Convex Functions: Part 1

Lecture 7-1 - CMSE 382

Prof. Elizabeth Munch

Michigan State University
::
Dept of Computational Mathematics, Science & Engineering

Fri, Feb 27, 2026

This Lecture

Topics:

- Definition (vid),
- First order characterization (vid),
- Second order characterization (vid),
- Operations preserving convexity
  - Scaling, Summation, and Affine transformation (vid),
  - Composition (vid),
  - Point-wise maximum (vid)

Announcements:

- Enjoy Spring Break!

Section 1

Definition of Convex Function

Definition of convex functions

Definition

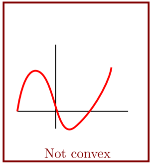
A function $f : C \to \mathbb{R}$ defined on a convex set $C \subset \mathbb{R}^n$ is

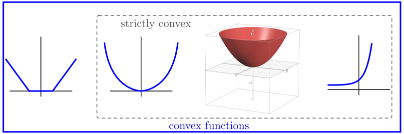
- **convex** if and only if

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$
$$\forall \mathbf{x}, \mathbf{y} \in C, \lambda \in [0,1].$$

- **strictly convex** if and only if

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) < \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$
$$\forall \mathbf{x} \neq \mathbf{y} \in C, \lambda \in (0,1).$$





Jensen inequality

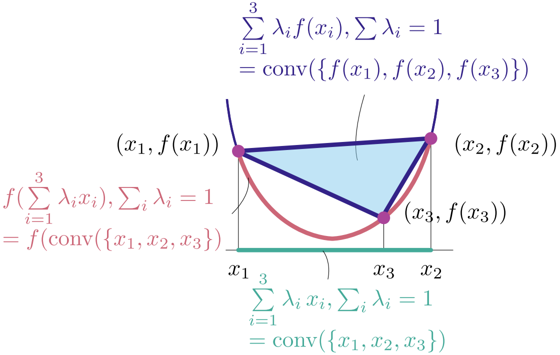
Theorem

Let $f : C \to \mathbb{R}$ be a convex function. Then we have

$$f\left(\sum_{i=1}^{m} \lambda_i \mathbf{x}_i\right) \leq \sum_{i=1}^{m} \lambda_i f(\mathbf{x}_i),$$

if $\lambda \in \Delta_m$ and $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in C$.

- $\lambda = (\lambda_1, \ldots, \lambda_m) \in \Delta_m$ means $\lambda_i \geq 0$ for all i and $\sum_{i=1}^{m} \lambda_i = 1$.



Example: Affine functions are convex

Example: Is the function (called affine function) $f(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} + b$, where $\mathbf{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}$ convex?

Answer:

Recall: A function $f : C \to \mathbb{R}$ defined on a convex set $C \subset \mathbb{R}^n$ is **convex** if and only if

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \le \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in C, \lambda \in [0,1], \textstyle\sum_i \lambda_i = 1.$$

$$\begin{aligned}
f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) &= \mathbf{a}^\top(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) + b \\
&= \lambda \mathbf{a}^\top \mathbf{x} + (1-\lambda)\mathbf{a}^\top \mathbf{y} + b \text{ (but } b = \lambda b + (1-\lambda) b) \\
&= \lambda \mathbf{a}^\top \mathbf{x} + (1-\lambda)\mathbf{a}^\top \mathbf{y} + \lambda b + (1-\lambda) b \\
&= \lambda(\mathbf{a}^\top \mathbf{x} + b) + (1-\lambda)(\mathbf{a}^\top \mathbf{y} + b) \\
&= \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})
\end{aligned}$$

So an affine function is convex.

Section 2

First order characterization

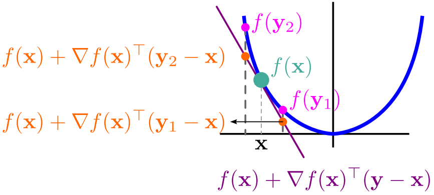
The gradient inequality

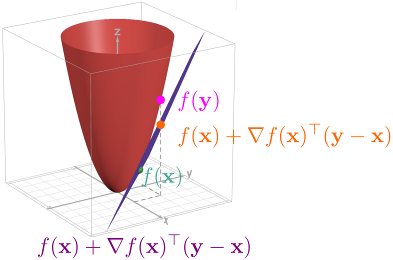
Theorem

If $f : C \to \mathbb{R}$ is continuously differentiable then:

- f *is convex if and only if* $f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^\top(\mathbf{y} - \mathbf{x}), \forall \mathbf{x}, \mathbf{y} \in C$.
- f *is strictly convex if and only if* $f(\mathbf{y}) > f(\mathbf{x}) + \nabla f(\mathbf{x})^\top(\mathbf{y} - \mathbf{x}), \forall \mathbf{x} \neq \mathbf{y} \in C$.

