Convex Function Cont’ and Convex Sets

Lecture 4 for 18660/18460: Optimization

Guannan Qu

January 22, 2026

Admin Stuff

- HW 1 has been released and due Feb 5
- Quiz for Lecture 4 out, due Jan 27 before lecture
- Reminder about Recitation Session
  - Friday, January 23, 10:00 AM – 11:00 AM ET
  - Topic: linear algebra + calculus (helpful for HW1)
  - Zoom only (with recording)

Recall: Operations Preserving Convexity

- Nonnegative linear combination

$f_1 \; f_2 \cdots f_m$ convex, $a_1 \; a_2 \cdots a_m \geq 0$ $\Big\} \Rightarrow a_1 f_1 + a_2 f_2 + \cdots + a_m f_m$ convex

- Compositions

  - $h(Ax+b)$ is convex if h is convex
  - $h(g(x))$ is convex if h is nondecreasing & convex and g convex

- Pointwise maximization **Let's cover now**

Operations Preserving Convexity

dom(f) = dom(f₁) ∩ dom(f₂) . . . ∩ dom(fₘ)

Pointwise maximization:

If $f_1, f_2, \ldots, f_m$ is convex, then $f(x) = \max(f_1(x), \ldots, f_m(x))$ is convex.

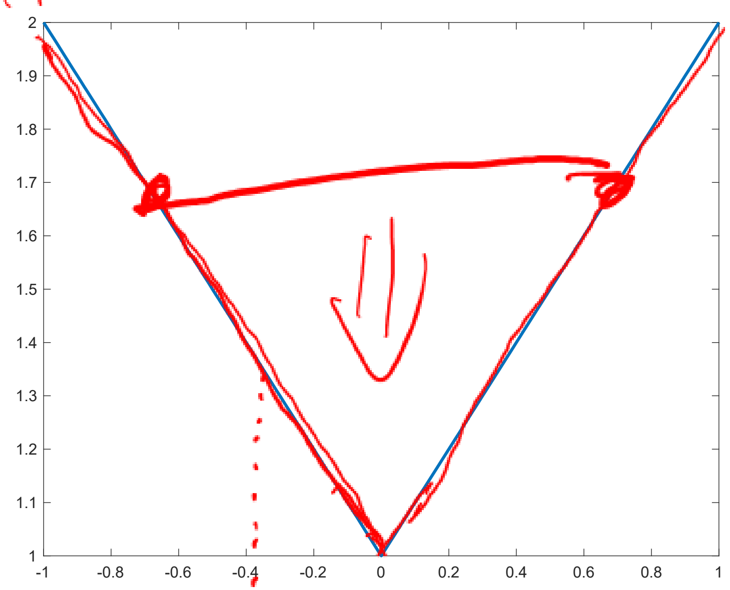
dom(f₁) dom(f₂)

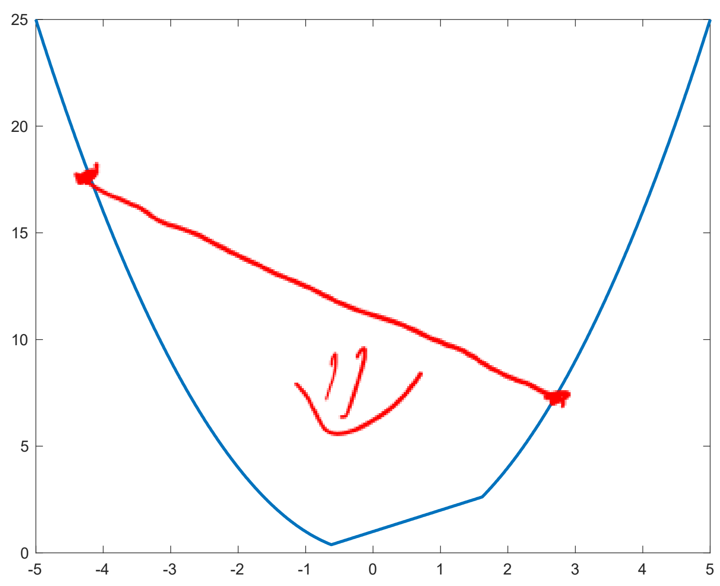
$f(x) = \max(x+1, -x+1)$

$f(x) = \max(x+1, x^2)$

y = −x + 1

y = x + 1





Operations Preserving Convexity

Pointwise maximization:
If $f_1, f_2, \dots, f_m$ is convex, then $f(x) = \max(f_1(x), \dots, f_m(x))$ is convex.

Proof: we use definition. Given $t \in [0,1]$ and x, y,

want to show $f(tx + (1-t)y) \le tf(x) + (1-t)f(y)$

Let i_x to be the index s.t. $f(x) = f_{i_x}(x)$

Let i_y ...... $f(y) = f_{i_y}(y)$

Let i ...... $f(tx + (1-t)y) = f_i(tx + (1-t)y)$

$f(tx + (1-t)y) = f_i(tx + (1-t)y) \le tf_i(x) + (1-t)f_i(y)$

$\le t\,\underline{f_{i_x}(x)} + (1-t)f_{i_y}(y)$

$= tf(x) + (1-t)f(y) \Rightarrow f$ is convex

$\left(f_{i_x}(x) = \max_j f_j(x) \ge f_i(x) \right)$

Operations Preserving Convexity

Pointwise maximization:

