EECS 127/227AT | Optimization Models in Engineering | UC Berkeley | Fall 2023

Discussion 6

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

$$\begin{aligned}
\vec{a}^\top(\theta\vec{x}_1 + (1-\theta)\vec{x}_2) &= \theta(\vec{a}^\top\vec{x}_1) + (1-\theta)(\vec{a}^\top\vec{x}_2) && (2)\\
&= \theta b + (1-\theta)b && (3)\\
&= b. && (4)
\end{aligned}$$

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

$$\begin{aligned}
\vec{a}^\top(\theta\vec{x}_1 + (1-\theta)\vec{x}_2) &= \theta(\vec{a}^\top\vec{x}_1) + (1-\theta)(\vec{a}^\top\vec{x}_2) && (5)\\
&\le \theta b + (1-\theta)b && (6)\\
&= b. && (7)
\end{aligned}$$

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

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

$$C_1, C_2 \text{ are convex} \implies C = C_1 \cap C_2 \text{ is convex} \tag{8}$$

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 $\theta\vec{x}_1 + (1-\theta)\vec{x}_2 \in C$. A special case of this is when $C_1 \cap C_2 = \emptyset$.

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 nonempty 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 whenever $\vec{x} \neq \vec{y}$ and $\theta \notin \{0,1\}$.

Definition. A function $f : \mathbb{R}^n \to \mathbb{R}$ is concave if $\mathrm{dom}(f)$ is a nonempty 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 whenever $\vec{x} \neq \vec{y}$ and $\theta \notin \{0,1\}$.

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 dom(f) is a nonempty open set and f is differentiable. Then f is convex if and only if

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

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

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

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

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

$$f(\vec{x}) = \max(f_1(\vec{x}), f_2(\vec{x})), \tag{15}$$

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

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. Further, it is nonempty by assumption.

$$\mathrm{epi}(f) = \{(\vec{x}, t) \mid \vec{x} \in \mathrm{dom}(f), f(\vec{x}) \leq t\} \tag{16}$$

$$= \{(\vec{x}, t) \mid \vec{x} \in \mathrm{dom}(f), \max(f_1(\vec{x}), f_2(\vec{x})) \le t\} \tag{17}$$
$$= \{(\vec{x}, t) \mid \vec{x} \in \mathrm{dom}(f_1) \cap \mathrm{dom}(f_2), f_1(\vec{x}) \le t \text{ and } f_2(\vec{x}) \le t\} \tag{18}$$
$$= \{(\vec{x}, t) \mid \vec{x} \in \mathrm{dom}(f_1), f_1(\vec{x}) \le t\} \cap \{(\vec{x}, t) \mid \vec{x} \in \mathrm{dom}(f_2), f_2(\vec{x}) \le t\} \tag{19}$$
$$= \mathrm{epi}(f_1) \cap \mathrm{epi}(f_2) \tag{20}$$

Because f_1 and f_2 are convex, then $\mathrm{epi}(f_1)$ and $\mathrm{epi}(f_2)$ are nonempty convex sets. 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.

(b) **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 $\mathrm{dom}(g) = \{t \in \mathbb{R} \mid \vec{x}+t\vec{v} \in \mathrm{dom}(f)\}$. Also, $\mathrm{dom}(g)$ is nonempty because $\vec{0} \in \mathrm{dom}(g)$.

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 $\lambda \in [0, 1]$:

$$g(\lambda t_1 + (1-\lambda)t_2) = f(\vec{x} + (\lambda t_1 + (1-\lambda)t_2)\vec{v}) \tag{21}$$
$$= f(\lambda(\vec{x} + t_1\vec{v}) + (1-\lambda)(\vec{x} + t_2\vec{v})) \tag{22}$$
$$\le \lambda f(\vec{x} + t_1\vec{v}) + (1-\lambda) f(\vec{x} + t_2\vec{v}) \tag{23}$$
$$= \lambda g(t_1) + (1-\lambda) g(t_2) \tag{24}$$

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:

$$g(\lambda 1 + (1-\lambda)0) = g(\lambda) \le \lambda g(1) + (1-\lambda) g(0) \tag{25}$$
$$f(\vec{x}_2 + \lambda(\vec{x}_1 - \vec{x}_2)) \le \lambda f(\vec{x}_2 + 1(\vec{x}_1 - \vec{x}_2)) + (1-\lambda) f(\vec{x}_2 + 0(\vec{x}_1 - \vec{x}_2)) \tag{26}$$
$$f(\lambda\vec{x}_2 + (1-\lambda)\vec{x}_2) \le \lambda f(\vec{x}_1) + (1-\lambda) f(\vec{x}_2) \tag{27}$$

Therefore f is convex.

