Convex Functions

Lecture 2 for 18660/18460: Optimization

Guannan Qu

January 15, 2026

Admin Stuff

- Quiz for today's lecture is available now and is due before next Tue's lecture

Quiz for Lecture 2 (due Jan 20 before lecture)
Not available until Jan 15 at 12:20pm **Due** Jan 20 at 11am 1 pt 1 Question

Admin Stuff

See Piazza Post for more details.

We will be holding our recitation of the semester Friday from 10-11AM ET. The first session is designed to review essential **Linear Algebra and Calculus** concepts to help you prepare for Homework 1.

Logistics

- **Date:** Friday, January 23
- **Time:**
  - 10:00 AM – 11:00 AM ET
  - 4:00 PM – 5:00 PM Kigali
- **Location:** Zoom (Same link as Office Hours)
  - **Link:** Join Zoom Meeting

Recall

Convex optimization

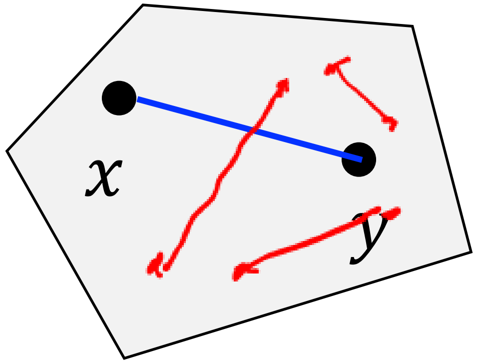
$$\min f(x) \quad \text{Convex objective functions}$$

