5. Linear Inequalities and Elimination

- Searching for certificates
- Projection of polyhedra
- Quantifier elimination
- Constructing valid inequalities
- Fourier-Motzkin elimination
- Efficiency
- Certificates
- Farkas lemma
- Representations
- Polytopes and combinatorial optimization
- Efficient representations

Searching for Certificates

Given a feasibility problem

$$\text{does there exist } x \text{ such that } f_i(x) \leq 0 \text{ for all } i = 1, \ldots, m$$

We would like to find certificates of infeasibility. Two important methods include

- Optimization
- Automated inference, or *constructive* methods

In this section, we will describe some constructive methods for the special case of *linear equations and inequalities*

Polyhedra

A set $S \subset \mathbb{R}^n$ is called a *polyhedron* if it is the intersection of a finite set of closed halfspaces

$$S = \left\{ x \in \mathbb{R}^n \mid Ax \leq b \right\}$$

- A bounded polyhedron is called a *polytope*
- The *dimension* of a polyhedron is the dimension of its affine hull

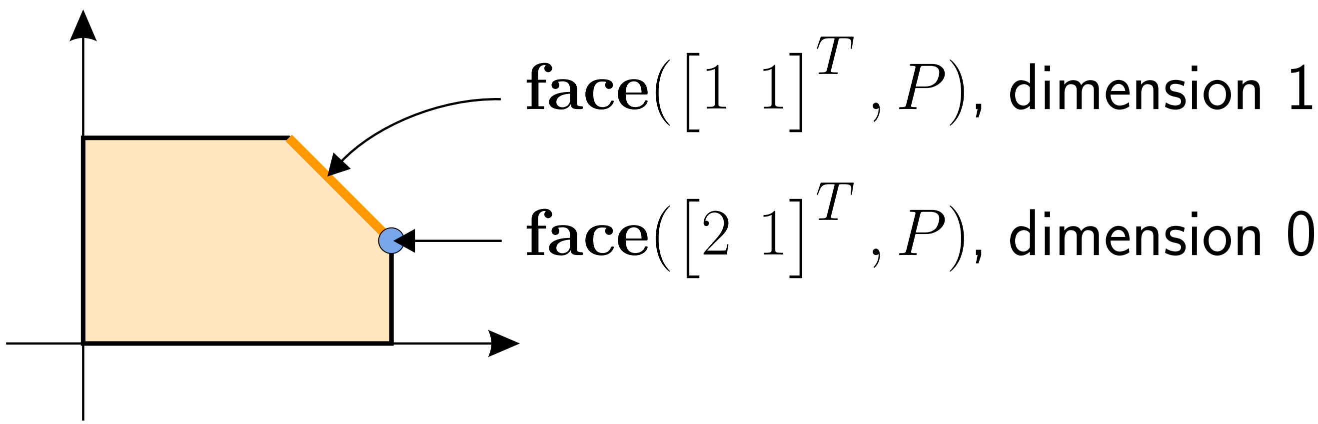
$$\mathbf{affine}(S) = \left\{ \lambda x + \nu y \mid \lambda + \nu = 1,\; x, y \in S \right\}$$

- If $b = 0$ the polyhedron is a cone
- Every polyhedron is convex

Faces of Polyhedra

given $a \in \mathbb{R}^n$, the corresponding *face* of polyhedron P is

$$\mathbf{face}(a, P) = \left\{ x \in P \mid a^T x \geq a^T y \text{ for all } y \in P \right\}$$



- Faces of dimension 0 are called *vertices*
  1 *edges*
  $d-1$ *facets*, where $d = \dim(P)$
- Facets are also said to have *codimension* 1

