LINEAR PROGRAMMING AND EXTENSIONS

R. Chandrasekaran
UT Dallas

April 10, 1997

Chapter 1

Geometry of LP

The most important concepts in this area are *convexity of sets and functions*; *topological properties of closed and open sets*; the notion of *bounded and unbounded sets*, etc. We will define the terms used and show the relationships first in pure geometric terms and then relate these to the corresponding algebraic notions.

Convex Subsets in $\mathbf{R}^n$: *A* set $X \subset \mathbf{R}^n$ is said to be *nonconvex* if there are two points in the set such that the line segment joining them is not entirely in the set; equivalently, if

$$\exists\, x \in X, y \in X; \lambda \in (0,1) \ni \lambda x + (1-\lambda)y \notin X)$$

A set is *convex* if it is not nonconvex. The set $L = \{z : z = \lambda x + (1-\lambda)y\}$ is the *straight line* passing through points x and y. If in addition $\lambda \in [0,1]$, then L is called the *line segment* with x and y as its terminals. "Clearly", L is a convex set (under either condition). When $\lambda \in [0,1]$, a point $z \in L$ is called a *convex combination* of x and y — and is a point somewhere between x and y on the line joining the two. If $S = \{x^1, x^2, ..., x^n\}$ is a set of points, then $z = \sum_j \lambda_j x^j$ is a convex combination of the points in S if $\lambda_j \geq 0$; and $\sum_j \lambda_j = 1$. The set of all such points z is called the *convex hull* of the set S. (*Indeed this notion can be suitably extended even when the set S is not finite*; *of course*, *the sum would be replaced by an integral*).

Facts:

1. If S and T are convex sets, then so is $S \cap T$. Indeed, if S_α is a collection of convex sets for $\alpha \in A$, then $\cap_{\alpha \in A} S_\alpha$ is a convex set. *In order to avoid special cases*, *it is customary to assert* (*by default*) *that ϕ is a convex set.*

2. If $S =$ is convex, $x^i \in S; 1 \leq i \leq n$ and $\lambda_i \geq 0$; $\sum_i \lambda_i = 1$, then $\sum_i \lambda_i x_i \in S$. (The proof of this is by induction on n. What would you do if the number of points selected to form the convex combination is not finite ?)

3. $H = \{x : cx = \alpha\}$ is called a *hyperplane*. $H^- = \{x : cx \leq \alpha\}$, $H^{--} = \{x : cx < \alpha\}$, $H^+ = \{x : cx \geq \alpha\}$ and $H^{++} = \{x : cx \geq \alpha\}$ are called *halfspaces*. H^- and H^+ are *closed* and the other two are *open*. All of these are convex sets. Intersection of any number of these (*an example would be the set of feasible solutions of a linear program*) is a convex set; if the number of these is finite the set is said to be *polyhedral*. Intersection of a finite number of hyperplanes is called an *affine* set. Affine sets are of the form $\{x : Ax = b\}$; the dimension of the subspace $\{x : Ax = 0\}$ is the dimension of the above affine set.

4. $D = \{x : \sum_i (x_i - a_i)^2 \leq \alpha^2\}$ is called a *disc* with center a and radius α, and is convex. The set $D^0 = \{x : \sum(x_i - a_i)^2 < \alpha^2\}$, is the *interior* of the disc D and is also convex. Generalization of the disc to higher dimensions is called a *hypersphere*. If $\exists$ finite $\alpha > 0 \ni S \subset D$, then the set S is said to be *bounded*. Bounded convex polyhedral sets are called *polytopes*.

5. The dimension of a convex set is the dimension of the smallest (in terms of set inclusion) affine set containing it.

6. $x \in S$ is called a *boundary point* of S if $\forall\ \alpha > 0$, the disc with center x and radius α has points both inside the disc and outside it. A boundary point need not belong to the set under consideration. If S is a set that contains all its boundary points is said to be *closed*. If $\bar{S}$ the complement of S is closed then S is said to be *open*. Another way to define an open set is to require that $\forall$ points in the set, $\exists\ \alpha > 0$ such that disc centered at the point with radius α is in the set. *Please avoid the confusion between the terms "closed' and "bounded"*.

7. The set of feasible solutions to an LP is closed, convex, and polyhedral; but it is not necessarily bounded.

8. A point $x^0 \in S$ is *not* an *extreme point (vertex)* of the convex set S if $\exists$ points x^1, and x^2 and $\lambda \in (0, 1) \ni x^0 = \lambda x^1 + (1 - \lambda)x^2$. Thus, x^0 is an extreme point of S if for every straight line L passing through it the segment $L \cap S$ has x^0 as one end. If L is a straight line and S is a convex set $\ni$ for every point $\in L \cap S$, the above statement is true for all lines $L' \neq L$, then we say that $L \cap S$ is an *edge* of the set S; if in this case, $L \cap S$ has two end points these are extreme points of S and are said to be *adjacent*. If $L \cap S$ is an edge which has only one end (hence is a semi-infinite line) then it is called a *ray* starting from the end and this end is an extreme point of S. If it has no end at all, then it can be shown that the set S has no extreme points at all. The disc D has an infinite number of extreme points while D^0 has none; open sets do not have extreme points. A necessary and sufficient condition for a closed convex set to have extreme points is that there be no subset that is an affine set of dimension ≥ 1.