$$\text{s.t.} \quad x \in C \quad \text{Convex constraint set}$$

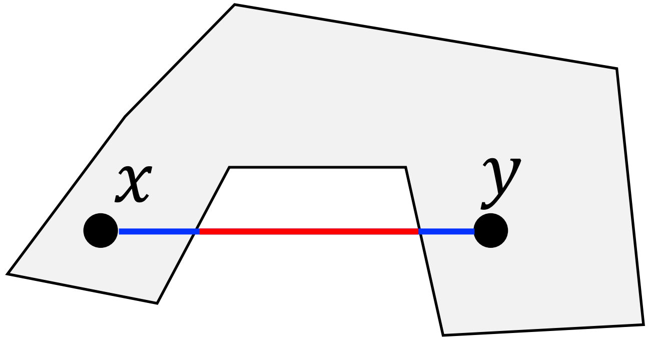
Recall

Convex set: set $C \subseteq \mathbb{R}^n$ such that

$$x, y \in C \Rightarrow tx + (1-t)y \in C, \text{ for all } t \in [0,1]$$



Convex



Nonconvex

Outline for Today

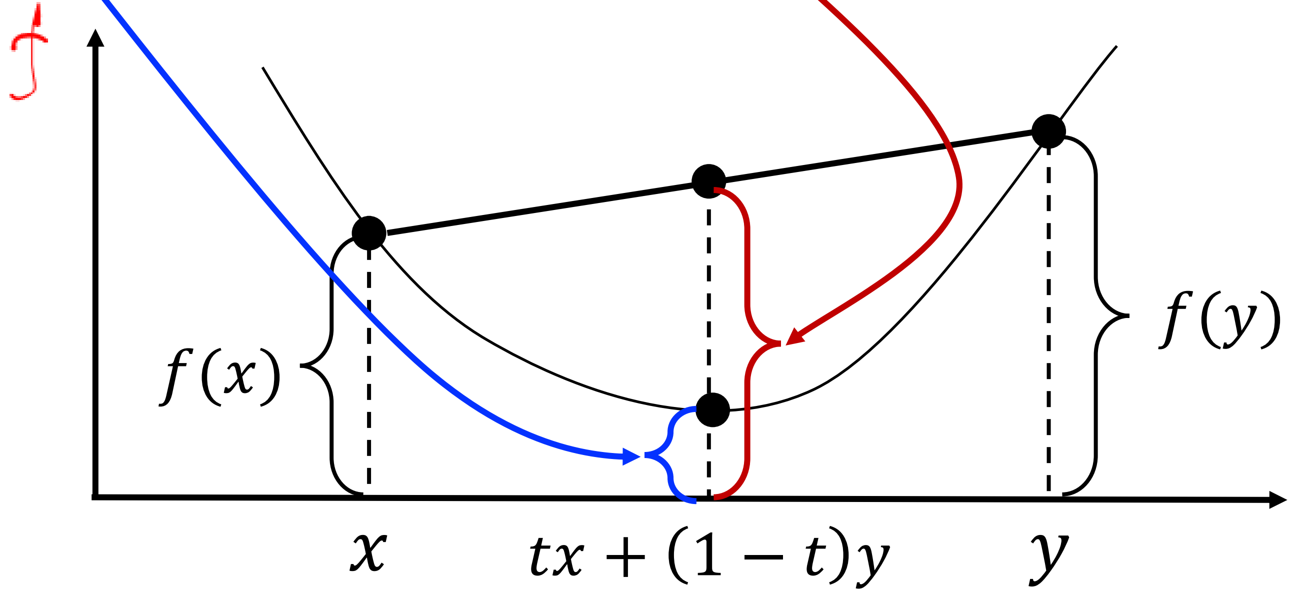
Convex functions

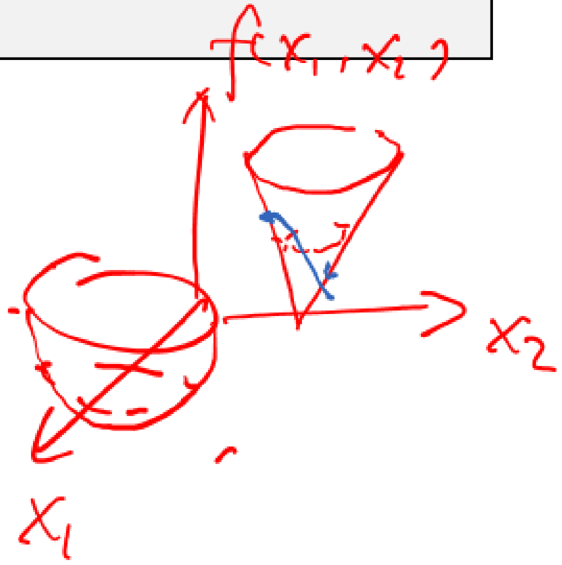
- Definition of convex functions
- Show that many commonly seen functions are convex
- Provide a “toolbox” that tells whether a function is convex or not

Recall

Convex function: function $f : \mathbb{R}^n \to \mathbb{R}$ such that $\mathrm{dom}(f) \subseteq \mathbb{R}^n$ is convex, and for all $x, y \in \mathrm{dom}(f)$

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y), \text{ for all } t \in [0,1]$$





Recall

Convex function: function $f: \mathbb{R}^n \to \mathbb{R}$ such that $\mathrm{dom}(f) \subseteq \mathbb{R}^n$ is convex, and for all $x, y \in \mathrm{dom}(f)$

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y), \text{ for all } t \in [0,1]$$

- We say $f: \mathbb{R}^n \to \mathbb{R}$ to mean f takes variables of dimension n
- However, f may not defined on all points in $\mathbb{R}^n$
  - E.g. $\log x$ and $\sqrt{x}$ are not defined on the entire $\mathbb{R}$
- $\mathrm{dom}(f)$ means the subset on which f is defined
  - E.g. $(0, +\infty)$ for $\log x$ and $[0, +\infty)$ for $\sqrt{x}$
- When $\mathrm{dom}(f)$ is the whole space $\mathbb{R}^n$, we don't state it explicitly.

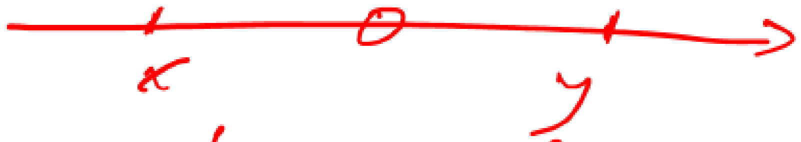
Recall

Convex function: function $f: \mathbb{R}^n \to \mathbb{R}$ such that $\text{dom}(f) \subseteq \mathbb{R}^n$ is convex, and for all $x, y \in \text{dom}(f)$

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y), \text{ for all } t \in [0,1]$$

$\hookrightarrow tx + (1-t)y \in \text{dom}(f)$

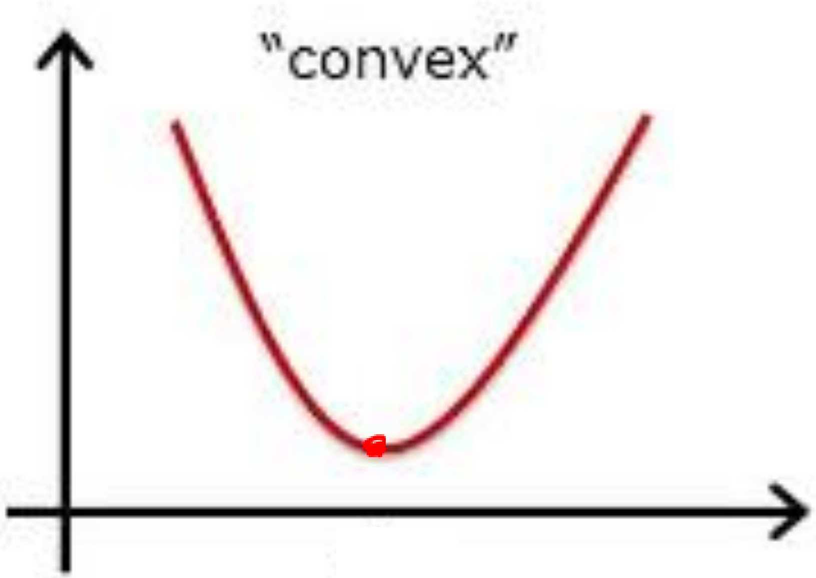
The domain has to be convex!



