Lecture #5 Notes Summary

Basic notions of Linear Programming, Duality, and Algorithms

Introduction

Definition 1 (Linear Programming Problem)**.** A *Linear Programming Problem* (or simply a Linear Program, abbreviated LP) is an optimization problem of the form

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & \forall i \in [m],\ g_i(x) \leq 0, \end{aligned} \tag{1}$$

where $f, g_1, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ are affine functions.

Note that maximization problems can also be handled as LP, since $\max_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) = -\min_{\boldsymbol{x}\in\mathcal{X}} -f(\boldsymbol{x})$.

Definition 2 (Feasibility, Boundedness)**.**

- A vector $\boldsymbol{x} \in \mathbb{R}^n$ is called a *feasible solution* for Problem (1) if $\forall i \in [m],\ g_i(\boldsymbol{x}) \leq 0$.
- Problem (1) is *infeasible* if the feasibility set $\{\boldsymbol{x} \in \mathbb{R}^n :\ \forall i \in [m],\ g_i(\boldsymbol{x}) \leq 0\}$ is empty. In this case we say that the value of Problem (1) is $+\infty$.
- Problem (1) is *unbounded* if there exists no $M \in \mathbb{R}$ such that $f(\boldsymbol{x}) \geq M$ for all feasible $\boldsymbol{x}$.

Proposition 1. *Let $\boldsymbol{x}^*$ be a* local optimum *of Problem* (1) *(i.e. $\boldsymbol{x}^*$ is feasible and there is a ball $\mathcal{B}$ centered at $\boldsymbol{x}^*$ such that for all feasible $\boldsymbol{x} \in \mathcal{B}$, $f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$). Then, $\boldsymbol{x}^*$ is an optimal solution of Problem* (1) *(i.e., for all feasible $\boldsymbol{x}$ we have $f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$, so $\boldsymbol{x}^*$ is also a* global optimum*).*

Sketch of the proof. Assume the contrary. Let $\boldsymbol{z}^*$ be feasible and such that $f(\boldsymbol{z}^*) < f(\boldsymbol{x}^*)$. The function f is linear on the segment $[\boldsymbol{x}^*, \boldsymbol{z}^*]$, which is included in the feasibility region by convexity. So $f(\boldsymbol{x}^*+\epsilon(\boldsymbol{z}^*-\boldsymbol{x}^*)) = (1-\epsilon)f(\boldsymbol{x}^*) + \epsilon f(\boldsymbol{z}^*) < f(\boldsymbol{x}^*)$ for all $\epsilon > 0$ contradicts the local optimality of $\boldsymbol{x}^*$. □

Definition 3 (Canonical/Standard form)**.** A LP is in *canonical form* (resp. *standard form*) if it is written as:

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} \quad & A\boldsymbol{x} \geq \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}. \end{aligned} \qquad \Bigg(\text{resp.} \qquad \begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} \quad & A\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}. \end{aligned} \Bigg),$$

where the rows of the matrix $A \in \mathbb{R}^{m\times n}$ are independent (rank $A = m$). The vector $\boldsymbol{c} \in \mathbb{R}^n$ is the cost vector, and $\boldsymbol{b} \in \mathbb{R}^m$ is the right-hand-side vector. The inequalities are understood componentwise, i.e. $\forall i \in [m],\ (A\boldsymbol{x})_i \geq b_i$, and $\forall j \in [n], x_j \geq 0$.

Proposition 2. *Any LP can be put into canonical form (resp. standard form).*

Sketch of the proof. The reduction to canonical/standard form relies on the following two techniques:

- Inequalities can be transformed into equalities by *introducing slack variables*

$$A\boldsymbol{x} \leq \boldsymbol{b} \iff A\boldsymbol{x} + \boldsymbol{s} = \boldsymbol{b} \text{ for some } \boldsymbol{s} \geq \boldsymbol{0}.$$

- Any unconstrained variables x can be expressed as the difference $x^+ - x^-$ of two nonnegative variables $(x^+ \geq 0, x^- \geq 0)$.

