Convexity

Oscar Dalmau
dalmau@cimat.mx

Centro de Investigación en Matemáticas
CIMAT A.C. Mexico

February 4, 2020

Outline

Convex set

Definition:(Convex set) A subset $\Omega \subset \mathbb{R}^n$ is said to be convex if for every pair of points $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$ and $\alpha \in (0,1)$, the point $\boldsymbol{x} = \alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2$ is in $\mathbb{R}^n$, i.e., $\boldsymbol{x} \in \mathbb{R}^n$.

Convex set: Examples

Definition:(Convex set) A subset $\Omega \subset \mathbb{R}^n$ is said to be convex if for every pair of points $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$ and $\alpha \in (0,1)$, the point $\boldsymbol{x} = \alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2$ is in $\mathbb{R}^n$, i.e., $\boldsymbol{x} \in \mathbb{R}^n$.

- The interception of two convex sets is convex
- The segment line between two points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is convex.
- A line is a convex set
- A half-space is convex
- The set $\{\boldsymbol{x}|\ \boldsymbol{x}^T \mathbf{A} \boldsymbol{x} \leq r\}$, where $\mathbf{A}$ is positive definite, is a convex set

Convex set: Examples

- The interception of two convex sets is convex

Let $\mathbf{A}$, $\mathbf{B}$ be two convex sets. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{A} \cap \mathbf{B}$ then $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{A}$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{B}$ (by definition of $\cap$). Then, by the convexity of $\mathbf{A}$ and $\mathbf{B}$, for all $\alpha \in [0, 1]$

$$\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y} \in \mathbf{A} \quad \text{and} \quad \alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y} \in \mathbf{B}$$

Therefore, (by definition of $\cap$)

$$\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y} \quad \in \quad \mathbf{A} \cap \mathbf{B}$$

Then $\mathbf{A} \cap \mathbf{B}$ is convex!

Convex set: Half space

- A half-space $\mathbf{H}$ is convex

A half-space can be written as the set $\mathbf{H} = \{\boldsymbol{x}|\ \boldsymbol{a}^T\boldsymbol{x} \geq b\}$, for an appropriate vector $\boldsymbol{a}$ and real number b. Let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbf{H}$ then

$$\begin{aligned} \boldsymbol{a}^T\boldsymbol{x}_1 &\geq b \\ \boldsymbol{a}^T\boldsymbol{x}_2 &\geq b \end{aligned}$$

then for all $\alpha \in [0, 1]$

$$\begin{aligned} \alpha\boldsymbol{a}^T\boldsymbol{x}_1 &\geq \alpha b \\ (1-\alpha)\boldsymbol{a}^T\boldsymbol{x}_2 &\geq (1-\alpha)b \\ \boldsymbol{a}^T(\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) &\geq \alpha b + (1-\alpha)b = b \end{aligned}$$

Then $\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2 \in \mathbf{H}$ and we conclude that $\mathbf{H}$ is convex

Convex set: Examples

- The set $\mathbf{H} = \{\boldsymbol{x} | \; \boldsymbol{x}^T \mathbf{A} \boldsymbol{x} \leq r\}$, where $\mathbf{A}$ is positive definite, is a convex set

We need to prove that $\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2 \in \mathbf{H}$ for $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbf{H}$ and for all $\alpha \in [0,1]$, i.e., we should prove that

$$[\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2]^T \mathbf{A} [\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2] \quad \leq \quad r$$

