Convex Sets cont' and Convex Optimization Problem

Lecture 5 for 18660/18460: Optimization

Guannan Qu

Jan 27, 2026

Admin Stuff

- Quiz for lecture 5 out, due on Jan 29 before lecture
- A 10-min video tutorial on CVX (along with python and matlab examples) has been uploaded to canvas

Recall

$\frac{\mu}{2}\|x\|_2^2 = \frac{\mu}{2}(x_1^2 + x_2^2 + \cdots + x_n^2) = \frac{\mu}{2} x^T x$

- Strong convexity For $\mu > 0$, μ-strongly convex $\Rightarrow$

  $\rightarrow$ $f(x) - \frac{\mu}{2}\|x\|^2$ is convex

  $\rightarrow$ $f(y) \geq f(x) + [\nabla f(x)]^T (y-x) + \frac{\mu}{2}\|y-x\|^2$

  $\rightarrow$ $\nabla^2 f(x) \succeq \mu I$ [All eig value of $\nabla^2 f(x) \geq \mu$]

- Smoothness For $L > 0$, L-Smoothness $\Rightarrow$

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

  $\rightarrow$ $f(y) \leq f(x) + [\nabla f(x)]^T (y-x) + \frac{L}{2}\|y-x\|^2$

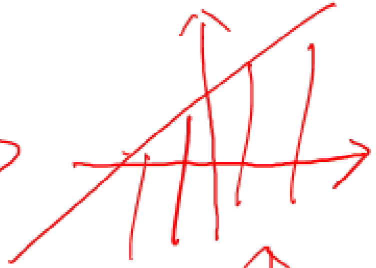
  $\rightarrow$ $\nabla^2 f(x) \preceq L I$

$\boxed{\frac{L}{\mu} \text{ condition number}}$

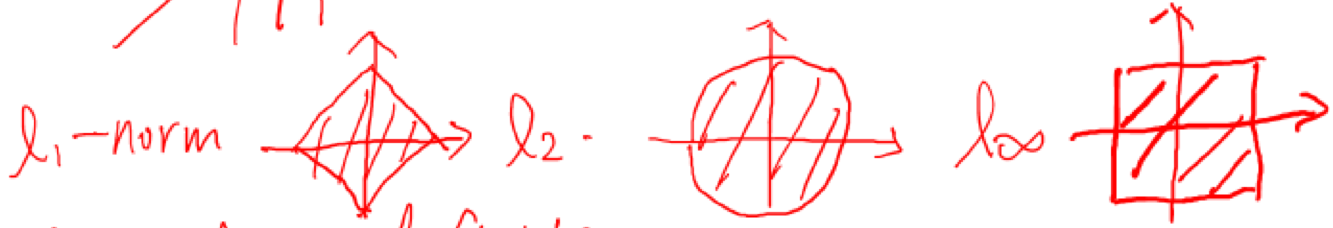
Recall

$\{x : f(x) \le 0\}$ is convex if f convex

- Convex constraint sets
  - ~~Halfspaces~~ $\{x : a^T x + b \le 0\}$



  - Norm balls $\{x : \|x\| \le 1\}$



  - Ellipsoids $\{x : x^T A x \le 1\}$ A positive definite
  - ~~Hyperplanes~~ $\cancel{\{x : a^T x + b = 0\}}$

- The intersection of two convex sets are still convex!

≥ 2

Equality Constraint

$\begin{cases} f(x) \le 0 \\ f(x) \ge 0 \end{cases}$

Given function f, we can also form equality constraint set

$$\{x : f(x) = 0\} \;\; = \{x : f(x) \le 0\} \cap \{x : -f(x) \le 0\}$$

Equality Constraint

Lemma. For any affine function $f(x) = a^\top x + b$, the following equality constraint set

$$\{x: f(x) = 0\} = \{x: a^\top x + b = 0\}$$

is convex.

Proof: both $\{x: \overbrace{a^\top x + b}^{\text{affine, convex}} \le 0\}$ and $\{x: \overbrace{-a^\top x - b}^{\text{affine, convex}} \le 0\}$ are convex.

Equality Constraint

Lemma. For any affine function $f(x) = a^{\top}x + b$, the following equality constraint set

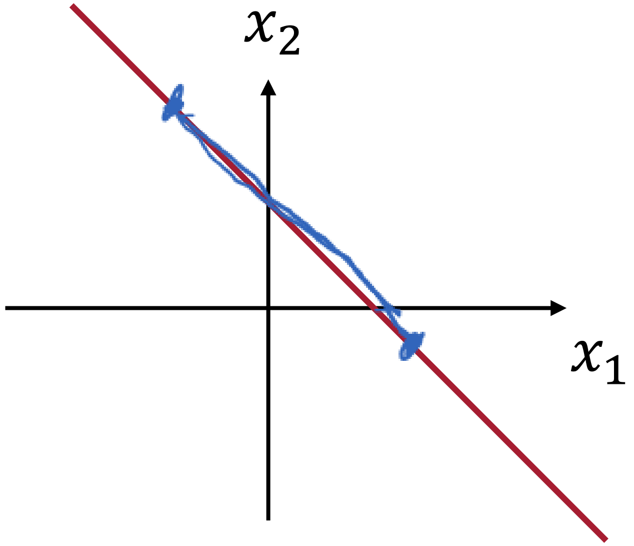
$$\{x: f(x) = 0\} = \{x: a^{\top}x + b = 0\}$$

is convex. Such a set is called as a hyperplane.

