16 Sums of squares on the hypercube (continued)

Let $H_n = \{-1,1\}^n$ and let $h(n,k) = \binom{n}{0} + \cdots + \binom{n}{k}$ be the dimension of the space of polynomials of degree at most k on H_n. Define $\Sigma_{2k}(H_n)$ be the cone of polynomials in $\{-1,1\}^n$ that are k-sos:

$$\Sigma_{2k}(H_n) = \left\{ (p_S)_{\substack{S\subseteq[n]\\|S|\leq 2k}} \text{ s.t. } \sum_{\substack{S\subseteq[n]\\|S|\leq 2k}} p_S x^S \text{ is } k\text{-sos} \right\}. \tag{1}$$

$\Sigma_{2k}(H_n)$ is a proper cone living in the space of polynomials of degree at most $2k$ in $\{-1,1\}^n$, which has dimension $h(n,2k)$. We saw last lecture that

$$\Sigma_{2k}(H_n) = \pi(\mathbf{S}_+^{h(n,k)})$$

where π is the projection map

$$\pi(Q)_S = \sum_{\substack{U,V\subseteq[n]\\|U|,|V|\leq k\\U\triangle V=S}} Q_{U,V} \quad \forall S\subseteq[n], |S|\leq 2k. \tag{2}$$

The following proposition gives an expression for the dual of $\Sigma_{2k}(H_n)$:

Proposition 16.1. *The dual cone of* $\Sigma_{2k}(H_n)$ *is*

$$\Sigma_{2k}(H_n)^* = \left\{ (y_S)_{\substack{S\subseteq[n]\\|S|\leq 2k}} \text{ s.t. } \left[y_{U\triangle V}\right]_{|U|,|V|\leq k} \text{ positive semidefinite} \right\}. \tag{3}$$

Proof. Since $\Sigma_{2k}(H_n)$ is given as the projection of a positive semidefinite cone (of which we know the dual) we are going to use the following very simple lemma:

Lemma 16.1. *For any linear map* $\pi : \mathbb{R}^n \to \mathbb{R}^m$ *and set* X *we have* $(\pi(X))^* = \{y\in\mathbb{R}^m : \pi^*(y) \in X^*\}$.

Proof. We first show $\subseteq$: if $y \in (\pi(X))^*$ then $\langle y, \pi(x)\rangle \geq 0$ for all $x\in X$ and so $\pi^*(y)\in X^*$. Conversely, if $\pi^*(y)\in X^*$ then $\langle \pi^*(y), x\rangle \geq 0$ for all $x \in X$ and so $\langle y,\pi(x)\rangle \geq 0$ for all $x\in X$ which means that $y \in (\pi(X))^*$. □

Now to compute the dual of $\Sigma_{2k}(H_n)$ we just need to compute the adjoint of π defined in (2). Since $\pi : \mathbf{S}^{h(n,k)} \to \mathbb{R}^{h(n,2k)}$, the adjoint is a linear map $\pi^* : \mathbb{R}^{h(n,2k)} \to \mathbf{S}^{h(n,k)}$ and it must satisfy for any $y\in\mathbb{R}^{h(n,2k)}$ and $Q\in\mathbf{S}^{h(n,k)}$:

$$\langle \pi^*(y), Q\rangle = \langle y, \pi(Q)\rangle.$$

Note that

$$\begin{aligned}\langle y,\pi(Q)\rangle = \sum_{|S|\leq 2k} y_S \pi(Q)_S &= \sum_{|S|\leq 2k} \sum_{\substack{U,V\subseteq[n]\\|U|,|V|\leq k\\U\triangle V=S}} y_S Q_{U,V}\\ &= \sum_{\substack{U,V\subseteq[n]\\|U|,|V|\leq k}} y_{U\triangle V} Q_{U,V} \qquad = \langle Y, Q\rangle\end{aligned}$$

where $Y = [y_{U \triangle V}]_{|U|,|V| \leq k}$. Thus we have

$$\pi^*(y) = \Big[y_{U \triangle V} \Big]_{|U|,|V| \leq k}.$$

Combining Lemma 16.1 with the expression for π^* and the fact that the positive semidefinite cone is self-adjoint we get (3) as desired. □

Remark 1 (Interpretation in terms of moments of measures on the hypercube)**.** *If μ is a probability measure on H_n, we can consider its moments: $y_S = \mathbb{E}_\mu[x^S] = \mathbb{E}_\mu[\prod_{i \in S} x_i]$ for $S \subseteq [n]$. In this case, the matrix $\pi^*(y) = \Big[y_{U \triangle V} \Big]_{|U|,|V| \leq k}$ is nothing but $\mathbb{E}_\mu \Big[m(x)m(x)^T \Big] \succeq 0$ where $m(x) = [x^S]_{|S| \leq k}$.*

Application: maximum cut problem Recall the maximum cut problem, which consists in finding the cut with the largest value in a graph $G = (V, E)$. In Lecture 8 we formulated this problem as:

$$\begin{array}{ll} \text{maximise} & x^T L_G x \\ \text{subject to} & x \in \{-1, 1\}^n \end{array} \tag{4}$$

where

$$x^T L_G x = \frac{1}{2} \sum_{ij \in V} w_{ij} (x_i - x_j)^2$$

is the Laplacian of G. Problem (4) has the same optimal value as:

$$\begin{array}{ll} \text{minimise} & \gamma \\ \text{subject to} & \gamma - x^T L_G x \text{ nonnegative on } \{-1, 1\}^n \end{array} \tag{5}$$