Projection of Polytopes

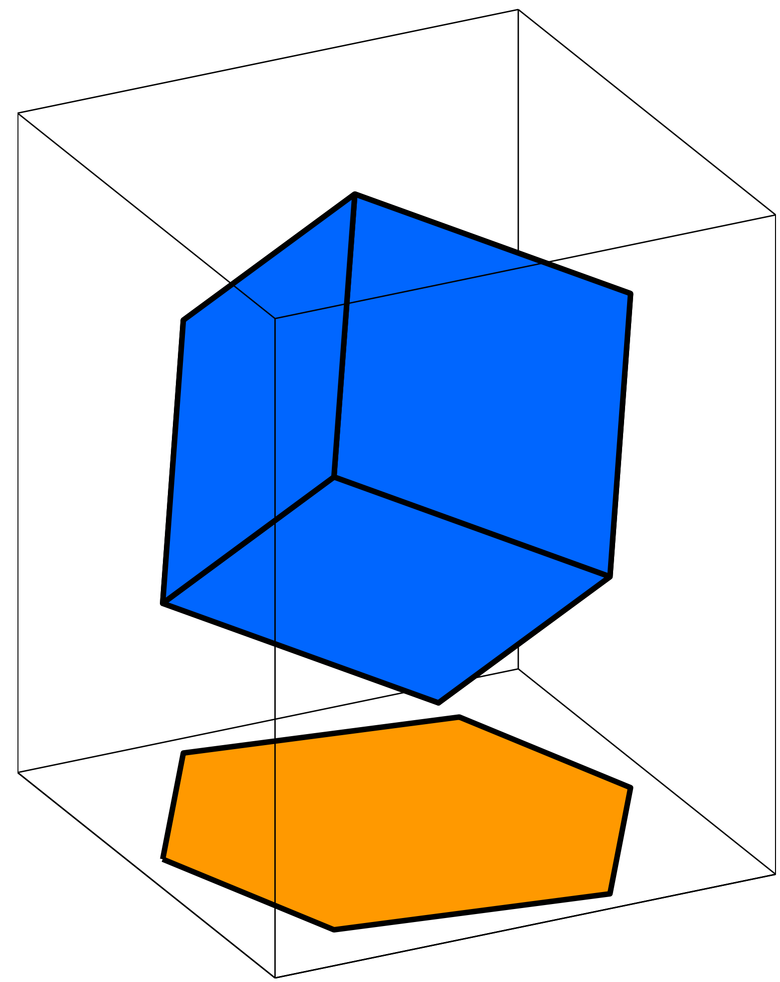
Suppose we have a polytope

$$S = \left\{ x \in \mathbb{R}^n \mid Ax \leq b \right\}$$

We'd like to construct the projection onto the hyperplane

$$\left\{ x \in \mathbb{R}^n \mid x_1 = 0 \right\}$$

Call this projection $P(S)$



In particular, we would like to find the *inequalities* that define $P(S)$

Projection of Polytopes

We have

$$P(S) = \left\{ x_2 \mid \text{there exists } x_1 \text{ such that } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in S \right\}$$

- Our objective is to perform *quantifier elimination* to remove the existential quantifier and find a *basic semialgebraic* representation of $P(S)$

- Alternatively, we can interpret this as finding *valid inequalities* that do not depend on x_1; i.e., the intersection

$$\mathbf{cone}\{f_1, \ldots, f_m\} \cap \mathbb{R}[x_2, \ldots, x_n]$$

This is called the *elimination cone* of valid inequalities

Projection of Polytopes

- Intuitively, $P(S)$ is a polytope; what are its vertices?

