Polytopes and the simplex method

Bill Casselman
University of British Columbia
cass@math.ubc.ca

A region in Euclidean space is called **convex** if the line segment between any two points in the region is also in the region. For example, disks, cubes, lines, single points, and the empty set are convex but circles, doughnuts, and coffee cups are not.

Convexity is the core of much mathematics. Whereas arbitrary subsets of Euclidean space can be extremely complicated, convex regions are relatively simple. For one thing, any closed convex region is the intersection of all affine half-spaces $f \geq 0$ containing it. One can therefore try to approximate it by the intersection of finite sets of half-spaces—in other words, by **polytopes**.

How to describe a convex set? The major consideration in the theory of convex regions is that there are two complementary ways commonly used to do it: as the convex hull of a set of points, or as the set of points satisfying a given set of affine inequalities. For example, the unit disk is the convex hull of the unit circle, but it is also the set of points (x, y) such that

$$x \cos\theta + y \sin\theta \leq 1$$

for all $0 \leq \theta < 2\pi$. The central computational problem involving convex sets is to transfer back and forth between these two modes. The point is that in solving problems, sometimes one mode is convenient and sometimes the other.

In the first part of this essay I'll discuss general convex bodies, and then go on to prove the very basic result that the convex hull of any finite set of points can also be specified in terms of a finite set of affine inequalities. In this, I shall follow [Weyl:1935]. This simple result has many consequences.

As far as I know, convex bodies were first systematically examined by Minkowski, who used them in spectacular ways in [Minkowski:1910]. [Weyl:1935] is a pleasant account of the basic facts, and I follow his exposition in the first part.

Defining a polytope by affine inequalities and defining it as a convex hull are dual modes of specification. Thus we arrive immediately at the most basic computational problem concerned with convex polytopes—*given a polytope specified in one of these ways, to find an explicit specification in the other*. More specifically, if a polytope is specified as the intersection of a finite number of half-spaces $f_i(x) \geq 0$, how can one find its vertices and extremal rays? If a region is specified as the convex hull of a finite set of points, how can one find its vertices among this finite set? How can one find a finite number of half-spaces whose intersection is the polytope?

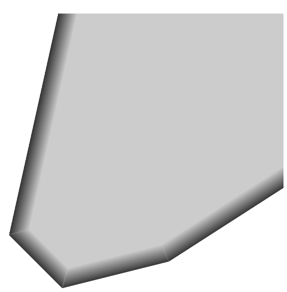
There are several other questions one might pose. A **face** of a polytope is any affine polytope on its boundary. A **wall** is a proper face of maximal dimension.

- Is the region bounded?
- What is its dimension (taken to be -1 if the region is empty)?
- What is a minimal set of affine inequalities defining it? Or, in other words, what are its walls?
- What is the maximum value of a linear function on the polytope (which might be infinite)?
- How to describe the face where a linear function takes its maximum value?
- What is its facial structure?
- What is the dual cone of linear functions bounded on the polytope?

The answer to some of these can be extremely complex. Answering them reduces in general to answering a few basic ones. For example, to find out whether one region is contained in the other it suffices that you know the characterization of one in terms of affine inequalities, but the other in terms of extremal points. Finding the intersection of two convex bodies is trivial if they are specified by affine inequalities, and finding the convex hull of their union is easiest if the regions are given as convex hulls.

These problems are well known to arise in applications of mathematics. Among other things, convexity is a crucial feature in the way we perceive things in 3D, and hence problems related to convexity arise in computer graphics. But in fact they also arise in trying to understand the geometry of configurations in high dimensions, which comes up surprisingly often even in the most theoretical mathematical investigations. For one thing, acquiring a picture of an object in high dimensions, or navigating regions of high dimensional space, often amounts to understanding the structure of convex regions.

Answering these questions can be daunting, since the complexity of problems increases drastically as dimension increases. There are now many algorithms known to answer such questions, but for most practical purposes the one of choice is still the **simplex method**. This is known to require time exponential in the number of inequalities for certain somewhat artificial problems. But on the average, as is shown in a series of papers beginning apparently with [Smale:1983] and including most spectacularly [Spielman-Teng:2004], it works well. It is also more flexible than competitors. A polytope is by definition the intersection of a finite number of half-spaces in a Euclidean space. Thus cubes are polytopes but spheres are not. The simplest example, and the model in some sense for all others, is the coordinate octant $x_i \ge 0$. Another simple example is the bounded simplex $x_i \ge 0$, $\sum x_i \le 1$. This essay will try to explain how to deal with polytopes in practical terms. The literature on this is extensive, and I'll not be innovative. Since linear programming is commercially applicable, little of the literature is aimed at quite the same audience as this note is. (But compare §3.2 in [Ziegler:2007], titled 'Linear programming for geometers', or [Knuth:2005].)



What is new here? Nothing much. One thing that might be a little new is my fussiness in the discussion of affine structures. My main reference for the simplex method has been the well known text [Chvátal:1983], but my notation is more geometric than his. In this exposition there are no tableaux, no slack variables, and no dictionaries. I use only terminology familiar to all mathematicians—bases, coordinate systems, matrices. The audience for expositions of linear programming is not, by and large, made up of mathematicians, and the standard notation in linear programming is, for most mathematicians, rather bizarre. I follow more geometric conventions.

Contents

Part I. Geometry

1. Affine spaces

If $\boldsymbol{V}$ is a vector space of dimension n, an **affine space** V associated to it is a smooth space on which it acts simply transitively. It is a principal homogeneous space for $\boldsymbol{V}$. Given a choice of origin P, an affine space becomes isomorphic to $\boldsymbol{V}$: $v \mapsto P+v$. Without further data, it doesn't make sense to add elements of V, or to multiply one by a scalar, but for P, Q in V the difference $Q-P$ is an element of $\boldsymbol{V}$. The fundamental operations on V are the weighted means

$$\mu_s(P,Q) = (1-s)P + sQ = P + s(Q-P)\,.$$

The point is that convex subsets of V are well defined by the affine structure, and indeed affine spaces are the natural structures into which convex sets embed. Affine spaces occur naturally—if $\boldsymbol{V}$ is a hyperplane through the origin in a vector space, then any parallel hyperplane is an affine space with respect to it, but not generally a vector space.

DUALITY.

1.1. Lemma. *Suppose f to be an $\mathbb{R}$-valued function on V. The following are equivalent:*

(a) *the function $f(P+v) - f(P)$ is a linear function on $\boldsymbol{V}$ independent of P;*
(b) *we have*

$$f((1-s)P + sQ) = (1-s)f(P) + sf(Q)$$

for all P, Q in V, s in $\mathbb{R}$;
(c) *for any finite set (P_i) in V*

$$f\left(\sum c_i P_i\right) = \sum c_i f(P_i)$$

whenever $\sum c_i = 1$.

In these circumstances, f is said to be an **affine function** on V. The function specified in (a) is the **gradient** ∇_f of f. It is characterized by the equation

$$f(Q) - f(P) = \nabla_f(Q-P)\,. \qquad \textbf{(1.2)}$$

Proof. Only one point might not be trivial. Condition (b) implies (c) because

$$\sum\nolimits_{i\le n} c_i P_i = (1-c_n)\left(\frac{\sum_{i\le n-1} c_i P_i}{1-c_n}\right) + c_n P_n\,.$$

A choice of origin in V makes it isomorphic to $\boldsymbol{V}$, and then every function satisfying these conditions is of the form $f + c$ where f is a linear function on $\boldsymbol{V}$ and c is a constant. Let $\mathcal{V}$ be the space of such functions. It contains the constants as a canonical subspace. What is the quotient? According to Lemma 1.1 we have a short eact sequence

(1.3)
$$0 \longrightarrow \mathbb{R} \xrightarrow{\alpha} \mathcal{V} \xrightarrow{\nabla} \widehat{\boldsymbol{V}} \longrightarrow 0\,.$$

Here $\widehat{\boldsymbol{V}}$ is the linear dual of the vector space $\boldsymbol{V}$. This sequence splits. Suppose given a point P in V. The function $f \mapsto f(P)$ is a linear surjection from $\mathcal{V}$ to $\mathbb{R}$ commuting with the inclusion of $\mathbb{R}$. What is the corresponding map backwards on the right? Given a function φ in $\widehat{\boldsymbol{V}}$, the function

$$f(Q) = \varphi(Q - P)$$

is an affine function on V.

The space $\widehat{\mathcal{V}}$ is the linear dual of $\mathcal{V}$. The dual of the short exact sequence (1.3) is hence

$$0 \longrightarrow \boldsymbol{V} \xrightarrow{\widehat{\nabla}} \widehat{\mathcal{V}} \xrightarrow{\widehat{\alpha}} \mathbb{R} \longrightarrow 0\,.$$

Here I identify $\mathbb{R}$ with its dual through the pairing xy. The injection on the left takes v in $\boldsymbol{V}$ to the linear function on $\mathcal{V}$ that takes an affine function f to the constant $f(P + v) - f(P)$, which we know to be independent of P.

More interesting is the canonical embedding

$$\iota\colon\, V \longrightarrow \widehat{\mathcal{V}}\,,$$

which takes P to the linear function $f \mapsto f(P)$. The image of V in $\widehat{\mathcal{V}}$ is the hyperplane $\widehat{\alpha}^{-1}(1)$. If we have chosen an identification of V with $\boldsymbol{V}$, then $\widehat{\mathcal{V}}$ may be identified with $\boldsymbol{V} \oplus \mathbb{R}$, and $\iota(V)$ with the points $(v, 1)$.

The points of V in $\widehat{\mathcal{V}}$ are interpreted as locations, while the points of $\boldsymbol{V}$ in the same space are interpreted as displacements. This agrees with the usual conventions of physics!

COORDINATES. Affine coordinates on an affine space (say of dimension n) are determined by an affine basis, which is to say a choice of origin in V together with a basis of $\boldsymbol{V}$. A change of coordinates from (x_i) to (y_i) is defined by a substitution

$$x = Ay + a$$

in which A is an $n \times n$ matrix and a a column vector of dimension n. The pair A and a are canonical. This corresponds to a canonical right action on affine bases by certain matrices in $\mathrm{GL}_{n+1}(\mathbb{R})$:

$$[\,u \quad P\,]\begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix} = [\,u \cdot A \quad P + u \cdot a\,]\;.$$

Thus A is canonically in GL_n and a in $\mathbb{R}^n$ is the vector of coordinates of the new origin $(y = 0)$ in the old coordinate system. This equation can also be expressed as

$$\begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix}\begin{bmatrix} y \\ 1 \end{bmatrix}\,.$$

Composition of coordinate changes becomes, conveniently, matrix multiplication.

The appearance of $\mathbb{R}^{n+1}$ should not be a surprise, considering the canonical embedding of V as a hyperplane in $\widehat{\mathcal{V}}$, which has dimension $n + 1$.

