6.896 Sublinear Time Algorithms	May 8, 2007
Lecture 0	
Lecturer: Ronitt Rubinfeld	*Scribe: Erdong Chen*

1 Linear Test

Last time, we showed: over some other domain, $\exists f$ s.t. $\delta_f = 2/9$, but f is 2/3-far from linear.

This time, we will consider boolean functions, and analyze it via fourier analysis.

For boolean functions $f : \{\pm 1\}^n \to \{\pm 1\}$, we define $\delta_f = Pr[f(x)f(y) \neq f(x \odot y)]$, where $\odot$ is coordinate-wise multiplication.

The multiplication rules:

X	1	-1
1	1	-1
-1	-1	1

Then we will define **linear function**: $\chi_S(x) = \sum_{i \in S} x_i$ for any $S \subset \{1..n\}$.

The following analysis are based on the above domains and functions.

It is easy to see the following,

$$f(x)f(y)f(x \odot y) = \begin{cases} 1 \text{ if test accepts on } x, y \\ -1 \text{ if test fails on } x, y \end{cases} \tag{1}$$

Equivalently, we have

$$\frac{1 - f(x)f(y)f(x \odot y)}{2} = \begin{cases} 0 \text{ if test accepts on } x, y \\ 1 \text{ if test fails on } x, y \end{cases} \tag{2}$$

We will prove that this is an indicator variable, so we have the rejection probability: $\delta_f = E[\frac{1-f(x)f(y)f(x\odot y)}{2}]$.

2 Fourier analysis on hypercube

Let's consider the first class of functions, vector space of all n-bit functions mapping to $\Re$, $G = \{g | g : \{\pm 1\}^n \to \Re\}$.

We also have $\dim(G) = 2^n$, and we also know that all functions can be written in combination of 2^n basis functions(We will not prove this in the class).

Let's see some examples.

Example 1

indicator function:

$$e_a(x) = \begin{cases} 1 \text{ if } x = a \\ 0 \text{ o.w.} \end{cases} \tag{3}$$

A function g(can be represented by a 2^n-bit vector) is a weighted combination of basis functions, i.e., $g = \sum_a e_a(x)g(a)$, where $g(a)$ is a scalar.

Example 1

Inner product: $< f, g >= \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} f(x)g(x)$.

Basis χ_S is orthonogonal w.r.t. inner product.

First, $< \chi_S, \chi_S >= \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} \chi_S(x)\chi_S(x) = 1$.

Secondly, if $S \neq T$, we have

$$< \chi_S, \chi_T >= \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} \chi_S(x)\chi_T(x) = \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} \prod_{i \in S} x_i \prod_{i \in T} x_i$$

$$= \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} \prod_{i \in S\Delta T} x_i (*)$$

If $i \in S \cap T$, $x_i^2 = 1$, so drop out. The symmetric difference $S\Delta T$ is non-empty, because $S \neq T$. pick $j \in S\Delta T$, we have

$$(*) = \frac{1}{2^n} \sum_{x \in \{\pm 1\}^n} \prod_{j \in S\Delta T} (x_j^{(1)} + x_j^{(-1)})$$

, where

$$x_j^{(1)} = (x_1, x_2, \ldots, x_{j-1}, 1, x_{j+1}, \ldots, x_n)$$
$$x_j^{(-1)} = (x_1, x_2, \ldots, x_{j-1}, -1, x_{j+1}, \ldots, x_n)$$

Thus, we have

$$(*) = \frac{1}{2^n} \sum_{\text{pairs } x^{(1)}, x^{(-1)} w.r.t. j} (1 * \prod_{i \in S\Delta T \setminus \{j\}} x_i^{(1)} + (-1) * \prod_{i \in S\Delta T \setminus \{j\}} x_i^{(-1)})$$

.

Because $x_i^{(1)} = x_i^{(-1)}$ for$i \in S\Delta T \setminus \{j\}$,we have $(*) = 0$.

Thus, any function f is expressible as linear combination of χ_S, a.k.a. "Fourier coefficients".

Def: $\hat{f}(S) \equiv < f, \chi_S >= \frac{1}{2^n} \sum_x f(x) \chi_S(x)$.

Theorem 1 $f(x) = \sum_S \hat{f}(S) \chi_S(x)$

We first consider the Fourier coefficients of linear functions. f is linear $\leftrightarrow \exists S \subseteq [n]$ s.t. $\hat{f}(S) = 1$ and $\forall T \neq S, \hat{f}(T) = 0$.