If f_i is convex for all $i \in I$ where I may be an infinite set, then

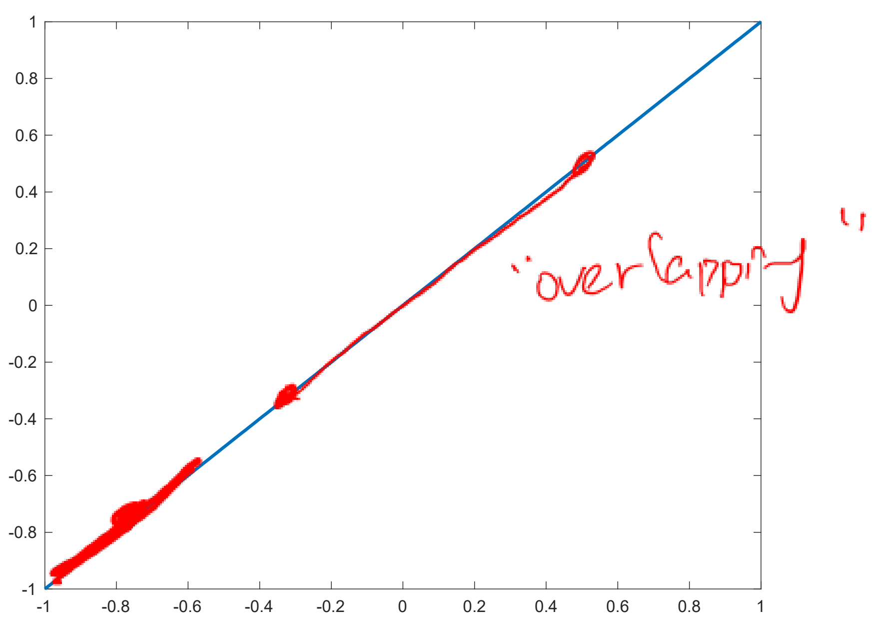
$$f(x) = \max_{i \in I} f_i(x)$$

is convex.

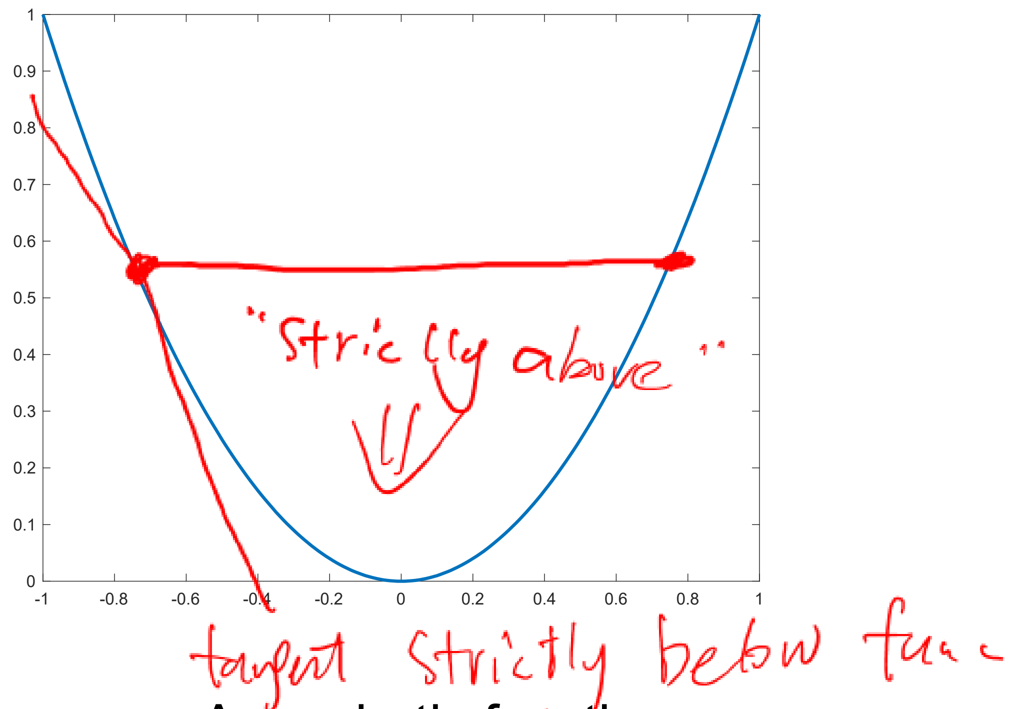
Proof: identical

Strict Convexity and Strong Convexity

Both these functions are convex, but they look different



An affine function



A quadratic function

Strict Convexity

strict

Definition. A function $f(x)$ is strictly convex if for any $x \neq y$, $t \in (0,1)$

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y).$$

Equivalent definition using first order condition:

strict

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

Equivalent definition using second order condition:

$$\nabla^2 f(y) \succ 0.$$

Strong Convexity

$\to L_2$ norm

Definition. For $\mu > 0$, a function $f(x)$ is μ-strongly convex if $f(x) - \frac{\mu}{2}\|x\|^2$ is convex

Equivalent definition using first order condition: for any x, y

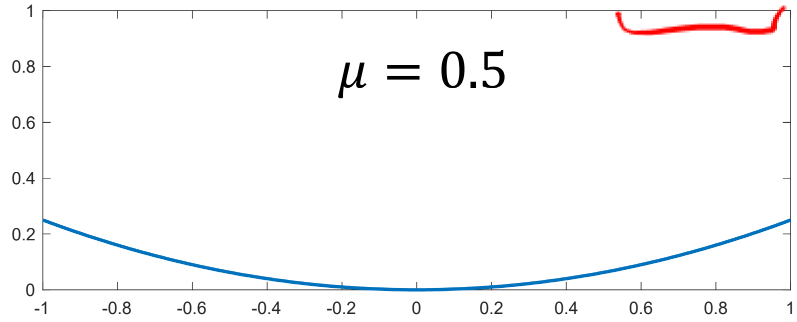
$$f(y) \geq f(x) + \nabla f(x)^\top (y - x) + \frac{\mu}{2}\|y - x\|^2.$$

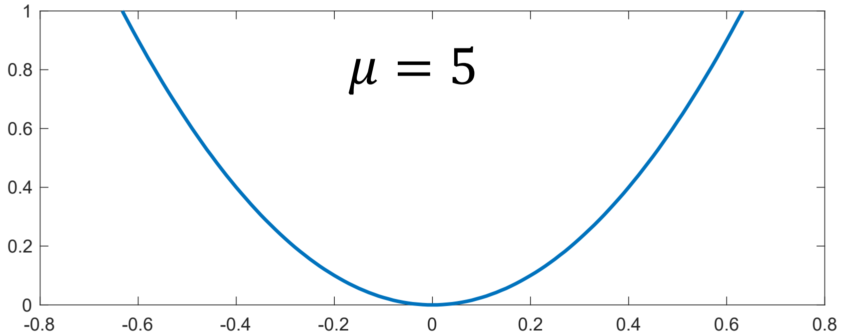
Equivalent definition using second order condition: for any x

$$\nabla^2 f(x) \succcurlyeq \mu I$$

How to understand strong convexity?