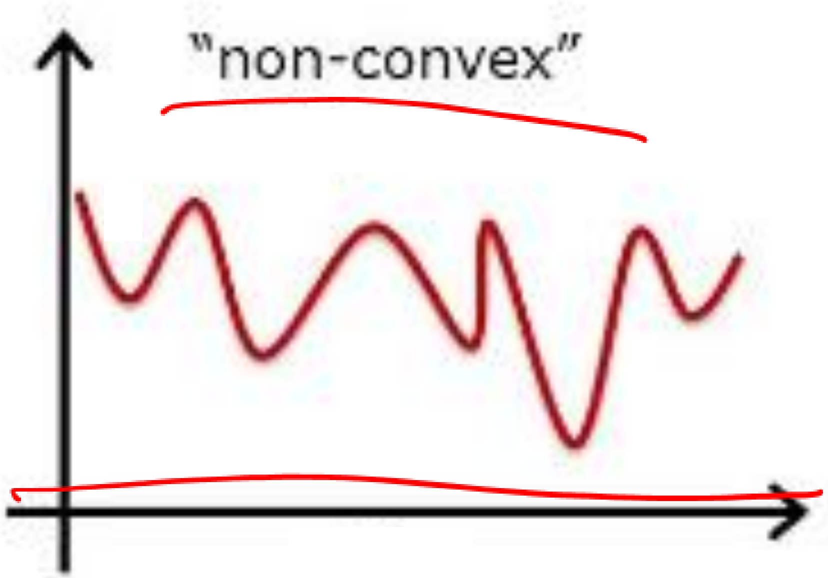
$\frac{1}{2}x + \frac{1}{2}y \notin \text{dom}(f)$

Why we care convexity?

Convex functions are easier to optimize!



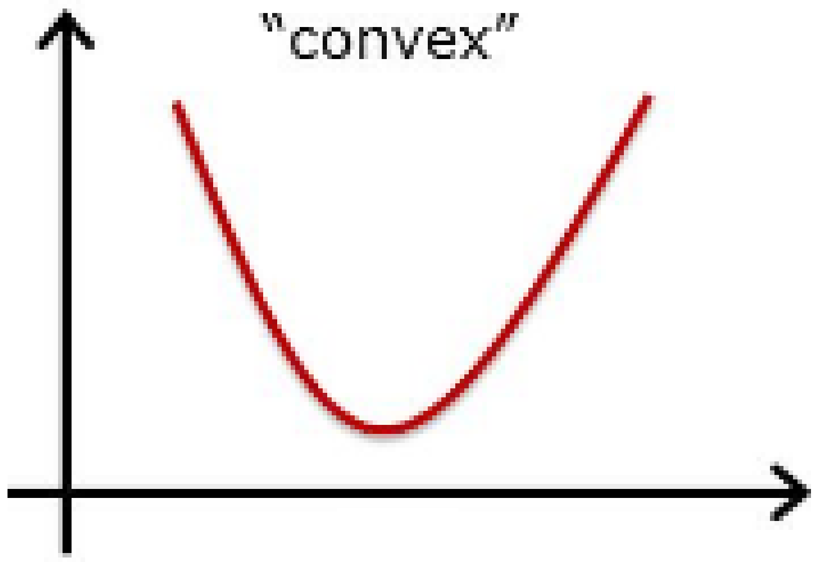
Single "Basin"



Multiple "Basins"

Why we care convexity?

Convex optimization already capture many real-world problems



Single "Basin"

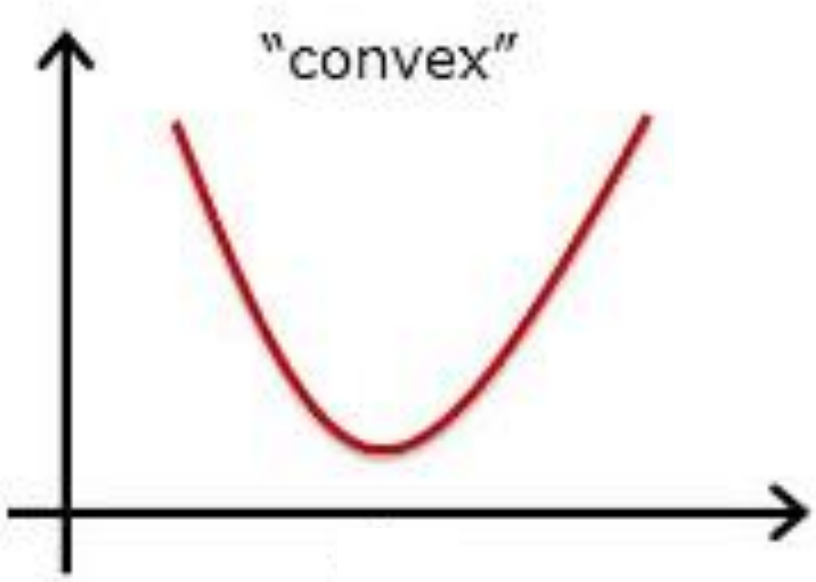
For example

- Common functions: linear, quadratic,…
- ML like linear/logistic regression, SVM…
- DC power flow, linear optimal control, …