The tangent hyperplanes of convex functions are always underestimates of the function.

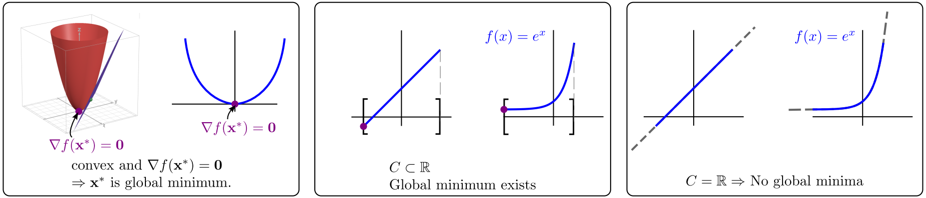




Theorem

Given a continuously differentiable convex function $f : C \to \mathbb{R}$ over convex set $C \subseteq \mathbb{R}^n$.

- *If $\nabla f(\mathbf{x}^*) = 0$ for some $\mathbf{x}^*$, then $\mathbf{x}^*$ is a global minimizer.*
- *If $C = \mathbb{R}^n$, then $\nabla f(\mathbf{x}^*) = 0$ if and only if $\mathbf{x}^*$ is a global minimum of f over $\mathbb{R}^n$.*



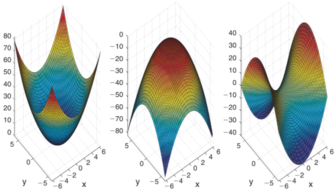
Convexity of quadratic functions

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}$ be the quadratic function given by $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} + 2b^T\mathbf{x} + c$, where $A \in \mathbb{R}^{n\times n}$ is symmetric, $b \in \mathbb{R}^n$, and $c \in \mathbb{R}$. Then:

- *f is convex if and only if $A \succeq 0$.*
- *f is strictly convex if and only if $A \succ 0$.*

$x^2 + y^2$ $\quad$ $-x^2 - y^2$ $\quad$ $x^2 - y^2$



A

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $\quad$ $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ $\quad$ $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Monotonicity of the gradient

Theorem

If $f : C \to \mathbb{R}$ is continuously differentiable over the convex set $C \subseteq \mathbb{R}^n$. Then f is convex over C if and only if

$$(\nabla f(\mathbf{x}) - \nabla f(\mathbf{y}))^\top (\mathbf{x} - \mathbf{y}) \geq 0 \; \forall x, y \in C$$

Section 3

Second order characterization

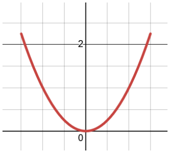
Second order characterization

Theorem

Let f be a twice continuously differentiable function over an open convex set $C \subseteq \mathbb{R}^n$.

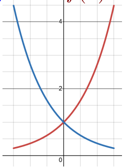
- *f is convex over C if and only if $\nabla^2 f(\mathbf{x}) \succeq 0$ for any $\mathbf{x} \in C$.*
- *If $\nabla^2 f(\mathbf{x}) \succ 0$ for any $\mathbf{x} \in C$, then f is strictly convex over C.*

$f(x) = x^2$



Strictly Convex
$\nabla f(x) = 2x$
$\nabla^2 f(x) = 2$

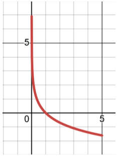
$f(x) = e^x$ $\quad f(x) = e^{-x}$



Strictly Convex

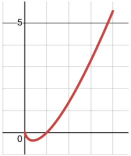
	e^x	e^{-x}
$\nabla f(x)$:	e^x	$-e^{-x}$
$\nabla^2 f(x)$:	e^x	e^{-x}

