CS 2429 - Foundations of Communication Complexity

Lecturer: Toniann Pitassi

1 Applications of Communication Complexity: Extended Formulations of Linear Programs

Linear programming is a very powerful tool for attacking hard combinatorial optimization problems. Methods such as the ellipsoid algorithm have shown that linear programming is solvable in polynomial time. Linear programming also plays a central role in the design of approximation algorithms. In fact, it is known that linear programming is P-complete, and this implies that if $NP = P$ then for every problem in NP, given an instance, it is possible (in polytime) to solve it via a polynomial-sized LP.

A large class of linear programs were identified by Yannakakis, and referred to as *extended formulations*. We emphasize that extended formulations of LPs to not capture all LPs for solving a given NP-hard problems, but nonetheless, they capture a large and useful family of LPs.

We will define extended formulations, and then prove that lower bounds on extended formulations follow from communication complexity lower bounds.

2 Definitions and Background

For a matrix A, let A_j denote the j^{th} column and A^j the j^{th} row. For any combinatorial optimization problem, we can encode its set of possible solutions as a set of points $X \subseteq \{0,1\}^n$ such that optimizing an instance of the problem becomes the problem of optimizing a linear objective function, $f(X)$ over the convex hull of these points. This convex hull, $P = conv(X)$ defines a polytope in n dimensional space, whose *vertices* are the feasible solutions to the problem and whose *facets* correspond to the constraints of the problem.

A (convex) polyhedron is a set $P \subseteq R^d$ that is the intersection of a finite collection of closed halfspaces. (That is, P is the set of solutions of a finite system of linear inequalities.) A face of P is a subset $F = \{x \in R^d \mid w^T x = \delta\}$ of P such that P satisfies the inequality $w^T x \leq \delta$. Note that face F is again a polyhedron. A *vertex* is a face of dimension 0. A *facet* is a face of dimension one less than P. An inequality $w^T x \leq \delta$ is called *facet-defining* if the face F it defines is a facet. A polytope is a bounded polyhedron, or equivalently, P is a polytope if and only if P is the convex hull $conv(V)$ of a finite set V of points.

This convex hull, $P = conv(X)$ defines a polytope in n dimensional space, whose *vertices* are the feasible solutions to the problem and whose *facets* (or sides) correspond to the constraints of the problem.

A *feasible solution* for an optimization problem with constraints $Ax \leq b$ is any point $x \in R^n$ which satisfies all of the constraints of the problem. The *size* of a linear program is the number of constraints it contains.

2.1 The permutahedron polytope

Examples. Consider the permutahedron consisting of the convex hull of all permutations of $[1, .., n]$. It is not too hard to see that it is defined by exponentially many constraints. (For all $S \subseteq [n]$ we have $\sum_{i \in S} x_i \geq 1+2+\ldots+|S| = |S|(|S|+1)/2$. However, it is possible to introduce new variables, and to write a new set of constraints over the original variables plus the new variables that is polynomial size, and such that the projection of this higher dimensional polytope down to the original variables is the permutahedron! That is, it is possible to rewrite the constraints (using more variables) in order to obtain a polynomial size *extended formulation* for the permutahedron LP.

To do this, Goemans showed that we can take any switching network that sorts $x_1, \ldots, x_n$. A switching network has n inputs, $x_1, \ldots, x_n$. It is visualized as a set of n horizontal wires. It also has a set of ordered comparator gates that connect pairs of wires. A comparator gate from wire i to wire j is drawn as a vertical arrow originating at wire i and pointing to wire j. There are two outputs to a comparator gate, the OR output (the start of the arrow) computes the OR of the two inputs, and the AND output (the end of the arrow) computes the AND of the two inputs. The effect of this gate is to sort i and j. Goeman's showed that if we have a comparator network for $x_1, \ldots, x_n$ with m gates, then we can introduce $2m$ new variables, $x_{n+1}, \ldots, x_{n+2m}$, corresponding to all possible subfunctions computed by the network, and a set of inequalities over the variables $x_1, \ldots, x_{n+2m}$ such that the projection of this polytope back to the original variables is the permutahedron. The inequalities express the fact that the output gates are $1, 2, .., n$, and that each intermediate gate is computing correctly. If the circuit has k comparators, we have

(1) for all $i \in [n]$, $x_{2k+i} = i$

(2) for any comparator gate with inputs x_i, x_j and outputs $x_k = min(x_i, x_j)$, and $x_l = max(x_i, x_j)$ we have $x_i + x_j = x_k + x_l$, $x_k \leq x_i$, $x_k \leq x_j$.

