Discussion 5

1. Convexity of Sets

Definition. A set C is convex if and only if the line segment between any two points in C lies in C:

$$C \text{ is convex} \iff \forall \vec{x}_1, \vec{x}_2 \in C,\ \forall \theta \in [0,1],\ \theta\vec{x}_1 + (1-\theta)\vec{x}_2 \in C \tag{1}$$

(a) Show that the **intersection of convex sets is convex**:

$$C_1, C_2 \text{ are convex} \implies C = C_1 \cap C_2 \text{ is convex} \tag{2}$$

Solution: Consider $\vec{x}_1, \vec{x}_2 \in C$ and $\theta \in [0,1]$. Then $\vec{x}_1, \vec{x}_2 \in C_1$ and $\vec{x}_1, \vec{x}_2 \in C_2$. Since C_1 and C_2 are convex we have, $\theta\vec{x}_1 + (1-\theta)\vec{x}_2 \in C_1$ and $\theta\vec{x}_1 + (1-\theta)\vec{x}_2 \in C_2$, which implies $\theta\vec{x}_1 + (1-\theta)\vec{x}_2 \in C$.

(b) Show that the following sets are convex:

i. **(OPTIONAL)** A vector subspace of $\mathbb{R}^n$.
Solution: If C is a vector subspace of $\mathbb{R}^n$ then $\forall \vec{x}_1, \vec{x}_2 \in C$, and $\forall \alpha, \beta \in \mathbb{R}$, $\alpha\vec{x}_1 + \beta\vec{x}_2 \in C$. So $\forall \vec{x}_1, \vec{x}_2 \in C$, $\forall \theta \in [0,1]$, $\theta\vec{x}_1 + (1-\theta)\vec{x}_2 \in C$.

ii. **(OPTIONAL)** A hyperplane, $\mathcal{L} = \{\vec{x} \mid \vec{a}^\top \vec{x} = b\}$.
Solution: $\forall \vec{x}_1, \vec{x}_2 \in H$, $\forall \theta \in [0,1]$:

$$\vec{a}^\top(\theta\vec{x}_1 + (1-\theta)\vec{x}_2) = \theta(\vec{a}^\top\vec{x}_1) + (1-\theta)(\vec{a}^\top\vec{x}_2) \tag{3}$$
$$= \theta b + (1-\theta)b \tag{4}$$
$$= b. \tag{5}$$

So, $\theta\vec{x}_1 + (1-\theta)\vec{x}_2 \in H$ and H is convex.
Other proof: an hyperplane is the intersection of two half-spaces, therefore it is convex.

iii. A halfspace, $\mathcal{H} = \{\vec{x} \mid \vec{a}^\top \vec{x} \leq b\}$.
Solution: $\forall \vec{x}_1, \vec{x}_2 \in H$, $\forall \theta \in [0,1]$:

$$\vec{a}^\top(\theta\vec{x}_1 + (1-\theta)\vec{x}_2) = \theta(\vec{a}^\top\vec{x}_1) + (1-\theta)(\vec{a}^\top\vec{x}_2) \tag{6}$$
$$\leq \theta b + (1-\theta)b \tag{7}$$
$$= b. \tag{8}$$

So, $\theta\vec{x}_1 + (1-\theta)\vec{x}_2 \in H$ and H is convex.

Definition. A function $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is affine if it is the sum of a linear function and a constant,

$$f(\vec{x}) = A\vec{x} + \vec{b}, \tag{9}$$

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

(c) **(OPTIONAL)** Prove that if $S \subseteq \mathbb{R}^n$ is convex, then the image of S under an affine function f,

$$f(S) = \{f(\vec{x}) \mid \vec{x} \in S\}, \tag{10}$$

is convex.

Solution: Let $\vec{y}_1, \vec{y}_2 \in f(S)$. This implies there exist $\vec{x}_1, \vec{x}_2 \in S$ such that $\vec{y}_1 = A\vec{x}_1 + \vec{b}$ and $\vec{y}_2 = A\vec{x}_2 + \vec{b}$.