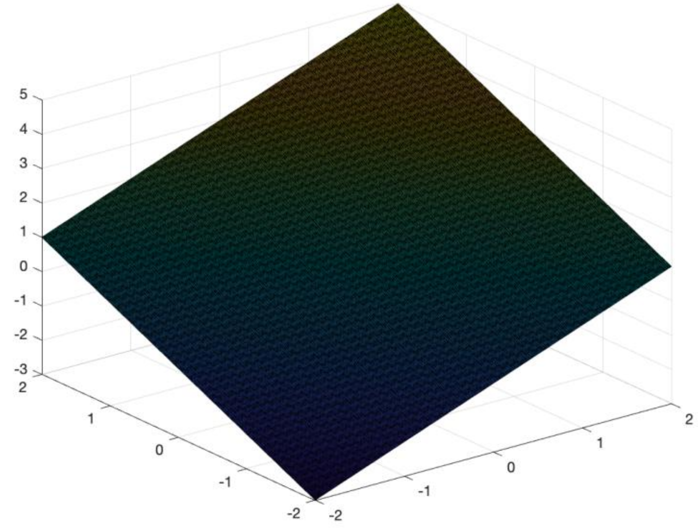
a b

$[1, 1]^{\top} x + 1 = 0$

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



$\{x \in \mathbb{R}^3 : x_1 + x_2 - x_3 + 1 = 0\}$

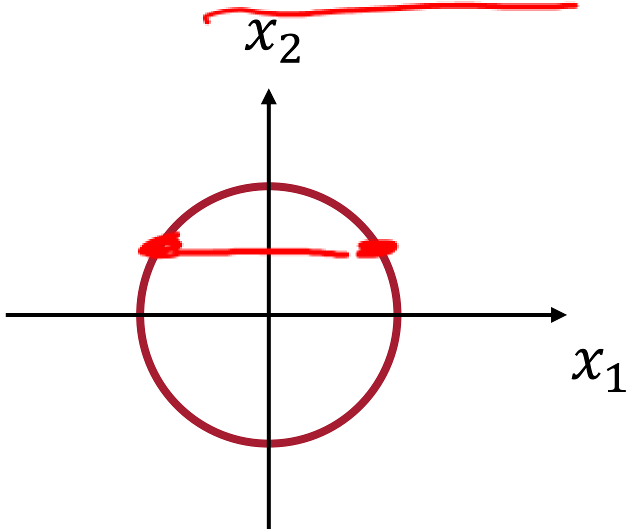


Equality Constraint

Alert! For non-affine functions $f(x)$, $\{x: f(x) = 0\}$ is generally NOT convex.

Examples:

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



Polyhedrons

$a_1^T x + b_1 \geq 0$

$-a_1^T x - b_1 \leq 0$

- Intersection of multiple affine inequality constraints

$x \in \mathbb{R}^n \quad a_i \in \mathbb{R}^n$

$$Ax + b = \begin{bmatrix} a_1^T x + b_1 \\ a_2^T x + b_2 \\ \vdots \\ a_m^T x + b_m \end{bmatrix} \leq 0$$

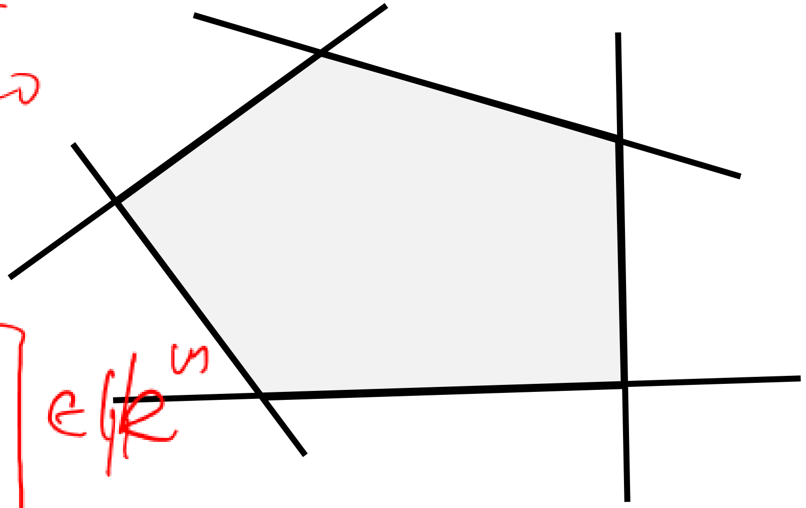
each element in vec ≤ 0

$$\left.\begin{aligned} a_1^\top x + b_1 &\leq 0 \\ &\dots \\ a_m^\top x + b_m &\leq 0 \end{aligned}\right\}$$

Short hand notation

$$Ax + b \leq 0$$

$$A = m \overset{n}{\begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix}} \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m$$



Polyhedrons

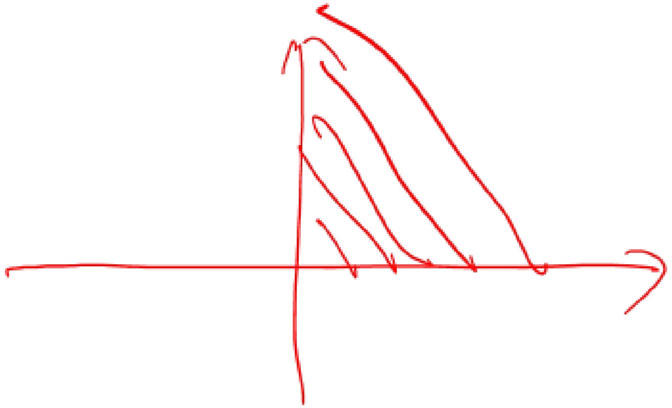
Example:

- Is the following set $\{x: Ax \leq b, Cx = d\}$ a polyhedron?

