3. Improving linear relaxations of integer programs.

Lecturer: Massimo Lauria

http://www.csc.kth.se/~lauria/sos14/

We discuss methods to improve the linear relaxations of integer programs. They are systematic ways to find tighter and tighter polytopes containing the convex hull of the integer solutions. The main idea is to lift*, that is, to augment the linear program to a higher dimensional space and then to* project *it down to the original space to obtain a polytope that is smaller than the one we started with. We shall see two different ways to do lift and project: namely the Lovász-Schrijver and Sherali-Adams hierarchies.*

Most of the material in this lecture was covered in the first lecture note on integer and linear programming.[1] Consequently, these lecture notes should be seen as a complement, and we will frequently point to the material in those notes.

[1] See the lecture notes for the first lecture.

We consider feasibility problems of the form

$$\begin{aligned} \text{find} \quad & x \\ \text{subject to} \quad & Ax \le b \\ & x \in \{0,1\}^n. \end{aligned} \tag{1}$$

The boolean constraint $x_i \in \{0,1\}$ is equivalent to $x_i^2 - x_i = 0$, thus problem (1) can be reformulated as

$$\begin{aligned} \text{find} \quad & x \\ \text{subject to} \quad & Ax \le b \\ & x_i^2 - x_i = 0 \qquad \text{for } i = 1, \ldots, n. \end{aligned} \tag{2}$$

As you can see it is a quadratic program. Since we know how to solve linear program we can study a simple linear relaxation of this problem, which is

$$\begin{aligned} \text{find} \quad & x \\ \text{subject to} \quad & Ax \le b \\ & 0 \le x \le 1, \end{aligned} \tag{3}$$

where the boolean constraints has been relaxed to a linear inequalities, enlarging the set of feasible solutions. Problem (3) is a pretty standard linear program, so there exist several efficient algorithms to find feasible solutions for it. If there is none, then problems (1) and (2) are unsolvable as well. However, a feasible solution for problem (3) may not be boolean.[2]

[2] if b is integral and A is a so called "total unimodular" matrix, then all feasible solutions are actually in integer hull.

The convex hull of a set C is denoted $\mathbf{conv}(C)$. We denote the feasible set of problem (3) as $\mathcal{P}$ and the convex hull of the integer solutions as $\mathcal{P}_I$, namely

$$\mathcal{P}_I = \mathbf{conv}(\mathcal{P} \cap \{0,1\}^n). \tag{4}$$

Gomory[3] introduces a way to remove spurious fractional solutions: if $a^T x \leq b$ is a valid inequality for $\mathcal{P}$ and a is an integer vector, then $a^T x \leq \lfloor b \rfloor$ is also valid over integer solutions (this new inequality is called an "integer cut"). We denote by $\mathcal{P}'$ the intersection of $\mathcal{P}$ and all the integer cuts of the above form. $\mathcal{P}'$ is also known as the *elementary closure* of $\mathcal{P}$. Iterating this procedure and we get a sequence of smaller and smaller polytopes

$$[0,1]^n \supseteq \mathcal{P}' \supseteq \mathcal{P}'' \supseteq \cdots \supseteq \mathcal{P}_I. \tag{5}$$

Eisenbrand and Shulz[4] proved that after $O(n^2 \log n)$ iteration the process produces the polytope $\mathcal{P}_I$. However, this procedure is not very useful for algorithmic optimization, because even optimizing on $\mathcal{P}'$ is NP-hard [5].

Lovász-Schrijver and Sherali-Adams hierarchies describe two sequences of relaxations on problem (1), each relaxation tighter than the previous one in the sequence. So each hierarchy induces a sequence of polytopes

$$\mathcal{P} = \mathcal{P}_0 \supseteq \mathcal{P}_1 \supseteq \cdots \supseteq \mathcal{P}_t \supseteq \cdots \supseteq \mathcal{P}_I. \tag{6}$$

The advantage with respect to the Gomory technique is that that the optimum over the polytopes in the sequence is easy to compute. Naturally, if $\mathcal{P}_t = \emptyset$ for any t, then $\mathcal{P}_I = \emptyset$. That is, no feasible integer solution exists.

[3] Ralpha E. Gomory. Outline of an algorithm for integer solutions to linear programs. *Bulletin of the American Mathematical Society*, 64(5):275–278, 1958

[4] Friederich Eisenbrand and Andreas S. Schulz. Bounds on the Chvátal rank of polytopes in the 0/1-cube. *Integer Programming and Combinatorial Optimization*, pages 137–150, 1999

[5] F. Eisenbrand. *A note on the membership problem for the first elementary closure of a polyhedron*. Max-Planck-Institut für Informatik, 1998

Lovász-Schrijver hierarchy

In the Lovász-Schrijver hierarchy[6] the linear program is augmented to a higher dimensional space of approximately n^2 dimensions. We introduce quadratic inequalities in the problem and, via a variable substitution, turn them into linear inequalities in the higher dimensional space. These inequalities are then projected back to the original space, yielding a tighter feasible set. This procedure is repeated several times to obtain tighter and tighter polytopes as in (6). The formal definition of Lovász-Schrijver is in the first lecture note.

