CS 598CSC: Combinatorial Optimization	**Lecture date: January 21, 2009**
Instructor: Chandra Chekuri	**Scribe: Sungjin Im**

1 Polyhedra and Linear Programming

In this lecture, we will cover some basic material on the structure of polyhedra and linear programming. There is too abundant material on this topic to be covered in a few classes, so pointers will be given for further reading. For algorithmic and computational purposes one needs to work with rational polyhedra. Many basic results, however, are valid for both real and rational polyhedra. Therefore, to expedite our exploration, we will not make a distinction unless necessary.

1.1 Basics

Definition 1. *Let $x_1, x_2, ..., x_m$ be points in $\mathbb{R}^n$. Let $x = \sum_{i=1}^m \lambda_i x_i$, where each $\lambda_i \in \mathbb{R}$, $1 \le i \le m$ is a scalar. Then, x is said to be a(n)*

1. *Linear combination (of x_i, $1 \le i \le m$) for arbitrary scalars λ_i.*
2. Affine *combination if $\sum_i \lambda_i = 1$.*
3. Conical *combination if $\lambda_i \ge 0$.*
4. *Convex combination if $\sum \lambda_i = 1$ and $\lambda_i \ge 0$ (affine and also canonical).*

In the following definitions and propositions, unless otherwise stated, it will be assumed that $x_1, x_2, ..., x_m$ are points in $\mathbb{R}^n$ and $\lambda_1, \lambda_2, ..., \lambda_m$ are scalars in $\mathbb{R}$.

Definition 2. *$x_1, x_2, ..., x_m$ are said to be linearly independent if $\sum_{i=1}^m \lambda_i x_i = 0 \Rightarrow \forall i \in [m]\ \lambda_i = 0$.*

Definition 3. *$x_1, x_2, ..., x_m$ are said to be affinely independent if the vectors $(x_i - x_1), i = 2, ..., m$ are linearly independent, or equivalently if $\sum_{i=1}^m \lambda_i x_i = 0$ and $\sum_{i=1}^m \lambda_i = 0 \Rightarrow \forall i \in [m]\ \lambda_i = 0$.*

The following proposition is easy to check and the proof is left as an exercise to the reader.

Proposition 4. *$x_1, x_2, ..., x_m$ are affinely independent if and only if the vectors $\binom{x_i}{1}$, $i = 1, 2, ..., m$, are linearly independent in $\mathbb{R}^{m+1}$.*

A set $X \subseteq \mathbb{R}^n$ is said to be a(n) subspace [affine set, cone set, convex set] if it is closed under linear [affine, conical, convex] combinations. Note that an affine set is a translation of a subspace. Given $X \subseteq \mathbb{R}^n$, we let $\texttt{Span}(X), \texttt{Aff}(X), \texttt{Cone}(X)$, and $\texttt{Convex}(X)$ denote the closures of X under linear, affine, conical, and convex combinations, respectively. To get an intuitive feel of the above definitions, see Figure 1.

Definition 5. *Given a convex set $X \subseteq \mathbb{R}^n$, the affine dimension of X is the maximum number of affinely independent points in X.*

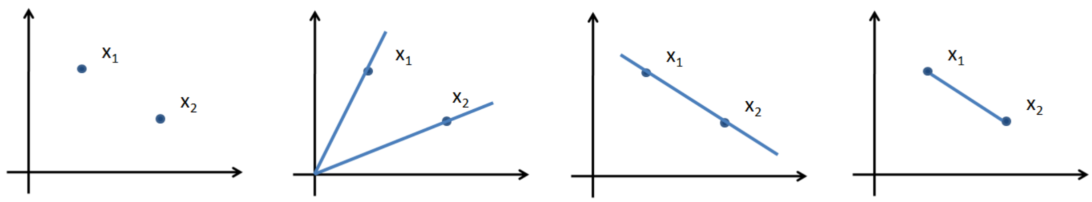


Figure 1: The subspace, cone set, affine set, and convex set of x_1, x_2 (from left to right). Note that the subspace is $\mathbb{R}^2$ and the cone set includes all points inside and on the two arrows.