It is not hard to see (although we will not prove it here) that $Proj_n(Q) = P$ and the size of Q is polynomial.

Intuitively, if it is a legal comparator circuit, then when we feed it some permutation of $[1, .., n]$, it will output the values in sorted order, and conversely, if the circuit outputs the values in sorted order and all of the intermediate variables are defined correctly, then the input must be a permutation of $[1, .., n]$.

(DRAW PICTURE)

2.2 Extended Formulations

Definition An extended formulation (EF) of a polytope $P \subset R^n$ defined by $Ax \leq b$ is a linear system:

$$Ex + Fy = g, y \geq 0$$

in the variables $(x, y) \in R^{n+r}$ where E, F are real matrices with n, r columns respectively and such that $x \in P$ (satisfies $Ax \leq b$) if and only if there exists y such that $Ex + Fy = g, y \geq 0$ holds.

The *size* of an extended formulation is the number of y variables, or equivalently the number of inequalities. The extension complexity of a polytope P, $xc(P)$, is the minimum number of inequalities in any extended formulation for P.

By using extended formulations, we can sometimes reduce the number of facets exponentially. When this can be done, we can run a standard LP algorithm to get a polynomial time algorithm. This is one of the most common approaches attempting to show that P equals NP. For example, the TSP polytope is the set of vectors corresponding to tours of K_n. A polysize extended formulation for the TSP polytope implies $P = NP$! We note that EFs do not characterize *all* linear programs for a given problem. The restriction is that the polytope must be independent of the instance, so the instance only shows up in the objective function. Thus lower bounds for EFs rules out a large family of polytime LP algorithms but does not rule out all of them. (Since linear programming is P-complete, ruling out all LPs would essentially be showing that $P \neq NP$.)

2.3 Brief History

(1) Yannakakis in 1990 showed that any symmetric EF for TSP or matching has exponential size. He also established the connection between EFs and communication complexity. He also established the equivalence between EF and nonnegative rank.

(2) Fiorini-et-al, 2012 showed that any EF for clique and TSP has exponential size. This is a reduction to set disjointness.

(3) Braun-et-al, 2012 extended the EF setup for approximation algorithms, and showed that any EF for approximating clique within $n^{1/2-\epsilon}$ has exponential size.

(4) Braverman, Moitra 2013, and Braun-Pakutta, 2013 showed that any EF approximating clique within $n^{1-\epsilon}$ has exponential size. These papers use information complexity techniques; the latter paper uses the notion of common information.

(5) Chan-et-al, 2013 showed that any EF that approximates Maxcut within $(2-\epsilon)$ has quasipolynomial size. This paper is a reduction to Sherali-Adams lower bounds via Fourier analysis. This is the only paper that I am aware of that looks at instances that are "natural". They also extend the EF setup to the SDP setting and prove lower bounds for symmetric EFs of SDPs.

(6) Rothvoss, 2013 showed that any EF for perfect matching has exponential size. This is particularly interesting since we do have LPs for solving perfect matching. This is a corruption lower bound.

All of the above results with the exception of (5) use communication complexity.

3 Lower Bounds on Extension Complexity

3.1 Yannakakis' Factorization Theorem

Yannakakis shows that the notion of EF which is a geometric parameter is equivalent to nonnegative rank of the associated slack matrix, which is algebraic parameter.

Definition Consider a polytope defined by the convex hull of a set of vertices, $P = conv(V)$, and dually defined by a set of facets $Q = \{x \in R^d \mid Ax \leq b\}$. We refer to P as the "inner" description of the polytope and Q as the outer description. Note that there can be more than one outer description of P.

The slack matrix, $S_{P,Q}$ for P, Q is a matrix with v rows and f facets, where v is the number of vertices of P, and f is the number of facets (inequalities) in Q. The entry $[i, j]$ is equal to $b_i - A_i x_j$, or the distance from the ith facet to the jth vertex.

Definition The nonnegative rank, $rank_+(M)$ of a matrix M is the smallest $r \in N$ such that M can be expressed as $M = FV$ where F, V are nonnegative matrices with intermediate dimension r.

In Yannakakis' seminal paper, he gave a factorization theorem which is the backbone for most of our extended formulation lower bounds. It shows that the nonnegative rank of the slack matrix $S_{P,Q}$ is equivalent to the extension complexity of P.

