1 The Hermite Polynomial and Fourier Coefficients

Let $\gamma$ be the 1-dimensional gaussian measure and $f : \mathbb{R} \to \mathbb{R}$ a function in $L^2(\gamma)$ such that the set of points $x \in \mathbb{R}$ where $f(X)$ is discontinuous has measure 0. Consider the probability measure $\{-1, 1\}^n_\theta$ on $n$ variables, where each variable equals -1 independently with probability $(1 - \theta)/2$ and equals +1 independently with probability $(1 + \theta)/2$. Let $f_n : \{-1, 1\}^n \to \mathbb{R}$ be a function such that $f_n(x_1, \ldots, x_n) = f(\sum_{i=1}^n (x_i - \theta)/\sqrt{n(1 - \theta^2)})$, and consider the basis of all symmetric functions $W_n^k(x_1, \ldots, x_n) = (1 - \theta^2)^{-k/2}(n_k)^{-1/2}(\sum_{S \subseteq [n]:|S| = k} \prod_{i \in S} (x_i - \theta))$. From the previous lecture, we know that $f_n(x_1, \ldots, x_n) = \sum_{k=0}^n \hat{f}_n(k)W_n^k(x_1, \ldots, x_n)$ and $f(X) = \sum_{k=0}^n \hat{f}(k)h_k(X)$, where $\hat{f}_n(k) = \langle f_n, W_n^k \rangle$, $\hat{f}(k) = \langle f, h_k \rangle$, and $h_k$ is the normalized $k$th Hermite polynomial. (See previous lecture for full definitions). We now prove the following theorem:

**Theorem 1** $\forall k \in \mathbb{N}$, $\lim_{n \to \infty} \hat{f}_n(k) = \hat{f}(k)$.

**Proof:** For notation it will be useful to define the random variable, $X_n = \sum_{i=1}^n (x_i - \theta)/\sqrt{n(1 - \theta^2)}$. To prove our theorem, we will prove $\lim_{n \to \infty} \langle h_k(X_n), f_n(x_1, \ldots, x_n) \rangle_\theta = \lim_{n \to \infty} \langle h_k(X_n), f_n(x_1, \ldots, x_n) \rangle_\theta = \lim_{n \to \infty} \langle h_k(X_n), f(X_n) \rangle_\theta = \langle h_k(X), f(X) \rangle_\gamma$. The second equality follows by definition as $f_n(x_1, \ldots, x_n) = f(X_n)$. The third equality follows by the central limit theorem, which implies that for a fixed $k$, $\lim_{n \to \infty} \langle h_k(X_n), f(X_n) \rangle_\theta = \langle h_k(X), f(X) \rangle_\gamma$. Therefore, we just need to prove $\lim_{n \to \infty} \langle h_k(X_n), f_n(x_1, \ldots, x_n) \rangle_\theta = \lim_{n \to \infty} \langle h_k(X_n), f_n(x_1, \ldots, x_n) \rangle_\theta$. To complete the proof, we prove the following statement by induction on $k$, which implies the statement above.

$$\lim_{n \to \infty} E_\theta[||(W_0^n(x_1, \ldots, x_n), \ldots, W_k^n(x_1, \ldots, x_n)) - (h_0(X_n), \ldots, h_k(X_n))||_2] = 0$$

The base case is trivial, as $W_0^n(x_1, \ldots, x_n) = h_0(X_n) = 1$. The inductive step can be proved by noting:

---

1By symmetric, we mean a function $f_n$ such that $f_n(x_1, \ldots, x_n) = f_n(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ for any permutation $\sigma$. 

---
1. By the central limit theorem, \( \lim_{n \to \infty} \langle h_i(X_n), h_j(X_n) \rangle_\theta = \delta_{i,j} \)

2. \( W^n_i(x_1, ..., x_n) \) is a symmetric polynomial of degree \( i \). (See Footnote 1 for definition of symmetric).

3. \( h_i(X_n) \) and \( W^n_i(x_1, ..., x_n) \) have positive coefficient for all monomials of highest degree.

We leave the formal proof of the inductive step to the reader.

\( \square \)

**Example:** To illustrate the use of this theorem, consider the majority function \( f_n(x_1, ..., x_n) = \text{Maj}(x_1, ..., x_n) \) and the uniform measure \( \{-1, 1\}^n \). Note that if we define \( f(X) = \text{sgn}(X) \), where

\[
\text{sgn}(X) = \begin{cases} 
-1 & \text{if } X < 0 \\
0 & \text{if } X = 0 \\
+1 & \text{if } X > 0 
\end{cases}
\]

then \( f_n(x_1, ..., x_n) = f((\sum_{i=1}^n x_i)/\sqrt{n}) \) and we can apply Theorem 1. Although computing \( \hat{f}_n(k) \) is difficult, Theorem 1 implies that if we can compute \( \hat{f}(k) \), then it will be a good estimate of \( \hat{f}_n(k) \) for large \( n \).