First, we note that $\boldsymbol{x}_1^T \mathbf{A} \boldsymbol{x}_2 \leq r$

$$\begin{aligned}
[\boldsymbol{x}_2 - \boldsymbol{x}_1]^T \mathbf{A} [\boldsymbol{x}_2 - \boldsymbol{x}_1] &\geq 0 \\
\boldsymbol{x}_2^T \mathbf{A} \boldsymbol{x}_2 + \boldsymbol{x}_1^T \mathbf{A} \boldsymbol{x}_1 &\geq 2\boldsymbol{x}_1^T \mathbf{A} \boldsymbol{x}_2 \\
\boldsymbol{x}_1^T \mathbf{A} \boldsymbol{x}_2 &\leq \frac{\boldsymbol{x}_2^T \mathbf{A} \boldsymbol{x}_2 + \boldsymbol{x}_1^T \mathbf{A} \boldsymbol{x}_1}{2}
\end{aligned}$$

then $\boldsymbol{x}_1^T \mathbf{A} \boldsymbol{x}_2 \leq r$ due to $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbf{H}$, ie, $\boldsymbol{x}_1^T \mathbf{A} \boldsymbol{x}_1 \leq r$ and $\boldsymbol{x}_2^T \mathbf{A} \boldsymbol{x}_2 \leq r$.

Convex set: Examples

- The set $\mathbf{H} = \{\boldsymbol{x} | \ \boldsymbol{x}^T \mathbf{A} \boldsymbol{x} \leq r\}$, where $\mathbf{A}$ is positive definite, is a convex set

We need to prove that $\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2 \in \mathbf{H}$ for $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbf{H}$ and for all $\alpha \in [0,1]$, i.e., we should prove that

$$[\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2]^T \mathbf{A} [\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2] \quad \leq \quad r$$

(comment) we can also prove that $\boldsymbol{x}_1^T \mathbf{A} \boldsymbol{x}_2 \leq r$ using the Cauchy-Schwartz inequality, i.e., defining $\boldsymbol{y}_i = \mathbf{A}^{1/2} x_i$

$$\begin{aligned} |\boldsymbol{x}_1^T \mathbf{A} \boldsymbol{x}_2| &= |\boldsymbol{y}_1^T \boldsymbol{y}_2| \leq \|\boldsymbol{y}_1\| \|\boldsymbol{y}_2\| \\ &= \sqrt{\|\boldsymbol{y}_1\|^2 \|\boldsymbol{y}_2\|^2} = \sqrt{\boldsymbol{y}_1^T \boldsymbol{y}_1 \boldsymbol{y}_2^T \boldsymbol{y}_2} \\ &= \sqrt{\boldsymbol{x}_1^T \mathbf{A} \boldsymbol{x}_1 \boldsymbol{x}_2^T \mathbf{A} \boldsymbol{x}_2} \leq \sqrt{r^2} = r \end{aligned}$$

Convex set: Examples

- The set $\mathbf{H} = \{\boldsymbol{x} |\ \boldsymbol{x}^T \mathbf{A} \boldsymbol{x} \leq r\}$, where $\mathbf{A}$ is positive definite, is a convex set

We need to prove that $\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2 \in \mathbf{H}$ for $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbf{H}$ and for all $\alpha \in [0,1]$, i.e., we should prove that

$$\begin{aligned} [\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2]^T \mathbf{A} [\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2] &= \alpha^2 \boldsymbol{x}_1^T \mathbf{A} \boldsymbol{x}_1 + (1-\alpha)^2 \boldsymbol{x}_2^T \mathbf{A} \boldsymbol{x}_2 \\ &\quad + 2\alpha(1-\alpha)\boldsymbol{x}_1^T \mathbf{A} \boldsymbol{x}_2 \\ &\leq [\alpha^2 + 2\alpha(1-\alpha) + (1-\alpha)^2]r \\ &\leq [\alpha + (1-\alpha)]^2]r = r \end{aligned}$$

then $\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2 \in \mathbf{H}$, ie, $\mathbf{H}$ is convex

Convex function

Definition: (Convex function)

1. A function $f(\boldsymbol{x})$ defined over a convex set Ω is said to be convex if for every pair of points $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$ and $\alpha \in (0, 1)$, it holds the inequality

$$f[\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2] \quad \leq \quad \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2),$$

   If $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$ and

$$f[\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2] \quad < \quad \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2)$$

   then $f(\boldsymbol{x})$ is said to be strictly convex.

