

- Given vectors  $x_1, \dots, x_k \in \mathbb{R}^n$  and  $\lambda_1, \dots, \lambda_k \geq 0$  with  $\sum_{i=1}^k \lambda_i = 1$ , we call  $x = \sum_{i=1}^k \lambda_i x_i$  a *convex combination* of  $x_1, \dots, x_k$ .

For a point  $p$  belonging to line segment  $p'p''$ , there exists a  $\alpha \in [0, 1]$  such that  $p = \alpha p' + (1 - \alpha)p''$ . That is,  $p$  is a convex combination of  $p'$  and  $p''$ .

- A set  $S$  is called a *convex set* if for every two points  $x, y \in S$ , every point on the line segment joining  $x$  and  $y$  is contained in  $S$ .



Equivalently, a set  $S$  is convex if for every set  $S'$  of points located in  $S$ , the convex combination of points in  $S'$  also belongs to  $S$ .

The intersection of any number of convex sets in  $\mathbb{R}^n$  is convex.

- Let  $S \subseteq \mathbb{R}^n$  be a convex set. The function  $f : S \rightarrow \mathbb{R}^1$  is a *convex function* in  $S$  if for any two points  $x, y \in S$ ,  $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$  for  $\lambda \in \mathbb{R}^1$  and  $0 \leq \lambda \leq 1$ . Ex.  $x^2, e^x, |x|$ .
- The *convex hull*  $CH(S)$  of a set  $S$  is the smallest convex set that contains  $S$ . That is, the convex hull of  $S$  is the intersection of all convex sets that contain  $S$ .

Let  $CC(S)$  be the set that precisely comprises of all the convex combinations of points in  $S$ . Then,  $CC(S) = CH(S)$ .

For any set  $S$  and  $S' \subseteq S$ ,  $CH(S') \subseteq CH(S)$ .

- Caratheodory's theorem*: For a point  $x \in \mathbb{R}^d$ , suppose  $x$  is expressed as a convex combination of  $k > d + 1$  number of points in  $\mathbb{R}^d$ . Then,  $x$  can be expressed as a convex combination of at most  $d + 1$  points.
- An *extreme point* of a convex set  $S$  is a point  $p$  in  $S$  which cannot be expressed as a convex combination of points in  $S - \{p\}$ . In other words, a point  $p \in S$  is an extreme point of  $S$  if and only if  $p$  is a vertex of  $CH(S)$ .
- A *vector space* or *linear space* in  $\mathbb{R}^n$  is a nonempty subset of  $\mathbb{R}^n$  closed under vector addition and scalar multiplication. The *dimension* of a linear space  $S$  is the maximum number of linearly independent vectors in  $S$ .

An *affine space*  $A$  in  $\mathbb{R}^n$  is the space resulting from adding a fixed vector  $t$  to all the elements of a linear space  $S$ , i.e.,  $A = \{t + y \mid y \in S\}$ . (For example, a hyperplane not necessarily passing through the origin is an affine space.) The dimension of  $A$  is the dimension of  $S$ . If  $B$  is an arbitrary subset of  $\mathbb{R}^n$ , then the dimension of  $B$  is the smallest dimension of any affine space containing  $B$ .

- The vectors  $v_1, \dots, v_r$  are *linearly independent* if  $\sum_i \alpha_i v_i = 0$  and  $\forall_i \alpha_i \in \mathbb{R}$  imply that  $\forall_i \alpha_i = 0$ .

The vectors  $v_1, \dots, v_r$  are *affinely independent* if  $\sum_i \alpha_i v_i = 0, \forall_i \alpha_i \in \mathbb{R}$ , and  $\sum_i \alpha_i = 0$  together imply that  $\forall_i \alpha_i = 0$ .

- A *hyperplane* in  $\mathbb{R}^n$  is an affine subspace of dimension  $n - 1$ ; in other words, it is the set of all solutions of a linear equation of the form  $a_1x_1 + \dots + a_nx_n = b$ , where  $a_1, \dots, a_n$  are not all 0.