[6] L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0–1 optimization. *SIAM Journal on Optimization*, 1:166, 1991

Example 1. We want to improve the following linear relaxation:

$$\begin{aligned} \text{find} \quad & x \\ \text{subject to} \quad & x \geq 1/2 \\ & 0 \leq x \leq 1, \end{aligned} \tag{7}$$

which has infinitely many fractional solutions and just one boolean solution, $x = 1$. Since $x \geq 1/2$ and $x \leq 1$, the following quadratic inequality holds.

$$(1 - x)(x - 1/2) = 3x/2 - x^2 - 1/2 \geq 0. \tag{8}$$

We add a new variable y to represent the quadratic term x^2, and the inequality (8) becomes linear with respect to x and y.

$$3x/2 - y - 1/2 \geq 0. \tag{9}$$

The definition of y also turns $x^2 - x = 0$ into $y = x$. We then project $\mathbb{R}^2$ back into $\mathbb{R}$, using that $y = x$, we get the program

$$\begin{aligned} \text{find} \quad & x \\ \text{subject to} \quad & x \geq 1 \\ & 0 \leq x \leq 1, \end{aligned} \tag{10}$$

which has only the boolean solution $x = 1$. Notice that the inequality $x \geq 1$ did not hold in the original linear program, and that it has been deduced by a the linearization of quadratic constraints.

Proof system interpretation

The Lovász-Schrijver hierarchy can be seen as a *proof system* for determining whether a polytope contains an integer solution. We consider the polytope $\mathcal{P}$ induced by the problem

$$\begin{aligned} \text{find} \quad & x \\ \text{subject to} \quad & Ax \leq b \\ & 0 \leq x_i \leq 1 \quad \text{for } i = 1, \ldots, n \\ & x_i^2 - x_i = 0 \quad \text{for } i = 1, \ldots, n. \end{aligned} \tag{11}$$

The inequalities are the *axioms* of the proof system. Inference rules allow to introduce quadratic inequalities via multiplication and quadratic axioms. Then it is possible to combine inequalities to cancel the quadratic terms and to get back to linear inequalities. To derive a specific inequality we may have go to back and forth several time between degree one and two. The *rank of the inequality* is, informally, the number of times we have to do that. [7] The original problem (11) is infeasible if we can derive the inequality $1 \leq 0$.

[7] A formal definition of the system is in the lecture notes of the first lecture.

Geometric interpretation

We introduce the matrix variable $Y \in \mathbb{R}^{(n+1)\times(n+1)}$ and we enforce the following constraints on the entries of Y, based on the inequalities (11).[8]

[8] Y has row and columns indexed from 0 to n. Y_{ij} is the element in the ith row and jth column. $Y^{(j)}$ is the jth columns of Y.

$$\begin{aligned} & Y_{ji} = Y_{ij} && \text{for every } i, j = 0, \ldots, n, \\ & Y_{0i} = Y_{i0} = Y_{ii} && \text{for } i = 0, \ldots, n \\ & AY^{(j)} - bY_{0j} \leq 0 && \text{for } j = 0, \ldots, n. \\ & A(Y^{(0)} - Y^{(j)}) - b(Y_{00} - Y_{0j}) \leq 0 && \text{for } j = 0, \ldots, n. \end{aligned} \tag{12}$$

This new set of inequalities represents a polytope in $\mathbb{R}^{(n+1)\times(n+1)}$. Now we consider its projection on variables $Y_{00}, Y_{01}, \ldots Y_{0n}$, and its restriction $Y_{00} = 1$. We get a new polytope in the original domain by interpreting each Y_{0j} as x_j, for $j = 1, \ldots, n$. This lift-and-project process defines the Lovász-Schrijver operator $N(\cdot)$, so the new polytope is called $N(\mathcal{P})$.

Observe that the positive combinations of the constraints (12) are, up to syntactic translation, all the quadratic inequality derivable in rank 1 for the initial linear inequalities that describe $\mathcal{P}$. The projection on $Y_{00}, Y_{01}, \ldots Y_{0n}$ picks the valid quadratic inequalities that mention only $Y_{00}, Y_{01}, \ldots Y_{0n}$, and so they correspond to the linear inequalities of rank 1. The syntactic translation amounts to the introduction the auxiliary variable x_0, to the variable substitution $x_i x_j = Y_{ij}$ for every $i, j = 0, \ldots, n$, and to the assignment $x_0 = 1$.

In the end we have that $N(\mathcal{P}) = \mathcal{P}_1$, and more generally that $N(\mathcal{P}_t) = \mathcal{P}_{t+1}$. Thus $N^t(\mathcal{P}) = \mathcal{P}_t$.