□

Note that the feasibility space of any LP is of the form $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} \leq \boldsymbol{b}\}$, i.e. $\mathcal{P}$ is the intersection of the halfspaces $(A\boldsymbol{x})_i \leq b_i$. We say that $\mathcal{P}$ is a *convex polyhedron.* A fundamental result from polyhedral geometry gives an alternative representation of $\mathcal{P}$ based on its extreme points and extreme rays (we won't prove this result here).

Definition 4 (extreme points)**.** A point $\boldsymbol{x}$ of a convex polyhedron $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} \leq \boldsymbol{b}\}$ is called a *vertex* (or an *extreme point*) if $\boldsymbol{x} = \frac{1}{2}\boldsymbol{x_1} + \frac{1}{2}\boldsymbol{x_2}$ for some $\boldsymbol{x_1}, \boldsymbol{x_2} \in \mathcal{P}$ implies $\boldsymbol{x} = \boldsymbol{x_1} = \boldsymbol{x_2}$.

Theorem 3 (affine Minkowski-Weyl theorem)**.** *Let $\mathcal{P}$ be a nonempty subset of $\mathbb{R}^n$. The following statements are equivalent:*

(i) $\mathcal{P}$ is a convex polyhedron ($\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} \leq \boldsymbol{b}\}$ for some $(A, \boldsymbol{b}) \in \mathbb{R}^{m\times n} \times \mathbb{R}^m$)

(ii) $\mathcal{P}$ is generated its (finite) set of extreme points $\boldsymbol{a_1}, \ldots, \boldsymbol{a_p} \in \mathbb{R}^n$ and extreme rays $\boldsymbol{c_1}, \ldots, \boldsymbol{c_q} \in \mathbb{R}^n$:

$$\mathcal{P} = \operatorname{conv}(\boldsymbol{a_1}, \ldots, \boldsymbol{a_p}) + \operatorname{cone}(\boldsymbol{c_1}, \ldots, \boldsymbol{c_q}) = \Big\{ \sum_{i=1}^{p} \alpha_i \boldsymbol{a_i} + \sum_{j=1}^{q} \lambda_j \boldsymbol{c_j} : \ \boldsymbol{\alpha} \geq \boldsymbol{0}, \sum_i \alpha_i = 1, \ \boldsymbol{\lambda} \geq \boldsymbol{0} \Big\}.$$

Definition 5 (basic feasible solution)**.** Consider a LP in standard form (with equality constraints). A feasible vector $\boldsymbol{x}$ is called a *basic feasible solution* (bfs) if there is a set $I \subseteq [n]$ of cardinality m such that $x_i = 0$ for all $i \notin I$ and the $m \times m$ matrix A_I is invertible (where A_I denotes the submatrix of A formed with the columns indexed in I).

Note that every choice of I uniquely defines the nonzeros of $\boldsymbol{x}$, i.e. $\boldsymbol{x}_I = A_I^{-1}\boldsymbol{b}$, so I is a valid subset defining a bfs iff $\boldsymbol{x}_I \geq \boldsymbol{0}$. Geometrically, bfs correspond to extreme points of the feasibility polyhedron:

Theorem 4. *Let $\mathcal{P} = \{\boldsymbol{x} \in \mathbb{R}^n : \ \boldsymbol{x} \geq \boldsymbol{0}, A\boldsymbol{x} = \boldsymbol{b}\}$ be the feasibility polyhedron of a LP in standard form. Then, the set of bfs of the LP coincides with the set of vertices of $\mathcal{P}$.*

Proof. cf. Exercise 3 of the Worksheet #5. □

Theorem 5 (Fundamental theorem of Linear Programming)**.** *Every LP has the following properties:*

(i) If it has no optimal solution, then it is either infeasible or unbounded.

(ii) If it is feasible, then there exists a bfs.

(iii) If it has an optimal solution (i.e. the minimum is attained), then it has an optimal bfs.

