The Geometry of Linear Programming*

Matthew Hough

March 20, 2023

Contents

*Based heavily on Chapter 2 in [1]

1 Polyhedra

We define a polyhedron (plural: polyhedra) as follows:

Definition 1.1 (polyhedron)**.** A **polyhedron** is a set that can be described in the form

$$\{x \in \mathbb{R}^n : Ax \geq b\}$$

where A is an $m \times n$ matrix and b is a vector in $\mathbb{R}^m$.

We also note that a set of the form

$$\{x \in \mathbb{R}^n : Ax = b,\ x \geq 0\}$$

is a polyhedron but written in **standard form**.

Definition 1.2 (bounded set)**.** A set $S \subseteq \mathbb{R}^n$ is **bounded** if there exists a constant K s.t. the absolute value of every component of every element of S is less than or equal to K.

Definition 1.3 (hyperplane/halfspace)**.** Let a be a nonzero vector in $\mathbb{R}^n$ and let b be a scalar.

1. The set $\{x \in \mathbb{R}^n : a^T x = b\}$ is called a **hyperplane**.
2. The set $\{x \in \mathbb{R}^n : a^T x \geq b\}$ is called a **halfspace**.

Note that a hyperplane is the boundary of a corresponding halfspace. Also, the vector a in the definition of the hyperplane is perpendicular to the hyperplane itself. One can see this by taking points x, y belonging to the same hyperplane and observing that $a^T x = a^T y \implies a^T(x - y) = 0$. This implies that a is orthogonal to any direction vector confined to the hyperplane.

Observe that $Ax \geq b$ is equivalent to requiring $a_i^T x \geq b_i$ for each $i \in [m]$, where a_i are the rows of A. It follows that a polyhedron is the intersection of finitely many halfspaces.

2 Extreme points, vertices, and basic feasible solutions

Definition 2.1 (extreme point)**.** Let P be a polyhedron. A vector $x \in P$ is an **extreme point** of P if we cannot find two vectors $y, z \in P$ both different from x, and a scalar $\lambda \in [0, 1]$ such that

$$x = \lambda y + (1 - \lambda) z$$

That is to say, if x lies on any line segment within P, it can only be an endpoint of that line segment. Or more succinctly, $x \notin \mathrm{relint}[y, z]$ for any $y, z \in P$ distinct from x.

Definition 2.2 (vertex)**.** Let P be a polyhedron. A vector $x \in P$ is a **vertex** of P if there exists some $c \in \mathbb{R}^n$ s.t. $c^T x < c^T y$ for all y satisfying $y \in P, y \neq x$.

Geometrically, this is saying that x is a vertex of P iff we can construct a hyperplane with P on one side of it s.t. it meets P only at the point x. Another way to think of this is that we can balance a hyperplane on the point x in a way that the hyperplane does not touch any other point in P.

The above two definitions are too generic however to be useful in the design of algorithms. We want a definition that relies on a representation of a polyhedron in terms of linear constraints, and which reduces to an algebraic test. Consider a polyhedron $P \subseteq \mathbb{R}^n$ defined in terms of the linear equality and inequality constraints:

$$\begin{aligned} a_i^T x &\geq b_i, \quad i \in M_1, \\ a_i^T x &\leq b_i, \quad i \in M_2, \\ a_i^T x &= b_i, \quad i \in M_3, \end{aligned}$$

where M_1, M_2 and M_3 are finite index sets, each a_i is a vector in $\mathbb{R}^n$, and each b_i is a scalar.

Definition 2.3 (active constraint)**.** If a vector x^* satisfies $a_i^T x^* = b_i$ for some $i \in M_1$, M_2 or M_3, we say that the corresponding constraint is **active** at x^*.

Observe that if there are n constraints that are active at a vector x^*, then x^* satisfies a certain system of n linear equations in n unknowns. This system has a unique solution iff these n equations are linearly independent.