But how do we optimize on $\mathcal{P}_t$ using only the constraints that describe of $\mathcal{P}$? The *ellipsoid algorithm* solves linear programs and does not need an explicit description of the polytope, it just need a separation oracle, as mentioned in the proof sketch of Theorem 10 in the first lecture. A *separator oracle* for a convex set (in this case $N^t(\mathcal{P})$) is an efficient algorithm that, given a point (in this case Y), either verifies that the solution is in the polytope or finds an hyperplane that separates the set from the point. More concretely it finds an inequality that is valid for the set but that is violated by the point.

The separator oracle for $N^t(\mathcal{P})$ does the following tests:

1. $Y = Y^T$,
2. $Y_{0i} = Y_{i0} = Y_{ii}$, for $i = 0, \ldots, n$,
3. $Y^{(i)} \in N^{t-1}(\mathcal{P})$ for $i = 0, \ldots, n$,
4. $Y^{(0)} - Y^{(i)} \in N^{t-1}(\mathcal{P})$ for $i = 0, \ldots, n$?

Assuming to have a separator oracle for $N^{t-1}(\mathcal{P})$ we can find a violated inequality for Y easily, using these checks.

A separator oracle for $N^0(\mathcal{P})$ just checks if one of the initial inequalities is falsified. The one for $N^t(\mathcal{P})$ requires up to $O(n)$ calls to the separator of $N^{t-1}(\mathcal{P})$, so in total a separator oracle for $N^t(\mathcal{P})$ has running time $n^{\mathcal{O}(t)}$.

Sherali-Adams hierarchy

The Sherali-Adams hierarchy is another family of a lift and project relaxations. However, in the Sherali-Adams we only do a single step of lifting and then a projection. See the description in the lecture notes for the first lecture.

Proof system interpretation

The Sherali-Adams hierarchy can be interpreted as a proof system for determining whether some polynomial inequalities

$$p_1 \geq 0, p_2 \geq 0, \ldots, p_m \geq 0 \tag{13}$$

hold simultaneously. A Sherali-Adams proof of $p \geq 0$ is and equation of the form

$$\sum_{l=1}^{m} p_l g_l + \sum_i (x_i^2 - x_i) h_i + g_0 = p \tag{14}$$

where each

$$g_l = \alpha_l \prod_{i \in A_l} x_i \prod_{j \in B_l} (1 - x_j) \tag{15}$$

with $\alpha_l \geq 0$ and $A_l \cap B_l = \emptyset$, and each h_i is an arbitrary polynomial. If we go back to the original problem (1), and to its relaxation (3), then polynomials

p_l for $l = 1, \dots, m$ correspond to the inequalities $Ax \le b$. Problem (1) are infeasible if we can get $p = -1$ in proof (14).

The *rank of a Sherali-Adams* proof is defined in different ways in literature.[9] Here, the rank of a Sherali-Adams proof is its degree, thus

$$N^{t-1}(\mathcal{P}) \supseteq S_t(\mathcal{P}), \tag{16}$$

where $S_t(\mathcal{P})$ is the polytope generated by the linear inequalities derivable with a Sherali-Adams proof of degree t.

[9] See the notes of the first lecture.

Geometric interpretation

We can see $S_t(\mathcal{P})$ as the projection of a polytope in a larges space, as we did for Lovász-Schrijver.

We add the auxiliary variable x_0 and introduce the variable $Y \in \mathbb{R}^s$ where $s = \binom{n}{\le t}$. We perform a syntactic substitution such that every monomial of degree at most t in the original variables in transformed into a linear inequality on the Y variables, for example,

$$x_2^2 x_3 x_5 = Y_{\{2,3,5\}}.$$

Consider the inequalities of degree t derivable as in (14), linearized by the substitution. We denote as $\mathcal{E}$ the set of points in $\mathbb{R}^s$ that satisfies them. If we project $\mathcal{E}$ on $(Y_\emptyset, Y_{\{1\}}, \dots, Y_{\{n\}})$ and then we fix $Y_\emptyset = 1$ we get exactly the polytope $S_t(\mathcal{P})$.

For more details on Sherali-Adams relaxations, please see page 9 of notes of the first lecture.

References

[Eis98] F. Eisenbrand. *A note on the membership problem for the first elementary closure of a polyhedron*. Max-Planck-Institut für Informatik, 1998.

[ES99] Friederich Eisenbrand and Andreas S. Schulz. Bounds on the Chvátal rank of polytopes in the 0/1-cube. *Integer Programming and Combinatorial Optimization*, pages 137–150, 1999.

[Gom58] Ralpha E. Gomory. Outline of an algorithm for integer solutions to linear programs. *Bulletin of the American Mathematical Society*, 64(5):275–278, 1958.

[LS91] L. Lovász and A. Schrijver. Cones of matrices and set-functions and 0–1 optimization. *SIAM Journal on Optimization*, 1:166, 1991.