Proof. Consider an arbitrary LP of the form (1), and assume that it is feasible. Its feasibility polyhedron $\mathcal{P}$ is nonempty, and by Minkowski-Weyl we have $\mathcal{P} = \left\{ \sum_{i=1}^p \alpha_i \boldsymbol{a_i} + \sum_{j=1}^q \lambda_j \boldsymbol{c_j} : \ \boldsymbol{\alpha} \geq \boldsymbol{0}, \sum_i \alpha_i = 1, \ \boldsymbol{\lambda} \geq \boldsymbol{0} \right\}$, where the $\boldsymbol{a_i}$ are the vertices of $\mathcal{P}$ and the $\boldsymbol{c_j}$ are the extreme rays. Every $\boldsymbol{a_i}$ is feasible, and so by Theorem 4 we have the existence of a bfs, which proves (ii). Now, assume that the problem is bounded. Clearly, we must have $f(\boldsymbol{c_j}) \geq 0$ for all $j \in [q]$ (otherwise we see that the problem is unbounded by letting $\lambda_j \to \infty$). For a feasible $\boldsymbol{x} = \sum_{i=1}^p \alpha_i \boldsymbol{a_i} + \sum_{j=1}^q \lambda_j \boldsymbol{c_j}$ we have $f(\boldsymbol{x}) = \sum_{i=1}^p \alpha_i f(\boldsymbol{a_i}) + \sum_{j=1}^q \lambda_j f(\boldsymbol{c_j})$ (linearity of f). So decreasing the value of a λ_j can only decrease the value of $f(\boldsymbol{x})$, and the optimal value of Problem (1) is $\inf_{\boldsymbol{x} \in \mathcal{P}} f(\boldsymbol{x}) = \inf_{\boldsymbol{\alpha} \in \Delta_p} \sum_{i=1}^p \alpha_i f(\boldsymbol{a_i})$, where Δ_p is the standard simplex of $\mathbb{R}^p$. This is a minimization problem for a continuous function over a compact set, so the minimum is attained, which proves (i). Moreover if $f(\boldsymbol{a_i}) = \max_{i' \in [p]} f(\boldsymbol{a_{i'}})$ we know that a solution of this problem consists in putting all the *weight* on the i^{th} coordinate, i.e. $\alpha_i = 1$ and $j \neq i \Rightarrow \alpha_j = 0$, which gives the optimal bfs $\boldsymbol{x^*} = \boldsymbol{a_i}$ and (iii) is proved. □

Duality

We start with an informal derivation of the dual of a LP in standard form: define the Lagrangian $\mathcal{L}(\boldsymbol{x}, \boldsymbol{y}) := \boldsymbol{c}^T \boldsymbol{x} + \boldsymbol{y}^T (\boldsymbol{b} - A\boldsymbol{x})$, and note that

$$\max_{\boldsymbol{y} \in \mathbb{R}^m} \ \mathcal{L}(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} \boldsymbol{c}^T \boldsymbol{x} & \text{if } A\boldsymbol{x} = \boldsymbol{b} \\ +\infty & \text{otherwise.} \end{cases}$$

So we can rewrite the standard LP problem as follows:

$$\min\{\boldsymbol{c}^T \boldsymbol{x} | A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \in \mathbb{R}^n_+\} = \min_{\boldsymbol{x} \geq \boldsymbol{0}} \ \max_{\boldsymbol{y} \in \mathbb{R}^m} \ \mathcal{L}(\boldsymbol{x}, \boldsymbol{y}).$$

The *dual* of the LP is obtained by permuting the order of the min and the max. The minimization step with respect to $\boldsymbol{x} \geq \boldsymbol{0}$ yields:

$$\min_{\boldsymbol{x} \in \mathbb{R}^n_+} \ \mathcal{L}(\boldsymbol{x}, \boldsymbol{y}) = \min_{\boldsymbol{x} \geq \boldsymbol{0}} \ \boldsymbol{b}^T \boldsymbol{y} + \boldsymbol{x}^T (\boldsymbol{c} - A^T \boldsymbol{y}) = \begin{cases} \boldsymbol{b}^T \boldsymbol{y} & \text{if } \boldsymbol{c} - A^T \boldsymbol{y} \geq \boldsymbol{0} \\ -\infty & \text{otherwise.} \end{cases}$$