Theorem 2.4. Let x^* be an element of $\mathbb{R}^n$ and let $I = \{i : a_i^T x = b_i\}$ be the set of indices of constraints that are active at x^*. Then, TFAE:

1. There exist n vectors in the set $\{a_i : i \in I\}$ which are linearly independent.
2. The span of the vectors a_i where $i \in I$, is all of of $\mathbb{R}^n$.
3. The systems of equations $a_i^T x = b_i$ where $i \in I$, has a unique solution.

Proof. Suppose that the vectors $a_i, i \in I$ span $\mathbb{R}^n$. Then the span of these vectors has dimension n. Hence, n of these vectors form a basis of $\mathbb{R}^n$, and are therefore linearly independent. Conversely, suppose that n of the vectors a_i, $i \in I$ are linearly independent. Then, the subspace spanned by these n vectoes is n-dimensional, so must be equal to $\mathbb{R}^n$. We have shown the equivalence of 1. and 2.

Now suppose the system of equations $a_i^T x = b_i$, $i \in I$ has multiple solutions, say x^1 and x^2. Then the nonzero vector $d = x^1 - x^2$ satisfies $a_i^T d = 0$ for all $i \in I$. Since d is orthogonal to every vector a_i, $i \in I$, d cannot be a linear combination of these vectors, so $a_i, i \in I$ cannot span $\mathbb{R}^n$. Conversely, if $a_i, i \in I$ do not span $\mathbb{R}^n$, choose a nonzero vector d which is orthogonal to the subspace spanned by these vectors. If x satisfies $a_i^T x = b_i$ for all i, we also must have $a_i^T (x + d) = b_i$ for all i. This proves the equivalence of 2. and 3. □

Note. If we say that certain constraints are linearly independent, we means that the corresponding vectors a_i are linearly independent.

We now give an algebraic definition of a ”corner point” as a feasible solution at which there are n linearly independent active constraints.

Definition 2.5 (basic solution)**.** Consider a polyhedron P defined by linear equality and inequality constraints, and let x^* be an element of $\mathbb{R}^n$.

1. The vector x^* is a **basic solution** if:

   (a) All equality constraints are active;
   (b) Out of the constraints that are active at x^*, there are n of them that are linearly independent.

2. If x^* is a basic solution that satisfies all of the constraints, we say it is a **basic feasible solution** (BFS).

Suppose we want to look for basic solutions of a polyhedron. Of course, we can first impose the equality constraints for feasibility. Then we can require that enough additional constraints are active, so we get a total of n linearly independent active constraints. Once we have n linearly independent active constraints, by Theorem 2.4 a unique vector x^* is determined. However, this procedure provides no guarantee on the feasibility of x^* since some of the inactive constraints could be violated. In the case that some of the inactive constraints are violated, we have a basic (but not basic feasible) solution.

Theorem 2.6. Let P be a nonempty polyhedron and let $x^* \in P$. Then TFAE:

1. x^* is a vertex;
2. x^* is an extreme point;
3. x^* is a basic feasible solution.

Proof. For the purposes of this proof, we assume wlog that P is represented in terms of constraints of the form $a_i^T x \geq b_i$ and $a_i^T x = b_i$.

Suppose $x^* \in P$ is a vertex. By definition, there exists some $c \in \mathbb{R}^n$ such that $c^T x^* < c^T y$ for every $y \in P$ with $y \neq x^*$. If $y \in P, z \in P$, $y \neq x^*, z \neq x^*$, and $\lambda \in [0, 1]$, $c^T x^* < c^T y$ and $c^T x^* < c^T z$. It follows that $c^T x^* < c^T(\lambda y + (1 - \lambda)z)$ and, therefore, $x^* \neq \lambda y + (1 - \lambda)z$. Thus, x^* is an extreme point.