The hyperplane with equation  $a_1x_1 + \dots + a_nx_n = b$  induces two *closed half-spaces*:  $\{x \in \mathbb{R}^n : x_1 + \dots + a_nx_n \leq b\}$  and  $\{x \in \mathbb{R}^n : x_1 + \dots + a_nx_n \geq b\}$ .

- A *polyhedron* (a.k.a., *H-polyhedron* or *polyhedral set*) in  $\mathbb{R}^n$  is a set of type  $P = \{x \in \mathbb{R}^n : Ax \leq b\}$  for some matrix  $A \in \mathbb{R}^{m \times n}$  and some vector  $b \in \mathbb{R}^m$ . That is, the convex set obtained by the intersection of a finite number of affine halfspaces is a polyhedron. If  $A$  and  $b$  are rational, then  $P$  is a *rational polyhedron*. A bounded H-polyhedron is called a *H-polytope* (a.k.a., *polytope*).

A subset  $P$  of  $\mathbb{R}^n$  is called a *polytope* (a.k.a., *V-polytope*) if  $P$  is the convex hull of a finite number of vectors.

*Minkowski-Weyl's theorem*: A subset  $P$  of  $\mathbb{R}^n$  is a polytope if and only if it is a bounded polyhedron.

- The *dimension of a convex polyhedron*  $P \subseteq \mathbb{R}^n$  is the smallest dimension of an affine subspace containing  $P$ . Equivalently, it is the largest  $d \leq n$  for which  $P$  contains points  $x_1, \dots, x_d$  such that the  $d$ -tuple of vectors  $(x_1 - x_0, \dots, x_d - x_0)$  is linearly independent. That is, the dimension is the maximum number of affinely independent points in  $P$  minus 1. When the dimension of  $P$  is  $n$  then  $P$  is *full-dimensional*.
- For a hyperplane  $H$ , if  $H \cap P \neq \emptyset$  and  $P$  is entirely contained in one of the closed half-spaces defined by  $H$  then  $H$  is called a *supporting hyperplane of  $P$* . Formally, let  $P = \{x : Ax \leq b\}$  be a nonempty polyhedron. If  $c$  is a nonzero vector for which  $\delta = \max\{cx : x \in P\}$  is finite, then  $\{x : cx = \delta\}$  is called a *supporting hyperplane of  $P$* .
- A *face of  $P$*  is  $P$  itself or the intersection of  $P$  with a supporting hyperplane of  $P$ .

$F$  is a face of  $P$  if and only if  $F = \{x \in P : A'x = b'\} \neq \emptyset$  for some subsystem  $A'x \leq b'$  of  $Ax \leq b$ .

A point  $x$  for which  $\{x\}$  is a face is called a *vertex of  $P$* . That is, a vertex is a 0-dimensional face. An edge is a 1-dimensional face.

A *facet of  $P$*  is a maximal face distinct from  $P$ . An inequality  $cx \leq \delta$  is *facet-defining* for  $P$  if  $cx \leq \delta$  for all  $x \in P$  and  $\{x \in P : cx = \delta\}$  is a facet of  $P$ .

- The *rank lemma*: Let  $P = \{x \mid Ax \leq b\}$  be a polyhedron in  $\mathbb{R}^n$  and let  $z \in P$ . Also, let  $A_z$  be the submatrix of  $A$  consisting of those rows  $a_i$  of  $A$  for which  $a_i z = b_i$ . Then  $z$  is an extreme point of  $P$  if and only if  $\text{rank}(A_z) = n$ .

A polyhedron  $P$  is *pointed* if it has at least one vertex.

- *Asymptotic upper bound theorem*: A  $d$ -dimensional polytope with  $n$  vertices has at most  $O(n^{\lfloor d/2 \rfloor})$  facets and no more than  $O(n^{\lfloor d/2 \rfloor})$  faces in total.

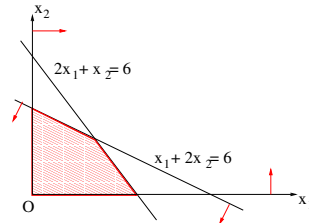
Since any polytope formed with  $f$  half-spaces can have at most  $f$  facets, in  $d$ -dimensions, the maximum number of vertices a polytope has is  $\Theta(f^{\lfloor d/2 \rfloor})$ .

