Feasibility and Representability of Polytopes

Jesús A. De Loera, UC Davis

July 15, 2009

Is there any solution of $Ax \geq b$?

- we say that the system of inequalities $Ax \geq b$ is **feasible** if there is at least one x that satisfies all the inequalities. We wish to know when and certify the feasibility/infeasibility of polyhedra.
- Analogously, in linear algebra,
  Fredholm's Lemma: $\{x : Ax = b\}$ is non-empty if and only if $\{y : y^T A = 0, y^T b = -1\}$ is empty.
  Such a vector y is a **mathematical proof** that $Ax = b$ has no solution.
- We will prove today
  Farkas Lemma: A polyhedron $\{x : Ax \leq b\}$ is non-empty if and only if there is no solution $\{y : y^T A = 0, y^b < 0, y \geq 0\}$.
- We will give an (inefficient) algorithmic proof of Farkas lemma using an algorithm that decides whether a polyhedron is feasible: **Fourier-Motzkin' algorithm**

Fourier-Motzkin Algorithm

INPUT: Polyhedron $P = \{x : Ax \leq b\}$
OUTPUT: Yes/No depending whether P is empty or not.

- If P is described in a single variable x, P is feasible if

$$\max(b_i/a_i \ : b_i/a_i < 0) \leq \min(b_j/a_j \ : b_j/a_j > 0)$$

- Else we eliminate leading variable (x_1). Re-write the inequalities to be regrouped in 3 groups:

$x_1 + (a_i')^T x' \leq b_i'$, **(if coefficient of a_{i1} is positive) (TYPE I)**

$x_1 + (a_j')^T x' \leq b_j'$, **(if coefficient of a_{j1} is negative) (TYPE II)**

$(a_k')^T x' \leq b_k'$, **(if coefficient of a_{k1} is zero) (TYPE III)**

Here $x' = (x_2, x_3, \ldots, x_n)$.

Fourier-Motzkin continued

Add all possible pairs of inequalities of (TYPE I) and (TYPE II). Create new system (with fewer variables):

- $$(a'_j + a'_i)^T x' \leq (b_j + b_i) \ \ \textbf{for } i \textbf{ of type I and } j \textbf{ of type II}$$

  $$\textbf{Keep equations of type III} (a'_k)^T x' \leq b'_k$$

- Original system of inequalities has a solution if and only if the system $(*)$ is feasible **WHY?**

  $(*)$ **is equivalent to** $(a'_j)^T x - b_j \leq b_i - (a'_i)^T x'$, **and** $(a'_k)^T x' \leq b'_k$

  If we find $x_2, x_3, \ldots, x_n$ satisfying $(*)$, find

  $$max((a'_j)^T x - b_j) \leq x_1 \leq min(b_i - (a'_i)^T x').$$

- Process ends when we have a single variable.

Proof of Farkas Lemma

- Reduce one more time until we have no variables. New system becomes

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leq \begin{bmatrix} b'_1 \\ b'_2 \\ \vdots \\ b'_n \end{bmatrix}$$

  Polyhedron $\{x : Ax \leq b\}$ infeasible $\iff$ if $b'_i < 0$ for some i.
- Rewriting and addition steps correspond to row operations on the original matrix A.

  $0 = MAx \geq Mb = b'$, **with matrix** M **with non-negative entries**
- Set $y^T = (e_i)^T M$, with e_i standard i-th unit vector then

$$0 = y^T A, \; y^T b < 0, \; \text{and } y \geq 0.$$

More on Farkas I

Here is another form of Farkas lemma:

- **Corollary:**
  $\{x : Ax = b, x \geq 0\} = \emptyset \iff \{y : y^T A \geq 0, y^T b < 0\} \neq \emptyset.$
- **proof** $\{x : Ax = b, x \geq 0\} \neq \emptyset \iff \{x : Ax \leq b, \ -Ax \leq -b, \ -Ix \leq 0\} \neq \emptyset$. By previous version of Farkas, this happens if and only if no solution exists of $y^T = [y_1 \ y_2 \ y_3]^T$ with

$$[y_1 \ y_2 \ y_3]^T \begin{bmatrix} A \\ -A \\ -I \end{bmatrix} = 0, \ [y_1 \ y_2 \ y_3]^T \begin{bmatrix} b \\ -b \\ 0 \end{bmatrix} < 0, \ y^T \geq 0$$

- The vector $y_1 - y_2$ has the desired property.

More on Farkas II

Here is another form of Farkas lemma:

- **Corollary:**
  $\{x : Ax \leq b, x \geq 0\} \neq \emptyset \iff$ When $y^T A \geq 0,$ then $y^T b \geq 0$
- **proof Necessity:** We know $x \geq 0$, $Ax = b$, if in addition $y^T A \geq 0$ then $y^T b = y^T Ax \geq 0$.
  Sufficiency: Suppose if $y^T A \geq 0,$ then $y^T b \geq 0$ but assume $\nexists x \geq 0$ with $Ax = b$. From the previous corollary, $\exists y$ with $y^T A \geq 0, y^T b < 0$. Therefore $0 \leq y^T b < 0$ which is a contradiction.
- There are many more consequences and variations of Farkas lemma (ALL theory of Linear Optimization based on it!!!).