Suppose $x^* \in P$ is not a basic feasible solution and let $I = \{i : a_i^T x^* = b_i\}$. Since x^* is not a BFS, there do not exist n linearly independent vectors in the family $a_i, i \in I$. Thus, the vectors $a_i, i \in I$, lie in a proper subspace of $\mathbb{R}^n$, and there exists some nonzero vector $d \in \mathbb{R}^n$ such that $a_i^T d = 0$, for all $i \in I$. Let $\epsilon > 0$ be sufficiently small and consider the vectors $y = x^* + \epsilon d$ and $z = x^* - \epsilon d$. Notice that $a_i^T y = a_i^T x^* = b_i$, for $i \in I$. Furthermore, for $i \notin I$, we have $a_i^T x^* > b_i$ and, provided that ϵ is small enough, we will also have $a_i^T y > b_i$ (it suffices to choose ϵ so that $\epsilon |a_i^T d| < a_i^T x^* - b_i$ for all $i \notin I$). Thus, when ϵ is small enough, $y \in P$ and, by a similar argument, $z \in P$. We finally notice that $x^* = (y + z)/2$, which implies that x^* is not an extreme point.

Let x^* be a BFS and let $I = \{i : a_i^T x^* = b_i\}$. Let $c = \sum_{i \in I} a_i$. We then have

$$c^T x^* = \sum_{i \in I} a_i^T x^* = \sum_{i \in I} b_i$$

Furthermore, for any $x \in P$ and any i, we have $a_i^T x \geq b_i$, and

$$c^T x = \sum_{i \in I} a_i^T x \geq \sum_{i \in I} b_i \tag{2.1}$$

This shows that x^* is an optimal solution to the problem of minimizing $c^T x$ over the set P. Furthermore, equality holds in (2.1) iff $a_i^T x = b_i$ for all $i \in I$. Since x^* is a BFS, there are n linearly independent constraints that are active at x^*, and x^* is the unique solution t the system of equations $a_i^T x = b_i$, $i \in I$. It follows that x^* is the unique minimizer $c^T x$ over the set P and, therefore, x^* is a vertex of P. $\square$

Corollary 2.7. Given a finite number of linear inequality constraints, there can only be a finite nunber of basic or basic feasible solutions.

Proof. Consider a system of m linear inequality constraints imposed on a vector $x \in \mathbb{R}^n$. At any basic solution, there are n linearly independent active constraints. Since any n linearly independent active constraints define a unique point, it follows that different basic solutions correspond to different sets of n linearly independent active constraints. Therefore, the number of basic solutions is bounded above by the number of ways that we can choose n constraints out of a total m, which is finite. $\square$

Although the number of basic, and therefore, basic feasible solutions is guaranteed to be finite, it can be very large. For example, take the unit cube:

$$\{x \in \mathbb{R}^n : 0 \leq x_i \leq 1,\ i \in [n]\}$$

The unit cube is defined in terms of $2n$ constraints (n for $0 \leq x_i$ and n for $x_i \leq 1$), but it has 2^n basic feasible solutions (the corners of the unit cube).

Definition 2.8 (adjacent basic solution)**.** Two distinct basic solutions to a set of linear constraints in $\mathbb{R}^n$ are said to be **adjacent** if we can find $n - 1$ linearly independent constraints that are active at both of them. If two adjacent basic solutions are also feasible, then the line segment that joins them is called an **edge** of the feasible set.

3 Polyhedra in standard form

Consider the aforementioned standard form polyhedron

$$P = \{x \in \mathbb{R}^n : Ax = b,\ x \geq 0\}$$

Let A have dimensions $m \times n$ and note m is the number of equality constraints. We will make the assumption that the m rows of the matrix are linearly independent. Since the rows are n-dimensional, this requires that $m \leq n$.

We will see that when P is nonempty, linearly dependent rows of A correspond to redundant constraints that can be discarded; therefore our linear independence assumption can be made without loss of generality.

Recall that at any basic solution, there must be n linearly independent constraints that are active. We know that every basic solution satisfies the equality constraints $Ax = b$, which provides us with m active constraints. Because of our assumption on the rows of A, these are linearly independent. In order to obtain a total of n active constraints, we need to choose $n - m$ of the variables x_i and set them to zero. This makes the constraints $x_i \geq 0$ active.

