Lecture 13: Symmetric cone factorizations CSE 599I, Spring 2021
Instructor: James R. Lee

1 Factorizations of the slack matrix

Recall that QML_+^n is the set of nonnegative quadratic multilinear functions $f : \{0,1\}^n \to \mathbb{R}_+$, and the (infinite) matrix $\mathcal{M}_n : \mathsf{QML}_+^n \times \{0,1\}^n \to \mathbb{R}_+$ given by $\mathcal{M}_n(f,x) := f(x)$ is a slack matrix for the correlation polytope CORR_n. Thus understanding $\mathrm{rk}_+(\mathcal{M}_n)$ and $\mathrm{rk}_{\mathsf{psd}}(\mathcal{M}_n)$ is equivalent to understanding the LP and SDP extension complexities of CORR_n.

1.1 Cones of positive functions

Let $L^2(\{0,1\}^n)$ denote the vector space of functions $f : \{0,1\}^n \to \mathbb{R}$. Consider nonnegative functions $q_1, q_2, \ldots, q_r : \{0,1\}^n \to \mathbb{R}_+$, and the set of $f \in \mathrm{cone}(q_1, q_2, \ldots, q_r)$ that can be written $f = \lambda_1 q_1 + \cdots + \lambda_r q_r$ for $\lambda_1, \ldots, \lambda_r \geqslant 0$. All such functions are nonnegative, and we have the folowing characterization.

Lemma 1.1. *It holds that*

$$\mathrm{rk}_+(\mathcal{M}_n) = \min\left\{r : \mathsf{QML}_+^n \subseteq \mathrm{cone}(q_1, \ldots, q_r) \textit{ for some } q_1, \ldots, q_r : \{0,1\}^n \to \mathbb{R}_+\right\}.$$

Proof. Suppose that $\mathsf{QML}_+^n \subseteq \mathrm{cone}(f_1, \ldots, f_r)$. Then every $f \in \mathsf{QML}_+^n$ can be written as

$$f = \lambda_1(f) f_1 + \cdots + \lambda_r(f) f_r,$$

hence

$$\mathcal{M}_n(f,x) = f(x) = \langle \lambda(f), q(x) \rangle,$$

where $\lambda(f) = (\lambda_1(f), \ldots, \lambda_r(f))$ and $q(x) = (q_1(x), \ldots, q_r(x))$. Thus $\mathrm{rk}_+(\mathcal{M}_n) \leqslant r$. Reversing the construction yields the other direction. □

So we can think of $\mathrm{rk}_+(\mathcal{M}_n)$ as the minimum number of "axioms" that generate all the true quadratic inequalities on $\{0,1\}^n$.

1.2 SOS cones

For a subspace $\mathcal{U} \subseteq L^2(\{0,1\}^n)$, define the *sum of squares cone over* $\mathcal{U}$ by

$$\mathsf{sos}(\mathcal{U}) := \mathrm{cone}(q^2 : q \in \mathcal{U}).$$

Note that every $f \in \mathsf{sos}(\mathcal{U})$ is nonnegative on $\{0,1\}^n$ and, moreover, there is a certificate of the form

$$f = \sum_{i=1}^{k} q_i^2, \quad q_1, \ldots, q_k \in \mathsf{sos}(\mathcal{U}).$$

Lemma 1.2. *It holds that*

$$\mathrm{rk}_{\mathsf{psd}}(\mathcal{M}_n) \leqslant \min\{\dim(\mathcal{U}) : \mathsf{QML}_+^n \subseteq \mathsf{sos}(\mathcal{U})\} \leqslant \mathrm{rk}_{\mathsf{psd}}(\mathcal{M}_n)^2.$$

Proof. Suppose that $\mathsf{QML}_+^n \subseteq \mathsf{sos}(\mathcal{U})$ and $\dim(\mathcal{U}) = r$. Let $q_1, \dots, q_r : \{0,1\}^n \to \mathbb{R}$ be a basis for $\mathcal{U}$, and define $Q : \{0,1\}^n \to \mathbb{S}_r^+$ by

$$Q(x)_{ij} := q_i(x)q_j(x).$$

Given $p \in \mathcal{U}$, we can write $p = \sum_{i=1}^r \lambda_i q_i$. Define the rank-1 PSD matrix $\Lambda(p^2)_{ij} = \lambda_i \lambda_j$. Note that

$$\mathrm{Tr}(\Lambda(p^2)Q(x)) = \sum_{i,j} \lambda_i \lambda_j q_i(x) q_j(x) = p(x)^2.$$