- Use quadratic function $\frac{\mu}{2}\|x\|^2 = \frac{\mu}{2} x^\top x$ as a reference to tell how "curved" functions are





Strong Convexity

Definition. For $\mu > 0$, a function $f(x)$ is μ-strongly convex if $f(x) - \frac{\mu}{2}\|x\|^2$ is convex

Equivalent definition using first order condition: for any x, y

$$f(y) \geq f(x) + \nabla f(x)^\top (y - x) + \frac{\mu}{2}\|y - x\|^2.$$

Equivalent definition using second order condition: for any x

$$\nabla^2 f(x) \succcurlyeq \mu I$$

How to understand strong convexity?

- Use quadratic function $\frac{\mu}{2}\|x\|_2^2 = \frac{\mu}{2} x^\top x$ as a reference to tell how "curved" functions are
- Why choosing $\frac{\mu}{2}\|x\|^2$ as the reference? - its Hessian is μI, a simple matrix

More on Strong Convexity

Definition. For $\mu > 0$, a function $f(x)$ is μ-strongly convex if $f(x) - \frac{\mu}{2}\|x\|^2$ is convex

Equivalent definition using first order condition: for any x, y

$$f(y) \geq f(x) + \nabla f(x)^\top (y - x) + \frac{\mu}{2}\|y - x\|^2.$$

Equivalent definition using second order condition: for any x

$$\nabla^2 f(x) \succcurlyeq \mu I$$

Strong convexity of general quadratic functions $f(x) = \frac{1}{2}x^\top Ax$

$$\nabla^2 f(x) = A \succcurlyeq \sigma_{min}(A) I$$

smallest eigvalue of A

When A is positive definite (i.e. $\sigma_{min}(A) > 0$), $f(x) = \frac{1}{2}x^\top Ax$ is $\sigma_{min}(A)$ -strongly convex

Comparison btw Strict/Strong convexity

strict convexity	μ-strong convexity
$\nabla^2 f(x) \succ 0, \forall x \in dom(f)$	$\nabla^2 f(x) \succcurlyeq \mu I, \forall x \in dom(f)$

Example: $f(x) = \frac{1}{x}$ over $(0, +\infty)$

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

$f''(x) = \frac{2}{x^3} > 0$ for $x \in (0, +\infty)$

So f is strictly convex

But f is NOT μ-strongly convex $\forall \mu$

Pick μ, $f''(x) = \frac{2}{x^3} < \mu$ for sufficiently large x

Example: $f(x) = \frac{1}{2} a x^2$ over $\mathbb{R}$, $a > 0$

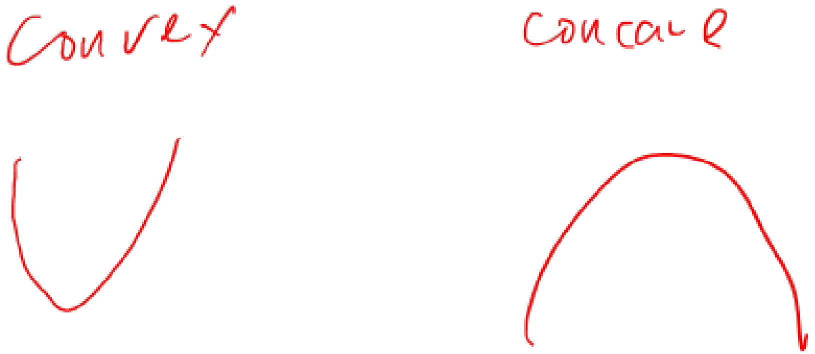
$f''(x) = a \geq \mu > 0$

take $\mu = a$

$\Rightarrow f$ is a-strongly convex

Concavity

A function f is concave if $-f$ is convex
A function f is strictly concave if $-f$ is strictly convex
A function f is μ-strongly concave if $-f$ is μ-strongly convex



	Definition	First-order condition	Second-order condition
concave	$f(tx+(1-t)y) \geq tf(x)+(1-t)f(y)$, $\forall x,y,\forall t \in [0,1]$	$f(y) \leq f(x)+\nabla f(x)^\top (y-x), \forall x,y$	$\nabla^2 f(x) \preccurlyeq 0$, $\forall x$
convex	$f(tx+(1-t)y) \leq tf(x)+(1-t)f(y)$, $\forall x,y,\forall t \in [0,1]$	$f(y) \geq f(x)+\nabla f(x)^\top (y-x), \forall x,y$	$\nabla^2 f(x) \succcurlyeq 0$, $\forall x$
strictly concave	$f(tx+(1-t)y) > tf(x)+(1-t)f(y)$, $\forall x \neq y,\forall t \in (0,1)$	$f(y) < f(x)+\nabla f(x)^\top (y-x), \forall x \neq y$	$\nabla^2 f(x) \prec 0$, $\forall x$
strictly convex	$f(tx+(1-t)y) < tf(x)+(1-t)f(y)$, $\forall x \neq y,\forall t \in (0,1)$	$f(y) > f(x)+\nabla f(x)^\top (y-x), \forall x \neq y$	$\nabla^2 f(x) \succ 0$, $\forall x$
μ-strongly concave	$f(x)+\frac{\mu}{2}\|x\|^2$ is concave	$f(y) \leq f(x)+\nabla f(x)^\top (y-x) - \frac{\mu}{2}\|y-x\|^2$, $\forall x,y$	$\nabla^2 f(x) \preccurlyeq -\mu I$, $\forall x$
μ-strongly convex	$f(x)-\frac{\mu}{2}\|x\|^2$ is convex	$f(y) \geq f(x)+\nabla f(x)^\top (y-x) + \frac{\mu}{2}\|y-x\|^2$, $\forall x,y$	$\nabla^2 f(x) \succcurlyeq \mu I$, $\forall x$

