DATA130026 Optimization
Solution of Assignment 13

1. For each of the following functions on $\mathbb{R}^n$, explain how to calculate a subgradient at a given x.

   (a) $f(x) = \sup_{0 \le t \le 1} p(t)$, where $p(t) = x_1 + x_2 t + \ldots + x_n t^{n-1}$.

   (b) $f(x) = x_{[1]} + x_{[2]} + \ldots + x_{[k]}$, where $x_{[i]}$ denotes the ith largest elements of x.

   (c) $f(x) = \|Ax - b\|_2 + \|x\|_2$ where $A \in \mathbb{R}^{m \times n}$.

   Solution.

   (a) Define the active set $T := \{t \mid p(t) = f(x)\}$, then the subdifferential of f at x is

   $$\partial f(x) = \operatorname{conv} \bigcup \{\partial_x p(t) \mid t \in T\}.$$

   $\forall\, \tilde{t} \in T$, we have

   $$\partial_x p(\tilde{t}) = (1, \tilde{t}, \ldots, \tilde{t}^{n-1})^T.$$

   Therefore, we can easily obtain a subgradient

   $$g(x) = (1, \tilde{t}, \ldots, \tilde{t}^{n-1})^T.$$

   (b) Define $A_k := \left\{a \in \mathbb{R}^n \mid a_i \in \{0,1\}, i = 1, \ldots, n,\ \sum_{i=1}^n a_i = k\right\}$. Then

   $$f(x) = \sup_{a \in A_k} a^T x.$$

   Define the active set $\tilde{A} := \{a \in A_k \mid a^T x = f(x)\}$, then the subdifferential of f at x is

   $$\partial f(x) = \operatorname{conv} \bigcup \{\tilde{a} \mid \tilde{a} \in \tilde{A}\}.$$

   Therefore, we can easily obtain a subgradient $g(x) = \tilde{a}$.
   Moreover, $\tilde{a}$ satisfies

   $$\tilde{a}_i = \begin{cases} 1, & \text{if } x_i \in \{x_{[1]}, \ldots, x_{[k]}\}, \\ 0, & \text{if } x_i \notin \{x_{[1]}, \ldots, x_{[k]}\}. \end{cases}$$

   (c) We know that

   $$\partial \|Ax - b\|_2 = \begin{cases} \dfrac{A^T(Ax - b)}{\|Ax - b\|_2}, & \text{if } Ax - b \neq 0, \\ \{A^T g \mid \|g\|_2 \le 1\}, & \text{if } Ax - b = 0. \end{cases}$$

   $$\partial \|x\|_2 = \begin{cases} \dfrac{x}{\|x\|_2}, & \text{if } x \neq 0, \\ \{h \mid \|h\|_2 \le 1\}, & \text{if } x = 0. \end{cases}$$

Then the subdifferential of f at x is

$$\partial f(x) = \begin{cases} \dfrac{A^T(Ax-b)}{\|Ax-b\|_2} + \dfrac{x}{\|x\|_2}, & \text{if } Ax-b \neq 0 \text{ and } x \neq 0, \\ \left\{ \dfrac{A^T(Ax-b)}{\|Ax-b\|_2} + h \;\middle|\; \|h\|_2 \leq 1 \right\}, & \text{if } Ax-b \neq 0 \text{ and } x = 0, \\ \left\{ A^T g + \dfrac{x}{\|x\|_2} \;\middle|\; \|g\|_2 \leq 1 \right\}, & \text{if } Ax-b = 0 \text{ and } x \neq 0, \\ \left\{ A^T g + h \mid \|g\|_2 \leq 1, \|h\|_2 \leq 1 \right\}, & \text{if } Ax-b = 0 \text{ and } x = 0. \end{cases}$$

Therefore, we can easily obtain a subgradient by setting $g = h = 0$:

$$g(x) = \begin{cases} \dfrac{A^T(Ax-b)}{\|Ax-b\|_2} + \dfrac{x}{\|x\|_2}, & \text{if } Ax-b \neq 0 \text{ and } x \neq 0, \\ \dfrac{A^T(Ax-b)}{\|Ax-b\|_2}, & \text{if } Ax-b \neq 0 \text{ and } x = 0, \\ \dfrac{x}{\|x\|_2}, & \text{if } Ax-b = 0 \text{ and } x \neq 0, \\ 0, & \text{if } Ax-b = 0 \text{ and } x = 0. \end{cases}$$

2. (subgradient of the maximum eigenvalue function). Consider the function $f : \mathbb{S}^n \to \mathbb{R}$ given by $f(\mathbf{X}) = \lambda_{\max}(\mathbf{X})$ (recall that $\mathbb{S}^n$ is the set of all $n \times n$ symmetric matrices). Let $\mathbf{X} \in \mathbb{S}^n$ and let $\mathbf{v}$ be a normalized eigenvector of $\mathbf{X}$ ($\|\mathbf{v}\|_2 = 1$) associated with the maximum eigenvalue of $\mathbf{X}$. Show that

$$\mathbf{v}\mathbf{v}^T \in \partial f(\mathbf{X}) \tag{1}$$

Solution. To show this, note that for any $\mathbf{Y} \in \mathbb{S}^n$

$$\begin{aligned} \lambda_{\max}(\mathbf{Y}) &= \max_{\mathbf{u}} \left\{ \mathbf{u}^T \mathbf{Y} \mathbf{u} : \|\mathbf{u}\|_2 = 1 \right\} \\ &\geq \mathbf{v}^T \mathbf{Y} \mathbf{v} \\ &= \mathbf{v}^T \mathbf{X} \mathbf{v} + \mathbf{v}^T (\mathbf{Y} - \mathbf{X}) \mathbf{v} \\ &= \lambda_{\max}(\mathbf{X}) \|\mathbf{v}\|_2^2 + \operatorname{Tr}\left(\mathbf{v}^T (\mathbf{Y} - \mathbf{X}) \mathbf{v} \right) \\ &= \lambda_{\max}(\mathbf{X}) + \operatorname{Tr}\left(\mathbf{v}\mathbf{v}^T (\mathbf{Y} - \mathbf{X}) \right) \\ &= \lambda_{\max}(\mathbf{X}) + \left\langle \mathbf{v}\mathbf{v}^T, \mathbf{Y} - \mathbf{X} \right\rangle \end{aligned}$$

establishing (1).