1 Linear Programming Definitions (based on Karloff's book).

A linear program in general form is the problem of minimizing a linear function subject to a finite number of equality and inequality constraints. Where $x, c \in \mathbb{R}^n$, A is a $k \times n$ matrix, $b \in \mathbb{R}^k$, A' is $l \times n$, and $b' \in \mathbb{R}^l$, the following is a general linear program with n variables, k equality constraints, and l inequality constraints. The redundant notation $x_j \lesseqgtr 0$ means that variable x_j is not sign-constrained. Either k or l can be 0.

$$\begin{aligned} &\min c^T x \\ \text{subject to}\quad &Ax = b \\ &A'x \geq b' \end{aligned}$$

$$x_1 \geq 0 \quad x_2 \geq 0 \quad \cdots \quad x_r \geq 0 \quad x_{r+1} \lesseqgtr 0 \quad \cdots \quad x_n \lesseqgtr 0$$

Let us look at three special classes of linear programs.

Standard form: $\begin{aligned} &\min c^T x, \\ \text{s.t. } &Ax = b \\ &x \geq 0 \end{aligned}$ in which every variable is sign-constrained and only equality constraints are allowed, and

Canonical form: $\begin{aligned} &\min c^T x, \\ \text{s.t. } &Ax \geq b \\ &x \geq 0 \end{aligned}$ a form which allows only inequality constraints and in which all variables are sign-constrained. If all entries in the arrays c, A, and b are non-negative we have a **covering linear program**.

Pseudo-packing form : $\begin{aligned} &\max c^T x, \\ \text{s.t. } &Ax \leq b \end{aligned}$ a maximization version in which only inequality constraints are allowed. If in addition every variable is sign-constrained (this can be written as $-x_i \leq 0$), and all entries in the arrays c, A, and b are non-negative we have a **packing linear program**.

All four forms are equivalent in that any one can easily and quickly be converted to any other. Clearly, a linear program in standard , canonical form , or pseudo-packing form is already in general form, so to prove equivalence it suffices to show that general-form linear programs can be converted to standard , canonical, and pseudo-packing forms.

2 Linear Algebra and Geometry Refresher

Conventions. In this class, all vectors will be column vectors, unless explicitly defined to be row vectors. Sometimes $v_1, v_2, ..., v_n$ will denote the components of n-vector v, sometimes n distinct vectors. Whenever we mean the latter we will say so explicitly. The jth column of an array A (as a column vector) will be denoted A_j; its ith row, *as a row vector*, will be denoted A^i.

A *vector space* or *linear space* S (in $\mathbb{R}^n$) is a nonempty subset of $\mathbb{R}^n$ closed under vector addition and scalar multiplication. Vectors $v_1, \ldots, v_r$ are *linearly independent* if and only if whenever $\sum c_i v_i = 0$ and $c_i \in \mathbb{R}$, then $c_i = 0$ for all i. Vectors that are not linearly independent are *linearly dependent.* A vector v is a *linear combination* of vectors $v_1, \ldots, v_r$ if there exists $c_i \in \mathbb{R}$ such that

$v = \sum_{i=1}^{r} c_i v_i$. Alternatively, a set of vectors is linearly independent if none is a linear combination of the others. The dimension *dim*(S) of a linear space S is the maximum number of linearly independent vectors in S.

An *affine space* (in $\mathbb{R}^n$) is the translate of a linear space. Formally, $A \subseteq \mathbb{R}^n$ is an affine space if and only if $A = \{t + y \mid y \in S\}$ for a fixed n-vector t and linear space S. The dimension of A is defined to be $\dim(S)$. By extension, if B is an arbitrary subset of $\mathbb{R}^n$, $\dim(B)$ is defined to be the smallest dimension of any affine space containing B. For example, if $v_1, v_2, \ldots, v_k$ are arbitrary n-vectors, because

$$\{v_1, \ldots, v_k\} \subseteq \{v_1 + \sum_{i=2}^{k} \alpha_i (v_i - v_1) \mid \alpha_i \in \mathbb{R}\},$$

the dimension of $\{v_1, \ldots, v_k\}$ is at most $k - 1$.

If A is an $m \times n$ matrix, we define

$$\textit{column space}(A) = \{Ax \mid x \in \mathbb{R}^n\},$$

and

$$\textit{rank}(A) = \dim(\text{column space}(A)).$$

We say that A has *full rank* if rank(A) is its smaller dimension. *Nullspace*(A) is the vector space $\{x \mid Ax = 0\}$ and the *nullity* of A is the dimension of nullspace(A).

The first several theorems are standard in the theory of linear algebra.

Theorem 2.1 *Rank(A) is simultaneously the maximum number of linearly independent rows in A and the maximum number of linearly independent columns in A.*

Theorem 2.2 *Nullity*(A)+ *rank*$(A) = n$*, if A is* $m \times n$*.*

Let A be an $m \times n$ matrix and suppose $At = b$. The set of all solutions to the linear system $Ax = b$ is

$$\{x \mid Ax = b\} = \{t + y \mid Ay = 0\} = \{t + y \mid y \in \text{nullspace } (A)\};$$

all solutions can be found by adding one particular solution to the nullspace. Using Gaussian Elimination, we can solve a linear system, transform a matrix to upper echelon form, compute its rank and determinant, find its inverse if it's nonsingular, and so on. Gaussian Elimination runs in $O(n^3)$ steps on our exact-arithmetic model. (On a RAM, Gaussian Elimination *can* be implemented so as to run in polynomial time).