$Ax - b \leq 0 \quad Cx - d \leq 0$ and $-Cx + d \leq 0$

Intersect of 2 groups of affine ineq $\Rightarrow$ polyhedron

- Is the following set $\{x \in \mathbb{R}^n: x_i \geq 0, i = 1,2, \dots n\}$ a polyhedron?
  - Yes, aka the "non-negative orthant"



Revisit: Resource Allocation

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

s.t.

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

affine ineq. constraint

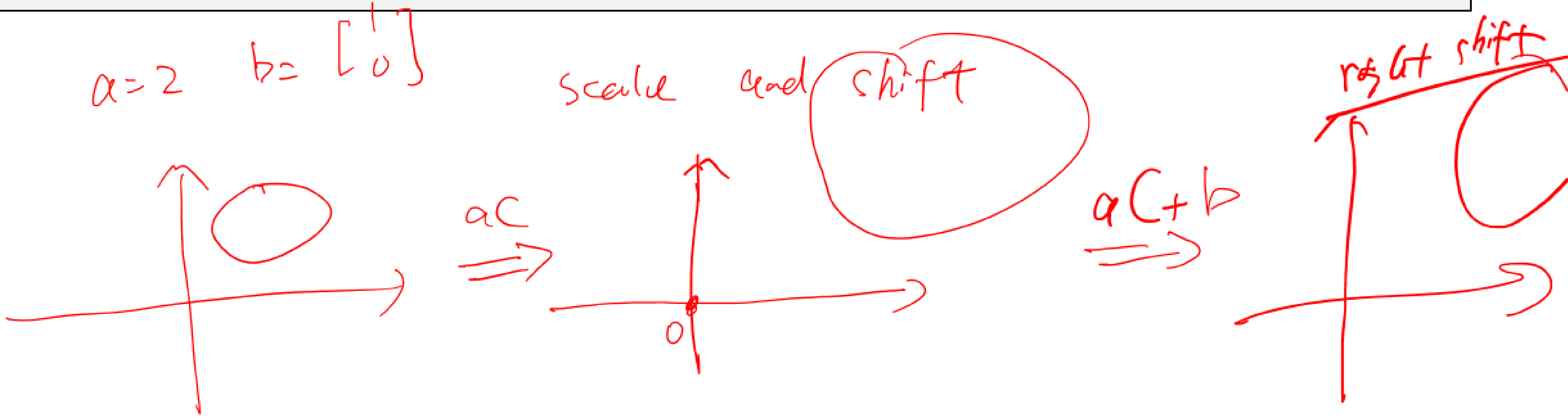
Is this constraint set convex? Yes

Operations Preserving Set Convexity

scalar vector

Scaling and translation: if $C \subset \mathbb{R}^n$ is convex, then for $a \in \mathbb{R}$, $b \in \mathbb{R}^n$, the following is convex.

$$aC + b = \{ax + b : x \in C\}$$



Operations Preserving Set Convexity

for any convex set D must exist convex func. g for which $D = \{y: g(y) \le 0\}$

$g(y) = \text{dist}(y, D)$

m-dim

Affine preimages: if $D \subset \mathbb{R}^m$ is convex, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, then the following is convex.

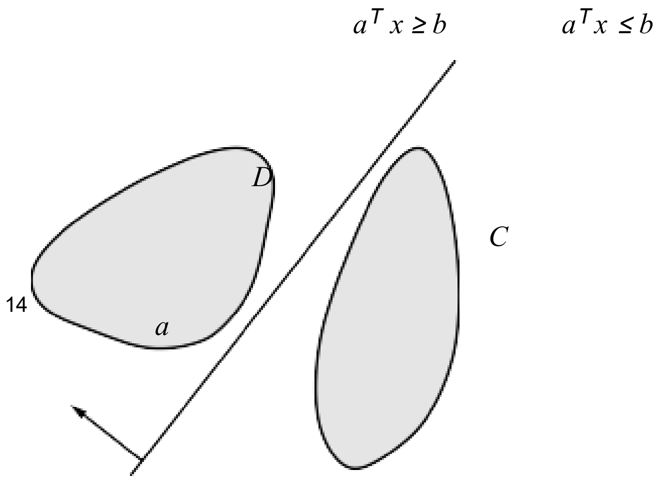
$$C = \{x \in \mathbb{R}^n : A\,x + b \;\in\; D\}$$

Prf: Let's prove for the case $D = \{y \in \mathbb{R}^m : g(y) \le 0\}$ for convex function g

