Extended Formulations

Introduction[1]

Consider a polytope $P = \{x \in \mathbb{R}^n \mid Ax \le b\}$, where A is a $m \times n$ matrix for $m \ge n$. The size of a polytope will be the number inequalities in $Ax \le b$[2]. For our main theorem, we assume that none of the inequalities in $Ax \le b$ are implicit equalities. We also assume that there are no redundant constraints[3]. Thus, every constraint in $Ax \le b$ corresponds to a facet [Theorem 8.1. [Sch98]].

Although ellipsoid method solves LP in polynomial time, it is impractical and alternatives such as Simplex or interior point method require knowing all the constraints. Hence, we are interested in expressing a polytope using polynomially many inequalities (in the number of variables). It is possible that the size is exponential even after removing the redundant constraints. We could then try to extend a polytope to a higher dimension, i.e, increase the number of variables, and hope that the size is polynomial in the new number of variables. See an example below.

Example 1 (Spanning Tree Polytope). Let P_{span} be the convex hull of the characteristic vectors of the spanning trees of the complete graph K_n. The characteristic vector is a vector $x \in \mathbb{R}^{\binom{n}{2}}$ such that each entry $x_{i,j}$ corresponds to an edge $\{i, j\}$. Then, P_{span} as described by Edmonds [Edm71] is :

- $\sum_{\{i,j\}\in\binom{V}{2}} x_{i,j} = n-1,$
- $\sum_{\{i,j\}\in\binom{A}{2}} x_{i,j} \le |A|-1 \qquad$ for all $A \subseteq V$,
- $0 \le x_{i,j} \le 1 \qquad$ for all $\{i, j\} \in \binom{V}{2}$.

The number of variables and the size of this description is $\mathcal{O}(n^2)$ and $\mathcal{O}(2^n)$ respectively[4]. By introducing $\mathcal{O}(n^3)$ new variables, Martin [Mar91] showed that P_{span} has an extended formulation of size $\mathcal{O}(n^3)$.

Example 2 (Parity polytope). Let $PP = CH\{x \in \{0,1\}^n \mid \sum_{i=1}^n x_i = k \text{ for some odd } k\}$. A description of this polytope was given by Jeroslow [Jer75]:

[1]This presentation is based on [Yan91]. There are of course notes on this topic by many others: here, here, here and here.

[2]The alternative is to define the size as number of equalities plus the inequalities. But Fiorini et al. [FMP+15] comment that it makes little difference.

[3]By Khachiyan's theorem, implicit equalities and redundant constraints can be found in polynomial time. See the remark below Theorem 13.4 [Sch98].

[4]Verify that none of the constraints are redundant.

- $\sum_{i\in A} x_i - \sum_{i\notin A} x_i \le |A| - 1$ for all even subsets A of $\{1,\dots,n\}$,
- $0 \le x_i \le 1$ for all i.

The number of variables is n but the size of this description is exponential. Next, we can increase the number of variables and reduce the size to a polynomial. For this, write $PP = CH\{\cup_{k \text{ odd}} S_k\}$, where $S_k = \left\{x \in \{0,1\}^n \mid \sum_{i=1}^n x_i = k\right\}$. We need two observations: (i) $PP = CH\{\cup_{k \text{ odd}} CH\{S_k\}\}$. (ii) If $y \in CH\{S_k\}$[5], then $\sum_{i=1}^n y_i = k$.

By Observation (i), $x \in PP$ iff there exists α_k's and $y_k \in CH\{S_k\}$ such that $x = \sum_{k \text{ odd}} \alpha_k y_k$, where

- $\sum_{k \text{ odd}} \alpha_k = 1$.

Let $y_{i,k}$ denote entries of y_k, and define $z_k = \alpha_k y_k$. Let $z_{i,k}$ denote entries of z_k. Then,

- $x_i = \sum_{k \text{ odd}} z_{i,k}$ for all $i = 1,\dots,n$.

By Observation (ii), $\sum_{i=1}^n y_{i,k} = k$ for each y_k, which can also be written as

- $\sum_{i=1}^n z_{i,k} = \alpha_k k$ for all odd k.

Since each $y_k \in [0,1]^n$, i.e, each $y_{i,k} \in [0,1]$, we have

- $0 \le z_{i,k} \le \alpha_k$ for all $i = 1,\dots,n$ and odd k.

These four constraints describe PP. The number of variables and the number of inequalities are both $\mathcal{O}(n^2)$.