Write affine functions as row vectors, so

$$f(x) = [\,\nabla_f \;\; c\,] \begin{bmatrix} x \\ 1 \end{bmatrix} ,$$

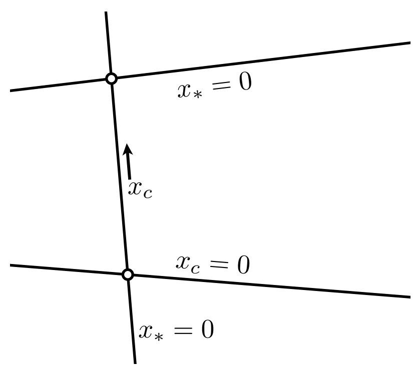
in which ∇_f is the linear part of f, and the vector representing it is made up of the coordinates of f in the current coordinate system. Then

$$[\,\nabla_f \;\; c\,] \begin{bmatrix} x \\ 1 \end{bmatrix} = [\,\nabla_f \;\; c\,] \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y \\ 1 \end{bmatrix}$$

so that its expression in y coordinates is

$$[\,\nabla_f \;\; c\,] \begin{bmatrix} A & a \\ 0 & 1 \end{bmatrix} .$$

PIVOTING. Affine coordinate changes are the basic processes in the simplex method. The essence of the method is to move from one vertex of a polytope to another. (Actually, this isn't quite what it does, as we shall see, but it is what it would like to do. What it actually does is to move from one coordinate system to another, and it might do this without changing the vertex itself.) That is to say, it involves a change of coordinates by a particular kind of affine transformation known as a **pivot**.



In this, we are given a coordinate system (x_i) and some affine functions. We are going to swap one of these affine functions, say F, for one of the variables x_c. At the same time, we want to express the other given affine functions in terms of the new variables. This reduces to the following problem: we are given

$$x_* = \sum_1^n a_j x_j + a_{n+1} ,$$

which we want to replace x_c. To do this, we write

$$x_c = (1/a_c)\Big(x_* - \sum_{j \neq c} a_j x_j - a_{n+1}\Big)$$

Of course, we must assume that $a_c \neq 0$.

Suppose now that we are given another affine function

$$f = \sum\nolimits_j b_j x_j + b_{n+1}$$

which we wish to convert to an expression with F replacing x_c. The new coordinates of x_* itself are the same as the old ones of x_c, since it just replaces x_c as a basis element. But then

$$\begin{aligned} f &= (b_c/a_c)\Big(x_* - \sum_{j \neq c} a_j x_j - a_{n+1}\Big) + \sum_{j \neq c} b_j x_j + b_{n+1} \\ &= (b_c/a_c)x_* + \sum\nolimits_{j \neq c} \big(b_j - (a_j/a_c)b_c\big)x_j + \big(b_{n+1} - (a_{n+1}/a_c)b_c\big) . \end{aligned}$$

In effect, we are

(a) dividing the c-th entry of each f by a_c,
(b) for each $j \neq c$ (including $j = n+1$), replacing the j-th entry b_j of f by $b_j - pb_c$, with $p = a_j/a_c$.

We do this to all of the given affine functions, which we arrange as rows of a matrix. This amounts to simple column operations. Since column operations are equivalent to multiplication by certain $(n+1) \times (n+1)$ matrices on the right, this is consistent with earlier remarks on coordinate change. Explicitly, the matrix by which we finally multiply is that obtained by performing the same column operations on the identity matrix I_{n+1}. These effects accumulate.

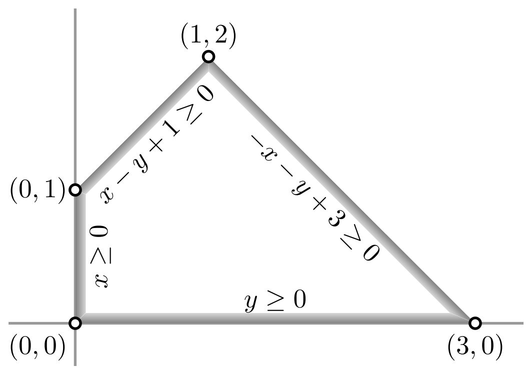
Example. Start with the inequalities

$$\begin{aligned} x &\geq 0 \\ y &\geq 0 \\ -x - y + 3 &\geq 0 \\ -x + y - 1 &\geq 0 \end{aligned}$$

associated to the vertex $(0,0)$ of the region below. This corresponds to the matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 3 \\ -1 & 1 & -1 \end{bmatrix}$$

Pivot to the corner $(3,0)$, changing the variable x to $x_* = -x - y + 3$:



The column operations amount to:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 3 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 3 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 2 & -4 \end{bmatrix}$$

The inequalities expressed in the new coordinate system are

$$\begin{aligned} x_* - y + 3 &\geq 0 \\ y &\geq 0 \\ x_* &\geq 0 \\ -x_* + 2y - 4 &\geq 0\,. \end{aligned}$$

2. Duality

It is a little troublesome to say much about arbitrary convex regions. For example, the union of an open disk with any subset of its circular boundary is convex. But closed convex sets behave somewhat less erratically, and the following elementary result is therefore useful as well as elementary.

2.1. Proposition. *The closure of a convex region is convex.*

SEPARATION THEOREMS. I begin now with some observations on arbitrary closed convex regions. The following result is the basic fact:

2.2. Proposition. *Given a closed convex region Ω and a point P not in it, there exists an affine function f such that $f(\Omega) \ge 0$ and $f(P) < 0$.*

Geometrically, this says that there exists a hyperplane strictly separating P from Ω.

This result remains true (in the guise of the Hahn-Banach Theorem) even for locally convex topological vector spaces of infinite dimension. The proof I give is valid for Hilbert spaces.

Proof. Choose P as origin and assign V a Euclidean metric. It is to be proved that there exists an affine function $f(x)$ such that $f(P) = 0$, $f(\Omega) \ge c > 0$.

Since Ω is closed, there exists a point Q in Ω at minimal distance from P. I claim that the hyperplane passing through Q and perpendicular to the line through P and Q will do. The equation of this hyperplane is

$$x \bullet Q = Q \bullet Q\,,$$

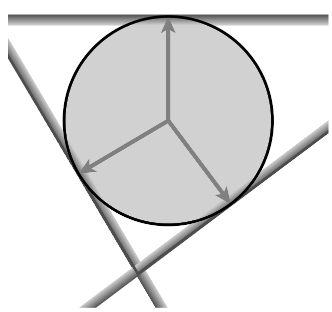
so that we must prove that $R \bullet Q - Q \bullet Q = (R - Q) \bullet Q \ge 0$ for all R in Ω. But if R lies in Ω then so does $Q + t(R - Q)$ for all t in $[0, 1]$, and by definition of Q its length is at least that of Q. This implies that

$$2t(R - Q) \bullet Q + t^2(R - Q) \bullet (R - Q) \ge 0$$

for all small t, which implies in turn that $(R - Q) \bullet Q \ge 0$. ▮

2.3. Corollary. *Any closed convex set in V is the intersection of all half-spaces containing it.*

For example, the unit disk in the plane is the intersection of all half-planes $x \cos\theta + y \sin\theta \le 1$ as θ ranges over $[0, 2\pi)$.



There is a special case worth calling attention to. If U is a vector space, a (homogeneous) **cone** in U is a subset stable under multiplication by $c \ge 0$.

2.4. Corollary. *If Ω is a closed convex cone in the vector space **V** then it is the intersection of all linear half-spaces containing it.*

Proof. If f is an affine function, let ∇_f be its linear gradient $f - f(0)$. According to Corollary 2.3, the corollary will follow from the claim that $f(\Omega) \ge 0$ if and only if (i) $\nabla_f(\Omega) \ge 0$ and (ii) $f(0) \ge 0$. Since

$$f(v) = \nabla_f(v) + f(0)\,,$$

one way is trivial. As for the other, we have

$$f(cv) = c\nabla_f(v) + f(0)\,.$$

If $f(\Omega) \ge 0$ then $f(0) \ge 0$ since 0 is contained in Ω. If v lies in Ω then cv also lies in Ω for all $c > 0$. Taking c off to ∞, we deduce that $\nabla_f(v) \ge 0$.

LINEAR DUALS. There are two notions of duality related to convex sets. If Ω is a closed convex cone of a vector space U, its **linear dual** $\Omega_0^\vee$ is the set of all linear functions f on U such that $f(\Omega) \ge 0$. It is a closed convex cone in $\widehat{U}$.

For example, if (e_i) is a basis of V, $(\widehat{e}_i)$ the dual basis, and Ω the cone generated by the e_i for $1 \le i \le d$, then $\Omega_0^\vee$ is the set of all sums

$$\sum_{i=1}^{d} c_i \widehat{e}_i + \sum_{d+1}^{n} c_i \widehat{e}_i \quad (c_i \ge 0 \text{ for } i \le d, \text{ arbitrary for } i > d)\,.$$

Note that the space of linear functions vanishing on Ω is the same as the space of translations in the dual space leaving $\Omega_0^\vee$ invariant.

The linear dual has also a linear dual. There is a canonical map from a cone Ω to the linear dual of $\Omega_0^\vee$. The simplest statement of duality is this:

2.5. Proposition. *If Ω is a closed convex cone, the canonical map from Ω to the dual of $\Omega_0^\vee$ is a bijection.*

Proof. It is an injection because the canonical map from V to the dual of $\widehat{V}$ is an isomorphism. That it is a surjection is a consequence of Corollary 2.4.

2.6. Corollary. *The map taking Ω to $\Omega_0^\vee$ is a bijection between closed convex cones in V and those in $\widehat{V}$.*

AFFINE DUALS. If Ω is a closed convex subset of V, its **affine dual** $\Omega^\vee$ is the set of all affine functions f such that $f(\Omega) \ge 0$. It is a closed convex cone in $\mathcal{V}$, which I recall to be the space of affine functions on V. The extreme cases are (i) $\Omega = \emptyset$ with $\Omega^\vee = \mathcal{V}$, and (ii) $\Omega = V$ with $\Omega^\vee = \{0\}$.

If C is a closed cone in $\mathcal{V}$, I define its **point dual** $C^\vee$ to be the subset of all x in V such that $f(x) \ge 0$ for all f in C. It is closed and it is convex as well, according to Lemma 1.1. Any x in Ω is of course a point in the point dual of $\Omega^\vee$.

The following is just a rewording of Corollary 2.3.

2.7. Proposition. *If Ω is a closed convex subset of V, then the embedding of Ω into the point dual of $\Omega^\vee$ is a bijection.*

Duality for affine convex sets is not as simple as it is for cones, since the dual of an affine convex set is not again an affine convex set, but rather a cone. However, this is not difficult to deal with. I recall that $\widehat{\mathcal{V}}$ contains canonical copies of $\boldsymbol{V}$ and V, the latter being the affine space $\widehat{\alpha}^{-1}(1)$. The set Ω may be identified with a closed convex subset of this copy of V. Define $\overline{\Omega}$ to be the closure of the cone over this copy of Ω. If Ω is bounded, this will be in fact the cone over Ω, but if Ω is unbounded this will no longer be the case. The closure of the cone will contain in addition points of $\boldsymbol{V}$.

I recall that a ray in any affine space is the set of all $P + tv$ with v in $\boldsymbol{V}$, $t \ge 0$.

2.8. Lemma. *Suppose ρ to be a ray in $\boldsymbol{V}$. The following are equivalent:*

(a) the ray ρ is contained in $\overline{\Omega}$;
(b) for some point P in Ω, the ray $P + \rho$ is contained in Ω;
(c) for any point P in Ω, the ray $P + \rho$ is contained in Ω.

2.9. Proposition. *The subset $\overline{\Omega}$ in $\widehat{\mathcal{V}}$ is the same as the linear dual of $\Omega^\vee$.*

I call the set $\overline{\Omega}$ the **homogenization** of Ω in the copy of V in $\widehat{\mathcal{V}}$. The requirement of closure is necessary, as is shown by the example of $[0, \infty)$ in $\mathbb{R}$, whose homogenization in $\mathbb{R}^2$ is the region $x \ge 0$, $y \ge 0$.

