Lecture 21 – ME6402, Spring 2025
A Brief Introduction to Convex Optimization

Maegan Tucker

March 27, 2025

Georgia Tech

Goals of Lecture 21

- Define optimization problems
- Define convex functions and sets
- Define convex optimization problems

Additional Reading

- S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

Optimization Problems

We often encounter problems of the form

$$\begin{aligned} \text{minimize}_x \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \qquad i = 1, \dots, m \end{aligned}$$

where:

- $x \in \mathbb{R}^n$ is an optimization variable,
- f_0 is the objective function, and
- $f_i(x)$ are constraint functions.

The *optimal value* of $f_0(x)$ is the (limit of the) smallest value obtained by $f_0(x)$ on the *feasible set*. A point that achieves the optimal value (*i.e.*,argmin) is an *optimal point*.

- Of course, if we are instead interested in maximizing a function $\tilde{f}_0(x)$, we simply define $f_0(x) = -\tilde{f}_0(x)$ to change to a minimization problem.
- Equality constraint $f(x) = 0$ is allowed by including two constraints: $f(x) \leq 0$ and $-f(x) \leq 0$.

Example

Minimum effort stabilization from CLF:

Given system $\dot{x} = f(x) + g(x)u$ and CLF $V(x)$, use optimization-based controller

$$
\begin{aligned}
k(x) = &\text{argmin}_u \quad \|u\|^2 \\
&\text{subject to} \quad \frac{\partial V}{\partial x}(f(x) + g(x)u) \leq -\varepsilon A(x),
\end{aligned}
$$

- ε is user chosen
- $A(x)$ is some positive definition function. $A(x) = x^T x$ or $A(x) = V(x)$ are common choices
- Generally **cannot** consider a strict inequality constraint like $\dot{V}(x) < 0$, hence the need for $\varepsilon A(x)$

Example 2

Finding polynomial Lyapunov functions:

Given system $\dot{x} = f(x)$, solve

$$k(x) = \text{argmin}_c \quad 0$$
$$\text{subject to} \quad V(x) \geq \varepsilon_1 A(x)\ \forall x$$
$$\frac{\partial V}{\partial x} f(x) \leq -\varepsilon_2 A(x)\ \forall x$$

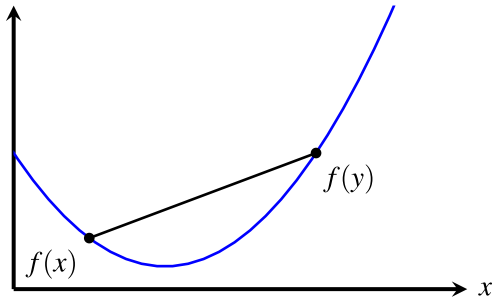
where, *e.g.*, $x \in \mathbb{R}^2$, $V(x) = c_1 x_1^4 + c_2 x_1^3 x_2 + c_3 x_1^2 x_2^2 + c_4 x_1 x_2^3 + c_5 x_2^4 + c_5 x_1^3 + \ldots + c_{n-2} x_1 + c_{n-1} x_2 + c_n$

- No cost $\implies$ feasibility question
- "$\forall x$" $\implies$ infinite, uncountable number of constraints

Convex functions and sets

A *convex function* $f : \mathbb{R}^n \to \mathbb{R}$ satisfies for all x, y and all $0 \leq \theta \leq 1$:

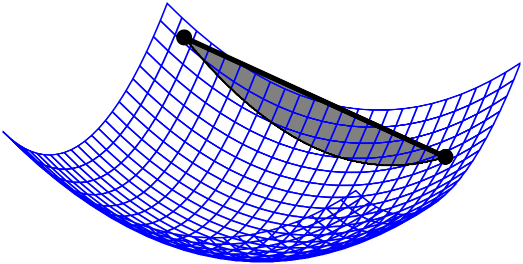
$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y).$$



Convex functions and sets

A *convex function* $f : \mathbb{R}^n \to \mathbb{R}$ satisfies for all x, y and all $0 \leq \theta \leq 1$:

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y).$$



Example

$f(x) = c^T x$ for fixed $c \in \mathbb{R}^n$:

- Convexity:

$$f(\theta x + (1-\theta)y)$$
$$\leq \theta f(x) + (1-\theta)f(y).$$

Example

$f(x) = c^T x$ for fixed $c \in \mathbb{R}^n$:

$$\begin{aligned} f(\theta x + (1-\theta)y) &= c^T(\theta x + (1-\theta)y) \\ &= \theta c^T x + (1-\theta)c^T y \\ &= \theta f(x) + (1-\theta)f(y), \end{aligned}$$

so f is convex (satisfies the required inequality with equality for all $\theta \in [0,1]$).

- Convexity:

$$\begin{aligned} & f(\theta x + (1-\theta)y) \\ & \leq \theta f(x) + (1-\theta)f(y). \end{aligned}$$

First Order and Second Order Tests for Convexity

Fact. When f is once differentiable, f is convex if and only if $f(y) \geq f(x) + \nabla f(x)^T (y - x)$ for all x, y.

Fact. When f is twice differentiable, f is convex if and only if $\nabla^2 f(x) \succeq 0$ for all x.

- The notation $M \succeq 0$ or $M \succ 0$ for a square symmetric matrix M means that M is *positive (semi)definite (PD/PSD)*. Recall that M is PSD (respectively, positive definite) if $x^T M x \geq 0$ (respectively, $x^T M x > 0$ for all x). In previous lectures, we did not use this additional notation.

Example 1

Example. Consider the quadratic function

$$f(x) = \frac{1}{2}x^T P x + q^T x + r, \qquad P = P^T$$

Then $\nabla^2 f(x) = P$ for all x, so quadratic functions are convex if and only if $P \geq 0$, *i.e.*, P is a positive semidefinite matrix.

