Lecture notes Mathematical Programming

Kateřina Vokálová

21. 5. 2024

1 Cut and correlation polytopes

Definition 1.1. A *polytope* $P \subseteq \mathbb{R}^n$ is the convex hull of a finite set of points in $\mathbb{R}^n$. It can also be viewed as a bounded set defined by a finite number of linear constraints (halfspaces in $\mathbb{R}$).

$$P = \text{conv}(\{v_1, v_2, \dots, v_k\})$$

$$P = \{x \in \mathbb{R}^n | Ax \leq b \text{ for } A \in \mathbb{R}^{r \times n}, b \in \mathbb{R}^r\}$$

Remark. The number of vertices defining a polytope may be exponential in the number of halfspaces. Consider the hypercube H_d in $\mathbb{R}^d$ which needs d inequalities $0 \leq x_i \leq 1$, but 2^d vertices ($H_d = \text{conv}(\{0,1\}^d)$).

Definition 1.2. For graph $G = (V, E)$ and cut $E' \subseteq E$ define its *incidence vector* $\chi^{E'}$ of size $|E|$ as

$$\chi_e^{E'} = \begin{cases} 1 & e \in E' \\ 0 & e \notin E' \end{cases}$$

and define the *cut polytope* as

$$\text{CUT}(G) := \text{conv}\{\chi^{E'} | \ E' \subseteq E \text{ is an edge cut}\}.$$

If G is the complete graph K_n, we simply denote $\text{CUT}(K_n)$ by CUT_n.

Definition 1.3. We define the *correlation polytope* as

$$\text{CORR}(n) = \text{conv}\{bb^T | \ b \in \{0,1\}^n\}$$

The polytope lies in $\mathbb{R}^{n^2}$. The feasible point of this polytope is a matrix $x \in \mathbb{R}^{n \times n}$.

2 Extension complexity and rectangle covering bound

Definition 2.1. The *extension complexity* of a polytope P, denoted by $\text{xc}(P)$, is the smallest number of facets of polytope $Q \subseteq \mathbb{R}^m$ such that P is a projection of Q.

Alternatively, the *extension complexity*, $\text{xc}(P)$, is the minimum number of inequalities required to describe Q, even when allowed to use auxiliary variables or extended formulations.

Definition 2.2. Let P be a polytope as defined in Definition 1.1. Then $S \in \mathbb{R}^{r \times k}$ defined as $S_{ij} := b_i - A_i v_j$ is the *slack matrix of* P w.r.t. $Ax \leq b$ and V.

Definition 2.3. We define the support matrix $\mathrm{supp}(S)$ for the slack matrix S as

$$\mathrm{supp}(S)_{i,j} = \begin{cases} 1 & S_{i,j} \neq 0 \\ 0 & S_{i,j} = 0 \end{cases}$$

Definition 2.4. A *rectangle* is the cartesian product of a set of row indices and a set of column indices. The *rectangle covering bound* is the minimum number of rectangles needed to cover all the 1-entries of $\mathrm{supp}(S)$.

Definition 2.5. *Monochromatic 1-rectangle* is a rectangle, which has all elements ones.

Theorem 2.1 (Yannakakis)**.** *Let M be any matrix with nonnegative real entries and $\mathrm{supp}(M)$ its support matrix. Then $\mathrm{rk}_+(M) \geq$ rectangle covering bound for $\mathrm{supp}(M)$.*

Theorem 2.2 (Yannakakis from the Lecture 11)**.** *For polytope P and its slack matrix S it holds*

$$\mathrm{xc}(P) = \mathrm{rk}_+(S).$$

3 $\mathbf{CORR}_{n-1} \approx \mathbf{CUT}_n$

Lemma 3.1. *For all $a \in \{0,1\}^n$, the inequality*

$$\left\langle 2\mathrm{diag}(a) - aa^T, x \right\rangle \leq 1$$

is valid for $x \in \mathrm{CORR}_n$.