We can define a hierarchy of semidefinite relaxations for the maximum cut problem via:

$$v_k = \min \quad \gamma \quad : \quad \gamma - x^T L_G x \text{ is } k\text{-sos on } \{-1, 1\}^n. \tag{6}$$

One can verify that $v_1 \geq v_2 \geq \cdots \geq v_n = \text{maxcut}(G)$ where $\text{maxcut}(G)$ is the value of the maximum cut (i.e., the optimal value of (4)). The equality $v_n = \text{maxcut}(G)$ follows from the fact that any nonnegative function on $\{-1, 1\}^n$ is n-sos (see second bullet point of Example 15.1). Let us compute the dual of the problem that defines v_k. First note that v_k can be expressed as:

$$v_k = \min_{\gamma \in \mathbb{R}} \quad \gamma \quad : \quad \gamma - x^T L_G x \in \Sigma_{2k}(H_n). \tag{7}$$

Let $y \in \Sigma_{2k}(H_n)^*$ be our dual variable. For any such y and feasible $\gamma \in \mathbb{R}$ of (7) we have:

$$\langle y, \gamma - x^T L_G x \rangle \geq 0. \tag{8}$$

It is important to note that x in Equation (8) plays the role of an *indeterminate*. The coefficients of $\gamma - x^T L_G x$ in the basis of square-free monomials are given by:

$$\begin{array}{rcl} S = \emptyset & : & \gamma - \text{Tr}(L_G) \\ S = \{i\} \;\; (i \in [n]) & : & 0 \\ S = \{i, j\} \;\; (i \neq j) & : & -2(L_G)_{ij} \end{array}$$

This implies that:

$$\langle y, \gamma - x^T L_G x \rangle = y_\emptyset (\gamma - \text{Tr}(L_G)) - 2 \sum_{1 \leq i < j \leq n} y_{\{i,j\}} (L_G)_{ij}$$

The objective function of (7) was γ. Thus for any $y \in \Sigma_{2k}(H_n)^*$ satisfying $y_\emptyset = 1$ we have the following lower bound on the optimal value of (7):

$$\gamma \geq y_\emptyset \operatorname{Tr}(L_G) + 2 \sum_{1 \leq i < j \leq n} y_{\{i,j\}} (L_G)_{ij}.$$

Note that the right-hand side of the inequality above can be rewritten as:

$$\sum_{1 \leq i,j \leq n} y_{\{i\} \triangle \{j\}} (L_G)_{ij}$$

The dual problem of (7) consists in finding the best lower bound, and so the dual problem is:

$$\max \quad \sum_{1 \leq i,j \leq n} y_{\{i\} \triangle \{j\}} (L_G)_{ij} \quad \text{s.t.} \quad \Big[y_{U \triangle V} \Big]_{|U|,|V| \leq k} \succeq 0, \quad y_\emptyset = 1. \tag{9}$$

Case $k = 1$: Let us look at the problem (9) when $k = 1$. This can be written as:

$$\max_{\substack{y_1,\ldots,y_n \\ y_{ij}\ (i<j)}} \quad \sum_{1 \leq i,j \leq n} y_{\{i\} \triangle \{j\}} (L_G)_{ij} \quad \text{s.t.} \quad \begin{bmatrix} 1 & y_1 & y_2 & \cdots & y_n \\ y_1 & 1 & y_{12} & \cdots & y_{1n} \\ y_2 & y_{12} & 1 & \cdots & y_{2n} \\ \vdots & & & \ddots & \vdots \\ y_n & y_{1n} & y_{2n} & \cdots & 1 \end{bmatrix} \succeq 0. \tag{10}$$

Note that the variables $y_1, \ldots, y_n$ do not play a role in the objective function. It is not difficult to show that (10) has the same optimal value as:

$$\max_{y_{ij} (i<j)} \quad \sum_{1 \leq i,j \leq n} y_{\{i\} \triangle \{j\}} (L_G)_{ij} \quad \text{s.t.} \quad \begin{bmatrix} 1 & y_{12} & \cdots & y_{1n} \\ y_{12} & 1 & \cdots & y_{2n} \\ \vdots & & \ddots & \vdots \\ y_{1n} & y_{2n} & \cdots & 1 \end{bmatrix} \succeq 0. \tag{11}$$

Indeed: first note that if $\{y_i\ (i = 1, \ldots, n), y_{ij}\ (1 \leq i < j \leq n)\}$ is feasible for (10) then $\{y_{ij}\ (1 \leq i < j \leq n)\}$ is feasible for (11) and has the same objective function value. Conversely, if $\{y_{ij}\}$ is feasible for (11) then letting $y_1 = \cdots = y_n = 0$ we get a feasible point of (10) with the same objective function value. This shows that (10) and (11) have the same optimal value. Observe that (11) is exactly the semidefinite relaxation we defined in Lecture 8 for the maximum cut:

$$\max \quad \operatorname{Tr}(L_G Y) \quad \text{s.t.} \quad Y \succeq 0, \quad Y_{ii} = 1\ \forall i = 1, \ldots, n.$$