1 Convexity of Sets

A set $C \subseteq \mathbb{R}^n$ is called convex if the line segment between any two points in $C$ lies in $C$, i.e., for any $x_1, x_2 \in C$ and any $\theta$ with $0 \leq \theta \leq 1$, we have,

$$\theta x_1 + (1 - \theta)x_2 \in C.$$

(a) Show that the following sets are convex:

i A vector subspace $V$ of $\mathbb{R}^n$.

ii A hyperplane $L$ in $\mathbb{R}^n$, given by $L := \{x : a^\top x = b\}$, where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$.

iii A halfspace $H$ in $\mathbb{R}^n$, given by $H := \{x : a^\top x \leq b\}$, where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$.

iv A norm-ball $B$ in $\mathbb{R}^n$, given by $B := \{x : \|x - x_c\| \leq r\}$, where $x_c \in \mathbb{R}^n$, $r \geq 0$ and $\|\cdot\|$ is a norm on $\mathbb{R}^n$.

(b) Operations that preserve convexity of sets:

i Intersection

1. Show that convexity is preserved under intersection, i.e. if $S_1$ and $S_2$ are convex subsets of $\mathbb{R}^n$, then $S := S_1 \cap S_2$ is convex.

2. Show that a polyhedron is convex. A polyhedron is the solution set of a finite number of linear inequalities. For example $P := \{x \in \mathbb{R}^2 \mid x_2 \leq x_1, x_1 \geq 0, x_1 \leq 1\}$ is a polyhedron.

ii Mapping under an affine function

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called affine if it is a sum of a linear function and a constant, i.e.

$$f(x) = Ax + b,$$

for $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

1. Prove that if $S \subseteq \mathbb{R}^n$ is convex, then the image of $S$ under $f$,

$$f(S) := \{f(x) : x \in S\},$$

which is a subset of $\mathbb{R}^m$, is convex.

Example: Let $e_i, 1 \leq i \leq n$ denote the standard basis vectors in $\mathbb{R}^n$. As an example of what we have just shown, the projection of a convex set in $\mathbb{R}^n$ onto the span of any subset of the standard basis vectors is convex. E.g. if $S \subseteq \mathbb{R}^3$ is convex, then $T := \{(x_1, x_2) \in \mathbb{R}^2 : (x_1, x_2, x_3) \in S \text{ for some } x_3 \in \mathbb{R}\}$ is convex.
2. Show that the inverse image of a convex set $C \subseteq \mathbb{R}^m$ under the affine function $f(x) = Ax + b$ is convex. Here, the inverse image of $C$ is defined to be the subset of $\mathbb{R}^n$ given by

$$f^{-1}(C) := \{x : f(x) \in C\}.$$
2 Convexity of Functions

Recall the following definitions and facts involving convex functions:

- A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *convex* if $\text{dom}(f)$ is a convex set and if, for all $x, y \in \text{dom}(f)$ and $0 \leq \theta \leq 1$, we have

  \[ f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y). \]  

  \hspace{1cm} (1)

  The function $f$ is called *strictly convex* if the inequality in (1) is strict whenever $\theta \in (0, 1)$ and $x \neq y$.

- A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *concave* if $\text{dom}(f)$ is a convex set and if, for all $x, y \in \text{dom}(f)$ and $0 \leq \theta \leq 1$, we have

  \[ f(\theta x + (1 - \theta)y) \geq \theta f(x) + (1 - \theta)f(y). \]

  The function $f$ is called *strictly concave* if the inequality is strict when $\theta \in (0, 1)$ and $x \neq y$.

- A function $f$ is concave if and only if $-f$ is convex and strictly concave iff $-f$ is strictly convex. An affine function is both convex and concave.

It is useful to keep in mind several alternative characterizations of convexity:

- **First order condition**
  Suppose $f$ is differentiable at every point in its domain (in particular, for this to make sense, $\text{dom}(f)$ would need to be an open set). Then $f$ is convex if and only if we have

  \[ f(y) \geq f(x) + \nabla f(x)^\top (y - x), \]

  for all $x, y \in \text{dom}(f)$.

- **Second order condition**
  Suppose $f$ is twice differentiable at every point in its domain (so, in particular, $\text{dom}(f)$ will need to be an open set when one applies this criterion). Then $f$ is convex if and only if its Hessian $\nabla^2 f(x)$ is positive semidefinite at every $x \in \text{dom}(f)$.