The *length* of a vector x is $\|x\| = \sqrt{x^T x}$.

Definition. A *hyperplane* in $\mathbb{R}^n$ is

$$H = \{x \mid a_1 x_1 + \cdots + a_n x_n = b\},$$

where not all $a_i = 0$. (It is easy to see that dim$(H) = n - 1$, since if, say, $a_1 \neq 0$, H can be written as $[b/a_1 \;\; 0 \;\; 0 \;\; \cdots \;\; 0]^T$ plus the nullspace of $1 \times n$ matrix $[a_1 \;\; a_2 \;\; \cdots \;\; a_n]$.) A *halfspace* is the set $\{x \mid a_1 x_1 + \cdots + a_n x_n \geq b\}$ where not all $a_i = 0$.

We say a set $T \subseteq \mathbb{R}^n$ is *bounded* if there is a real r such that $\|x\| \leq r$ for all $x \in T$.

Definition. A *polyhedron* is the intersection of finitely many halfspaces. A bounded nonempty polyhedron is called a *polytope*.

Example. The set $\{x \mid Ax \leq b, \}$ is the intersection of finitely many halfspaces, as is $\{x \mid Ax = b,\ x \geq 0\}$. If they are nonempty and bounded, they're polytopes.

We say a sequence of equality or inequality constraints is *linearly independent* if the sequence of coefficient vectors is linearly independent. The right-hand sides and whether the constraints are equalities or inequalities are ignored.

Definition. If $x, y, p \in \mathbb{R}^n$ and $p = \lambda x + (1 - \lambda)y,\ 0 \leq \lambda \leq 1$, then p is the *convex combination* of x and y. More generally, p is the *convex combination* of points $x_1, x_2, \ldots, x_r$ if $p = \sum_{i=1}^{r} \lambda_i x_i$, $\lambda_i \geq 0$, and $\sum_{i=1}^{r} \lambda_i = 1$. Point p is the *strict convex combination* of points $x_1, x_2, \ldots, x_r$ if $p = \sum_{i=1}^{r} \lambda_i x_i,\ \lambda_i > 0$, and $\sum_{i=1}^{r} \lambda_i = 1$.

Definition. $S \subseteq \mathbb{R}^n$ is *convex* if whenever $x, y \in S$, $\lambda x + (1 - \lambda)y \in S$ for all $\lambda \in [0, 1]$. That is, the line segment between x and y lies in S.

Definition. Given a set of vectors S, the *convex hull* of S is the set of all convex combinations of vectors of S.

It is easy to see that (a) the intersection of arbitrarily many convex sets is convex, (b) a convex combination of points in a convex set is itself in the set, and (c) halfspaces are convex sets.

3 Vertices

Definition. If P is a polyhedron, $v \in P$ is an *extreme point* or *vertex* if v cannot be written as the strict convex combination of two distinct points $x, y \in P$.

If v is extreme, one can't "slide" along any line through v and stay within P. Not even a 1-dimensional ball around v lies entirely in P.

A constraint is *tight* if it is satisfied with equality.

Lemma 3.1 *Suppose $P = \{x \in \mathbb{R}^n | Ax \leq b\}$ is a polyhedron. Then v is an extreme point of P if and only if there are n linearly independent constraints among the constraints $Ax \leq b$ that are tight at v.*

Since an equality is a pair of inequalities, the same result holds even if P is defined by both inequalities and equalities.

Say a minimization problem is *unbounded* if for each real B, there is a feasible point of cost less than B. We define unbounded maximization problems similarly.

Theorem 3.2 *In a pseudo-packing LP with A of rank n, suppose that $p \in F$, where $F = \{x \in \mathbb{R}^n | Ax \leq b\}$ is the set of feasible solutions. Then either the pseudo-packing LP, instance is unbounded or there is a vertex v of F satisfying $c^T v \geq c^T p$.*

Corollary 3.3 *If in a pseudo-packing LP with A of rank n, $F := \{x \in \mathbb{R}^n | Ax \leq b\}$ satisfies $F \neq \emptyset$, and there is a B with $c^T x \leq B$ for all $x \in F$, then there is an optimal vertex. Provided that the cost is bounded, linear programming is a finite problem!*

Note that this proof yields a polynomial-time algorithm to find a vertex v such that $c^T v \geq c^T p$.

Theorem 3.4 *If $P = \{x \in \mathbb{R}^n | Ax \leq b\}$ is a polytope, with A of rank n, then every $x \in P$ can be written as a convex combination of at most $n + 1$ vertices.*

In other words, a polytope is the convex hull of its vertices. There is a reverse to this statement.

