Lift and Project Techniques

Jacob Denson

University of Alberta

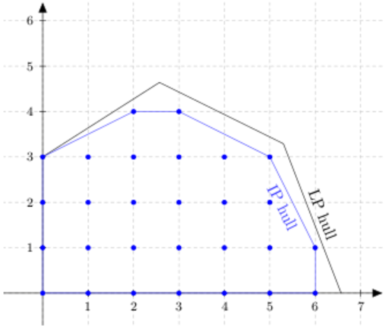
December 9, 2016

What are Lift and Project Techniques Used?

- Integer Linear Programs.
- NP hard, but described in an environment where an easy approximation algorithm is almost immediate.
- Lift and Project techniques provide a method for tightening the approximation algorithms.

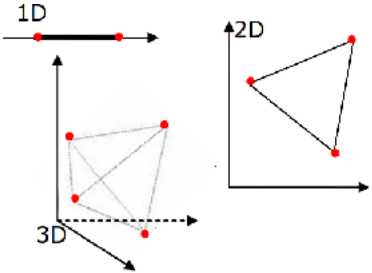
Tightening the Relaxation Space

- Best relaxation on convex hull of integer coordinates.
- Not normally feasible in polynomial time.



Lifting and Projecting

- Lift Solution Polyhedra to higher dimensional space, in such a way that the polyhedron is more accurate to integer solutions.
- Solve problem in higher dimnsional space.
- Project back into original space, to get more accurate solution.



Notation

- We will be performing zero-one linear optimization on a set
$$K = \{x \in \{0,1\}^n : Ax \leq b\}$$
- The first relaxation is
$$K_0 = \{x \in \mathbf{R}^n : Ax \leq b\}$$
- w.l.o.g, assume that $0 \leq x_i \leq 1$ for all $x \in K_0$.

Quadratic Programming

- Can solve zero-one integer programs using quadratic programming, by implicitly assuming $x_i(1 - x_i) = 0$, or more explicitly, $x_i^2 = x_i$.
- Quadratic programming is not P time, but relaxation are...
- Replace products of variables with new variables, approximately behaving like the products due to linear constraints.

Sherali-Adams Polynomials

- For disjoint $J, K \subset [n]$, let

$$P_{J,K}(x) = \left(\prod_{j\in J} x_j\right)\left(\prod_{k\in K}(1-x_k)\right)$$

  of order m if $|J \cup K| = m$.
- Exponentially many polynomials, but only $O(n^m)$ of order m.
- For $x \in K$, $P_{J,K}(x) \geq 0$ for all J, K.

The Sherali-Adams Lifting Technique

1. Multiply all constraints by all $P_{J,K}$, for $P_{J,K}$ order m, multiplying the number of constraints by $O(n^m)$.
2. Introduce new constraints $P_{J,K}(x) \geq 0$.
3. Perform reduction $x_i^2 = x_i$ to reduce redundant inequalities.
4. Approximate polynomial program by linear program by introducing new variables w_J, for $J \subset [n]$, which are swapped out for $\left(\prod_{j \in J} x_j\right)$ to make the constraints linear.
5. Let $w_{\{i\}} = x_i$, so we can project back solutions.
6. Note that we could consider the J and K as ordered tuples rather than sets, in which case we should have $y_{J,K} = y_{K,J}$, which hints at a semidefinite programming solution. This is what modern lift and project methods do.

Example

- The independent set problem on a graph $G = (V, E)$ can be described by the 0-1 integer linear program

$$\max \sum x_i$$

$$x_i + x_j \leq 1 \ (ij \in E)$$

- Then m'th lift has the substituted constraints

$$\sum_{T' \subset T} (-1)^{|T'|} [w_{S \cup T' \cup \{j\}} + w_{S \cup T' \cup \{i\}} - w_{S \cup T'}] \leq 0$$

$$0 \leq \sum_{T' \subset T} (-1)^{|T'|} w_{S \cup T' \cup \{i\}} \leq \sum_{T' \subset T} (-1)^{|T'|} w_{S \cup T'}$$

where $ij \in E$, and $|S| + |T| \leq m$.