If f is an affine function on V, it extends to a unique linear function $\overline{f}$ on $\widehat{\mathcal{V}}$.

2.10. Proposition. *The linear dual of the homogenization $\overline{\Omega}$ is made up of the $\overline{f}$ for f in $\Omega^\vee$.*

TRANSLATION INVARIANCE AND SUPPORT. It may happen that a convex region Ω is invariant under translation by vectors in some linear subspace of $V_\circ$. Define $\mathrm{Inv}(\Omega)$ to be the linear space of all such vectors, the invariance subspace associated to Ω. The following is immediate.

2.11. Lemma. *If Ω is a closed convex set, then $\mathrm{Inv}(\Omega) \neq \{0\}$ if and only if Ω contains a line.*

2.12. Proposition. *If Ω is any closed convex set of V, then $\mathrm{Inv}(\Omega)$ is the same as $\mathrm{Inv}(\overline{\Omega})$.*

A closed convex set Ω is called **nondegenerate** if $\mathrm{Inv}(\Omega) = \{0\}$. If Ω is a closed convex region with $U = \mathrm{Inv}(\Omega)$, its image Ω/U in V/U will be nondegenerate and Ω will be isomorphic to $U \times (\Omega/U)$. The analysis of convex regions thus reduces easily to that of nondegenerate ones.

If Ω is any convex subset of V, its **affine support** is the smallest affine subspace containing it. I'll call the dimension of that subspace the **dimension** of Ω. The dimension will be n if and only if its interior is non-empty.

2.13. Proposition. *If the cone Ω is nondegenerate of dimension n, then so is $\Omega_0^\vee$.*

There is a strong relationship between the support of a convex set and that of the invariance subspace of its dual. The support of Ω is the sub

2.14. Proposition. *The invariance subspace of $\Omega^\vee$ is the same as the space of affine functions vanishing on Ω.*

Let π be the canonical projection from $\mathcal{V}$ to $\widehat{\boldsymbol{V}}$.

2.15. Proposition. *We have*

$$\pi(\Omega^\vee)^\perp = \mathrm{Inv}(\Omega)\,.$$

3. Convex hulls

3.1. Proposition. *If Ω is a closed convex set, then a point P in Ω is on its boundary if and only if there exists an affine function f with $f(P) = 0$, $f(\Omega) \geq 0$.*

Proof. One way is trivial.

As for the other, suppose P to be in the boundary of Ω. Choose an origin and a Euclidean metric on V. Since P is on the boundary, there exists a sequence of points P_i converging to P but not in Ω. For each of these there exists an affine function $f_i(x) = v_i \bullet x + c_i$ with $\|v_i\| = 1$ and $f_i(\Omega) \geq 0$, $f_i(P_i) = 0$. One can choose a subsequence so that the v_i convergem say to v, and then the c_i also converge, say to c. Then $f(x) = v \bullet x + c$ satisfies the conditions $f(\Omega) \geq 0$, $f(P) = 0$.

The set of all such f is a convex cone called the **tangent cone** at P.

COMPACT HULLS. The **convex hull** of a set C is the set of all weighted means of elements of C with non-negative weights. It is the smallest convex set containing C.

The following first appeared in [Steinitz:1913] (p. 155), after an earlier special case had been dealt with in [Carathéodory:1907]. Both these works were primarily interested in questions about conditional convergence!

3.2. Proposition. (Carathéodory's Theorem) *Suppose X to be a nonempty subset of $\mathbb{R}^n$. If $y \neq 0$ is a non-negative linear combination of points in X, it may be expressed as a positive linear combination of linearly independent points of X.*

Proof. The proof is an algorithm.

• Suppose $v = \sum c_i x_i$ with $c_i > 0$, x_i in X. Let N be the number of x_i. If the x_i are linearly independent, exit.

Otherwise, there exists an equation

$$\sum a_i x_i = 0\,.$$

Negating if necessary, we may assume one $a_i > 0$. Then for any b

$$v = \sum (c_i - b\,a_i)x_i\,.$$

The line $y = c_i - xa_i$ in $\mathbb{R}^2$ starts off at $(0, c_i)$, and has slope $-a_i$. At least one of them crosses the positive x-axis; suppose the first one crosses at $(b, 0)$. Then all $c_i - ba_i$ are non-negative, and at least one is 0. This gives an expression for y involving a smaller number of x_i. Loop to •.

The procedure cannot go on forever, since $y \neq 0$ and a single element of X is a linearly independent set.

3.3. Corollary. *In $\mathbb{R}^n$, any element in the convex hull of a finite set X is in the convex hull of some subset of at most $n + 1$ elements.*

In other words, every element in the convex hull of X is contained in a simplex whose vertices are in X.

Proof. Apply the previous result to the homogenization of X in $\mathbb{R}^{n+1}$.

3.4. Corollary. *The convex hull of a compact set C is compact.*

In infinite dimensional topological vector space, it is not even necessary that the convex hull of a compact set be closed.

Proof. It is the image of the compact set $C^{n+1} \times \Delta_{n+1}$ with respect to a continuous map, where

$$\Delta_{n+1} = \{(c_i) | c_i \geq 0,\ \sum c_i = 1\}\,.$$

EXTREMAL POINTS. An **extremal point** of a closed convex set Ω is a point that does not lie in the interior of a line segment between two points of Ω.

3.5. Lemma. *Suppose H to be a tangent hyperplane of Ω. Then an extremal point of $H \cap \Omega$ must also be an extremal point of Ω.*

Proof. Suppose H to be the hyperplane $f(P) = 0$. $x = (1 - t)y + tz$ with y, z in Ω, $t > 0$. But then $f(x) = (1-t)f(y)+tf(z)$. If either $f(y)$ or $f(z)$ is $> c$ then the other must be $< c$, which is not possible. So y and z both lie in $H \cap \Omega$.

If $\mathrm{Inv}()\Omega) \neq \{0\}$, then Ω will not have any extremal points. But this the only case in which there are none.

3.6. Proposition. *A closed convex set Ω has an extremal point if and only if $\mathrm{Inv}(\Omega) = \{0\}$.*

Proof. By induction. It is clear for $n = 0$ or 1.

Suppose P in Ω. Any line through P must exit Ω somewhere, say at Q. Since Q is on the boundary, there exists a hyperplane containing Q with Ω on one side. By induction, $H \cap \Omega$ possesses an extremal point, and by Lemma 3.5 it is extremal for Ω.

In one case, there are lots of extremal points, in a strong sense.

3.7. Proposition. *Any bounded, closed, convex set subset Ω of $\mathbb{R}^n$ is the convex hull of its extremal points.*

Proof. If $n = 0$ or $n = 1$ this is immediate. Otherwise, suppose $n \geq 2$, and suppose given P in Ω. Let ℓ be a line through P. If P is extremal, we are through. Otherwise P lies in the interior of a line segment between two points of Ω. This line must exit Ω in either direction, since Ω is bounded, say at points $P_0 \neq P_1$. Each of these is a boundary point, so by Proposition 3.1 there exist hyperplanes H_i with Ω sandwiched in between them and and H_i containing P_i. By induction, P_i is in the convex hull of $H_i \cap \Omega$. Apply Lemma 3.5.

3.8. Proposition. *Suppose Ω to be a closed convex set. A point P of Ω is extremal if and only if there exists a hyperplane H, with Ω on one side, such that $H \cap \Omega = \{P\}$.*

Proof. The only point to be verified is that if P is extremal then there exists a hyperplane H such that $H \cap \Omega = \{P\}$.

Now to prove the Proposition. Let S be the unit sphere around P, $\Sigma = S \cap \Omega$. Then Σ is compact. Since $\mathrm{Inv}(\Omega) = \{0\}$ the set $\Omega \notin \{P\}$ is convex, so that the convex hull $\overline{\Sigma}$ of Σ is also disjoint from P. Hence there exists an affine function f such that $f(P) = 0$, $f(\overline{\Sigma}) > 0$. But Ω is in the cone generated by Σ generates Ω, so that $f(x) > 0$ for $x \neq P$ in Ω.

If ρ is a ray contained in the convex set Ω, it is **extremal** if any line segment that contains a point of ρ lies entirely in ρ.

3.9. Corollary. *Any nondegenerate closed convex set is the convex hull of its extremal points and rays.*

Proof. Let $\overline{\Omega}$ be the cone over Ω in $V \oplus \mathbb{R}$. The origin is a vertex on it. There then exists a hyperplane H with $H \cap \overline{\Omega} = \{0\}$. Any translate of H will intersect Ω in a closed bounded convex set. It will be the convex hull of its extremal points. Those that do not lie on the hyperplane $x_{n+1} = 0$ will correspond to extremal points of Ω, while those on that hyperplane will be parallel to extremal rays of Ω.

3.10. Corollary. *Suppose Ω to be a closed convex cone of dimension n. Then it is the intersection of all regions $f(P) \geq 0$ as f ranges over generators of extremal rays of $\Omega^\vee$.*

Proof. Under the asumption that $\dim(\Omega) = n$, $\Omega^\vee$ will be a proper cone. It will then possess compact slices, and one of these will be the convex hull of its extremal points. Therefore any f with $f(\Omega) \geq 0$ will be in the cone generated by these. Therefore the intersection of all $f \geq 0$ for f in $\Omega^\vee$ is the same as that with f is extremal.

CONVEX SETS ON SPHERES. In order to understand unbounded convex sets, it is convenient to relate them to convex sets on spheres.

If P and Q are two points inside an open hemisphere on $\mathbb{S}^n$, there is a unique shortest arc PQ of a great circle connecting them. A subset Γ of an open hemisphere is said to be convex if PQ is contained in Γ whenever P and Q are in Γ. Let $\overline{\Gamma}$ be the cone over Γ with vertex at the origin. Since Γ is contained in an open hemisphere, there exist hyperplanes through the origin whose intersection with $\overline{\Gamma}$ is just $\{0\}$. I'll say that any hyperplane parallel to one of these is transverse to $\overline{\Gamma}$. Then Γ is convex if and only if every transverse slice of $\overline{\Gamma}$, which is compact, is convex.

What does this have to do with affine convex sets? Suppose C to be a nondegenerate closed cone in the half-space $\alpha \geq 0$ in $\widehat{\mathcal{V}}$. The rays in C are in bijection with a set of points in the upper hemisphere, through the map $x \mapsto x/\|x\|$. The map taking C to its spherical projection is a bijection of nondegenerate cones in the upper half-space with convex set in the sphere whose support is contained in some open hemisphere.

4. Introduction to polytopes

Continue to suppose V to be an affine space, say of dimension n.

A **polytope** in V is a region bounded by a finite number of hyperplanes. It is also the convex hull of a finite number of points and rays. Transferring back and forth between these representations is a primary task in programming. The simplex method starts with a region defined by a finite set of affine inequalities, and finds vertices of that region. The next part of this essay will explain it at length. In this section I'll explain an algorithmic proof of a dual theorem of [Weyl:19345/50]. Suppose X to be any finite set of points in $\mathbb{R}^n$. Assume that it contains at least n linearly independent points. A linear hyperplane is said to be extremal with respect to X if (1) it contains at least $n-1$ linearly independent points of X and (2) all the points of X lie in one of the closed half-spaces determined by it.