Theorem 3.1. Consider the constraints $Ax = b$ and $x \geq 0$. Assume that the $m \times n$ matrix A has linearly independent rows. A vector $x \in \mathbb{R}^n$ is a basic solution iff we have $Ax = b$, and there exist indices $B(1), \ldots, B(m)$ such that

1. The columns of $A_{B(1)}, \ldots, A_{B(m)}$ are linearly indpendent.

2. If $i \neq B(1), \ldots, B(m)$, then $x_i = 0$.

Proof. See [1, Theorem 2.4]. □

In view of Theorem 3.1, all basic solutions ot a standard form polyhedron can be constructed according to the following procedure:

1. Choose m linearly independent columns $A_{B(1)}, \ldots, A_{B(m)}$.

2. Let $x_i = 0$ for all $i \neq B(1), \ldots, B(m)$.

3. Solve the system of m equations $Ax = b$ for the unknowns $x_{B(1)}, \ldots, x_{B(m)}$.

If a basic solution constructed according to this procedure is nonnegative, then it is feasible, and it is a basic feasible solution. Conversely, since every basic feasible solution is a basic solution, any BFS can be obtained from this procedure.

Notation. If x is a basic solution, the variables $x_{B(1)}, \ldots, x_{B(m)}$ are called **basic variables**; the remaining variables are called **nonbasic**. The columns $A_{B(1)}, \ldots, A_{B(m)}$ are called the **basic columns** and, since they are linearly independent, they form a basis of $\mathbb{R}^m$.

We say two bases are **distinct** if they have different sets $\{B(1), \ldots, B(m)\}$ of basic indices; if two bases involve the same set of indices in a different order, they will be considered the same basis.

If we arrange the m basic variables next to each other in a matrix, we obtain an $m \times n$ matrix B, called the **basis matrix**. This matrix must be invertible, since we require the basic columns to be linearly independent. We can similarly define a vector x_B with the values of the basic variables. Thus,

$$B = \begin{bmatrix} A_{B(1)} & A_{B(2)} & \ldots & A_{B(m)} \end{bmatrix}, \quad x_B = \begin{bmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{bmatrix}.$$

We can determine the basic variables by solving the equation $Bx_B = b$. Observe this system of equations has a unique solution, since B has linearly independent columns.

Example 3.2. Let the constraint $Ax = b$ be of the form

$$\begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 6 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 8 \\ 12 \\ 4 \\ 6 \end{bmatrix}$$

Let us choose A_4, A_5, A_6, A_7 as our basic columns. Note that they are linearly independent and the corresponding basis matrix is the identity. We then have the basic solution $x = (0, 0, 0, 8, 12, 4, 6)$ which is nonnegative and therefore a BFS. Another basis can be obtained by choosing the columns A_3, A_5, A_6, A_7 (note that they are linearly independent). The corresponding basic solution is $x = (0, 0, 4, 0, -12, 4, 6)$, which is not feasible because $x_5 < 0$.

Suppose now that there was an eighth column A_8, identical to A_7. Then the two sets of columns $\{A_3, A_5, A_6, A_7\}$ and $\{A_3, A_5, A_6, A_8\}$ coincide. On the other hand, the corresponding sets of basic indices, which are $\{3, 5, 6, 7\}$ and $\{3, 5, 6, 8\}$, are different and we have two different bases, according to our conventions.

Note. Different basic solutions must correspond to different bases, because a basis uniquely determines a basic solution. However, two different bases may lead to the same basic solution. In other words, if we have two different basic solutions, they cannot have come from the same basis. However, it is possible that two different bases give rise to the same basic solution.

Definition 3.3 (adjacent basis)**.** We say that two bases are **adjacent** if they share all but one basic column.

It is not hard to check that adjacent basic solutions can always be obtained from two adjacent bases. Conversely, if two adjacent bases lead to distinct basic solutions, then the two basic solutions are adjacent.

Example 3.4. Referring back to Example 3.2, the bases $\{A_4, A_5, A_6, A_7\}$ and $\{A_3, A_5, A_6, A_7\}$ are adjacent. The corresponding basic solutions $x = (0, 0, 0, 8, 12, 4, 6)$ and $x = (0, 0, 4, 0, -12, 4, 6)$. are adjacent since $n = 7$ and we have a total of six common linearly independent active constraints; these are $x_1 \geq 0, x_2 \geq 0$, and the four equality constraints.