- *Separation theorem*: Let  $S \subseteq \mathbb{R}^n$  be a finite set and let  $v \in \mathbb{R}^n \setminus \text{convexhull}(S)$ . Then there exists an inequality  $w^T x \leq t$  that separates  $v$  from  $\text{convexhull}(S)$ , that is,  $w^T s \leq t$  for all  $s \in \text{convexhull}(S)$  but  $w^T v > t$ .

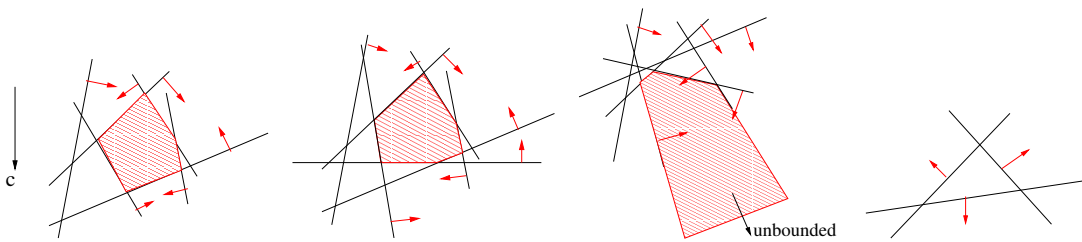
- *Theorem of the alternatives:* The system  $Ax \leq b$  has no solution  $x \in \mathbb{R}^n$  if and only if there exists  $y \in \mathbb{R}^m$  such that  $y \geq 0$ ,  $A^T y = 0$  and  $b^T y < 0$ .

*Farkas' lemma:* The system  $Ax = b$  has a nonnegative solution if and only if there is no vector  $y$  satisfying  $y^T A \geq 0$  and  $y^T b < 0$ .

- The linear constraints together form a convex feasible region, which is a convex polyhedra.



- The LP  $\min\{c^T x : Ax \geq b, x \geq 0\}$  is *unbounded* (from below) if  $\forall \lambda \in \mathbb{R}$ , there exists a feasible  $x$  such that  $c^T x < \lambda$ . (Analogous definition applies to maximization LP as well.) When the feasible region is empty, then the LP is said to be *infeasible*. Any linear program either (i) has an optimal solution with a finite objective value, (ii) is infeasible, or (iii) is unbounded.



- A subset of columns  $B$  of the constraint matrix  $A$  is called a *basis* if the matrix of columns corresponding to  $B$  is linearly independent.

The *basic solution* corresponding to  $B$  is a vector  $x \in \mathbb{R}^n$  with

$$x_j = 0 \text{ if } j \notin B, \text{ and}$$

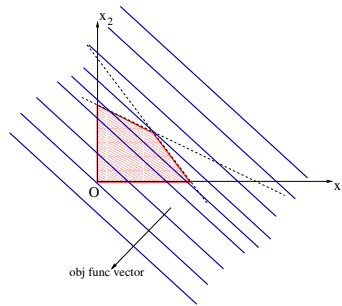
$$x_{j_k} = \text{the } k^{\text{th}} \text{ component of } B^{-1}b, \text{ for } k = 1, \dots, m.$$

If  $x$  is both basic and feasible, then  $x$  is called a *basic feasible solution*.

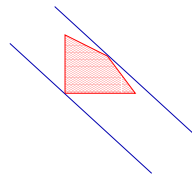
Let  $A$  be a  $m \times n$  matrix with full row rank. Then, every feasible  $x$  to  $P = \{x : Ax = b, x \geq 0\}$  is a basic feasible solution if and only if  $x$  is an extreme point solution. (Note that after introducing slack variables for each constraint, any LP can be transformed to a linear program in *standard form*:  $\{x : Ax = b, x \geq 0\}$ .)

