Reading for This Lecture

- C&T Chapter 1
Review from Last Time

- Recall that a mathematical model consists of:
  - Decision variables (with domains)
  - Constraints (functions of the variables with domains)
  - Objective Function (maximize or minimize)
  - Parameters and Data

The general form of a *mathematical optimization model* is:

\[
\begin{align*}
\text{min or max } & \quad f(x_1, \ldots, x_n) \\
\text{s.t. } & \quad g_i(x_1, \ldots, x_n) \begin{cases} 
\leq \\
eq \\
\geq 
\end{cases} b_i \\
( & \quad x_1, \ldots, x_n \quad \in \quad X
\end{align*}
\]

where \( X \) may be a discrete set. Without loss of generality, we may assume that all constraints are of the "\( \geq \)" type.
Categorizing Mathematical Optimization Problems

• Mathematical optimization problems can be categorized along several fundamental lines.
  – Constrained vs. Unconstrained
  – Convex vs. Nonconvex
  – Linear vs. Nonlinear
  – Discrete vs. Continuous

• What is the importance of these categorizations?
  – Knowing what category an instance is in can tell us something about how difficult it will be to solve.
  – Different solvers are designed for different categories.
Unconstrained Optimization

• When $M = \emptyset$ and $X = \mathbb{R}^n$, we have an *unconstrained optimization problem*.

• Unconstrained optimization problems will not generally arise directly from applications.

• They do, however, arise as *subproblems* when solving mathematical optimization problems.

• In unconstrained optimization, it is important to distinguish between the *convex* and *nonconvex* cases.
Linear Optimization Problems

• A linear optimization problem is one that can be written in a form in which the functions $f$ and $g_i, i \in M$ are all linear and $X = \mathbb{R}^n$.

• In general, a linear optimization problem is one that can be written as

\[
\begin{align*}
\text{minimize} & \quad c^\top x \\
\text{s.t.} \quad & a_i^\top x \geq b_i \quad \forall i \in M_1 \\
& a_i^\top x \leq b_i \quad \forall i \in M_2 \\
& a_i^\top x = b_i \quad \forall i \in M_3 \\
& x_j \geq 0 \quad \forall j \in N_1 \\
& x_j \leq 0 \quad \forall j \in N_2
\end{align*}
\]

• Equivalently, a linear optimization problem can be written as

\[
\begin{align*}
\text{minimize} & \quad c^\top x \\
\text{s.t.} \quad & Ax \geq b
\end{align*}
\]

• Generally speaking, linear optimization problems are “easy” to solve.
Nonlinear Optimization Problems

- A **nonlinear optimization problem** is any mathematical optimization problem that cannot be expressed as a linear optimization problem.

- Usually, this terminology also assumes $X = \mathbb{R}^n$.

- Note that by this definition, it is not always obvious whether a given instance is really nonlinear.

- In general, nonlinear optimization problems are difficult to solve to global optimality.
Special Case: Convex Optimization Problems

- A *convex optimization problem* is a nonlinear optimization problem in which the objective function $f$ is convex and the feasible region

$$\mathcal{F} = \{x \in \mathbb{R}^n \mid g_i(x) \geq b_i\}$$

is a convex set.

- In practice, convex optimization problems are usually “easy” to solve.
Special Case: Quadratic Optimization Problems

• A **quadratic optimization problem** is one in which all of the functions $f$ and $g_i$ for $i \in M$ are quadratic functions.

• Often, the term **quadratic optimization problem** refers specifically to an optimization problem of the form

$$\begin{align*}
\text{minimize} \quad & \frac{1}{2} x^\top Q x + c^\top x \\
\text{s.t.} \quad & Ax \geq b
\end{align*}$$

• Because $x^\top Q x = \frac{1}{2} x^\top (Q + Q^\top) x$, we can assume without loss of generality that $Q$ is **symmetric**.

• The objective function of the above optimization problem is then convex if and only if $Q$ is **positive semidefinite**, i.e., $y^\top Q y \geq 0$ for all $y \in \mathbb{R}^n$.

• There are specialized methods for solving convex quadratic optimization problems efficiently.
Special Case: Integer Optimization Problems

- When $X = \mathbb{Z}^n$, we have an integer optimization problem.
- When $X = \mathbb{Z}^r \times \mathbb{R}^{n-r}$, we have a mixed integer optimization problem.
- By convention, all functions are assumed to be linear in these cases unless otherwise specified.
- If some of the functions are nonlinear, then we have a mixed integer nonlinear optimization problem.
- All mathematical optimization problem with integer variables are difficult to solve in general.
Probability Review

• Stochastic optimization is essentially mathematical optimization with “random” parameters.

• Therefore, we’ll need to dig out just a little probability theory.

• The symbol $\omega$ will denote the outcome of a random experiment.

• The set of all possible outcomes, called the sample space, will generally be denoted $\Omega$.

• Subsets of $\Omega$ are called events.
**Probability spaces**

- Let $\mathcal{A}$ be a set of events.
- A probability measure (or distribution) $P$ is a function that indicates the probability that each event $A \in \mathcal{A}$ will occur.
- Probability measures must satisfy certain axioms and have the following basic properties.
  - $0 \leq P(A) \leq 1$
  - $P(\Omega) = 1$, $P(\emptyset) = 0$
  - $P(A_1 \cup A_2) = P(A_1) + P(A_2)$ if $A_1 \cap A_2 = \emptyset$.
- The triple $(\Omega, \mathcal{A}, P)$ is called a *probability space*. 
Random Variables

• A random variable $\xi$ on a probability space $(\Omega, \mathcal{A}, P)$ is a function $\xi : \Omega \rightarrow \mathbb{R}$ such that $\{\omega | \xi(\omega) \leq x\} \in \mathcal{A}$ for all finite $x$.

• $\xi$ has a *cumulative distribution* given by $F_\xi(x) = P(\xi \leq x)$.

• *Discrete random variables* are those that take on a finite number of values $\xi^k, k \in K$.

• Random variables have an associated *probability density function*.

• For a discrete random variable the density function $f(\xi^k) \equiv P(\xi = \xi^k)$

• A continuous random variables has density $f$ with the property

$$P(a \leq \xi \leq b) = \int_a^b f(\xi) d\xi$$

$$= \int_a^b dF(\xi)$$

$$= F(b) - F(a)$$
Expectation and Variance

• The *Expected value* of $\xi$ is
  
  - $E(\xi) = \sum_{k \in K} \xi^k f(\xi^k)$ (Discrete)
  - $E(\xi) = \int_{-\infty}^{\infty} f(\xi) d\xi = \int_{-\infty}^{\infty} dF(\xi)$. (Continuous)

• *Variance* of $\xi$ is $\text{Var}(\xi) = E(\xi - E(\xi)^2)$. 