Finally, given $f \in \mathsf{QML}_+^n$, write $f = \sum_{i=1}^k c_i p_i^2$ for some $p_1, \dots, p_k \in \mathsf{sos}(\mathcal{U})$. Then we have

$$f(x) = \mathrm{Tr}(\Lambda(f)Q(x)),$$

where $\Lambda(f) = \sum_{i=1}^k c_i \Lambda(p_i^2)$. We conclude that if $\mathsf{QML}_+^n \subseteq \mathsf{sos}(\mathcal{U})$, then $\mathrm{rk}_{\mathsf{psd}}(\mathcal{M}_n) \leqslant \dim(\mathcal{U})$.

For the other side of the inequality, assume $\mathrm{rk}_{\mathsf{psd}}(\mathcal{M}_n) \leqslant r$, and write $f(x) = \mathrm{Tr}(P(f)Q(x))$, where $\{P(f), Q(x) : f \in \mathsf{QML}_+^n, x \in \{0,1\}^n\} \subseteq \mathbb{S}_r^+$. Decomposing $P(f) = \sum_{i=1}^r \alpha_i(f)\alpha_i(f)^T$ and $Q(x) = \sum_{i=1}^r q_i(x)q_i(x)^T$, we have

$$\mathrm{Tr}(P(f)Q(x)) = \sum_{i,j} \langle a_i(f), q_j(x)\rangle^2,$$

implying that if $\mathcal{U} = \mathrm{span}\{(q_j(x))_i : i, j \in \{1, 2, \dots r\}\}$, then $\mathsf{QML}_+^n \subseteq \mathsf{sos}(\mathcal{U})$. □

2 Symmetric nonnegative factorizations

Let $\mathcal{S}_n$ denote the group of permutations on $\{1, \dots, n\}$. Then every $\pi \in \mathcal{S}_n$ acts on a function $f : \{0,1\}^n \to \mathbb{R}$ in the natural way, by permuting the coordinates:

$$\pi f(x_1, \dots, x_n) = f(x_{\pi(1)}, \dots, x_{\pi(n)}).$$

Note that QML_+^n is invariant under the action of $\mathcal{S}_n$: $f \in \mathsf{QML}_+^n \iff \pi f \in \mathsf{QML}_+^n$. So it is natural to study symmetric cones in Lemma 1.1 and Lemma 1.2.

2.1 Symmetric axioms are juntas

As we will see now, there are not many small symmetric sets of axioms. For every $x \in \{0,1\}^n$, let $|x| = x_1 + \cdots + x_n$ denote its Hamming weight. A *k-junta* $f : \{0,1\}^n \to \mathbb{R}$ is a function that depends on only k of its coordinates. Say that f is an *almost k-junta* if there are coordinates $i_1, \dots, i_k \in [n]$ such that

$$f(x_1, \dots, x_n) = g(x_{i_1}, \dots, x_{i_k}, |x|),$$

for some function $g : \{0,1\}^k \times \mathbb{R} \to \mathbb{R}$. For $1 \leqslant k \leqslant n$, let $\tilde{\mathcal{J}}_k$ denote the cone generated by all *nonnegative* almost k-juntas.

Lemma 2.1. *Consider $f_1, \dots, f_r : \{0,1\}^n \to \mathbb{R}_+$ and suppose that the family $\mathcal{F} = \{f_1, \dots, f_r\}$ is invariant under the coordinate action of $\mathcal{S}_n$. If $r \leqslant \binom{n}{k}$ for some $k < n/4$, then* $\mathrm{cone}(f_1, \dots, f_r) \subseteq \tilde{\mathcal{J}}_k$.

To prove the lemma, we need a few basic facts about group actions. Suppose a finite group G acts on a set $\mathcal{X}$ by permutations. For $x \in \mathcal{X}$, one defines $\mathsf{Stab}(x) := \{g \in G : gx = x\}$ and $\mathsf{Orb}(x) := \{gx : g \in G\}$. (More generally for $S \subseteq \mathcal{X}$, define $\mathsf{Stab}(S) := \{g \in G : gS = S\}$.) Then the orbit-stabilizer theorem implies that

$$|\mathsf{Stab}(x)| \cdot |\mathsf{Orb}(x)| = |G|, \quad \forall x \in \mathcal{X}. \tag{2.1}$$

We also need the following lemma.