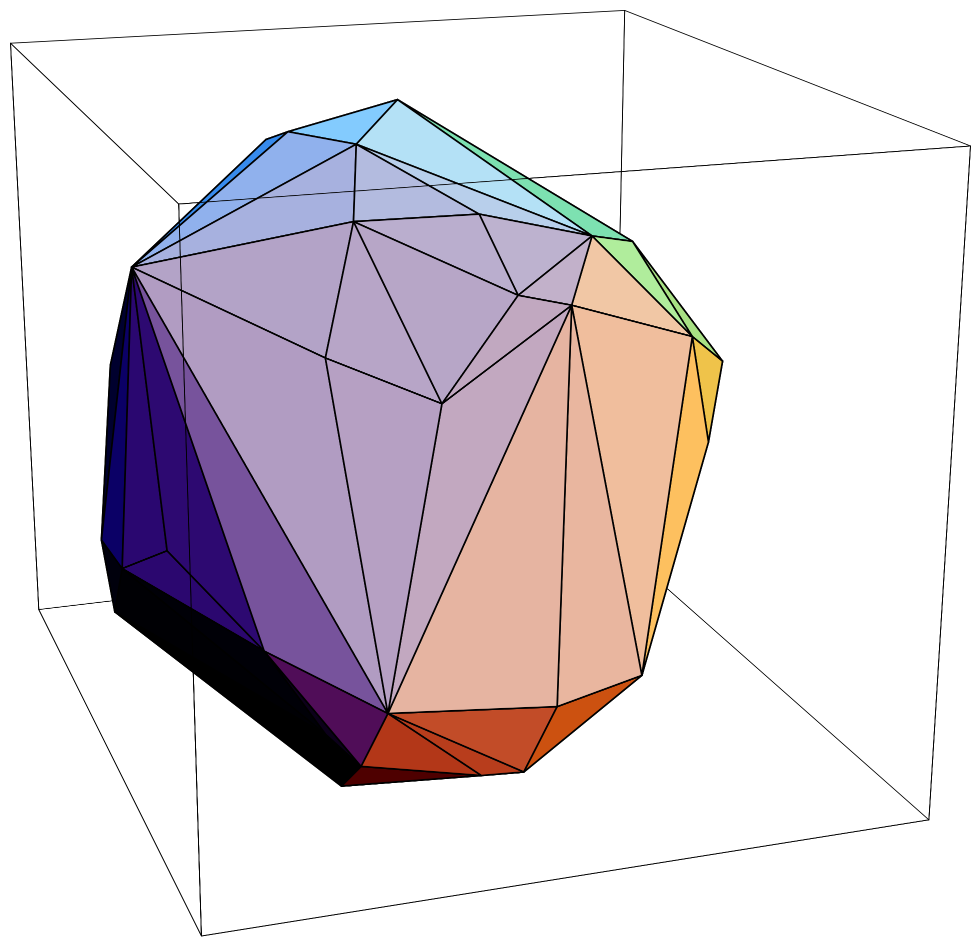
> Every face of $P(S)$ is the projection of a face of S

- Hence every vertex of $P(S)$ is the projection of some vertex of S
- What about the facets?
- So one algorithm is
  - Find the vertices of S, and project them
  - Find the convex hull of the projected points

But how do we do this?

Example

- The polytope S has dimension 55, 2048 vertices, billions of facets
- The 3d projection $P(S)$ has 92 vertices and 74 facets