Extended formulation

To formalize this idea of extending a polytope to a higher dimension, we define the following.

Definition 1. *Let $P = \{x \in \mathbb{R}^n \mid Ax \le b\}$, and $Q = \{(x,y) \in \mathbb{R}^{n+s} \mid Bx + Cy \le d\}$. Then, Q is an extended formulation of P if $P = \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^s \text{ s.t. } (x,y) \in Q\}$. The extension complexity $xc(P)$ of P is the minimum size of an extended formulation of P* [6].

One way to prove P = NP is to express any NP-complete problem by a LP, and show that it has an extended formulation of polynomial size. Yannakakis [Yan91] shows that matching and TSP polytope cannot be expressed by a symmetric LP of subexponential size. This demonstrates that LP approach cannot determine P = NP. Later Fiorini et al. [FMP+15] show that TSP polytope cannot be expressed by a LP of superpolynomial size, and Rothvoss [Rot17] shows that matching polytope has exponential extension complexity which also implies exponential lower bound for TSP polytope. Note that exponential lower bound does not mean that the linear optimization problem over that LP

[5]Note $y \in [0,1]^n$.

[6]There is also a different definition. We say that a polytope Q expresses P if there exists a linear map π such that $\pi(Q) = P$. It can be shown that these two notions are equivalent, i.e, $xc(P)$ is the minimum of the number of facets over all polytopes that express P [Yan91, FMP+15].

cannot be solved in polytime. E.g. checking whether a graph has a perfect matching can be done polynomial time, but the decision version of TSP is NP-complete.

Slack matrix

A polytope P in $\mathbb{R}^n$ with m facets and v vertices can be described in two ways. Let $\{x \in \mathbb{R}^n \mid \langle A_i, x\rangle \le b_i\}$ for $i = 1,\dots,m$ be the halfspaces corresponding to the facets. If A is a matrix whose rows are the vectors $A_1,\dots,A_m$, then P can be expressed as an intersection of these halfspaces,

$$P = \{x \in \mathbb{R}^n \mid Ax \le b\}.$$

Let $x^1,\dots,x^v$ be the vertices of P. Then, P can also be expressed as a convex hull of these vertices,

$$P = \mathrm{CH}(x^1,\dots,x^v).$$

Next, we define slack matrix.

Definition 2. *The slack matrix S associated with a polytope P is a $m \times v$ matrix, whose rows are indexed by the facets and columns are indexed by the vertices, such that its $(i,j)^{th}$ entry is the slack of j^{th} vertex in the i^{th} constraint given by*

$$S_{i,j} = b_i - \langle A_i, x^j\rangle.$$

Observe that S is a nonnegative matrix.

Nonnegative rank

Let us first recall a definition of rank of a matrix. A matrix A has rank r if r is the smallest integer such that A can be written as a sum of r rank 1 matrices[7], i.e., if A is a $m \times n$ matrix over $\mathbb{R}$ of rank r, then there exists vectors $u_1,\dots,u_r \in \mathbb{R}^m$ and $v_1,\dots,v_r \in \mathbb{R}^n$ such that $A = \sum_{i=1}^r u_i v_i^T$. Define U to be the $m \times r$ matrix whose columns are $u_1,\dots,u_r$, and V to be the $r \times n$ matrix whose rows are $v_1^T,\dots,v_r^T$. Then, $A = UV$. Thus we have

Definition 3. *Let $A \in \mathbb{R}^{m\times n}$. Then,*

$$rank(A) = \min\{r \mid \exists U \in \mathbb{R}^{m\times r}, V \in \mathbb{R}^{r\times n} \text{ such that } A = UV\}.$$

Similar to this is the notion of nonnegative rank denoted by rank_+ for nonnegative matrices.

Definition 4. *Let $A \in \mathbb{R}^{m\times n}_{\ge 0}$. Then,*

$$rank_+(A) = \min\{r \mid \exists U \in \mathbb{R}^{m\times r}_{\ge 0}, V \in \mathbb{R}^{r\times n}_{\ge 0} \text{ such that } A = UV\}.$$

It is easy to see that $\mathrm{rank}(A) \le \mathrm{rank}_+(A)$ for a nonnegative matrix A[8].

