Optimization Theory and Algorithm	Lecture 7 - 10/08/2021
Lecture 7	
Lecturer: Haishan Ye	*Scribe: Haishan Ye*

Edited by: Junbo Hao

1 Simplex Method for Linear Programming

1. Fundamental Theorem of Linear Programming

2. Simplex Method

Recall the standard form of linear programming:

$$\min_{\mathbf{x}} \ \mathbf{c}^\top \mathbf{x},$$
$$s.t. \ \ A\mathbf{x} = \mathbf{b},$$
$$\mathbf{x} \succeq 0,$$

where $A \in^{m\times n}$ with $m \leqslant n$ is of full row rank, **c** and **x** are n-dimension vector.

Theorem 1.1. *A linear programming whose feasible domain is not not empty. Then its optimum is either unbounded or attained at least one vertex of the feasible domain.*

Definition 1.2. Feasible domain: $P = \{\mathbf{x} \in^n \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geqslant 0\}$.

Definition 1.3. Hyperplane: $a^\top x = \beta$, with $a, x \in^{n\times 1}$. Closed Half space: $a^\top x \leqslant \beta$. Polyhedral: Intersection of a finite number of closed half space. Polytope: Bounded polyhedral.

Note that the feasible domain P of the linear programming is a polyhedral because $a_i^\top x = b_i$ is the intersection of $a_i^\top x \geqslant b_i$ and $a_i^\top x \leqslant b_i$ and $x_i \geqslant 0$ is a closed half space.

Furthermore, P is convex since closed half space is convex and the intersection still reserves the convexity.

Definition 1.4. Extreme point of P is the point that can not be expressed by the convex combination of other points.

Theorem 1.5. *P is convex polyhedral and $x \in P$ is a vertex if and only if x is a extreme point of P.*

Theorem 1.6. *$x \in P$ is a extreme point of P if and only if columns of A with respect to positive x_i are linearly independent.*

Proof. Denote that

$$\mathbf{x} = \begin{bmatrix} \bar{\mathbf{x}} \\ 0 \end{bmatrix} \text{ with } \bar{\mathbf{x}} = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} > 0, \text{ and } \bar{A} = [A_1, \dots, A_p]. \tag{1}$$

It is easy to check that $A\mathbf{x} = \bar{A}\bar{\mathbf{x}} = \mathbf{b}$.

Proof by contradiction. Assume that $\mathbf{x}$ is an extreme point but $\bar{A}$ is linearly dependent. Since $\bar{A}$ is linearly dependent, there exist a $\bar{\mathbf{w}} \neq 0$ such that $\bar{A}\mathbf{w} = 0$. Therefore, there exist a small number ϵ such that $\bar{\mathbf{x}} \pm \epsilon\bar{\mathbf{w}} \geqslant 0$ and $\bar{A}(\bar{\mathbf{x}} \pm \epsilon\bar{\mathbf{w}}) = \bar{A}\bar{\mathbf{x}} = \mathbf{b}$. Letting

$$\mathbf{y}_1 = \begin{bmatrix} \bar{\mathbf{x}} + \epsilon\bar{\mathbf{w}} \\ 0 \end{bmatrix}, \text{ and } \mathbf{y}_2 = \begin{bmatrix} \bar{\mathbf{x}} - \epsilon\bar{\mathbf{w}} \\ 0 \end{bmatrix}.$$

It is easy to check that $\mathbf{x} = \frac{\mathbf{y}_1 + \mathbf{y}_2}{2}$ and $y_1, y_2 \in P$. That is $\mathbf{x}$ can be expressed by the convex combination of $\mathbf{y}_1$ and $\mathbf{y}_2$, which contradicts with the fact that $\mathbf{x}$ is an extreme point of P.

Now we assume that $\bar{A}$ is linearly independent but $\mathbf{x}$ is not an extreme point of P. Then we can represent $\mathbf{x}$ as

$$\mathbf{x} = \lambda\mathbf{y}_1 + (1-\lambda)\mathbf{y}_2, \ \mathbf{y}_1 \neq \mathbf{y}_2 \ \lambda \in (0,1), \ \mathbf{y}_1, \mathbf{y}_2 \geqslant 0.$$

By the form of $\mathbf{x}$ shown in Eqn. (1), it holds that

$$\mathbf{y}_1 = \begin{bmatrix} \bar{\mathbf{y}}_1 \\ 0 \end{bmatrix}. \tag{2}$$

Now,

$$\mathbf{x} - \mathbf{y}_1 = \lambda\mathbf{y}_1 + (1-\lambda)\mathbf{y}_2 - \mathbf{y}_1 = -(1-\lambda)(\mathbf{y}_1 - \mathbf{y}_2) \neq 0 \tag{3}$$

where the last inequality is because $\mathbf{y}_1 \neq \mathbf{y}_2$ and $\lambda < 1$. Therefore

$$A(\mathbf{x} - \mathbf{y}_1) = \bar{A}(\bar{\mathbf{x}} - \bar{\mathbf{y}}_1) = \mathbf{b} - \mathbf{b} = 0,$$

which contradicts the assumption A is linearly independent. ■

Managing extreme points algebraically Let A be an $m \times n$ matrix with, we say A has full rank (full row rank) if A has m linearly independent columns. In this, we can rearrange

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix} \begin{matrix} \leftarrow \text{basic variables} \\ \leftarrow \text{non-basic variables} \end{matrix} \quad A = [\underbrace{B}_{\text{Basis}}, \underbrace{N}_{\text{non-basis}}]. \tag{4}$$

Definition 1.7. If we set $\mathbf{x}_N$ to zero and $\mathbf{x}_B$ is the solution of $B\mathbf{x}_B = b$, then we say $\mathbf{x}$ is a basic solution. If $\mathbf{x}_B \geqslant 0$, then $\mathbf{x}$ is a basic feasible solution.