4.1. Theorem. *If X is any finite set of $\mathbb{R}^n$ that contains n linearly independent points, then the conical convex hull of X is the set of points x where $f(x) \geq 0$ for all linear f defining extremal hyperplanes.*

The proof will be constructive. It will suggest an algorithm to produce a list of all extremal hyperplanes H_α. Incidentally, if we have such a list $\{\alpha_i\}$, then any weighted sum $\sum c_i\alpha_i$ with each $c_i > 0$ will define a hyperplane transverse to the cone generated by X.

So suppose given a finite set X of points in $\mathbb{R}^n$, and let Ω be the homogeneous convex hull of the set X. We assume X to contain vertices of an open simplex of $\mathbb{R}^n$. It is to be shown that Ω is the intersection of all extremal half-spaces of X. I follow loosely the argument of [Weyl:1935/50].

- It is not necessary to assume that the dimension of the support of X is n, since the general case can be reduced to that by computing this support explicitly. To do this, place the vectors in X as columns in a matrix and then apply elementary row operations to it to make it row-reduced. At the end, certain columns in the matrix will be a basis for the support, and the remaining ones will record coordinates with respect to this basis.

∘ ————- ∘

One half of the theorem is trivial, since Ω is certainly contained in the intersection of all extremal half-spaces. It remains to be shown that if a point lies in the intersection of extremal half-spaces then it lies in Ω. It should be kept in mind that there may be no extremal half-spaces for X. In this case, the theorem asserts that the convex hull of the cone generated by X is all of $\mathbb{R}^n$.

The proof proceeds by induction on $n \geq 1$. If $n = 1$ then the points of X span a non-trivial segment. If this segment contains 0 in its interior, then the homogeneous convex hull is all of $\mathbb{R}^n$, and there are no linear functions cx taking a non-negative value on all of X except 0 itself. Otherwise, the homogeneous convex hull is a half-line with 0 on its boundary, and again the result is plain.

Now assume the Theorem to be true for dimensions up to $n-1$, and suppose X to have dimension n. There are two steps to the argument. The first step either produces an extremal half-space or demonstrates that the homogeneous convex hull of X is all of $\mathbb{R}^n$. The second proves the theorem when it is known that extremal half-spaces exist.

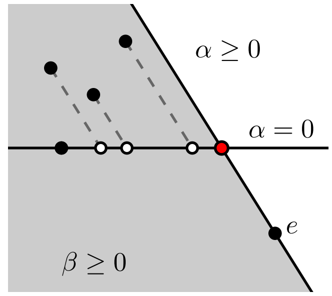
Step 1. Suppose $\alpha = 0$ to be any hyperplane containing $n-1$ linearly independent points of X. Either all points of X lie on one side of this hyperplane, say the side $\alpha \geq 0$, or they do not. If they do, then this hyperplane is extremal, and we can pass on to the second step. Suppose they do not, so that there are points of X on both sides of the hyperplane. A new extremal half-space $\beta \geq 0$ will be found that contains at least one more point of X, so we can set $\alpha = \beta$ and loop.

At this point we have in hand the hyperplane $\alpha = 0$, which is not extremal. Let e be a point of X with $\langle \alpha, e\rangle < 0$. By changing coordinates if necessary, we may assume α to be the coordinate function x_n, and $e = (0, 0, \ldots, -1)$.

- If the hyperplane is specified by a basis, we choose that basis together with e as a new basis, hence a new coordinate system. If it is only given by a list of points in X that lie on it, then a basis can be found through row reduction.

∘ ————- ∘

Let Y be the coordinate projection onto the plane $x_n = 0$ of the points $X^{\geq 0}$ of X with $x_n \geq 0$. The new set Y is non-degenerate of dimension $n-1$, by assumption on α. We may therefore apply induction. Either it possesses an extremal half-space $\langle \beta, x\rangle \geq 0$ with $\langle \beta, x\rangle = b_1x_1 + \cdots + b_{n-1}x_{n-1}$ or it does not. If it does, the half-space $\langle \beta, x\rangle \geq 0$ contains the point e in addition to the points in $X^{\geq}$. The hyperplane $\beta = 0$ contains $n-2$ linearly independent points on the hyperplane $\alpha = 0$ together with e, and also contains $n-1$ linearly independent points, so we may loop, replacing α by β.



In the second case, when Y does not possess an extremal half-space, we can apply the induction assumption. Every point in the hyperplane $\langle \alpha, x \rangle = 0$ must be in the convex hull of Y. Adding positive multiples of e, we see that every point where $x_n \leq 0$ is in the convex hull. Since we have eliminated the case where all points of X lie on one side of $\alpha = 0$, we can find a point with $x_n > 0$. A similar argument implies that every point where $x_n \geq 0$ is also there. We conclude that in this second case no extremal hyperplane exists.

Step 2. At this point we know there exist extremal half-spaces containing Ω. Let $\mathfrak{H}$ be the set of all of them, and let A be a set of linear functions α such that each one of them is defined by some inequality $\alpha \geq 0$. It is to be shown that if $\langle \alpha, p \rangle \geq 0$ for all α then p is in Ω.

Choose (arbitrarily) a point e in the interior of the homogeneous convex hull of $X = \{x_i\}$.

- Possible because we can locate a simplex whose vertices are in X.

Thus $\langle \alpha, e \rangle > 0$ for every α in Δ. Let p be an arbitrary point in the intersection of extremal half-spaces, so that also $\langle \alpha, p \rangle \geq 0$ for all α in Δ. We want to show that p is in Ω. We may assume that $p \neq e$.

The set of all

$$\lambda e + \mu p$$

with λ, $\mu \geq 0$ is a two-dimensional wedge. Any point of the same form with $\lambda < 0$, $\mu \geq 0$ will be in the plane spanned by this wedge, outside this wedge on the far side of the line through p. Since

$$\langle \alpha, p - \lambda e \rangle = \langle \alpha, p \rangle - \lambda \langle \alpha, e \rangle$$

and $\langle \alpha, e \rangle < 0$, the ray of all $q = p - \lambda e$ with $\lambda > 0$ will find itself eventually in the region where $\alpha < 0$. It will cross the hyperplane $\alpha = 0$ when

$$\lambda = \frac{\langle \alpha, p \rangle}{\langle \alpha, e \rangle}\,.$$

If λ is chosen to be the least value of $\langle \alpha, p \rangle / \langle \alpha, e \rangle$ as α ranges over Δ, then

(1) $p = q + \lambda e$ where $\lambda \geq 0$ and e is in the homogeneous convex hull of X;
(2) $\langle \lambda, q \rangle \geq 0$ for all α in Δ;
(3) $\langle \alpha, q \rangle = 0$ for some α in Δ.

In order to prove that p is in the homogeneous convex hull of X, it suffices to show that q is in it.

We want to apply the induction hypothesis to q. Let Y be the set of all x in X on the hyperplane H_α where $\alpha = 0$, which contains q. Because α is an extremal support, there exist at least $n - 1$ linearly independent points in Y, so that it is non-degenerate in $n - 1$ dimensions. Choose coordinates so that $\alpha = x_n$. By the induction assumption, the convex hull of Y is the intersection of all its extremal half-spaces in H_α. That q lies in this intersection will follow from:

4.2. Lemma. *Any extremal hyperplane for $Y = X \cap H_\alpha$ in H_α can be extended to an extremal hyperplane for X in $\mathbb{R}^n$.*

In effect, we are making this hyperplane in H_α a hinge.

Proof of the Lemma. Suppose

$$\langle \beta, y \rangle = b_1 x_1 + \cdots + b_{n-1} x_{n-1} \geq 0$$

be one of the extremal half-spaces for Y in H_α. The function β will vanish on a subset $\Upsilon \subseteq Y$ of essential dimension $n-2$. Any extension of β to X will be of the form

$$\langle B, x \rangle = b_1 x_1 + \cdots + b_{n-1} x_{n-1} + c x_n$$

for some constant c.

I claim that c can be chosen so that $B \geq 0$ on X and $B = 0$ meets at least one point x in $Y - X$. This means no more and no less than that it defines an extremal half space for X. The conditions on B are that

$$\langle B, x \rangle = b_1 x_1 + \cdots + b_{n-1} x_{n-1} + c x_n \geq 0$$

for all x in $X - Y$, and that $\langle B, x \rangle = 0$ for at least one of these x. These require that

$$-c \leq \frac{b_1 x_1 + \cdots + b_{n-1} x_{n-1}}{x_n}$$

for all x in $X - Y$, with equality in at least one case. But X is finite and $x_n > 0$ on $X - Y$, so the right hand side will achieve a minimum for some $c < 0$ as x varies.

• Implementing this proof as an algorithm that finds all extremal faces for a give X is not difficult. It goes by induction on dimension. The first step is straightforward, and at the end of it we have an external hyperplane along with a coordinate system designating it as $x_n = 0$. In the second step, we assemble a collection of external hyperplanes by adding to the ones we already have those external hyperplanes that hinge at the eternal faces of the given external hyperplanes, according to the proof of Lemma 4.2, until we have all of them.

∘ ———— ∘

Part II. The simplex method

5. Introduction

The algorithm suggested in an earlier section is an example of one of the two basic methods currently in use for analyzing convex bodies. Such methods are called **incremental**. In the remainder of this essay I'll explain the **simplex** method.

The task of the simplex method is dual to the one just outlined. It starts with a set of affine inequalities $f_i(x) \geq 0$, and the first main goal is to locate a vertex of the region defined by them. From that point on, several tasks are possible, as we shall see.

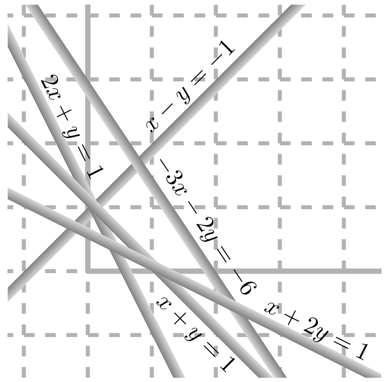
There is one complication, since there might not exist any vertices at all. This is for one of two reasons: (i) the space of translations leaving the region invariant is non-trivial; (ii) the region is actually empty. The process will uncover these possibilities as it goes.

• The process starts with a number of inequalities. We put them in uniform format, and then make up a matrix whose rows are the functions involved.

EXAMPLE. I'll interpolate my account of the general process with a running example. It starts with:

$$\begin{aligned} 2x + y &\geq 1 \\ x + 2y &\geq 1 \\ x + y &\geq 1 \\ x - y &\geq -1 \\ 3x + 2y &\leq 6\,. \end{aligned}$$

In pictures:



It is important to realize that the picture is deceptive, since in 3D you get a global view of 2D geometrical structures. *The principal virtue of the simplex method is that you act only on local information.* In fact, the simplex method can be used to build up some kind of global view of structures in high dimensions.

Here is the uniform format:

$$\begin{array}{rcrl} f_1(x,y) & = & 2x+y-1 & \geq 0 \\ f_2(x,y) & = & x+2y-1 & \geq 0 \\ f_3(x,y) & = & x+y-1 & \geq 0 \\ f_4(x,y) & = & x-y+1 & \geq 0 \\ f_5(x,y) & = & -3x-2y+6 & \geq 0 \end{array}$$

and here the matrix:

$$F = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ -2 & -3 & 6 \end{bmatrix}$$

○————- ○

• The next steps replace Ω by the homogenization of its quotient $\Omega/\mathrm{Inv}(\Omega)$. The point is to wind up with a cone that possesses the origin as a vertex. One can do these steps in either order, since the invariance subspaces $\mathrm{Inv}(\Omega)$ and $\mathrm{Inv}(\overline{\Omega})$ are the same. It is slightly less intuitive, but technically more convenient, to first homogenize. This means replacing the constants in the affine functions f_i by multiples of x_{n+1}, and adding $x_{n+1} \geq 0$ itself to the requirements. In terms of matrices, this means only adding a final row corresponding to x_{n+1}.