So, we can now define the dual of Problem (P): $\min\{\boldsymbol{c}^T \boldsymbol{x} | A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \in \mathbb{R}^n_+\}$:

Definition 6 (Dual of a LP)**.** The dual of a problem (P) written in standard form is

$$\begin{aligned} \max_{\boldsymbol{y} \in \mathbb{R}^m} \quad & \boldsymbol{b}^T \boldsymbol{y} \\ \text{s.t.} \quad & A^T \boldsymbol{y} \leq \boldsymbol{c} \end{aligned} \tag{D}$$

Proposition 6. *The dual of* (D) *is* (P).

Proof. cf. Exercise 5 of the Worksheet #5. □

Theorem 7 (The weak duality theorem)**.** *Let* $\boldsymbol{x}$ *be a feasible solution for* (P) *and* $\boldsymbol{y}$ *be a feasible solution for* (D)*. Then,* $\boldsymbol{b}^T \boldsymbol{y} \leq \boldsymbol{c}^T \boldsymbol{x}$*. In particular, we have* $\text{OPT}(D) \leq \text{OPT}(P)$.

Proof. $\boldsymbol{b}^T \boldsymbol{y} = \boldsymbol{x}^T A^T \boldsymbol{y} = \sum_i x_i (A^T \boldsymbol{y})_i \leq \sum_i x_i c_i = \boldsymbol{c}^T \boldsymbol{x}$, where the inequality follows from $\boldsymbol{x} \geq \boldsymbol{0}$ and $\boldsymbol{c} \geq A^T \boldsymbol{y}$. The inequality $\text{OPT}(D) \leq \text{OPT}(P)$ follows from an analysis of all finite/infinite values for $\text{OPT}(P)$ and $\text{OPT}(D)$ (for example, (P) unbounded $\Rightarrow \text{OPT}(P) = -\infty \Rightarrow$ (D) infeasible $\Rightarrow \text{OPT}(D) = -\infty$). □

But we have a much stronger result:

Theorem 8 (The strong duality theorem)**.** *Let P and D be a pair of primal and dual LPs. If P or D is feasible, then* $\mathrm{OPT}(P) = \mathrm{OPT}(D)$.

The proof of this result relies on an important lemma:

Lemma 9 (Farkas' Lemma)**.** *Let $A \in \mathbb{R}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$. Then, exactly one of these two statements is true:*

1. $\exists \boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \geq \boldsymbol{0},\ A\boldsymbol{x} = \boldsymbol{b}$;
2. $\exists \boldsymbol{y} \in \mathbb{R}^m : A^T\boldsymbol{y} \geq \boldsymbol{0},\ \boldsymbol{b}^T\boldsymbol{y} < 0$.

Proof. The proof of the lemma will be done in Exercise 4 of the worksheet #5, and we only prove the theorem here. The theorem says that one of the following 4 cases must occur: (i) (P) and (D) are infeasible; (ii) (P) is unbounded, (D) is infeasible, and $\mathrm{OPT}(P) = \mathrm{OPT}(D) = -\infty$; ($iii$) (P) is infeasible, (D) is unbounded, and $\mathrm{OPT}(P) = \mathrm{OPT}(D) = +\infty$; ($iv$) Both (P) and (D) have an optimal solution, and their optimal value coincide. The case (i) can indeed occur, but there is nothing to prove. The cases (ii) and (iii) were already proved in the weak duality theorem. So we must only prove the last (and most interesting) case.

