ME 6402 – Lecture 22

CONVEX OPTIMIZATION PROBLEMS

April 1 2025

Overview:

- Introduce important classes of convex optimization problems

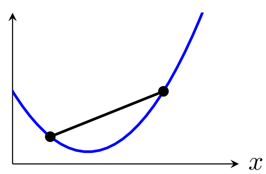
Additional Reading:

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

Recall: Convex functions and sets

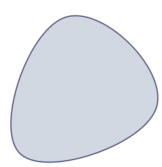
Definition: *Convex function.* A convex function $f : \mathbb{R}^n \to \mathbb{R}$ satisfies the following condition for all $x, y \in \mathbb{R}^n$ and all $0 \leq \theta \leq 1$:

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



Definition: *Convex Set.* A convex set $\mathcal{C}$ satisfies the following condition:

$$\text{if } x, y \in \mathcal{C} \text{ then } \theta x + (1-\theta)y \in \mathcal{C} \text{ for all } 0 \leq \theta \leq 1$$



Example 1:

The probability simplex is the set of vectors $x \in \mathbb{R}^n$ such that $x \geq 0$ and $\mathbb{1}^T x = 1$. It is a convex set.

Let x_1 and x_2 be two elements of the probability simplex. For any $0 \leq theta \leq 1$,

$$\theta x_1 + (1-\theta)x_2 \geq 0$$

and

$$\mathbb{1}^T(\theta x_1 + (1-\theta)x_2) = \theta \mathbb{1}^T x_1 + (1-\theta)\mathbb{1}^T x_2 = \theta + (1-\theta) = 1$$

Example 2:

The set of symmetric matrices in $\mathbb{R}^{n\times n}$ is a vector space (what is its dimension?). The subset of symmetric positive semidefinite matrices is a convex subset of this vector space. In fact, for any P.S.D. matrices X_1 and X_2, and any $\theta_1 \geq 0$, and $\theta_2 \geq 0$, $\theta_1 X_1 + \theta_2 X_2$ is also P.S.D.

$$x^T(\theta_1 X_1 + \theta_2 X_2)x = \theta_1 \underbrace{x^T X_1 x}_{\geq 0} + \theta_2 \underbrace{x^T X_2 x}_{\geq 0} \geq 0$$

for all x.

Convex Optimization

Recall the optimization problem

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0, \quad i = 1, \dots, m \end{aligned}$$

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:

Least Squares is a convex optimization problem:

$$\underset{x}{\text{minimize}} \quad \|Ax - b\|_2^2$$

We know that $\|\cdot\|_2$ is convex because it is a norm (we proved this in the last lecture), $\|\cdot\|_2^2$ is also convex (convince yourself of this), and composition with affine transformation preserves convexity (we also proved this in the last lecture).

- This is an *unconstrained* optimization problem since there are no constraints
- Optimization problems rarely have closed form solutions, but the least squares problem does: $x = (A^T A)^{-1} A^T b$

Linear Optimization Programs (LP)

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^T x \\ \text{s.t.} \quad & a_i^T x \leq b_i, \quad i = 1, \dots, m \end{aligned}$$

- Linear programs are a class of optimization problems that can be solved very efficiently
- If feasible set is compact, then vertices of feasible region contain optimal points

Quadratic Optimization Programs (QP)

Quadratic costs give rise to quadratic optimization problems. Quadratic programs (QPs) are a class of optimization problems that have quadratic cost with affine constraints:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \frac{1}{2}x^T P x + q^T x + r \\ \text{s.t.} \quad & a_i^T x \le b_i, \quad i = 1, \ldots, m \end{aligned}$$

where P is positive semidefinite, $P \succeq 0$.

Quadratically Constrained Quadratic Optimization Programs (QCQP)

A QCQP has quadratic cost with quadratic constraints:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \frac{1}{2}x^T P_0 x + q_0^T x + r_0 \\ \text{s.t.} \quad & \frac{1}{2}x^T P_i x + q_i^T x + r_i, \quad i = 1, \ldots, m \end{aligned}$$

where all P_i's are positive semidefinite.

- Least squares is a QP because $\|Ax - b\|_2^2 = x^T A^T A x - 2b^T A x + b^T b$, with $A^T A \succeq 0$.
- All LPs are QPs, all QPs are QCQPs.

Second-Order Cone Programs (SOCP)

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f^T x \\ \text{s.t.} \quad & \|A_i x + b_i\|_2 \le c_i^T x + d_i, \quad i = 1, \ldots, m \end{aligned}$$

- If all c_i's are zero, then SOCP reduces to QCQP
- If all A_i's are zero, then SOCP reduces to LP

Linear matrix inequalities (LMIs) and semidefinite programs (SDPs)

A twist: Instead of scalar inequality ($\leq$) in constraints, what if we allowed for matrix inequalitie ($\preceq$)?

First form:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^T x \\ \text{s.t.} \quad & x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0 \end{aligned}$$

where $F_1, \ldots, F_n, G$ are all symmetric matrices.