Quiz Results

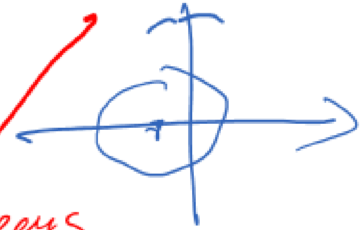
Suppose $f(x)$ is a strictly convex function. Select all that is true:

(A) $dom(f)$ is a convex set	100 %
(B) $-f(x)$ is strictly concave	100 %
(C) $f(x)$ is strongly convex	18 %
(D) $f(x)$ is an affine function	0 %

Quiz Results

(D) $M = \begin{bmatrix} \ddots & & \\ & M_{ii} & \\ & & \ddots \end{bmatrix}$ $M_{ii} < 0$ $x^T M x = M_{ii} < 0$

$x = [0 \dots 0, 1, 0 \dots 0]$ ↑ ith entry   M is not PSD



M is PSD iff $x^T M x \geq 0$ $\forall x$

Which of the following conditions imply that a symmetric matrix M is NOT positive semi-definite? Select all that is true.

means at least one eigvalue < 0

(A) M's minimum eigenvalue is -1 < 0	100 %
(B) M has rank 1	0 %
(C) There exists i, j $(i \neq j)$ such that $M_{i,j} < 0$	0 %
(D) There exists i such that $M_{i,i} < 0$	82 %

(B) rank 1 means M has one non-zero eigvalue, all of the eigvalues $= 0$

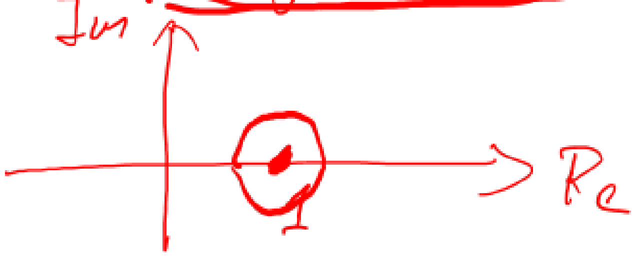
$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ rank 1 psd

(C) $\begin{bmatrix} 1 & -0.1 \\ -0.1 & 1 \end{bmatrix}$ psd

$\begin{bmatrix} 1 & -0.99 \\ -0.99 & 1 \end{bmatrix}$

"Gershgorin Disk"

$\text{Radius}(i) = \sum_{j \neq i} |a_{ij}|$



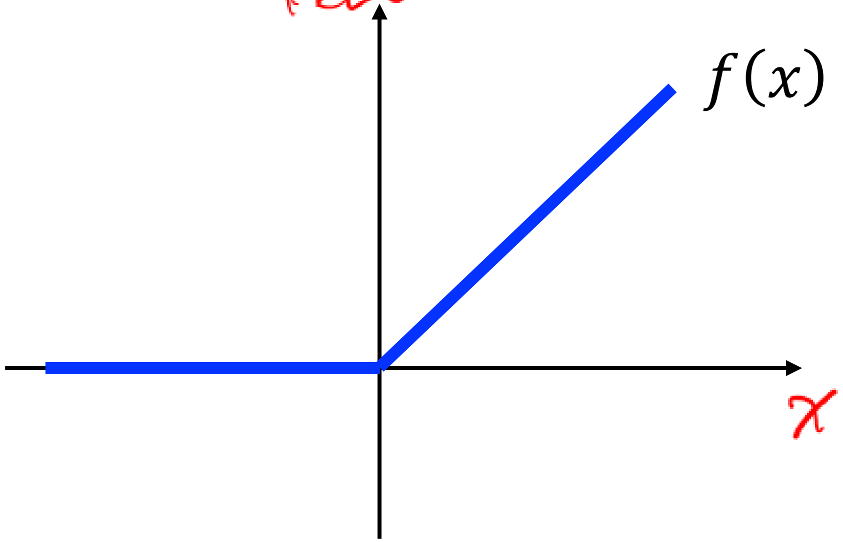
Smoothness

Definition. A function $f(x)$ is L-smooth if it is differentiable and its gradient is L-Lipschitz continuous

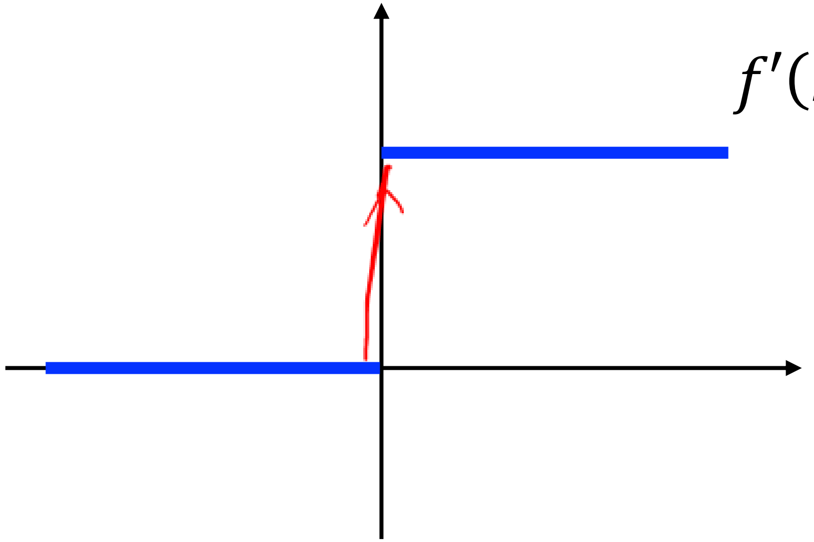
$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$$

