Linear Programming in Combinatorics

Amritanshu Prasad

The Institute of Mathematical Sciences, Chennai.

Homi Bhabha National Institute, Mumbai.
Email address: amri@imsc.res.in

CHAPTER 1

Introduction to Linear Programming

1. Feasibility and Optimization

A linear program in equational form consists of a set of variables, $\mathbf{x} = (x_1, \dots, x_n)$, an $m \times n$ matrix with real entries A, a bound vector $\mathbf{b} = (b_1, \dots, b_m)$, and an objective vector $\mathbf{c} = (c_1, \dots, c_n)$. The *linear program* is the problem:

$$\text{(LP)} \qquad \text{maximize } \mathbf{c}^T\mathbf{x} \text{ subject to } \mathbf{x} \geq 0 \text{ and } A\mathbf{x} = \mathbf{b}.$$

The set $P = P(A, \mathbf{b})$ of all vectors satisfying $\mathbf{x} \geq 0$ and $A\mathbf{x} = \mathbf{b}$ is called the polytope of all *feasible solutions*. The function $\mathbf{x} \mapsto \mathbf{c}^T\mathbf{x}$ is called the *objective function*. An *optimal solution* is a vector $\mathbf{x}_0 \in P(A, \mathbf{b})$ such that $\mathbf{c}^T\mathbf{x} \leq \mathbf{c}^T\mathbf{x}_0$ for every $\mathbf{x} \in P(A, \mathbf{b})$. Sometimes we will only be interested in the set of feasible solutions, which does not depend on the objective vector.

Assume without loss of generality that $\mathbf{b}$ lies in the column space of A, and that the rows of A are linearly independent. For a subset $B \subset [n]$, let A_B denote the subatrix of A consisting of columns from B. We say that B is a *basic set* if B has m elements and A_B has rank m. For $\mathbf{x} \in \mathbf{R}^n$ define:

$$\operatorname{supp}(\mathbf{x}) = \{1 \leq j \leq n \mid x_j \neq 0\}.$$

Definition 1.1 (Basic feasible solution). A *basic feasible solution* to (LP) is a feasible solution $\mathbf{x} \in P(A, \mathbf{b})$ such that $\operatorname{supp}(\mathbf{x})$ is contained in a basic set.

Clearly, a feasible solution $\mathbf{x}$ is basic if and only if the submatrix $A_{\operatorname{supp}(\mathbf{x})}$ has linearly independent columns.

Example 1.2 (The Birkhoff polytope). Take $n = d^2$, indexing the d^2 variables as $\mathbf{x} = (x_{ij})_{1 \leq i,j \leq d}$, a square array of size d. As constraints, say that the row sums and column sums of $\mathbf{x}$ are all equal to 1, i.e.,

$$\sum_i x_{ij} = 1 \text{ for } j = 1, \dots, d$$
$$\sum_j x_{ij} = 1 \text{ for } j = 1, \dots, d.$$

These $2d$ constraints are not independent–the sum of the row sum constaints is equal to the sum of the column sum constraints, which is the same as the constraint that all entries of the matrix add up to d. Removing, say the last column constraint gives a $(2d-1) \times d^2$ matrix A of rank $2d - 1$.

When $d = 2$, the equation is:

$$(1) \qquad \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Every submatrix of A with three columns is non-singular. Thus there are four possible basic sets, but only two basic solutions given by the 2×2 permutation matrices.

In general, what are the possible basic sets and basic feasible solutions? Let $\sigma \in S_d$ and suppose $\mathbf{x}$ is the permutation matrix $x_{ij} = \delta_{i\sigma(i)}$. Then $\mathrm{supp}(\mathbf{x}) = \{(i, \sigma_i) \mid i \in [d]\}$. The first d rows of the corresponding column vectors of A are just the coordinate vectors of $\mathbf{R}^d$. Therefore each permutation matrix is a basic feasible solution. What basic sets does it correspond to? Can you show that there are no more basic feasible solutions?

Given a subset S of $[d]^2$, let $\mathbf{c} = (c_{ij})$ be the array whose (i, j)th entry is 1 if $(i, j) \in S$, and 0 otherwise. What is the correspondence between subsets $S \subset [d]^2$ and the basic feasible solutions that maximize the corresponding objective function?