Theorem 1 *Consider a polytope defined by the convex hull of a set of points, $P = conv(V)$, and also defined by a set of facets $Q = \{Ax \leq b\}$. Let $S_{P,Q}$ denote the slack matrix of the polytope with respect to P, Q. Then the nonnegative rank of $S_{P,Q}$ is equal to $xc(P)$.*

We note that this is actually an equivalence, but here we will just prove the direction that we need in order to prove lower bounds on extension complexity.

Let $Q = \{Ax \leq b\}$, and let $Ex + Fy = g, y \geq b$ be an extended formulation of Q. Consider the following "extended" slack matrix, M. The first set of rows will correspond to each equation $e^i x + f^i y = g_i$; the second set of rows will correspond to each equation $y_i \geq b_i$. And the last set of rows will be one for each inequality from Q, $A^i x \leq b_i$. For each $u \in conv(V)$, consider a corresponding pair (u, w) such that $Eu + Fw = g$, $w \geq b$. (Such a pair exists since $Ex + Fy = g, y \geq b$ is an extended formulation of $Ax \leq b$. The columns of M are labelled with all such pairs (u, w) where $u \in conv(V)$. The entry (i, j) of M where i corresponds to an inequality (or equality) and j corresponds to a pair (u, w) will give the slack – the distance between the point (u, w) in the extended polytope to the inequality. Note that the entire first set of columns contains all 0's – since these are equalities. Since $Ax \leq b$ follows from $Ex + Fy = g, y \geq b$ (in the sense that any assignment satisfying the latter inequalities also satisfies the inequality $Ax \leq b$), by Farkas' Lemma, $Ax \leq b$ can be written as a nonnegative linear combination of $Ex + Fy = g, y \geq b$. Thus the third group of rows of M can be written as a nonnegative linear combination of the top two rows of M. Since the very top row is all 0's, the third group of rows of M can be written as a nonnegative linear combination of the rows $y_i \geq b_i$. Thus the nonnegative rank of M is no greater than the number of rows in the second group, or in other words, the number of y variables.

Since the slack matrix $S_{P,Q}$ is a subset of the matrix M, it follows that the nonnegative rank of $S_{P,Q}$ is also at most the number of y variables.

Notice that the same argument works for *any* set of inequalities Q' that are implied by Q. That is, let $P = conv(V)$ be a polytope, and let Q' be a set of linear inequalities such that every point in P satisfies all linear inequalities of Q'. (So the polytope defined by Q' contains the polytope defined by $P = Q$.) Then a lower bound on the nonnegative rank of $S_{P,Q'}$ implies the same lower bound on the extension complexity of P.

3.2 The Clique Polytope

The clique polytope (for graphs with n vertices) is the convex hull of the following set of 2^n vectors: all vectors $v \in \{0, 1\}^{n^2}$ that describe a clique on some subset $S \subseteq [n]$ (and have no other edges).

Given a graph $G = (V, E)$ over $V \subseteq [n]$, $w(G)$ is defined as follows.

(1) $e_{i,i} = 1$ if $i \in V$ and 0 otherwise

(2) $e_{i,j} = 0$ if $i, j \in V$ and $(i, j) \in E$;

(3) $e_{i,j} = -1$ if $i, j \in V$ and $i, j) \notin E$;

(4) otherwise $e_{i,j} = 0$.

It is not too hard to see that for each graph G, $< w(G), x > \leq cliquenum(G')$ and the vector x achieving equality corresponds to a maximal clique in G.

For a graph G, the linear program (with exponentially many constraints) for clique has variables $x_{i,j}$, $i, j \in [n]$. We want to maximize $< w(G), x >$ subject to the following constraints: For each graph G', $< w(G'), x > \leq cliquenum(G')$. Note that the x achieving the max value will be a maximal clique in G.

Let Q^{all} be the polytope defined by the above inequalities. Clearly $P \subseteq Q^{all}$. We want to show that there is no polysize EF for Q^{all}. Given Q^{all} defined by $Ax \leq b$ as described above, an EF is $Ex + Fy = b'$, $y \geq 0$ such that the projection to x is Q^{all}.

By the factorization theorem, it suffices to prove lower bounds on the nonnegative rank of $SM(P, Q^{all})$.

4 Nonnegative Rank and Communication Complexity

Let M be our slack matrix, and suppose that it has rank r. Then it can be written as the sum of r rank one, nonnegative matrices, $M_1 + \ldots + M_r$. The support of each M_i is a combinatorial rectangle. Thus the nonnegative rank of M is at least the number of rectangles needed to cover the support of S. Thus, it suffices to show that M has high nondeterministic communication complexity.