[7] By singular value decomposition of a real matrix A of size $m\times n$ and rank r, there exists two orthogonal matrices Q and R such that $A = Q\Sigma R^T$, where Σ is a nonnegative diagonal matrix of size $m \times n$ such that its first r diagonal entries are nonzero. Let $Q_1,\dots,Q_m$ and $R_1,\dots,R_n$ be the columns vectors of Q and R respectively. Then, $A = \sum_{i=1}^r Q_i \Sigma_{i,i} R_i^T$. Let $u_i = \Sigma_{i,i} Q_i$ and $v_i = R_i$.

[8] Find examples of nonnegative matrices for which there is a large gap between its rank and nonnegative rank.

Yannakakis's theorm

Theorem 5. *If $P = \{x \in \mathbb{R}^n \mid Ax \le b\}$ is a polytope with slack matrix S, then $\text{rank}_+(S) = xc(P)$.*

Proof. *Claim 1*: "$\text{rank}_+(S) \ge xc(P)$".

Suppose that $\text{rank}_+(S) = r$, then we'll show that $xc(P) \le r$, i.e, there is an extended formulation of P of size r.

Since $\text{rank}_+(S) = r$, there exists matrices U and V of sizes $m \times r$ and $r \times v$ respectively such that $S = UV$. Let $U_1, \dots, U_m \in \mathbb{R}^r_{\ge 0}$ be the rows of U, and $V_1, \dots, V_v \in \mathbb{R}^r_{\ge 0}$ be the columns of V. Introducing a new vector $y \in \mathbb{R}^r$, consider a new polytope $Q = \{(x, y) \in \mathbb{R}^{n+r} \mid Ax + Uy = b, y \ge 0\}$ of size r. We'll show next that Q is an extended formulation of P, i.e, $P = \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^r \text{ s.t. } (x, y) \in Q\}$. This requires showing two inclusions.

Claim 1.1: "$P \supseteq \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^r \text{ s.t. } (x, y) \in Q\}$".

Consider a $x \in \mathbb{R}^n$ such that there exists $y \in \mathbb{R}^r$ and $(x, y) \in Q$. Then, they satisfy $Ax + Uy = b, y \ge 0$. Since y and U are nonnegative, clearly $Ax \le b$. Thus, $x \in P$.

Claim 1.2: "$P \subseteq \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^r \text{ s.t. } (x, y) \in Q\}$".

To show that the convex hull definition will be useful. Let $x^1, \dots, x^v$ be the vertices of P. Suppose we show for an arbitrary vertex $x^j \in P$ that there exists a $y_j \in \mathbb{R}^r$ such that $Ax^j + Uy_j = b$ and $y_j \ge 0$. Then, let $x \in P$ be an arbitrary point. It can be written as a convex combination $x = \sum_{j=1}^v c_j x^j$. Let $y = \sum_{j=1}^v c_j y_j$. Note since y_j's are nonnegatives, y is also nonnegative. Then, $Ax + Uy = \sum_{j=1}^v c_j (Ax^j + Uy_j) = (\sum_{j=1}^v c_j) b = b$. Thus $(x, y) \in Q$ proving the inclusion. Now consider an arbitrary vertex $x^j \in P$, and let $y_j = V_j$ (j^{th} column of V). Then,

$$Ax^j + Uy_j = Ax^j + UV_j = \begin{bmatrix} A_1 x^j + U_1 V_j \\ A_2 x^j + U_2 V_j \\ \vdots \\ A_m x^j + U_m V_j \end{bmatrix} = \begin{bmatrix} A_1 x^j + S_{1,j} \\ A_2 x^j + S_{2,j} \\ \vdots \\ A_m x^j + S_{m,j} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = b.$$

Thus, $(x^j, y_j) \in Q$ and y_j is too since V is nonnegative.

This completes the proof that Q is an extended formulation of P of size r and of Claim 1.

Claim 2: "$\text{rank}_+(S) \le xc(P)$".

Suppose that $xc(P) = r$, then we'll show that $\text{rank}_+(S) \le r$, i.e, the slack matrix S can be written as a product of $m \times r$ and $r \times v$ matrices.

Since $xc(P) = r$, there is an extended formulation $Q = \{(x, y) \mid Bx + Cy \le d\}$ of P with r inequalities, i.e, B is an $r \times n$ matrix and $d \in \mathbb{R}^r$. By definition, $P = \{x \in \mathbb{R}^n \mid \exists y \text{ s.t. } Bx + Cy \le d\}$. Next, we need a result:

Claim 2.1: Given A, b, suppose $P = \{x \mid Ax \le b\}$ is not empty. Further suppose that $\langle c, x\rangle \le \delta$ for all $x \in P$. Given $\delta' \le \delta$, show that the linear inequality $\langle c, x\rangle \le \delta'$ is a non-negative linear combination of the inequalities in $Ax \le b$[9].