Assume that Problem (D) is feasible and bounded (so it has an optimal solution). Let $\boldsymbol{y^*}$ be an optimal solution of Problem (D), and denote by I the set of indices where the inequality constraint is tight: $I = \{i \in [n] :\ \boldsymbol{a_i}^T\boldsymbol{y^*} = c_i\}$, where $\boldsymbol{a_i}$ stands for the i^{th} column of A. Denote by $\bar{A}$ the matrix formed by the columns $\{\boldsymbol{a_i} : i \in I\}$. We claim that $\boldsymbol{b} \in \mathrm{cone}(\{\boldsymbol{a_i} : i \in I\})$, i.e. $\boldsymbol{b} = \bar{A}\boldsymbol{\lambda}$ for some nonnegative vector $\boldsymbol{\lambda}$. Otherwise, by Farkas' lemma, there exists a vector $\boldsymbol{d}$ such that $\bar{A}^T\boldsymbol{d} \geq \boldsymbol{0},\ \boldsymbol{b}^T\boldsymbol{d} < 0$. But we will see that this contradicts the optimality of $\boldsymbol{y^*}$: let $\boldsymbol{y'} = \boldsymbol{y^*} - \epsilon\boldsymbol{d}$ for an $\boldsymbol{\epsilon} > 0$. For all $i \in I$ we have $\boldsymbol{a_i}^T\boldsymbol{y'} = c_i - \epsilon\boldsymbol{a_i}^T\boldsymbol{d} < c_i$, and if ϵ is sufficiently small we also have $\boldsymbol{a_i}^T\boldsymbol{y'} < c_i$ for all $i \notin I$ (this constraint was not tight so we don't violate it if ϵ is small enough). So $\boldsymbol{y'}$ is a feasible solution for (D) and $\boldsymbol{b}^T\boldsymbol{y'} = \boldsymbol{b}^T\boldsymbol{y^*} - \epsilon\boldsymbol{b}^T\boldsymbol{d} > \boldsymbol{b}^T\boldsymbol{y^*}$, a contradiction.

So, we can write $\boldsymbol{b} = \sum_{i\in I}\lambda_i\boldsymbol{a_i}$ for some nonnegative scalars λ_i. Define the vector $\boldsymbol{x} \in \mathbb{R}^n$ by setting $x_i = \lambda_i$ if $i \in I$, and $x_i = 0$ otherwise, so that $A\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \geq \boldsymbol{0}$, i.e., $\boldsymbol{x}$ is a feasible solution for (P). Finally, $\boldsymbol{c}^T\boldsymbol{x} = \sum_{i\in I} c_i x_i = \sum_{i\in I}(\boldsymbol{a_i}^T\boldsymbol{y^*})x_i = \boldsymbol{y^*}^T\big(\sum_{i\in I} x_i\boldsymbol{a_i}\big) = \boldsymbol{b}^T\boldsymbol{y^*}$. By the weak duality theorem, $\boldsymbol{x}$ is an optimal solution for (P). □

Theorem 10 (Complementary Slackness)**.** *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be feasible solutions for the pair of primal-dual problems (P) and (D) (in standard form as in Definition 6). The following statements are equivalent:*

(i) $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *are optimal for* (P) *and* (D)*;*

(ii) $\boldsymbol{c}^T\boldsymbol{x} = \boldsymbol{b}^T\boldsymbol{y}$*;*

(iii) $\boldsymbol{x}^T(\boldsymbol{c} - A^T\boldsymbol{y}) = 0$ $\Big($*which can be rewritten as* $\quad \forall i \in [n], \qquad (x_i = 0)\ \text{ OR }\ (A^T\boldsymbol{y})_i = c_i \quad\Big)$.

Proof. We know that (i) and (ii) are equivalent from the duality theorems. Now, since $\boldsymbol{x}$ is feasible we have $\boldsymbol{b}^T\boldsymbol{y} = \boldsymbol{x}^TA^T\boldsymbol{y}$, which is equal to $\boldsymbol{x}^T\boldsymbol{c}$ iff $\boldsymbol{x}^T(\boldsymbol{c} - A^T\boldsymbol{y}) = 0$, so (ii) $\iff$ (iii). □

Algorithms to solve LP

In this section we briefly review the three main families of algorithms to solve a LP.