To compute \( \hat{f}(k) = \langle f, h_k \rangle_\gamma \), first note that \( f \) is an odd function and \( h_k \) is an even function when \( k \) is even. Therefore, \( \hat{f}(k) = 0 \) for even \( k \), and we only need to compute \( \hat{f}(k) \) for odd \( k \). For odd \( k \):
\begin{align*}
\hat{f}(k) &= \langle f, h_k \rangle_\gamma = (2/\sqrt{k!}) \int_0^\infty H_k(x) d\gamma(x) \\
&= (-2/\sqrt{2\pi k!}) \int_0^\infty \frac{d^k}{dx^k}(e^{-x^2/2}) dx \\
&= (-2/\sqrt{2\pi k!}) \cdot \left( \frac{d^{k-1}}{dx^{k-1}}(e^{-x^2/2}) \right|_0^\infty \\
&= \sqrt{2/(\pi k!)} \cdot H_{k-1}(0) \\
&= \sqrt{2/(\pi k!)} \cdot (k-1)!/(2^{(k-1)/2} \cdot ((k-1)/2)!) \\
&= \sqrt{2/(\pi k)} \cdot \sqrt{(k-1)!/((k-1)/2)!} \\
&= \sqrt{2/(\pi k)} \cdot \sqrt{((k-1)/2)^{2k-1}} \\
&\approx \sqrt{2/(\pi k)} \cdot \sqrt{(2^{k-1} / \sqrt{\pi (k-1)/2})^{2k-1}} \\
&\approx \sqrt{2/(\pi k)} \cdot \sqrt{1/\sqrt{\pi (k-1)/2}} \\
&= \Theta(k^{-3/4})
\end{align*}

In the third to last step, we use the approximation \((\frac{m}{m/2}) \approx 2^{m}/\sqrt{\pi m/2}\).

With this estimate of \(\hat{f}(k)\), it follows that \(\sum_{r,r>k} \hat{f}^2(r) = \Theta(k^{-1/2})\). Then since \(\sum_{k=0}^\infty \hat{f}^2(r) = |f|^2_2 = 1\), we can conclude \(\sum_{r,r\leq k} \hat{f}^2(r) = 1 - \Theta(k^{-1/2})\) for large \(n\). This observation implies that the Fourier coefficients of the majority function are largely concentrated on the coefficients of low degree polynomials.

### 2 Influence

#### 2.1 Definition and Examples

**Definition 2** Let \(f \in L^2(\prod_{i=1}^n \mu_i)\). The influence of the \(i\)th variable is defined as follows:

\[
I_i(f) = \mathbb{E}_{\prod_{j \neq i} \mu_j}[\text{Var}_{\mu_i}[f]]
\]

**Example:** Let \(f : \{-1,1\}^n \rightarrow \{-1,1\}\) be a function, and let \(\{-1,1\}_{0}^n\) be our measure (i.e. \(\mu_i = \{-1,1\}_0\) for all \(i \in [n]\)). For \(x \in \{-1,1\}^n\), we define \(x^{\oplus i}\) to be the operation that flips the \(i\)th coordinate of \(x\) (i.e. \(x^{\oplus i}\) returns \(x' \in \{-1,1\}^n\), such that \(x'_i = -x_i\) and \(x'_j = x_j\) for all \(j \neq i\)). It is not difficult to show the following lemma:
Lemma 3 \(I_i(f) = P_{\prod_{j \neq i} \mu_j}[f(x) \neq f(x^{\oplus i})].\)

Proof: Consider all the variables of \(x \in \{-1, 1\}^n\) as fixed except the \(i\)th coordinate. Then

\[
\text{Var}_{\mu_i}[f(x)] = \begin{cases} 
1 & \text{if } f(x) \neq f(x^{\oplus i}) \\
0 & \text{if } f(x) = f(x^{\oplus i}) 
\end{cases}
\]

When we no longer assume \(x_j\) is fixed for \(j \neq i\), then \(\text{Var}_{\mu_i}[f(x)]\) can be thought of as an indicator random variable \(M_f\) that is 1 if \(f(x) \neq f(x^{\oplus i})\) and 0 otherwise. Then the proof is trivial as:

\[
I_i(f) = E_{\prod_{j \neq i} \mu_j}[\text{Var}_{\mu_i}[f]] = E_{\prod_{j \neq i} \mu_j}[M_f] = P_{\prod_{j \neq i} \mu_j}[f(x) \neq f(x^{\oplus i})].
\]

\(\square\)

Exercise 4 (1 point) Suppose \(f\) only attains values \(a\) and \(b\), and our measure is \(\{-1, 1\}_\theta^n\). Write \(I_i(f)\) in terms of \(a, b, \theta,\) and \(\Pr[f(x) \neq f(x^{\oplus i})].\)

Included below are some examples of influence. Unless otherwise stated, assume \(x_1, \ldots, x_n\) are drawn from measure \(\prod_{i=1}^n \mu_i.\)

Example: Let \(f(x_1, \ldots, x_n) = g(x_1)\). Then applying the definition of influence, we have:

\[
I_i(f) = \begin{cases} 
\text{Var}_{\mu_i}[g] & \text{if } i = 1 \\
0 & \text{if } i > 1 
\end{cases}
\]