Simple Example

$$-4x_1 - x_2 \leq -9 \qquad (1)$$

$$-x_1 - 2x_2 \leq -4 \qquad (2)$$

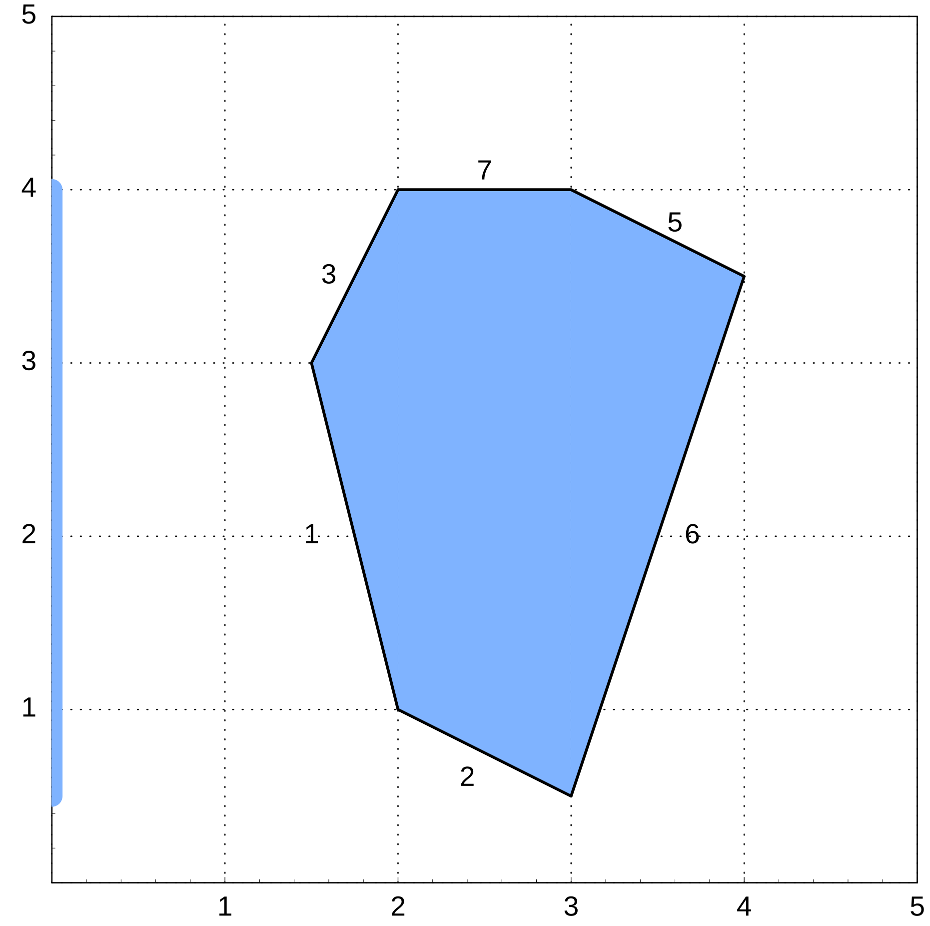
$$-2x_1 + x_2 \leq 0 \qquad (3)$$

$$-x_2 - 6x_2 \leq -6 \qquad (4)$$

$$x_1 + 2x_2 \leq 11 \qquad (5)$$

$$6x_1 + 2x_2 \leq 17 \qquad (6)$$

$$x_2 \leq 4 \qquad (7)$$



Constructing Valid Inequalities

We can generate new valid inequalities from the given set; e.g., if

$$a_1^T x \le b_1 \qquad \text{and} \qquad a_2^T x \le b_2$$

then

$$\lambda_1(b_1 - a_1^T x) \; + \; \lambda_2(b_2 - a_2^T x) \ge 0$$

is a valid inequality for all $\lambda_1, \lambda_2 \ge 0$

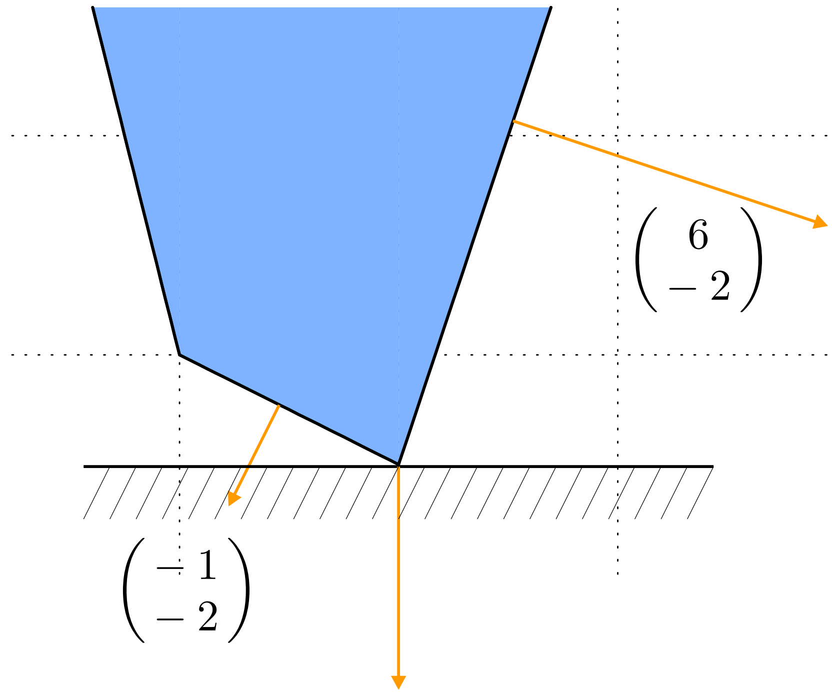
Here we are applying the inference rule, for $\lambda_1, \lambda_2 \ge 0$

$$f_1, f_2 \ge 0 \qquad \Longrightarrow \qquad \lambda_1 f_1 + \lambda_2 f_2 \ge 0$$

Constructing Valid Inequalities

For example, use inequalities (2) and (6) above

$$-x_1 + 2x_2 \leq -4$$
$$6x_1 - 2x_2 \leq 17$$



Pick $\lambda_1 = 6$ and $\lambda_2 = 1$ to give

$$6(-x_1 - 2x_2) + (6x_1 - 2x_2) \leq 6(-4) + 17$$
$$-2x_2 \leq 1$$

- The corresponding vector is in the *cone* generated by a_1 and a_2
- If a_1 and a_2 have opposite sign coefficients of x_1, then we can pick some element of the cone with x_1 coefficient zero.