EXAMPLE. First

$$\begin{array}{rcrl} \overline{f}_1(x,y,z) & = & 2x+y-z & \geq 0 \\ \overline{f}_2(x,y,z) & = & x+2y-z & \geq 0 \\ \overline{f}_3(x,y,z) & = & x+y-z & \geq 0 \\ \overline{f}_4(x,y,z) & = & x-y+z & \geq 0 \\ \overline{f}_5(x,y,z) & = & -3x-2y+6z & \geq 0 \\ \overline{f}_6(x,y,z) & = & z & \geq 0 \end{array}$$

and then

(5.1)
$$\overline{F} = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ -2 & -3 & 6 \\ 0 & 0 & 1 \end{bmatrix}$$

○————- ○

• We apply column reduction to this matrix.

This does two things: it gives us (1) a basis for the subspace $U = I(\Omega)$ of translations leaving the system invariant, and (2) a basis for a subspace transverse to U—i.e. a **slice** of the projection from V to V/U. This is one of the things explained in the appendix on matrix manipulations. The point is that the column reduction is carried out through right multiplication by an $(n+1) \times (n+1)$ matrix R. In effect, this amounts to a coordinate change. Say that the rank of F is d. Then the first d columns of R are a basis of the slice, and the last $(n+1)-d$ columns are a basis of the invariance translations. The rows of the reduced matrix are the coordinates of the original affine inequalities, but in the new coordinate system. This also is explained in more detail in the appendix on matrix manipulations.

EXAMPLE. After column reduction:

$$\overline{F} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & -3 \\ 4 & 3 & -13 \\ 1 & 1 & -3 \end{bmatrix}$$

with the reducing matrix

$$R = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 1 & 1 & -3 \end{bmatrix}$$

The matrix $\overline{F}$ can be interpreted as asserting linear relationships among the original f_i, for example

$$f_4 = 2f_1 - 3f_3 .$$

In this example, the space U is trivial, since the rank of F is 3.

∘———∘

• At this point, we transfer all consideration to the slice. The point of the quotient reduction and the homogenization is that *we now have a vertex of a region related closely to the original* Ω *to work with, namely the origin.* We now have as well a coordinate system with the property that the region $\overline{\Omega}$ is contained in the coordinate octant. (Since this octant is a simplicial cone, this is presumably the origin of the nomenclature.) In other words, in this coordinate system the cone we are looking at contains inequalities $y_i \ge 0$ for $1 \le i \le d$. This is called a coordinate system **adapted** to the region, and the corresponding basis is called a **vertex basis**.

There are now two possibilities: (1) the conditions imposed by the inequalities are inconsistent, in which case the region is empty or (2) the region possesses a vertex. If it is not empty, we want to find an explicit vertex along with an adapted coordinate system.

• The next major step will attempt to locate a vertex on the hyperplane $x_{n+1} = 1$. This will be done by first turning the cone defined by the linear inequalities into the affine region $x_{n+1} \le 1$ or $1 - x_{n+1} \ge 0$. This gives us a new matrix whose rows are of dimension $n+2$.

EXAMPLE.

$$\overline{F} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & -3 & 0 \\ 4 & 3 & -13 & 0 \\ 1 & 1 & -3 & 0 \\ -1 & -1 & 3 & 1 \end{bmatrix}$$

∘———∘

We then then try to maximize the function x_{n+1}. This will involve pivoting through a sequence of vertex bases. The next few sections will explain this in detail, but I can summarize very briefly what happens.

At this point, there will be d of the original affine functions that make up a vertex basis at the origin, say $y_i = f_{n_i}$ for $1 \le i \le d$. Write each function in the matrix as

$$f_i = \sum_j c_{i,j} y_j + c_i \, .$$

Suppose the function we want to maximize (x_{n+1} in the original coordinates) is, in the current coordinate system, equal to

$$x_{n+1} = \sum \gamma_j y_j \, .$$

(1) Pick some c such that $\gamma_c > 0$. This is always possible, since x_{n+1} is one of the functions of the system.
(2) Make a list of all k such that $c_{k,c} < 0$. If the list is empty, we have to stop. There is no edge of the cone leading out to the hyperplane $x_{n+1} = 0$, and the region Ω is empty.
(3) Otherwise, extract from this list a smaller one, of all those for which the ratio $-c_k/c_{k,c}$ is a minimum. (Note that $c_k \ge 0$ because the origin is actually in $\overline{\Omega}$.) Choose one of these values of k, say $k = r$. As we shall see later there will be rules to specify which c and r are chosen if there are not unique choices.
(3) Replace y_c in the basis by f_r. Make the proper substitutions in the matrix of inequalities to get a new matrix. (This is a pivot.)
(4) If $c_r = 0$, we have not actually changed vertices. In this case, we loop back to (1).
(5) Otherwise, the new basis is a vertex on the hyperplane $x_{n+1} = 1$, which is what we want, since the slice of $\overline{\Omega}$ where $x_{n+1} = 1$ is the same as the original Ω. In that case, we make the function $1 - x_{n+1}$ one of the coordinates, and then set $x_{n+1} = 1$ to get an affine vertex basis in n dimensions.

EXAMPLE. In the continuing example, we can choose $c = 1$. Then there is exactly one r such that $c_{r,c} < 0$, which means we swap f_1 with $1 - z$. We transfer immediately along an edge to $z = 1$, finally arriving at the matrix

$$\begin{bmatrix} -1 & -1 & 3 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -2 & -2 & 3 & 2 \\ -4 & -1 & -1 & 4 \\ -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

I'll recall how to interpret this matrix. It represents a vertex basis of the affine convex set defined by the inequalities (5.1) , together with inequalities defining the associated homogeneous cone, but truncated at $z = 1$. Except that the affine coordinate system is different from the original one. The coordinate functions are now f_7 (a. k. a. $1 - z$), f_2, and f_3.

$$\begin{aligned} -f_7 - f_2 + 3\overline{f}_3 + 1 &\ge 0 \\ \overline{f}_2 &\ge 0 \\ \overline{f}_3 &\ge 0 \\ -2f_7 - 2\overline{f}_2 + 3\overline{f}_3 + 2 &\ge 0 \\ -4f_7 - \overline{f}_2 - \overline{f}_3 + 4 &\ge 0 \\ -f_7 + 1 &\ge 0 \\ f_7 &\ge 0 \end{aligned}$$

The point where all the basis functions vanishes is in fact the domain Ω, since all the constants in these equations are non-negative. We can obtain the equations defining the original Ω by setting $f_7 = 0$, and then eliminating those which are only constants. This give us

$$\begin{aligned} -f_2 + 3f_3 + 1 &\geq 0 \\ f_2 &\geq 0 \\ f_3 &\geq 0 \\ -2f_2 + 3f_3 + 2 &\geq 0 \\ -f_2 - f_3 + 4 &\geq 0 \end{aligned}$$

The origin of the coordinate system is the intersection of the lines $f_2 = x + 2y = 1$, $f_3 = x + y = 1$ in the original coordinate system. The vertex we are looking at, in the original coordinate system, is $(1, 0)$.

6. Vertex bases

Before I explain exactly how the simplex method goes, I'll explain in this section and the next the basic structure and process involved.

Suppose Ω to be in $\mathbb{R}^n$, defined by a finite set of inequalities $f_i \geq 0$, and suppose ω to be a vertex of Ω. This means that that there exists a hyperplane H intersecting Ω in ω, or equivalently that Ω is not invariant under any translations. The practical consequence of this is that there exists a system of coordinates is adapted to ω and Ω. This means that

- the point ω is the origin of the system;
- the coordinate functions are among the affine functions in the inequalities defining Ω.

In particular, Ω is contained in the coordinate octant. Being a vertex basis is a local property—it depends purely on the geometry of the regions $f_i \geq 0$ in a neighbourhood of the origin.

I'll call the coordinate function in such a coordinate system a **vertex basis**. *This is the basic structure (no pun intended) in the simplex method*. As we shall see, if we are given any set of inequalities to start with, the first major step in any exploration of a convex polytope Ω is to decide whether it is empty or not, and if it is not to find a vertex and am associated vertex basis. The point is that any vertex of the polytope may be reached by progressing through a succession of vertex bases by following edges of the bases one encounters. One does this by **pivoting**, which I'll explain in the next section.

I should point out, however:

> *In the standard literature on linear programming, the set of functions that are* **not** *coordinates make up what is called the basis. I am not sure where or why this usage originated. I follow more standard mathematical terminology and call the basis the set of functions in the inequalities that are coordinate functions. They are, after all, a basis for the space of all linear functions. This is consistent with my geometrical approach, whereas the literature often seems to shun geometrical notions.*

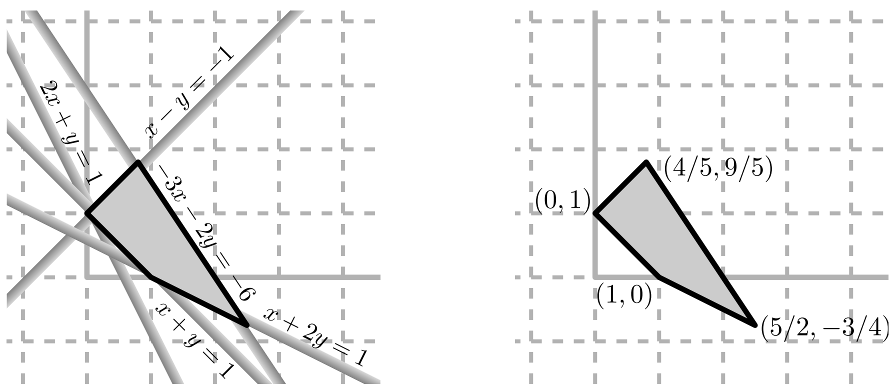
The simplest vertices are those for which Ω is locally a simplicial cone—given locally at the origin exactly by inequalities $f \geq 0$ for f in the basis. These are called **simple vertices**, and many results are known only for them. But this does not always occur. This simple situation is the generic case, in the sense that perturbing a given system almost always turns any given polytope into one all of whose vertices are simplicial. A vertex not of this form is called **singular**. Keep in mind that it is the system that is singular, not necessarily the polytope itself. In two dimensions, for example, all vertices are geometrically simplicial, but systems of inequalities may nonetheless be singular.

For example, suppose given the system of inequalities

$$\begin{aligned} 2x + y &\geq 1 \\ x + 2y &\geq 1 \\ x + y &\geq 1 \\ x - y &\geq -1 \\ 3x + 2y &\leq 6\,. \end{aligned}$$

Note that the origin is not in Ω.

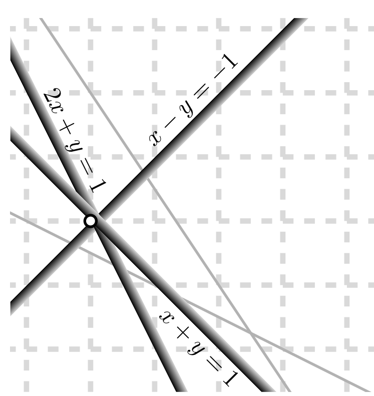
We have the following picture, where the shading along a line $f = c$ indicates the side where $f - c \geq 0$. The convex region defined by the inequalities is outlined in heavy lines.