Every basic feasible solution corresponds to one and only one vertex. And, there exists a basic feasible solution corresponding to every vertex of  $F$ . However, the correspondance between bases and basic feasible solutions is not one to one: there can be many bases, each of which corresponds to the same basic feasible solution. When more than two different bases correspond to the same basic feasible solution  $x$ , then  $x$  contains more than  $n - m$  zeros; hence,  $x$  is said to be a *degenerate* solution.

- The objective function is a family of parallel hyperplanes; the objective function value is fixed along any one such hyperplane. Significantly, for the set  $S$  of points satisfying the given constraints,  $\max\{w^T x : x \in S\} = \max\{w^T x : x \in \text{bd}(\text{CH}(S))\}$ .



An optimal solution corresponding to a basic feasible solution is called a *basic optimal solution*. As every supporting hyperplane contains at least one vertex, there exists a vertex at which optimum is guaranteed to occur.



Let  $P = \{x : Ax \geq b, x \geq 0\}$  and assume that  $\min\{c^T x : x \in P\}$  is finite. Then, for every  $x \in P$ , there exists an extreme point solution  $x' \in P$  such that  $c^T x' \leq c^T x$ , i.e., there is always an extreme point optimal solution.

Hence, LP is a combinatorial optimization problem even though it seem to be a continuous optimization problem! (Some say it lies on the boundary of continous and discrete optimization problems.)

- *Local optima is the global optima for any LP*: Consider an optimization problem with feasibility constraints  $F$  and objective function  $f$ , where  $F \subseteq R^n$  is a convex set and  $f$  is a convex function defined over the domain  $F$ . Then, the global optimality equals the local optimality whenever the locality is defined with respect to an  $\epsilon$ -radius Euclidean closed ball  $B$ , that is, if  $f(x)$  is locally optimum (say, minimum), then  $\forall y \in B, f(x) \leq f(y)$ .
- *Dantzig's simplex algorithm*:

A variable whose value is currently set to 0 is called a nonbasic variable; and, a variable whose value is currently not equal to 0 is called a basic variable. At any time of execution of the simplex algorithm, every variable is either a basic or a nonbasic variable.

In every iteration of simplex algorithm, one nonbasic variable becomes a basic variable and one basic variable becomes a nonbasic variable; for two adjacent vertices of a LP polytope, the basic and nonbasic sets are identical except for these. Hence, in every iteration, choose the nonbasic variable that gives the fastest rate of increase in the objective function value as the entering basic variable, and choose the basic variable that most limits the increase in the value of the entering basic variable as the leaving basic

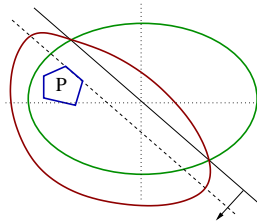
variable. Significantly, the updated solution vector in any iteration is guaranteed to improve the objective function value; besides, the new solution vector corresponds to a vertex of the LP polytope.

If the objective function value cannot be increased by increasing the values of any of the variables on which it depends, then the simplex must have arrived to optimality.

- (1) corresponding to a vertex of the LP polytope, initialize the basic variables.
- (2) while (the current solution is not optimal)
  - (i) choose entering and leaving basic variables
  - (ii) update the current solution vector

Using the asymptotic upper bound theorem, the worst-case time was proven to be exponential in the number of variables. Klee & Minty provided instances (known as Minty's cube) on which the algorithm indeed visits all the vertices of the polytope. However, it does well in practice. Specifically, under various probability distributions over the input, the simplex algorithm is proven to be taking polynomial time in the average-case.

