3. Improving linear relaxations of integer programs.

Lecturer: Massimo Lauria

Disclaimer: this lecture notes has not yet been reviewed by the mail lecturer. It is released as is for the convenience of the students.
We discuss two methods of improving linear relaxations of integer programs, namely the Lovász-Schrijver and Sherali-Adams hierarchies. They allow for a systematic way of finding 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 where a smaller polytope than the one we started with is obtained. This procedure of lift and project are done in different ways for the different hierarchies, as we shall see.*

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

This lecture contained much of the material given in the notes on integer and linear programming [1]. Consequently, these lecture notes should be seen as a complement to those.

[1] See notes from first lecture on course webpage.

We consider feasibility problems of the form

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

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

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

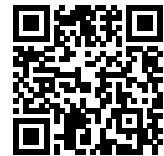A simple linear relaxation of this problem is

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

where the boolean constraint has been relaxed into a linear inequality constraint. Problem (3) is convex and there exists efficient algorithm for solving it. If problem (3) lacks a feasible solution, we can state that problems (1) and (2) lack one as well. However, if a feasible solution is found it is rare that it is the boolean solution.

We denote the feasible set of problem (3) as $\mathcal{P}$ and the convex hull[2] of the integer solution as $\mathcal{P}_I$, we then get

[2] The convex hull of a set C is denoted **conv**(C).

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

Gomory[3] introduces a way to cut fractional solutions by noticing that 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. We denote by $\mathcal{P}_1$ the intersection of $\mathcal{P}$ and the additional constraints of the above form, also known as the elementary closure of $\mathcal{P}$. Iterating this procedure and we get a sequence of smaller and smaller polytopes $\mathcal{P} \supseteq \mathcal{P}_1 \supseteq \cdots \supseteq \mathcal{P}_k \cdots$. It was proved that for some $k = O(n^2 \log n)$, $\mathcal{P}_k = \mathcal{P}_I$. However, this procedure is not very useful for optimization because even optimizing in $\mathcal{P}_1$ is already NP-hard.

The objective of Lovász-Schrijver and Sherali-Adams hierarchies is to improve the initial problem relaxation (3). The relaxation is improved in the sense that the feasible set gets smaller and therefore closer to the convex hull of the integer solution.

In the Lovász-Schrijver and Sherali-Adams hierarchies each relaxation system induces

$$\mathcal{P} = \mathcal{P}_0 \supseteq \mathcal{P}_1 \supseteq \cdots \supseteq \mathcal{P}_t \supseteq \cdots \supseteq \mathcal{P}_I. \tag{5}$$

Naturally, if $\mathcal{P}_t = \varnothing$ for any t, then $\mathcal{P}_I = \varnothing$. That is, no feasible integer solution exists.

Lovász-Schrijver hierarchy

In the Lovász-Schrijver hierarchy[4] the linear program is augmented to a higher dimensional space of approximately n^2 dimensions. Quadratic inequalities are introduced which, via a variable substitution, become linear inequalities in the higher dimensional space. These inequalities are then projected back onto the original variables, yielding a tighter feasible set. This procedure can be repeated in an iterative manner to obtain tighter and tighter sets, as depicted in (5).

[4] 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:

$$\begin{aligned} \text{find} \quad & x \\ \text{subject to} \quad & x \geq 1/2 \\ & 0 \leq x \leq 1, \end{aligned} \tag{6}$$

which has infinitely many fractional solutions and one boolean solution, $x = 1$. We introduce the quadratic inequality

$$(1-x)(x-1/2) = 3x/2 - x^2 - 1/2 \geq 0. \tag{7}$$

By substituting x^2 with y, inequality (7) becomes linear with respect to x and y. We then project y onto x via the relation $x_i^2 - x_i = 0$, for $i = 0, \ldots, n$, that is $y = x$. The new feasibility problem is

$$\begin{aligned} \text{find} \quad & x \\ \text{subject to} \quad & x \geq 1 \\ & 0 \leq x \leq 1, \end{aligned} \tag{8}$$

which only has the boolean solution $x = 1$.

         Scribe: Mariette Annergren

Proof system interpretation

The Lovász-Schrijver hierarchy can be seen as a *proof system* for determining if a given set contains an integer solution. We consider the case when the given set is $\mathcal{P}$. Consequently, the constraints

$$Ax \leq b, \ 0 \leq x \leq 1, \text{ and } x_i^2 - x_i = 0, \text{ for } i = 0, \ldots, n, \qquad (9)$$

are the *axioms* of the proof system. The system allows for quadratic inequalities to be introduced via multiplication and positive combinations of the axioms. How many of these modifications that are done to introduce a particular linear inequality is related to the *rank of the inequality*[5]. The original problems (1) and (2) are infeasible if we get the linear inequality $1 \leq 0$.

[5] For a definition and related results, see Definition 2 and Theorem 3 in notes from the first lecture on course webpage.