The following theorem will show that with no loss of generality, we may assume that A has foll row rank.

Theorem 3.5. Let $P = \{x : Ax = b, x \geq 0\}$ be a nonempty polyhedron, where A is a matrix of dimensions $m \times n$, with rows $a_1^T, \ldots, a_m^T$. Suppose that $\mathrm{rank}(A) = k < m$ and that the rows $a_{i_1}^T, \ldots, a_{i_k}^T$ are linearly independent. Consider the polyhedron

$$Q = \{x : a_{i_1}^T x = b_{i_1}, \ldots, a_{i_k}^T = b_{i_k},\ x \geq 0\}$$

Then $Q = P$.

Proof. See [1, Theorem 2.5]. □

Notice that the polyhedron Q in the theorem above is in standard form: namely, $Q = \{x : Dx = f, x \geq 0\}$, where D is a $k \times n$ submatrix of A, with rank equal to k, and f is a k-dimensional subvector of b. We conclude that as long as the feasible set is nonempty, a linear programming problem in standard form can be reduced to an equivalent standard form problem.

Example 3.6. Consider the (nonempty) polyhedron defined by the constraints

$$\begin{aligned} 2x_1 + x_2 + x_3 &= 2 \\ x_1 + x_2 &= 1 \\ x_1 + x_3 &= 1 \\ x_1, x_2, x_3 &\geq 0 \end{aligned}$$

Observe that $\mathrm{rank}(A) = 2$, since the last two rows $(1, 1, 0)$ and $(1, 0, 1)$ are linearly independent, but the first row is equal to the sum of the other two rows. So the first constraint is redundant and after it is removed, we still have the same polyhedron.

4 Degeneracy

With the definition of a basic solution that we are using, we must have n linearly independent active constraints. This allows for the possibility that the number of active constraints is greater than n. Of course, in n dimensions, no more than n of them can be linearly independent. In this case, we say we have a **degenerate basic solution**. In other words, at a degenerate basic solution, the number of active constraints is greater than the minimum necessary.

Definition 4.1 (degenerate basic solution)**.** A basic solution $x \in \mathbb{R}^n$ is said to be degenerate if more than n of the constraints are active at x.

Example 4.2. Consider the polyhedron P defined by the constraints

$$\begin{aligned} x_1 + x_2 + 2x_3 &\leq 8 \\ x_2 + 6x_3 &\leq 12 \\ x_1 &\leq 4 \\ x_2 &\leq 6 \\ x_1, x_2, x_3 &\geq 0 \end{aligned}$$

The vector $x = (2, 6, 0)$ is a nondegenerate basic feasible solution, because there are exactly three active and linearly independent constraints, namely $x_1 + x_2 + 2x_3 \leq 8$, $x_2 \leq 6$, and $x_3 \geq 0$. The vector $x = (4, 0, 2)$ is a degenerate basic feasible solution, because there are four active constraints, three of them linearly independent, namely $x_1+x_2+2x_3 \leq 8$, $x_2+6x_3 \leq 12$, $x_1 \leq 4$, and $x_2 \geq 0$.

For standard form polyhedra, we have the following definition of degeneracy:

Definition 4.3. Consider the standard form polyhedra $P = \{x \in \mathbb{R}^n : Ax = b,\ x \geq 0\}$ and let x be a basic solution. Let m be the number of rows A. The vector x is a degenerate basic solution if more than $n - m$ of the components of x are zero.

Essentially what is happening in degeneracy is that we pick a basic solution by picking n linearly independent constraints to be satisfied with equality, and we realize that certain other constraints are also satisfied with equality.

