MATH 233B: Polyhedral techniques in combinatorial optimization
Instructor: Jan Vondrák
Lecture date: March 2, 2017 **Scribe: Yuval Wigderson**

Recall that for a polytope
$$P = \{x \in \mathbb{R}^n : Ax \leq b\}$$
an extended formulation is
$$Q = \{(x, y) \in \mathbb{R}^{n+r} : Cx + Dy \leq d\}$$
such that
$$P = \{x \in \mathbb{R}^n : \exists y, (x, y) \in Q\}$$
and the complexity of Q is the number of inequalities in Q. We want to prove that some polytopes can't have extended formulation that are too simple.

Theorem 1 (Yannakakis '91) *For*
$$P = conv(\{x_1, \ldots, x_v\}) = \{x : Ax \leq b\}$$
where A has m inequalities, let $S \in \mathbb{R}^{m \times v}$ be the "slack matrix" of P, defined by
$$S_{ij} = b_i - A_i x_j$$
In other words, S_{ij} is the amount of slack that the ith constraint has at the jth vertex. Then the minimum complexity of an extended formulation of P is exactly the nonnegative rank of S, namely
$$\begin{aligned}\text{rank}_+(S) :&= \min\{r : S = UV, U \in \mathbb{R}_+^{m \times r}, V \in \mathbb{R}_+^{r \times v}\}\\ &= \min\{r : \exists U_1, \ldots, U_m, V_1, \ldots, V_v \in \mathbb{R}_+^r; S_{ij} = U_i \cdot V_j\}\end{aligned}$$

Example: Let P be a unit square in the plane. Then up to numbering the vertices and the constraints, we have
$$S = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$
Then $\text{rank}(S) = 3$. However, $\text{rank}_+(S) = 4$.

Proof of Theorem 1. Suppose we have the slack matrix S and $\text{rank}_+(S) = r$, and we write $S_{ij} = U_i V_j$ for $U_i, V_j \in \mathbb{R}_+^r$. Let U be the matrix whose rows are U_i. We define the following extended formulation of $P = \{x : Ax \leq b\}$:
$$Q = \{(x, y) \in \mathbb{R}^{n+r} : Ax + Uy = b, y \geq 0\}$$
This description has lots of equalities, but only r inequalities, so has complexity r (recall that we only count inequalities because every equality can be used to just eliminate a variable, and thus shouldn't be counted in the complexity).

To check that this is really an extended formulation, suppose that x_j is the jth vertex of P. Then we define $y_j = V_j$, and we observe that $(x_j, y_j) \in Q$ since $y_j \geq 0$ and

$$A_i x_j + U_i y_j = A_i x_j + U_i \cdot V_j = A_i x_j + S_{ij} = b_i$$

Additionally, suppose that $x \notin P$. Then there is a constraint that is violated, say $A_i x > b_i$. Therefore, for any $y \geq 0$, we have that

$$A_i x + U_i y \geq A_i x > b_i$$

and thus $(x, y) \notin Q$. Thus, we indeed have that

$$P = \{x : \exists y, (x, y) \in Q\}$$

and thus this is an extended formulation, and it has complexity $\leq r$.

For the converse, suppose that there is some extended formulation $Q = \{(x, y) : Bx + Cy \leq d\}$, and say that we have r inequalities. For each vertex x_j of P, fix y_j such that $(x_j, y_j) \in Q$. Define a vector

$$V_j = d - Bx_j - Cy_j \in \mathbb{R}^r_+$$

Each inequality $A_i x \leq b_i$ is implied by the system of inequalities $Bx + Cy \leq d$. By the Farkas lemma (or LP duality), there is a nonnegative linear combination of the inequalities in $Bx+Cy \leq d$ that produces $A_i x \leq b_i$. In other words, there is a vector U_i so that

$$U_i^T \begin{pmatrix} B & C & d \end{pmatrix} = \begin{pmatrix} A_i & 0 & b \end{pmatrix}$$

such that $U_i \geq 0$, and where this equation is an equation of block matrices. We now have that

$$U_i \cdot V_j = U_i \cdot (d - Bx_j - Cy_j) = b_i - A_i x_j$$

Thus, we conclude that $\text{rank}_+(S) \leq r$. □

So now the problem is to lower bound the nonnegative rank. Returning to the square example from before, how do we prove that $\text{rank}_+(S) \geq 4$. The nonnegative rank problem is equivalent to writing

$$S = \sum_{i=1}^{r} u_i v_j^T$$

where $u_i, v_j \geq 0$. We claim that $\text{supp}(u_i) \times \text{supp}(v_j) \subseteq \text{supp}(S)$, namely that this block can contain no zeros. The reason is that if there is a zero in such a block, then we can never make it nonzero in this sum, as we're adding more nonnegative things. So in the case of our matrix S, we observe that any block can cover at most 2 nonzero entries, since every 2×2 or 1×3 block contains both zeros and ones. Since we have 8 nonzero entries, we need at least 4 nonnegative rank 1 matrices, and thus $\text{rank}_+(S) = 4$.

We will prove a lower bound for the extension complexity of the correlation polytope.

