Introduction to Optimization

1 Basic Optimization

Optimization is the selection of the best configuration of variables from a set of feasible configurations. More formally, given a **feasible set** of configurations $\mathcal{X}$, we select the value of the **decision variable** $\boldsymbol{x} \in \mathcal{X}$ such that the **objective function** $f(\boldsymbol{x})$ is minimized. We write an optimization problem as:

1.1 Continuous Optimization

In **continuous optimization**, we optimize a continuous objective function $f(\boldsymbol{x})$ with $\mathcal{X}$ a connected subset of Euclidean space. Thus, we encode $\mathcal{X}$ as a set of inequality constraints on $\boldsymbol{x}$:

$$\begin{aligned}&\min f(\boldsymbol{x})\\ &\text{s.t. } g_i(\boldsymbol{x}) \leq b_i \forall i \in C\end{aligned}$$

Continuous optimization problems are **NP-**hard in general, but with convexity constraints introduced they are solvable in polynomial time.

Definition 1. A **convex set** is a set $\mathcal{X}$ with $\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y} \in \mathcal{X}$ $\forall \alpha \in [0,1]; \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$. That is, the line connecting any two points in the set is also in the set.

Definition 2. A **convex function** is a function $f : \mathcal{X} \to \mathbb{R}$ such that $f(\alpha\boldsymbol{x}+(1-\alpha)\boldsymbol{y}) \leq \alpha f(\boldsymbol{x})+(1-\alpha)f(\boldsymbol{y})$ $\forall \alpha \in [0,1]; \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$. That is, the value of a function between two points is always at or below the line created by those two points.

If $f(\boldsymbol{x})$ is a convex function and $\mathcal{X}$ is a convex set, we have a **convex optimization** problem.

Example 3. A common convex optimization problem is **least squares regression**: Given a dataset $(\boldsymbol{a_i}, b_i)\}_{i=1}^m$ with $a_i \in \mathbb{R}^n$ and $b_i \in \mathbb{R}$ $\forall i$, find the linear function $f : \mathbb{R}^n \to \mathbb{R}$ with $f(\boldsymbol{a}) = \boldsymbol{x}^\intercal \boldsymbol{a}, \boldsymbol{x} \in \mathbb{R}^n$ which minimizes the least squares error. This problem can be simply written as:

$$\min \|\mathbf{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2$$

Example 4. Another convex optimization problem is **min-cost flow**: Given a graph $G = (V, E)$ with a cost $c_e \in \mathbb{R}^+$ and capacity $d_e \in \mathbb{R}^+$ on each edge e, find the minimum cost routing of r units of flow from the source s to the sink t. As opposed to the last example, here we must utilize inequality constraints.

$$\begin{aligned}\min\ & \sum_{e\in E} c_e x_e\\ \text{s.t.}\ & \sum_{e\leftarrow v} x_e = \sum_{e\to v} x_e, v \in V \setminus \{s,t\}\\ & \sum_{e\leftarrow s} x_e = r\\ & x_e \leq d_e\\ & x_e \geq 0\end{aligned}$$

This is a linear optimization which generalizes to a convex optimization if c_e is dependent upon x_e.

1.2 Combinatorial Optimization

In **combinatorial optimization**, we optimize over a finite feasible set $\mathcal{X}$.

Example 5. A common combinatorial optimization problem is **shortest path**: Given a graph $G = (V, E)$ with a cost $c_e \in \mathbb{R}^+$ on each edge e, find the minimum cost path from s to t.

Example 6. Another combinatorial optimization problem is the **traveling salesman problem**. Given V with metric $d(u, v)$ between $u, v \in V$, find the minimum length tour.

Oftentimes, it is useful to encode a continuous optimization problem as a combinatorial optimization problem, and vice versa. For example, shortest path can be encoded as a minimum cost flow problem with unit capacities and $r = 1$.

