Lecture 32: Convolution

Warm-up Exercise

- Before we begin, let us start with some warm-up exercises
- Suppose $f : \{0,1\}^n \to \mathbb{R}$ is a real-valued function
- Think: $\widehat{f}(0)$ represents the mean of f, when the input of $f(\cdot)$ is chosen uniformly at random, i.e., $\widehat{f}(0) = \mathbb{E}\left[f(U)\right]$
- Think: $\sum_{S \neq 0} \widehat{f}(S)^2$ represents the variance of the random variable $f(U)$

- For two functions $f, g : \{0,1\}^n \to \{+1,-1\}$ the inner-product $\langle f, g \rangle$ measures the "similarity" of the two functions
- Think: If f, g agree at $(1-\varepsilon)$ fraction of inputs and they differ in the remaining ε fraction of the inputs then $\langle f, g \rangle = (1-2\varepsilon)$

Outline of today's lecture

- In today's lecture we shall study about an important property of the Fourier basis functions that makes them special, namely, additive homomorphism
- This additive homomorphism property shall help us prove interesting properties of an important technical tools in Fourier analysis called Convolution

Additive Homomorphism

- Let $f\colon \{0,1\}^n \to \mathbb{R}$ be an arbitrary function
- We say that f exhibits "additive homomorphism" if, for all $x, y \in \{0,1\}^n$, we have

$$f(x+y) = f(x) \cdot f(y)$$

- Observe that all the Fourier basis functions χ_S satisfy this additive homomorphism property

Discussion of "What is a Fourier Basis"

- Let $F = \{f_0, f_1, \ldots, f_{N-1}\}$ be a set of functions $\{0,1\}^n \to \mathbb{R}$ such that
  1. **Orthonormality.** The functions in F are orthonormal with respect to an "inner-product"
  2. **Symmetry.** For all $i \in \{0, \ldots, N\}$ and $x \in \{0,1\}^n$, we have $f_i(x) = f_x(i)$
  3. **Additive Homomorphism.** For all $x, y \in \{0,1\}^n$, and $i \in \{0, \ldots, N-1\}$, we have $f_i(x+y) = f_i(x) \cdot f_i(y)$
- Any analysis that we perform in this course extends to any basis F with the properties mentioned above
- Think: These properties imply that $f_0(x) = 1$, for all $x \in \{0,1\}^n$!

Intuition of the Convolution Operator

- Let X, Y be probability distributions over $\{0,1\}^n$
- Consider the following algorithm

1. Sample $x \sim X$ and sample $y \sim Y$
2. Output $z = x \oplus y$

- Note that the output of this algorithm is a distribution over the sample space $\{0,1\}^n$. Let us represent the output distribution of this algorithm by Z (also referred to as the distribution $X \oplus Y$)
- Question: What is the $\mathbb{P}[Z = z]$?
  - Note that x can be anything in $\{0,1\}^n$. However, given x and z, there is a unique $y = x \oplus z$ such that $x \oplus y = z$
  - So, we have

$$\mathbb{P}[Z = z] = \sum_{x \in \{0,1\}^n} \mathbb{P}[X = x]\,\mathbb{P}[Y = x \oplus z]$$

- The distribution Z is (a scaling) of the convolution of the distributions X and Y.

Convolution

1. Let $f, g\colon \{0,1\}^n \to \mathbb{R}$ be two functions
2. The convolution of f and g is the function $(f * g)\colon \{0,1\}^n \to \mathbb{R}$ defined as follows

$$(f * g)(x) = \frac{1}{N} \sum_{y \in \{0,1\}^n} f(y) \cdot g(x - y)$$

3. Note that if X and Y are two function representing probability distributions over $\{0,1\}^n$, then $N(X * Y)$ is the function corresponding to the probability distribution $X \oplus Y$
4. Note that the Convolution is a bilinear operator!

Fourier Transform of Convolution I

- Given two function f and g, we are interested in expressing the function $\widehat{(f*g)}$ using the functions $\widehat{f}$ and $\widehat{g}$
- We shall prove the following result

Lemma

For all functions $f, g\colon \{0,1\}^n \to \mathbb{R}$ and $S \in \{0,1\}^n$, we have

$$\widehat{(f*g)}(S) = \widehat{f}(S)\cdot\widehat{g}(S)$$

- We shall provide a direct proof for this result

$$\begin{aligned}\widehat{(f*g)}(S) &= \frac{1}{N}\sum_{x\in\{0,1\}^n}(f*g)(x)\chi_S(x)\\ &= \frac{1}{N}\sum_{x\in\{0,1\}^n}\frac{1}{N}\sum_{y\in\{0,1\}^n} f(y)g(x-y)\chi_S(x)\end{aligned}$$

Fourier Transform of Convolution II

$$= \frac{1}{N^2} \sum_{x \in \{0,1\}^n} \sum_{y \in \{0,1\}^n} f(y) g(x-y) \chi_S(y) \chi_S(x-y)$$

The final step above is a consequence of the additive homomorphism of the function χ_S. Let us continue with the simplification.

$$\begin{aligned}
\widehat{(f * g)}(S) &= \frac{1}{N^2} \sum_{x \in \{0,1\}^n} \sum_{y \in \{0,1\}^n} f(y) g(x-y) \chi_S(y) \chi_S(x-y) \\
&= \frac{1}{N^2} \sum_{y \in \{0,1\}^n} f(y) \chi_S(y) \sum_{r \in \{0,1\}^n} g(r) \chi_S(r) \\
&= \left(\frac{1}{N} \sum_{y \in \{0,1\}^n} f(y) \chi_S(y) \right) \left(\frac{1}{N} \sum_{r \in \{0,1\}^n} g(r) \chi_S(r) \right) \\
&= \widehat{f}(S) \cdot \widehat{g}(S)
\end{aligned}$$

Fourier Transform of Convolution III

- We can succinctly summarize this result as follows: $\widehat{(f * g)} = \widehat{f} \cdot \widehat{g}$
- Exercise: Express $\widehat{f \cdot g}$ using $\widehat{f}$ and $\widehat{g}$

Offset of a Function

- Let $f\colon \{0,1\}^n \to \mathbb{R}$
- Define $g\colon \{0,1\}^n \to \mathbb{R}$ as $g(x) = f(x-c)$, for some $c \in \{0,1\}^n$
- We are interested in expressing $\widehat{g}(S)$ using $\widehat{f}(S)$ and c

$$\begin{aligned}
\widehat{g}(S) &= \frac{1}{N}\sum_{x\in\{0,1\}^n} g(x)\chi_S(x) \\
&= \frac{1}{N}\sum_{x\in\{0,1\}^n} f(x-c)\chi_S(x-c)\chi_S(c) \\
&= \chi_S(c)\frac{1}{N}\sum_{x\in\{0,1\}^n} f(x-c)\chi_S(x-c) \\
&= \chi_c(S)\cdot\widehat{f}(S)
\end{aligned}$$

Offset of a Function II

- That is, we conclude that $\widehat{g} = \chi_c \cdot \widehat{f}$. Recall that $\chi_c(S) \in \{+1, -1\}$. So, $\chi_c(S) \cdot \widehat{f}(S)$ is either $\widehat{f}(S)$ or $-\widehat{f}(S)$.
- Intuition: If g is an offset of the function f, then $\widehat{g}$ is a "twisting/rotation" of the function $\widehat{f}$. So, by studying the magnitudes of the Fourier transform, we can study the function "independent of their offsets"
- Additional Perspective: In fact, this result also implies that g can be rewritten as $N(\widehat{\chi_c} * f)$