(c) **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:

$$f = \sum_{i=1}^{n} w_i f_i \tag{28}$$

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:

$$\begin{aligned}\nabla^2 f &= \nabla^2\left(\sum_{i=1}^{n} w_i f_i\right) && (29)\\ &= \sum_{i=1}^{n} w_i \nabla^2 f_i && \text{(linearity of } \nabla^2\text{)} \quad (30)\end{aligned}$$

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

$$\begin{aligned}\forall \vec{y}, \forall \vec{x} \quad \vec{y}^\top(\nabla^2 f(\vec{x}))\vec{y} &= \vec{y}^\top(\sum_{i=1}^{n} w_i \nabla^2 f_i(\vec{x}))\vec{y} && (31)\\ &= \sum_{i=1}^{n} w_i \vec{y}^\top(\nabla^2 f_i(\vec{x}))\vec{y} && (32)\\ &\geq 0 && (\vec{y}^\top(\nabla^2 f_i(\vec{x}))\vec{y} \geq 0, \text{ because } f_i \text{ is convex}) \quad (33)\end{aligned}$$

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

3. Convexity of Constraint Sets

Let $f_1, \ldots, f_m, h_1, \ldots, h_p \colon \mathbb{R}^n \to \mathbb{R}$ be functions. Let $S \subseteq \mathbb{R}^n$ be defined as

$$S \doteq \left\{ \vec{x} \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} f_i(\vec{x}) \leq 0 & \forall i = 1, \ldots, m \\ h_j(\vec{x}) = 0 & \forall j = 1, \ldots, p \end{array} \right\}. \quad (34)$$

Show that if $f_1, \ldots, f_m$ are convex functions, and $h_1, \ldots, h_p$ are affine functions, then S is a convex set.

Solution: Let $\vec{x}, \vec{y} \in S$ and let $\theta \in [0, 1]$. Then for any $i = 1, \ldots, m$, we have

$$\begin{aligned} f_i(\theta\vec{x} + (1-\theta)\vec{y}) &\leq \theta \underbrace{f_i(\vec{x})}_{\leq 0} + (1-\theta)\underbrace{f_i(\vec{y})}_{\leq 0} \\ &\leq 0. \end{aligned}$$

And for any $j = 1, \ldots, p$, we have

$$\begin{aligned} h_j(\theta\vec{x} + (1-\theta)\vec{y}) &= \theta \underbrace{h_j(\vec{x})}_{=0} + (1-\theta)\underbrace{h(\vec{y})}_{=0} \\ &= 0. \end{aligned}$$

Thus $\theta\vec{x} + (1-\theta)\vec{y} \in S$. Thus S is convex.

4. Properties of Convex Functions

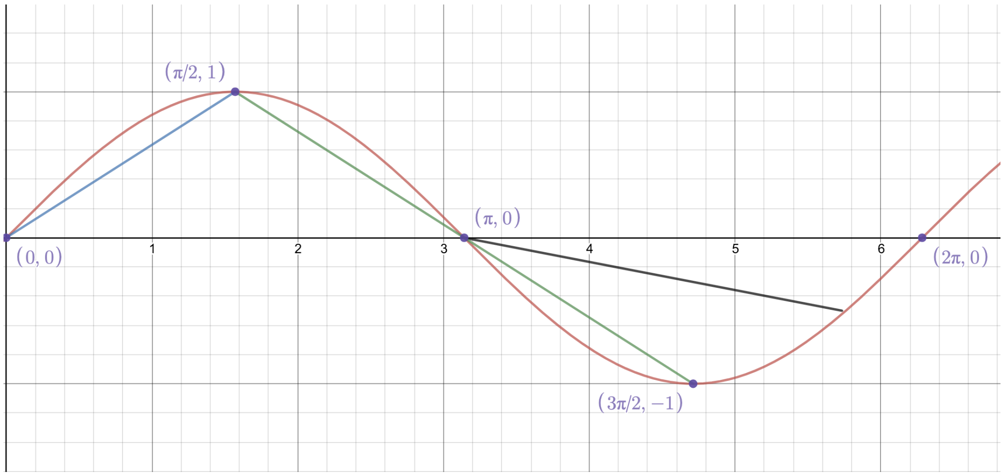
In this exercise, we examine convexity and what it represents graphically.

(a) In what region between $[0, 2\pi]$ is $\sin(x)$ a convex function? In what region between $[0, 2\pi]$ is $\sin(x)$ a concave function? Give a region between $[0, 2\pi]$ where $\sin(x)$ is neither convex nor concave.

Solution: The function $\sin(x)$ is convex (in fact, strictly convex) between $[\pi, 2\pi]$; similarly, it is concave (in fact, strictly concave) between $[0, \pi]$. It is non-convex and non-concave for any interval between $[0, 2\pi]$ that is not a subset of the two aforementioned intervals.

(b) Plot $\sin(x)$ between $[0, 2\pi]$. For each of the 3 intervals defined above in part (a), draw a chord to illustrate graphically on what regions the function is convex, concave, and neither convex nor concave.

Solution:



In the region $[0, \pi]$, the function is concave and all chords (e.g., the *blue* chord above) lie below the function. In the region $[\pi, 2\pi]$, the function is convex and all chords (e.g., the *black* chord above) lie above the function. When considering the full region $[0, 2\pi]$, or any region that is not a subset of the two regions above, chords (like the example *green* chord above) do not lie strictly above or strictly below the function.

(c) Show that for all $x \in [0, \frac{\pi}{2}]$,

$$\frac{2}{\pi}x \leq \sin x \leq x. \tag{35}$$

Solution: From part (a), we know that $\sin(x)$ is concave on $[0, \frac{\pi}{2}]$, and thus every value lies below every tangent and above every chord that can be defined in the region.

In the region $[0, \frac{\pi}{2}]$, $\sin(x)$ can therefore be upper bounded by its tangent at 0 (the identity function $f(x) = x$) and lower bounded by the chord between $(0, \sin(0))$ and $(\pi/2, \sin(\pi/2))$ (the linear function $\frac{2}{\pi}x$).

Note that we could establish different upper and lower bounds as well; all values of $\sin(x)$ lie below any tangent line of the function, and values within the span of a chord lie above that chord.