Motivation: Magic Squares

A **magic square** is a square grid of non-negative real numbers such that the rows, columns, and diagonals all add up to the same value.

0	2	1
2	1	0
1	0	2

Magic Squares are closed under non-negative linear combinations

$$3\times\begin{array}{|c|c|c|}\hline 0&2&1\\\hline 2&1&0\\\hline 1&0&2\\\hline\end{array}=\begin{array}{|c|c|c|}\hline 0&6&3\\\hline 6&3&0\\\hline 3&0&6\\\hline\end{array}$$

$$\begin{array}{|c|c|c|}\hline 0&2&1\\\hline 2&1&0\\\hline 1&0&2\\\hline\end{array}+\begin{array}{|c|c|c|}\hline 1&2&0\\\hline 0&1&2\\\hline 2&0&1\\\hline\end{array}=\begin{array}{|c|c|c|}\hline 1&4&1\\\hline 2&2&2\\\hline 3&0&3\\\hline\end{array}$$

Question: Is there a finite set of $n \times n$ magic squares so that we can express every other possible magic square as a linear non-negative combination?

YES!

There are four such 3×3 magic squares:

0	2	1
2	1	0
1	0	2

2	0	1
0	1	2
1	2	0

1	2	0
0	1	2
2	0	1

1	0	2
2	1	0
0	2	1

IMPORTANT: There is an algorithm for computing a minimal such set of magic squares for $n \times n$ magic squares. These magic squares are the **extreme rays** of the cone of magic squares.

Polyhedral Cones

A set $\mathcal{C} \subseteq \mathbb{R}^n$ is a **cone** if it is closed under addition and multiplication by a positive constant.

- A set $\mathcal{C} \subseteq \mathbb{R}^n$ is a **inequality constrained** cone if $\mathcal{C} = \{x \in \mathbb{R}^n : Ax \geq 0\}$ for some matrix A.
- A set $\mathcal{C} \subseteq \mathbb{R}^n$ is a **finitely generated** cone if $\mathcal{C} = \{\lambda B : \lambda \in \mathbb{R}^k_+\}$ for some matrix B.

Theorem (Minkowski-Weyl)

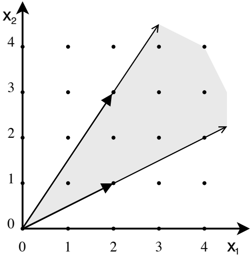
A cone $\mathcal{C} \subseteq \mathbb{R}^n$ is finitely constrained if and only if it is finitely generated.

- The set of *extreme rays* of the cone is the *minimal* set of generators of a cone.

FUNDAMENTAL QUESTION: how do we convert between the two repre

Example

Consider the following cone $\mathcal{C}$ and its two representations:



- $\mathcal{C} = \{x \in \mathbb{R}^2 : 3x_1 - 2x_2 \geq 0, -x_1 + 2x_2 \geq 0\}$.
- $\mathcal{C} = \{\lambda_1(2,1) + \lambda_2(2,3) : \lambda_1, \lambda_2 \in \mathbb{R}_+\}$.

Another great reason to solve this question

- **Theorem**: (Weyl-Minkowski's Theorem): For a convex subset P of $\mathbb{R}^d$ the following statements are equivalent:
    - P is an **H-polyhedron**, i.e., P is given by a system of linear inequalities $P = \{x : Ax \geq b\}$.
    - P is a **V-polyhedron**, i.e., For finitely many vectors $v_1, \ldots, v_n$ and $r_1, \ldots, r_s$ we can write

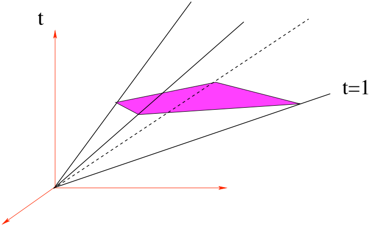
$$P = conv(v_1, v_2, \ldots, v_n) + cone(r_1, r_2, \ldots, r_s)$$

  Here $R + S$ denotes the Minkowski sum of two sets, $R + S = \{r + s : r \in R, s \in S\}$.

- We need to design an efficient algorithm for the conversion between the H-polyhedron and V-polyhedron!

Reduction to the case of Cones!!

- We can reduce this problem to problem of transforming between the two representations of a cone. From an H-polyhedron construct a cone from the polytope as follows:



- Observe: If the original polytope was given by inequalities $Ax \geq b$ then the cone is given by inequalities $\bar{A}y \geq 0$, where $\bar{A}$ is the extended matrix $[A, -b]$ and $y = (x, t)$.
- Enough to solve Weyl-Minkowski's Theorem for cones:

The Double Description Method (Motzkin-Raiffa-Thompson-Thrall 1953)

For a convex subset P of $\mathbb{R}^d$ is a cone if any of the following occurs:

- P is an **H-cone**, i.e., P is given by a system of linear inequalities $P = \{x : Ax \geq 0\}$.
- P is a **V-cone**, i.e., For finitely many vectors $r_1, \ldots, r_s$ we can write

$$P = cone(r_1, r_2, \ldots, r_s)$$

This is equivalent to (Matrix form!!):
For a convex subset P of $\mathbb{R}^d$ is a cone if any of the following occurs:

- P is an **H-cone**, i.e., $\exists$ matrix A such that $P = \{x : Ax \geq 0\}$.
- P is a **V-cone**, i.e., $\exists$ matrix R such that $P = \{x : x = Ry,\ y \geq 0\}$

We say the pair (A, R) is a **double description pair** (DD-pair).