9. Let S be a convex set. If H is a hyperplane $\ni\ S \subset H^+$ and $S \cap H \neq \phi$, then H is said to be a *supporting* hyperplane of S; if S is polyhedral then, $S \cap H$ is called a *face* of S. Faces of the highest dimension are called *facets*.

10. A set S that satisfies $[x \in S \Longrightarrow \lambda x \in S \ \forall\ \lambda \geq 0]$ is called a *cone*. If it is convex and is a cone then it is called a *convex cone*. A convex cone that does not contain any subspace of dimension ≥ 1 is said to be *pointed*. If it is polyhedral set then it is called a *polyhedral cone*.

Theorem 1 *Let S be a convex set and $x \notin \hat{S}$ where $\hat{S}$ is the closure of S (the smallest closed set containing S). Then, $\exists$ a hyperplane $H \ni \hat{S} \subset H^{++}$ and $x \in H^{--}$. (This is known as the* separating hyperplane theorem*).*

Let x be a boundary point of a convex set S. Then $\exists$ a supporting hyperplane passing through x for S.

Theorem 2 *A closed convex set is the intersection of all its supporting half-spaces.*

Theorem 3 *Let S be subset of $\mathbf{R}^n$. Any point in the convex hull of S can be written as a convex combination of at most $n+1$ points in S. (This result is often called* Caratheodory's theorem*).*

Theorem 4 *A closed bounded (compact) convex set is the convex hull of its extreme points. (This is known as* Krein-Millman theorem*).*

Functions: A function $f : \mathbf{R^n} \rightarrow \mathbf{R}$ is said to be *convex* if $\forall\ x^1,\ x^2$, and $\lambda\ \in (0,1)\ f(\lambda x^1 + (1-\lambda)x^2) \leq \lambda f(x^1) + (1-\lambda) f(x^2)$. f is *concave* if $-f$ is convex (or equivalently the inequality goes in the opposite direction). A function that is both convex and concave is a *linear* function. Another equivalent definition of convex functions is the following: f is convex iff $\{(x,z) : f(x) \leq z\}$ is a convex set; this set is known in the literature as the *epigraph* of f. A simple consequence of this is the fact that the sets $\{x : f(x) \leq \alpha\}$ are convex $\forall \alpha$; these sets are known as the *level sets*. If f satisfies only the second condition then f is said to be *quasiconvex*. If the sets $\{x : f(x) \geq \alpha\}$ are convex $\forall\ \alpha$, then f is said to be *quasiconcave*. A function that is both quasiconvex and quasiconcave is said to be *quasimonotone*. An example of such functions of a single variable is the monotone function.

1.1 Properties

1. $\sum \lambda_i f^i(x)$ is convex if each f^i is convex and $\lambda_i \geq 0\ \forall\ i$.

2. $\sup_\nu\ f_\nu(x)$ is a convex function if each of a family of functions f_ν is convex.

3. If f is convex function and x^0 is a *local minimum* of f over a convex set S, then it is a *global minimum*.

4. If f is convex over S, then it is continuous in the relative interior of S.

5. If f is a differentiable function of a single variable, then f' is monotone nondecreasing is both necessary and sufficient for convexity; if f is twice differentiable then $f'' \geq 0$ is a necessary and sufficient condition for convexity. If f is a function of several variables then these are modified to:

$$f(x) - f(x^0) \geq (x - x^0)\nabla f(x^0)$$

and the *hessian* H is *positive semidefinite*. (*A matrix M is positive semidefinite* ($P.S.D$) $\Longleftrightarrow x^t M x \geq 0 \; \forall \, x$).

6. F quasiconvex $\Longleftrightarrow \quad f(\lambda x^1 + (1-\lambda)x^2) \leq \max[f(x^1), f(x^2)] \; \forall \; x^1, \; x^2$, $\lambda \in [0,1]$.