Geometric interpretation

We introduce the matrix variable $Y \in \mathbb{R}^{(n+1) \times (n+1)}$ and, given $\mathcal{P}$, we enforce the following constraints on Y[6]:

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

$$
\begin{aligned}
Y_{ji} &= Y_{ij}, \quad \text{for every } i, j = 0, \ldots, n, \\
Y_{0i} &= Y_{i0} = Y_{ii}, \quad \text{for } i = 0, \ldots, n, \\
\sum_{j=1}^{n} A^{(j)} Y_{ij} &- b Y_{i0} \leq 0, \quad \text{for } i = 0, \ldots, n.
\end{aligned}
\qquad (10)
$$

If Y satisfies the constraints (10), it satisfies every derivable linear inequality of rank 1. We can see that this is true by introducing the auxiliary variable $x_0 = 1$ and performing the variable substitution $x_i x_j = Y_{ij}$ for every $i, j = 0, \ldots, n$. We let any terms of degree zero be represented by Y_{00}.

Once we have a new set of inequalities expressed in Y, we project these back onto the original variables x. [7]

[7] The projection of Y onto x is defined as the Lovász-Schrijver operator. It is denoted $N(\cdot)$. We then have $N(\mathcal{P}_t) = \mathcal{P}_{t+1}$ and $N^t(\mathcal{P}) = \mathcal{P}_t$.

To optimize over a linear program we can, for example, use the *ellipsoid algorithm*. To use this algorithm it is sufficient to have a separation oracle for the considered convex set, as mentioned in the proof sketch of Theorem 8[8]. A separator in the Lovász-Schrijver hierarchy to determine $N^t(\mathcal{P})$ is based on the following checks: Is

[8] See notes from the first lecture on course webpage.

1. $Y = Y^T$,

2. $Y_{0i} = Y_{i0} = Y_{ii}, \quad \text{for } i = 0, \ldots, n,$

3. $Y^{(i)} \in N^{t-1}(\mathcal{P}) \quad \text{for } i = 0, \ldots, n,$

4. $Y^{(0)} - Y^{(i)} \in N^{t-1}(\mathcal{P}) \quad \text{for } i = 0, \ldots, n$?

If we can find a violated inequality, we can find a separating hyperplane and thus construct a tighter polytope $\mathcal{P}_t$. This is done with only one call to the separator oracle, meaning that the operations $N^t(\mathcal{P})$ requires $n^{\mathcal{O}(t)}$ calls to the oracle.

Sherali-Adams hierarchy

The Sherali-Adams hierarchy can, as the Lovász-Schrijver hierarchy, be seen as a lift and project relaxation. However, in the Sherali-Adams hierarchy we do not have to start with linear inequalities. Consequently, we may end up with higher and higher orders of the inequalities for each iteration. It is not until the last iteration that we project the problem back onto linear inequalities.

Proof system interpretation

The Sherali-Adams hierarchy can be interpreted as a proof system for determining if the set of polynomial inequalities

$$p_1 \geq 0, p_2 \geq 0, \ldots, p_m \geq 0. \tag{11}$$

holds. The Sherali-Adams proof of $p \geq 0$ is

$$\sum_{l=1}^{m} p_l g_l + \sum_i (x_i^2 - x_i) h_i + g_0 = p \tag{12}$$

where each

$$g_l = \alpha_l \prod_{i \in A_l} x_i \prod_{j \in B_l} (1 - x_j) \tag{13}$$

with $\alpha_l \geq 0$ and $A_l \cap B_l = \emptyset$, and each h_i is an arbitrary polynomial. For the problem statement we consider here, problem (3), the polynomials p_l for $l = 1, \ldots, m$ correspond to the inequality constraint $Ax \leq b$. The original problems (1) and (2) are infeasible if we get $p = -1$.

The *rank of a Sherali-Adams* proof is defined in different ways in literature[9]. Based on the definitions of the rank, we can state the following:

$$N^t(\mathcal{P}) \supseteq S_{t+1}(\mathcal{P}), \tag{14}$$

where $S_t(\mathcal{P})$ is the polytope we get through a proof of rank at most t.

[9] See notes from the first lecture on course webpage.

Geometric interpretation

Similar to the Lovász-Schrijver hierarchy, we add the auxiliary variable $x_0 = 1$ and introduce the variable $Y \in \mathbb{R}^s$ where $s = \binom{n}{\leq t}$. We perform a variable substitution such that every monomial in p is mapped into an element of Y, for example,

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

We denote the set of points satisfying all the new linear inequalities in $\mathbb{R}^s$ as $\mathcal{E}$. We can then state that $\mathcal{E}$ satisfies every derivable linear inequality of rank at most t, after projecting over, with some abuse of notation, $(Y_{\emptyset}, Y_1, \ldots, Y_n)$.

For an example of a Sherali-Adams relaxation, please see page 9 of notes from the first lecture on course webpage.

Scribe: Mariette Annergren

References

[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.