In this example, the system is singular at $(0, 1)$ but not at the other three vertices of the region. The singularity at $(0, 1)$ arises because the three bounding lines

$$\begin{aligned} 2x + y - 1 &= 0 \\ x + y - 1 &= 0 \\ x - y + 1 &= 0 \end{aligned}$$

all pass through that point. Any two of the corresponding affine functions will form a vertex basis. Only the last two lines actually bound the region. But any two of these three might be the local coordinate system of a vertex basis.



Suppose we choose

$$\boldsymbol{x} = 2x + y - 1, \quad \boldsymbol{y} = x + y - 1\,.$$

We can express this as:

$$\begin{bmatrix} 2 & 1 & -1 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} .$$

The pair $(\boldsymbol{x}, \boldsymbol{y})$ make up a vertex basis. In the new system, the inequalities become

$$\begin{aligned} \boldsymbol{x} &\ge 0 \\ \boldsymbol{y} &\ge 0 \\ -\boldsymbol{x} + 3\boldsymbol{y} + 1 &\ge 0 \\ 2\boldsymbol{x} - 3\boldsymbol{y} &\ge 0 \\ -\boldsymbol{x} - \boldsymbol{y} + 4 &\ge 0 \,. \end{aligned}$$

7. Pivoting

The basic step in using the simplex method is moving from one vertex basis to another—and, hopefully, from one vertex to another—along a single edge of the polytope. (Or rather, to be precise, along what you might call a virtual edge.) Geometrically we are moving from one vertex to a neighbouring one, except that in the case of a singular vertex where more than the generic number of faces meet the move might travel along an edge of length 0. In all cases, this basic step is called a pivot. It is executed by pivoting the matrix F.

The explicit problem encountered in pivoting is this: we are given a vertex basis and a variable x_c in the basis. The x_c-axis represents an edge (or possibly a virtual edge, as I'll explain in a moment) of the polytope. We want to locate the vertex of the polytope next along that edge, and replace the old vertex basis by one adapted to that new vertex. The variable x_c is part of the original basis. The new basis will be obtained by replacing that variable by one of the affine functions f_r in the system of inequalities defining the polytope. In applications, the exiting variable x_c is chosen according to various criteria which depend on the goal in mind, as we shall see in the next section. Sometimes there are several choices for f_r as well.

There are conditions on the possible f_r. For one thing, the new vertex must be also actually a vertex of the region Ω concerned. This puts a bound on how far out along the x_c-axis we can go. Suppose

$$f(x) = \sum a_i x_i + a_{n+1} \,.$$

We know already that $a_{n+1} \ge 0$, because our vertex does lie in Ω. On the x_c-axis

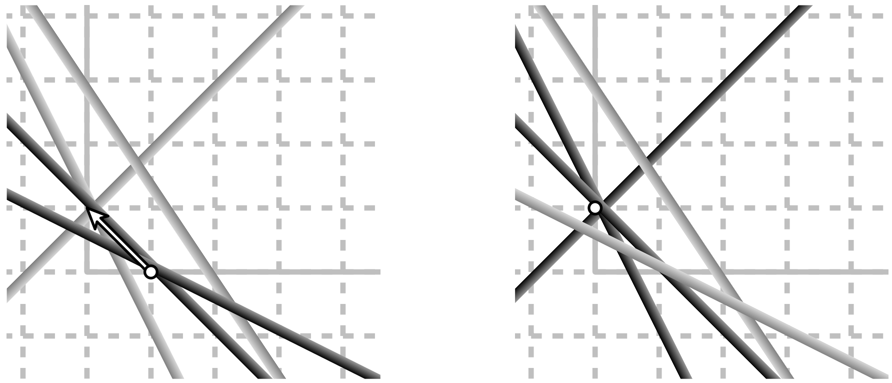
$$f(x) = a_c x_c + a_{n+1} \,.$$

If $a_c \ge 0$, there is no limit to x_c. But if $a_c < 0$, we must have

$$a_c x_c + a_{n+1} \ge 0, \quad x_c \le \frac{a_{n+1}}{-a_c} \,.$$

This has to hold for all the functions f_r, so *the new value of x_c is the minimum value of $-a_{r,n+1}/a_{r,c}$ for those r for which $a_{r,c} < 0$*. Given this choice of variable x_c, and without further conditions imposed, the possible choices for the new basis variable will be any of the f_r producing this minimum value of x_c.

In the example we have already looked at, we could start with coordinates $x + 2y = 1$, $x + y = 1$ at the vertex $(1, 0)$ and move along the edge $x + y = 1$, with $x_c = x + 2y - 1$. We can only go as far as the vertex $(0, 1)$, because beyond there both of the affine functions $x - y + 1$ and $2x + y - 1$ become negative. We can choose either one of them to replace $x + 2y - 1$ in the vertex coordinates.



When a pivot is made, the current coordinate basis is the given vertex basis, and in the course of the pivot it is changed to the new vertex basis. This involves expressing all the f_i in terms of the new basis.

How does it work in general? Given a vertex basis, any of the f_i may be expressed as

$$f_i = a_i + \sum_j a_{i,j} x_j$$

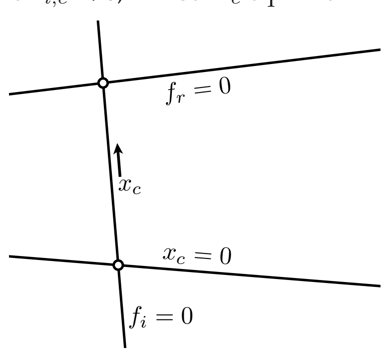
with the $a_i \geq 0$ since the origin is assumed, by definition of a vertex basis, to lie in the region $f_i \geq 0$. We propose to move out along the directed edge where all basic variables but one of them, say x_c, are equal to 0, and in the direction where x_c takes non-negative values. This amounts to moving along an edge of the polytope. We shall move along this edge until we come to the vertex at its other end, if there is one, or take into account the fact that this edge extends to infinity. So we must find the maximum value that x_c can take along the line it is moving on. Along this edge the other x_j vanish, and we have each

$$f_i = a_i + a_{i,c} x_c$$

There are two cases to deal with:

(a) If $a_{i,c} \geq 0$ there is no restriction on how large x_c can be;
(b) if $a_{i,c} < 0$ then we can only go as far as $x_c = -a_i/a_{i,c}$.

Therefore if all the $a_{i,c}$ are non-negative then the edge goes off to infinity, but otherwise the value of x_c is bounded by all of the $-a_r/a_{r,c}$. In that case, we choose r such that $a_{r,c} < 0$ and $-a_r/a_{r,c}$ takes that minimum value among those where $a_{i,c} < 0$, and set x_c equal to it.



We also change bases—replacing x_c by f_r. The minimum value might in fact be 0, in which case we change vertex frames (i.e. virtual vertices) without changing the geometric vertex. This sort of **virtual move** will happen only if the vertex is singular, which means as I have said that more than the generic number of faces meet in that vertex, and the maximal value of x_c will be 0.

Thus our choice of r is made so that

$$a_i - (a_r/a_{r,c}) a_{i,c} \geq 0$$

for all i, since if $a_{i,c} \geq 0$ there is no condition (a_i, $a_{i,c}$ and $-a_r/a_{r,c}$ are all non-negative) and if $a_{i,c} < 0$ then by choice of r

$$-\frac{a_i}{a_{i,c}} \geq -\frac{a_r}{a_{r,c}} \,.$$

When we change basis, we have to change all occurrences of x_c (which becomes just a function f_c in our system) by substitution. In short, we pivot.

Sometimes, as in the example, there will be several possible choices of f_r, if the target vertex isn't simplicial. Indeed, in practice when pivoting is done in an application there may be several possible edges to move out along—i.e. which variable x_c to exit from the basis. The most important thing in all cases is to prevent looping. There are two commonly used methods to do this. One is a method of choosing both x_c and f_r, while the other is fussy about only f_r. I'll explain these in the next section.

8. Finding a vertex basis to start

As we shall see later, and as I have already suggested, nearly all analysis of a polytope starts with a vertex basis. But if we are given just a set of defining inequalities, it might not be evident how to find one. As I have mentioned, we must deal locally. We cannot get an overall view of our region, but just inspect what is going on in the neighbourhood of the current origin.

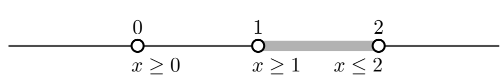
At this point, we have at hand a vertex basis for the homogenized version of the original system of affine inequalities. It is from this that we shall find a vertex basis for the original system (or find that the region is in fact empty).

One of the more elegant features of the simplex method is that this will turn out to be a special case of the original problem that originally motivated the simplex method, that of finding the maximum value of a linear function on a polytope, given an initial vertex basis. I postpone the detailed explanation of how this works. At any rate, in our situation we are given linear coordinates of the function that was our original homogeneous variable x_{n+1}. Our given vertex is the origin, where $x_{n+1} = 0$. What we want to locate is a vertex of the original inhomogeneous system, which corresponds to the slice of our cone at x_{n+1}. So we add to system the affine inequality $x_{n+1} \leq 1$, and find the maximum value of x_{n+1} subject to the new conditions. If the maximum value is 1, we have located a good vertex basis for the original region. If not, the solution set of the original inequalities is empty.

It's a little difficult to picture what is going on, except in very low dimensions. Suppose the original equations were

$$\begin{aligned} x &\geq 1 \\ x &\geq 1 \\ x &\leq 2 \end{aligned}$$

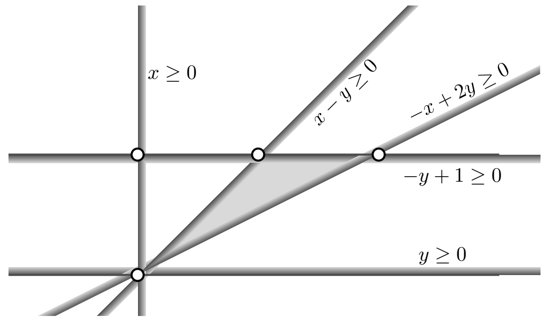
in one dimension. The important point is to realize that the origin of this coordinate system is not a vertex of the polytope—the function $x - 1$ takes a negative value at $x = 0$.



These become the system

$$\begin{aligned} x &\geq 0 \\ x - y &\geq 0 \\ -x + 2y &\geq 0 \\ y &\geq 0 \\ -y + 1 &\geq 0 \,. \end{aligned}$$

in two dimensions, after adding the new equations in y.



For a vertex basis at the origin we have six choices—any two functions out of the four that vanish at 0: x, $x - y$, $-x + 2y$, y. What we get by elimination depends on the order in which we write the equations. Assuming $y \geq 0$, $-y + 1 \geq 0$ to be last, we are reduced to three choices. The choice $\boldsymbol{x} = x$, $\boldsymbol{y} = x - y$ (so $y = \boldsymbol{x} - \boldsymbol{y}$) gives us the new system

$$\begin{aligned} \boldsymbol{x} &\geq 0 \\ \boldsymbol{y} &\geq 0 \\ \boldsymbol{x} - 2\boldsymbol{y} &\geq 0 \\ \boldsymbol{x} - \boldsymbol{y} &\geq 0 \\ -\boldsymbol{x} + \boldsymbol{y} + 1 &\geq 0 \,. \end{aligned}$$

In this case, the simplex method finds that there is indeed a vertex basis satisfying the original inequalities.