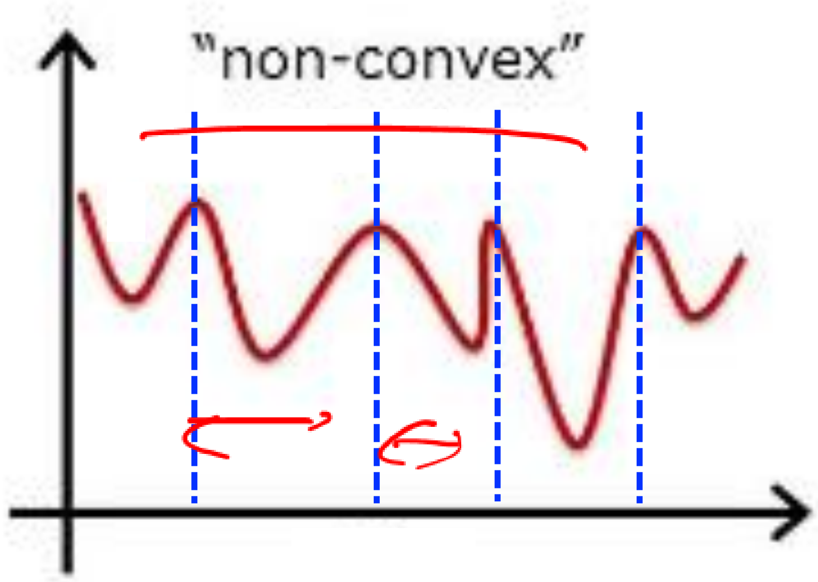
We will give many examples in the class

Why we care convexity?

Convexity serves as the basis to study non-convex optimization



Single "Basin"



Each individual region is convex

We will discuss non-convex optimization in the final lectures.

Outline for Today

Convex functions

- Definition of convex functions
- Show that many commonly seen functions are convex
- Provide a “toolbox” that tells whether a function is convex or not

Linear and Affine Functions

Convex function: function $f\colon \mathbb{R}^n \to \mathbb{R}$ such that $\mathrm{dom}(f) \subseteq \mathbb{R}^n$ is convex, and

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y), \quad \text{for all } t \in [0,1]$$

What about linear functions: $f(x) = a^\top x$, where $a, x \in \mathbb{R}^n$?

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \qquad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$a^\top x = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

Linear and Affine Functions

Convex function: function $f\colon \mathbb{R}^n \to \mathbb{R}$ such that $\mathrm{dom}(f) \subseteq \mathbb{R}^n$ is convex, and

$$f(tx + (1-t)y) \le tf(x) + (1-t)f(y), \text{ for all } t \in [0,1]$$

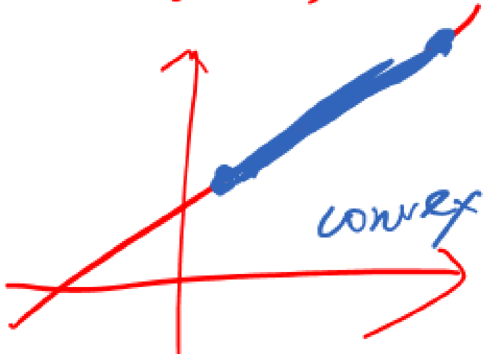
What about affine functions: $f(x) = a^\top x + b$, where $a, x \in \mathbb{R}^n$, $b \in \mathbb{R}$?

Proof:

"Inform" proof:

if $n=1$

$f(x) = ax+b$



$$\text{LHS} = a^T[tx + (1-t)y] + b$$

$$= t \cdot a^T x + (1-t)a^T y + b$$

$$\text{RHS} = t(a^T x + b) + (1-t)(a^T y + b)$$

$$= ta^T x + tb + (1-t)a^T y + (1-t)b$$

$$= ta^T x + (1-t)a^T y + b$$

"="

Quadratic Functions

1-d: $A = a \quad x^T A x = a x^2$

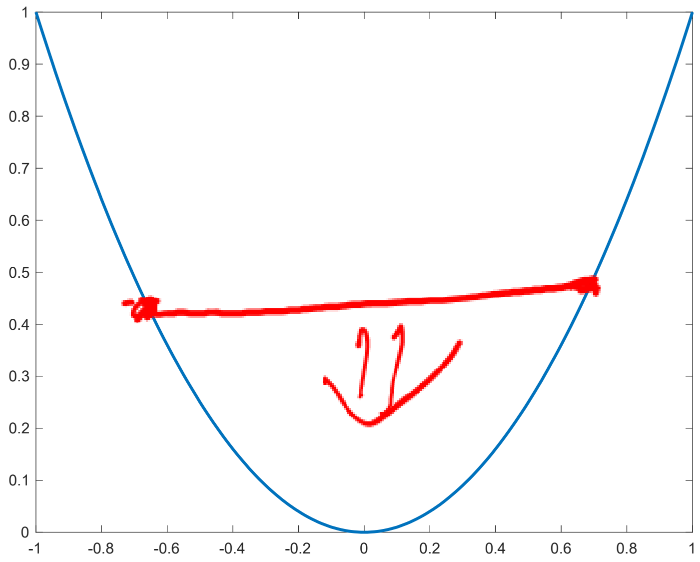
2-d: $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad x^T A x = A_{11} x_1^2 + A_{22} x_2^2 + A_{12} x_1 x_2 + A_{21} x_2 x_1$

$A_{12} = A_{21}$.

What about quadratic functions: $f(x) = x^\top A x$ for symmetric matrix A?