The Simplex Algorithm

Very good in practice, returns an optimal bfs, but things can go badly in theory

- Start from a basic feasible solution $\boldsymbol{x^{(0)}}$
- At the k^{th} iteration, move from this vertex to an adjacent vertex of the feasibility polyhedron $\mathcal{P}$, by selecting an edge which is a *descent direction* for the objective $\boldsymbol{c}^T\boldsymbol{x}$. More precisely:
  - If there is no such descent direction, we have found a local (and hence global) optimum.
  - The operation of moving along an edge corresponds to choosing a variable x_j that was not in the basis of $\boldsymbol{x^{(k)}}$ (i.e., $x_j^{(k)} = 0$), and to increase it while decreasing the other basic variables.
  - A simple formula gives the step length to reach the vertex $\boldsymbol{x^{(k+1)}}$ along the chosen edge, or finds that the descent direction is in fact a ray of $\mathcal{P}$, i.e. the problem is unbounded.

This algorithm is very efficient in practice, although one can construct examples where the simplex algorithm visits all the vertices of $\mathcal{P}$ before reaching the solution (i.e., an exponential number of iterations is needed).

The Ellipsoid Algorithm

Useful in theory to obtain complexity bounds, but poor performance in practice

- Reduce to a strict feasibility problem: find $\boldsymbol{x} \in \mathcal{P} = \{\boldsymbol{x} : A\boldsymbol{x} < \boldsymbol{b}\}$
- Start with an ellipsoid $\mathcal{E}_0$ centered at $\boldsymbol{x^{(0)}}$ that contains $\mathcal{P}$
- At each step $k = 0, 1, \ldots, K$:
  - If $\boldsymbol{x^{(k)}} \in \mathcal{P}$, the algorithm stops and returns $\boldsymbol{x^{(k)}}$
  - Otherwise, find a constraint $\boldsymbol{a_i}^T\boldsymbol{x} \leq b_i$ violated by $\boldsymbol{x^{(k)}}$
  - Find the minimum-volume ellipsoid $\mathcal{E}_{k+1}$ (centered at $\boldsymbol{x^{(k+1)}}$) that contains the half-ellipsoid

$$\mathcal{E}_k \cap \{\boldsymbol{x} : \boldsymbol{a_i}^T\boldsymbol{x} \leq \boldsymbol{a_i}^T\boldsymbol{x^{(k)}}\}.$$

It can be shown that for rational-valued A and $\boldsymbol{b}$, there is a number K of iterations **polynomial in the input size** such that $\mathcal{P}$ is guaranteed to be empty if $\boldsymbol{x^{(K)}} \notin \mathcal{P}$, and the reduction of the LP to a strict feasibility problem can also be done with polynomial number of iterations.

Barrier Method

Strong algorithm in theory and in practice, it often beats the simplex for large scale problems

- Find a point $\boldsymbol{x^{(0)}}$ such that $A\boldsymbol{x^{(0)}} = \boldsymbol{b}$
- Choose a sequence $\beta_0, \beta_1, \ldots, \beta_k, \ldots$ of positive numbers decreasing to 0
- At each step $k = 0, 1, \ldots, K$, find an approximate solution $\boldsymbol{x^{(k+1)}}$ of the problem

$$\min_{\boldsymbol{x}} \left\{\boldsymbol{c}^T\boldsymbol{x} - \beta_k \sum_{j=1}^{n} \ln x_j \middle|\ A\boldsymbol{x} = \boldsymbol{b}\right\},$$

  by making one (or several) *Newton steps*, starting from $\boldsymbol{x^{(k)}}$.

It was shown that for some sequence $\beta_k = \beta_0/\alpha^k$ (with $\alpha > 1$), we have $\boldsymbol{c}^T\boldsymbol{x^{(K)}} \leq \mathrm{OPT}(P) + \epsilon$ after $K = O\big(\log \frac{n}{\epsilon}\big)$ iterations, and the iterations themselves are polynomial in the input size of the LP. This type of algorithm is harder to implement than the simplex, and it does not always provide a bfs.