Lecture 6: Junta degree and a hardness amplification
CSE 599S: Entropy optimality, Winter 2016
Instructor: James R. Lee **Last updated:** January 24, 2016

1 Non-negative rank and positivity certificates

Recall the matrix $\mathcal{M}_n : \mathsf{QML}_+^n \times \{0,1\}^n \to \mathbb{R}_+$ from last lecture, defined by $\mathcal{M}_n(f,x) = f(x)$. Our goal is to prove a lower bound on $\mathrm{rank}_+(\mathcal{M}_n)$, and hence on $\bar{\gamma}(\mathrm{CUT}_n)$.

If $r = \mathrm{rank}_+(\mathcal{M}_n)$, it means we can write

$$f(x) = \mathcal{M}_n(f,x) = \sum_{i=1}^{r} A_i(f) B_i(x) \tag{1.1}$$

for some functions $A_i : \mathsf{QML}_+^n \to \mathbb{R}_+$ and $B_i : \{0,1\}^n \to \mathbb{R}_+$. (Here we are using a factorization $\mathcal{M}_n = AB$ where $A_{f,i} = A_i(f)$ and $B_{x,i} = B_i(x)$.)

More succinctly, we can write $f = \sum_{i=1}^r A_i(f) B_i$. Thus the low-rank factorization gives us a "proof system" for QML_+^n. Every $f \in \mathsf{QML}_+^n$ can be written as a conic combination of the functions $B_1, B_2, \ldots, B_r$, thereby certifying its positivity (since the B_i's are positive functions).

Let's think about natural families $\mathcal{B} = \{B_i\}$ of "axioms." Observe that QML_+^n is invariant under the natural action of S_n (the symmetric group on $\{1,\ldots,n\}$), where a permutation $\sigma : [n] \to [n]$ acts by permuting the coordinates:

$$\sigma f(x_1,\ldots,x_n) = f(x_{\sigma(1)},\ldots,x_{\sigma(n)})\,.$$

Thus we might expect that our family $\mathcal{B}$ should share this invariance. Once we entertain this expectation, there are natural small families of axioms to consider: The space of non-negative k-juntas for some $k \ll n$. (See Section 1.0.1 for exercises that explain why these are essentially the only small symmetric families of axioms.)

A *k-junta* $b : \{0,1\}^n \to \mathbb{R}$ is a function whose value only depends on k of its input coordinates. For a subset $S \subseteq \{1,\ldots,n\}$ with $|S| = k$ and an element $z \in \{0,1\}^k$, let $q_{S,z} : \{0,1\}^n \to \{0,1\}$ denote the function given by $q_{S,z}(x) = 1$ if and only if $x|_S = z$ (where we use $x|_S$ to denote the ordered restriction of x to the coordinates in S).

We let $\mathcal{J}_k = \{q_{S,z} : |S| \leqslant k, z \in \{0,1\}^{|S|}\}$. Observe that $|\mathcal{J}_k| \leqslant O(n^k)$. Let us also define $\mathrm{cone}(\mathcal{J}_k)$ as the set of all non-negative combinations of functions in $\mathcal{J}_k$.

Exericse (0 points) 1.1. Show that $\mathrm{cone}(\mathcal{J}_k)$ is precisely the set of all non-negative combinations of non-negative k-juntas.

If it were true that $\mathsf{QML}_+^n \subseteq \mathrm{cone}(\mathcal{J}_k)$ for some k, we could immediately conclude that $\mathrm{rank}_+(\mathcal{M}_n) \leqslant |\mathcal{J}_k| \leqslant O(n^k)$ by writing $\mathcal{M}_n$ in the form (1.1) where now $\{B_i\}$ ranges over the elements of $\mathcal{J}_k$ and $\{A_i(f)\}$ gives the corresponding non-negative coefficients that follow from $f \in \mathcal{J}_k$.

1.0.1 Symmetric families of axioms

Exercise (1 point) 1.2. Consider first the following lemma [Yannakakis 1991].

Lemma 1.3. *Let H be a subgroup of the symmetric group S_n with $|H| \geqslant |S_n|/\binom{n}{d}$ for some $d < n/4$. Then there exists a set $J \subseteq [n]$ such that $|J| \leqslant d$ and such that H contains all the even permutations that fix the elements of J.*

Using this lemma, prove the following. Let Q be a family of functions mapping $\{0,1\}^n$ to $\mathbb{R}$ and such that Q is invariant under the action of S_n, i.e. for every $\pi \in S_n$,

$$Q = \{x \mapsto q(\pi x) : q \in Q\} ,$$

where πx permutes the coordinates of x according to π.