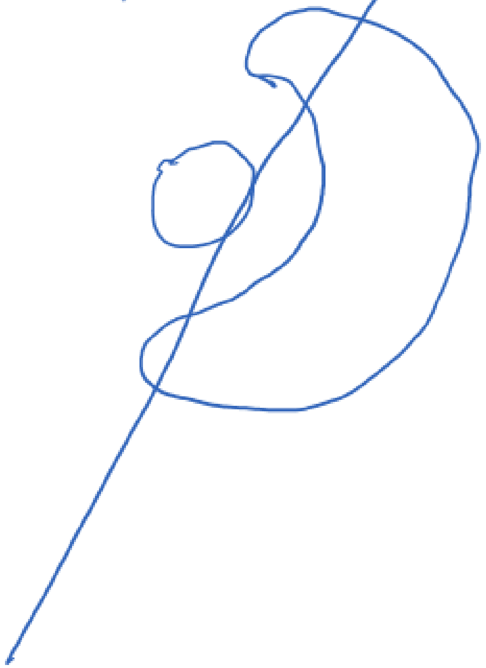
$Ax+b \in D \iff g(Ax+b) \le 0$ define $f(x) = g(Ax+b)$ convex. (composition of g and affine map)

$C = \{x : \underbrace{f}_{\text{convex}}(x) \le 0\}$ is convex

Separating Hyperplane Theorem



Non-Example



Theorem: if C, D are nonempty convex sets with $C \cap D = \emptyset$, then there exists a, b such that

$$C \subseteq \{x : a^\top x \leq b\} \qquad D \subseteq \{x : a^\top x \geq b\}$$

Summary

- Strong convexity, smoothness, condition number
- Convex constraint sets
  - Halfspaces
  - Norm balls
  - Ellipsoids
  - Hyperplanes
  - Polyhedrons
- Operations preserving convexity
  - Intersection
  - Scaling and translation
  - Preimage

$x^2+1=0$

Quiz Results

Correct Answer

Select all the sets that are convex.

(A) $\{x \in \mathbb{R}^n : x_1 + x_2 + \ldots + x_n \geq 0\}$	88 %
(B) $\{x \in \mathbb{R}^2 : x_1^2 + x_2^4 \leq 1\}$	100 %
(C) $\{x \in \mathbb{R}^2 : x_1^2 + x_2^3 \leq 1\}$	6 %
(D) $\{x \in \mathbb{R}^n : x^\top A x = 0\}$ where A is positive definite.	81 %

$f(x) = x_1^2 + x_2^4 - 1$

(D) the set only contains $\{0\}$

when A is p.d. $A = A^{\frac{1}{2}} \cdot A^{\frac{1}{2}}$ $\quad A^{\frac{1}{2}}$ is invertable

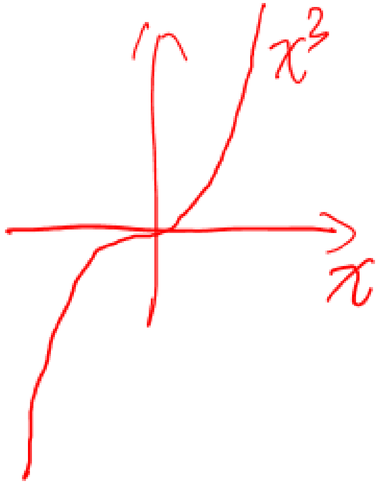
$x^T A x = 0 \Rightarrow x^T A^{\frac{1}{2}} A^{\frac{1}{2}} x = 0 \Rightarrow \|A^{\frac{1}{2}} x\|^2 = 0$

$\Rightarrow A^{\frac{1}{2}} x = 0 \Rightarrow x = 0$

(B) $f(x) = \underbrace{x_1^2}_{\text{convex}} + \underbrace{x_2^4}_{\text{convex}} - 1 \Rightarrow$ convex

$(x_2^4)'' = (4x_2^3)' = 12x_2^2 \geq 0$

(C) $f(x) = x_1^2 + \boxed{x_2^3}$ not convex



$(x_2^3)'' = (3x_2^2)' = 6x_2 \begin{cases} \geq 0 & \text{if } x_2 \geq 0 \\ < 0 & \text{if } x_2 < 0 \end{cases}$

non-convex

Convex Optimization Problem

- Convex program
- Feasibility
- Optimality
- Common types of convex programs

Typical Formulation of Convex Program

Decision Variable

$$\min_x f(x)$$

Objective func. where f convex

s.t.

$$a_1^\top x + b_1 = 0$$
$$\cdots$$
$$a_m^\top x + b_m = 0$$

Equality constraints
Usually written as $Ax + b = 0$

$$g_1(x) \le 0$$
$$\cdots$$
$$g_k(x) \le 0$$

Inequality constraints
Each $g_i(x)$ is convex function

$$x \in C$$

Additional constr. represented as set

Typical Formulation of Convex Program

$$\min_x f(x)$$ Objective func. where f convex

s.t. $$a_1^\top x + b_1 = 0$$
$$\cdots$$
$$a_m^\top x + b_m = 0$$

Equality constraints
Usually written as $Ax + b = 0$

$$g_1(x) \le 0$$
$$\cdots$$
$$g_k(x) \le 0$$

Inequality constraints
Each $g_i(x)$ is convex function

$$x \in C$$ Additional constr. represented as set

Flexibility in formulation

Often group

$$\begin{aligned} c_1^\top x + d_1 &\le 0 \\ c_2^\top x + d_2 &\le 0 \\ &\cdots \\ c_p^\top x + d_p &\le 0 \end{aligned}$$

into

$$Cx + d \le 0$$

Typical Formulation of Convex Program

$$\min_x f(x)$$

Objective func. where f convex

s.t.