2. If $\phi(\boldsymbol{x})$ is defined over a convex set Ω and $f(\boldsymbol{x}) = -\phi(\boldsymbol{x})$ is convex, then $\phi(\boldsymbol{x})$ is said to be concave. If $f(\boldsymbol{x})$ is strictly convex, $\phi(\boldsymbol{x})$ is strictly concave.

Convex function: Jensen's inequality

- Jensen's inequality: If $f(\cdot)$ is a real convex function, $x_1, x_2, \ldots, x_n$ are numbers in its domain, and $a_1, a_2, \ldots, a_n$, positive weights then:

$$f\left(\frac{\sum_{i=1}^{n} a_i x_i}{\sum_{i=1}^{n} a_i}\right) \leq \frac{\sum_{i=1}^{n} a_i f(x_i)}{\sum_{i=1}^{n} a_i}$$

- If $a_i = \frac{1}{n}$ then $\sum_{i=1}^{n} a_i = 1$ and

$$f\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) \leq \frac{\sum_{i=1}^{n} f(x_i)}{n}$$

- If $x_1, x_2, \ldots, x_n$ are positive numbers then

$$\frac{x_1 + x_2 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \cdots x_n}$$

by taking $f(x) = -\log(x)$ which is a convex function.

Convexity of linear combination of convex functions

Theorem 2.1

If $f_1(\boldsymbol{x}), f_2(\boldsymbol{x})$ are convex functions on the convex set Ω then $f(\boldsymbol{x}) = af_1(\boldsymbol{x}) + bf_2(\boldsymbol{x})$, with $a, b \geq 0$, is convex on the set Ω.

Proof: Let $\alpha \in (0,1)$ then

$$\begin{aligned} f_1(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) &\leq \alpha f_1(\boldsymbol{x}) + (1-\alpha) f_1(\boldsymbol{y}) \\ f_2(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) &\leq \alpha f_2(\boldsymbol{x}) + (1-\alpha) f_2(\boldsymbol{y}) \end{aligned}$$

$$\begin{aligned} f(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) &= af_1(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) + bf_2(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \\ &\leq a(\alpha f_1(\boldsymbol{x}) + (1-\alpha) f_1(\boldsymbol{y})) + \\ &\quad\; b(\alpha f_2(\boldsymbol{x}) + (1-\alpha) f_2(\boldsymbol{y})) \\ &= \alpha(af_1(\boldsymbol{x}) + bf_2(\boldsymbol{x})) + (1-\alpha)(af_1(\boldsymbol{y}) + bf_2(\boldsymbol{y})) \\ &= \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) \end{aligned}$$

Then $f(\boldsymbol{x})$ is convex

Relation between convex functions and convex sets

Theorem 2.2

If $f(\boldsymbol{x})$ is a convex function on a convex set Ω, then the set $S = \{\boldsymbol{x} : \boldsymbol{x} \in \Omega, f(\boldsymbol{x}) \leq l\}$ is convex for every real number l.

Proof: If $\boldsymbol{x}, \boldsymbol{y} \in \Omega$, then $f(\boldsymbol{x}) \leq l$ and $f(\boldsymbol{y}) \leq l$ from the definition of Ω. Since $f(\boldsymbol{x})$is convex

$$f(\alpha\boldsymbol{y} + (1-\alpha)\boldsymbol{x}) \leq \alpha f(\boldsymbol{y}) + (1-\alpha)f(\boldsymbol{x}) \leq \alpha l + (1-\alpha)l \leq l$$

for $\hat{\boldsymbol{x}} = \alpha\boldsymbol{y} + (1-\alpha)\boldsymbol{x}$ and $\alpha \in (0, 1)$, Therefore $\hat{\boldsymbol{x}} \in \Omega$, then Ω is convex.

Relation between convex functions and convex sets

Theorem 2.2

If $f(\boldsymbol{x})$ is a convex function on a convex set Ω, then the set $S = \{\boldsymbol{x} : \boldsymbol{x} \in \Omega, f(\boldsymbol{x}) \leq l\}$ is convex for every real number l.