Lemma 2 *$forall S \subseteq [n], \hat{f}(S) = 1 - 2dist(f, \chi_S)$, where $dist(f, \chi_S) = Pr_x[f(x) \neq \chi_S(x)]$.*

Proof $2^n \hat{f}(S) = \sum f(x)\chi_S(x) = \sum_{x \text{ s.t. } f(x)=\chi_S(x)} 1 + \sum_{x \text{ s.t. } f(x) \neq \chi_S(x)} (-1) = 2^n(1 - \text{dist}(f, \chi_S(x))) - 2^n \text{dist}(f, \chi_S) = 2^n(1 - 2\text{dist}(f, \chi_S(x)))$. ■

Example 2

$\forall x, f(x) = -1$

$\forall S \neq \emptyset$, for $j \in S$, $f(x_1, x_2, \ldots, x_{j-1}, 1, x_{j+1}, \ldots, x_n) = a$, and $f(x_1, x_2, \ldots, x_{j-1}, -1, x_{j+1}, \ldots, x_n) = -a$.

thus $\text{dist}(f, \chi_S) = 1/2$, so $\hat{f}(S) = 0$.

On the other hand, $\text{dist}(f, \chi_\emptyset) = 1, \hat{f}(\emptyset) = -1$.

Lemma 3 *Two distinct linear functions differ on exactly half points*

Proof for $S \neq T, 0 =< \chi_S, \chi_T >= 1 - 2\text{dist}(\chi_S, \chi_T)$, so $\text{dist}(\chi_S, \chi_T) = 1/2$ ■

$< f, g >=< \sum \hat{f}(S)\chi_S, \sum \hat{g}(T), \chi_T >= \sum_{S,T} \hat{f}(S)\hat{g}(T) < \chi_S, \chi_T >= \sum_S \hat{f}(S)\hat{g}(S)$, it is called "**Plancherel**".

$< f, f >= \sum_S \hat{f}(S)^2$, it is called "**Parseval's Identity**".

f is boolean(i.e. $f(x) \in \{\pm 1\}, < f, f >= \frac{1}{2^n} \sum \hat{f}(x)^2 = 1$. We'll call it "**Boolean Parseval's**"

Theorem 4 *f is δ_f-close to some linear function.*

Proof

If $S = T = U, \chi_S(x)\chi_T(y)\chi_U(x \odot y) = \prod_{i \in S} x_i y_i (x \odot y)_i = \prod_{i \in S} x_i y_i x_i y_i = 1$.

If $\neg(S = T = U)$, then either $S \neq U$ or $T \neq U$.

$E_{x,y}[\chi_S(x)\chi_T(y)\chi_U(x \odot y)] = E_{x,y}[\prod_{i \in S} x_i \prod_{j \in T} y_j, \prod_{k \in U} x_k \prod l \in U y_l] = E_{x,y}[\prod_{i \in S \Delta U} x_i \prod_{j \in T \Delta U} y_j] = E_x[\prod_{i \in S \Delta U} x_i] E_y[\prod_{j \in T \Delta U} y_j](**)$

If $S \neq U$, then $S \Delta U \neq \emptyset$, so $E_x[\chi_{S \Delta U}(x)] = 0$. If $T \neq U$, then $E_y[\chi_{T \Delta U}(y)] = 0$.

Thus, $(**) = 0$ because either $S \neq U$ or $T \neq U$.

$$E_{x,y}[f(x)f(y)f(x \odot y)] = E_{x,y}[\sum_S \hat{f}(S)\chi_S(x) \sum_T \hat{f}(T)\chi_T(x) \sum_U \hat{f}(U)\chi_U(x \odot y)]$$

$$= E_{x,y}[\sum_{S,T,U} \hat{f}(S)\hat{f}(T)\hat{f}(U)\chi_S(x)\chi_T x \chi_U(x \odot y)]$$

$$= \sum_{S,T,U} \hat{f}(S)\hat{f}(T)\hat{f}(U) E_{x,y}[\chi_S(x)\chi_T x \chi_U(x \odot y)]$$

because the proof above

$$= \sum_S \hat{f}(S)^3$$

$$\leq (\max_S \hat{f}(S)) \sum_{S'} \hat{f}(S')^2$$

by Boolean Parseval's, we have $\hat{f}(S')^2 = 1$,

$$= \max_S \hat{f}(S)$$

$$= 1 - 2 \min_S \mathrm{dist}(f, \chi_S)$$

so $\delta_f \geq 1/2 - 1/2(1 - 2\max_S \mathrm{dist}(f, \chi_S)) = \min_S \mathrm{dist}(f, \chi_S)$. ■