Example 2

Example. Any norm $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is convex:

$$\|\theta x + (1-\theta)y\| \leq \|\theta x\| + \|(1-\theta)y\| = \theta\|x\| + (1-\theta)\|y\|$$

- Recall that a norm $\|\cdot\|$ satisfies:

1. $\|x+y\| \leq \|x\| + \|y\|$ (Triangle inequality)
2. $\|ax\| = |a|\|x\|$
3. if $\|x\| = 0$ then $x = 0$

Example 3

Example. If f is convex, then

$$g(x) = f(Ax + b)$$

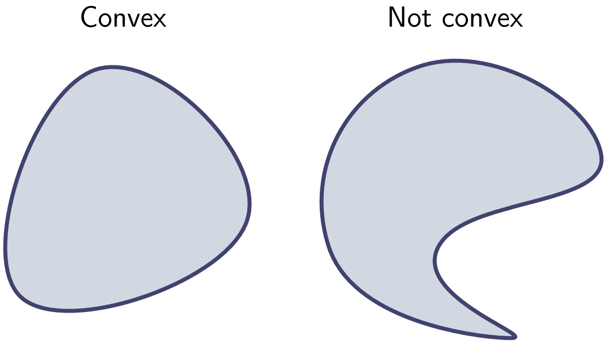
is convex for any A, b:

$$\begin{aligned} g(\theta x + (1-\theta)y) &= f(A(\theta x + (1-\theta)y) + b) \\ &= f(\theta(Ax+b) + (1-\theta)(Ay+b)) \\ &\leq \theta f(Ax+b) + (1-\theta)f(Ay+b) \\ &= \theta g(x) + (1-\theta)g(y). \end{aligned}$$

Convex Sets

A *convex set* $\mathcal{C}$ satisfies

whenever $x_1, x_2 \in \mathcal{C}$, then $\theta x_1 + (1-\theta)x_2 \in \mathcal{C}$ for all $0 \leq \theta \leq 1$.



Example: Convex Sets as Sublevel Sets of Convex Functions

Example. Any α-*sublevel* set $C_\alpha = \{x : f(x) \leq \alpha\}$ of a convex function is convex.

Proof. Choose $x, y \in C_\alpha$ so that $f(x) \leq \alpha$ and $f(y) \leq \alpha$. By convexity, $f(\theta(x) + (1-\theta)y) \leq \alpha$ for any $0 \leq \theta \leq 1$, and hence $\theta x + (1-\theta)y \in C_\alpha$. □

The converse does *not* hold.

Convex Optimization

Optimization problem from Slide 2:

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \qquad i = 1, \ldots, m \end{aligned}$$

The above optimization problem is *convex* if f_0 and all f_i's are convex.

- In this case, the feasible set is a convex set.

Example: Equality Constraints

Example: Convex optimization problems may include equality constraints, but only if they are affine, *i.e.*, of the form $Ax+b=0$.

Proof. To include an equality constraint $f_i(x)=0$, we add $f_i(x)\le 0$ and $-f_i(x)\le 0$ as inequality constraints. To be convex, we require $f(x)$ and $-f(x)$ to be convex. The only such functions are affine. To see this, we assume f_i is differentiable. Then convexity of f_i and $-f_i$ means:

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

and

$$-f_i(y)\ge -f_i(x)-\nabla f(x)^T(y-x),$$

so that $f_i(y)=f_i(x)+\nabla f_i(x)^T(y-x)$ for any x, y, *i.e.*, f_i is affine. $\square$

Feasibility of Convex Problems

<u>Theorem:</u> For a convex optimization problem, a feasible point x is optimal if and only if $\nabla f_0(x)^T(y-x) \geq 0$ for all feasible y.

Proof. (if) Since f_0 is convex, for any x, y.

$$f_0(y) \geq f_0(x) + \nabla f_0(x)^T(y-x).$$

Let x be a feasible point such that $\nabla f_0(x)^T(y-x) \geq 0$ for all feasible y. Then for any feasible $y \neq x$, using (**??**), $f_0(y) \geq f_0(x)$ and x is optimal.

Feasibility of Convex Problems

<u>Theorem:</u> For a convex optimization problem, a feasible point x is optimal if and only if $\nabla f_0(x)^T(y-x) \geq 0$ for all feasible y.

Proof. (only if) Now suppose x is optimal but there is some feasible y such that $\nabla f_0(x)^T(y-x) < 0$. The point $z_\theta = \theta y + (1-\theta)x$ must also be feasible since the feasible set is convex. For small θ, $f(z_\theta) < f(x)$ since $\frac{d}{d\theta} f_0(z_\theta)|_{\theta=0} = \nabla f_0(z_\theta)^T(y-x)\big|_{\theta=0} = \nabla f_0(x)^T(y-x) < 0$.

□

Optimality for Unconstrained Convex Optimization Problems

When all y are feasible, the above condition reduces to: x is optimal if and only if

$$\nabla f_0(x) = 0.$$

Example

Consider

$$\text{minimize}_x \quad \frac{1}{2}x^T P x + q^T x + r$$

where $P \succeq 0$. Then x is optimal if and only if $Px + q = 0$. Three cases:

1. If $q \notin \text{Range}(P)$, no solution. In this case, objective function is unbounded (below)
2. If P is nonsingular (i.e.,$P \succ 0$), then $x^* = -P^{-1}q$ is unique solution
3. If P is singular but $q \in \text{Range}(P)$, then set of optimal points is affine subspace $\{x \mid Px = -q\} = \{x^* + y \mid y \in \text{Null}(P), x^* \text{ is any vector such that } Px^* = -q\}$.