Note: If the function $f(\boldsymbol{x}) = \boldsymbol{x}^T \mathbf{A} \boldsymbol{x}$, where $\mathbf{A}$ is positive definite matrix, is a convex function then the set $H = \{\boldsymbol{x} | \ \boldsymbol{x}^T \mathbf{A} \boldsymbol{x} \leq r\}$ is also convex.

Relation between convex functions and convex sets

Theorem 2.2

If $f(\boldsymbol{x})$ is a convex function on a convex set Ω, then the set $S = \{\boldsymbol{x} : \boldsymbol{x} \in \Omega, f(\boldsymbol{x}) \leq l\}$ is convex for every real number l.

$f(\boldsymbol{x})$ is convex due to

$$
\begin{aligned}
f[\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}] &= [\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}]^T \mathbf{A}[\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}] \\
&= \alpha^2\boldsymbol{x}^T\mathbf{A}\boldsymbol{x} + (1-\alpha)^2\boldsymbol{y}^T\mathbf{A}\boldsymbol{y} \\
&\quad +2\alpha(1-\alpha)\boldsymbol{x}^T\mathbf{A}\boldsymbol{y} \\
&\leq \alpha^2\boldsymbol{x}^T\mathbf{A}\boldsymbol{x} + (1-\alpha)^2\boldsymbol{y}^T\mathbf{A}\boldsymbol{y} \\
&\quad +\alpha(1-\alpha)(\boldsymbol{x}^T\mathbf{A}\boldsymbol{x} + \boldsymbol{y}^T\mathbf{A}\boldsymbol{y}) \\
&= \alpha\boldsymbol{x}^T\mathbf{A}\boldsymbol{x} + (1-\alpha)\boldsymbol{y}^T\mathbf{A}\boldsymbol{y} \\
&= \alpha f(\boldsymbol{x}) + (1-\alpha)f(\boldsymbol{y})
\end{aligned}
$$

Property of convex functions relating to gradient

Theorem 2.3

If $f(\boldsymbol{x}) \in \mathcal{C}^1$, then $f(\boldsymbol{x})$ is convex over a convex set Ω if and only if $f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + g(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x})$ for all $\boldsymbol{x}$ and $\boldsymbol{y} \in \Omega$, where $\boldsymbol{g}(\boldsymbol{x})$ is the gradient of $f(\boldsymbol{x})$.

Proof: The proof of this theorem consists of two parts. First we prove that if $f(\boldsymbol{x})$ is convex, the inequality holds. Then we prove that if the inequality holds, $f(\boldsymbol{x})$ is convex. The two parts constitute the necessary and sufficient conditions of the theorem.

Property of convex functions relating to gradient

Theorem

If $f(\boldsymbol{x}) \in \mathcal{C}^1$, then $f(\boldsymbol{x})$ is convex over a convex set Ω if and only if $f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + g(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x})$ for all $\boldsymbol{x}$ and $\boldsymbol{y} \in \Omega$, where $\boldsymbol{g}(\boldsymbol{x})$ is the gradient of $f(\boldsymbol{x})$.

Proof: (a) If $f(\boldsymbol{x})$ is convex, then for all α in the range $0 < \alpha < 1$

$$
\begin{aligned}
f(\alpha\boldsymbol{y} + (1-\alpha)\boldsymbol{x}) &\leq \alpha f(\boldsymbol{y}) + (1-\alpha) f(\boldsymbol{x}) \\
f(\boldsymbol{x} + \alpha(\boldsymbol{y} - \boldsymbol{x})) &\leq \alpha[f(\boldsymbol{y}) - f(\boldsymbol{x})] + f(\boldsymbol{x}) \\
\frac{f(\boldsymbol{x} + \alpha(\boldsymbol{y} - \boldsymbol{x})) - f(\boldsymbol{x})}{\alpha} &\leq f(\boldsymbol{y}) - f(\boldsymbol{x}) \\
\lim_{\alpha \to 0} \frac{f(\boldsymbol{x} + \alpha(\boldsymbol{y} - \boldsymbol{x})) - f(\boldsymbol{x})}{\alpha} &\leq f(\boldsymbol{y}) - f(\boldsymbol{x})
\end{aligned}
$$