- Kachiyan's *Ellipsoid algorithm*:



- (1) start with an ellipsoid  $E$  that contains all the basic feasible solutions of the given LP
- (2) invoke a *separation oracle*  $O$ , which is given as input, with center  $c$  of the ellipsoid  $E$ :
  - if  $c$  is feasible,  $O$  returns a constraint  $C: \sum_j d_j x_j \leq \sum_j d_j c_j$  for the objective function  $\sum_j d_j x_j$
  - else  $O$  returns a *violated constraint*  $C: \sum_j a_j x_j \leq b$  for the center  $c$ , i.e., the current set of basic feasible solutions and  $c$  are on different sides of  $C$
  - in both the cases, one of the hyperspaces  $H_C$  defined by  $C$  contains a basic optimal solution (if one exists for the given LP)
- (3) compute a minimum volume ellipsoid  $E'$  that contains the region common to both  $E$  and  $H_C$
- (4) set  $E := E'$ ; and go to step (2) unless the ellipsoid is sufficiently small so that it can contain at most one basic feasible solution, if any (if one exists, then this must be a basic optimal solution)

This algorithm takes  $O(n^6 Lt)$  time, where  $n$  is the number of variables,  $L$  is the number of bits in encoding the input instance, and  $t$  is the time complexity of the separation oracle; the time does not depend on the number of constraints! In practice, this algorithm does not perform as good as the simplex algorithm. And, due to square-root operations involved, rounding errors need to be handled carefully.

However, the user of this algorithm requires to provide a separation oracle: given a point  $p$  either determine that  $p$  is in the polytope  $P$  corresponding to the given LP or find a constraint  $a^T x \leq b$  that contains  $P$  but  $a^T p > b$ . For some LPs, oracles that do better than explicitly checking  $p$  versus each constraint are

possible, i.e., a polynomial-time oracle exists even when the number of constraints is exponential! Hence, subjected to the efficiency of separation oracle provided, the Ellipsoid method take only polynomial time even when the number of constraints is exponential.

Ex. relaxed LP for arborescence rooted at  $r$ :

$$\begin{aligned} \min \quad & \sum_{(i,j) \in E} c_{ij} x_{ij} \\ \sum_{(i,j) \in E, i \in S, j \in V-S} x_{ij} & \geq 1 \quad \text{for every } S \subseteq V, r \in S \quad \text{--- (1)} \\ 0 \leq x_{ij} & \leq 1 \quad \text{for every } (i, j) \in E \end{aligned}$$

separation oracle: detecting a violated constraint among inequalities in (1) is equivalent to finding a  $r$ - $t$  min-cut for some  $t \in V - \{r\}$  such that the cut capacity is less than one when the weight of arc  $(i, j)$  is set to  $x_{ij}$  for every arc  $(i, j)$ ; other constraints can be checked explicitly

- Karmakar's *interior point algorithm*: takes  $O(n^{3.5} L^2 (\lg L) (\lg \lg L))$  time; in practice, competes with (or, does better than) the simplex algorithm. Here,  $n$  is the number of variables, and  $L$  is the number of bits in encoding the input instance. In practice, this algorithm does not perform as good as the simplex algorithm. And, like in Ellipsoid algorithm, there are rounding errors due to square-root operations.

The hunt is on for a strongly polynomial time algorithm.

- *Primal-dual LP pairs* -

<p>primal LP:</p> $\begin{aligned} \min \quad & \sum_{i=1}^n c_i x_i \\ a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n & = b_i \quad i \in M \\ a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n & \geq b_i \quad i \in \overline{M} \\ x_j & \geq 0 \quad j \in N \\ x_j & <> 0 \quad j \in \overline{N} \end{aligned}$	<p>dual LP:</p> $\begin{aligned} \max \quad & \sum_{j=1}^m b_j y_j \\ y_i & <> 0 \quad i \in M \\ y_i & \geq 0 \quad i \in \overline{M} \\ a_{1j}y_1 + a_{2j}y_2 + \dots + a_{mj}y_m & \leq c_j \quad j \in N \\ a_{1j}y_1 + a_{2j}y_2 + \dots + a_{mj}y_m & = c_j \quad j \in \overline{N} \end{aligned}$
--	--

- *Weak duality theorem*: Let  $x^*, y^*$  be feasible solutions of primal and dual LPs respectively. Assuming the primal LP is a minimization LP,

$$\sum_{i=1}^m b_i y_i^* \leq \sum_{j=1}^n c_j x_j^*$$

A consequence of weak duality theorem: If  $\sum_{i=1}^m b_i y_i^* = \sum_{j=1}^n c_j x_j^*$  then  $x^*$  and  $y^*$  are optimal solutions to the primal and dual LPs respectively.