1.2 Polyhedra, Polytopes, and Cones

Definition 6 (Hyperplane, Halfspace)**.** *A* hyperplane *in $\mathbb{R}^n$ is the set of all points $x \in \mathbb{R}^n$ that satisfy $a \cdot x = b$ for some $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$. A* halfspace *is the set of all points x such that $a \cdot x \leq b$ for some $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$.*

Definition 7 (Polyhedron)**.** *A* Polyhedron *in $\mathbb{R}^n$ is the intersection of finitely many halfspaces. It can be equivalently defined to be the set $\{x \mid Ax \leq b\}$ for a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $b \in \mathbb{R}^{m \times 1}$.*

Definition 8 (Polyhedral cone)**.** *A* polyhedral cone *is $\mathbb{R}^n$ the intersection of finitely many halfspaces that contain the origin, i.e. $\{x \mid Ax \leq 0\}$ for a matrix $A \in \mathbb{R}^{m \times n}$.*

Definition 9 (Polyotpe)**.** *A* polytope *is a* bounded *polyhedron.*

Note that a polyhedron is a convex and closed set. It would be illuminating to classify a polyhedron into the following four categories depending on how it looks.

1. Empty set (when the system $Ax \leq b$ is infeasible.)
2. Polytope (when the polyhedron is bounded.)
3. Cone
4. (Combination of) Cone and Polytope

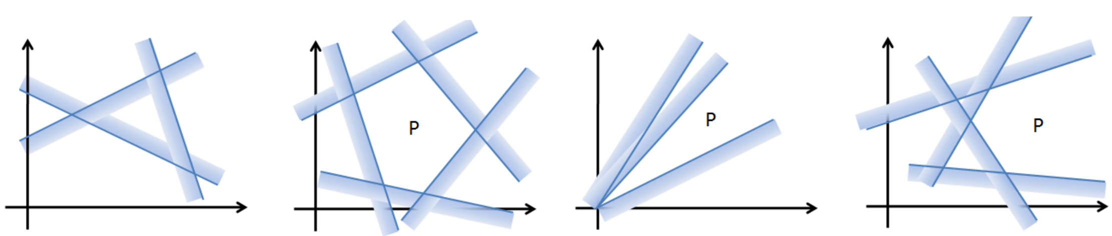
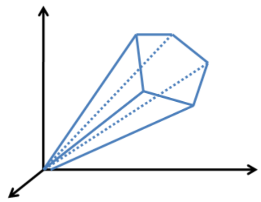


Figure 2: Examples of polyhedrons, left to right: Empty, Polytope, Cone, Combination of cone and polytope.

What "combination of cone and polytope" means will become clear soon in Theorem 12. For the examples, the reader is referred to Figure 2. In 2-D, a cone can have only two "extreme rays," while in 3-D there is no bound on the number of extreme rays it can have.



For the most of part, we will be largely concerned with polytopes, but we need to have a better understanding of polyhedra first. Although it is geometrically "obvious" that a polytope is the convex hull of its "vertices," the proof is quite non-trivial. We will state the following three theorems without proof.

Theorem 10. *A bounded polyhedron is the convex hull of a finite set of points.*

Theorem 11. *A polyhedral cone is generated by a finite set of vectors. That is, for any $A \in \mathbb{R}^{m \times n}$, there exists a finite set X such that $\{x = \sum_i \lambda_i x_i \mid x_i \in X, \lambda_i \geq 0\} = \{x \mid Ax \leq 0\}$.*

Theorem 12. *A polyhedron $\{x \mid Ax \leq b\}$ can be written as the Minkowski sum of a polytope Q and a cone C, i.e. $P = Q + C = \{x + y \mid x \in Q, y \in C\}$.*

One can (geometrically) obtain the Minkowski sum of a polytope Q and a cone C by sweeping the origin of the cone C over the polytope Q. If the polyhedron P is pointed (has at least one "vertex"), the decomposition is, in fact, modulo scaling factor unique. Further the cone C above is $\{x \mid Ax \leq 0\}$, or equivalently the set of unbounded directions in P. The cone C is called the characteristic cone or the recession cone of P.