Property of convex functions relating to gradient

Theorem

If $f(\boldsymbol{x}) \in \mathcal{C}^1$, then $f(\boldsymbol{x})$ is convex over a convex set Ω if and only if $f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + g(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x})$ for all $\boldsymbol{x}$ and $\boldsymbol{y} \in \Omega$, where $\boldsymbol{g}(\boldsymbol{x})$ is the gradient of $f(\boldsymbol{x})$.

Proof: (a) If $f(\boldsymbol{x})$ is convex, then for all α in the range $0 < \alpha < 1$

$$\begin{aligned}
\lim_{\alpha \to 0} \frac{f(\boldsymbol{x} + \alpha(\boldsymbol{y} - \boldsymbol{x})) - f(\boldsymbol{x})}{\alpha} &\leq f(\boldsymbol{y}) - f(\boldsymbol{x}) \\
\nabla f(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) &\leq f(\boldsymbol{y}) - f(\boldsymbol{x}) \\
f(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) &\leq f(\boldsymbol{y})
\end{aligned}$$

This concludes the first part, i.e. the inequality holds.

Property of convex functions relating to gradient

(b) As the inequality $f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + g(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x})$ holds, then, for $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \Omega$

$$\begin{aligned} f(\boldsymbol{x}) + g(\boldsymbol{x})^T(\boldsymbol{y}_1 - \boldsymbol{x}) &\leq f(\boldsymbol{y}_1) \\ f(\boldsymbol{x}) + g(\boldsymbol{x})^T(\boldsymbol{y}_2 - \boldsymbol{x}) &\leq f(\boldsymbol{y}_2) \end{aligned}$$

therefore

$$\begin{aligned} \alpha f(\boldsymbol{x}) + \alpha g(\boldsymbol{x})^T(\boldsymbol{y}_1 - \boldsymbol{x}) &\leq \alpha f(\boldsymbol{y}_1) \\ (1-\alpha) f(\boldsymbol{x}) + (1-\alpha) g(\boldsymbol{x})^T(\boldsymbol{y}_2 - \boldsymbol{x}) &\leq (1-\alpha) f(\boldsymbol{y}_2) \\ f(\boldsymbol{x}) + g(\boldsymbol{x})^T(\alpha\boldsymbol{y}_1 + (1-\alpha)\boldsymbol{y}_2 - \boldsymbol{x})) &\leq \alpha f(\boldsymbol{y}_1) + (1-\alpha) f(\boldsymbol{y}_2) \end{aligned}$$

taking $\boldsymbol{x} = \alpha\boldsymbol{y}_1 + (1-\alpha)\boldsymbol{y}_2$ one concludes that $f(\alpha\boldsymbol{y}_1 + (1-\alpha)\boldsymbol{y}_2) \leq \alpha f(\boldsymbol{y}_1) + (1-\alpha) f(\boldsymbol{y}_2)$ i.e., $f(\boldsymbol{x})$ is convex.

Property of convex functions relating to the Hessian

Theorem 2.4

A function $f(\boldsymbol{x}) \in \mathcal{C}^2$ is convex over a convex set Ω if and only if the Hessian $\mathbf{H}(\boldsymbol{x})$ of $f(\boldsymbol{x})$ is positive semidefinite for $\boldsymbol{x} \in \Omega$.