We want to show that $\lambda\vec{y}_1 + (1-\lambda)\vec{y}_2 \in f(S)$ for $0 \leq \lambda \leq 1$.

Since S is convex we have $\lambda\vec{x}_1 + (1-\lambda)\vec{x}_2 \in S$. Further $A(\lambda\vec{x}_1 + (1-\lambda)\vec{x}_2) + \vec{b} = \lambda\vec{y}_1 + (1-\lambda)\vec{y}_2$. This shows that $\lambda\vec{y}_1 + (1-\lambda)\vec{y}_2 \in f(S)$.

2. Convexity of Functions

Definition. A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is convex if $\mathrm{dom}(f)$ is a convex set and if for all $\vec{x}, \vec{y} \in \mathrm{dom}(f)$ and $\theta \in [0, 1]$, we have,

$$f(\theta\vec{x} + (1-\theta)\vec{y}) \leq \theta f(\vec{x}) + (1-\theta) f(\vec{y}). \tag{11}$$

The function f is strictly convex if the inequality is strict.

Definition. A function $f : \mathbb{R}^n \to \mathbb{R}$ is concave if $\mathrm{dom}(f)$ is a convex set and if for all $\vec{x}, \vec{y} \in \mathrm{dom}(f)$ and θ with $0 \leq \theta \leq 1$, we have,

$$f(\theta\vec{x} + (1-\theta)\vec{y}) \geq \theta f(\vec{x}) + (1-\theta) f(\vec{y}). \tag{12}$$

The function f is strictly concave if the inequality is strict.

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

Property: Jensen's inequality. The inequality in Equation (11) is known as **Jensen's Inequality**. This can be extended to convex combinations of more than one point. If f is convex, and $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k \in \mathrm{dom}(f)$, and $\theta_1, \theta_2, \ldots, \theta_k \geq 0$ with $\sum_{i=1}^k \theta_i = 1$ then,

$$f(\theta_1\vec{x}_1 + \theta_2\vec{x}_2 + \cdots + \theta_k\vec{x}_k) \leq \theta_1 f(\vec{x}_1) + \theta_2 f(\vec{x}_2) + \cdots + \theta_k f(\vec{x}_k). \tag{13}$$

Property: first order condition. Suppose f is differentiable. Then f is convex if and only if $\mathrm{dom}(f)$ is convex and

$$f(\vec{y}) \geq f(\vec{x}) + \nabla f(\vec{x})^\top(\vec{y} - \vec{x}), \tag{14}$$

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

Property: Second order condition. Suppose f is twice differentiable. Then f is convex if and only if, $\mathrm{dom}(f)$ is convex and the Hessian of f, $\nabla^2 f(\vec{x})$, is positive semi-definite for all $\vec{x} \in \mathrm{dom}(f)$.

(a) Under what condition on $A \in \mathbb{R}^{n\times n}$, where A is symmetric, is the function $f : \vec{x} \to \vec{x}^\top A\vec{x}$ convex?

Solution: We have $\nabla^2 f(x) = 2A$ and for f to be convex we require A to be positive semi-definite.

(b) **(OPTIONAL) Restriction to a line**.

Show that a function f is convex if and only if for all $\vec{x} \in \mathrm{dom}(f)$ and all $\vec{v}$, the function $g : \mathrm{dom}(g) \to \mathbb{R}$ given by $g(t) = f(\vec{x} + t\vec{v})$ is convex for $\mathrm{dom}(g) = \{t \in \mathbb{R} \mid \vec{x} + t\vec{v} \in \mathrm{dom}(f)\}$.

Solution: In the first direction: assume f is convex and consider $\vec{x} \in \mathrm{dom}(f)$, $\vec{v}$ and the function $g : \mathrm{dom}(g) \to \mathbb{R}$ given by $g(t) = f(\vec{x} + t\vec{v})$ where $\mathrm{dom}(g) = \{t \in \mathbb{R} \mid \vec{x} + t\vec{v} \in \mathrm{dom}(f)\}$.

