EECS 127/227AT Optimization Models in Engineering
Spring 2020 — Discussion 4

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:

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

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

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

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: $\forall \vec{x}_1, \vec{x}_2 \in H$, $\forall \theta \in [0,1]$:

$$\begin{aligned}
\vec{a}^\top(\theta \vec{x}_1 + (1-\theta)\vec{x}_2) &= \theta(\vec{a}^\top \vec{x}_1) + (1-\theta)(\vec{a}^\top \vec{x}_2) \\
&= \theta b + (1-\theta) b \\
&= b.
\end{aligned}$$

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

$$\begin{aligned}
\vec{a}^\top(\theta \vec{x}_1 + (1-\theta)\vec{x}_2) &= \theta(\vec{a}^\top \vec{x}_1) + (1-\theta)(\vec{a}^\top \vec{x}_2) \\
&\le \theta b + (1-\theta) b \\
&= b.
\end{aligned}$$

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

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

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

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

(c) **[Optional] Conservation of convexity through affine transformation**. Prove that if $S \subseteq \mathbb{R}^n$ is convex, then the image of S under an affine function f,

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

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 $0 \le \lambda \le 1$.
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 $f :: \mathbb{R}^n \to \mathbb{R}$ 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,

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

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 $0 \le \theta \le 1$, we have,

$$f(\theta\vec{x} + (1-\theta)\vec{y}) \ge \theta f(\vec{x}) + (1-\theta)f(\vec{y}).$$

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 (1) 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 $\theta_1, \theta_2, \ldots, \theta_k \ge 0$ with $\sum_{i=1}^{k} \theta_i = 1$ then,

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

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

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

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

Solution: In the first direction: assume f is convex and consider $\vec{x} \in \text{dom}(f)$, $\vec{v}$ and the function $g : \text{dom}(g) \to \mathbb{R}$ given by $g(t) = f(\vec{x}+t\vec{v})$ where $\text{dom}(g) = \{t \in \mathbb{R} \mid \vec{x}+t\vec{v} \in \text{dom}(f)\}$. Because f is convex, $\text{dom}(f)$ is convex, therefore $\text{dom}(g)$ is also convex. For $t_1, t_2 \in \text{dom}(g)$ and $\lambda \in [0, 1]$:

$$\begin{aligned}
g(\lambda t_1 + (1-\lambda)t_2) &= f(\vec{x} + (\lambda t_1 + (1-\lambda)t_2)\vec{v}) \\
&= f(\lambda(\vec{x} + t_1\vec{v}) + (1-\lambda)(\vec{x} + t_2\vec{v})) \\
&\le \lambda f(\vec{x} + t_1\vec{v}) + (1-\lambda)f(\vec{x} + t_2\vec{v}) \\
&= \lambda g(t_1) + (1-\lambda)g(t_2)
\end{aligned}$$

Therefore g is convex.

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

Because g is convex:

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

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 $f_1, \dots, f_n$ are n convex functions from $\mathbb{R}^n$ to $\mathbb{R}$ and $w_1, \dots, w_n \in \mathbb{R}_+$ are n positive scalars, then the function:

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

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

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

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

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

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

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

(d) **[Optional] Point-wise maximum** Show that if f_1 and f_2 are convex functions then their pointwise maximum f, defined by

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

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

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

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

Because f_1 and f_2 are convex, then $\text{epi}(f_1)$ and $\text{epi}(f_2)$ are convex. Because convexity of sets is preserved under intersection, $\text{epi}(f)$ 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,

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

is convex.

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

3. Disproving convexity: Finding counter-examples

Though we spend a lot of time in this course learning how to prove convexity of sets and functions, in practical scenarios we may not have a mathematical representation of a set/function and so it is not possible to prove convexity. Instead, we may be able to represent this set/function in terms of a query $Q(\vec{x})$ that returns some information about the element $\vec{x}$ in relation to the set/function. For example, instead representing the set $S = \{\vec{x} \mid \text{some condition on } \vec{x}\}$ we only have $Q(\vec{x})$ which returns whether or not $\vec{x} \in S$.

In these cases we can **disprove** convexity by showing that one or more of the properties of convex sets/functions are violated by finding counterexamples. In this problem we will see how we can disprove convexity for sets/functions given limited information that can be accessed via certain types of queries.

(a) **Disproving convexity of set S (Proving non-convexity of set S)**
Assume that we know that the set lies within some $\mathcal{D}$.
Query: $Q(\vec{x})$: For $\vec{x} \in \mathcal{D}$ that returns True if $\vec{x} \in S$ and False if $\vec{x} \notin S$. How can you use Q to check/disprove convexity of S?

Solution: Choose $\vec{x}$ and $\vec{y}$ randomly in $\mathcal{D}$ and if both lie in S then check if $(\vec{x} + \vec{y})/2$ lies in S. We can choose any point on line segment joining $\vec{x}, \vec{y} \in S$ instead of the mid-point.

(b) **Disproving convexity of function f (Proving non-convexity of function f).**
Assume that we know $\text{dom}(f)$, denoted as $\mathcal{D}$ and that $\mathcal{D}$ is convex.

i. Query: $G(\vec{x})$:For $\vec{x} \in \mathcal{D}$, returns function value $f(\vec{x})$.
How can you use G to check/disprove convexity of f?
Solution: Get $G(\vec{x}), G(\vec{y})$ for $\vec{x}, \vec{y} \in \mathcal{D}$ and then check if $G(\frac{\vec{x}+\vec{y}}{2}) \leq \frac{G(\vec{x})+G(\vec{y})}{2}$. Can also check for other points on line segment joining $\vec{x}$ and $\vec{y}$.

ii. Query: $H(\vec{x})$: For $\vec{x} \in \mathcal{D}$, returns $f(\vec{x})$ and $\nabla f(\vec{x})$. (Here we assume that f is differentiable). How can you use H to check/disprove convexity of f?
Solution: Check first order condition $f(\vec{y}) \geq f(\vec{x}) + \nabla f(\vec{x})^\top (\vec{y} - \vec{x})$.