Linear Programming

Jack Yang

December 13, 2020

Linear Programming

A *linear programming problem* may be defined as the problem of *maximizing or minimizing a linear function subject to linear constraints*. The constraints maybe equalities or inequalities.

Linear programs are problems that can be expressed in standard matrix form as

$$\begin{aligned} \text{Manimize} \quad & c^T x \\ \text{s.t.} \quad & Ax \leq b \\ \text{and} \quad & x \geq 0 \end{aligned}$$

Here we assume that the matrix A has a full row rank.

Motivation. Is LP Solvable in Polynomial Time?

Last time: Simplex Algorithm works well in general, but suffers from exponential-time computational complexity.

Klee-Minty example shows simplex method may have to visit every vertex to reach the optimal one.

Simplex operations only check adjacent extreme points, but it takes many iterations.

Klee-Minty: Worst Case Analysis

Reference: Victor Klee, George J. Minty, "How good is the simplex algorithm?" in (O. Shisha edited) Inequalities, Vol. III (1972), pp. 159-175.

In 2-Dimensional

$$\begin{aligned} \min \quad & -x_2 \\ \text{s.t.} \quad & x_1, x_2 \geq 0 \\ & x_1 \leq 1 \\ & x_2 \geq \varepsilon \cdot x_1 \qquad 0 < \varepsilon < \frac{1}{2} \\ & x_2 \leq 1 - \varepsilon \cdot x_1 \end{aligned}$$

Generalization in d-Dimension

$$
\begin{aligned}
\min \quad & -x_d \\
\text{s.t.} \quad & x_i \geq 0 \text{ for all } 1 \leq i \leq d \\
& x_1 \leq 1 \\
& x_2 \geq \varepsilon \cdot x_1 \\
& x_2 \leq 1 - \varepsilon \cdot x_1 \\
& \quad \vdots \\
& x_d \geq \varepsilon \cdot x_{d-1} \\
& x_d \leq 1 - \varepsilon \cdot x_{d-1}
\end{aligned}
$$

There are $2^d - 1$ iterations. Hence, in theory, the simplex algorithm is not a polynomial-time algorithm. It is an *exponential-time algorithm*.

Interior Point Method

The basic idea is to approach optimal solutions from the interior of the feasible domain.

Compare with Simplex Algorithm, Interior Point Method takes more complicated operations in each iteration to find a better moving direction, while iterates fewer times.

Interior Point

Here are two criteria for a point x to be an interior feasible solution.

- $Ax = b$: every linear constraint is satisfied
- $x > 0$: every component is positive

Optimal Solution

A current feasible solution is optimal if and only if "no feasible direction at this point is a good direction."

A **feasible direction** at a current solution is a direction that allows it to take a small movement while staying to be interior feasible.

A **good direction** at a current solution is a direction that leads it to a new solution with a lower objective value.

Feasible Direction Descent

The process is somewhat like gradient descent algorithm.

1. Determine a starting point $x_0 \in \mathbb{R}^d$ such that $x_0 \in X$ where X is a polyhedral set. Set $k = 0$.
2. Determine a search direction $p_k \in \mathbb{R}^d$ such that p_k is a feasible descent direction.
3. Determine a step length $\alpha_k > 0$ such that $f(x_k + \alpha_k p_k) < f(x_k)$ and $x_k + \alpha_k p_k \in X$.
4. Let $x_{k+1} = x_k + \alpha_k p_k$.
5. If a termination criterion is fulfilled, then stop. Otherwise, let $k := k + 1$ and go to Step 2.

Observations from Feasible Direction

$$x_{k+1} = x_k + \alpha_k p_k \quad Ax_k = b, x_k > 0$$

There is no problem to stay interior if the step-length is small enough.

To maintain feasiblity, we need

$$\begin{aligned} Ax_{k+1} &= b \\ Ax_k + \alpha_k Ap_k &= b \end{aligned}$$

So, $Ap_k = 0$, which means $p_k \in \mathcal{N}(A)$, the null space of A.

Observations from Good Direction

$$c^T x_{k+1} \leq c^T x_k$$
$$c^T x_k + \alpha_k c^T p_k \leq c^T x_k$$

Then,

$$c^T p_k \leq 0.$$

To check optimality, recall that "no feasible direction at this point is a good direction."

At a current solution, we check no $p_k \in \mathbb{R}^d$ with $Ap_k = 0$ can make $c^T p_k \leq 0$.

Neyman-Pearson Lemma

Theorem (Neyman-Pearson)

Let $Y_1, \ldots, Y_n$ have a joint distribution depending on the parameters ϑ and consider testing $H_0 : \vartheta = \vartheta_0$ against $H_1 : \vartheta = \vartheta_1$. Then the test with the highest power among all tests with significance level α has rejection region

$$\mathcal{R} = \{y : \lambda(y) \leq a\}$$

where $\lambda(y) = \dfrac{L(\vartheta_0 \mid y)}{L(\vartheta_1 \mid y)}$ is the likelihood ratio and a is a constant determined by α, that is a such that $P(\lambda(Y) \leq a \mid H_0) = \alpha$.

Application: Uniformly Most Powerful Test

Suppose that $Y_1, Y_2, \ldots, Y_n \sim \text{Poisson}(\lambda)$ are independent. We want to test

$$H_0 : \lambda = \lambda_0 \text{ vs } H_1 : \lambda = \lambda_1$$

where $\lambda_1 > \lambda_0$. We reject H_0 if

$$\ln\left(\frac{\lambda_0}{\lambda_1}\right) \cdot \sum_{i=1}^{n} y_i \leq c$$

where c is a constant that is chosen to give a significance level α. It follows that the rejection region is of the form

$$\mathcal{R} = \{y : \bar{y} \geq d\}$$

where d is a constant that is chosen to give a significance level α.