$f(x) = -\ln x$



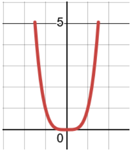
Strictly Convex
Over $x \in \mathbb{R}_{++}$
$\nabla f(x) = -1/x$
$\nabla^2 f(x) = 1/x^2$

$f(x) = x \ln x$



Strictly Convex
Over $x \in \mathbb{R}_{++}$
$\nabla f(x) = \ln x + 1$
$\nabla^2 f(x) = 1/x$

$f(x) = x^4$



Strictly Convex
$\nabla f = 4x^3$
$\nabla^2 f = 12x^2$
$\nabla^2 f(0) = 0$

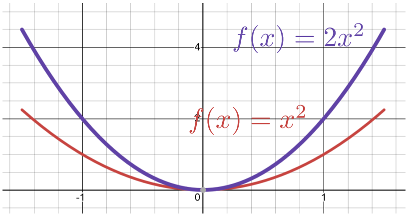
Section 4

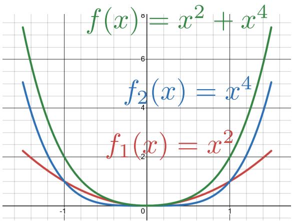
Operations Preserving Convexity

Convexity under summation and multiplication by non-negative scalars

Theorem

1. *Let f be a convex function defined over a convex set $C \subseteq \mathbb{R}^n$ and let $\alpha \geq 0$. Then αf is a convex function over C.*
2. *Let $f_1, f_2, \ldots, f_p$ be convex functions over a convex set $C \subseteq \mathbb{R}^n$. Then the sum function $f_1 + f_2 + \cdots + f_p$ is convex over C.*





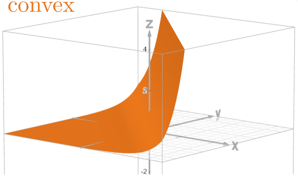
Theorem

Let $f : C \to \mathbb{R}$ be a convex function defined on a convex set $C \subseteq \mathbb{R}^n$. Let $A \in \mathbb{R}^{n\times m}$ and $\mathbf{b} \in \mathbb{R}^n$. Then the function g defined by

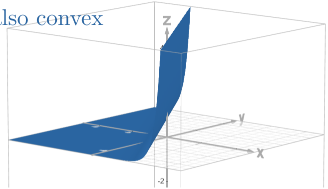
$$g(\mathbf{y}) = f(A\mathbf{y} + \mathbf{b}) \tag{1}$$

is convex over the convex set $D = \{\mathbf{y} \in \mathbb{R}^m : A\mathbf{y} + \mathbf{b} \in C\}$.

$f(x) = \exp(x + y)$
is convex



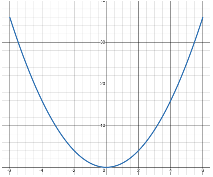
$f(x) = \exp(3x + 2y)$
also convex



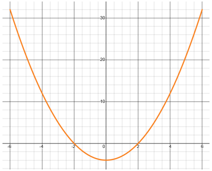
Convexity under composition

Is convexity preserved under composition? Not always!

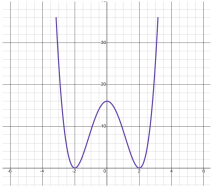
$g(x) = x^2$
(convex)



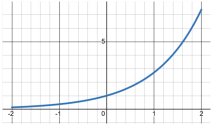
$h(x) = x^2 - 4$
(convex)



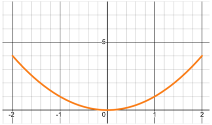
$g(h(x)) = (x^2 - 4)^2$
(NOT convex)



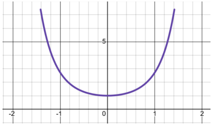
$g(x) = \exp(x)$
(convex)



$h(x) = x^2$
(convex)



$g(h(x)) = (\exp(x))^2$
(Convex)