Proof: (a) If $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{d}$ where $\boldsymbol{y}$ and $\boldsymbol{x}$ are arbitrary points in Ω, then the Taylor series yields

$$f(\boldsymbol{y}) = f(\boldsymbol{x}) + g(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) + \frac{1}{2}\boldsymbol{d}^T\mathbf{H}(\boldsymbol{x} + \alpha\boldsymbol{d})\boldsymbol{d}$$

where $0 \leq \alpha \leq 1$.
Now if $\mathbf{H}(\boldsymbol{x})$ is positive semidefinite everywhere in Ω, then $\frac{1}{2}\boldsymbol{d}^T\mathbf{H}(\boldsymbol{x} + \alpha\boldsymbol{d})\boldsymbol{d} \geq 0$ then $f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + g(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x})$.
Therefore, from a previous Theorem, $f(\boldsymbol{x})$ is convex.

Property of convex functions relating to the Hessian

Theorem

A function $f(\boldsymbol{x}) \in \mathcal{C}^2$ is convex over a convex set Ω if and only if the Hessian $\mathbf{H}(\boldsymbol{x})$ of $f(\boldsymbol{x})$ is positive semidefinite for $\boldsymbol{x} \in \Omega$.

Proof: (b) by contradiction, If $\mathbf{H}(\boldsymbol{x})$ is not positive semidefinite everywhere in Ω, then there exist a point $\boldsymbol{x}$ and at least a $\boldsymbol{d}$ such that $\boldsymbol{d}^T\mathbf{H}(\boldsymbol{x} + \alpha\boldsymbol{d})\boldsymbol{d} < 0$ using also the Taylor's theorem

$$f(\boldsymbol{y}) = f(\boldsymbol{x}) + g(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) + \frac{1}{2}\boldsymbol{d}^T\mathbf{H}(\boldsymbol{x} + \alpha\boldsymbol{d})\boldsymbol{d}$$

one concludes that

$$f(\boldsymbol{y}) < f(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x})$$

and therefore, from a previous Theorem, $f(\boldsymbol{x})$ is nonconvex. Finally, $f(\boldsymbol{x})$ is convex if and only if $\mathbf{H}(\boldsymbol{x})$ is positive semidefinite everywhere in Ω.

Properties of strictly convex functions

For a strictly **strictly convex**, we have the following theorem, which is analogous to the previous theorems.

Theorem 2.5

1. *If $f(\boldsymbol{x})$ is a* **strictly convex** *function on a convex set Ω, then the set $S = \{\boldsymbol{x} : \boldsymbol{x} \in \Omega \text{ for } f(\boldsymbol{x}) < K\}$ is convex for every real number K.*
2. *If $f(\boldsymbol{x}) \in \mathcal{C}^1$, then $f(\boldsymbol{x})$ is* **strictly convex** *over a convex set if and only if $f(\boldsymbol{y}) > f(\boldsymbol{x}) + g(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x})$ for all $\boldsymbol{x}$ and $\boldsymbol{y} \in \Omega$ where $\boldsymbol{g}(\boldsymbol{x})$ is the gradient of $f(\boldsymbol{x})$.*
3. *A function $f(\boldsymbol{x}) \in \mathcal{C}^2$ is* **strictly convex** *over a convex set Ω if and only if the Hessian $\mathbf{H}(\boldsymbol{x})$ is positive definite for $\boldsymbol{x} \in \Omega$.*

Relation between local and global minimizers in convex functions

Theorem 3.1

If $f(\boldsymbol{x})$ is a convex function defined on a convex set Ω, then

1. *the set of points S where $f(\boldsymbol{x})$ is minimum is convex;*
2. *any local minimizer of $f(\boldsymbol{x})$ is a global minimizer.*

Note: It is assumed that the set S of minimums is not empty

Relation between local and global minimizers in convex functions

Theorem

If $f(\boldsymbol{x})$ is a convex function defined on a convex set Ω, then

1. the set of points S where $f(\boldsymbol{x})$ is minimum is convex;
2. any local minimizer of $f(\boldsymbol{x})$ is a global minimizer.