Lemma 1.3. *For each basic subset $B \subset [d]$, there exists at most one basic feasible solution $\mathbf{x}$ with $\mathrm{supp}(\mathbf{x}) \subset B$.*

Proof. Let $\mathbf{x}_B$ denote the vector $(x_i)_{i \in B}$. The matrix A_B is non-singular, so the equation $A_B\mathbf{x}_B = \mathbf{b}$ has at most one solution. Solutions $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$ with $\mathrm{supp}(x) \subset B$ are in bijection with solutions of $A_B\mathbf{x}_B = \mathbf{b}$ (set $x_j = 0$ for $j \notin B$ to get $\mathbf{x}$ from $\mathbf{x}_B$). Therefore $A\mathbf{x} = \mathbf{b}$ also has at most one solution. □

Remark 1.4. The same basic feasible solution could be obtained from different basic sets. For example, each basic solution for (1) corresponds to two basic sets. Also not every basic set B admits a basic feasible solution. For example, in Example 1.2, $B = \{(1, 1), (1, 2), (1, 3), (2, 1), (3, 1)\}$ is a basic set with no feasible solution.

Theorem 1.5 (Existence of basic optimal solutions). *For a linear program in equational form:*

$$\text{maximize } \mathbf{c}^t\mathbf{x} \text{ subject to } A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0$$

if there is at least one feasible solution, and the objective function is bounded above on $P(A, \mathbf{b})$, then there exists at least one optimal solution. Among the optimal solutions there is at least one basic solution.

Proof. We claim that, for any feasible solution $\mathbf{x}_0$, there exists a basic feasible solution $\mathbf{x}$ with $\mathbf{c}^T\mathbf{x} \geq \mathbf{c}^T\mathbf{x}_0$. This implies that an optimal solution, if it exists, will be basic. Suppose $\mathbf{x}$ is a feasible solution. Among all feasible solutions $\mathbf{x}$ with $\mathbf{c}^T\mathbf{x} \geq \mathbf{c}^T\mathbf{x}_0$ choose one with support of minimal cardinality and call it $\tilde{\mathbf{x}}$. If $A_{\mathrm{supp}(\tilde{\mathbf{x}})}$ is non-singular then $\tilde{\mathbf{x}}$ is basic and we are done. Otherwise, there exists a vector $\mathbf{y} \in \mathbf{R}^n$ with $\mathrm{supp}(\mathbf{y}) \subset \mathrm{supp}(\mathbf{x})$ such that $A\mathbf{y} = 0$. Replacing $\mathbf{y}$ by $-\mathbf{y}$ if necessary, assume that $\mathbf{c}^T\mathbf{y} \geq 0$.

We claim that we may further assume that $\mathbf{y}$ has at least one *negative* coordinate. Suppose that all the coordinates of $\mathbf{y}$ are non-negative. If $\mathbf{c}^T\mathbf{y} = 0$, then we can replace $\mathbf{y}$ with $-\mathbf{y}$. If $\mathbf{c}^T\mathbf{y} > 0$ and all coordinates of $\mathbf{y}$ are positive, then $\tilde{\mathbf{x}} + t\mathbf{y}$ is a feasible solution for all $t > 0$. The objective function $\mathbf{c}^T(\tilde{\mathbf{x}} + t\mathbf{y})$ grows unboundedly as t grows, contradicting its boundedness.

Thus $\mathbf{y}$ has at least one negative coordinate, hence it is possible to choose a value $t > 0$ such that $\tilde{\mathbf{x}} + t\mathbf{y}$ is a feasible solution with $\mathbf{c}^T(\tilde{\mathbf{x}} + t\mathbf{y}) \geq \mathbf{c}^T\tilde{\mathbf{x}}$ and $\mathrm{supp}(\tilde{\mathbf{x}} + t\mathbf{y})$ is strictly smaller than $\mathrm{supp}(\tilde{\mathbf{x}})$. This contradicts the minimality condition on the cardinality of $\mathrm{supp}(\tilde{\mathbf{x}})$.

The set of basic feasible solutions is finite. The element of this set that maximizes the objective function must therefore be an optimal solution. □

DEFINITION 1.6 (Vertex). Let $P \subset \mathbf{R}^n$ be convex closed set. An element $\mathbf{v} \in P$ is said to be a *vertex* of P if there exists $\mathbf{c} \in \mathbf{R}^n$ such that $\mathbf{c}^T\mathbf{x}$ attains its maximum uniquely at $\mathbf{v}$.

Theorem 1.5 says that every vertex of $P(A, \mathbf{b})$ is a basic feasible solution. The converse is also true:

THEOREM 1.7. *The basic feasible solutions to* (LP) *are precisely the vertices of* $P(A, \mathbf{b})$.