Because f is convex, $\mathrm{dom}(f)$ is convex, therefore $\mathrm{dom}(g)$ is also convex. For $t_1, t_2 \in \mathrm{dom}(g)$ and $\lambda \in [0, 1]$:

$$\begin{aligned}
g(\lambda t_1 + (1-\lambda)t_2) &= f(\vec{x} + (\lambda t_1 + (1-\lambda)t_2)\vec{v}) && (15)\\
&= f(\lambda(\vec{x} + t_1\vec{v}) + (1-\lambda)(\vec{x} + t_2\vec{v})) && (16)\\
&\leq \lambda f(\vec{x} + t_1\vec{v}) + (1-\lambda)f(\vec{x} + t_2\vec{v}) && (17)\\
&= \lambda g(t_1) + (1-\lambda)g(t_2) && (18)
\end{aligned}$$

Therefore g is convex.

In the other direction: Consider $\vec{x}_1, \vec{x}_2 \in \mathrm{dom}(f)$ and $\lambda \in [0, 1]$. Define $g : t \to f(\vec{x}_2 + t(\vec{x}_1 - \vec{x}_2))$. g is convex and $0 \in \mathrm{dom}(g)$ and $1 \in \mathrm{dom}(g)$, so $[0, 1] \in \mathrm{dom}(g)$. Therefore $\lambda\vec{x}_1 + (1-\lambda)\vec{x}_2 \in \mathrm{dom}(f)$ and $\mathrm{dom}(f)$ is convex.

Because g is convex:

$$\begin{aligned}
g(\lambda 1 + (1-\lambda)0) = g(\lambda) &\leq \lambda g(1) + (1-\lambda)g(0) && (19)\\
f(\vec{x}_2 + \lambda(\vec{x}_1 - \vec{x}_2)) &\leq \lambda f(\vec{x}_2 + 1(\vec{x}_1 - \vec{x}_2)) + (1-\lambda)f(\vec{x}_2 + 0(\vec{x}_1 - \vec{x}_2)) && (20)\\
f(\lambda\vec{x}_2 + (1-\lambda)\vec{x}_2) &\leq \lambda f(\vec{x}_1) + (1-\lambda)f(\vec{x}_2) && (21)
\end{aligned}$$

Therefore f is convex.

(c) **(OPTIONAL) Non-negative weighted sum**.

Show that the non-negative weighted sum of convex functions is convex: i.e. if $f_1, \ldots, f_n$ are n convex functions from $\mathbb{R}^n$ to $\mathbb{R}$ and $w_1, \ldots, w_n \in \mathbb{R}_+$ are n positive scalars, then the function:

$$f = \sum_{i=1}^{n} w_i f_i \tag{22}$$

is convex. To make the question easier, you can assume that the functions $f_1, \ldots, f_n$ are twice-differentiable.

Solution: Check convexity by using the second order condition. First, the weighted sum of twice-differentiable function is also twice-differentiable:

$$\begin{aligned}
\nabla^2 f &= \nabla^2 \left(\sum_{i=1}^{n} w_i f_i\right) && (23)\\
&= \sum_{i=1}^{n} w_i \nabla^2 f_i && \text{(linearity of } \nabla^2) \quad (24)
\end{aligned}$$

Next we check that $\nabla^2 f$ is PSD.

$$\begin{aligned}
\forall\vec{y}, \forall\vec{x} \quad \vec{y}^\top(\nabla^2 f(\vec{x}))\vec{y} &= \vec{y}^\top(\sum_{i=1}^{n} w_i \nabla^2 f_i(\vec{x}))\vec{y} && (25)\\
&= \sum_{i=1}^{n} w_i \vec{y}^\top(\nabla^2 f_i(\vec{x}))\vec{y} && (26)\\
&\geq 0 && (\vec{y}^\top(\nabla^2 f_i(\vec{x}))\vec{y} \geq 0, \text{ because } f_i \text{ is convex}) \quad (27)
\end{aligned}$$

So $\forall\vec{x}$, $\nabla^2 f(\vec{x})$ is PSD, so f is convex.

