ such that $\varphi(x^*, u) < \infty$, then

$$\nu(u) = \inf_{x \in \mathbb{R}^n} \varphi(x, u) \leq \varphi(x^*, u) < \infty.$$

$(\Leftarrow)$ $\varphi(\cdot, u) > -\infty$ is obvious. On the other hand, $\nu(u) < \infty$ implies that $\exists x$ with

$$\varphi(x, u) \leq \nu(u) + 1 < \infty.$$

(b) Notice

$$\begin{aligned} \langle w, u \rangle - \nu(u) =& \langle 0, x \rangle + \langle w, u \rangle - \inf_{z \in \mathbb{R}^n} \varphi(z, u) \text{ for } \forall x \in \mathbb{R}^n \\ =& \sup_{z \in \mathbb{R}^n} \left\{ \langle 0, x \rangle + \langle w, u \rangle - \varphi(z, u) \right\} \text{ for } \forall x \in \mathbb{R}^n \\ =& \sup_{z \in \mathbb{R}^n} \left\{ \langle 0, z \rangle + \langle w, u \rangle - \varphi(z, u) \right\}. \end{aligned}$$

Then

$$
\begin{aligned}
\nu^*(w) = &\sup_{u\in\mathbb{R}^m} \{\langle w, u\rangle - \nu(u)\} \\
= &\sup_{u\in\mathbb{R}^m} \left\{ \sup_{z\in\mathbb{R}^n} \{\langle 0, z\rangle + \langle w, u\rangle - \varphi(z,u)\} \right\} \\
= &\sup_{u\in\mathbb{R}^m, z\in\mathbb{R}^n} \langle 0, z\rangle + \langle w, u\rangle - \varphi(z,u) \\
= &\varphi^*\left(0, w\right).
\end{aligned}
$$

(c)

$$
\begin{aligned}
\text{The optimal value of } (D_u) = &\max_{w\in\mathbb{R}^m} \{\langle w, u\rangle - \varphi^*(0,w)\} \\
= &\max_{w\in\mathbb{R}^m} \{\langle w, u\rangle - \nu^*(w)\} \\
= &\nu^{**}(u) \le \nu(u).
\end{aligned}
$$

13.

14. Define $\varphi(t) = f(x + t(y-x))$, then $\varphi(0) = f(x)$, $\varphi(1) = f(y)$, $\varphi'(t) = \langle \nabla f(x + t(y-x)), y - x\rangle$ and $\varphi''(t) = (y-x)^T \nabla^2 f(x+t(y-x))(y-x)$. Recall the Taylor formula with integral remainder on φ at $t = 0$,

$$
\varphi(1) = \varphi(0) + \varphi'(0) + \int_0^1 (1-t)\varphi''(t)dt.
$$

That is to say,

$$
\begin{aligned}
f(y) = &f(x) + \langle \nabla f(x), y - x\rangle + \int_0^1 (1-t)(y-x)^T \nabla^2 f(x + t(y-x))(y-x)dt \\
= &f(x) + \langle \nabla f(x), y - x\rangle + \int_0^1 (1-t)(y-x)^T \nabla^2 f(x)(y-x)dt \\
&+ \int_0^1 (1-t)(y-x)^T \left[\nabla^2 f(x + t(y-x))(y-x) - \nabla^2 f(x)\right](y-x)dt \\
\le &f(x) + \langle \nabla f(x), y - x\rangle + (y-x)^T \nabla^2 f(x)(y-x) \int_0^1 (1-t)dt \\
&+ \int_0^1 (1-t)(y-x)^T \left\|\nabla^2 f(x + t(y-x))(y-x) - \nabla^2 f(x)\right\|(y-x)dt \\
\le &f(x) + \langle \nabla f(x), y - x\rangle + \frac{1}{2}(y-x)^T \nabla^2 f(x)(y-x) + L_f \left\|y - x\right\|^3 \int_0^1 (1-t)tdt \\
= &f(x) + \langle \nabla f(x), y - x\rangle + \frac{1}{2}(y-x)^T \nabla^2 f(x)(y-x) + \frac{L_f}{6} \left\|y - x\right\|^3.
\end{aligned}
$$