PROOF. Let $B \subset [n]$ be a basic subset $\mathbf{v}$ be the basic feasible solution to (LP) with respect to B. Define $\mathbf{c}$ to be the vector with $c_j = 0$ for $j \in B$, and $c_j = -1$ otherwise. Then $\mathbf{c}^T\mathbf{v} = 0$, and by Lemma 1.3, $\mathbf{c}^T\mathbf{x} < 0$ for every $\mathbf{x} \in P$. □

DEFINITION 1.8 (General form of a linear program). A more general form of a linear program involves linear inequalities and equalities. As before take A to be an $m \times n$ matrix with real entries, $\mathbf{b} \in \mathbf{R}^m$, and $\mathbf{c} \in \mathbf{R}^n$. A general linear program has the form:

$$\text{(GLP)} \qquad \text{optimize } \mathbf{c}^T\mathbf{x} \text{ subject to } a_{i1}x_1 + \cdots + a_{in}x_n \; R_i \; b_i \text{ for } i = 1, \ldots, m,$$

where R_i is one of the three symbols $\leq$, $\geq$, and $=$, and the word optimize is replaced by either maximize, or minimize. A basic feasible solution is one that is defined by equalities in n linearly independent constraints (which could be equality or inequality to begin with).

EXAMPLE 1.9. The cube can be defined by the inequalities:

$$0 \leq x_i \leq 1, \text{ for } i = 1, 2, 3.$$

The inequality $x_i \leq 1$ can be turned into an equality by introducing *slack variables* y_i, and writing:

$$x_i \geq 0, \; y_i \geq 0, \; x_i + y_i \leq 1 \text{ for } i = 1, 2, 3.$$

The linear program in equational form is equivalent to the original, more general one, in the sense that there is a bijection amongst their feasible solutions that maps vertices to vertices (why?). What are the basic subsets? What are the basic feasible solutions?

EXERCISE 1.10. Rewrite the n-simplex:

$$\Delta_n = \{(x_1, \ldots, x_n) \mid 0 \leq x_1 \leq x_2 \leq \cdots x_n \leq 1\}$$

in equational form. Determine

EXERCISE 1.11. Express the hyperoctahedron:

$$H_n = \{\mathbf{x} \mid -1 \leq x_1 + \cdots + x_n \leq 1\}$$

in equational form.

2. The Simplex Method

The simplex method begins with a basic set B for which there exists a basic feasible solution. Using the relations imposed by $A\mathbf{x} = \mathbf{b}$, each of the variables x_j, where $j \notin B$, can be expressed in terms of the basic variables. The objective function can also be expressed in terms of the basic variables. Suppose that $B = \{j_1, \ldots, j_m\}$ and $\bar{B} = \{k_1, \ldots, k_{n-m}\}$ is the complement of B in $[n]$. Then the system of equations $A\mathbf{x} = \mathbf{b}$ can be rearranged as:

$$A_B \mathbf{x}_B = \mathbf{b} - A_{\bar{B}} \mathbf{x}_{\bar{B}}. \tag{2}$$

Since A_B is invertible, the basic variables can be expressed in terms of the non-basic ones:

$$\mathbf{x}_B = A_B^{-1}(\mathbf{b} - A_{\bar{B}} \mathbf{x}_{\bar{B}}).$$

Indeed, the basic feasible point is computed by setting $\mathbf{x}_{\bar{B}} = 0$ in the above equation. Since each basic variable is expressed in terms of the non-basic variables in (2), the objective function can be expressed in terms of the non-basic variables only.

This data is represented in terms of a *simplex tableau*:

$$\begin{array}{ccc} \mathbf{x}_B & = & \mathbf{d} - D\mathbf{x}_{\bar{B}} \\ \hline \mathbf{c}^T\mathbf{x} & = & e - \mathbf{e}^T \mathbf{x}_{\bar{B}}. \end{array} \tag{T}$$

Here:

$$\begin{aligned} \mathbf{d} &= A_B^{-1}\mathbf{b} \\ D &= A_B^{-1} A_{\bar{B}} \\ e &= \mathbf{c}_B^T \mathbf{d} \\ \mathbf{e}^T &= \mathbf{c}_B^T D - \mathbf{c}_{\bar{B}}^T. \end{aligned}$$

There are m equations above, one for each basic variable, expressing it as an affine linear combination of the basic variables. The last line of (T) is simply the objective function expressed in terms of the basic variables. The information contained in (T) is equivalent to the information in (LP). But (T) gives a sort-of parametrization of $P(A, b)$. Note that $\mathbf{x}$ is determined by $\mathbf{x}_B$ and $\mathbf{x}_{\bar{B}}$. We have:

> Each vector $\mathbf{x} \in P(A, b)$ corresponds to a unique vector $\mathbf{x}_{\bar{B}} \in \mathbf{R}_{\geq 0}^{n-m}$ with $\mathbf{x}_B$ computed by using (T) is non-negative. Under this correspondence the basic feasible solution corresponding to B (if it exists) comes from $\mathbf{x}_{\bar{B}} = 0$.