Proof (a) If f^* is a minimum value of $f(\boldsymbol{x})$, then $S = \{\boldsymbol{x} : f(\boldsymbol{x}) \leq f^*,\ \boldsymbol{x} \in \Omega\}$ is convex by virtue of a previous Theorem, due to $f(\boldsymbol{x})$ is convex.

Relation between local and global minimizers in convex functions

Theorem

If $f(\boldsymbol{x})$ is a convex function defined on a convex set Ω, then

1. the set of points S where $f(\boldsymbol{x})$ is minimum is convex;
2. any local minimizer of $f(\boldsymbol{x})$ is a global minimizer.

Proof

(b) Let $\boldsymbol{x}^* \in \Omega$ be a local minimizer and for all $\boldsymbol{y} \in \Omega$. Then, we can choose $\alpha \in (0, 1]$ such that

$$f(\boldsymbol{x}^*) \quad \leq \quad f[\boldsymbol{x}^* + \alpha(\boldsymbol{y} - \boldsymbol{x}^*)]$$

ie, when $\boldsymbol{x}^* + \alpha(\boldsymbol{y} - \boldsymbol{x}^*)$ if sufficiently close $\boldsymbol{x}^*$

Relation between local and global minimizers in convex functions

Theorem

If $f(\boldsymbol{x})$ is a convex function defined on a convex set Ω, then

1. the set of points S where $f(\boldsymbol{x})$ is minimum is convex;
2. any local minimizer of $f(\boldsymbol{x})$ is a global minimizer.

Proof

(b)

$$f(\boldsymbol{x}^*) \leq f[\boldsymbol{x}^* + \alpha(\boldsymbol{y} - \boldsymbol{x}^*)]$$

by convexity and simplifying

$$\begin{aligned} f(\boldsymbol{x}^*) &\leq (1-\alpha)f(\boldsymbol{x}^*) + \alpha f(\boldsymbol{y}) \\ \alpha f(\boldsymbol{x}^*) &\leq \alpha f(\boldsymbol{y}) \end{aligned}$$

then $f(\boldsymbol{x}^*) \leq \alpha f(\boldsymbol{y})$ for all $\boldsymbol{y} \in \Omega$, ie, $\boldsymbol{x}^*$ is a global minimizer.

Relation between local and global minimizers in convex functions

Theorem

If $f(\boldsymbol{x})$ is a convex function defined on a convex set Ω, then

1. the set of points S where $f(\boldsymbol{x})$ is minimum is convex;
2. any local minimizer of $f(\boldsymbol{x})$ is a global minimizer.

Summary

The local minimizers are located in a convex set, and all are global minimizers.

Existence of a global minimizer in convex functions

Theorem 3.2

If $f(\boldsymbol{x}) \in \mathcal{C}^1$ is a convex function on a convex set Ω and there is a point $\boldsymbol{x}^*$ such that $\boldsymbol{g}(\boldsymbol{x}^*)^T \boldsymbol{d} \geq 0$ where $\boldsymbol{d} = \boldsymbol{x} - \boldsymbol{x}^*$ for all $\boldsymbol{x} \in \Omega$, then $\boldsymbol{x}^*$ is a global minimizer of $f(\boldsymbol{x})$.

Proof As $f(\boldsymbol{x})$ is a convex function

$$f(\boldsymbol{x}) \quad \geq \quad f(\boldsymbol{x}^*) + \boldsymbol{g}(\boldsymbol{x}^*)^T(\boldsymbol{x} - \boldsymbol{x}^*)$$

Since $\boldsymbol{g}(\boldsymbol{x}^*)^T(\boldsymbol{x} - \boldsymbol{x}^*) \geq 0$ we have $f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$ and so $\boldsymbol{x}^*$ is a local minimizer. By the previous Theorem, $\boldsymbol{x}^*$ is also a global minimizer.

Note: Similarly, if $f(\boldsymbol{x})$ is a *strictly convex function* and $\boldsymbol{g}(\boldsymbol{x}^*)^T \boldsymbol{d} > 0$ then $\boldsymbol{x}^*$ is a *strong global minimizer*.