- Characterizes rate of change of gradients
- Example of a non-smooth function

ReLU



$f(x) = \max(x, 0)$



$f'(x)$

Smoothness

f_1 L_1 Smooth
f_2 L_2 smooth

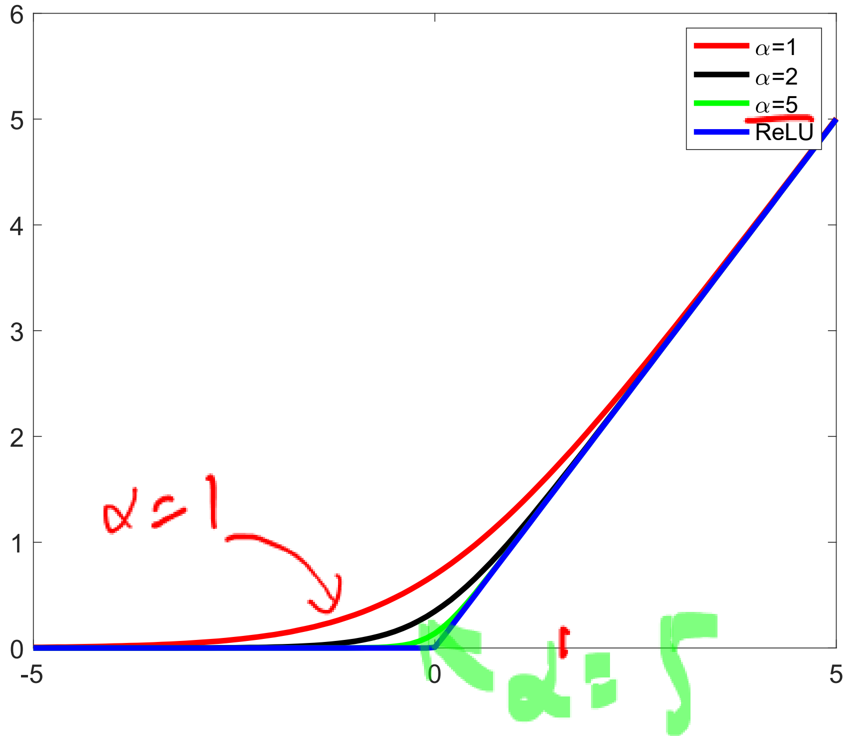
$a_1 f_1 + a_2 f_2$

$|a_1| L_1 + |a_2| L_2$ smooth

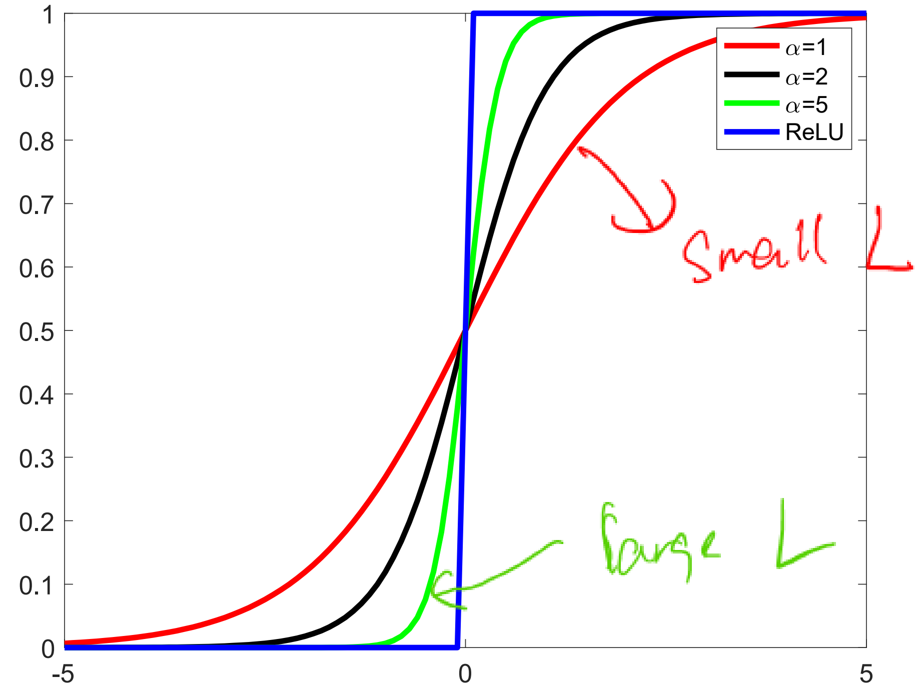
Definition. A function $f(x)$ is L-smooth if it is differentiable and its gradient is L-Lipschitz continuous

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$$

$$f(x) = \frac{1}{\alpha}\log(1 + e^{\alpha x})$$



$$f'(x)$$



Smoothness

Definition. A function $f(x)$ is L-smooth if it is differentiable and its gradient is L-Lipschitz continuous

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$$

Why we care?

- Non-smooth functions are usually "harder" to optimize than smooth functions
- Smooth functions with larger L are "harder" to optimize

Smoothness

$(0, +\infty)$

dom(f)

Definition. A function $f(x)$ is L-smooth if it is differentiable and its gradient is L-Lipschitz continuous

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$$

First order condition: for any x, y

$$f(y) \leq f(x) + \nabla f(x)^{\top}(y - x) + \frac{L}{2}\|y - x\|^2.$$

Second order condition: for any x

$$\nabla^2 f(x) \preccurlyeq LI \Leftarrow \text{all eigenvalue of } \nabla^2 f \leq L$$

Strong Convexity vs Smoothness

μ-strong convexity	L-smooth
$\nabla^2 f(x) \succcurlyeq \mu I, \forall x \in dom(f)$	$\nabla^2 f(x) \preccurlyeq LI, \forall x \in dom(f)$