Now consider a simple subset of all graphs, $SIMPLE = \{G_a \mid a \subset [n]\}$ where $G_a = (V_a, E_a)$, $V_a = a$ and $E_a = \emptyset$. That is, the vertices of G_a are the vertices in a, and G_a has no edges.

Let Q' be the set of inequalities $< w(G_a, x > \leq cliquenum(G_a)$ for all $G_a \in SIMPLE$. Note that the clique number of G_a is 1 for all $G_a \in SIMPLE$.

We will show that UDISJ is embedded in $SM(P, Q')$ and therefore embedded in $SM(P, Q^{all})$.

Let $Clique_b$ be a graph consisting of a clique on all vertices in b and no other edges.

Consider $G_a \in SIMPLE$ and $Clique(b)$. That is, G_a has no edges and the vertex set is a, and $Clique(b)$ is a clique over the vertices in b. If $a \cap b$ is empty, then $< w(G_a), Clique(b) >= 0$ and if $a \cap b$ has size 1, then $< w(G_a), Clique(b) >= 1$.

Thus, we see that UDISJ is embedded in $SM(P, Q')$ and therefore the nonnegative For any pair (a, b) with the promise that either a, b are disjoint or they intersect in exactly one element,

4.1 Extended Formulations for Approximate Linear Programs

We can generalize the above setup for extended formulations of linear programs that produce approximate solutions.

A linear encoding of a combinatorial optimization problem is a pair $(\mathcal{L}, \mathcal{O})$. $\mathcal{L} \subseteq \{0, 1\}^*$ is a set of feasible solutions, and $\mathcal{O} \subseteq R^*$ is a set of admissible objective functions. An instance is a pair (d, w) where d is a nonnegative integer and $w \in \mathcal{O} \cap R^d$. To solve the instance we want to find $x \in \mathcal{L} \cap \{0,1\}^d$ such that $w^T x$ is maximized. (Or in the case of a minimization problem, is minimzed.)

$(\mathcal{L}, \mathcal{O})$ defines a pair of nested convex sets $P \subseteq Q$, where $P = conv(x \mid x \in \mathcal{L} \cap \{0,1\}^d)$ and $Q = \{x \in R^d \mid \forall w \in \mathcal{O} \cap R^d,\ w^T x \leq max\{w^T x \mid x \in P\}\}$.

A ρ-approximate EF is $Ex + Fy = g$, $y \geq 0$ such that

$$max\{w^T x \mid Ex + Fy = g,\ y \geq 0\} \geq max\{w^T x \mid x \in P\}\ \forall w \in R^d$$

and

$$max\{w^T x \mid Ex + Fy = g,\ y \geq 0\} \leq \rho max\{w^T x \mid x \in P\}\ \forall w \in \mathcal{O} \cap R^d$$

Letting $K = \{x \in R^d \mid \exists y \in R^r s.t. Ex + Fy = g,\ y \geq 0\}$, this is equivalent to saying that there is an extended formulation whose projection is K, where $P \subseteq K \subseteq \rho Q$.

With this more general setup, it is not hard to see that we can again define a slack matrix $S_{P,\rho Q}$, corresponding to P, Q, and the same argument as before essentially shows that the nonnegative rank of the slack matrix is equal to the ρ-approximate extension complexity of clique.

Going back to our clique problem, let us consider the same instances that we did before. Now whenever a is disjoint from b, the entry corresponding to (a, b) will have value ρ if $a \cap b$ has size 1, and will have value $\rho - 1$ if $a \cap b$ has size 0. Thus we cannot appeal to nondeterministic communication complexity of the support of the slack matrix since these entries are never 0. Nonetheless, the nonnegative rank (for this specific submatrix) can be lower bounded by a discrepancy argument, and also by an information complexity argument. This leads to a stronger result, showing that clique cannot be ρ approximated by a subexponential-size extended formulation, for ρ linear in n.

we see that a ρ-approximate Extende

4.2 An Equivalent Communication Complexity Problem

Above we saw lower bounds on the nonnegative rank of $S_{P,Q}$ (and hence exponential lower bounds on the extension complexity) by lower bounding the nondeterministic communication complexity of $S_{P,Q}$.

Here we will give a communication complexity model that is actually *equivalent* to nonnegative rank.

Recall that the problem is: Alice is given a vertex v and Bob is given a face $A^i x \leq b_i$ and then want to determine the distance of v from the inequality. The minimum complexity of a protocol for solving this problem in expectation equals $\log(rank_+(S_{P,Q})) + O(1)$.