1   Boolean Fourier analysis

In this lecture, we introduce Fourier analysis of boolean functions. As a motivation of this study, we will then use Fourier analysis to complete the proof of the Goemans-Williamson algorithmic gap instance for Max-Cut (from the previous lecture).

The functions that we deal with are of the form $f : \{-1, +1\}^n \to \mathbb{R}$. Boolean Fourier analysis involves representing such functions on the hypercube as multilinear polynomials. A multilinear polynomial in the variables $x_1, \cdots, x_n$ is a function of the form $\sum_{S \subseteq [n]} c_S \cdot \Pi_{i \in S} x_i$ where $c_S \in \mathbb{R}$ for all $S \subseteq [n]$. A *monomial* is a term of the form $\Pi_{i \in S} x_i$ for some $S \subseteq [n]$; we use the notation $\chi_S(x) \doteq \Pi_{i \in S} x_i$.

Proposition 1.1. *Any function $f : \{-1, +1\}^n \to \mathbb{R}$ can be expressed as a multilinear polynomial $\sum_{S \subseteq [n]} \hat{f}(S) \cdot \chi_S(x)$ that agrees with f on $\{-1, +1\}^n$. Here $\hat{f}(S) \in \mathbb{R}$ for all $S \subseteq [n]$.*

Proof. For any $S \subseteq [n]$, let $b(S)$ denote the vector in $\{-1, +1\}^n$ with $+1$s at positions S and -1s everywhere else. Define function $g(x) \doteq \sum_{S \subseteq [n]} f(b(S)) \cdot \Pi_{i \in S}(\frac{1+x_i}{2}) \cdot \Pi_{j \notin S}(\frac{1-x_j}{2})$. Observe that g agrees with f on $\{-1, +1\}^n$. Furthermore, the products in each term of the summation in g can be expanded to get a multilinear polynomial, which is the desired representation of f. $\qquad\square$

The representation of a boolean function f as in the above proposition is called its *Fourier expansion*. The coefficients $\hat{f}(S)$s are called *Fourier coefficients*. In the following, we study some properties of the Fourier expansion of boolean functions.

Theorem 1.2 (Parseval)**.** *If $f : \{-1, +1\}^n \to \mathbb{R}$ then $\sum_{S \subseteq [n]} \hat{f}(S)^2 = E_{x \in \{-1,+1\}^n}[f(x)^2]$.*

Proof. Expand $f(x)^2$ as follows (throughout the proof we assume $x \in \{-1, +1\}^n$).

$$
\begin{aligned}
f(x)^2 &= (\textstyle\sum_{S \subseteq [n]} \hat{f}(S) \cdot \chi_S(x)) \cdot (\textstyle\sum_{T \subseteq [n]} \hat{f}(T) \cdot \chi_T(x)) \\
&= \textstyle\sum_{S,T \subseteq [n]} \hat{f}(S)\hat{f}(T) \cdot \chi_{S \oplus T}(x)
\end{aligned}
$$

The last equality follows from the fact that $\chi_S(x) \cdot \chi_T(x) = \chi_{S \oplus T}(x)$ for $x \in \{-1, +1\}^n$. We also need the following simple fact.

Proposition 1.3.

$$
E[\chi_U(x)] = \begin{cases} 0 & U \neq \phi \\ 1 & U = \phi \end{cases}
$$

Proof. If $U = \phi$, then $\chi_U(x) = 1$ for all x, and hence $E[\chi_U(x)] = 1$. If $U \neq \phi$, then $E[\chi_U(x)] = E[\Pi_{i \in U} x_i] = \Pi_{i \in U} E[x_i] = 0$ since the expectation is taken over x such that each $x_1, \cdots, x_n$ is picked independently from $\{-1, +1\}$. $\square$

Now taking expectation over x uniformly chosen from $\{-1, +1\}^n$ and using Proposition 1.3,

$$E[f(x)^2] = \sum_{S,T \subseteq [n]} \hat{f}(S)\hat{f}(T) \cdot E[\chi_{S \oplus T}(x)] = \sum_{S \subseteq [n]} \hat{f}(S)^2$$

$\square$

It is easy to see (using Proposition 1.3) that:

Proposition 1.4. *If* $f : \{-1, +1\}^n \to \mathbb{R}$ *and* $S \subseteq [n]$, *then* $\hat{f}(S) = E_{x \in \{-1,+1\}^n}[f(x) \cdot \chi_S(x)]$.

A pair A, B of random $\{-1, +1\}$ variables are call *ρ-correlated random bits* if $E[A] = E[B] = 0$ and $E[A \cdot B] = \rho$. Note that this definition is symmetric in A and B. The above definition is also equivalent to choosing A uniformly from $\{-1, +1\}$ and then flipping the sign of A independently with probability $\frac{1}{2} - \frac{1}{2}\rho$ to obtain B.

A pair x, y of random $\{-1, +1\}^n$ variables are said to be *ρ-correlated* if for every $i \in [n]$, (x_i, y_i) are ρ-correlated random bits that are independent of $\{x_j \mid j \neq i\} \cup \{y_j \mid j \neq i\}$.