Minkowski-Weyl Algorithmic version

- **Proposition:** For any pair of matrices A, R, (A, R) is a DD-pair of cone C if and only if (R^T, A^T) is a double description pair (of the polar cone of C).
  Proof: Use Farkas lemma.
- **A first algorithm proof of Minkowski-Weyl's theorem:** Let R be a matrix defining a V-cone, C, thus

$$C = \{x : x = Ry, \; y \geq 0\}.$$

  By Fourier-Motzkin we can eliminate all variables y from above system.
  The resulting system of inequalities is written as $Ax \geq 0$ (since Fourier-Motzkin respects the direction of inequalities).
  This proves that every V-cone can be written as an H-cone.
  By previous lemma we are done to prove the converse.
  WARNING: Not an efficient algorithm.

The Double description Method I

- Suppose A is an $m \times d$ matrix, defines cone $C = \{x : Ax \geq 0\}$.
- Let A_K denote the submatrix of A given by rows in index set K.
- Suppose we found already a matrix R which is DD pair with A_K. From a new row index $i \notin K$ construct new DD pair $(A_{K \cup \{i\}}, R')$ (**but HOW?**):
- Partition the column index set J of R into three parts:
  - $J^+ = \{j \in J : A_i r_j > 0\}$
  - $J^0 = \{j \in J : A_i r_j = 0\}$
  - $J^- = \{j \in J : A_i r_j < 0\}$

  We recover the new R' from the following lemma:

The Double description Method II

- **Lemma:** The pair $(A_{K\cup\{i\}}, R')$ is a DD pair, when the matrix R' is given by the $d \times J'$ matrix such that
  - the index set is $J' = J^+ \cup J^0 \cup (J^+ \times J^-)$, and
  - the new columns are $r_{jj'} = (A_i r_j) r_{j'} - (A_i r_{j'}) r_j$ for each $(j, j') \in J^+ \times J^-$.
- **Proof:** Let $C(A_{K\cup\{i\}}) = \{x : A_{K\cup\{i\}} x \geq 0\}$ and $C(R') = \{x : x = R'y,\ y \geq 0\}$. We wish $C(A_{K\cup\{i\}}) = C(R')$.
- Clearly $C(R') \subset C(A_{K\cup\{i\}})$ because $r_{jj'} \in C(A_{K\cup\{i\}})$.
- Take $x \in C(A_{K\cup\{i\}})$. Then

$$x = \sum_{j\in J} \lambda_j r_j, \quad \text{with } \lambda_j \geq 0$$

- If there is no $\lambda_k > 0$ for $k \in J^-$ then $x \in C(R')$ already. Thus assume such λ_k exists.

The Double description Method III

- Therefore since $A_i x \geq 0$ there must also be $\lambda_h > 0$ with $h \in J^+$.
- Substract a suitable multiple of $r_{kh} = (A_i r_h) r_k - (A_i r_k) r_h$ from $x = \sum_{j \in J} \lambda_j r_j$
- We are left with a new expression of x with smaller non-zero coefficients. This process can be repeated as long as $\lambda_k > 0$ with $k \in J^-$ exists.
- So in finitely many steps we must get read of all such λ at which point we have $x \in C(R')$.

The Double description Method IV

- We can refine the above construction, finding a matrix R' which has no redundant columns!!
- We say r_j is a **extreme ray** if it cannot be written as a non-negative combination of two other rays.
  Thus all we need to do is throw away columns of the matrix which are not extreme rays. How to tell???
- **Lemma:** Let $Z(x)$ be the set of indices of inequalities such that $A_i x = 0$. A ray r is an extreme ray of the cone $\{x : x \in \mathbb{R}^d,\ Ax \geq 0\} \iff$ the rank of the submatrix $A_{Z(r)} = d - 1$.
- How to do the initial DD pair?? Select a maximal submatrix A_K with linearly independent rows of A.
- Initial matrix R is the solution to $A_K R = I$. WHY? $rank(A) = d$ then A_K must be square then $R = A_K^{-1}$. Then (A_K, R) is DD pair since $A_K x \geq 0 \iff A_K^{-1} y, y \geq 0$.

The Double description Method V

- The double description method has a dual version called the **Beneath-Beyond method**.
- DD is practical for low dimensions (see CDD).
- The size of intermediate polytopes can be very very sensitive to the order in which the subspaces are introduced.
- D. Bremner (1999) showed a family of polytopes for which the double description method is exponential.

Thank you