Show that if $d < n/4$ and $|Q| < \binom{n}{d}$, then each $q \in Q$ can be written

$$q(x_1, \dots, x_n) = q'(x_{i_1}, \dots, x_{i_d}, x_1 + x_2 + \cdots + x_n) \tag{1.2}$$

for some $q' : \{0,1\}^d \times \mathbb{N} \to \mathbb{R}$. In other words, every $q \in Q$ depends on at most d coordinates and possibly also the value $\sum_{i=1}^n x_i$.

Exercise (1 point) 1.4. Use the preceding exercise to show the following. Suppose that $\mathsf{QML}_{2n}^+ \subseteq \text{cone}(Q)$ for some family Q that is invariant under the action of S_n, and such that $|Q| < \binom{2n}{d}$ for some $d < n/2$. Then $\mathsf{QML}_+^n \subseteq \text{cone}(\mathcal{J}_d)$. This shows that, invariant families of axioms of a given size, one cannot do much better than $\mathcal{J}_d$.

[Hint: Given $q \in \mathsf{QML}_n^+$, define $f \in \mathsf{QML}_{2n}^+$ by $f(x,y) = q(x)$. Now apply Exercise 1.2 to Q to investigate the structure of f.]

1.1 Junta degree and the dual cone

Clearly $\mathsf{QML}_+^n \subseteq \text{cone}(\mathcal{J}_n)$. We will now see that juntas cannot yield a smaller set of axioms for QML_+^n. Combined with Exercise 1.4, this shows that if $\mathsf{QML}_+^n \subseteq \text{cone}(Q)$ and Q is a family of non-negative functions that is invariant under the action of S_n (see Exercise 1.2), then $|Q| > c^n$ for some $c > 1$.

Theorem 1.5. *Consider the function $f : \{0,1\}^n \to \mathbb{R}_+$ given by $f(x) = (x_1 + x_2 + \cdots + x_n - 1)^2$. Then $f \notin \text{cone}(\mathcal{J}_{n-1})$.*

Proof. Suppose we write $f = \sum_{i=1}^N q_i$ where each q_i is non-negative. Clearly if $\sum_{i=1}^n x_i = 1$, then $f(x_1, \dots, x_n) = 0$, hence $q_i(x_1, \dots, x_n) = 0$ for every i. But if $q_i \in \mathcal{J}_{n-1}$, then there is some coordinate on which it does not depend. Without loss, suppose it is the first coordinate. Then $0 = q_i(1, 0, \dots, 0) = q_i(0, 0, \dots, 0)$. But $f(0, 0, \dots, 0) = 1$. We conclude that $f \notin \mathcal{J}_{n-1}$. □

Let us now prove this in a more roundabout way by introducing a few definitions. First, for $f : \{0,1\}^n \to \mathbb{R}_+$, define the *junta degree of f* to be

$$\deg_J(f) = \min\{k : f \in \text{cone}(\mathcal{J}_k)\} .$$

Since every f is an n-junta, we have $\deg_J(f) \leqslant n$.

Now because $\{f : \deg_J(f) \leqslant k\}$ is a cone (spanned by $\mathcal{J}_k$), there is a universal way of proving that $\deg_J(f) > k$. Say that a functional $\varphi : \{0,1\}^n \to \mathbb{R}$ is *k-locally positive* if for all $|S| \leqslant k$ and $z \in \{0,1\}^{|S|}$, we have

$$\sum_{x \in \{0,1\}^n} \varphi(x) q_{S,z}(x) > 0 .$$

These are precisely the linear functionals separating a function f from cone($\mathcal{J}_k$): We have $\deg_J(f) > k$ if and only if there is a k-locally positive functional φ such that $\sum_{x\in\{0,1\}^n} \varphi(x)f(x) < 0$. (This follows by the characterization of Exercise 1.1 together with the hyperplane separation theorem of [Lecture 5, Exercise 2.5].) Now we are ready to prove Theorem 1.5 in a different way.

Second proof of Theorem 1.5. We will use an appropriate k-locally positive functional. Define

$$\varphi(x) = \begin{cases} -1 & |x| = 0 \\ 1 & |x| = 1 \\ 0 & |x| > 1\,, \end{cases}$$

where $|x|$ denotes the hamming weight of $x \in \{0,1\}^n$.