- **Restriction to a line**
  A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex iff for all $x \in \text{dom}(f)$ and all $v \in \mathbb{R}^n$, the function $g : \mathbb{R} \to \mathbb{R}$ given by $g(t) = f(x + tv)$, with its domain defined to be $\text{dom}(g) := \{ t : x + tv \in \text{dom}(f) \}$, is convex.

- **Epigraph is convex**
  A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if its epigraph, defined as $\text{epi}(f) := \{(x, t) : x \in \text{dom}(f), f(x) \leq t\}$ is a convex subset of $\mathbb{R}^{n+1}$. 
Figure 1: A graph of the function $f$ in Problem 2(e). The function can be thought of as taking the value $\infty$ outside its domain, $\text{dom}(f) = [-1, 1]$. Note that the function jumps at $-1$ and at $1$ in the sense that $\lim_{x \uparrow 1} f(x) = 1 \neq f(1) = 1.5$ and $\lim_{x \downarrow -1} f(x) = 1 \neq f(-1) = 1.25$. The existence of such jumps is emphasized by the open circles and the closed circles in the diagram. With the interpretation that the function equals $\infty$ outside the closed interval $[-1, 1]$ we would also have $\lim_{x \downarrow 1} f(x) = \infty \neq f(1) = 1.5$ and $\lim_{x \uparrow -1} f(x) = \infty \neq f(-1) = 1.25$.

(a) Show that if $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex, $x_1, x_2, \ldots, x_k \in \text{dom}(f)$, and $\theta_1, \theta_2, \ldots, \theta_k \geq 0$ with $\sum_{i=1}^{k} \theta_i = 1$, then we have

$$f(\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k) \leq \theta_1 f(x_1) + \theta_2 f(x_2) + \cdots + \theta_k f(x_k).$$

This result is known as Jensen’s inequality for discrete probability distributions.

(b) Find a condition on the symmetric matrix $A \in \mathbb{R}^{n \times n}$ in order to ensure that $f(x) = x^T A x$ is a convex function from $\mathbb{R}^n$ to $\mathbb{R}$.

(c) Show that the following operations preserve convexity:

i. **Nonnegative weighted sum:** If $f$ and $g$ are convex functions on $\mathbb{R}^n$ with domains $\text{dom}(f)$ and $\text{dom}(g)$ respectively, then $af + bg$ for $a, b \geq 0$ is a convex function on $\mathbb{R}^n$ with domain $\text{dom}(f) \cap \text{dom}(g)$, assuming $\text{dom}(f) \cap \text{dom}(g) \neq \emptyset$.

ii. **Pointwise maximum:** If $f$ and $g$ are convex functions on $\mathbb{R}^n$ with domains $\text{dom}(f)$ and $\text{dom}(g)$ respectively, then so is their pointwise maximum $h$, defined as

$$h(x) := \max\{f(x), g(x)\},$$

with domain $\text{dom}(f) \cap \text{dom}(g)$, assuming $\text{dom}(f) \cap \text{dom}(g) \neq \emptyset$. 

(d) Show that \( f : S^n_{++} \to \mathbb{R} \) defined by \( f(X) = \log \det(X) \) is concave. Note that we can think of \( f \) as a function on \( S^n \) (which is a vector space that contains \( S^n_{++} \)), with \( \text{dom}(f) = S^n_{++} \).

**Hint:** Use the “restriction to a line” condition to check for convexity.

(e) Consider the function \( f : \mathbb{R} \to \mathbb{R} \) with \( \text{dom}(f) = [-1, 1] \), given by

\[
    f(x) = \begin{cases} 
        1.5 & \text{if } x = 1, \\
        |x| & \text{if } -1 < x < 1, \\
        1.25 & \text{if } x = -1.
    \end{cases}
\]

This is illustrated in Figure 1.

Show that \( f \) is a convex function.
3 Square-to-linear function

Consider the square-to-linear function

\[ f(x, y) := \begin{cases} \frac{x^T x}{y} & \text{if } y > 0, \\ \infty & \text{otherwise,} \end{cases} \]

defined for \((x, y) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} : y > 0\).

(a) Show that \(\text{dom}(f)\) is a convex subset of \(\mathbb{R}^{n+1}\).

(b) Show that the Hessian of \(f\) at \((x, y) \in \text{dom}(f)\) is given by

\[
\nabla^2 f(x, y) = \frac{2}{y^3} \begin{bmatrix} y^2 I_n & -yx \\ -yx^T & x^T x \end{bmatrix}.
\]

(c) Show that \(\nabla^2 f(x, y)\) is PSD at every \((x, y) \in \text{dom}(f)\) and thereby conclude that \(f\) is a convex function.