(d) **Point-wise maximum**.

Show that if f_1 and f_2 are convex functions then their pointwise maximum f, defined by

$$f(\vec{x}) = \max(f_1(\vec{x}), f_2(\vec{x})), \tag{28}$$

with $\mathrm{dom}(f) = \mathrm{dom}(f_1) \cap \mathrm{dom}(f_2)$, is also convex.

Solution: Because f_1 and f_2 are convex, then $\mathrm{dom}(f_1)$ and $\mathrm{dom}(f_2)$ are convex sets. Because convexity of sets is preserved under intersection, $\mathrm{dom}(f) = \mathrm{dom}(f_1) \cap \mathrm{dom}(f_2)$ is also convex.

$$\begin{aligned}
\mathrm{epi}(f) &= \{(\vec{x}, t) \mid \vec{x} \in \mathrm{dom}(f), f(\vec{x}) \leq t\} && (29)\\
&= \{(\vec{x}, t) \mid \vec{x} \in \mathrm{dom}(f), \max(f_1(\vec{x}), f_2(\vec{x})) \leq t\} && (30)\\
&= \{(\vec{x}, t) \mid \vec{x} \in \mathrm{dom}(f_1) \cap \mathrm{dom}(f_2), f_1(\vec{x}) \leq t \text{ and } f_2(\vec{x}) \leq t\} && (31)\\
&= \{(\vec{x}, t) \mid \vec{x} \in \mathrm{dom}(f_1), f_1(\vec{x}) \leq t\} \cap \{(\vec{x}, t) \mid \vec{x} \in \mathrm{dom}(f_2), f_2(\vec{x}) \leq t\} && (32)\\
&= \mathrm{epi}(f_1) \cap \mathrm{epi}(f_2) && (33)
\end{aligned}$$

Because f_1 and f_2 are convex, then $\mathrm{epi}(f_1)$ and $\mathrm{epi}(f_2)$ are convex. Because convexity of sets is preserved under intersection, $\mathrm{epi}(f)$ is convex. Because of the equivalence between the convexity of functions and the convexity of their epigraphs, f is convex.

(e) Show that a piece-wise linear function that can be written as,

$$f(\vec{x}) = \max(\vec{a}_1^\top \vec{x} + \vec{b}_1, \vec{a}_2^\top \vec{x} + \vec{b}_2, ..., \vec{a}_m^\top \vec{x} + \vec{b}_m), \tag{34}$$

is convex.

Solution: $f(\vec{x})$ is the point-wise maximum of affine (hence convex) functions and is therefore convex.

3. Convexity and composition of functions

Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$. Define the composition of f with g as $h = f \circ g : \mathbb{R}^n \to \mathbb{R}$ such that $h(\vec{x}) = f(g(\vec{x}))$.

(a) Show that if f is convex and non decreasing and g is convex, then h is convex.

Solution:

$$\begin{aligned}
h(\lambda\vec{x} + (1-\lambda)\vec{y}) &= f(g(\lambda\vec{x} + (1-\lambda)\vec{y})) \\
&\leq f(\lambda g(\vec{x}) + (1-\lambda)(g(\vec{y}))) && (g \text{ convex and } f \text{ nondecreasing})\\
&\leq \lambda f(g(\vec{x})) + (1-\lambda) f(g(\vec{y})) && (f \text{ convex})\\
&= \lambda h(\vec{x}) + (1-\lambda) h(\vec{y})
\end{aligned}$$

So h is convex.

(b) Show that there exists f non decreasing and g convex, such that $h = f \circ g$ is not convex.

Solution: Take $n = 1$, $f(x) = \log(()x)$ and $g(x) = x$. Then $h(x) = \log(()x)$ is not convex.

(c) Show that there exists f convex and g convex such that $h = f \circ g$ is not convex.

Solution: Take $n = 1$, $f(x) = -x$ and $g(x) = x^2$, then $h(x) = -x^2$ is not convex.