$$a_1^\top x + b_1 = 0$$
$$\cdots$$
$$a_m^\top x + b_m = 0$$

Equality constraints
Usually written as $Ax + b = 0$

$$g_1(x) \le 0$$
$$\cdots$$
$$g_k(x) \le 0$$

Inequality constraints
Each $g_i(x)$ is convex function

$$x \in C$$

Additional constr. represented as set

Flexibility in formulation

RHS doesn't need to be 0, e.g.

$$h(x) \le \tilde{h}(x)$$

Which is equivalent to

$$h(x) - \tilde{h}(x) \le 0$$

Typical Formulation of Convex Program

$$\min_x f(x)$$ — Objective func. where f convex

s.t.

$$a_1^\top x + b_1 = 0$$
$$\cdots$$
$$a_m^\top x + b_m = 0$$

Equality constraints
Usually written as $Ax + b = 0$

$$g_1(x) \leq 0$$
$$\cdots$$
$$g_k(x) \leq 0$$

Inequality constraints
Each $g_i(x)$ is convex function

$$x \in C$$

Additional constr. represented as set

Flexibility in formulation

Can also include $\geq$, e.g.

$$h(x) \geq 0$$

Which is equivalent to

$$-h(x) \leq 0$$

Often do this for affine functions

Typical Formulation of Convex Program

$$\min_x f(x)$$ — Objective func. where f convex

s.t.

$$a_1^\top x + b_1 = 0$$
$$\cdots$$
$$a_m^\top x + b_m = 0$$

Equality constraints
Usually written as $Ax + b = 0$

$$g_1(x) \le 0$$
$$\cdots$$
$$g_k(x) \le 0$$

Inequality constraints
Each $g_i(x)$ is convex function

$$x \in C$$

Additional constr. represented as set

Implicit domain constr.

$x \in dom(f) \cap (\cap_{i=1}^k dom(g_k))$

e.g. $f(x) = \log x$

implies $dom(f) = (0, +\infty)$

Typical Formulation of Convex Program

Sometimes, the domain constraint and other simple constraints are written here.

$$\min_{x \in \mathbb{R}^n, x \geq 0} \; f(x)$$

Objective func. where f convex

$$\text{s.t.} \quad a_1^\top x + b_1 = 0$$
$$\cdots$$
$$a_m^\top x + b_m = 0$$

Equality constraints
Usually written as $Ax + b = 0$

log x - 1

(0, +∞)

$$g_1(x) \leq 0$$
$$\cdots$$
$$g_k(x) \leq 0$$

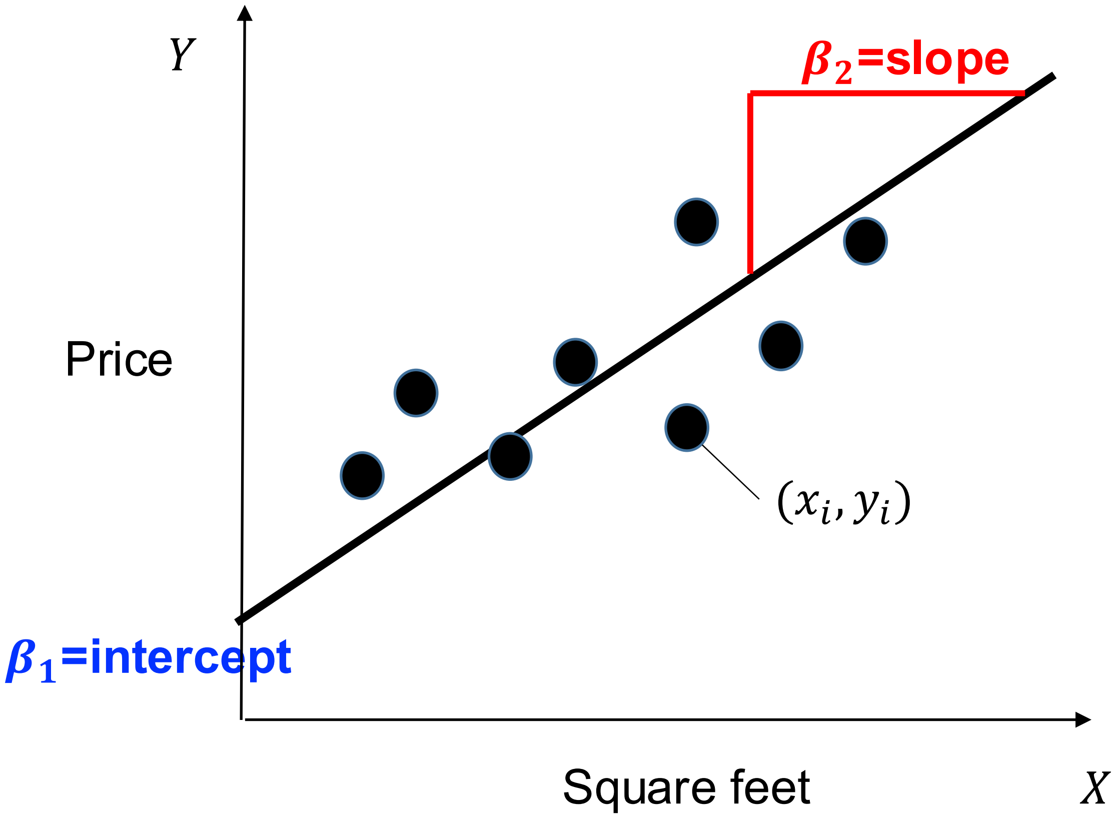
Inequality constraints
Each $g_i(x)$ is convex function