Fourier-Motzkin Elimination

Write the original inequalities as

$$\left.\begin{array}{r} \dfrac{x_2}{4} + \dfrac{9}{4} \\ -2x_2 + 4 \\ \dfrac{x_2}{2} \\ -6x_2 + 6 \end{array}\right\} \leq x_1 \leq \left\{\begin{array}{l} -2x_2 - 11 \\ -\dfrac{x_2}{3} + \dfrac{17}{6} \end{array}\right.$$

along with $x_2 \leq 4$

Hence every expression on the left hand side is less than every expression on the right, for every $(x_1, x_2) \in P$

Together with $x_2 \leq 4$, this set of pairs specifies exactly $P(S)$

The Projected Set

This gives the following system of inequalities for $P(S)$

$$x_2 \leq 5 \qquad -x_2 \leq 1 \qquad 0 \leq 7 \qquad -x_2 \leq -\frac{1}{2} \qquad x_2 \leq 4\tfrac{2}{5}$$

$$x_2 \leq 17 \qquad -x_2 \leq \tfrac{4}{5} \qquad -x_2 \leq -\tfrac{1}{2} \qquad x_2 \leq 4$$

- There are many redundant inequalities
- $P(S)$ is defined by the tightest pair

$$-x_2 \leq -\tfrac{1}{2} \qquad x_2 \leq 4$$

- When performing repeated projection, it is very important to eliminate redundant inequalities

Efficiency of Fourier-Motzkin elimination

If A has m rows, then after elimination of x_1 we can have no more than

$$\left\lfloor \frac{m^2}{4} \right\rfloor \text{ facets}$$

- If $m/2$ inequalities have a positive coefficient of x_1, and $m/2$ have a negative coefficient, then FM constructs exactly $m^2/4$ new inequalities
- Repeating this, eliminating d dimensions gives

$$\left\lfloor \frac{m}{2} \right\rfloor^{2^d} \text{ inequalities}$$

- Key question: how many are redundant? i.e., does projection produce exponentially more facets?

Inequality Representation

Constructing such inequalities corresponds to multiplication of the original constraint $Ax \le b$ by a positive matrix C

In this case

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 4 & 0 & 0 \\ 6 & 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 6 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad A = \begin{bmatrix} -4 & -1 \\ -1 & -2 \\ -2 & 1 \\ -1 & -6 \\ 1 & 2 \\ 6 & -2 \\ 0 & 1 \end{bmatrix} \qquad b = \begin{bmatrix} -9 \\ -4 \\ 0 \\ -6 \\ 11 \\ 17 \\ 4 \end{bmatrix}$$

Inequality Representation

The resulting inequality system is $CAx \le Cb$, since

$$x \ge 0 \text{ and } C \ge 0 \qquad \Longrightarrow \qquad Cx \ge 0$$

We find

$$CA = \begin{bmatrix} 0 & 7 \\ 0 & -14 \\ 0 & 0 \\ 0 & -14 \\ 0 & 5 \\ 0 & 2 \\ 0 & -4 \\ 0 & -38 \\ 0 & 1 \end{bmatrix} \qquad Cb = \begin{bmatrix} 35 \\ 14 \\ 7 \\ -7 \\ 22 \\ 34 \\ 5 \\ -19 \\ 4 \end{bmatrix}$$

Feasibility

In the example above, we eliminated x_1 to find

$$-x_2 \leq -\frac{1}{2} \qquad\qquad x_2 \leq 4$$

We can now eliminate x_2 to find

$$0 \leq \frac{7}{2}$$

which is obviously true; it's valid for every $x \in S$, but happens to be independent of x