Polynomials Cover Polynomials of Lower Order

Lemma

The Constraints for Polynomials of order $m+1$ imply the constraints for smaller orders.

Proof.

If $p \in [n] - (J \cup K)$, and $P_{J,K}$ has order m, then

$$P_{J+p,K}(x) + P(J, k+P) = x_p P_{J,K}(x) + (1-x_p) P_{J,P} = P_{J,P}$$

If $P_{J+p,K} \geq 0$ and $P_{J,p+K} \geq 0$, then $P_{J,P} \geq 0$. $\square$

- Thus if a program has N constraints originally, the m'th Sherali-Adams lift gives us $O(Nn^m)$ constraints, and shows the algorithm runs in polynomial time.
- This theorem also implies the solution sets $K_0 \supset K_1 \supset \dots$ are decreasing, where K_i is the solution set obtained from K by relaxing, and adding constraints over polynomials of order i.

Integral Optimality

Lemma

If $x \in \{0,1\}^n$, then $(x,w) \in Z_m = \{(x,w) : P_{J,K}(x,w) \geq 0\}$ holds if and only if $w_J = \prod_{j \in J} x_j$.

Proof.

One direction is easy. The other is proved by induction, and a case by case analysis, which is easy since we are working in $\{0,1\}^n$. By the previous lemma, if $(x,w) \in Z_m$, then $(x,w) \in Z_{m-1}$, so $w_J = \prod_{j \in J} x_j$ for all $|J| < m$. Now suppose $|J| = m$. Then if $k \neq l$ are elements of J, then

$$P_{J-\{k\},\{k\}}(x,w) = \left(\prod_{j \neq k} x_j\right) - w_J \geq 0$$

$$P_{J-\{k,l\},\{k,l\}}(x,w) = (1 - x_k - x_l)\left(\prod_{j \neq k,l} x_j\right) + w_J \geq 0$$

Integral Optimality (Continued)

Proof.

Thus

$$(x_k + x_l - 1) \prod_{j \neq k,l} x_j \leq w_J \leq \prod_{j \neq k} x_j = x_l \prod_{j \neq k,l} x_j$$

If $\prod_{j \neq k,l} x_j = 0$, then we obtain the inequality $0 \leq w_J \leq 0$, so

$$w_J = 0 = \prod_{j \in J} x_j$$

Otherwise, we prove by a case by case analysis on x_k and x_l. $\square$

Remark

- Implies consistency of the lifting, so that projected integral results make sense back in the origin solution.
- Thus, if we let K_m be the solution set above, where we agree to abuse notation and think of it as a set of (x, w) and as a set of x, then $K_0 \supset K_1 \supset \cdots \supset K_n \supset K$.
- K_0 is just the linear relaxation.

How good is the approximation?

- Still only convex, so it can't be 'that' good.
- Turns out that K_n is the best convex set possible, but optimizing over K_n is $O(n^n)$.
- More difficult to prove, because we are talking about the geometry of a set in $\mathbf{R}^n$ rather than $\{0, 1\}^n$.

A Different Coordinate System

Lemma

For a fixed (x, w), the linear endomorphism T generated by

$$w_J \mapsto P_{J,J^c}(x, w) := y_J$$

has inverse

$$w_J = \sum_{K \subset J^c} y_{J \cup K} = \sum_{J \subset K} y_K$$

Proof.

Just algebra... First note that

$$P_{J,J^c}(x, w) = \sum_{K \subset J^c} (-1)^{|K|} w_{J \cup K}$$

Then just manipulate the terms of the sum above. $\square$

Extreme points of Z_m

Lemma

The extreme points of

$$Z = \{(x, w) : (\forall J : P_{J,J^c}(x, w) \geq 0)\}$$

are zero-one integral in projection.

Proof.