Many facts about polyhedra and linear programming rely on (in addition to convexity) variants of Farkas' lemma that characterizes when a system of linear inequalities do not have solution. The simplest proof for one variant is via Fourier-Motzkin elimination that is independently interesting and related to the standard Gauss-Jordan elimination for solving system of linear equations.

1.2.1 Fourier-Motzkin Elimination

Let $P = \{x | Ax \leq b\} \subseteq \mathbb{R}^n$ be a polyhedron. For k in $[n]$,
we let $P^k = \{(x_1, .., x_{k-1}, x_{k+1}, ..., x_n) \mid (x_1, x_2, ..., x_n) \in P\}$ be the *projection* of P along the x_k-axis.

Theorem 13. *P^k is a polyhedron.*

Proof. We derive a set of inequalities that describe P^k. We do this by considering the inequalities in $Ax \leq b$ and eliminating the variables x_k as follows. Partition the inequalities in $Ax \leq b$ into three sets:

$$S^+ = \{i \in [m] \mid a_{ik} > 0\}, S^- = \{i \in [m] \mid a_{ik} < 0\}, \text{and } S^0 = \{i \in [m] \mid a_{ik} = 0\}.$$

Define a new set of inequalities consisting of S_0 and one new inequality for each pair (i, ℓ) in $S^+ \times S^-$:

$$a_{ik}(\sum_{j=1}^{n} a_{\ell j} x_j) - a_{\ell k}(\sum_{j=1}^{n} a_{ij} x_j) \leq a_{ik} b_\ell - a_{\ell k} b_i.$$

Note that the combined inequality does not have x_k. We now have a total of $|S^0| + |S^+||S^-|$ new inequalities. Let $P' = \{x' \in \mathbb{R}^{n-1} \mid A'x' \leq b'\}$ where $A'x' \leq b'$ is the new system of inequalities in variables $x_1, x_2, ..., x_{k-1}, x_{k+1}, ..., x_n$. We prove the theorem by showing that $P^k = P'$.

We first show the easier direction: $P^k \subseteq P'$. Consider any point $z \in P^k$. By definition of P^k, there exists $x \in P$ such that $Ax \leq b$ and x's projection along x_k-axis is z. It is easy to see that z satisfies the new system since the new one was obtained in a way oblivious to x_k, the real value of x's k_{th} coordinate.

We now show that $P' \subseteq P^k$. Without loss of generality, assume $k = 1$. Consider any $x' = (x_2, x_3, ..., x_n) \in P'$. We want to show that there exists $x_1 \in \mathbb{R}$ such that $Ax \leq b$, where $x = (x_1, x_2, ..., x_n)$. For simple notation, define $C_i = b_i - \sum_{j=2}^{n} a_{ij}x_j$ for $i \in [m]$. Note that $Ax \leq b$ can be rewritten as

$$a_{i1}x_1 \leq C_i, \forall i \in [m]. \tag{1}$$

Observe that x satisfies all inequalities consisting of S^0, since the new system as well includes those constraints. Thus we can refine our goal to show

$$\begin{array}{rl} \exists x_1 \text{ s.t.} & a_{i1}x_1 \leq C_i, \forall i \in S^+ \cup S^-. \\ \Leftrightarrow & \max_{\ell \in S^-} \frac{C_\ell}{a_{\ell 1}} \leq x_1 \leq \min_{i \in S^+} \frac{C_i}{a_{i1}}. \end{array}$$

It is easy to observe that this is equivalent to

$$\begin{array}{rrcll} & \frac{C_\ell}{a_{\ell 1}} & \leq & \frac{C_i}{a_{i1}} & , \forall (i,\ell) \in S^+ \times S^- \\ \Leftrightarrow & 0 & \leq & a_{i1}C_\ell - a_{\ell 1}C_i & , \forall (i,\ell) \in S^+ \times S^- \\ \Leftrightarrow & A'x' & \leq & b' & \end{array}$$

And we know that $A'x' \leq b'$ since $x' \in P'$, completing the proof. □

From Fourier-Motzkin elimination we get an easy proof of one variant of Farkas' lemma.