1.2 LP Related Results

1. The set of feasible solutions of a LP is a convex polyhedral set.

2. Let the LP be in the form: $[\min cx : x \geq 0; Ax = b]$. Extreme points of the feasible set are precisely the basic feasible solutions. The simplex algorithm "*moves*" from one extreme point to an adjacent one that is "*better*" along the edge joining them. Unbounded feasible region implies the existence of nonzero solutions to the system: $[x \geq 0; Ax = 0]$. These are the rays of the feasible set and form a convex cone. The extreme rays correspond to the extreme points of the set $\{x : x \geq 0; Ax = 0; \sum x_j = 1\}$. If the LP is unbounded then $\exists$ an extreme ray with $cx < 0$ and this is produced by the simplex algorithm. A degenerate pivot in the simplex algorithm means that we have not "moved" at all; but are at the same solution point but with different bases. *Thus, though there is no "geometric" movement there is an algebraic pretension of such movement.*

3. We can geometrically "*show*" that if the LP has an optimal solution, then it has an extreme point that is optimal – *provided* $\exists$ *extreme feasible solutions*. For this purpose, we need the following powerful result: Any point x^0 in a convex closed polyhedral set S can be expressed as the sum of a convex combination of its extreme points P^i and a nonnegative combination of its extreme rays R^j. "Clearly" if the problem is bounded, then there are no "*profitable*" rays – rays with $cx < 0$. Hence,

$$cx^0 = \sum_i \lambda_i c P^i + \sum_j \mu_j c R^j ; \lambda_i \geq 0; \sum \lambda_i = 1; \mu_j \geq 0$$

Therefore, $cx^0 \geq \sum \lambda_i c P^i$ and hence $\exists \, P^* \ni \; cP^* \leq cx^0$; P^* is the required point. *Please note that this proof can be altered to suit quasiconcave functions in general.*

4. The set of optimal solutions to an LP is a convex polyhedral set that is closed. Indeed, it is the intersection of the set of feasible solutions and a supporting hyperplane. This alone can be used to show that $\exists$ an extreme point optimal solution if an optimal solution and an extreme point solution exist.

5. Consider the LP:

$$[\min cx : x \geq 0; Ax = b] \tag{1.1}$$

   The set $S = \{b : 1.1 \text{ is feasible}\}$ is a closed convex polyhedral cone; it is the cone *generated* by the columns of A and some authors denote it by $pos(A)$. *A polyhedral cone can be defined in two ways*: (i) *it is generated by a finite number of generators or extreme rays*; (ii) *it is the intersection of a finite number of half spaces passing through the origin.* The set $T = \{b : 1.1 \text{has an optimal solution}\}$ is also a closed convex polyhedral cone. *Either or both of these cones may be empty.*

6. The set of feasible solutions to an LP of the form above may be thought of as a subset of $\mathbf{R^{n-m}}$ where A is an m x n matrix with full row rank. In this case a nondegenerate extreme point is one where $n - m$ hyperplanes meet; a degenerate one is where more meet.

7. The optimal value of the above LP may be thought of as a function of A, b, and c. It is convex in b (over the feasible set of b) for fixed A and c; convex in c for fixed A and b. Both these are piecewise linear. *One important question deals with the number of pieces for special cases.*

1.3 Proofs

1. $x, y \in S \cap T \implies x, y \in S$, and hence $\lambda x + (1 - \lambda)y \in S$; similarly $\lambda x + (1 - \lambda)y \in T$ and hence it is in $S \cap T$. Hence the result. The assertion about $\cap_{\alpha \in A} S_\alpha$ follows in a similar manner.

2. "Clearly" the assertion is true by definition for $n = 2$. Assume that it is true for $n \leq k$. Let $x =_i^k \sum_{=1}^{+1} \lambda_i x^i = \lambda_1 x_1 + (1 - \lambda_1)_i^k \sum_{=2}^{+1} \mu_i x_i = \lambda_1 x_1 + (1 - \lambda_1)z$ where $\mu_i = \lambda_i/(1 - \lambda_1) \geq 0$, and $\sum_{k=2}^{k+1} \mu_i = 1$. (*Please note that we have assumed that* $\lambda_1 > 0$; *but this is without loss since not all* λ_i *can be 0*). Since z is a convex combination of k members of S, $z \in S$; hence $x \in S$. *Please note that this proof will not generalize to the infinite case which needs a different kind of proof.*

3. To show that H is a convex set: Let $x, y \in H$. Hence $cx = \alpha$, and $cy = \alpha$. Thus, $\lambda cx + (1 - \lambda)cy = c(\lambda x + (1 - \lambda)y) = \alpha$. Hence H is convex. H^+, H^{++}, H^-, and H^{--}*are* shown to be convex in a similar manner. *Indeed, this is the most primitive way of showing a set to be convex.*

4. To show that D is convex: Let $x, y \in D$. Hence $\sum(x_i - a_i)^2 \leq \alpha^2$; and $\sum(y_i - a_i)^2 \leq \alpha^2$. Need to show that $\sum[\lambda(x_i + (1-\lambda)y_i) - a_i]^2 \leq \alpha^2$. $\sum[\lambda(x_i + (1-\lambda)y_i) - a_i]^2 = \sum[\lambda(x_i - a_i) + (1-\lambda)(y_i - a_i)]^2$. $= \lambda^2 \sum(x_i - a_i)^2 + (1-\lambda)^2 \sum(y_i - a_i)^2 + 2\lambda(1-\lambda)\sum(x_i - a_i)(y_i - a_i)$. $= \lambda \sum(x_i - a_i)^2 + (1-\lambda)\sum(y_i - a_i)^2 - \lambda(1-\lambda)\sum[(x_i - a_i) - (y_i - a_i)]^2$ $\leq \alpha^2$.