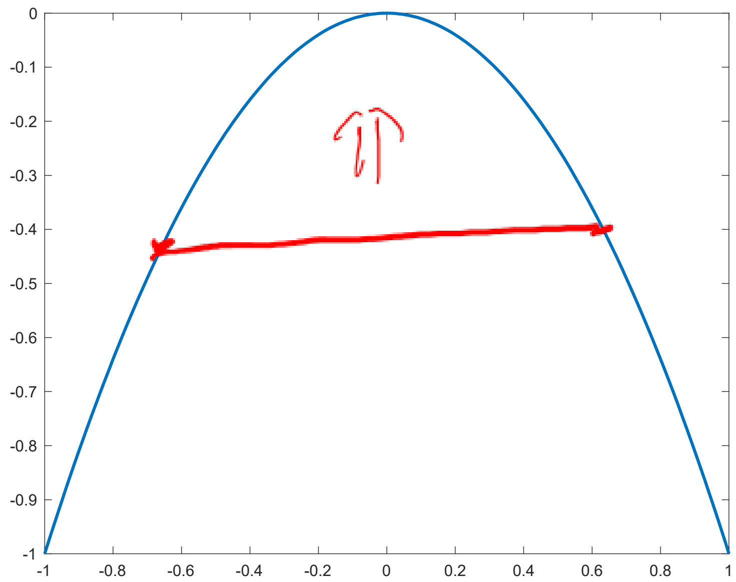
“Proof” by picture for 1-d case: $f(x) = ax^2$

$a = 1$



Convex

$a = -1$



not Convex

Quadratic Functions

Lemma. For $x \in \mathbb{R}^n$ and symmetric matrix $A \in \mathbb{R}^{n \times n}$, function $f(x) = x^\top Ax$ is convex if matrix A is positive semi-definite.

p.s.d.

Recall

A symmetric matrix A is positive semi-definite if either one of the following holds:

- For any $x \in \mathbb{R}^n$, $x^\top Ax \geq 0$
- All the eigenvalues of A is nonnegative

$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$

$(tx+(1-t)y)^T A (tx+(1-t)y)$

$\leq t\, x^T A x + (1-t) y^T A y$

How to prove the lemma?

We will use first/second order condition for convexity.

Tool 1: First Order Condition for Convexity

Lemma. A differentiable function $f(x)$ is convex if and only if for any x, y,

$$f(y) \geq f(x) + \nabla f(x)^{\top}(y - x).$$

How to interpretate this?

Why can this lemma help us show quadratic functions are convex?

How to prove this lemma?

Tool 1: First Order Condition for Convexity

① $g(y)$ is affine

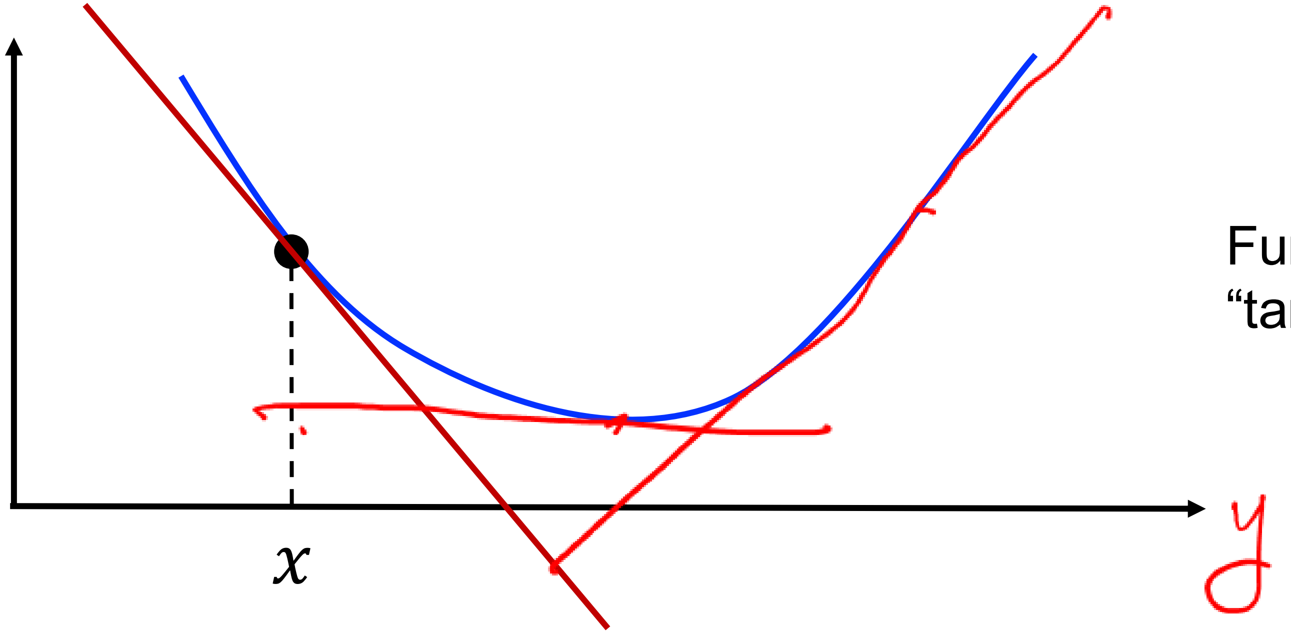
② $g(x) = f(x) + \nabla f(x)^T(x-x)$
$= f(x)$

Lemma. A differentiable function $f(x)$ is convex if and only if for any x, y,

$$f(y) \geq \underbrace{f(x) + \nabla f(x)^{\top}(y-x)}_{g(y)}.$$

③ $g'(y) = f'(x)$

How to interpretate this? Fix x, think of both left hand side and right hand side as function of y



Function value dominates "tangent lines!"