- *Strong duality theorem*: The primal LP has an optimal solution if and only if the dual LP has an optimal solution. At optimality, cost of primal LP equals to the cost of dual LP.
- *Complementary slackness conditions*: A pair  $x^*, y^*$  respectively feasible in a primal-dual LP pair is optimal if and only if

$$\forall_i y_i ((a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n) - b_i) = 0, \text{ and}$$

$$\forall_j (c_j - (a_{1j}y_1 + a_{2j}y_2 + \dots + a_{mj}y_m))x_j = 0.$$

The former conditions are known as the primal complementary slackness conditions whereas the latter are the dual complementary slackness conditions.

- For primal and dual LPs  $P$  and  $D$ , one of the following four statements must hold: (i) both  $P$  and  $D$  are feasible, (ii)  $P$  is infeasible and  $D$  is unbounded, (iii)  $P$  is unbounded and  $D$  is infeasible, and (iv) both  $P$  and  $D$  are infeasible.
- If the coefficients in the constraint matrix, objective function, and the RHS vector are all non-negative in the LP, then that LP is called a *covering LP*. The dual of a covering LP is called a *packing LP*.

- Let  $P$  be a polytope and let  $x$  be an extreme point solution of  $P$ , then  $x$  is *integral* if each coordinate of  $x$  is an integer. The polytope  $P$  is called *integral* if every extreme point of  $P$  is integral.
- The integer linear programming (ILP) is a linear program in which the variables are restricted to be integers. The decision version of the ILP problem is NP-complete.
- A square, integer matrix  $M$  is called *unimodular (UM)* if  $\det(M) = \pm 1$ . An integer matrix  $M$  is called *totally unimodular (TUM)* if every square, nonsingular submatrix of  $M$  is unimodular. Examples for TUM matrices include node-arc incidence matrix of a directed graph (occurs in ILPs corresponding to shortest path, maximum flow, Hitchcock problems) and the node-edge incidence matrix of an undirected bipartite graph (occurs in ILPs of maximum weighted bipartite perfect matching).

If  $A$  is TUM and  $b$  is integral, then the basic solutions corresponding to both LPs  $\{x : Ax = b, x \geq 0, x \text{ integer}\}$  and  $\{x : Ax = b, x \geq 0\}$  are same.

Analogously, if  $A$  is TUM and  $b$  is integral, then the basic solutions corresponding to both LPs  $\{x : Ax \leq b, x \geq 0, x \text{ integer}\}$  and  $\{x : Ax \leq b, x \geq 0\}$  are same.

- For a minimization (resp. maximization) ILP, the *integrality gap* of an ILP is defined as the supremum of  $\frac{OPT(ILP)}{OPT(LP_{relaxed})}$  (resp.  $\frac{OPT(ILP)}{OPT(LP_{relaxed})}$ ) over all problem instances.

The integrality gap enforces a limit on the approximation power of our relaxation: It is not possible, at least for any algorithm that derives its performance guarantee by comparing its value to that of the LP relaxation, to get an approximation factor better than the integrality gap of that ILP.

Ex: Consider the following LP relaxation for the vertex cover ILP:

$$\begin{aligned} \min \sum_{v \in V} x_v \quad & \text{s.t.} \\ x_v + x_w & \geq 1 \quad \forall (v, w) \in E \\ x_v & \geq 0 \quad \forall v \in V \end{aligned}$$

For  $K_n$ , an integral solution need to choose  $n - 1$  vertices. Again, a feasible fractional solution assigns  $x_v = \frac{1}{2}$  to all vertices, with the objective value  $\frac{n}{2}$ . Therefore, the integrality gap for  $K_n$  is at least  $\frac{n-1}{n/2}$ , which is equal to  $2(1 - \frac{1}{n})$ , i.e., the integrality gap approaches 2 as  $n$  tends to infinity.

References:

Combinatorial Optimization by C. H. Papadimitriou and K. Steiglitz.

Lectures on Discrete Geometry by J. Matousek.