EE464 Fourier-Motzkin Elimination

projection of polytopes

suppose we have a polytope

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

we'd like to construct the projection onto

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

call this projection $P(S)$

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

  (if it's the projection of an edge, then its the projection of the endpoints of the edge also)

- 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?

projection of polytopes

what are the facets of $P(S)$?

we'll see that the number of facets can increase enormously

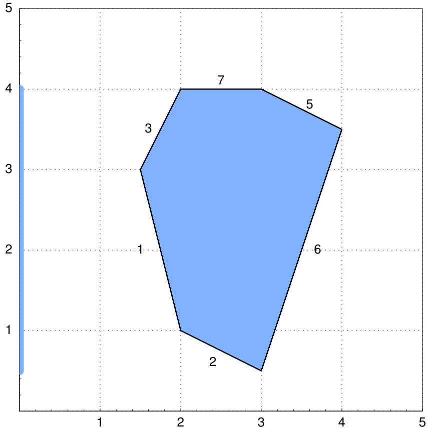
an upper bound; if $S \subset R^n$, and has m facets, then the projection onto R^{n-1} has less than

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

facets

example

$$-4x_1 - x_2 \leq -9 \qquad (1)$$
$$-x_1 - 2x_2 \leq -4 \qquad (2)$$
$$-2x_1 + x_2 \leq 0 \qquad (3)$$
$$-x_2 - 6x_2 \leq -6 \qquad (4)$$
$$x_1 + 2x_2 \leq 11 \qquad (5)$$
$$6x_1 + 2x_2 \leq 17 \qquad (6)$$
$$x_2 \leq 4 \qquad (7)$$



valid inequalities

we know 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

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

projection

we'd like to find the inequalities that define the projection $P(S)$

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

some other ways to say this

- we'd like to find *valid inequalities* that do not depend on x_1; i.e., the intersection

  $$\operatorname{cone}\{f_1, \dots, f_m\} \cap \mathbb{R}[x_2, \dots, x_n]$$

  which we might call the *elimination cone*

- we'd like to perform *quantifier elimination* to remove the *there exists* and find a *basic semialgebraic* representation of $P(S)$

Fourier-Motzkin elimination

- this procedure was invented by Fourier (1827) and rediscovered by Dines (1918) and Fourier (1936)
- similar to Gaussian elimination (1800)

we can generate inequalities of the form

$$(\lambda_1 a_1^T + \cdots + \lambda_m a_m^T)x \leq \lambda_1 b_1 + \cdots + \lambda_m b_m$$

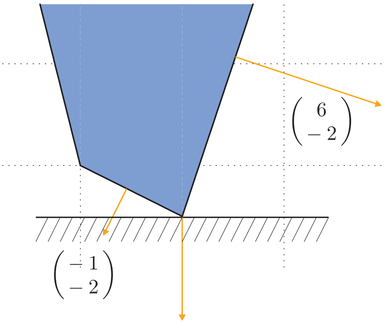
the idea is to combine pairs of inequalities that cancel x_1

since $\lambda_i \geq 0$, the members of each pair need opposite signed coefficients of x_1

example

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

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



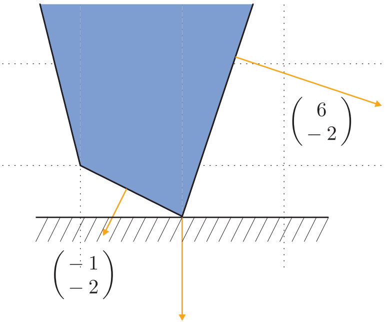
pick $\lambda_1 = 6$ and $\lambda_2 = 1$ to give

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

example

- any such positive linear combination of inequalities passes through corresponding vertex (in 2d)

  since that point satisfies both original inequalities with equality, it will also satisfy the new inequality with equality



- the corresponding vector is in the *cone* generated by a_1 and a_2

  so if a_1 and a_2 have opposite sign coefficients of x_1, then we can pick some element of the cone with coefficient zero.

Fourier-Motzkin theorem

the Fourier-Motzkin theorem says

- take all pairs of inequalities with opposite sign coefficients of x_1, and for each generate a new valid inequality that eliminates x_1
- also take all inequalities from the original set which do not depend on x_1 (i.e., (7) in this example)

this collection of inequalities defines exactly the projection of S onto $x_1 = 0$

matrix notation

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

in this case

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

matrix notation

the resulting inequality system is $CAx \le Cb$ holds, since

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

we find

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

the projection

this gives the system of inequalities for $P(S)$ as

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

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

- there are many redundant inequalities
- the tightest pair are

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

  and these define $P(S)$

extension

- since all the generated inequalities are valid, we know that they define a polytope that contains $P(S)$
- how do we know that this is actually $P(S)$?
- in other words, given x_2 satisfying the generated inequalities, when can we find an *extension* x_1 such that $(x_1, x_2) \in S$?

extension

view the original inequalities as

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

along with $x_2 \leq 4$

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

extension

if x_2 satisfies every inequality with one expression from the LHS and one from the RHS, then we must have

$$\max \begin{Bmatrix} \dfrac{x_2}{4} + \dfrac{9}{4} \\ -2x_2 + 4 \\ \dfrac{x_2}{2} \\ -6x_2 + 6 \end{Bmatrix} \leq \min \begin{Bmatrix} -2x_2 - 11 \\ -\dfrac{x_1}{3} + \dfrac{17}{3} \end{Bmatrix}$$

hence we can choose x_1 such that the original inequalities hold

- hence there is always an extension, and this proves the Fourier-Motzkin theorem

using Fourier-Motzkin

- by changing coordinates, and repeated application of FM, we can project a polytope onto any subspace of $\mathbb{R}^n$

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

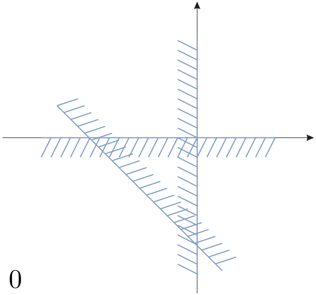
$$0 \leq -2$$

the we'd have derived a contradiction, and the original system of inequalities must 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

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

and subsequently eliminating x_2 gives

$$0 \leq -2$$

which is a contradiction

matrix notation

in matrix notation we have $A = \begin{bmatrix} -1 & 0 \\ 1 & 1 \\ 0 & -1 \end{bmatrix}$ and $b = \begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix}$

eliminating x_1 is multiplication

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

and similarly to eliminate x_2 we form

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

matrix notation

the final elimination is

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

so 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)

here λ is a *certificate* of infeasibility

Farkas Lemma

so Fourier-Motzkin gives a proof of Farkas lemma (Farkas 1894)

the primal problem is

$$\exists\, x \quad Ax \le b$$

the dual problem is a strong alternative

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

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

- does not require any compactness or topology
- 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: $\exists\, x \;\; Ax = b$
- dual: $\exists\, \lambda \;\; \lambda^T A = 0, \lambda^T b \neq 0$

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

linear programming

we can also use Farkas lemma to solve linear programs!

formulate the standard LP as the feasibility problem

$$\begin{aligned} c^T x &\le t \\ Ax &\le b \end{aligned}$$

and do a bisection search on t, testing feasibility via Fourier-Motzkin

of course, this is very inefficient compared to simplex or interior-point methods

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)

- solving the inequalities using interior-point methods is much faster than testing feasibility using FM
- allows floating-point arithmetic
- a similar speed advantage is obtained by directly solving the linear equations in p-satz and n-satz refutations