If we had arrived instead at

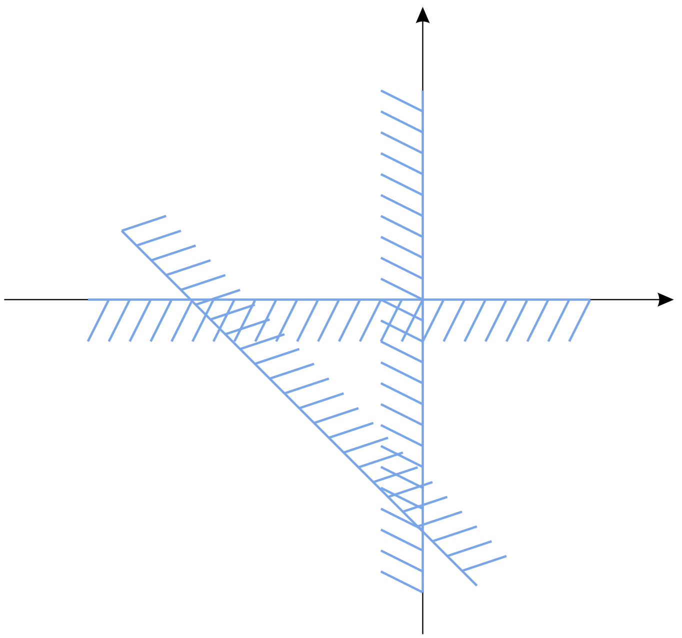
$$0 \leq -2$$

then we would have derived a contradiction, and the original system of inequalities would therefore be *infeasible*

Example

Consider the infeasible system

$$x_1 \geq 0$$
$$x_2 \geq 0$$
$$x_1 + x_2 \leq -2$$



Write this as $-x_1 \leq 0 \qquad x_1 + x_2 \leq -2 \qquad -x_2 \leq 0$

Eliminating x_1 gives

$$x_2 \leq -2 \qquad -x_2 \leq 0$$

Subsequently eliminating x_2 gives the contradiction

$$0 \leq -2$$

Inequality Representation

The original system is $Ax \leq b$ with $A = \begin{bmatrix} -1 & 0 \\ 1 & 1 \\ 0 & -1 \end{bmatrix}$ and $b = \begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix}$

To eliminate x_1, multiply

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} (Ax - b) \leq 0$$

Similarly to eliminate x_2 we form

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} (Ax - b) \leq 0$$

Certificates of Infeasibility

The final elimination is

$$\begin{bmatrix}1 & 1 & 1\end{bmatrix}(Ax - b) \leq 0$$

Hence we have found a vector λ such that

- $\lambda \geq 0$ (since its a product of positive matrices)
- $\lambda^T A = 0$ and $\lambda^T b < 0$ (since it gives a contradiction)

Fourier-Motzkin constructs a *certificate* of infeasibility; the vector λ

- Exactly decides feasibility of linear inequalities
- Hence this gives an extremely inefficient way to solve a linear program

Farkas Lemma

Hence Fourier-Motzkin gives a proof of Farkas lemma

The primal problem is

$$\exists x \quad Ax \leq b$$

The dual problem is a strong alternative

$$\exists \lambda \quad \lambda^T A = 0, \; \lambda^T b < 0, \; \lambda \geq 0$$

The beauty of this proof is that it is *algebraic*

- It does not require any compactness or topology
- It works over general fields, e.g. $\mathbb{Q}$,
- It is a *syntactic proof*, just requiring the axioms of positivity

Gaussian Elimination

We can also view Gaussian elimination in the same way

- Constructing linear combination of rows is *inference*

  Every such combination is a valid equality

- If we find $0x = 1$ then we have a proof of infeasibility

The corresponding strong duality result is

- Primal: $\exists\, x \quad Ax = b$
- Dual: $\exists\, \lambda \quad \lambda^T A = 0, \lambda^T b \neq 0$

Of course, this is just the usual range-nullspace duality

Computation

One feature of FM is that it allows *exact rational arithmetic*

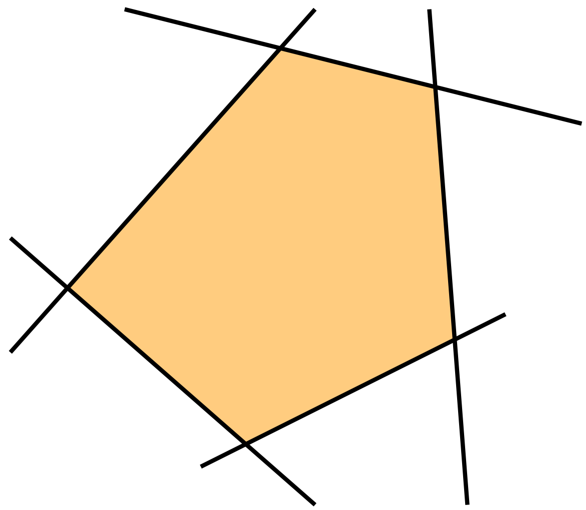
- Just like Groebner basis methods
- Consequently very slow; the numerators and denominators in the rational numbers become large
- Even Gaussian elimination is slow in exact arithmetic (but still polynomial)