This process will be a bit unusual from the standpoint of linear programming, as we shall see later. All moves except the final one will be what I call virtual. What we are doing, roughly, is moving around in the 'web' of vertex frames defined by the homogenized functions f_i.

ANOTHER EXAMPLE. An example in which $\Omega = \emptyset$ might be instructive. Start with the inequalities

$$\begin{aligned} x &\geq 1 \\ x &\leq 0 \,. \end{aligned}$$

This gives us the homogeneous matrix

$$F = \begin{bmatrix} 1 & -1 \\ -1 & 0 \\ 0 & 1 \end{bmatrix},$$

and then the affine system

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

with basis f_1, z in 3 dimensions. In this case, maximizing z doesn't do all that well. The variable z is the only choice to exit the basis. The only choice for one to enter it is f_2, but (of course) the move is only virtual. It results in the matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

But now all coefficients of z are negative,, telling us that $\Omega = \emptyset$.

9. Tracking coordinate changes

As we move among vertices, changing from one vertex basis to another, we'll want to keep track of where we are *in the original coordinate system*. We can recover the location of a vertex from the basis, since it amounts to solving a system of linear equations. But this is a lot of unnecessary work. We can instead maintain some data as we go along that make this a much simpler task. *What data do we have to maintain in order to locate easily the current vertex?* A slightly better question: *What data do we have to maintain in order to change back and forth between the current basis and the original one?*

Keeping track of location comes in two halves: changing back and forth between the original coordinate system and the one associated to the homogenized problem, in which we located the invariance subspace and wound up with the cone; changing back and forth between that basis and the current vertex basis. The first is done by a linear coordinate change, and after that, by affine changes, substituting f_r for x_c. Thus is done by exactly the same column operations that are performed on F, and in programs R is often tacked onto the bottom of F to make this convenient. Its rows are not listed in either of the two lists accompanying the data structure, of course.

10. Finding maximum values

The original problem that motivated the introduction of the simplex method in linear programming was to maximize an affine function $z = c + \sum c_i x_i$ subject to a set of conditions $f_i = a_i + \sum a_{i,j} x_j \ge 0$ that determine a convex region Ω. In doing this, one proceeds from one vertex of Ω to another, always increasing the value of z as one goes. This is kept up until it is seen either that z is unbounded on Ω, or that no choice of maximizing edge is possible and the maximal value of z has been found.

Well, not quite. What one really does is proceed, not from one vertex to another, but from one vertex basis to another. This takes place through a pivot operation. If the edge has length 0, the vertex itself does not change although the basis does, and the value of z does not in fact change. If the move is virtual—does not change vertices—then something has to be done to avoid going around in circles.

The problem is that in the simplex method one is always proceeding only on the basis of local information. There is no way to peek ahead to see where it is going. So one must find a way to pivot, based only on local information, that somehow avoids cycling. There have been a number of methods devised to do this, of which I shall discuss two.

Ata ny rate, if the process does not terminate then it must cycle around in a loop. Since the function z increases if the distance traversed along an edge is positive, it will cycle at a fixed vertex. In this case, all exiting and entering variables will vanish at that vertex, so that we may therefore ignore those functions that do not vanish at that vertex. For convenience, we may also assume that the function z to be maximized is 0 at the given vertex. Finally, there must be at least two vertex bases in the loop.

BLAND'S RULE. At each step, one affine function f_i in the current basis is replaced by another. One valid rule, introduced in [Bland:1977], is to choose the exiting and entering functions to have the least possible subscript. It is covered in many places in the literature, but most of them are simple variations of one another. I'll follow the explanation in [Camarena:2016], which is a variant of the treatment on pages 37–38 of Chvátal's book.

10.1. Proposition. *Following Bland's rule, the simplex method will terminate.*

Proof. In several steps.

Step 1. Let S be the set of all indices of the given affine functions. At any stage p, let $B = B_p$ be the indices of the functions f_i in the current basis. Then we can write

$$f_i = \sum\nolimits_B f_{i,j} f_j$$
$$z = \sum\nolimits_B \zeta_i f_i \,.$$

With these assumptions (particularly that all the f_i vanish at the given vertex), the rules for pivoting are not complicated.

10.2. Lemma. *Suppose that at some given stage the function involved are expressed as above. To make the next pivot:*

(a) choose i minimal with $\zeta_i > 0$;
(b) choose j minimal such that $f_{j,i} < 0$;
(c) replace f_i in the basis by f_j.

Step 2. Now suppose that a cycle does occur while following Bland's rule. Because it is a cycle, every function that enters the basis at one point must leave it somewhere else. Hence:

10.3. Lemma. *The following are equivalent:*

(a) the function f_s enters the basis at one vertex in the loop and exits the basis at another;
(b) the function f_s enters the basis at some vertex in the loop;
(c) the function f_s exits the basis at some vertex in the loop.

Call s **fickle** if one of these conditions is verified.

Step 3. Let a be greatest among the fickle indices, and suppose that f_a enters the basis at some stage p (i.e. in going from stage p in the loop to its successor). Write at this stage

$$\begin{aligned} f_i &= \sum\nolimits_B f_{i,j} f_j \\ z &= \sum\nolimits_B \zeta_i f_i \end{aligned}$$

with $B = B_p$.

Let b be such that f_b exits the basis in that step.

(i) $b < a$.
∘ Since according to Lemma 10.3 the index b is fickle.
(ii) b is least with $\zeta_b > 0$.
(iii) a is least with $f_{a,b} < 0$, and $f_{s,b} \geq 0$ for all other fickle s.
∘ If $f_{s,b} < 0$ then necessarily $s > a$, but a is maximal.

Step 4. Suppose that at some later stage q the function f_a leaves the basis, and suppose that at this point

$$z = \sum\nolimits_{B_q} \zeta_i^* f_i \,.$$

10.4. Lemma. *In these circumstances*

$$\zeta_b = \zeta_b^* + \sum \zeta_i^* f_{i,b} \,,$$

in which the sum is over elements i of B_q that are not in B_p.

This equation is the essential link between two stages. Note that the elements i in the sum are necessarily fickle.

Proof. Set P_t to be the point on the f_b axis such that

$$f_i(P_t) = \begin{cases} t & \text{if } i = b \\ 0 & \text{if } i \neq b \text{ is in } B. \end{cases}$$

Then

$$f_j(P_t) = f_{j,b} t$$

for any j in S, and

$$\begin{aligned} z(P_t) &= \zeta_b t \\ &= \zeta_b^* t + \sum\nolimits_j \zeta_j^* f_j(P_t) \\ &= \zeta_b^* t + \sum\nolimits_j \zeta_j^* f_{j,b}(P_t) \,. \end{aligned}$$

Equate coefficients, and keep in mind that the j-th term in the sum is 0 if j is in B_p. ▯

Step 5. We now have

(iv) $\zeta_a^* > 0$, and a is least in B_q with this property.
∘ Because a exits the basis at stage q, following Bland's rule.
(v) $\zeta_s^* \leq 0$ for all other fickle s in B.
∘ Because a is maximal among all fickle indices.

Step 6. Since $\zeta_b > 0$ but $\zeta_b^* \leq 0$, Lemma 10.4 implies that $\zeta_c^* f_{c,b} > 0$ for some c. Let's look at f_c more closely.

(vi) The index c is fickle.
∘ The index c is not in B_p because the sum in Lemma 10.4 ranges only over the complement of B_p. It is in B_q because $\zeta_c^* \neq 0$.

This implies that $c \leq a$. But

(vii) The index c is not equal to a.
∘ Because $\zeta_c^* f_{c,b} > 0$ while $\zeta_a^* f_{a,b} < 0$.

This implies in turn that $f_{c,b} > 0$ because among fickle indices only $f_{a,b} < 0$. Since $\zeta_c^* f_{c,b} > 0$, we deduce that $\zeta_c^* > 0$. But this is a contradiction, since according to Bland's rule only $\zeta_a^* > 0$. ▯

PERTURBATION. Bland's rule is very simple to implement. Not so simple is the **perturbation** method, which has as variant the **lexicographic** method. This also is explained in Chvátal's book. The variant is the practical implementation of the first, which is simpler to understand. In this, as soon as we have found an affine basis, we add to each of the 'slack' variables f_i a symbolic constant ε_i with these subject to conditions

$$1 \gg \varepsilon_1 \gg \varepsilon_2 \gg \cdots \gg \varepsilon_\ell > 0\,,$$

in which ℓ is the number of non-basic inequalities. Then whenever we pivot, the constant terms in all the variables and the objective function will be linear combinations of numbers and these ε_i. In effect, we are setting the constants to be some kind of ε-numbers $c_0 + \sum c_i\varepsilon_i$, which we can add, subtract, and multiply by scalars. Such numbers are compared in the natural way, which amounts to lexicographic ordering of the arrays. Furthermore, with these adjustments th system becomes simplicial—none of the functions f_i (i.e. no variable other than the basis variables) will ever evaluate to 0 at the vertex, because the different linear combinations of the ε_i will always be linearly independent (pivoting preserves this property, and the matrix starts with I). Thus, no pivots will ever be virtual.

At the end, set all $\varepsilon_i = 0$. This method has the property that we are allowed to go out along any coordinate edge, and need only make a choice among the functions that will become part of the basis. This happens only when there are ties. This feature plays a role in the process discovered by Avis and Fukuda that I shall mention later on.

There is an equivalent description common in the literature, which also uses lexicographic ordering of arrays. The best account I know is in [Apte:2012], although more cryptic accounts are easy enough to find. It does not use ε-numbers, but expands the matrices of dictionaries instead. I prefer the version given here, which seems to be more intuitive even if slightly less efficient.

The advantage of either one of these lexicographic strategies is that it is only the choice of new variable to set as coordinate that uses it. This makes it exactly suitable for the `lrs` methodfrom top down.

11. *Appendix.* Matrix manipulations

I first recall how to interpret matrices. I want to be careful in distinguishing vectors from their coordinate representations.

Suppose V to be a vector space of dimension n. Suppose (α_i) to be a basis of V. I shall often write it as a matrix:

$$\alpha = [\,\alpha_1 \quad \ldots \quad \alpha_n\,]\ .$$

Then to every vector x is associated a coordinate array—a column matrix $x_\alpha = (x_i)$ with n entries:

$$x = \sum x_i\alpha_i = [\,\alpha_1 \quad \ldots \quad \alpha_n\,]\begin{bmatrix} x_1 \\ \ldots \\ x_n \end{bmatrix} = \alpha\cdot x_\alpha\,.$$

These conventions are consistent. If $V = \mathbb{R}^n$, which has a standard basis, any other basis α will be specified by a matrix whose i-column is the coordinate vector of α_i with respect to that standard basis.