$$x \in C$$

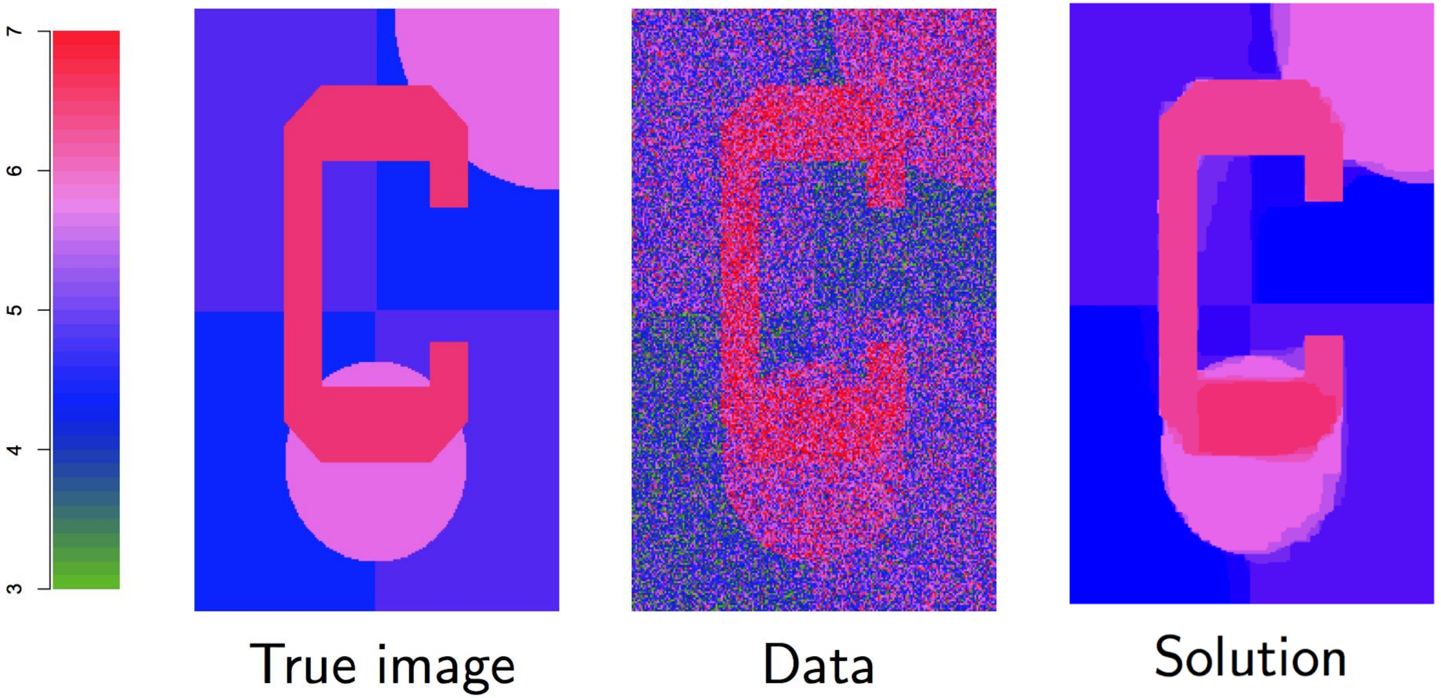
Additional constr. represented as set

Revisit: Linear Regression

$$\min_{\beta_1,\beta_2} \sum_{i=1}^{n} (y_i - \beta_1 - \beta_2 x_i)^2$$



Revisit: Denoising



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

θ_i stays close to y_i penalize changes in adjacent pixels

Feasibility

Definition. A convex program is feasible if its constraint set is non-empty. Otherwise, the convex program is infeasible.