Theorem 14 (Theorem of Alternatives)**.** *Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. For the system $Ax \leq b$, exactly one of the following two alternatives hold:*

- *The system is feasible.*
- *There exists $y \in \mathbb{R}^m$ such that $y \geq 0$, $yA = 0$ and $yb < 0$.*

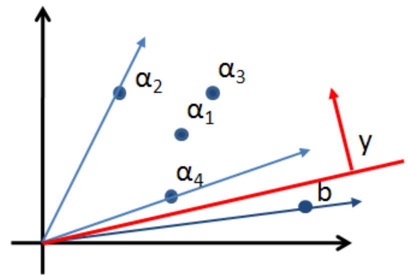
What the theorem says is that if the system of inequalities $Ax \leq b$ is infeasible then there is a proof (certificate) of this which can be obtained by taking non-linear combination of the inequalities (given by $y \geq 0$) to derive a contradiction of the following form: $0 = yA \leq yb < 0$.

Proof of [Theorem 14] Suppose that there exists a vector $y' \geq 0$ s.t. $y'A = 0$ and $y' \cdot b < 0$ and a vector x' such that $Ax' \leq b$. Then it easily follows that $0 \leq y'Ax' \leq y'b$, since $y' \geq 0$, which is a contradiction to the fact that $y'b < 0$.

Conversely, suppose $Ax \leq b$ is infeasible. Let $P = \{x \mid Ax \leq b\}$. We eliminate variables $x_1, x_2, ..., x_n$ (we can choose any arbitrary order) to obtain polyhedra $P = Q_0, Q_1, Q_2, ..., Q_{n-1}, Q_n$. Note that Q_{i+1} is non-empty iff Q_i is, and that Q_{n-1} has only one variable and Q_n has none. Note by the Fourier-Motzkin elimination procedure the inequalities describing Q_i are non-negative combination of the inequalities of P; this can be formally shown via induction. Thus, Q_n is empty iff we have derived an inequality of the form $0 \leq C$ for some $C < 0$ at some point in the process. That inequality gives the desired $y \geq 0$. □

Two variant of Farkas' lemma that are useful can be derived from the theorem of alternatives.

Theorem 15. *$Ax = b$, $x \geq 0$ has no solution iff $\exists y \in \mathbb{R}^m$ s.t. $A^T y \geq 0$ and $b^T y < 0$.*



The above theorem has a nice geometric interpretation. Let $\alpha_1, \alpha_2, ..., \alpha_n$ be the columns of A viewed as vectors in $\mathbb{R}^m$. Then $Ax = b, x \geq 0$ has a solution if and only if b is in the cone generated by $\alpha_1, \alpha_2, ..., \alpha_n$; here the combination is given by $x \geq 0$. So b is either in the $\texttt{Cone}(\alpha_1, \alpha_2, ..., \alpha_n)$ or there is a hyperplane separating b from $\alpha_1, \alpha_2, ..., \alpha_n$.

In fact the theorem can be strengthened to show that the hyperplane can be chosen to be one that spans $t-1$ linearly independent vectors in $\alpha_1, \alpha_2, ..., \alpha_n$, where $t = \texttt{rank}(\alpha_1, \alpha_2, ..., \alpha_n, b)$.