- The inequality above is called a linear matrix inequality (LMI)
- An optimization problem is a *semidefinite program* (SDP) if the constraints are LMIs and the cost is linear
- When $F_1, \ldots, F_n, G$ are actually scalars, we recover a standard affine constraint $f^T x + g \leq 0$ (LP)

Let's check that the constraint $x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0$ leads to a convex feasible set. To do this, let $x_1, x_2, \ldots, x_n$ and $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n$ be two sets satisfying the semidefinite inequality. Then,

$$\begin{aligned} & (\theta x_1 + (1-\theta)\hat{x}_1)F_1 + \cdots + (\theta x_n + (1-\theta)\hat{x}_n)F_n + G \\ & = \theta(x_1 F_1 + \cdots + x_n F_n + G) + (1-\theta)(\hat{x}_1 F_1 + \cdots + \hat{x}_n F_n + G) \\ & \preceq 0 \end{aligned}$$

Thus the constraint leads to a convex feasible set.

Note: Multiple LMIs can be combined into one LMI via block diagonalization:

$$\begin{aligned} x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0 \\ x_1 \hat{F}_1 + x_2 \hat{F}_2 + \cdots + x_n \hat{F}_n + \hat{G} \preceq 0 \end{aligned}$$

is the same as

$$x_1 \begin{bmatrix} F_1 & 0 \\ 0 & \hat{F}_1 \end{bmatrix} + \cdots + x_n \begin{bmatrix} F_n & 0 \\ 0 & \hat{F}_n \end{bmatrix} + \begin{bmatrix} G & 0 \\ 0 & \hat{G} \end{bmatrix} \preceq 0$$

Second form[1]:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \text{trace}(CX) \\ \text{s.t.} \quad & \text{trace}(A_i X) = b_i, \quad i = 1, \ldots, m \\ & X \succeq 0 \end{aligned}$$

- These two forms can be shown to be equivalent

[1] Recall that the first form of our SDP was

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^T x \\ \text{s.t.} \quad & x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0 \end{aligned}$$

- Seemingly more general constraints can be reduced to LMI constraints of the form above.
- In particular, matrix variables that appear linearly in semidefinite constraints are allowed.

Again, let's check that the constraints $\text{trace}(A_i X) = b_i$, with $X \succeq 0$ leads to a feasible set. To do this, let X_1 and X_2 both be feasible. Then,

$$\theta X_1 + (1-\theta)X_2 \succeq 0$$

and

$$\begin{aligned}
&\text{trace}(A_i(\theta X_1 + (1-\theta)X_2)) \\
&= \theta\,\text{trace}(A_i X_1) + (1-\theta)\,\text{trace}(A_i X_2) \\
&= \theta b_i + (1-\theta) b_i = b_i
\end{aligned}$$

for all $0 \leq \theta \leq 1$.

LMI Examples

Example 1: The Lyapunov inequality is given by $L(X) = A^T X + XA$ and we know A is Hurwitz if and only if there exists $X \succ 0$ such that $L(X) \prec 0$. $L(X)$ is linear in X. To see that, consider $X = aX_1 + bX_2$ and notice that

$$\begin{aligned}
L(X) &= A^T(aX_1 + bX_2) + (aX_1 + bX_2)A \\
&= a(A^T X_1 + X_1 A) + b(A^T X_2 + X_2 A) \\
&= aL(X_1) + bL(X_2)
\end{aligned}$$

Thus, $L(X) \preceq -\varepsilon I$ for some $\varepsilon > 0$ is a LMI constraint in the variable X.

Example 2: Consider $\dot{x} = A(t)x$ where $A(t)$ switches from among the set $\{A_1, \ldots, A_m\}$

- Even if all A_i are Hurwitz, stability is not guaranteed
- How could we prove asymptotic stability of $x = 0$?

One approach: Find a common Lyapunov function $V(x) = x^T P x$ that works for all A_is. Pose as an SDP:

$$\begin{aligned}
\underset{P}{\text{minimize}} \quad & \text{trace}(P) \\
\text{s.t.} \quad & PA_i + A_i^T P \preceq -\varepsilon I, \quad i = 1, \ldots, m \qquad P \succeq I
\end{aligned}$$

Solving convex optimization problems

Even though analytic solutions to convex optimization problems rarely exist, solvers have become so good and so fast that it is common to think of exact solutions to convex optimization problems as being readily available.

- CVX, CVXPY, CVXOPT, YALMIP are all basic purpose packages for solving convex optimization problems[2]
- Specialized functions such as MATLAB's `quadprog` for specific classes of problems (quadratic, in this case)

[2] Note: As a student, you all have free axis to GITHUB Copilot which can help you try out new coding languages faster by providing help with syntax.

Example: CVX provides easy coding of convex optimization problems in MATLAB (or CVXPY in Python). For example, consider the following Least Squares QP:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \|Ax - b\|_2^2 \\ \text{s.t.} \quad & Cx \le d \end{aligned}$$

this can be implemented as

```
cvx_begin
  variable x(n)
  minimize( norm(A*x - b, 2) )
  subject to
    C*x <= d
cvx_end
```