$$\min_{x \in \mathbb{R}} f(x)$$

s.t. $x = 0$

$x = 1$

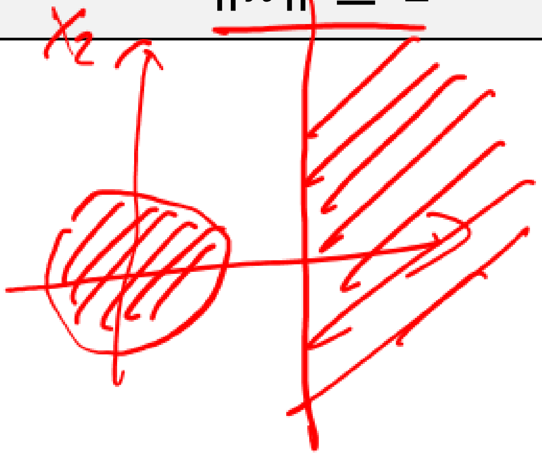
infeasible

infeasible

$$\min_{x \in \mathbb{R}^2} f(x)$$

s.t. $x_1 \geq 2$

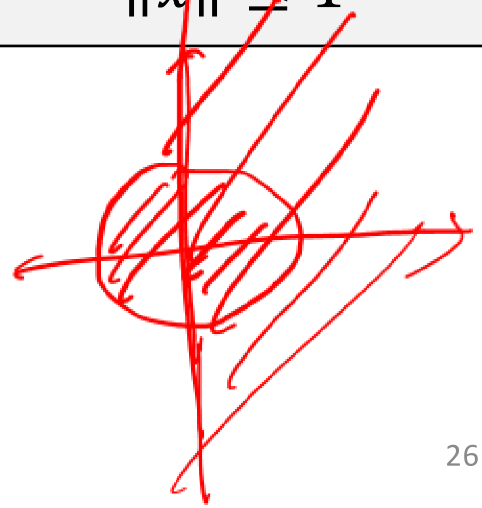
$\|x\| \leq 1$



$$\min_{x \in \mathbb{R}^2} f(x)$$

s.t. $x_1 \geq 0$

$\|x\| \leq 1$



Unconstrained convex program

Definition. A convex program is unconstrained if there are no constraints other than the implicit domain constraint.

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

$$\min_{\beta_1,\beta_2}\sum_{i=1}^{n}(y_i-\beta_1-\beta_2 x_i)^2$$

Optimality

Definition. Given a convex program

optimizer

minimizer

$$\min_{x \in C} f(x)$$

x^* is an (global) optimal solution if for any feasible point $x \in C$, we have

$$f(x^*) \le f(x)$$

The corresponding (global) optimal value of the convex program is $f^* = f(x^*)$

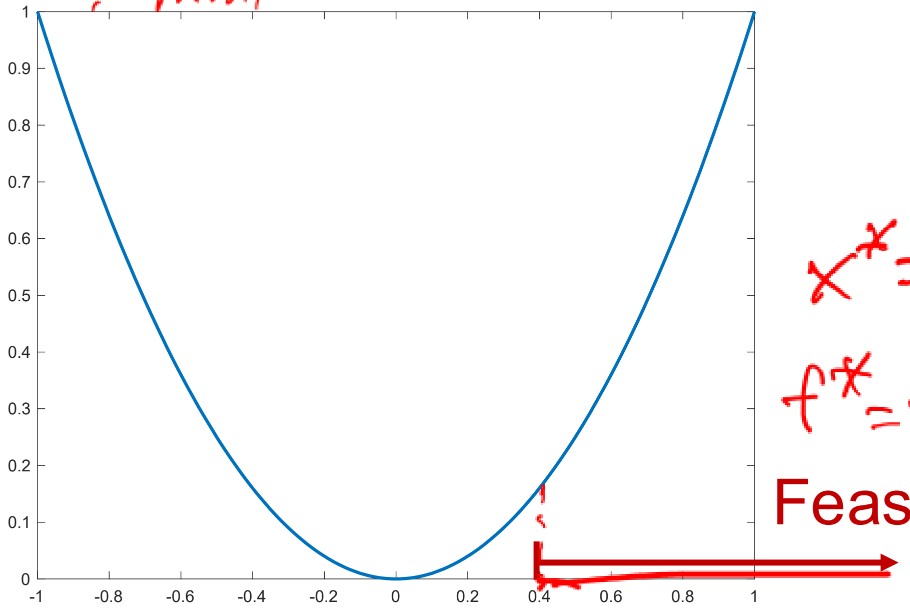
optimum, minimum

$$\min_{x \in \mathbb{R}} x^2$$

s.t. $\quad x \geq 0.4$

Optimal solution: $x^* = 0.4$

Optimal value: $f^* = 0.16$



$x^* = 0.4$

$f^* = f(x^*) = 0.4^2$

Feasible region

Optimality

Definition. Given a convex program

$$\min_{x \in C} f(x)$$

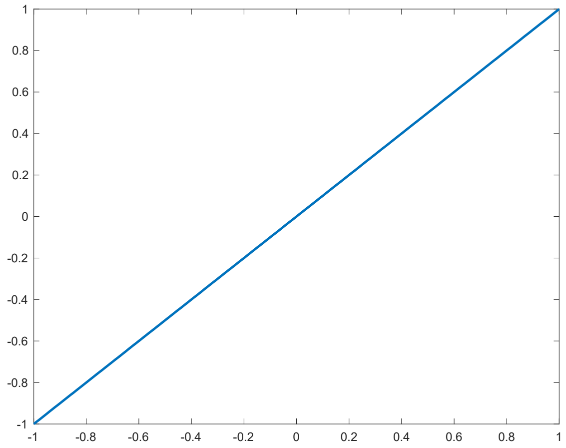
x^* is an (global) <u>optimal solution</u> if for any feasible point $x \in C$, we have

$$f(x^*) \leq f(x)$$

The corresponding (global) <u>optimal value</u> of the convex program is $f^* = f(x^*)$

Alert: optimal solution may not even exist!