Lemma 2.2. *Let H be a subgroup of $\mathcal{S}_n$ with $|H| \geqslant k!(n-k)!$ for some $k < n/4$. Then there is a set $S \subseteq [n]$ with $|S| \leqslant k$ such that H contains all even permutations in $\mathsf{Stab}(S)$.*

Proof. The orbits of H partition $[n] = B_1 \cup B_2 \cup \cdots \cup B_s$. Write them so that $|B_1| \geqslant |B_2| \geqslant \cdots \geqslant |B_s|$. It holds that

$$|H| \leqslant \prod_{i=1}^{s} |B_1|!$$

In particular, if $|H| \geqslant k!(n-k)!$, it must be that $|B_1| \geqslant n-k$. Define $T := B_1, S := [n] \setminus T, m := |T|$. It holds that H is a semidirect product $H_1 \rtimes H_2$, where H_1 acts transitively on T, and H_2 acts on S.

If there is some nontrivial partition $T = T_1 \cup T_2$ that H_1 preserves, then we obtain

$$|H| = |H_1||H_2| \leqslant 2\left((m/2)!\right)^2 (n-m)! < (n-k)!\,k!$$

for $k < n/4$. Hence it must be that H_1 preserves no such partition (it is said that H_1 *acts primitively on S*).

So H_1 is a primitive subgroup of $\mathcal{S}_m$. From basic group theory, we know that such a subgroup is either $\mathcal{S}_m$ or the alternating group $\mathcal{A}_m$ (the set of all *even* permutations of S), or else is substantially smaller:

$$|H| = |H_1||H_2| \leqslant \frac{m!}{\lfloor (m+1)/2 \rfloor!}(n-m)! < (n-k)!\,k!$$

for $k < n/4$. It follows that $H_1 = \mathcal{A}_m$ or $H_1 = \mathcal{S}_m$, hence H_1 contains all even permutations of T. Now $H = H_1 \rtimes H_2$ contains the normal subgroup $H_1 \times \{e_S\}$ where e_S is the identity permutation on S. □

Proof of Lemma 2.1. Since $\mathcal{F} = \{f_1, \ldots, f_r\}$ is itself invariant under the coordinate action of $\mathcal{S}_n$, it holds that $\mathsf{Orb}(f_1), \ldots, \mathsf{Orb}(f_r) \subseteq \mathcal{F}_r$, hence $|\mathsf{Orb}(f_i)| \leqslant r$ for each $i = 1, \ldots, r$.

Then (2.1) implies that $|\mathsf{Stab}(f_i)| \geqslant |\mathcal{S}_n|/r = n!/r$. Thus if $r < \binom{n}{k}$, we have $|\mathsf{Stab}(f_i)| \geqslant k!(n-k)!$. By Lemma 2.2, it holds that there is a set $J_i \subseteq [n]$ of coordinates with $|J_i| \leqslant k$ such that $\mathsf{Stab}(f_i)$ contains every even permutation that fixes J_i. We may assume that $|J_i| \geqslant 3$.

Now suppose there is an odd permutation π that fixes J_i but $\pi \notin \mathsf{Stab}(f_i)$. This implies there is some $x \in \{0,1\}^n$ such that $f_i(\pi x) \neq f_i(x)$. Since $|J_i| \geqslant 3$, there are two coordinates $a, b \in J_i$ such that $x_a = x_b$. If σ_{ab} is the permutation that swaps a and b, then $\sigma_{ab} x = x$. But $\pi\sigma_{ab}$ is an even permutation that fixes J_i, hence $\pi\sigma_{ab} \in \mathsf{Stab}(f_i)$, and

$$f_i(\pi\sigma_{ab}x) = f(x) = f(\pi x).$$

Thus $\mathsf{Stab}(f_i)$ actually contains *all permutations* that fix J_i.