EXAMPLE 2.1. Consider the Birkhoff polytope (Example 1.2) for $d = 3$. For convenience we abbreviate the variable indices (i, j) to ij. The matrix A has columns indexed by pairs in $[d] \times [d]$ written in increasing lexicographic order:

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Fix as objective function $x_{11} + x_{22} + x_{33}$. A basic subset is $B = \{11, 12, 13, 22, 31\}$ with basic solution given by $x_{13} = x_{22} = x_{31} = 1$, and all other coordinates zero.

The corresponding simplex tableau is:

$$\begin{array}{rcl} x_{11} & = & -x_{21}+x_{32}+x_{33} \\ x_{12} & = & x_{21}+x_{23}-x_{32} \\ x_{13} & = & 1-x_{23}-x_{33} \\ x_{22} & = & 1-x_{21}-x_{23} \\ x_{31} & = & 1-x_{32}-x_{33} \\ \hline \mathbf{c}^T\mathbf{x} & = & 1-2x_{21}-x_{23}+x_{32}+2x_{33}. \end{array}$$

In the above example, the objective function can be increased by increasing x_{32} or x_{33}. However, this increase should respect the constraints that all the variables are non-negative. The condition $x_{12} \geq 0$ (using the second equation, and leaving the values of x_{21} and x_{23} unchanged at 0) gives $x_{32} \leq 0$. Therefore, it is not feasible to increase x_{32}. However, it is feasible to increase x_{33}. The conditions $x_{13} \geq 0$ and $x_{31} \geq 0$ give $x_{33} \leq 1$. So we set $x_{33} = 1$, and recalculate all the basic variables, getting $x_{11} = x_{22} = x_{33} = 1$, and all other variables 0. We move x_{33} to the set of basic variables, and move x_{13} to the set of non-basic (which has now become 0), and use the equation:

$$x_{33} = 1 - x_{13} - x_{23}.$$

Using this we get a new tableau:

$$\begin{array}{rcl} x_{11} & = & 1-x_{21}+x_{32}+x_{23}-x_{13} \\ x_{12} & = & x_{21}+x_{23}-x_{32} \\ x_{22} & = & 1-x_{21}-x_{23} \\ x_{31} & = & x_{13}+x_{23}-x_{32}-x_{33} \\ x_{33} & = & 1-x_{23}-x_{13} \\ \hline \mathbf{c}^T\mathbf{x} & = & 3-2x_{13}-2x_{21}-3x_{23}+x_{32}. \end{array}$$

All the non-basic variables have negative coefficients, except x_{32}. However, the constrain $x_{12} \geq 0$ still does not allow us to increase x_{32} without changing any other non-basic variable. This suggests that we may have arrived at a maximum value for the objective function. Indeed, $x_{12} = x_{21} + x_{23} - x_{32} \geq 0$ implies that $x_{32} \leq x_{21} + x_{23}$, whene

$$\mathbf{c}^T\mathbf{x} = 3 - 2x_{13} - 2x_{21} - 3x_{23} + x_{32} \leq 3 - 2x_{13} - x_{21} - 2x_{23} \leq 3.$$

Therefore 3 is indeed a global maximum for the objective function, and is obtained uniquely at $x_{ij} = \delta_{ij}$.

An alternative approach would be to induct x_{32} into the set of basic variables, and remove x_{12}. Now the basic set is changed to $\{11, 22, 31, 32, 33\}$, but the basic feasible solution remains unchanged. This will result in the tableau:

$$\begin{array}{rcl} x_{11} & = & 1-x_{12}+x_{32}+2x_{23}-x_{13} \\ x_{22} & = & 1-x_{21}-x_{23} \\ x_{31} & = & x_{12}+x_{13}-x_{21}-x_{33} \\ x_{32} & = & x_{21}+x_{23}-x_{12} \\ x_{33} & = & 1-x_{23}-x_{13} \\ \hline \mathbf{c}^T\mathbf{x} & = & 3-2x_{13}-x_{21}-2x_{23}-x_{12}. \end{array}$$

In this case, all the coefficients of the objective function are negative, from which it immediately follows that $x_{ij} = \delta_{ij}$ is the unique global maximum.