Proof. We can show the inequality is satisfied for vertices $x = bb^T$ and by convexity, it is satisfied for every point of CORR_n. The inequality can be rewritten as

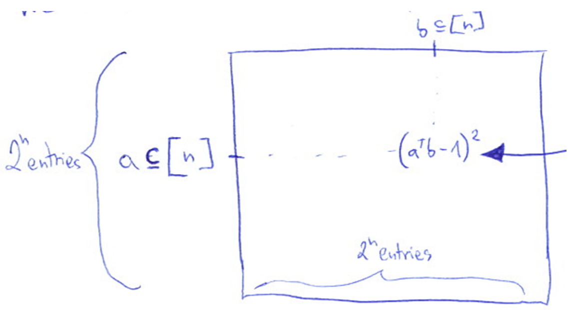
$$(1 - a^T b)^2 \geq 0$$

which trivially holds. $\square$

Definition 3.1. The slack matrix of the CORR_n is $2^n \times 2^n$ matrix $M^* = M^*(n)$

$$M^*_{a,b} = (a^T b - 1)^2.$$

Each row a represents a subset of $[n]$ (can be viewed as n-bit strings), the same applies to the column b.



Theorem 3.2. *For a slack matrix M^* of a correlation polytope CORR_n it holds that every 1-monochromatic rectangle cover of $\mathrm{supp}(M^*)$ has size $2^{\Omega(n)}$.*

Corollary 1.

$$\mathrm{rk}_+(M^*) \geq 2^{\Omega(n)}$$

Corollary 2.

$$\mathrm{xc}(\mathrm{CORR}_n) = 2^{\Omega(n)}$$

We have somewhat an equivalence between polytopes CUT_n and CORR_n formulated in the following theorem:

Theorem 3.3. *Let $M(n)$ denote the slack matrix of CUT_n, extended with a suitably chosen set of 2^n redundant inequalities. Then $M^*(n-1)$ occurs as a submatrix of $M(n)$ and hence*

$$\mathrm{xc}(\mathrm{CUT}_n) = 2^{\Omega(n)}.$$

The theorem can be also formulated as:

Lemma 3.4. *For every $n \geq 1$, the polytopes CUT_{n+1} and CORR_n are affine-equivalent.*

4 Lower bound on xc(CORR_n)

Definition 4.1. Define a *rectangle of the slack matrix M^**

$$D(n) = \{(a, b), a \subseteq [n], b \subseteq [n] |\ a \cap b = \emptyset\}.$$

There are only positions where a and b are different, so $a^T b = 0$ and $(a^T b - 1)^2 = (0-1)^2 = 1$. It is a monochromatic 1-rectangle.

Proposition 4.1.

$$|D(n)| = 3^n$$

Proof. We have n vertices and the vertex can have 3 states. Either it is in a, or it is in b, or it is neither in a nor b. □

Proposition 4.2 (Decomposition of D(n) into rectangles)**.** *Let us have the non-negative rank of matrix M^* $\mathrm{rk}_+(M^*) = k$. We can find $T \in \mathbb{R}^{2^n \times k}$ and $U \in \mathbb{R}^{k \times 2^n}$ such that $M^* = TU$. We will furtherly decompose M^* as*

$$M^* = T_1 U^1 + T_2 U^2 + \ldots + T_k U^k$$

with $T_i \in \mathbb{R}^{2^n \times 1}$ and $U_i \in \mathbb{R}^{1 \times 2^n}$.

For support matrices, it holds that

$$supp(M^*) = \bigcup_{i=1}^{k} supp(T_i U^i) = \bigcup_{i=1}^{k} supp(T_i) \times supp(U^i)$$

Moreover, if the all the matrices T_i, U^i are nonzero, we can define a rectangle $R_i = supp(T_i) \times supp(U^i)$ for each $T_i U^i$. These rectangles together cover the whole matrix $D(n)$

$$D(n) \subseteq \bigcup_{i=1}^{k} R_i.$$

Definition 4.2. A set $R \subseteq D(n)$ is *valid* if $\forall(a,b),(a',b') \in R : |a \cap b'| \neq 1$. For a valid R define two sets R_1, R_2 as following

$$R_1 := \{(a,b) \in R \text{ such that } n \in a, n \notin b\} \cup \{(a,b) \in R \text{ such that } (a \cup \{n\}, b) \notin R, n \notin b\}$$

$$R_2 := \{(a,b) \in R \text{ such that } n \notin a, n \in b\} \cup \{(a,b) \in R \text{ such that } (a, b \cup \{n\}) \notin R, n \notin a\}.$$

Lemma 4.3.

$$(a,b) \in R \Rightarrow (a,b) \in R_1 \cup R_2$$

Proof. Assume $a \cap b = \emptyset$ and we will split the proof into 3 cases: $n \in a$, $n \notin a, n \in b$, $n \notin a$ and $n \notin b$.

At first, $n \in a$ implies $n \notin b$, so $(a,b) \in R_1$.

Secondly, $n \notin a, n \in b$ implies $(a,b) \in R_2$.