Finally, note that if f_i is invariant under all permutations fixing J_i, then $f_i(x)$ is a function depending only on the coordinates of x in J_i and possibly $|x|$. Since $|J_i| \leqslant k$, that means f_i is an almost k-junta, completing the proof. □

Let $\mathcal{J}_k$ denote the cone generated by all nonnegative k-juntas $f : \{0,1\}^n \to \mathbb{R}_+$. At the expense of increasing n by a factor of 2, we can restrict our attention to such cones.

Lemma 2.3. *Suppose that $f_1, \ldots, f_r : \{0,1\}^{2n} \to \mathbb{R}_+$ and $\mathcal{F} = \{f_1, \ldots, f_r\}$ is invariant under the coordinate action of $\mathcal{S}_{2n}$. If $r < \binom{2n}{k}$ for some $k < n/2$ and $\mathsf{QML}_+^{2n} \subseteq \mathrm{cone}(f_1, \ldots, f_r)$, then $\mathsf{QML}_+^{n} \subseteq \mathcal{J}_{2k}$.*

Proof. For $x \in \{0,1\}^n$, define $\bar{x} \in \{0,1\}^n$ so that $\bar{x}_i = 1 - x_i$ for each $i \in [n]$, and define functions $\tilde{f}_1, \ldots, \tilde{f}_r : \{0,1\}^n \to \mathbb{R}_+$ by

$$\tilde{f}_i(x) := f_i(x, \bar{x}). \tag{2.2}$$

By Lemma 2.1, our assumptions imply that each f_i is an almost k-junta, i.e., depends on k coordinates of the input, and possibly its hamming weight. But note that $|(x,\bar{x})| = |x| + |\bar{x}| = n$. Therefore each $\tilde{f}_i$ is a genuine $2k$-junta. It follows that $\mathrm{cone}(\tilde{f}_1, \ldots, \tilde{f}_r) \subseteq \mathcal{J}_{2k}$.

Moreover, if $g \in \mathsf{QML}_+^n$, then $\hat{g}(x,y) := g(x)$ gives a function $\hat{g} \in \mathsf{QML}_+^{2n}$. By assumption, $\hat{g} = \lambda_1 f_1 + \cdots + \lambda_r f_r$ for some $\lambda_1, \ldots, \lambda_r \geqslant 0$, hence $g = \lambda_1 \tilde{f}_1 + \cdots + \lambda_r \tilde{f}_r$. Thus $\mathsf{QML}_+^n \subseteq \mathrm{cone}(\tilde{f}_1, \ldots, \tilde{f}_r)$, completing the proof. □

Inadequacy of juntas as axioms. To complete our study of the potential of small symmetric cones to capture QML_+^n, it suffices to show that QML_+^n is not contained in any junta cone.

Lemma 2.4. *Consider the function $f(x_1, \ldots, x_n) := (x_1 + x_2 + \cdots + x_n - 1)^2$. Then $f \in \mathsf{QML}_+^n$, but $f \notin \mathcal{J}_{n-1}$.*

This is straightforward: Suppose that q is an $(n-1)$-junta such that $q(e_1) = \cdots = q(e_n) = 0$. Then there is some coordinate $j \in [n]$ on which q does not depend and thus $q(0) = q(e_j) = 0$. On the other hand, $f(e_1) = \cdots = f(e_n) = 0$ but $f(0) = 1$. We conclude that $f \notin \mathcal{J}_{n-1}$.

2.2 Symmetric PSD factorizations

Let us now state the analog of Lemma 2.3 for symmetric PSD factorizations. For that, we need to generalize the notion of an SOS cone to obtain equality in Lemma 1.2. Let $L^2(\{0,1\}^n; \ell_2)$ denote the vector space of all functions $f : \{0,1\}^n \to \ell_2$, where ℓ_2 is the Hilbert space of square-summable sequences with norm $\|x\|_{\ell_2} = (\sum_{j \geqslant 1} |x_j|^2)^{1/2}$. For a subspace $\mathcal{U} \subseteq L^2(\{0,1\}^n; \ell_2)$, define

$$\mathsf{sos}(\mathcal{U}) := \mathrm{cone}\left(\|q\|_{\ell_2}^2 : q \in \mathcal{U}\right).$$

In other words, $f \in \mathsf{sos}(\mathcal{U})$ if and only if it can be written

$$f(x) = \sum_{i=1}^{m} \|q_i(x)\|_{\ell_2}^2$$

for some functions $q_1, q_2, \ldots, q_m \in \mathcal{U}$. The proof of the next result is similar to that of Lemma 1.2.