Observation 1. Degeneracy is not necessarily representation independent; it may depend on the particular representation of a polyhedron. For example, consider a nondegenerate BFS x^* of a standard form polyhedron $P = \{x : Ax = b,\ x \geq 0\}$, where A is $m \times n$. We know by Theorem 3.1, $n - m$ of the variables x_i^* are equal to zero. Let us now represent P in the form

$$P = \{x : Ax \geq b,\ -Ax \geq -b,\ x \geq 0\}.$$

Now at the BFS x^*, we have $n-m$ variables set to zero and an additional $2m$ inequality constraints satisfied with equality. We therefore have $n+m$ active constraints and x^* is degenerate. It follows that every BFS is degenerate under the second representation.

Still, it can be shown that if a BFS is degenerate under one particular standard form representation, then it is degenerate under every standard form representation of the same polyhedron.

5 Existence of extreme points

Definition 5.1 (containing a line)**.** A polyhedron $P \subseteq \mathbb{R}^n$ **contains a line** if there exists a vector $x \in P$ and a nonzero vector $d \in \mathbb{R}^n$ such that $x + \lambda d \in P$ for all scalars λ.

Theorem 5.2. Suppose the polyhedron $P = \{x \in \mathbb{R}^n : a_i^T x \geq b,\ i = 1, \dots, m\}$ is nonempty. Then, TFAE:

1. The polyhedron P has at least one extreme point.

2. The polyhedron P does not contain a line.

3. There exist n vectors out of the family $a_1, \dots, a_n$, which are linearly independent.

Proof. See [1, Theorem 2.6] □

Note. A bounded polyhedron does not contain a line. Similarly, the positive orthant does not contain a line. Since a polyhedron in standard form is contained in the positive orthant, it does not contain a line either.

Corollary 5.3. Every nonempty bounded polyhedron and every nonempty polyhedron in standard form has at least one basic feasible solution.

6 Optimality of extreme points

Theorem 6.1. Consider the linear programming problem of minimizing $c^T x$ over a polyhedron P. Suppose that P has at least one extreme point and that there exists an optimal solution. Then, there exists an optimal solution which is an extreme point of P.

Proof. Let Q be the set of all optimal solutions, which we have assumed to be nonempty. Let P be of the form $P = \{x \in \mathbb{R}^n : Ax \geq b\}$ and let v be the optimal value of the cost $c^T x$. Then $Q = \{x \in \mathbb{R}^n : Ax \geq b,\ c^T x = v\}$, which is clearly a polyhedron as well. Since $Q \subseteq P$, and since P contains no lines (Theorem 5.2), Q contains no lines either. Therefore, Q has an extreme point. Let x^* be an extreme point of Q. We will show that x^* is also an extreme point of P. Suppose for contradiction that x^* is not an extreme point of P. Then, there exists $y, z \in P$ distinct from x^*, and some $\lambda \in [0, 1]$ such that $x^* = \lambda y + (1 - \lambda) z$. It follows that $v = c^T x^* = \lambda c^T y + (1 - \lambda) c^T z$. Furthermore, since v is the optimal cost, $c^T y \geq v$ and $c^T z \geq v$. This implies that $c^T y = c^T z = v$ and therefore, $z, y \in Q$. But this is a contradiction because x^* is an extreme point of Q. It follows that x^* is an extreme point of P. In addition, since x^* belongs to Q, it is optimal. □

The above theorem applies to polyhedra in standard form, as well as to bounded polyhedra, since they do not contain a line.

The next result is stronger than Theorem 6.1. It shows that the existence of an optimal solution can be taken for granted, as long as the optimal cost is finite.

Theorem 6.2. Consider the linear programming problem of minimizing $c^T x$ over a polyhedron P. Suppose that P has at least one extreme point. Then, either the optimal cost is equal to $-\infty$, or there exists an extreme point which is optimal.

Proof. The proof idea is very similar to that of Theorem 6.1. See [1, Theorem 2.8]. $\square$

For a general linear programming problem, if the feasible set has no extreme points, then Theorem 6.2 does not apply directly. On the other hand, any linear programming problem can be transformed into an equivalent problem in standard form. This establishes the following corollary.

Corollary 6.3. Consider the linear programming problem of minimizing $c^T x$ over a nonempty polyhedron. Then, either the optimal cost is equal to $-\infty$ or there exists an optimal solution.

References

[1] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific, 1997.