Thus, each inequality $A_i x \le b_i$ in P is a nonnegative linear combination of inequalities in $Bx + Cy \le d$, i.e., there exists a nonnegative vector $p_i \in \mathbb{R}^r$ such that $p_i^T \begin{bmatrix} B & C & d \end{bmatrix} = \begin{bmatrix} A_i & 0 & b \end{bmatrix}$. In terms of slacks, we have $p_i^T(d - Bx - Cy) = b_i - A_i x$. Also, for each vertex x^j of P, fix a vector y_j such that $Bx^j + Cy_j \le d$ (existence due to extended formulation). Thus, for i^{th} inequality $A_i x \le b_i$ and j^{th} vertex x^j of P, we have $p_i^T(d - Bx^j - Cy_j) = b_i - A_i x^j = S_{i,j}$. Let U be a $m \times r$ matrix whose rows are p_i^T's, and V be a $r \times v$ matrix whose columns are $d - Bx^j - Cy_j$'s. Then, $UV = S$ where S is the slack matrix. This also shows $\text{rank}_+(S) \le r$, and proves Claim 2. □

Corollary 6. *[FMP+15] (a) If F is an extension of P, then $xc(F) \ge xc(P)$. (b) If Q is a face of P, then $xc(P) \ge xc(Q)$*

This result is useful for showing lower bounds. E.g. the correlation polytope has exponential extension complexity, and it can be shown that TSP polytope contains a face that is an extension of correlation polytope, which shows that TSP polytope also has exponential extension complexity [FMP+15].

Upper bound on $xc(P_{span})$

As an application of this theorem, we show that extension complexity of spanning tree polytope is $\mathcal{O}(n^3)$ by showing a bound on the nonnegative rank of its slack matrix. Let us consider only the second set of constraints from the description Example 1 for the slack matrix since the first and the third set of constraints can possibly increase the nonnegative rank only by a polynomial[10]. We had for all $A \subseteq V$, $\sum_{\{i,j\} \in \binom{A}{2}} x_{i,j} \le |A| - 1$. Let S be a slack matrix of size $2^n \times n^{n-2}$ such that its rows correspond to these constraints, i.e., the subsets of V, and columns correspond to spanning trees of K_n[11]. Given a subset of vertices A, and a spanning tree T, the slack of T in the constraint corresponding to A is

$$S_{A,T} = |A| - 1 - \sum_{\{i,j\} \in \binom{A}{2} \cap E(T)} x_{i,j},$$

where $E(T)$ is the set of edges of T. For each subset $A \subseteq V$, fix a vertex $k_A \in A$ arbitrarily. Now make T a rooted tree with k_A as the root. Then, the quantity

$$\sum_{\{i,j\} \in \binom{A}{2} \cap E(T)} x_{i,j}$$

is the number of nodes in $A \backslash \{k_A\}$ whose parent in T is also in A (one node for each edge). Note that k_A does not have a parent in T which means only $|A| - 1$ nodes in A

[9] This is a problem from Assignment 2 in this course.
[10] Verify this formally!
[11] By Cayley's formula, the number of spanning trees in K_n is n^{n-2}.

have a parent in T. Hence,

$$S_{A,T} = |A| - 1 - \sum_{\{i,j\} \in \binom{A}{2} \cap E(T)} x_{i,j}$$

is the number of nodes in $A \backslash \{k_A\}$ whose parent in T is not in A. Now, we introduce a variable $\lambda^T_{k_A,i,j}$ for each pair k_A and $\{i,j\}$ such that $k_A \notin \{i,j\}$[12], and defined as, for the given spanning tree T,

$$\lambda^T_{k_A,i,j} = \begin{cases} 1 & \text{if } j \text{ is the parent of } i \text{ in } T \text{ rooted at } k_A, \\ 0 & \text{otherwise,} \end{cases}$$

such that

$$S_{A,T} = \sum_{i \in A, j \notin A} \lambda^T_{k_A,i,j}.$$

Let $R = \{(k,i,j) \mid \{i,j\} \in \binom{V}{2}, k \notin \{i,j\}\}$, and $r = |R|$. Let U and W be matrices of sizes $2^n \times r$ and $r \times n^{n-2}$ respectively defined as follows: The rows and columns of U correspond to the subsets of A, and the elements of R respectively. Let U_A be the row vector corresponding to a subset A of V. Then, its $(k,i,j)^{th}$ entry is given by

$$U_{A,(k,i,j)} = \begin{cases} 1 & \text{if } k = k_A, i \in A, j \notin A \\ 0 & \text{otherwise.} \end{cases}$$