Optimization Approach

- Solving the inequalities using interior-point methods is much faster than testing feasibility using FM
- Allows floating-point arithmetic
- We will see similar methods for polynomial equations and inequalities

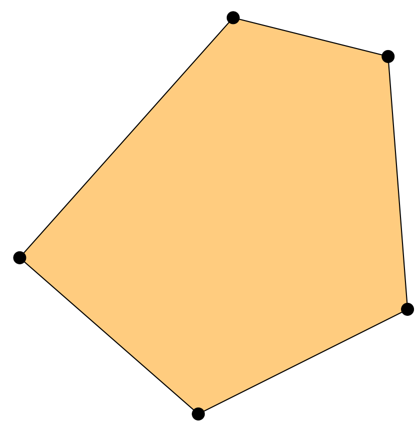
Representation of Polytopes

We can represent a polytope in the following ways



- *an intersection of halfspaces*, called an H-polytope

$$S = \left\{ x \in \mathbb{R}^n \mid Ax \leq b \right\}$$



- *the convex hull of its vertices*, called a V-polytope

$$S = \mathbf{co}\left\{ a_1, \ldots, a_m \right\}$$

Size of representations

In some cases, one representation is smaller than the other

- The n-cube

$$C_n = \left\{ x \in \mathbb{R}^n \mid -1 \leq x_i \leq 1 \text{ for all } i \right\}$$

  has $2n$ facets, and 2^n vertices

- The polar of the cube is the n-dimensional *crosspolytope*

$$\begin{aligned} C_n^* &= \left\{ x \in \mathbb{R}^n \mid \sum_i |x_i| \leq 1 \right\} \\ &= \mathbf{co}\,\{ e_1, -e_1, \ldots, e_n, -e_n \} \end{aligned}$$

  which has $2n$ vertices and 2^n facets

- Consequently *projection is exponential*.

Problem Solving using Different Representations

If S is a V-polytope

- *Optimization* is easy; evaluate $c^T x$ at all vertices
- To check *membership* of a given $y \in S$, we need to solve an LP; duality will give certificate of infeasibility

If S is an H-polytope

- *Membership* is easy; simply evaluate $Ay - b$

  The certificate of infeasibility is just the violated inequality
- *Optimization* is an LP

Polytopes and Combinatorial Optimization

Recall the MAXCUT problem

$$\begin{array}{rl} \text{maximize} & \mathbf{trace}(QX) \\ \text{subject to} & \mathbf{diag}\, X = 1 \\ & \mathbf{rank}(X) = 1 \\ & X \succeq 0 \end{array}$$

The *cut polytope* is the set

$$\begin{aligned} C &= \mathbf{co}\{\, X \in \mathbb{S}^n \mid X = vv^T,\ v \in \{-1,1\}^n \,\} \\ &= \mathbf{co}\{\, X \in \mathbb{S}^n \mid \mathbf{rank}(X) = 1,\ \mathbf{diag}(X) = 1,\ X \succeq 0 \,\} \end{aligned}$$

- Maximizing $\mathbf{trace}\, QX$ over $X \in C$ gives *exactly* the MAXCUT value
- This is equivalent to a *linear program*

MAXCUT

Although we can formulate MAXCUT as an LP, both the V-representation and the H-representation are exponential in the number of vertices

- e.g., for $n = 7$, the cut polytope has $116,764$ facets
  for $n = 8$, there are approx. $217,000,000$ facets

- Exponential description is not necessarily fatal; we may still have a polynomial-time *separation oracle*

- For MAXCUT, several families of valid inequalities are known, e.g., the *triangle inequalities* give LP relaxations of MAXCUT

Efficient Representation

- Projecting a polytope can dramatically change the number of facets
- Fundamental question: are polyhedral feasible sets the projection of higher dimensional polytope with fewer facets?
- If so, the problem is solvable by a simpler LP in higher dimensions

  The projection is performed *implicitly*

Convex Relaxation

- For any optimization problem, we can always construct an equivalent problem with a *linear* cost function
- Then, replacing the feasible set with its convex hull does not change the optimal value
- Fundamental question: how to efficiently construct convex hulls?