Since $\sum_J P_{J,J^c}(x, w) = 1$,

$$P_{\emptyset,[n]}(x, w) = 1 - \sum_{\emptyset \subsetneq J \subset [n]} P_{J,J^c}(x, w)$$

Thus

$$Z = \{(x, w) : (\forall J \neq \emptyset : P_{J,J^c}(x, w) \geq 0), \sum_{\emptyset \subsetneq J \subset [n]} P_{J,J^c}(x, w) \leq 1\}$$

Extreme points of Z_m (Continued)

Proof.

Now apply the invertible map T to Z, which maps Z into

$$S = \{y : \sum_{\emptyset \subsetneq J \subset [n]} y_J \leq 1, (\forall J \neq 0 : y_J \geq 0)\}$$

In particular, T preserves extreme points. The extreme points of S are $y_K = 1$ for some K, and $y_J = 0$ otherwise, or $y_J = 0$ for all J. Thus the extreme points of Z satisfy the same for their $P_{J,J^c}(x, w)$, but we have already seen how this implies that $x \in \{0, 1\}^n$. $\square$

Lemma

The map $J \mapsto w_J$ is monotone decreasing.

Proof.

If $J \subset K$, then since $J^c = K^c \cup (K - J)$,

$$w_K = \sum_{L \subset K^c} y_{L \cup K} \geq \sum_{L \cup J^c} y_{L \cup (K-J) \cup J} = w_J$$

so the lemma follows from the inversion map for T. $\square$

The Convex Hull K_n

Theorem

K_n is the convex hull of K.

Proof.

Suppose that $(x, w) \in K_n$ is an extreme point. Expand the constraints of A, so that

$$\sum a_i^j x_i \leq b_j$$

Then points $(x, w) \in K_n$ are defined by the constraints

$$\left(\sum a_i^j - b_j\right) P_{J,J^c}(x, w) \geq 0$$

$\square$

The Convex Hull K_n (Continued)

Proof.

Let $E = \{J : (\exists k : \sum_{j \in J} a_j^k - b_k < 0)\}$. If $J \in E$, then $P_{J,J^c}(x, w) = 0$, which if $J \in E^c$ we find $P_{J,J^c}(x, w) \geq 0$. Hence

$$K_n = \{(x, w) : (\forall J \in E : P_{J,J^c}(x, w) = 0), (\forall J \in E^c : P_{J,J^c}(x, w) \geq 0)\}$$

so that K_n is a subface of the set Z considered before. But we already verified that Z has binary extreme points, and this completes the proof. $\square$

Randomized Solutions on the Convex Hull

- The convex hull of K is a good place to apply randomized algorithms!
- If $x \in K_n$ is optimal, and $x = \sum a_i x_i$ with $x_i \in K$, then the random algorithm which chooses x_i with probability a_i achieves optimal value in expectation.
- But since we are working over the convex hull of K, the optimal points actually occur on the points in K!
- Still exponential time...
- The convex hull of K is a good place to study heirarchical lift and project techniques, like the more advanced Leserre semidefinite programming lift methods.

Extension

- There are smart ways of reducing the number of constraints we have – we don't need to use all $P_{J,K}(x, w)$.
- These methods can be extended in a fairly obvious way to any polynomial program to approximate it by linear programming, the details are easy.

Context of Work

- Based on Chvatal's work in 1973 on facetial inequalities.
- An explicit construction of what Chvatal found must exist.
- Work on K_1 had already been done to do work on the approximation of vertex cover, but Sherali-Adams generalized the techniques.

Applications

- Vertex Cover (Padberg).
- Zero-One Knapsack Problem - How do you maximize the value of your bag, where you can either take of leave an item (Balas, Zemel).
- Travelling Salesman (Crowder and Padberg).

Limitations

- The integrality gap is the ratio between the solution of a relaxed linear program, and the solution of an integer program. Tells us 'how good' the approximation is.
- The integrality gap for the lift has limitations.
- Karlin, Matheiu, Nguyen showed that lifting techniques for the knapsack problem can only achieve a gap of $2 - \varepsilon$ in polynomial time, whereas vertex cover has a $1 + \varepsilon$ approximation algorithm. Other lifting techniques work better.