The rows and the columns of W correspond to the elements of R and the spanning trees of K_n respectively. Let W_T be the column vector corresponding to a spanning tree T of K_n. Then, its $(k,i,j)^{th}$ entry is given by

$$W_{T,(k,i,j)} = \begin{cases} 1 & \text{if } j \text{ is the parent of } i \text{ in } T \text{ rooted at } k, \\ 0 & \text{otherwise.} \end{cases}$$

Now $\langle U_A, W_T \rangle$ is the number of pairs (i,j) such that $j \notin A$ is the parent of $i \in A$ in T rooted at k_A, where $i \neq j$ and $k_A \notin \{i,j\}$. In other words, it is the number of nodes in $A \backslash k_A$ whose parent in T rooted at k_A is not in A. This shows $S_{A,T} = \langle U_A, W_T \rangle$, $S = UW$ and $\text{rank}_+(S) \le r = \binom{n}{2}(n-2) = \mathcal{O}(n^3)$. Thus $xc(P_{\text{span}}) = \mathcal{O}(n^3)$. This upper bound is also due to Wong [Won80] and Martin [Mar91].

Lower bound on $xc(P_{span})$

Proposition 7. *If P is a full-dimensional polytope in $\mathbb{R}^n$, then $xc(P) \ge n+1$.*

Proof. The polytope P must have at least $n+1$ vertices, and $n+1$ constraints defining them. Consider the $(n+1) \times (n+1)$ submatrix S' of S corresponding to these constraints and vertices. Each vertex satisfies exactly n of these constraints as equalities since P is full-dimensional. Thus, every column and every row of S' has exactly one nonzero entry and n zero entries making $\text{rank}_+(S') = n+1$. Since $\text{rank}_+(S') \le \text{rank}_+(S) = xc(P)$, this proves the result. □

[12] Note $i \neq j$.

Recall that in the description of spanning tree polytope in Example 1 none of the constraints are redundant and there is one equality that is satisfied by all points. Thus P_{span} has dimension $\binom{n}{2}-1$. By the previous proposition, $xc(P_{\text{span}}) \geq \binom{n}{2} = \Omega(n^2)$. Recently, Khoshkhah and Theis [KT18] improved this lower bound by a logarithmic factor and comment that this is best possible using combinatorial methods. It is an open problem to improve any of these upper or lower bound[13].

References

[Edm71] Jack Edmonds. Matroids and the greedy algorithm. *Mathematical Programming*, 1(1):127–136, December 1971.

[FMP+15] Samuel Fiorini, Serge Massar, Sebastian Pokutta, Hans Raj Tiwary, and Ronald de Wolf. Exponential Lower Bounds for Polytopes in Combinatorial Optimization. *J. ACM*, 62(2):17:1–17:23, May 2015.

[Jer75] R. G. Jeroslow. On defining sets of vertices of the hypercube by linear inequalities. *Discrete Mathematics*, 11(2):119–124, January 1975.

[KT18] Kaveh Khoshkhah and Dirk Oliver Theis. On the combinatorial lower bound for the extension complexity of the Spanning Tree polytope. *Operations Research Letters*, 46(3):352–355, May 2018.

[Mar91] R. Kipp Martin. Using separation algorithms to generate mixed integer model reformulations. *Operations Research Letters*, 10(3):119–128, April 1991.

[Rot17] Thomas Rothvoss. The Matching Polytope has Exponential Extension Complexity. *J. ACM*, 64(6):41:1–41:19, September 2017.

[Sch98] Alexander Schrijver. *Theory of Linear and Integer Programming*. John Wiley & Sons, June 1998.

[Won80] Richard T Wong. Integer programming formulations of the traveling salesman problem. In *Proceedings of the IEEE international conference of circuits and computers*, volume 149, page 152. IEEE Press Piscataway NJ, 1980.

[Yan91] Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. *Journal of Computer and System Sciences*, 43(3):441–466, 1991.

[13] posed by Goemans at the 2010 Cargèse Workshop on Combinatorial Optimization.