1.1 (a) For each of the following convex functions, determine the subdifferential set at the specified point.

i. \( f(x_1, x_2, x_3) = \max\{|x_1|, |x_2|, |x_3|\} \) at \((x_1, x_2, x_3) = (0, 0, 0)\).

ii. \( f(x) = e^{\|x\|} \) at \( x = 0 \) (\( x \) is a scalar).

iii. \( f(x_1, x_2) = \max\{x_1 + x_2 - 1, x_1 - x_2 + 1\} \) at \((x_1, x_2) = (1, 1)\).

(b) For each of the following convex functions, explain how to calculate a subgradient at a given \( x \).

i. \( f(x) = \max_{i=1,\ldots,m} (a_i^T x + b_i) \).

ii. \( f(x) = \max_{i=1,\ldots,m} |a_i^T x + b_i| \).

iii. \( f(x) = \max_{i=1,\ldots,m} (-\log (a_i^T x + b_i)) \). You may assume \( x \) is in the domain of \( f \).

iv. \( f(x) = \sup_{0 \leq t \leq 1} p(t) \), where \( p(t) = x_1 + x_2 t + \cdots + x_n t^{n-1} \).

v. \( f(x) = x_{[1]} + \cdots + x_{[k]} \), where \( x_{[i]} \) denotes the \( i \)th largest element of the vector \( x \).

vi. \( f(x) = \inf_{A y \preceq b} \|x - y\|^2 \), \( i.e., \) the square of the distance of \( x \) to the polyhedron defined by \( A y \preceq b \). You may assume that the inequalities \( A y \preceq b \) are strictly feasible.

vii. \( f(x) = \sup_{A y \preceq b} y^T x \), \( i.e., \) the optimal value of an LP as a function of the cost vector. (You can assume that the polyhedron defined by \( A y \preceq b \) is bounded.)

1.2 Convex functions that are not subdifferentiable. Verify that the following functions, defined on the interval \([0, \infty)\), are convex, but not subdifferentiable at \( x = 0 \).

(a) \( f(0) = 1, \text{ and } f(x) = 0 \) for \( x > 0 \).

(b) \( f(x) = -\sqrt{x} \).

1.5 Subgradient optimality conditions for nondifferentiable inequality constrained optimization. Consider the problem

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq 0, \quad i = 1, \ldots, m,
\end{align*}
\]

with variable \( x \in \mathbb{R}^n \). We do not assume that \( f_0, \ldots, f_m \) are convex. Suppose that \( \tilde{x} \) and \( \tilde{\lambda} \geq 0 \) satisfy primal feasibility,

\[
\begin{align*}
f_i(\tilde{x}) & \leq 0, \quad i = 1, \ldots, m,
\end{align*}
\]

\[
\begin{align*}
f_0(\tilde{x}) & \leq \sup_{A y \preceq b} y^T \tilde{x}.
\end{align*}
\]
dual feasibility,
\[ 0 \in \partial f_0(\bar{x}) + \sum_{i=1}^{m} \tilde{\lambda}_i \partial f_i(\bar{x}), \]
and the complementarity condition
\[ \tilde{\lambda}_i f_i(\bar{x}) = 0, \quad i = 1, \ldots, m. \]
Show that \( \bar{x} \) is optimal, using only a simple argument, and definition of subgradient. Recall that we do not assume the functions \( f_0, \ldots, f_m \) are convex.

1.9 Conjugacy and subgradients. In this question, we show how conjugate functions are related to subgradients. Let \( f \) be convex and recall that its conjugate is \( f^*(v) = \sup_x \{ v^T x - f(x) \} \). Prove the following:

(a) For any \( v \) we have \( v^T x \leq f(x) + f^*(v) \) (this is sometimes called Young’s inequality).
(b) We have \( g^T x = f(x) + f^*(g) \) if and only if \( g \in \partial f(x) \).

Note that (you do not need to prove this) if \( f \) is closed, so that \( f(x) = f^{**}(x) \), result (b) implies the duality relationship that \( g \in \partial f(x) \) if and only if \( x \in \partial f^*(g) \) if and only if \( g^T x = f(x) + f^*(g) \).

1.10 If a function has a unique subgradient at a given point, is the function differentiable at that point? Provide a proof or a counter example.

1.12 Consider the function \( f : \mathbb{R}^2 \rightarrow \mathbb{R} \) given by
\[ f(x_1, x_2) = \max \left\{ \frac{1}{2} \|x\|^2 - x_1, \frac{1}{2} \|x\|^2 + x_1 \right\} \]

(a) Determine the subdifferential set \( \partial f(x) \) for \( x \in \mathbb{R}^2 \).
(b) Are the subgradients uniformly bounded over \( x \in \mathbb{R}^2 \)? Would your answer change if \( x \) is restricted to lie in the set \( X = \{ x \in \mathbb{R}^2 \mid \|x\| \leq 1 \} \)? If yes, provide a bound for the subgradient norms.

2.3 Matrix norm approximation. We consider the problem of approximating a given matrix \( B \in \mathbb{R}^{p \times q} \) as a linear combination of some other given matrices \( A_i \in \mathbb{R}^{p \times q}, i = 1, \ldots, n, \) as measured by the matrix norm (maximum singular value):
\[ \text{minimize} \quad \|x_1 A_1 + \cdots + x_n A_n - B\|. \]

(a) Explain how to find a subgradient of the objective function at \( x \).
(b) Generate a random instance of the problem with \( n = 5, p = 3, q = 6 \). Use CVX to find the optimal value \( f^* \) of the problem. Use a subgradient method to solve the problem, starting from \( x = 0 \). Plot \( f - f^* \) versus iteration. Experiment with several step size sequences.