Tool 1: First Order Condition for Convexity

$f(x) = x^T A x$

$\nabla f(x) = \underline{2Ax}$

1-d $\quad f(x) = ax^2$

$\underline{f'(x) = 2ax}$

Lemma. A differentiable function $f(x)$ is convex if and only if for any x, y,

$$\underbrace{f(y)}_{\text{LHS}} \geq \underbrace{f(x) + \nabla f(x)^\top (y - x)}_{\text{RHS}}.$$

Why can this lemma help us show quadratic functions are convex?

Given quadratic function $f(x) = x^\top A x$, we have $\underline{\text{LHS} = f(y) = y^T A y}$

$\text{RHS} = f(x) + \nabla f(x)^T (y - x)$

$= x^T A x + (2Ax)^T (y - x)$

$= x^T A x + 2x^T A (y - x)$

$= x^T A x + 2x^T A y - 2x^T A x$

$= \underline{2x^T A y - x^T A x}$

Why LHS $\geq$ RHS ?

- 1-d : LHS $-$ RHS

$= ay^2 - (2axy - ax^2)$

$= a(y^2 - 2xy + x^2)$

$= a(y - x)^2 \geq 0$

because A psd

- n-d : $\text{LHS} - \text{RHS} = (y - x)^T A (y - x) \geq 0$

Tool 1: First Order Condition for Convexity

Lemma. A differentiable function $f(x)$ is convex if and only if for any x, y,

$$f(y) \geq f(x) + \nabla f(x)^\top (y - x).$$

How to prove this lemma?

Tool 1: First Order Condition for Convexity

Lemma. A differentiable function $f(x)$ is convex if and only if for any x, y,

$$f(y) \geq f(x) + \nabla f(x)^\top (y - x).$$

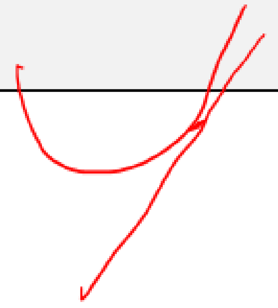
How to prove this lemma?

Summary for First-Order Condition

Lemma. A differentiable function $f(x)$ is convex if and only if for any x, y,

$$f(y) \geq f(x) + \nabla f(x)^{\top}(y - x).$$

How to interpretate this? function is above tangent line



Why can this lemma help us show quadratic functions are convex?

only involves 2 vars. as opposed t. x. y

How to prove this lemma?

TBP

Tool 2: Second Order Condition for Convexity

Lemma. A twice continuously differentiable function $f(x)$ is convex if for any x,

$$\nabla^2 f(x) \succcurlyeq 0.$$

This lemma helps show quadratic functions with PSD A are convex?

Given quadratic function $f(x) = x^\top A x$, we have $\nabla^2 f(x) = 2A$ $\succeq 0$

Tool 2: Second Order Condition for Convexity

Lemma. A twice continuously differentiable function $f(x)$ is convex if for any x,

$$\nabla^2 f(x) \succcurlyeq 0.$$

Proof. Using Taylor's formula, we have for any x, y, there exists $\alpha \in [0,1]$

$$f(y) = f(x) + \nabla f(x)^T (y-x) + \underbrace{\frac{1}{2}(y-x)^T \underbrace{\nabla^2 f(x+\alpha(y-x))}_{\succeq 0 \text{ psd}} (y-x)}_{\geq 0}$$

$$\geq f(x) + \nabla f(x)^T (y-x)$$

by 1st order condition $\Rightarrow$ f is convex

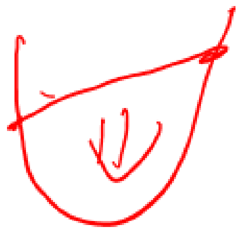
Quadratic Functions

Lemma. For $x \in \mathbb{R}^n$ and symmetric matrix $A \in \mathbb{R}^{n \times n}$, function $f(x) = x^\top A x$ is convex if matrix A is <u>positive semi-definite</u>.

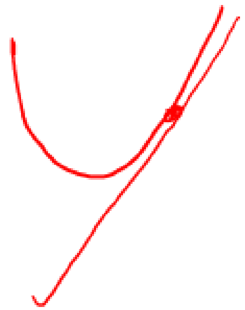
We have just proven it using the first/second order condition for convexity!

Summary: three ways to verify convexity

- By checking the definition: $f(tx+(1-t)y) \le tf(x)+(1-t)f(y)$



- First order condition: $f(y) \ge f(x) + \nabla f(x)^T (y-x)$



- Second order condition: $\nabla^2 f(x)$ psd

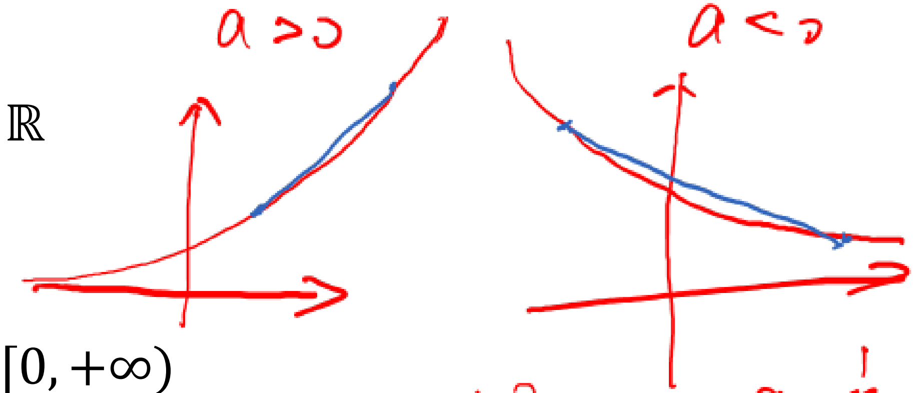
most used.