$\nabla^2 f(x) = A$

smallest eig of A

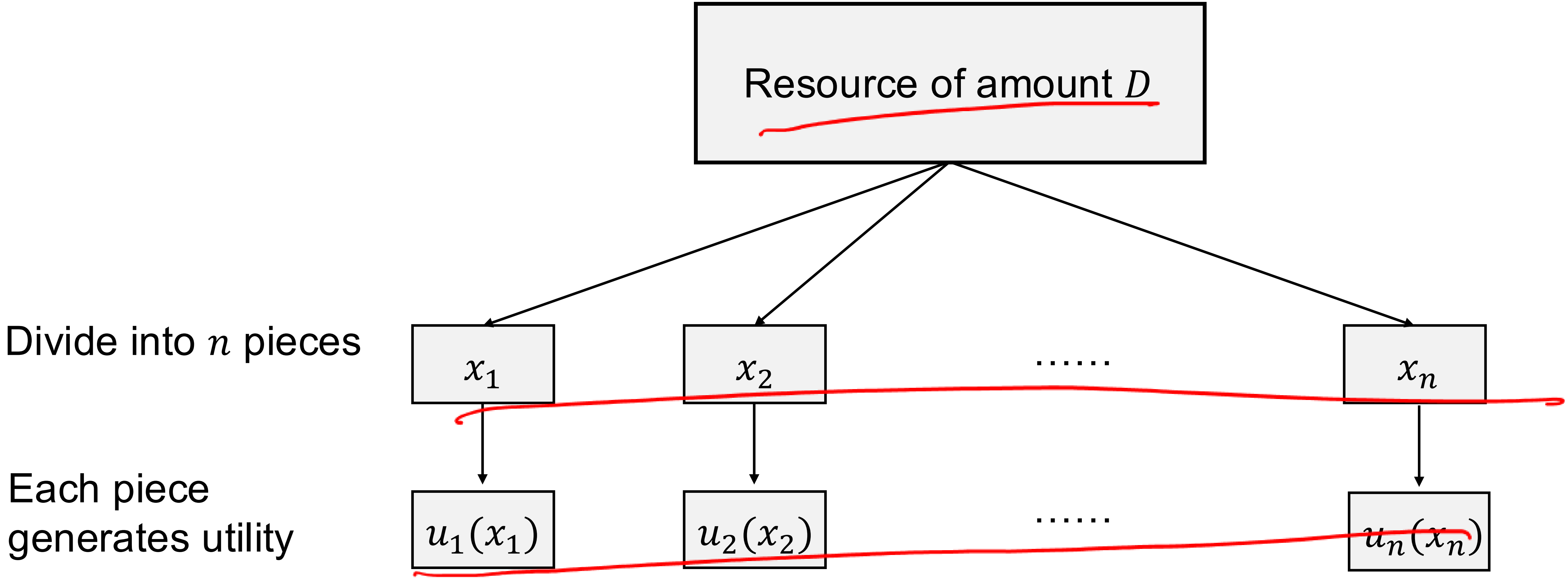
- Quadratic functions
  - We have shown that $f(x) = \frac{1}{2} x^\top A x$ is $\sigma_{min}(A)$-strongly convex when A positive definite
  - Since $\nabla^2 f(x) = A \preccurlyeq \sigma_{max}(A) I$, the function is also $\sigma_{max}(A)$-smooth

largest eig of A

- For a μ-strong convexity and L-smooth function, its **condition number** is defined as $\frac{L}{\mu}$
- Functions with larger condition number is harder to optimize using gradient descent
  - Will cover in detail in week 5

Convex Constraint Sets

Revisit: Resource Allocation



Goal: find an allocation $x_1, \dots, x_n$ to maximize total utility $u_1(x_1) + u_2(x_2) + \cdots + u_n(x_n)$

Revisit: Resource Allocation

$$\max_{x_1,\dots,x_n} g(x_1, \dots, x_n) = u_1(x_1) + u_2(x_2) + \cdots + u_n(x_n)$$

s.t.

$$\boxed{\begin{array}{l} x_i \geq 0, \forall i = 1,2, \dots, n \\ x_1 + x_2 + \cdots + x_n \leq D \end{array}}$$

Is this constraint set convex?

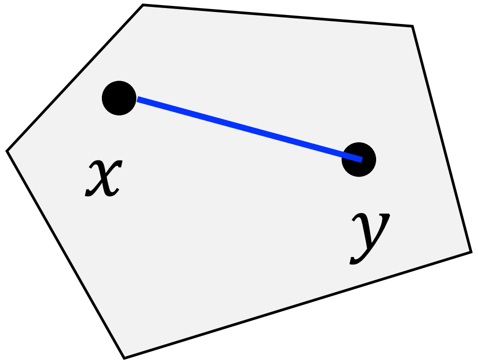
What you will learn…

- Basic convex sets
- A toolbox that can tell whether a set is convex or not

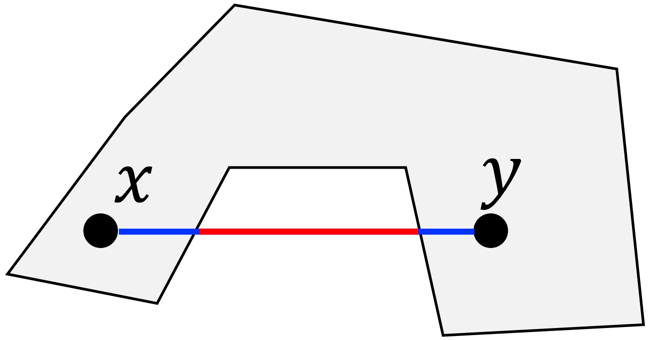
Recall

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

Inequality constraint sets

$C = \{x \mid f(x) \leq 0, x \in dom(f)\}$

Lemma. For any convex function f, the following level set,

$$C = \{x: f(x) \leq 0\}$$

is convex.

Proof: Want to prove C is convex

Pick arbitrary tow points $x, y \in C$

try to prove $tx + (1-t)y \in C$

Since $x, y \in C \Rightarrow$ $f(x) \leq 0$, $f(y) \leq 0$

$$\underbrace{f(tx+(1-t)y) \leq}_{\text{by convexity}} t \cdot \underbrace{f(x)}_{\leq 0} + (1-t)\underbrace{f(y)}_{\leq 0} \leq 0 \Rightarrow tx + (1-t)y \in C$$