Recall the the function f from the statement of the theorem and observe that by opening up the square, we have

$$\begin{aligned} \sum_{x\in\{0,1\}^n} \varphi(x)f(x) &= \sum_{x\in\{0,1\}^n} \varphi(x)\left(1 - 2\sum_i x_i + \sum_i x_i^2 + 2\sum_{i\neq j} x_i x_j\right) \\ &= \sum_{x\in\{0,1\}^n} \varphi(x)\left(1 - \sum_i x_i\right) = -1\,. \end{aligned} \tag{1.3}$$

Now consider some $S \subseteq \{1,\ldots,n\}$ with $|S| = k \leqslant n-1$ and $z \in \{0,1\}^k$. If $z = \mathbf{0}$, then

$$\sum_{x\in\{0,1\}^n} \varphi(x) q_{S,z}(x) = -1 + 1\cdot(n-k) \geqslant 0\,.$$

If $|z| > 1$, then the sum is 0. If $|z| = 1$, then the sum is non-negative because in that case $q_{S,z}$ is only supported on non-negative values of φ. We conclude that φ is k-locally positive for $k \leqslant n-1$. Combined with (1.3), this yields the statement of the theorem. □

Exercise (1 point) 1.6. Consider the *knapsack polynomial:* For $n \geqslant 1$ odd,

$$f(x) = \left(x_1 + x_2 + \cdots + x_n - \frac{n}{2}\right)^2 - \frac{1}{4}\,.$$

It is straightforward to check that $f(x) \geqslant 0$ for all $x \in \{0,1\}^n$. Define an appropriate locally positive functional to show that $\deg_J(f) \geqslant \lfloor \frac{n}{2} \rfloor$.

1.2 From juntas to general factorizations

So far we have seen that we cannot achieve a low non-negative rank factorization of $\mathcal{M}_n$ using k-juntas for $k \leqslant n-1$.

Remark 1.7. If one translates this into the setting of lift-and-project systems, it says that the k-round Sherali-Adams lift of the polytope

$$P = \left\{x \in [0,1]^{n^2} : x_{ij} = x_{ji},\ x_{ij} \leqslant x_{jk} + x_{ki} \quad \forall i,j,k \in \{1,\ldots,n\}\right\}$$

does not capture CUT_n for $k \leqslant n-1$.

In the next lecture, we will show that a non-negative factorization of $\mathcal{M}_n$ would lead to a k-junta factorization with k small (which we just saw is impossible). This will yield a lower bound on $\bar{\gamma}(\mathrm{CUT}_n)$.

For now, let us state the theorem we want to prove. We first define a submatrix of $\mathcal{M}_n$. Fix some integer $m \geqslant 1$ and a function $g : \{0,1\}^m \to \mathbb{R}_+$. Now define the matrix $M_n^g : \binom{[n]}{m} \times \{0,1\}^n \to \mathbb{R}_+$ given by

$$M_n^g(S, x) = g(x|_S)\,.$$

The matrix is indexed by subsets $S \subseteq [n]$ with $|S| = m$ and elements $x \in \{0,1\}^n$. Here, $x|_S$ represents the (ordered) restriction of x to the coordinates in S.

Theorem 1.8 (Chan-Lee-Raghavendra-Steurer 2013)**.** *For every $m \geqslant 1$ and $g : \{0,1\}^m \to \mathbb{R}_+$, there is a constant $C = C(g)$ such that for all $n \geqslant 2m$,*

$$\mathrm{rank}_+(M_n^g) \geqslant C \left(\frac{n}{\log n}\right)^{\deg_J(g)}\,.$$

Note that if $g \in \mathsf{QML}_m^+$ then M_n^g is a submatrix of $\mathcal{M}_n$. Since Theorem 1.5 furnishes a sequence of quadratic multi-linear functions $\{g_j\}$ with $\deg_J(g_j) \to \infty$, the preceding theorem tells us that $\mathrm{rank}_+(\mathcal{M}_n)$ cannot be bounded by any polynomial in n.

In fact, the groundbreaking work of [Fiorini, Massar, Pokutta, Tiwari, de Wolf 2012] showed earlier that $\mathrm{rank}_+(\mathcal{M}_n) \geqslant c^n$ for some constant $c > 1$. The advantage of Theorem 1.8 lies in its generality (allowing it to be extended to the setting of approximate lifts and semi-definite extended formulations).

Applying Theorem 1.8. We know that for every $g \in \mathsf{QML}_+^m$, we have $\bar{\gamma}(\mathrm{CUT}_{n+1}) = \mathrm{rank}_+(\mathcal{M}_n)$. Also fom Theorem 1.5, for every $m \geqslant 1$, we can find a function $g \in \mathsf{QML}_+^m$ such that $\deg_J(g) = m$.

Plugging this into Theorem 1.8 shows that for every fixed m,

$$\mathrm{rank}_+(\mathcal{M}_n) \geqslant \mathrm{rank}_+(M_n^g) \geqslant C(m) \left(\frac{n}{\log n}\right)^m\,.$$

In particular, we conclude that $\gamma(\mathrm{CUT}_n)$ cannot be bounded by any polynomial in n. One cannot obtain stronger bounds directly from Theorem 1.8 because the implicit constant C depends on the function g. Using a more delicate quantitative analysis, one can use the functions of Theorem 1.5 to achieve $\gamma(\mathrm{CUT}_n) \geqslant 2^{cn^{1/3}}$ for some constant $c > 0$. See [Lee-Raghavendra-Steurer 2015].