- Informal way: draw a plot!

Simple Univariate Functions

Exponential function: For any $a \in \mathbb{R}$, $f(x) = e^{ax}$ is convex over $\mathbb{R}$

$f'(x) = a e^{ax}$ $\quad f''(x) = a \cdot a e^{ax} = \underbrace{a^2}_{\geq 0} \underbrace{e^{ax}}_{> 0} \geq 0$

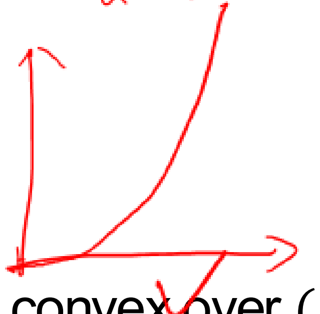


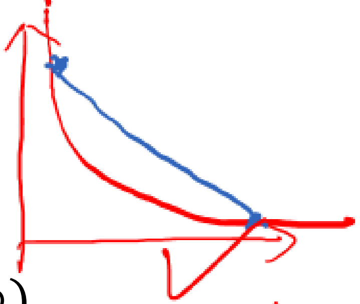
Power function: $f(x) = x^a$ is convex for $a \geq 1$ or $a \leq 0$ over $x \in [0, +\infty)$

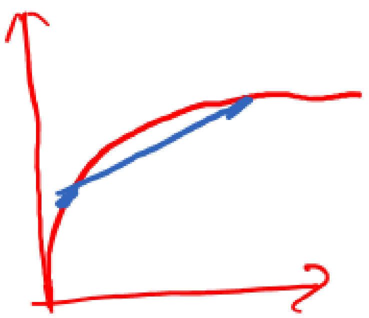
$f'(x) = a x^{a-1}$

$f''(x) = \underbrace{a \cdot (a-1)}_{\geq 0} \underbrace{x^{a-2}}_{\geq 0}$

if $a \geq 1$ or $a \leq 0$



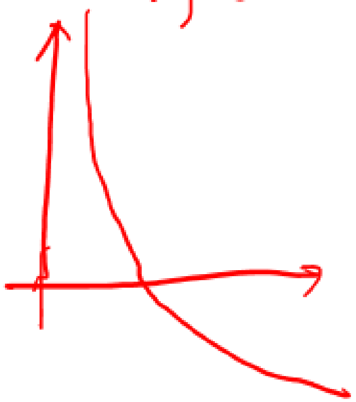




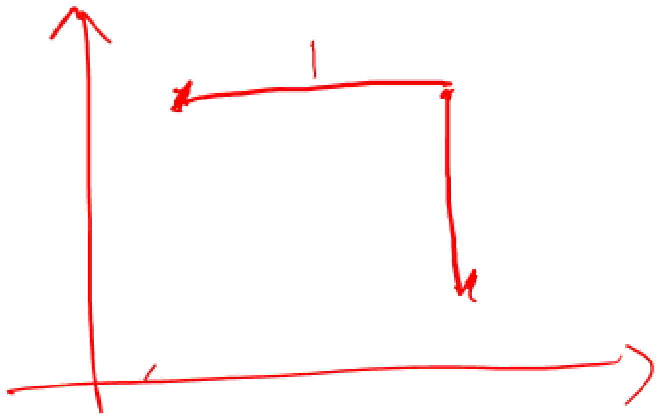
Negative logarithmic function: $f(x) = -\log x$ is convex over $(0, +\infty)$

$f'(x) = -\frac{1}{x}$

$f''(x) = -(-1) \cdot \frac{1}{x^2} = \frac{1}{x^2} > 0$



Norm Functions



Recall: ℓ_p norms for a vector $x \in \mathbb{R}^n$ is defined by

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \quad \text{for } p \geq 1$$

Special cases:

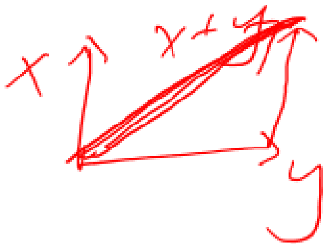
$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$$

Manhattan Distance

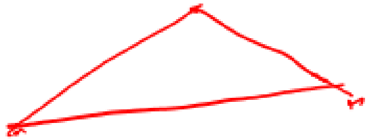
$$\underbrace{\|x\|_1 = \sum_{i=1}^{n} |x_i|}$$

$$\|x\|_\infty = \max_{i=1,\dots,n} |x_i|$$

Properties: Any norm $\|x\|$ satisfies
- For $a \in \mathbb{R}$, $\|ax\| = |a|\|x\|$
- $\|x\| = 0$ if and only if $x = 0$
- Triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$



Norm Functions



Lemma. All norm functions are convex!