Proof of [Theorem 15] We can rewrites $Ax = b, x \geq 0$ as

$$\begin{bmatrix} A \\ -A \\ -I \end{bmatrix} x \leq \begin{bmatrix} b \\ -b \\ 0 \end{bmatrix}$$

Hence by the Theorem of Alternatives, $Ax = b, x \geq 0$ is not feasible if and only if there exists a row vector $y' = \begin{bmatrix} u & v & w \end{bmatrix}$, where u, v are row vectors in $\mathbb{R}^m$ and w is a row vector in $\mathbb{R}^n$ such that

$$\begin{aligned} u, v, w &\geq 0 \\ uA - vA - w &= 0 \\ ub - vb &< 0 \end{aligned}$$

Let $y = u - v$. Note that $y \in \mathbb{R}^m$ is now not necessarily positive. From the second and third inequalities, we can easily obtain $A^T y \geq 0$ and $b^T y < 0$. $\square$

Another variant of Farkas' lemma is as follows and the proof is left as an exercise.

Theorem 16. *$Ax \leq b, x \geq 0$ has a solution iff $yb \geq 0$ for each row vector $y \geq 0$ with $yA \geq 0$.*

Another interesting and useful theorem is Carathéodory's Theorem

Theorem 17 (Carathéodory)**.** *Let $x \in \texttt{Convexhull}(X)$ for a finite set X of points in $\mathbb{R}^n$. Then $x \in \texttt{Convexhull}(X')$ for some $X' \subseteq X$ such that vectors in X' are affinely independent. In particular, $|X'| \leq n+1$.*

A conic variant of Carathéodory's Theorem is as follows.

Theorem 18. *Let $x \in \texttt{Cone}(X)$ where $X = \{x_1, x_2, ..., x_m\}$, $x_i \in \mathbb{R}^n$. Then $x \in \texttt{Cone}(X')$ for some $X' \subseteq X$ where vectors in X' are linearly independent. In particular, $|X'| \leq n$.*

Proof. Since $x \in \texttt{Cone}(X)$, $x = \sum_i \lambda_i x_i$ for some $\lambda_i \geq 0$. Choose a combination with mininum support, i.e. the smallest number of non-zero λ_i values. Let $X' = \{\lambda_i x_i \mid \lambda_i > 0\}$ and $I = \{i \mid \lambda_i > 0\}$. If vectors in X' are linearly independent, we are done. Otherwise, $\exists \alpha_i, i \in I$ s.t. $\sum_{i \in I} \alpha_i \lambda_i x_i = 0$. By scaling we can assume that $\forall i \in I, \alpha_i \leq 1$, and $\exists j \in I$ s.t. $\alpha_j = 1$. Then,

$$x = \sum_{i \in I} \lambda_i x_i = \sum_{i \in I} \lambda_i x_i - \sum_{i \in I} \alpha_i \lambda_i x_i = \sum_{i \in I} \lambda_i (1 - \alpha_i) x_i = \sum_{i \in I \setminus \{j\}} \lambda_i' x_i \quad (\lambda_i' \geq 0).$$

This contradicts the fact that we chose the conical combination for x with the least support. □

One can derive the affine version of Carathéodory's Theorem from the conical version, and the proof is left as an exercise.

1.3 Linear Programming

Linear programming is an optimization problem of the following form.

$$\begin{aligned} \max \;\; & c \cdot x \quad \text{(Primal-LP)} \\ Ax & \leq b \end{aligned}$$

The above is one of the standard forms. In other words, we wish to maximize a linear objective function over a *polyhedron*. Given an LP, there are three possibilities:

1. The polyhedron is *infeasible*.
2. The objective function can be made arbitrarily large in which case we can say it is *unbounded*.
3. There is a finite optimum value in which case we say it is *bounded*.