Thirdly, we will split the case $n \notin a, n \notin b$ into subcases.

- $(a \cup \{n\}, b) \notin R \Rightarrow (a,b) \in R_1$
- $(a, b \cup \{n\}) \notin R \Rightarrow (a,b) \in R_2$
- $(a \cup \{n\}, b) \in R$ and $(a, b \cup \{n\}) \in R \Rightarrow$ we have item at row $a \cup \{n\}$ and column b and another item at row a and column $b \cup \{n\}$, so we must have the item at row $a \cup \{n\}$ and column $b \cup \{n\}$ since it is rectangle. This implies $(a \cup \{n\}, b \cup \{n\}) \in R$ which is contradiction with the assumption that they are disjoint.

□

Lemma 4.4. *Let Q be a polyhedron having f facets such that* CORR_n *is an affine image of Q. Then there exists a covering of $D(n)$ of size f.*

Proof. By Definition 2.1 of the extension complexity and Theorem 2.2,

$$f = \mathrm{xc}(\mathrm{CORR}_n) = \mathrm{rk}_+(M^*).$$

In Proposition 4.2, we have found a rectangle covering of size $\mathrm{rk}_+(M^*)$, so it is a rectangle covering of size f. □

Theorem 4.5.

$$\mathrm{xc}(\mathrm{CORR}_n) \geq \left(\frac{3}{2}\right)^n$$

Proof. Let $\rho(n)$ be the largest cardinality of any valid subset of $D(n)$. Any covering of $D(n)$ must have size $\geq \frac{|D(n)|}{\rho(n)} = \frac{3^n}{\rho(n)}$. It remains to prove that $\rho(n) \leq 2^n$, which we will show by proving that $\rho(n) \leq 2\rho(n-1)$ holds for all $n \geq 1$. We will prove this by induction on n.

Let R be a valid subset with sets R_1, R_2. Define the function $f : R \to D(n-1)$ as

$$f((a,b)) := (a \setminus \{n\}, b \setminus \{n\})$$

$$f(R_1) = \{(a \setminus \{n\}, b \setminus \{n\}) | (a,b) \in R_1\}$$

$$f(R_2) = \{(a \setminus \{n\}, b \setminus \{n\}) | (a,b) \in R_2\}$$

The two parts of each R_i are disjoint when n is subtracted, so it holds that $|R_i| = |f(R_i)|$.

Then by Lemma 4.3 and using induction hypothesis, we have

$$|R| \leq |R_1| + |R_2| = |f(R_1)| + |f(R_2)| \leq \rho(n-1) + \rho(n-1) = 2\rho(n-1)$$

□

Corollary 3.

$$\operatorname{xc}(\mathrm{CUT}_n) \geq \left(\frac{3}{2}\right)^n$$

5 Under construction

Let P be polytope (by default let it be 0/1-polytope so all vertices lie on $\{0,1\}^d$), S is slack matrix of P.

Definition 5.1. Let C be a cut. Define $1,-1$ encoding C' for each edge e:

$$C'_e := \begin{cases} -1 & e \in C \\ +1 & e \notin C \end{cases}$$

Proposition 5.1. *The $1,-1$ encoding can be obtained from $0,1$ encoding by linear transformations.*

Proof. $0,1 \rightsquigarrow -\frac{1}{2},\frac{1}{2} \rightsquigarrow -1,1 \rightsquigarrow 1,-1$ □

Definition 5.2. Denote by $U \subseteq V$ one part of vertices of the graph using the cut C. Then define χ_U as following

$$(\chi_U)_v = \begin{cases} +1 & v \in U \\ -1 & v \notin U \end{cases}$$

Then we have $C' = \chi_U \chi_U^T$

Have the graph K_n with vertices numbered from 1 to n (so the numbers are from set $[n]$) and $x_{ij} := x_i x_j$ corresponding to the edge between vertices i and j.

$$(w^T x - 1)^2 = \left(\sum_{i\in[n]} w_i x_i - 1\right)^2 \geq 0$$

Have the quadratic polynomial with variable vector x

$$\left(w^T x - 1\right)^2 = \left(\sum_{i\in[n]} w_i x_i - 1\right)^2 = \left(\sum_{i} w_i x_i\right)^2 - 2\sum_{i\in[n]} w_i x_i + 1 = \sum_{i,j} w_i w_j x_i x_j - 2\sum_{i} w_i x_i + 1 =$$

$$= \sum_{i,j} w_i w_j x_{ij} - 2\sum_{i} w_i x_i + 1$$