5. The set of feasible solutions is closed because the limit points of a sequence of points in the set is in the set. The rest is easy.

1.3.1 Theorems

T1 Consider the problem $\min_{y \in \hat{S}} d(x, y)$; note $d(x.y) = \|x - y\|$. Since distance is a continuous nonnegative function and $\hat{S}$ is closed, the minimum is achieved at a point $y^* \in \hat{S}$, and $y^* \neq x$. Consider the halfspace $H^+ = \{y : 2(y^* - x)^t y \geq (y^* - x)^t (y^* + x)\}$. It is easy to show that $y^* \in H^+$ and $x \notin H^+$. For $y \in \hat{S}$, $\lambda y + (1-\lambda)y^* \in \hat{S}$ since S is convex. By the definition of y^*, we have:

$$\|[\lambda y + (1-\lambda)y^*] - x\|^2 = \|y^* + \lambda(y - y^*) - x\|^2 \geq \|y^* - x\|^2$$

This on simplification yields:

$$2\lambda(y - y^*)^t(y^* - x) + \lambda^2(y - y^*)^t(y - y^*) \geq 0 \; M \; \forall \; \lambda \in [0, 1]$$

Hence:

$$(y - y^*)^t(y^* - x) \geq 0 > (x - y^*)^t(y^* - x)/2$$

and therefore,
$2(y - y^*)^t(y^* - x) + (y^* - x)^t(y^* - x) > 0$, and hence
$(2y - y^* - x)^t(y^* - x) > 0$, and hence
$2y^t(y^* - x) > (y^* + x)^t(y^* - x)$ which is the required result. Hence $\hat{S} \subset H^{++}$, and $x \in H^{--}$.

Since x is a boundary point of S, $\exists$ a sequence of points $x_\nu \notin \hat{S} \ni \lim\ x_\nu = x$. (Choose each of these to be in smaller and smaller spheres centered at x and from outside $\hat{S}$). For x_ν, $\exists\ a_\nu \ni \|a_\nu\| = 1$, and $a_\nu^t y > a_\nu^t x_\nu \forall y \in \hat{S}$. Let a be a limit point of the sequence $\{a_\nu\}$. (Such points exist since all of these are in a compact set). Then $a^t y \geq a^t x \; \forall \; y \in \hat{S}$ with equality holding for x. This is the required supporting plane.

??If S is a closed convex set and T is the intersection of all its supporting halfspaces, then "clearly" $S \subseteq T$. If $x \in T - S$, then by theorem 1, $\exists\ H \ni S \subset H^{++}$ and $x \notin H^{++}$, and this is a contradiction to the definition of T.

1.3.2 Functions

1. Let $f(x) = \sup\ f_\nu(x)$. Since $f_\nu(\lambda x + (1-\lambda)y) \leq \lambda f_\nu(x) + (1-\lambda) f_\nu(y) \forall \nu$, the result follows.

2. Let $N(x^0)$ be the neighborhood of x^0. If $f(x^0) > f(x^1)$, then $\exists z = \lambda x^0 + (1-\lambda)x^1$ and $z \in N(x^0)$ with $f(z) < f(x^0)$.

1.4 LP Related Results

1. Let x be a basic feasible solution corresponding to the basis B (think of a basis as a set of linearly independent columns of A). By this we mean that $x^{NB} = 0$, and $Bx^B = b$ has a unique solution. If $x = \lambda y + (1-\lambda)z$, then $y^{NB} = z^{NB} = 0$; hence $y^B = z^B = x^B$. Hence $x = y = z$. Hence x is an extreme point.

Conversely, suppose x is an extreme point, and let B be the columns corresponding to positive components of x. Suppose these are dependent; $\exists\ z^B \neq 0 \ni Bz^B = 0$. By letting other components equal to 0 we get $z \neq 0 \ni Az = 0$. Consider the vector $x + \theta z$. Since $x^{NB} = z^{NB} = 0; (x+\theta z)^{NB} = 0\ \forall\ \theta$; also since $Az = 0; A(x+\theta z) = b\ \forall\ \theta$. Since $x^B > 0, \exists\ \theta > 0 \ni (x \pm \theta z)^B \geq 0$. This is a contradiction of extremeness of x. Hence the columns of B are independent and hence x is a basic feasible solution.