Theorem 3.5 *Let S be a finite set of vectors. Then its convex hull is a polytope.*

Theorem 3.6 *Let $P = \{x \in \mathbb{R}^n | Ax \leq b\}$ be a polytope, with A of rank n, and v be a vertex of P. Then there exists $c \in \mathbb{R}^n$ such that the linear program* $\max c^T x$, s.t. $Ax \leq b$ *has v as its unique optimum solution.*

4 Standard Form LPs: Basic Solutions

Definition. If A is an $m \times n$ matrix of rank m, any m linearly independent columns are a *basis*. (They form a basis of A's column space, $\mathbb{R}^m$). Sometimes we will use the columns as the basis, sometimes just their indices.

Definition. Consider the linear program $\begin{array}{c} \min c^T x \\ Ax = b \\ x \geq 0 \end{array}$ where A is of rank m and is $m \times n$. If columns $j_1 < j_2 < \cdots < j_m$ are a basis, the corresponding *basic solution* x is defined as follows and ignores c. Let B be the nonsingular matrix consisting of columns $j_1, j_2, \ldots, j_m$ of A. Any variable x_l where $l \notin \{j_1, j_2, \ldots, j_m\}$ is set to 0 and is called *nonbasic*. For $k = 1, \ldots, m$, x_{j_k} is set to the kth component of $B^{-1}b$ and is known as *basic*; the basic variables may or may not be 0. (The basic solution x is the unique vector satisfying $Ax = b$ subject to the condition that $x_l = 0$ for $l \notin \{j_1, \ldots, j_m\}$.) If x satisfies the sign constraints $x \geq 0$, x is a *basic feasible solution* or *bfs*.

BASIC SOLUTIONS ARE VERY IMPORTANT! - see Theorem 4.1 below.

If $m \times n$ matrix A has rank m, then $\begin{array}{c} Ax = b \\ x \geq 0 \end{array}$ always has a basic solution. If the feasible region is nonempty, we will see that it must have a bfs as well.

Notation. Let us use *LPS* to denote

$$\text{“}\begin{array}{c} \min c^T x, \\ \text{s.t. } Ax = b \\ x \geq 0 \end{array} \qquad \text{where } A \text{ is an } m \times n \text{ matrix of rank } m.\text{”}$$

We use F to denote the feasible set $\{x \mid Ax = b, \ x \geq 0\}$. Throughout, i will generally denote a row index, running from 1 to m, and j will generally denote a column index, running from 1 to n.

Theorem 4.1 *In LPS, suppose that $w \in \mathbb{R}^n$. Then w is a vertex of the feasible set F if and only if w is a basic feasible solution of the LPS instance.*

Corollary 4.2 *There are only finitely many vertices.*

Proof. Each is a bfs, which is determined by an m-element subset of the columns. The number of bfs's is therefore at most $\binom{n}{m}$. ■

Definition. A bfs is *degenerate* if it has more than $n - m$ zeroes.

Theorem 4.3 *If two different bases correspond to a single bfs v, then v is degenerate.*

(The converse is false.)

Theorem 4.4 *In LPS, suppose that $p \in F$. Then either the LPS instance is unbounded or there is a vertex v of F satisfying $c^T v \leq c^T p$.*

Corollary 4.5 *If in LPS $F \neq \emptyset$ and $c^T x \geq B$ for all $x \in F$, then there is an optimal vertex. Provided that the cost is bounded below, linear programming is a finite problem!*

5 Definitions for the ellipsoid algorithm

A set of n-vectors $u_1, u_2, ..., u_m$ is *orthonormal* if each has unit length and distinct vectors have 0 dot product. Alternatively, $u_1, u_2, ..., u_m$ are orthonormal if and only if $u_k^T u_l = 0$ if $k \neq l$, and $u_k^T u_k = 1$ for all k. A real square matrix U is ıorthogonal if $UU^T = I$, i.e., its rows (and hence columns) are orthonormal.

A complex number λ is an *eigenvalue* of A if $Ax = \lambda x$ for some nonzero complex vector x. The vector x is known as the *eigenvector corresponding to* λ.

Theorem 5.1 *(The* **Spectral Theorem***)*
If A is a real symmetric matrix, all of A's eigenvalues are real. Furthermore, if A's eigenvalues are $\lambda_1, \lambda_2, ..., \lambda_n$, then

$$A = U \Lambda U^{-1},$$

where

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

and U is an orthogonal matrix whose ith column is the eigenvector of A corresponding to λ_i, normalized to have length one.

If A is a real symmetric matrix, we say A is *positive definite* if

$$x^T A x > 0$$

for all nonzero, real n-vectors x.

Theorem 5.2 *Let A be real symmetric. The following are equivalent.*

(a) A is positive definite.

(b) All eigenvalues of A are positive.

(c) $A = QQ^T$ for a nonsingular $n \times n$ real matrix Q.

The *size* of a $m \times n$ matrix A with rational entries $A_{ij} = p_{ij}/q_{ij}$, with $p's$ and $q's$ integers, is defined as $size(A) = mn \log mn + \sum_i \sum_j \lceil \lg(|p_{ij}| + 1) \rceil + \lceil \lg(|q_{ij}| + 1) \rceil$. The size of an LP program is obtained by adding the sizes of A, b, and c.