Proposition 1.8. *A point* $\mathbf{x}$ *in* P *is an extreme point of* P *if and only if* $\mathbf{x}$ *is a basic feasible solution corresponding to some basis* B*.*

Proposition 1.9. *The polyhedron* P *has only a finite number of extreme points.*

Definition 1.10. A vector $\mathbf{d}$ is an extremal direction of P, if $\{\mathbf{x} \in^n | \ \mathbf{x} = \mathbf{x}^0 + \lambda \mathbf{d}, \lambda > 0\} \subset P$ for all $\mathbf{x}^0 \in P$.

Theorem 1.11 (Resolution Theorem)**.** *Let* $V = \left\{v^i \in^n | \ i \in I\right\}$ *be the set of all extreme point of* P *and* I *is a finite index set. Then* $\forall \mathbf{x} \in P$*, we have*

$$\mathbf{x} = \sum_{i \in I} \lambda_i v^i + \lambda \mathbf{d}, \tag{5}$$

where

$$\sum_i \lambda_i = 1, \lambda_i \geqslant 0,$$

and either $\mathbf{d} = 0$ *or* $\mathbf{d}$ *is a extreme direction.*

Theorem 1.12. *For a standard form LP, if its feasible domain* P *is nonempty, then the optimal objective value of* $z = \mathbf{c}^\top \mathbf{x}$ *over* P *is either unbounded below, or it is attained at (at least) an extreme point of* P*.*

Proof. By the resolution theorem, there are two cases:

Case 1, P has an extreme direction $\mathbf{d}$ such that $\mathbf{c}^\top \mathbf{d} < 0$. Then P is unbounded and $z \to -\infty$.

Case2, P does not have an extreme direction $\mathbf{d}$ such that $\mathbf{c}^\top \mathbf{d} < 0$. Then $\forall \mathbf{x} \in P$, either $\mathbf{x} = \sum_i \lambda_i v^i$ or $\mathbf{x} = \sum_i \lambda_i v^i + \bar{\mathbf{d}}$ with $\mathbf{c}^\top \bar{\mathbf{d}} \geqslant 0$.

In both cases, it holds that

$$\begin{aligned}
\mathbf{c}^\top \mathbf{x} =& \mathbf{c}^\top \left(\sum_i \lambda_i v^i \right) + \mathbf{c}^\top \bar{\mathbf{d}} \\
\geqslant& \sum_i \lambda_i (\mathbf{c}^\top v^i) \\
\geqslant& \min_i \mathbf{c}^\top v^i \\
=& \mathbf{c}^\top v^{\min}
\end{aligned}$$

■

Simplex Method Fundamental Matrix

$$M = \begin{bmatrix} B & N \\ 0 & I \end{bmatrix} \text{ and } M^{-1} = \begin{bmatrix} B^{-1} & -B^{-1}N \\ 0 & I \end{bmatrix} \tag{6}$$

It is easy to check that

$$M\mathbf{x} = \begin{bmatrix} B & N \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix} \tag{7}$$

$$\mathbf{x}(\lambda) = \mathbf{x} + \lambda \mathbf{d}_q, \tag{8}$$

with

$$\mathbf{d}_q = \begin{bmatrix} -B^{-1}A_q \\ \vdots \\ e_q \end{bmatrix} \tag{9}$$

Feasibility of $\mathbf{d}_q$? *Yes*! By

$$A\mathbf{x}(\lambda) = A(\mathbf{x} + \lambda \mathbf{d}_q) = A\mathbf{x} + [B, N] \begin{bmatrix} -B^{-1}N_q \\ e_q \end{bmatrix} = A\mathbf{x} = \mathbf{b}. \tag{10}$$

Definition 1.13 (reduced cost)**.** The quantity of $r_q = \mathbf{c}^\top \mathbf{d}_q = \mathbf{c}_q - \mathbf{c}_B^\top B^{-1} A_q$ is called a reduced cost with respect to the variable $\mathbf{x}_q$.

Theorem 1.14. *If $\mathbf{x} = [B^{-1}b; 0]$ is a basic feasible solution with B and $r_q < 0$, for some non-basic variable x_q, then $\mathbf{d}_q = [-B^{-1}A_q; e_q]$ leads to an improved objective function.*

Theorem 1.15. *If $\mathbf{x}$ is a basic feasible solution with $r_q \geqslant 0$ for all non-basic variables, then $\mathbf{x}$ is optimal solution.*

Proof. $\mathbf{x}$ is local optimum. Since linear programming is a convex optimization problem, the local optimum is the global one. ∎

How to choose step size λ Case 1: $\mathbf{d}_q \geqslant 0$, for all $\lambda > 0$.

Case 2: One $\mathbf{d}_q < 0$, $\lambda = \min_i \left\{ \frac{\mathbf{x}_i}{-\mathbf{d}_{q_i}} \mid \mathbf{d}_{q_i} < 0 \right\}$