Location of maximum of a convex function

Theorem 3.3

If $f(\boldsymbol{x})$ is a convex function defined on a bounded, closed, convex set Ω, then if $f(\boldsymbol{x})$ has a maximum over Ω, it occurs at the boundary of Ω.

Proof By contradiction, suppose the there is not a maximum at the boundary of Ω, then the maximum is a point $\boldsymbol{x}^*$ in the interior of Ω.
For any point $\boldsymbol{x}$ on the boundary of Ω we can draw a line through $\boldsymbol{x}^*$ which intersects the boundary at a two point $\boldsymbol{y}$, since Ω is bounded and closed. Then $f(\boldsymbol{x}^*) > f(\boldsymbol{x})$ and $f(\boldsymbol{x}^*) > f(\boldsymbol{y})$
On the other hand, we can find $\alpha \in (0, 1)$ such that $\boldsymbol{x}^* = \alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y}$ and since $f(\cdot)$ is convex

$$f(\boldsymbol{x}^*) \quad = \quad f[\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}] \leq \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) < f(\boldsymbol{x}^*)$$

which is a contradiction!

Location of maximum of a convex function

Theorem 3.4

If $f(\boldsymbol{x})$ is a convex function defined on a bounded, closed, convex set Ω, then if $f(\boldsymbol{x})$ has a maximum over Ω, it occurs at the boundary of Ω.

Proof If point $\boldsymbol{x}$ is in the interior of Ω, a line can be drawn through $\boldsymbol{x}$ which intersects the boundary at two points, say, $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, since Ω is bounded and closed. Since $f(\boldsymbol{x})$ is convex, some α exists in the range $0 < \alpha < 1$ such that
$\boldsymbol{x} = \alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2$, $f(\boldsymbol{x}) \leq \alpha f(\boldsymbol{x}_1) + (1-\alpha)f(\boldsymbol{x}_2)$

- If $f(\boldsymbol{x}_1) > f(\boldsymbol{x}_2)$, then
  $f(\boldsymbol{x}) \leq \alpha f(\boldsymbol{x}_1) + (1-\alpha)f(\boldsymbol{x}_2) < f(\boldsymbol{x}_1)$

and the maximum $\boldsymbol{x}_1$ is at the boundary.

Location of maximum of a convex function

- If $f(\boldsymbol{x}_1) < f(\boldsymbol{x}_2)$ then $f(\boldsymbol{x}) \leq \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2) < f(\boldsymbol{x}_2)$ and the maximum $\boldsymbol{x}_2$ is at the boundary.
- If $f(\boldsymbol{x}_1) = f(\boldsymbol{x}_2)$ $f(\boldsymbol{x}) \leq \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2) \leq f(\boldsymbol{x}_1)$ and $f(\boldsymbol{x}) \leq \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2) \leq f(\boldsymbol{x}_2)$ and the maximum $\boldsymbol{x}_1, \boldsymbol{x}_2$ are at the boundary.

Evidently, in all possibilities the maximizers occur on the boundary

Location of maximum of a convex function

Proof Since $f(\boldsymbol{x})$ is convex,

$$f(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) = f(\boldsymbol{x}^*) \quad \leq \quad \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2).$$

As $f(\boldsymbol{x}_1) < f(\boldsymbol{x}^*)$ and $f(\boldsymbol{x}_2) \leq f(\boldsymbol{x}^*)$ one obtains that

$$f(\boldsymbol{x}^*) \quad \leq \quad \alpha f(\boldsymbol{x}_1) + (1-\alpha) f(\boldsymbol{x}_2) < \alpha f(\boldsymbol{x}^*) + (1-\alpha) f(\boldsymbol{x}^*) = f(\boldsymbol{x}^*)$$

which is a contradiction! Therefore, if $f(\boldsymbol{x})$ has a maximum over Ω, then it should be located at the boundary of Ω.