Proof: use triangle inequality

$$\|tx + (1-t)y\| \le t\|x\| + (1-t)\cdot\|y\|$$

$\Rightarrow$ $f(x) = \|x\|$ is convex

Summary so far

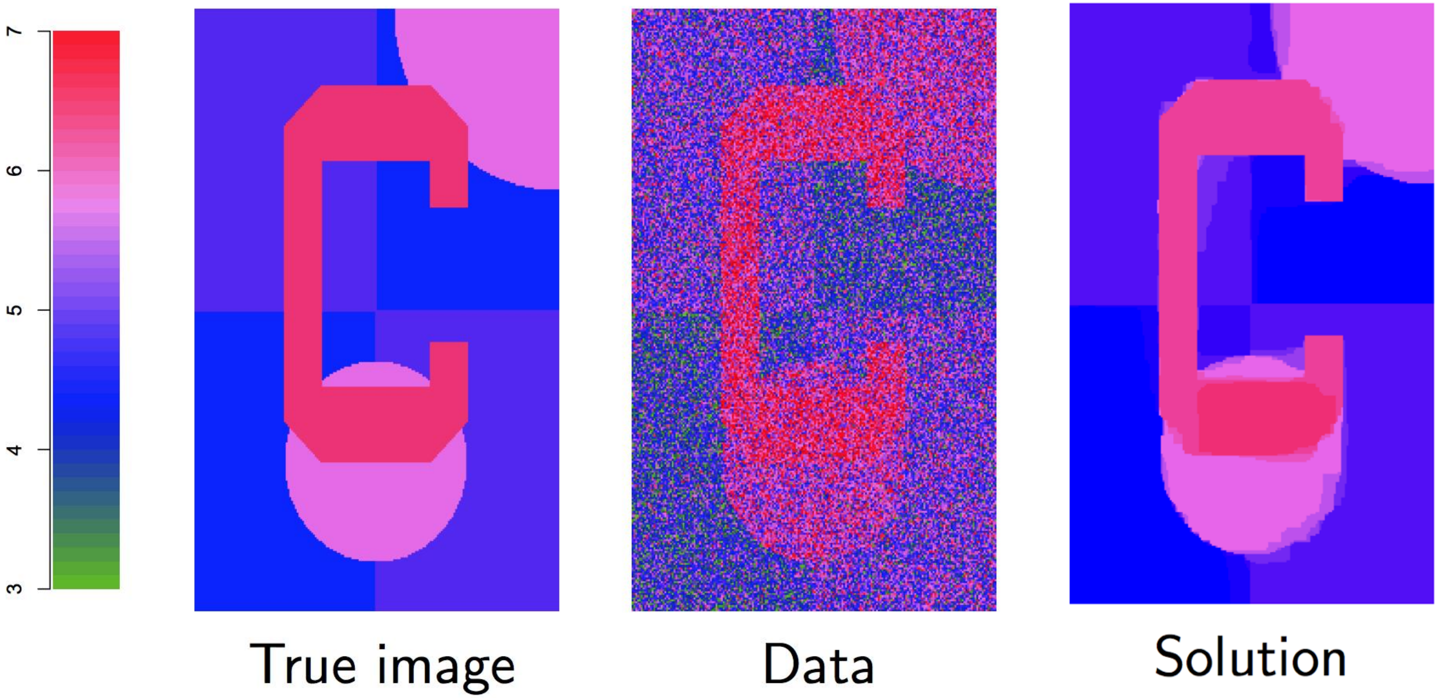
We have studied the convexity of

- Linear/affine functions
- Quadratic functions
- Simple univariate functions
- Norm functions

What about more sophisticated functions?

- e.g. $f(x) = x^2 + e^x$?
- e.g. $f(x) = \frac{1}{2} x^\top A x + b^\top x + c$?
- e.g. the denoising example last lecture

Revisit: Denoising



$$\min_{\theta_1,\ldots,\theta_n} \underbrace{\frac{1}{2}\sum_{i=1}^{n}(y_i-\theta_i)^2}_{\theta_i \text{ stays close to } y_i} + \lambda \underbrace{\sum_{(i,j)\ \text{adjacent}} |\theta_i-\theta_j|}_{\text{penalize changes in adjacent pixels}}$$

Revisit: Denoising

Is this function convex?

$$\min_{\theta_1,\ldots,\theta_n} \frac{1}{2}\sum_{i=1}^{n}(y_i - \theta_i)^2 + \lambda \sum_{(i,j)\ \text{adjacent}} |\theta_i - \theta_j|$$

Summary so far

We have studied the convexity of

- Linear/affine functions
- Quadratic functions
- Simple univariate functions
- Norm functions

What about more sophisticated functions?

- e.g. $f(x) = x^2 + e^x$?
- e.g. $f(x) = \frac{1}{2} x^\top A x + b^\top x + c$
- e.g. the denoising example last lecture
- Next: operations that preserve convexity!

Operations Preserving Convexity

Nonnegative linear combination
$f_1, \ldots, f_m$ convex implies $a_1 f_1 + \ldots + a_m f_m$ convex for any $a_1, \ldots a_m \geq 0$.

Proof: Checking Hessian

Consequence: $f(x) = x^2 + e^x$ is convex!

$f(x) = \frac{1}{2} x^\top A x + b^\top x + c$ is convex if A positive semi-definite!

Summary for Today

Typical functions and their convexity

- Affine, quadratic
- Simple univariate functions
- Norm functions

Ways to check convexity

- Definition
- First order and second order condition
- Operations that preserve convexity

Next lecture:

- More on operations that preserve convexity
- Strict and strong convexity, concavity
- Convex sets

Reminder: Quiz for today’s lecture due Jan 20 at 11AM