Definition 2 *The correlation polytope in dimension n is*

$$\mathbb{R}^{n^2} \supseteq P^{(n)}_{corr} = conv(\{xx^T : x \in \{0,1\}^n\})$$

Theorem 3 *The extension complexity of $P_{corr}^{(n)}$ is at least 1.5^n.*

Lemma 4 *For any $A \subseteq [n]$ there is a linear constraint valid for $P_{corr}^{(n)}$ such that it is tight for $\chi_B \chi_B^T$ if and only if $|A \cap B| = 1$.*

Proof: Consider the constraint

$$\left(\sum_{i \in A} x_i - 1\right)^2 \geq 0$$

This is certainly true for all of $P_{corr}^{(n)}$. It looks quadratic, but using the structure of the polytope, we can turn it into a linear constraint on the coordinate of $\mathbb{R}^{n^2}$. Expanding, this becomes

$$\sum_{i,j \in A} x_i x_j - 2 \sum_{i \in A} x_i + 1 \geq 0$$

Assume $x \in \{0,1\}^n$. Then if $y_{ij} = x_i x_j$, then y is a vertex of P_{corr}. Additionally, for $i = j$, we have that $y_{ii} = x_i^2 = x_i$. Therefore, the above inequality can be turned into

$$\sum_{i \in A} y_{ii} + \sum_{\substack{i,j \in A \\ i \neq j}} y_{ij} - 2 \sum_{i \in A} y_{ii} + 1 \geq 0$$

and we derive the final inequality

$$\sum_{\substack{i,j \in A \\ i \neq j}} y_{ij} - \sum_{i \in A} y_{ii} + 1 \geq 0$$

Then this is a linear inequality, and as we argued, it's valid for any vertex of $P_{corr}^{(n)}$, and thus for all of $P_{corr}^{(n)}$. Moreover, if we define $x = \chi_B \chi_B^T$, we see that the original inequality is tight if and only if $\sum_{i \in A} x_i = 1$, which happens if and only if $|A \cap B| = 1$. □

So let S be the submatrix of the slack matrix corresponding to only these constraints, so that it's a $2^n \times 2^n$ matrix. Its entries are

$$S_{A,B} = (|A \cap B| - 1)^2$$

for $A, B \subseteq [n]$.

Definition 5 *A rectangle $\mathcal{A} \times \mathcal{B}$ for $\mathcal{A}, \mathcal{B} \subseteq 2^{[n]}$ is* valid *if for all $A \in \mathcal{A}, B \in \mathcal{B}, |A \cap B| \neq 1$.*

Let

$$\mathcal{D}(n) = \{(A, B) : A, B \subseteq [n], A \cap B = \emptyset\}$$

Then we have $|\mathcal{D}(n)| = 3$.

Lemma 6 *Every valid rectangle covers at most 2^n entries in $\mathcal{D}(n)$.*

Note that this suffices, since it implies that covering all of the support of S will require at least $|\mathcal{D}(n)|/2^n = 1.5^n$ valid rectangles.

Fix a valid rectangle $\mathcal{R} = \mathcal{A} \times \mathcal{B}$. Suppose first that $A \cap B = \emptyset$ for all $A \in \mathcal{A}, B \in \mathcal{B}$. Then we can partition the elements of $[n]$ into those that might participate in some set of $\mathcal{A}$ and those that might participate in some set of $\mathcal{B}$. Then if these parts have size k and $n-k$, we get that

$$|\mathcal{R}| = |\mathcal{A}||\mathcal{B}| \leq 2^k 2^{n-k} = 2^k$$

In particular, $\mathcal{R}$ covers at most 2^n elements of $\mathcal{D}(n)$.

Now we deal with the general case. We will do this by induction on n. The inductive claim, as above, is that if $\mathcal{R} = \mathcal{A} \times \mathcal{B}$ is valid, then

$$|\mathcal{R} \cap \mathcal{D}(n)| \leq 2^n$$

We define

$$\mathcal{A}_1 = \{A \in \mathcal{A} : n \in A\} \cup \{A \in \mathcal{A} : n \notin A, A + n \notin \mathcal{A}\}$$

and analogously we define $\mathcal{B}_1$. We also define

$$\mathcal{A}_2 = \{A \in \mathcal{A} : n \notin A\}$$

and analogously $\mathcal{B}_2$. Finally, define

$$\mathcal{R}_1 = \mathcal{A}_1 \times \mathcal{B}_2 \qquad\qquad \mathcal{R}_2 = \mathcal{A}_2 \times \mathcal{B}_1$$

Then both $\mathcal{R}_1$ and $\mathcal{R}_2$ are valid rectangles, since they are subsets of another valid rectangle. Next, we claim that

$$\mathcal{R} \cap \mathcal{D}(n) \subseteq (\mathcal{R}_1 \cap \mathcal{D}(n)) \cup (\mathcal{R}_2 \cap \mathcal{D}(n))$$

Finally, we claim that the projection $\pi : 2^{[n]} \to 2^{[n-1]}$ given by dropping element n is injective on $\mathcal{A}_1, \mathcal{B}_1$. This is precisely by the definition of $\mathcal{A}_1$ and $\mathcal{B}_1$. Putting this all together, we get that

$$\begin{aligned}
|\mathcal{R} \cap \mathcal{D}(n)| &= |\mathcal{R}_1 \cap \mathcal{D}(n)| + |\mathcal{R}_2 \cap \mathcal{D}(n)| \\
&= |\pi(\mathcal{R}_1) \cap \mathcal{D}(n-1)| + |\pi(\mathcal{R}_2) \cap \mathcal{D}(n-1)| \\
&\leq 2^{n-1} + 2^{n-1} = 2^n
\end{aligned}$$

So the only thing that remains to be done is to prove the middle claim, that $\mathcal{R} \cap \mathcal{D}(n) \subseteq (\mathcal{R}_1 \cap \mathcal{D}(n)) \cup (\mathcal{R}_2 \cap \mathcal{D}(n))$.