A vector f in the dual space $\widehat{V}$ is canonically a linear function on V:

$$x \longmapsto \langle f, x\rangle\,.$$

If a basis α is chosen, it then corresponds to a matrix f_α with one row:

$$\langle f, x\rangle = f_\alpha\cdot x_\alpha = [\,f_1 \quad \ldots \quad f_n\,]\begin{bmatrix} x_1 \\ \ldots \\ x_n \end{bmatrix}\ .$$

What happens if we change basis, say from α to β? The basis β can be canonically expressed in terms of α:

(11.1)
$$\beta = \alpha \cdot R$$

in which R is an invertible $n \times n$ matrix. The i-th column of R is the coordinate vector of βi with respect to the basis α. This defines a canonical right action of GL_n on the set of bases, and makes it into a principal homogeneous space. Since

$$x = \alpha \cdot x_\alpha \ \text{ and } \alpha = \beta \cdot R^{-1}$$

we also have

$$x = \beta \cdot R^{-1} x_\alpha \,,$$

which means that the coordinate vector of x with respect to the basis β is

(11.2)
$$x_\beta = R^{-1} x_\alpha \,.$$

Similarly, for f in $\widehat{V}$

$$\langle f, x \rangle = f_\alpha x_\alpha = f_\alpha R \cdot R^{-1} x_\alpha = f_\beta x_\beta$$

so

(11.3)
$$f_\beta = f_\alpha R \,.$$

Now suppose that we are given vector spaces U of dimension m and V of dimension n, and a linear transformation τ from U to V. If we choose bases (α_i) and (λi) then τ gives rise to an $n \times m$ matrix T—the j-th column of T is the coordinate vector of $\tau(u_j)$ with respect to the basis (v_i). Then by definition

$$(\tau(x))_\lambda = T x_\alpha \,.$$

If a vector in U is represented by the column vector $x = (x_i)$, then Tx is the representation of its image in V. If a vector in $\widehat{V}$ is represented by the row vector f, its transpose image in $\widehat{U}$ is represented by fT.

Suppose R invertible, $\beta = \alpha R$. Then

$$T x_\alpha = TR \cdot R^{-1} x_\alpha = TR \cdot x_\beta$$

so TR is the matrix of τ with respect to the basis β. What is important is that the columns of all TR span the same space. Also, the space spanned by the rows of the TR is that of the image of the transpose ${}^t\tau$.

Familiar column operations may be applied to find the kernel of τ in U and its image in V. A matrix E is **weakly column reduced** if has these properties characterized recursively: (a) the first non-zero row of E has exactly one non-zero entry, which is 1; (b) the submatrix of E made up of the $e_{i,j}$ with $i \geq 1$, $j \geq 1$ is weakly column reduced. Here is a sample pattern:

$$\begin{bmatrix} \mathbf{1} & 0 & 0 & 0 & 0 \\ * & \mathbf{1} & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & \mathbf{1} & 0 & 0 \\ * & * & * & \mathbf{1} & 0 \\ * & * & * & * & 0 \\ * & * & * & * & 0 \end{bmatrix} .$$

If E is weakly column reduced of rank d, there is associated to it a unique sequence

$$r_1 < r_2 < \cdots < r_d$$

with the property that $e_{i,j} = 0$ if $i < r_j$ but $e_{r_j,j} = 1$. In the example above, the rank is 4 and this sequence is $1, 2, 4, 5$. Any matrix T may be reduced to a matrix that si weakly column reduced through multiplication on the right by an invertible matrix, say R. In this case, let d be the number of non-zero columns in TR. Then these columns are a basis of the image of T, and the last $m - d$ columns of R are a basis of the kernel of τ. The special rows are a basis of the inverse image of $\widehat{V}$ in $\widehat{U}$, and the i-th row is the corresponding coordinate vector of the inverse image of the i-th coordinate function in $\widehat{U}$.

A matrix that is weakly column reduced is called **column reduced** if in each row r_i there is only one non-zero entry. The following is a typical pattern:

$$\begin{bmatrix} \boldsymbol{1} & 0 & 0 & 0 & 0 \\ 0 & \boldsymbol{1} & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ 0 & 0 & \boldsymbol{1} & 0 & 0 \\ 0 & 0 & 0 & \boldsymbol{1} & 0 \\ * & * & * & * & 0 \\ * & * & * & * & 0 \end{bmatrix} .$$

Given any $n \times m$ matrix F of rank d, a standard algorithm in linear algebra will find a sequence of elementary column operations that will make F column reduced. This finds a matrix E that is column-reduced and an $n \times n$ invertible matrix R such that $E = F \cdot R$. In these circumstances the first d columns of R will be a basis for the subspace spanned by the columns of F, the rows r_i of F will be a basis for space of functions spanned by all the rows of F, and the rows of E will be the coordinates of the rows in terms of these special rows.

I'll next show some applications.

LINEAR INEQUALITIES. Suppose given a collection of linear functions f_i. Put them into a matrix whose rows have length n. Apply column reduction. In the end, certain rows will hold a copy of the identity matrix I_d, and will be a basis of the space spanned by the f_i. The remaining rows will hold the coordinates of the f_i with respect to that basis. The last $n - d$ rows will be 0, and if R is the matrix recording column operations its last $n - d$ rows will record a basis (in original coordinates) of the invariance translations.

In the end we shall be locating certain points specified in the new coordinate system. But we shall also need to know their coordinates in the original coordinates. We have

$$x_{\text{old}} = R x_{\text{new}} .$$

EXAMPLE. Suppose we start with the linear inequalities

$$\sum_{j=1}^{4} G_{i,j} x_j \ge 0 \quad (i = 1, \ldots, 4)$$

where

$$G = \begin{bmatrix} 1 & -1 & 0 & -1 \\ 0 & 1 & 1 & -1 \\ 1 & 0 & 2 & -1 \\ -1 & -2 & -6 & 1 \\ -9 & -1 & -23 & 6 \\ -2 & 0 & -5 & 1 \end{bmatrix} .$$

If we apply column reduction we get

$$G \cdot R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & -3 & 0 \\ 4 & 3 & -13 & 0 \\ 1 & 1 & -3 & 0 \end{bmatrix}$$

where

$$R = \begin{bmatrix} 2 & 2 & -1 & 3 \\ 1 & 2 & -1 & 2 \\ -1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The subspace U of translations leaving the system of inequalities invariant is the last column of R, and the nondegenerate ssystem of inequalities has as matrix the first 3 columns of $G \cdot R$.

ORTHOGONAL SPACES. Suppose v_i to be a collection of vectors in $\mathbb{R}^n$. *How to find the space of linear functions vanishing on its linear span?* Put the vectors as columns in a matrix V. We are looking for row vectors f of dimension n:

$$f \cdot V = 0\,.$$

Apply row reduction to this matrix, and let R be the matrix determining reduction. The result will be a matrix whose first d columns are linearly independent, with the last $n - d$ vanishing. The last $n - d$ rows of R will be a basis for the space of linear functions vanishing on the space spanned.

EXAMPLE. Let

$$V = \begin{bmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{bmatrix}$$

Then

$$V \cdot R = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

where

$$R = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 1 & 1 \end{bmatrix}.$$

Thus

$$f = [\,-1 \quad 1 \quad 1\,]$$

is a basis for the orthogonal space.

SUBSPACE BASES. Suppose $\{v_i\}$ to be a collection of vectors in $\mathbb{R}^n$. How to find a basis of the space they span, and an expression of the rest as linear combinations of these?

Put the vectors as columns in a matrix, and apply row-reduction. Certain of the columns will exhibit a copy of the identity matrix, say, of dimension d. The remaining columns will be the corresponding cooordinates of the rest. This is actually just a variant interpretation of the previous example.

Part III. Applications

12. Finding all vertices of a polygon

The traditional method from this point on is to use a stack to index and locate vertex bases, collecting them into equivalence classes, each corresponding to a vertex. Note that each vertex will correspond to a subset of the original set of functions, so our search is finite. Indeed, a good way to specify a vertex, at least when using exact arithmetic, is to tag a vertex by the set of functions vanishing there. On the other hand, it may happen that we have a singular vertex where several different sets of basic variables give rise to the same geometrical vertex. If we are doing floating point arithmetic, because of possible floating point round-off we should ignore this possibility, and proceed as if it doesn't occur. It shouldn't cause trouble—a single point might decompose into several, but they will all be close together and effectively indistinguishable.

Of course the region may be unbounded, in which case we should find the infinite edges of the polytope, too.

This method locates all edges and vertices by the simplex method, indexing the vertices by the set of vanishing variables, edges by the coordinate of the edge. The trouble is that obvious methods pass from one vertex basis to another, which means an apparent huge waste of time. Anyway, vertex bases are pushed onto the stack as they are met. They are distinguished by the subset of basic functions.

[Avis-Fukuda:1992] and [Avis:2000] describe a better method, called in its current version **lexicographic reverse search**. I hope to discuss in a later version of this essay.

13. Finding the dimension of a convex region

The dimension of a polytope makes sense in computation only when exact arithmetic is used, because the smallest of perturbations can change a set of low dimension to an open one.

Start with trying to find a vertex frame. If there is none, the dimension is -1. Next, throw away all but the f_i vanishing at the origin, so we are looking at the **tangent cone** of the vertex. Take as function to maximize the last basis variable; then find its maximum. If this is 0, we reduce the problem by eliminating that variable. Or we come across an infinite edge (axis of some x_c) out. In that case, throw away all those containing the variable x_c, in effect looking now at the tangent cone of that edge. Recursion: the dimension is one more than the dimension of that tangent cone. The whole process amounts to a succession of one or the other of these two strategies, applied recursing on rank. The case of rank one is straightforward—if the rank is 1, then in case (1) the dimension is 0 while in case (2) it is 1.

Part IV. References

1. David Avis, '`lrs`: a revised implementation of the reverse search vertex enumeration algorithm', pp. 177–198 in **Polytopes—Combinatorics and Computation**, edited by Gil Kalai and Günter Ziegler, Birkhäuser-Verlag, 2000.

2. —— and Komei Fukuda, 'A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra', *Discrete and Computational Geometry* **8** (1992), 295–313.

3. Robert G. Bland, 'New finite pivoting rules for the simplex method', *Mathematics of Operations Research* **2** (1977), 103–107.

4. Omar Antolin Camarena, 'Bland's rule guarantees termination', course notes for the spring of 2016, available at

`http://www.math.ubc.ca/~oantolin/math340/blands-rule.html`

5. Constantin Carathéodory, 'Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen', *Mathematische Annalen* **64** (1907), 95–115.

6. Vašek Chvátal, **Linear programming**, W. H. Freeman, 1983.

7. Donald E. Knuth, `lp.w`, a CWEB program in which Knuth lays out an introduction to linear programming. Posted August, 2005 at

`http://www-cs-faculty.stanford.edu/~uno/programs.html`

8. Jiří Matoušek and Bernd Gärtner, **Understanding and using linear programming**, Springer-Verlag, 2007.

9. Hermann Minkowski, 'Theorie der konvexen Körper, insbesondere Begründung ihres Oberflächenbegriffs'.

This has been published only as item XXV in volume II of **Gesammelte Abhandlungen**, Leipzig, 1911.

10. ——, **Geometrie der Zahlen**, Teubner, Berin, 1910.

11. Stephen Smale, 'On the average number of steps of the simplex method of linear programming', *Mathematical programming* **27** (1983), 241–262.

12. Daniel Spielman and Shang-Hua Teng, 'Smoothed analysis: why the simplex method usually takes polynomial time', *Journal of the ACM* **51** (2004), 385–463.

13. Ernst Steinitz, 'Bedingt konvergente Reihen und konvexe Systeme', *Journal für die Reine and Angewandte Mathematik* **143** (1913), 128–175.

14. Hermann Weyl, 'Elementare Theorie der konvexen Polyeder', *Commentarii Mathematici Helvetici* **7**, 1935, pp. 290–306.

There is a translation of this into English by H. Kuhn: 'The elementary theory of convex polytopes', pp. 3–18 in volume I of **Contributions to the theory of games**, *Annals of Mathematics Studies* **21**, Princeton Press, 1950.

15. Günter Ziegler, **Lectures on polytopes**, Springer, 2007.