\#

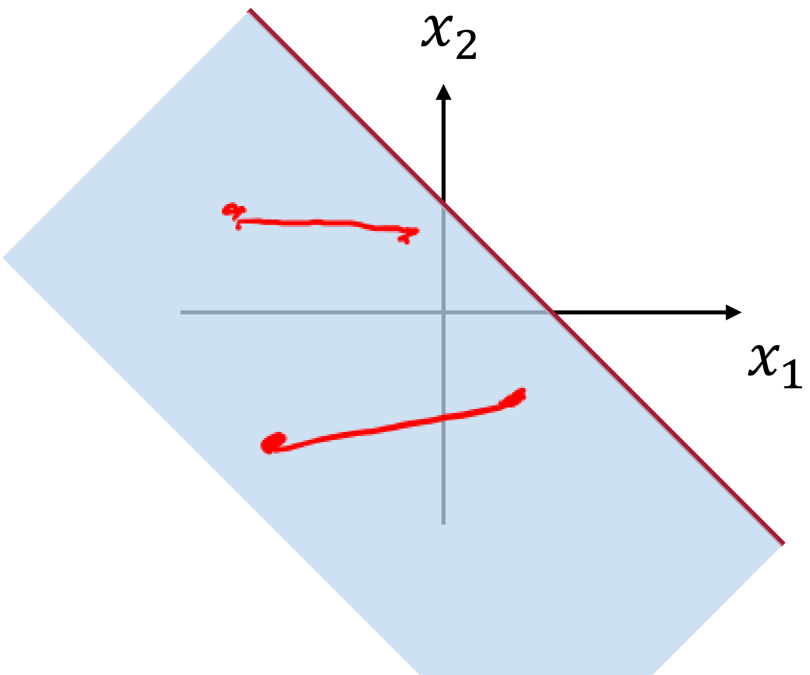
Inequality constraint sets

Lemma. For any convex function f, the following level set,

$$\{x: f(x) \leq 0\}$$

is convex.

Example: $\{x \in \mathbb{R}^2: x_1 + x_2 + 1 \leq 0\}$



Inequality constraint sets

Lemma. For any convex function f, the following level set,

$$\{x: f(x) \leq 0\}$$

is convex.

Halfspaces: $\{x \in \mathbb{R}^n: a^\top x + b \leq 0\}$

Inequality constraint sets

Lemma. For any convex function f, the following level set,

$$\{x: f(x) \leq 0\}$$

is convex.

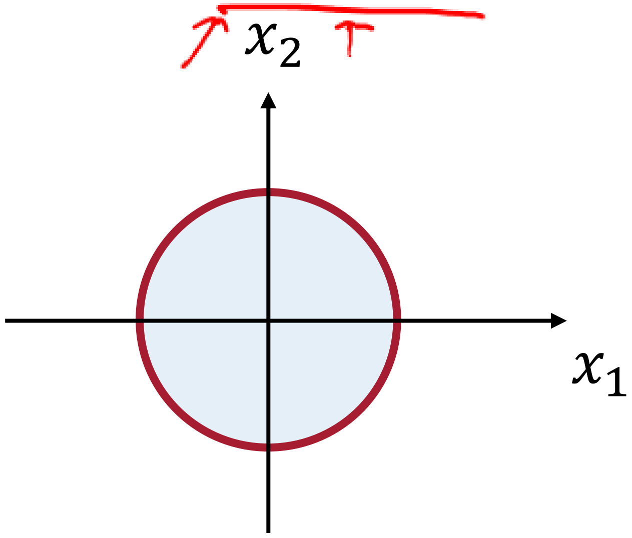
$\{x: x^\top Ax - 1 \leq 0\}$ for A positive semi-definite?

Inequality constraint sets

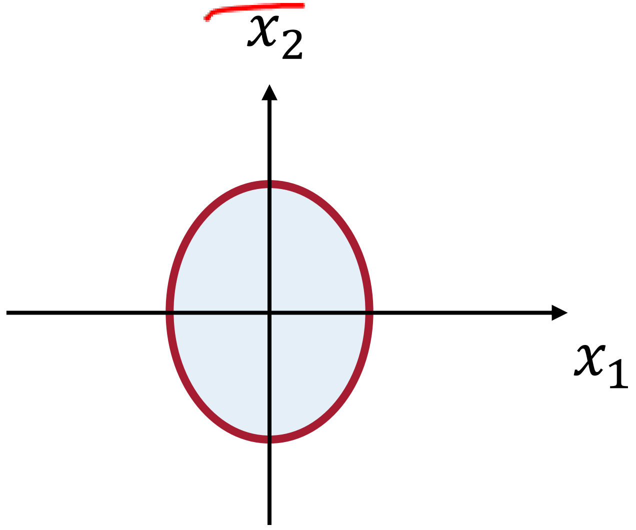
> **Lemma.** For any convex function f, the following level set,
> $$\{x: f(x) \leq 0\}$$
> is convex.

Example:

$\{x \in \mathbb{R}^2: x_1^2 + x_2^2 - 1 \leq 0\}$



$\{x \in \mathbb{R}^2: 2x_1^2 + x_2^2 - 1 \leq 0\}$



Inequality constraint sets

Lemma. For any convex function f, the following level set,

$$\{x: f(x) \leq 0\}$$

is convex.

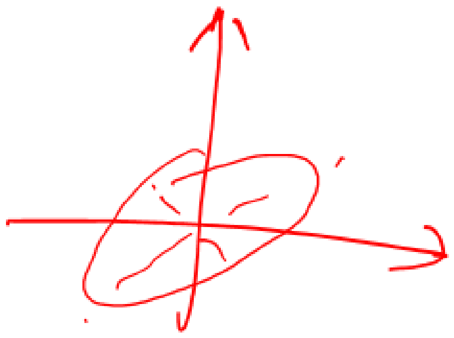
$(x-c)^T A (x-c)$

For positive definite A, $\{x: x^\top Ax - 1 \leq 0\}$ is an ellipsoid

$x^T A x \leq 1$

$y = Rx$
rotation

$y^T \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} y \leq 1 \quad = \lambda_1 y_1^2 + \lambda_2 y_2^2$



Inequality constraint sets

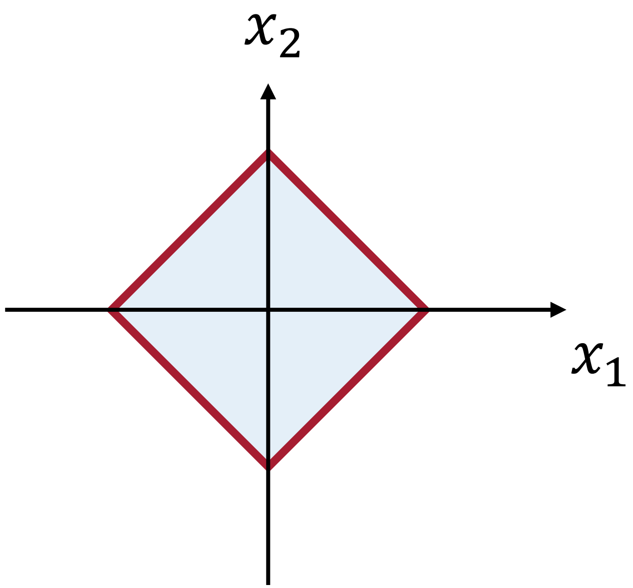
Lemma. For any convex function f, the following level set,

$$\{x: f(x) \leq 0\}$$

is convex.