$$\min_{x \in \mathbb{R}} x$$

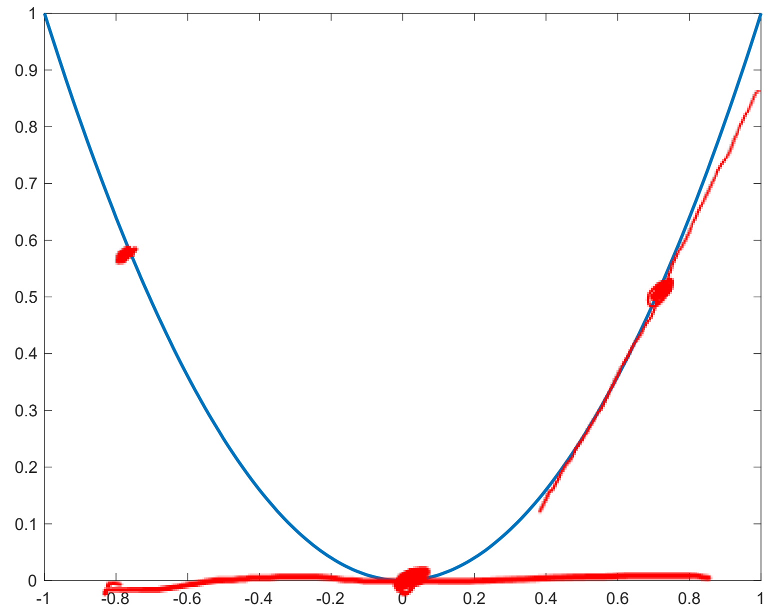


nonunique

First Order Optimality Condition

How to check whether a given x is an optimal solution?

$$f(x) = x^2$$



First Order Optimality Condition

$\min_{x \in \mathbb{R}} x = f(x)$ $\nabla f(x) = 1$

Theorem. For **unconstrained** convex optimization problem with differentiable objective $f(x)$, x^* is an optimal solution if and only if $\nabla f(x^*) = 0$.

Proof. $\Longrightarrow$: suppose x^* is an optimal solution, want to prove $\nabla f(x^*) = 0$

Proof by contradiction. Assume $\nabla f(x^*) \neq 0$

Define $g(t) = f(x^* - t\nabla f(x^*))$ $t \in \mathbb{R}$

$g'(t) = \nabla f(x^* - t\nabla f(x^*))^T \cdot [-\nabla f(x^*)]$

$g'(0) = \nabla f(x^*)^T \cdot [-\nabla f(x^*)] = -\|\nabla f(x^*)\|^2 < 0$

$g(t)$ is strictly decreasing at $t=0$.

Exist $t^* > 0$, s.t. $g(t^*) < g(0) \Rightarrow f(\underbrace{x^* - t^* \nabla f(x^*)}_{\text{new. better solution}}) < f(x^*)$

contradiction

First Order Optimality Condition

Theorem. For **unconstrained** convex optimization problem with differentiable objective $f(x)$, x^* is an optimal solution if and only if $\nabla f(x^*) = 0$.

Proof. ⇐: suppose $\nabla f(x^*) = 0$, want to prove x^* is an optimal solution

First Order Optimality Condition

Theorem. For **<u>unconstrained</u>** convex optimization problem with differentiable objective $f(x)$, x^* is an optimal solution if and only if $\nabla f(x^*) = 0$.

Summary of key proof idea:

- $f\big(x^* - t\nabla f(x^*)\big) < f(x^*)$ for small t, non-zero $\nabla f(x^*)$
  - Negative gradient direction is a "descent" direction!

- $\nabla f(x^*) = 0$ eliminates the possibility of obtaining lower cost by "descent".

First Order Optimality Condition: Examples

Quadratic cost functions:

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} x^\top A x + b^\top x + c$$

Calculate Gradient

Optimal x^* Satisfies

Leading to optimal solution

And optimal value

First Order Optimality Condition: Examples

Quadratic cost functions:

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} x^\top A x + b^\top x + c$$

Calculate Gradient $\quad \nabla f(x) = Ax + b$

Optimal x^* Satisfies $\quad \nabla f(x^*) = Ax^* + b = 0$

For general cost f, we can't solve $\nabla f(x^*) = 0$ directly. But the gradient is useful and form the basis of gradient based methods, which we will cover in lecture 9.

First Order Optimality Condition

What about constrained convex program?

Idea: the condition shall eliminate the possibility of obtaining lower cost by descent.

First Order Optimality Condition

Theorem. Given a convex program $\min_{x \in C} f(x)$ with differentiable objective $f(x)$, x^* is an optimal solution if and only if

$$\nabla f(x^*)^\top (x - x^*) \geq 0, \forall x \in C$$

If x^* lies in the interior of C, the condition is equivalent to $\nabla f(x^*) = 0$

First Order Optimality Condition

Theorem. Given a convex program $\min_{x \in C} f(x)$ with differentiable objective $f(x)$, x^* is an optimal solution if and only if

$$\nabla f(x^*)^\top (x - x^*) \geq 0, \forall x \in C$$

If x^* lies on the boundary of C, the condition means gradient direction is “aligned” with the set C.

Summary

- Standard ways to write a convex program
- Feasibility
- Optimality conditions