Example: Let \(f(x_1, \ldots, x_n) = g_1(x_1) \cdot g_2(x_2) \cdot \ldots \cdot g_n(x_n)\). Then applying the definition of influence and simplifying, we have:

\[
I_i(f) = \text{Var}_{\mu_i}[g_i] \cdot \prod_{j: j \neq i} E_{\mu_j}[g_j^2]
\]

Example: Assuming measure \(\{-1, 1\}_0^n\) and \(f(x_1, \ldots, x_n) = \text{Maj}(x_1, \ldots, x_n)\), then applying the definition of influence and using Lemma 3, we have:
\[ I_i(f) = \mathbf{P}_{\prod_{j:j \neq i} \mu_j}[(\sum_{j:j \neq i} x_j) = 0] \approx \sqrt{2/(\pi n)} \cdot (1 + o(1)) \]

### 2.2 Influences and Expansions

Next, we prove a general theorem about influence. Consider a function \( f \in L^2(\prod_{i=1}^n \mu_i) \), where \( f(x_1, \ldots, x_n) = \sum_{S: S \subseteq [n]} f_S(x_1, \ldots, x_n) = \sum_J \hat{f}(J) U_J(x_1, \ldots, x_n) \). Although not explicitly stated, \( J \) is a multi-index of size \( n \), \( U_J \in U^1 \otimes U^2 \otimes \ldots \otimes U^n \), and \( U^l \) is assumed to be a standard basis of \( \mu_l \) for all \( l \in [n] \). (See previous lectures for more details).

**Theorem 5** \( I_i(f) = \sum_{S \subseteq [n]: i \in S} |f_S|^2 = \sum_{J: J \neq 0} \hat{f}^2(J) \)

**Proof:** To prove the theorem, we first show \( \sum_{S \subseteq [n]: i \in S} |f_S|^2 = \sum_{J: J \neq 0} \hat{f}^2(J) \). Note that by definition \( f_S = \sum_{J: J \in J_S} \hat{f}(J) \cdot U_J \), where \( J_S \) is the set of multi-indices \( J \) such that \( J_k \neq 0 \) for all \( k \in S \) and \( J_k = 0 \) for all \( k \notin S \). Then \( |f_S|^2 = \sum_{J: J \in J_S} \hat{f}(J)^2 \), and \( \sum_{S \subseteq [n]: i \in S} |f_S|^2 = \sum_{S \subseteq [n]: i \in S} \sum_{J: J \in J_S} \hat{f}^2(J) = \sum_{J: J \neq 0} \hat{f}^2(J) \).

Now we only need to prove \( I_i(f) = \sum_{J: J \neq 0} \hat{f}^2(J) \). To prove this consider all variables other than \( x_i \) as fixed, and let us compute \( \text{Var}_{\mu_i}[f] \):

\[
\text{Var}_{\mu_i}[f] = \text{Var}_{\mu_i}[\sum_{J: J_i = 0} \hat{f}(J) U_J(x_1, \ldots, x_n) + \sum_{J: J_i \neq 0} \hat{f}(J) U_J(x_1, \ldots, x_n)]
= \mathbf{E}_{\mu_i}[\sum_{J: J_i \neq 0} \hat{f}(J) U_J(x_1, \ldots, x_n)^2]
= \sum_{J, K: J_i \neq 0, K_i \neq 0} \hat{f}(J) \hat{f}(K) \cdot \mathbf{E}_{\mu_i}[U_J \cdot U_K]
\]

To get from the first equation to the second, we note that \( \sum_{J: J_i = 0} \hat{f}(J) U_J(x_1, \ldots, x_n) \) is constant when all variables except \( x_i \) are fixed and \( \mathbf{E}_{\mu_i}[\sum_{J: J_i \neq 0} \hat{f}(J) U_J(x_1, \ldots, x_n)] = 0 \) because we started with a standard basis. To get from the second line to the third, note that the fourier coefficients \( \hat{f}(J) \) are constant.

Finally, note that by orthogonality \( \mathbf{E}_{\prod_{j \in [n]} \mu_j}[U_J \cdot U_K] = 1 \) if \( K = J \) and \( \mathbf{E}_{\prod_{j \in [n]} \mu_j}[U_J \cdot U_K] = 0 \) otherwise. Now plugging in definitions, the theorem is easy to see:
\[ I_i(f) = \mathbb{E}_{\prod_{j \neq i} \mu_j} [\text{Var}_{\mu_i}[f]] = \mathbb{E}_{\prod_{j \neq i} \mu_j} \left[ \sum_{J,K: J_i \neq 0, K_i \neq 0} \hat{f}(J) \hat{f}(K) \cdot \mathbb{E}_{\mu_i}[U_J \cdot U_K] \right] = \sum_{J,K: J_i \neq 0, K_i \neq 0} \hat{f}(J) \hat{f}(K) \cdot \mathbb{E}_{\prod_{j \in [n]} \mu_j}[U_J \cdot U_K] = \sum_{J: J_i \neq 0} \hat{f}(J)^2. \quad \square \]