Each linear program has its associated "dual" linear program. The LP we refer to by "dual" depends on the "starting" LP, which is called as the primal LP; in fact the dual of dual LP is exactly the same as the primal LP. Let us say that the following LP is the primal LP here.

$$\begin{aligned} \max \;\; & c \cdot x \\ Ax & \leq b \end{aligned}$$

We can "derive" the dual by thinking about how we can obtain an upper bound on the optimal value for the primal LP. Given the system $Ax \leq b$, any inequality obtained by non-negative combination of the inequalities in $Ax \leq b$ is a valid inequality for the system. We can represent a non-negative combination by a $m \times 1$ row vector $y \geq 0$.

Thus $yAx \leq yb$ is a valid inequality for $y \geq 0$. Take any vector $y' \geq 0$ s.t. $y'A = c$. Then such a vector gives us an upperbound on the LP value since $y'Ax = cx \leq y'b$ is a valid inequality. Therefore one can obtain an upperbound by minimizing over all $y' \geq 0$ s.t. $y'A = c$. Therefore the objective function of the primal LP is upperbounded by the optimum value of

$$\begin{aligned} \min \;\; & yb \quad \text{(Dual-LP)} \\ yA & = c \\ y & \geq 0 \end{aligned}$$

The above derivation of the Dual LP immediately implies the Weak Duality Theorem.

Theorem 19 (Weak Duality)**.** *If x' and y' are feasible solutions to Primal-LP and Dual-LP then $cx' \leq y'b$.*

Corollary 20. *If the primal-LP is unbounded then the Dual-LP is infeasible.*

Exercise: Prove that the dual of the Dual-LP is the Primal-LP.

The main result in the theory of linear programming is the following Strong Duality Theorem which is essentially a min-max result.

Theorem 21 (Strong Duality)**.** *If Primal-LP and Dual-LP have feasible solutions, then there exist feasible solutions x^* and y^* such that $cx^* = y^*b$.*

Proof. Note that by weak duality we have that $cx' \leq y'b$ for any feasible pair of x' and y'. Thus to show the existence of x^* and y^* it suffices to show that the system of inequalities below has a feasible solution whenever the two LPs are feasible.

$$\begin{aligned} cx &\geq yb \\ Ax &\leq b \\ yA &= c \\ y &\geq 0 \end{aligned}$$

We rewrite this as

$$\begin{aligned} Ax &\leq b \\ yA &\leq c \\ -yA &\leq -c \\ -y &\leq 0 \\ -cx + yb &\leq 0 \end{aligned}$$

and apply the Theorem of Alternatives. Note that we have inequalities in $n+m$ variables corresponding to the x and y variables. By expressing those variables as a vector $z = \begin{bmatrix} x \\ y^T \end{bmatrix}$, we have

$$\begin{bmatrix} A & 0 \\ 0 & A^T \\ 0 & -A^T \\ 0 & -I \\ -c & b^T \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq \begin{bmatrix} b \\ c^T \\ -c^T \\ 0 \\ 0 \end{bmatrix}$$

If the above system does not have a solution then there exists a vector $\begin{bmatrix} s & t & t' & u & v \end{bmatrix} \geq 0$, where s is a $m \times 1$ row vector, t, t' are $n \times 1$ row vectors, u is a $m \times 1$ row vector and v is a scalar such that

$$\begin{aligned} sA - v \cdot c &= 0 \\ tA^T - t'A^T - u + v \cdot b &= 0 \\ sb + t \cdot c - t \cdot c' &< 0 \end{aligned}$$

We replace $t - t'$ by w and note that now w is not necessarily positive. Hence we obtain that if the strong duality does not hold then there exist vectors $s, u \in \mathbb{R}^m, w \in \mathbb{R}^n, v \in \mathbb{R}$ such that

$$\begin{aligned} s, u, v &\geq 0 \\ sA - v \cdot c &= 0 \\ wA^T - u + v \cdot b &= 0 \\ sb + wc^T &< 0 \end{aligned}$$

We consider two cases.
Case 1: $v = 0$. (note that v is a scalar.)
In this case, the above system simplifies to

$$\begin{aligned} s &\geq 0 \\ sA &= 0 \\ wA^T &= 0 \\ sb + wc^T &< 0 \end{aligned}$$