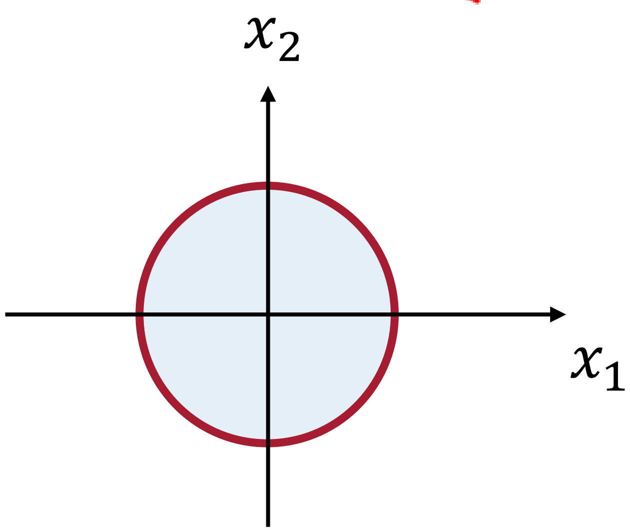
ℓ_1-norm ball

$\{x \in \mathbb{R}^2 : \|x\|_1 - 1 \leq 0\}$



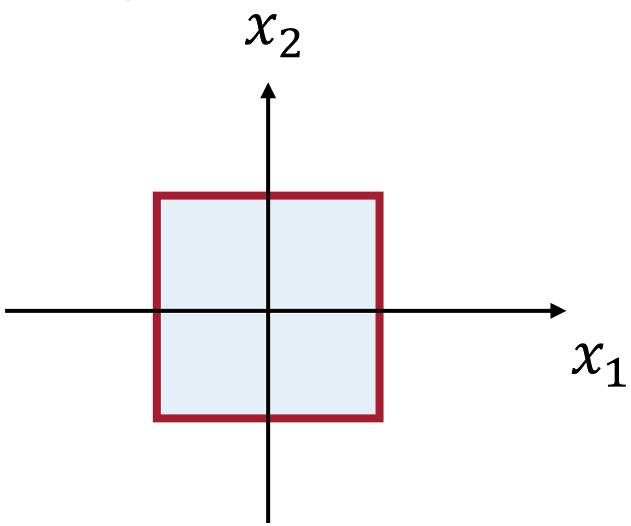
ℓ_2-norm ball

$\{x \in \mathbb{R}^2 : \|x\|_2 - 1 \leq 0\}$



ℓ_∞-norm ball

$\{x \in \mathbb{R}^2 : \|x\|_\infty - 1 \leq 0\}$

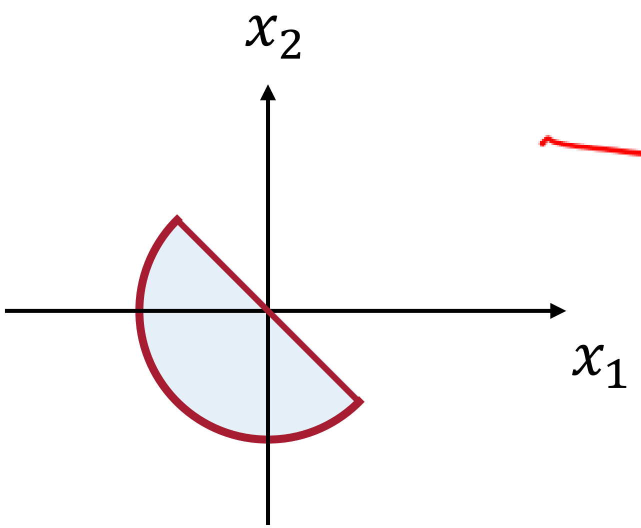


Intersection of multiple constraint sets

Lemma. For any two convex sets C_1, C_2, their intersection $C_1 \cap C_2$ is convex.
Specifically, for two convex functions f_1, f_2, the following set is convex:

$$\underbrace{\{x: f_1(x) \le 0\}}_{C_1} \cap \underbrace{\{x:\ f_2(x) \le 0\}}_{C_2} = \underbrace{\{x: f_1(x) \le 0, f_2(x) \le 0\}}_{C_1 \cap C_2}$$

Examples: ball intersected with half plane



Intersection of multiple constraint sets

Lemma. For any two convex sets C_1, C_2, their intersection $C_1 \cap C_2$ is convex. Specifically, for two convex functions f_1, f_2, the following set is convex:

$$\{x: \underbrace{f_1(x) \leq 0}_{C_1}\} \cap \{x: \underbrace{f_2(x) \leq 0}_{C_2}\} = \{\underbrace{x: f_1(x) \leq 0, f_2(x) \leq 0}_{C_1 \cap C_2}\}$$

Proof:

$$\underline{C_1 \cap C_2} = \{x: \underbrace{\max(\overset{\text{convex}}{f_1(x)}, \overset{\text{convex}}{f_2(x)})}_{\underline{f(x)} \text{ is convex}} \leq 0\}$$

Can generalize to intersection of more than two convex constraint sets

Summary

- Strict and Strong convexity, smoothness, condition number
- Convex constraint sets
  - Halfspaces
  - Norm balls
  - Ellipsoids

Next lecture:
More on convex set
convex optimization problems!