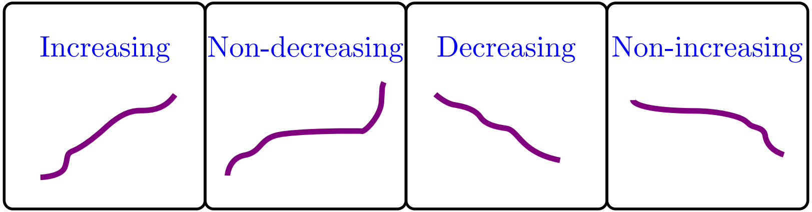


Monotonic function

Definition

A function $f : I \to \mathbb{R}$ where $I \subseteq \mathbb{R}$ is called:

- **Increasing** if for $x < y$ we have $f(x) < f(y)$ for all $x, y \in I$.
- **Non-decreasing** if for $x < y$ we have $f(x) \leq f(y)$ for all $x, y \in I$.
- **Decreasing** if for $x < y$ we have $f(x) > f(y)$ for all $x, y \in I$.
- **Non-increasing** if for $x < y$ we have $f(x) \geq f(y)$ for all $x, y \in I$.

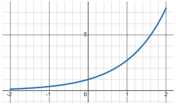


Convexity under composition with a non-decreasing convex function

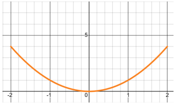
Theorem

Let $f : C \to \mathbb{R}$ be a convex function over the convex set $C \subseteq \mathbb{R}^n$. Let $g : I \to \mathbb{R}$ be a one-dimensional nondecreasing convex function over the interval $I \subseteq \mathbb{R}$. Assume that $f(C) \subseteq I$. Then $g(f(\mathbf{x}))$, $\mathbf{x} \in C$, is a convex function over C.

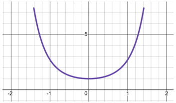
$g(x) = \exp(x)$
(convex)



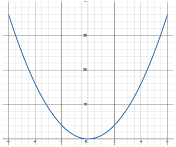
$h(x) = x^2$
(convex)



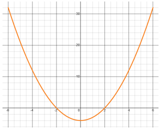
$g(h(x)) = (\exp(x))^2$
(Convex)



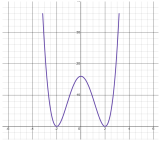
$g(x) = x^2$
(convex)



$h(x) = x^2 - 4$
(convex)



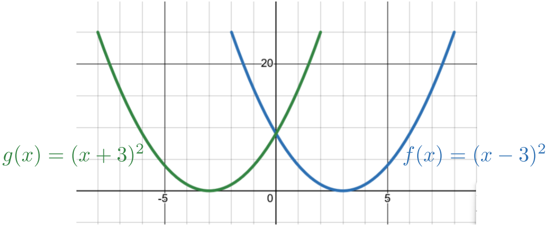
$g(h(x)) = (x^2 - 4)^2$
(NOT convex)

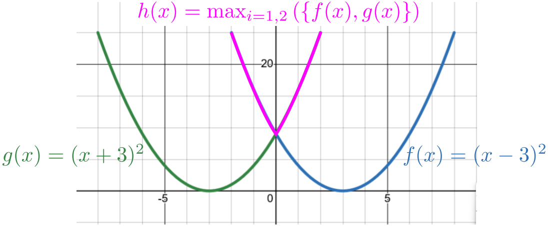


Pointwise maximum preserves convexity

Theorem

Let $f_1, \ldots, f_p : C \to \mathbb{R}$ be p convex functions over the convex set $C \subseteq \mathbb{R}^n$. Then the maximum function $f(\mathbf{x}) \equiv \max_{i=1,\ldots,p} f_i(\mathbf{x})$ is a convex function over C.





Groups - Round 3

Group 1
Lowell, Tianjuan, Lauryn, Atticus

Group 2
Alice, Aidan, Dev, Anthony

Group 3
Abigail, Michal, Breena, Andrew

Group 4
Kyle, Vinod, Dori, Joseph

Group 5
Yousif, Jamie, Jay, K.M Tausif

Group 6
Shanze, Saitej, Karen, Jack

Group 7
Arjun, Noah, Luis, Arya

Group 8
Morgan, Jonid, Sanskaar, Jake

Group 9
Quang Minh, Monirul Amin, Daniel, Ha

Group 10
Braedon, Dominic, Zheng, Lora

Group 11
Sai, Brandon, Purvi, Aaron

Group 12
Igor, Scott, Maye, Long