Since we have $y^*A = c$ and $sA = 0$, $y^* + \alpha s$ is a feasible solution for the dual for any scalar $\alpha \geq 0$. Similarly knowing that $Ax^* \leq b$ and $Aw^T = 0$ (from $wA^T = 0$), it follows that $x^* - \alpha w^T$ is feasible for the primal for any scalar $\alpha \geq 0$. Applying the Weak Duality Theorem, we have that $\forall \alpha \geq 0$,

$$\begin{aligned} & c(x^* - \alpha w^T) &\leq& \quad (y^* + \alpha s) \cdot b \\ \Rightarrow \quad & cx^* - y^*b &\leq& \quad \alpha(s \cdot b + c \cdot w^T) \end{aligned}$$

However, the LHS is fixed while the RHS can be made arbitrarily small because $s \cdot b + c \cdot w^T < 0$ and α can be chosen arbitrarily large.
Case 2: $v > 0$.
Let $s' = \frac{1}{v}(s), w' = \frac{1}{v}(w)$, and $u' = \frac{1}{v}(u)$. Then, we have

$$\begin{aligned} s', u' &\geq 0 \\ s'A &= c \\ w'A^T - u' &= -b \Rightarrow -A(w')^T = b - u' \leq b \ \text{ [Since } u' \geq 0.] \\ s'b + w'c^T &< 0 \end{aligned}$$

From the inequalities above, we observe that s' is dual feasible and $-w'$ is primal feasible. Thus by the Weak Duality, we have $-w' \cdot c^T \leq s'b$, contradicting that $s'b + w \cdot c^T < 0$.

Finally, we make a remark on where the contradiction really comes from for each case. For the first case where $v = 0$, note that the inequality $-cx + yb \leq 0$, which forces the optimal values for the primal and the dual meet each other, was never used. In other words, we got a contradiction because there do not exist feasible solutions satisfying both the primal and the dual. On the other hand, for the second case $v > 0$, we had feasible solutions for both LPs, but obtained a contradiction essentially from the assumption that the two optimal values are different. $\square$

Complementary Slackness is a very useful consequence of Strong Duality.

Theorem 22 (Complementary Slackness)**.** *Let x^*, y^* be feasible solutions to the primal and dual LPs. Then x^*, y^* are optimal solutions if and only if $\forall i \in [m]$, either $y_i^* = 0$ or $a_i x^* = b_i$.*

Proof. Let $\alpha_1, \alpha_2, ...\alpha_m$ be the row vectors of A. Suppose that the given condition is satisfied. Then we have $y^*Ax^* = \sum_{i=1}^m y_i^*(\alpha_i x^*) = \sum_{i=1}^m y_i^* b_i = y^*b$. Also we know that $cx^* = y^*Ax^*$ since y^* is a feasible solution for the dual. Thus we have $cx^* = y^*b$, and by the Weak Duality, we conclude that x^*, y^* are optimal.

Conversely, suppose that x^* and y^* both are optimal. Then by the Strong Duality and because x^*, y^* are feasible, we have that $y^*b = cx^* = y^*Ax^* \leq y^*b$. Thus we obtain the equality $y^*Ax^* = y^*b$, that is $\sum_{i=1}^m y_i^*(\alpha_i x^*) = \sum_{i=1}^m y_i^* b_i$. This equality forces the desired condition, since $\alpha_i x^* \leq b_i, y_i^* \geq 0$ because x^* and y^* are feasible solutions. □

References

[1] L. Schrijver. *Theory of Linear and Integer Programming (Paperback)*. Wiley, Chapters 7, 8, 1998.

[2] J. Lee. *A First Course in Combinatorial Optimization*. Cambridge University Press, 2004.

[3] M. Goemans. Lecture notes on linear programming and polyhedral combinatorics. http://www-math.mit.edu/ goemans/18433S09/polyhedral.pdf, 2009.