Lemma 2.5. *It holds that $\mathrm{rk}_{\mathsf{psd}}(\mathcal{M}_n) = \min\{\dim(\mathcal{U}) : \mathsf{QML}_+^n \subseteq \mathsf{sos}(\mathcal{U})\}$, where now the minimum is over all subspaces $\mathcal{U} \subseteq L^2(\{0,1\}^n; \ell_2)$.*

Let Q_d denote the subspace of all degree at most d multilinear polynomials $p : \{0,1\}^n \to \mathbb{R}$, i.e., those that can be written

$$p(x) = \sum_{S \subseteq [n], |S| \leqslant d} c_S \prod_{i \in S} x_i.$$

In other words, $Q_d = \mathrm{span}\left(\{\chi_S(x) : |S| \leqslant d\}\right)$, where $\chi_S(x) := \prod_{i \in S} x_i$.

Theorem 2.6. *Suppose that $\mathcal{U} \subseteq L^2(\{0,1\}^{2n}; \ell_2)$ with $\mathcal{U} = \mathrm{span}\left(q_1, \ldots, q_r\right)$ is such that $\{q_1, \ldots, q_r\}$ is invariant under the coordinate action of $\mathcal{S}_{2n}$. If $r < \binom{2n}{d}$ and $d < n/2$, then $\mathsf{QML}_+^n \subseteq \mathsf{sos}(Q_d)$.*

2.3 Equivariant factorizations

Consider a PSD factorization

$$\mathcal{M}_n(f, x) = \mathrm{Tr}(P(f)Q(x))$$

with $P : \mathsf{QML}_+^n \to \mathbb{S}_r^+, Q : \{0,1\}^n \to \mathbb{S}_r^+$.

Then Theorem 2.6 involves symmetric factorizations in the following sense: For every $\sigma \in \mathcal{S}_n$, there should exist some permutation matrix Π_σ such that

$$\mathcal{M}_n(f, \sigma x) = \mathcal{M}_n(\sigma f, x) = \mathrm{Tr}\left(P(f)\Pi_\sigma Q(x)\Pi_\sigma^\dagger\right).$$

There are ways to both strengthen and relax this notion. For instance, we have not exploited the full set of symmetries in QML_+^n. If we define $\pi_i(x_1, \dots, x_i, \dots x_n) = (x_1, \dots, 1 - x_i, \dots, x_n)$ then the maps $\pi_1, \dots, \pi_n$ generate an additional class of symmetries.

And, on the other hand, we could substantially weaken our notion of symmetry to reflect the fact that the PSD cone is invariant under a much broader class of transformations: Every congruence transformation $Q \mapsto AQA^\dagger$ where A is invertible. We examine an example.

Let $\mathcal{G}_n$ denote the group acting on $\{0,1\}^n$ by permutations, where $\mathcal{G}_n$ is generated by $\pi_1, \dots, \pi_n$ and the coordinate permutations $\sigma \in \mathcal{S}_n$. Suppose we have a factorization as above

$$\mathcal{M}_n(f, x) = \mathrm{Tr}(P(f)Q(x))$$

which is $\mathcal{G}_n$*-equivariant* in the following sense: For every $\sigma \in \mathcal{G}_n$, there is some invertible matrix A_σ such that

$$\mathcal{M}_n(f, \sigma x) = \mathrm{Tr}\left(P(f)A_\sigma Q(x)A_\sigma^\dagger\right).$$

It turns out that if $r < \binom{n}{d}$ for some $d < n/2$, then

$$\mathsf{QML}_+^n \subseteq \mathsf{sos}(\mathcal{U}_d),$$

where

$$\mathcal{U}_d = \mathrm{span}\left(\{\chi_S, \chi_{[n]}\chi_S : |S| \leqslant d\}\right).$$

Now by a trick similar to (2.2), we can lift our functions to $2n$ variables such that the parity $\chi_{[n]}(x, \bar{x})$ is always constant, yielding the following.

Theorem 2.7. *If $\mathcal{M}_{2n}$ has a $\mathcal{G}_n$-invariant PSD factorization of rank $r < \binom{2n}{d}$ and $d < n$, then $\mathsf{QML}_+^n \subseteq \mathsf{sos}(\mathcal{Q}_d)$.*