2 Linear Programming

2.1 Introduction to Linear Programming

2.1.1 Standard Form

Linear programming (LP) is a method to encode linear optimization problems in the form:

$$\begin{aligned} &\max \boldsymbol{c}^\intercal \boldsymbol{x} \\ &\text{s.t. } \boldsymbol{a_i}^\intercal \boldsymbol{x} \leq b_i, i \in C_1 \\ &\quad\;\; \boldsymbol{a_i}^\intercal \boldsymbol{x} = b_i, i \in C_2 \\ &\quad\;\; \boldsymbol{a_i}^\intercal \boldsymbol{x} \geq b_i, i \in C_3 \end{aligned}$$

It is advantageous to write LPs in standard form:

$$\begin{aligned} &\max \boldsymbol{c}^\intercal \boldsymbol{x} \\ &\text{s.t. } \boldsymbol{a_i}^\intercal \boldsymbol{x} \leq b_i, i \in 1, \ldots, m \\ &\quad\;\; x_j \geq 0, j \in 1, \ldots, n \end{aligned}$$

Note that an unconstrained variable x_j can be replaced by $x_j^+ - x_j^-$ where $x_j^+, x_j^- \geq 0$.

2.1.2 Geometric View

Each constraint in an LP is a halfspace in $\mathbb{R}^m$ whose intersection creates the feasible region. We wish to move as far as possible in the direction of $\boldsymbol{c}$ while staying in the feasible region. Rewriting an LP in standard form places the feasible region in the 1st hyperquadrant.

2.1.3 Applications

A fundamental application of LP is the **optimal production** problem. Given n products and m materials, where one unit of product j uses a_{ij} of material i. There are b_i units of material i available, and product j yields profit c_j per unit. How many of each product should be produced to maximize profit?

$$\begin{aligned} &\max \boldsymbol{c}^\intercal \boldsymbol{x} \\ &\text{s.t. } \boldsymbol{a_i}^\intercal \boldsymbol{x} \leq b_i, i = 1, \ldots, m \\ &\quad\;\; x_j \geq 0, j = 1, \ldots, n \end{aligned}$$

2.1.4 LP Facts

1. Feasible regions of LPs (i.e., polyhedrons) are convex.
2. Sets of optimal solutions of LPs are convex. In fact, the set of optimal solutions of an LP is a face of the polyhedron created by the feasible region. This face is the intersection of P with the hyperplane $\boldsymbol{c}^{\mathsf{T}}\boldsymbol{x} = OPT$.
3. At any vertex, n linearly independent constraints are satisfied with equality (i.e., are tight).
4. An LP either a) has an optimal solution, b) is unbounded, or c) is infeasible.

2.2 The Fundamental Theorem of Linear Programming

Theorem 7. *If an LP in standard form has an optimal solution, then it has a vertex optimal solution.*

Proof. For the sake of contradiction, suppose such an LP does not have a vertex optimal solution. Choose a non-vertex optimal solution $\boldsymbol{x}$ with the maximum number of tight constraints. Because this solution is not a vertex, $\exists \boldsymbol{y} \neq 0$ such that $\boldsymbol{x} \pm \boldsymbol{y}$ are both feasible. $\boldsymbol{y}$ is perpendicular to $\boldsymbol{c}$ because $\boldsymbol{x}$ is optimal and perpendicular to all tight constraints because if not, $\boldsymbol{x} \pm \boldsymbol{y}$ would not both be feasible.

Choose $\boldsymbol{y}$ such that $y_j < 0$ for some j. Let α be the largest constant such that $\boldsymbol{x} + \alpha\boldsymbol{y}$ is feasible. Such an α exists because either $x_j < 0$ or we exit the feasible region. Thus, at the point $\boldsymbol{x} + \alpha\boldsymbol{y}$, an additional constraint becomes tight, a contradiction. $\square$

Corollary 8. *If an LP in standard form has an optimal solution, then there is an optimal solution with at most m nonzero variables. For example, in the optimal production problem, there exists an optimal plan with at most m products.*