For any function $f : \{-1, +1\}^n \to \mathbb{R}$, the *noise stability of f at ρ* is defined to be the quantity $S_\rho(f) \doteq E_{x,y}[f(x) \cdot f(y)]$ where x, y are ρ-correlated random variables.

Proposition 1.5. $S_\rho(f) = \sum_{S \subseteq [n]} \hat{f}(S)^2 \cdot \rho^{|S|}$

Proof. Using the Fourier expansion of f, $f(x) \cdot f(y) = \sum_{S,T \subseteq [n]} \hat{f}(S)\hat{f}(T) \cdot \chi_S(x)\chi_T(y)$. So $S_\rho(f) = \sum_{S,T \subseteq [n]} \hat{f}(S)\hat{f}(T) \cdot E[\chi_S(x)\chi_T(y)]$. We now establish the following fact (when x, y are ρ-correlated):

$$E[\chi_S(x)\chi_T(y)] = \begin{cases} 0 & \text{if } S \neq T \\ \rho^{|S|} & \text{if } S = T \end{cases} \tag{1}$$

We can write $\chi_S(x)\chi_T(y) = (\Pi_{i \in S \setminus T} x_i)(\Pi_{i \in S \cap T} x_i y_i)(\Pi_{i \in T \setminus S} y_i)$. Since each pair x_i, y_i is independent of all other pairs x_j, y_j (for $j \neq i$), we have:

$$E[\chi_S(x)\chi_T(y)] = (\Pi_{i \in S \setminus T} E[x_i])(\Pi_{i \in S \cap T} E[x_i y_i])(\Pi_{i \in T \setminus S} E[y_i])$$

When $S \neq T$ either $S \setminus T$ or $T \setminus S$ is non-empty and since $E[x_i] = E[y_i] = 0$ for all $i \in [n]$, $E[\chi_S(x)\chi_T(y)] = 0$. When $S = T$, we have $E[\chi_S(x)\chi_T(y)] = \Pi_{i \in S} E[x_i y_i] = \rho^{|S|}$.

Now using Equation 1, we immediately obtain the proposition. $\square$

Proposition 1.6. *If* $f : \{-1, +1\}^n \to \mathbb{R}$ *and* $-1 \leq \rho \leq 0$, *then* $S_\rho(f) \geq \rho \cdot E[f(x)^2]$.

Proof. From Proposition 1.5, we have $S_\rho(f) = \sum_{S \subseteq [n]} \hat{f}(S)^2 \rho^{|S|} \geq \rho \sum_{S \subseteq [n]} \hat{f}(S)^2$ since $\rho \in [-1, 0]$. And from Theorem 1.2, we have $\sum_{S \subseteq [n]} \hat{f}(S)^2 = E[f(x)^2]$. $\square$

2

2 Finishing the proof of Max-Cut algorithmic gap

Recall the Max-Cut algorithmic gap instance G constructed in the previous lecture. The vertices are $V = \{-1, +1\}^n$ and the edge-distribution picks edge (x, y) where $x \in \{-1, +1\}^n$ uniformly and $y \in \{-1, +1\}^n$ is obtained by negating each component of x independently with probability $\frac{1}{2} - \frac{1}{2}\rho^*$. Observe that the edge distribution x, y corresponds to ρ^*-correlated random variables. Any feasible SDP solution is a mapping $F : \{-1, +1\}^n \to \mathbb{R}^d$ (for some dimension d) where $F(v)$ is a unit vector for each $v \in \{-1, +1\}^n$; the SDP value of such a solution is $\mathsf{Value}(F) = E_{x,y}[\frac{1}{2} - \frac{1}{2}\langle F(x), F(y)\rangle]$ (here x, y are ρ^*-correlated). We prove the following upper bound on the SDP value of instance G.

Theorem 2.1. $SDP(G) \leq \frac{1}{2} - \frac{1}{2}\rho^*$.

Consider any SDP solution (in dimension d) given by $F_j : \{-1, +1\}^n \to \mathbb{R}$ for $j = 1, \cdots, d$ such that $\sum_{j=1}^d F_j(x)^2 = 1$ for all $x \in \{-1, +1\}^n$. We will show that $\mathsf{Value}(F) \leq \frac{1}{2} - \frac{1}{2}\rho^*$, which is equivalent to showing $E_{x,y}[\langle F(x), F(y)\rangle] = \sum_{j=1}^d E_{x,y}[F_j(x) \cdot F_j(y)] \geq \rho^*$.

Since x, y are ρ^*-correlated, $E_{x,y}[F_j(x) \cdot F_j(y)] = S_{\rho^*}(F_j)$ for any $j \in [d]$. Now since $-1 \leq \rho^* \leq 0$, Proposition 1.6 implies that $S_{\rho^*}(F_j) \geq \rho^* \cdot E_x[F_j(x)^2]$ (for all $j \in [d]$). Thus, $\sum_{j=1}^d E_{x,y}[F_j(x) \cdot F_j(y)] \geq \sum_{j=1}^d \rho^* \cdot E_x[F_j(x)^2] = \rho^* \cdot E_x[\sum_{j=1}^d F_j(x)^2] = \rho^*$. This completes the proof of Theorem 2.1.