FOURIER ANALYSIS & METHODS 2020.03.09

JULIE ROWLETT

Abstract. Caveat Emptor! These are just informal lecture notes. Errors are inevitable! Read at your own risk! Also, this is by no means a substitute for the textbook, which is warmly recommended: *Fourier Analysis and Its Applications,* by Gerald B. Folland. He was the first math teacher I had at university, and he is awesome. A brilliant writer. So, why am I even doing this? Good question...

1. The generating function for the Hermite polynomials

This theory item is similar to the analogous result for the Bessel functions, but with a bit of a twist.

Theorem 1. *For any $x \in \mathbb{R}$ and $z \in \mathbb{C}$, the Hermite polynomials,*

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2},$$

satisfy

$$\sum_{n=0}^{\infty} H_n(x) \frac{z^n}{n!} = e^{2xz - z^2}.$$

Proof: The key idea with which to begin is to consider instead

$$e^{-(x-z)^2} = e^{-x^2+2xz-z^2}.$$

We consider the Taylor series expansion of this guy, with respect to z, viewing x as a parameter. By definition, the Taylor series expansion for

$$e^{-(x-z)^2} = \sum_{n \geq 0} a_n z^n,$$

where

$$a_n = \frac{1}{n!} \frac{d^n}{dz^n} e^{-(x-z)^2}, \quad \text{evaluated at } z = 0.$$

To compute these coefficients, we use the chain rule, introducing a new variable $u = x - z$. Then,

$$\frac{d}{dz} e^{-(x-z)^2} = -\frac{d}{du} e^{-u^2},$$

and more generally, each time we differentiate, we get a -1 popping out, so

$$\frac{d^n}{dz^n} e^{-(x-z)^2} = (-1)^n \frac{d^n}{du^n} e^{-u^2},$$

Hence, evaluating with $z = 0$, we have

$$a_n = \frac{1}{n!} (-1)^n \frac{d^n}{du^n} e^{-u^2}, \quad \text{evaluated at } u = x.$$

The reason it's evaluated at $u = x$ is because in our original expression we're expanding in a Taylor series around $z = 0$ and $z = 0 \iff u = x$ since $u = x - z$. Now, of course, we have

$$\frac{d^n}{du^n} e^{-u^2}, \quad \text{evaluated at } u = x = \frac{d^n}{dx^n} e^{-x^2}.$$

Hence, we have the Taylor series expansion

$$e^{-(x-z)^2} = e^{-x^2+2xz-z^2} = \sum_{n \geq 0} \frac{z^n}{n!} (-1)^n \frac{d^n}{dx^n} e^{-x^2}.$$

Now, we multiply both sides by e^{x^2} to obtain

$$e^{2xz-z^2} = e^{x^2} \sum_{n \geq 0} \frac{z^n}{n!} (-1)^n \frac{d^n}{dx^n} e^{-x^2}.$$

We can bring e^{x^2} inside because everything converges beautifully. Then, we have

$$e^{2xz-z^2} = \sum_{n \geq 0} \frac{z^n}{n!} e^{x^2} (-1)^n \frac{d^n}{dx^n} e^{-x^2}.$$

Voilà! The definition of the Hermite polynomials is staring us straight in the face! Hence, we have computed

$$e^{2xz-z^2} = \sum_{n \geq 0} \frac{z^n}{n!} H_n(x).$$

1.1. **Applications to best approximations.**

Exercise 1. *Find the polynomial of at most degree* 40 *which minimizes*

$$\int_{\mathbb{R}} |f(x) - P(x)|^2 e^{-x^2} dx,$$

where f is some function in the weighted $\mathcal{L}^2$ space on $\mathbb{R}$ with weight e^{-x^2}.

We know that the Hermite polynomials are an orthogonal basis for $\mathcal{L}^2$ on $\mathbb{R}$ with the weight function e^{-x^2}. We see that same weight function in the integral. Therefore, we can rely on the theory of the Hermite polynomials! Consequently, we define

$$c_n = \frac{\int_{\mathbb{R}} f(x) H_n(x) e^{-x^2} dx}{||H_n||^2},$$

where

$$||H_n||^2 = \int_{\mathbb{R}} H_n^2(x) e^{-x^2} dx = 2^n n! \sqrt{\pi}.$$

The polynomial we seek is:

$$P(x) = \sum_{n=0}^{40} c_n H_n(x).$$

Some variations on this theme are created by changing the weight function.

Exercise 2. *Find the polynomial of at most degree* 60 *which minimizes*

$$\int_{\mathbb{R}} |f(x) - P(x)|^2 e^{-2x^2} dx.$$

This is not the correct weight function for H_n. However, we can make it so. The correct weight function for $H_n(x)$ is e^{-x^2}. So, if the exponential has $2x^2 = (\sqrt{2}x)^2$, then we should change the variable in H_n as well. We will then have, via the substitution $t = \sqrt{2}x$,

$$\int_{\mathbb{R}} H_n(\sqrt{2}x) H_m(\sqrt{2}x) e^{-2x^2} dx = \int_{\mathbb{R}} H_n(t) H_m(t) e^{-t^2} \frac{dt}{\sqrt{2}} = 0, \quad n \neq m.$$

Moreover, the norm squared is now

$$\int_{\mathbb{R}} H_n^2(t) e^{-t^2} \frac{dt}{\sqrt{2}} = \frac{||H_n||^2}{\sqrt{2}} = \frac{2^n n! \sqrt{\pi}}{\sqrt{2}}.$$

Consequently, the functions $H_n(\sqrt{2}x)$ are an orthogonal basis for $\mathcal{L}^2$ on $\mathbb{R}$ with respect to the weight function e^{-2x^2}. We have computed the norms squared above. The coefficients are therefore

$$c_n = \frac{\int_{\mathbb{R}} f(x) H_n(\sqrt{2}x) e^{-2x^2} dx}{2^n n! \sqrt{\pi}/\sqrt{2}}.$$

The polynomial is

$$P(x) = \sum_{n=0}^{60} c_n H_n(\sqrt{2}x).$$

2. The Laguerre polynomials

The Laguerre polynomials come from understanding the quantum mechanics of the hydrogen atom. We shall not get into this[1]

Definition 2. *The Laguerre polynomials,*

$$L_n^\alpha(x) = \frac{x^{-\alpha} e^x}{n!} \frac{d^n}{dx^n} (x^{\alpha+n} e^{-x}).$$

We summarize their properties in the following

Theorem 3 (Properties of Laguerre polynomials)**.** *The Laguerre polynomials are an orthogonal basis for $\mathcal{L}^2$ on $(0, \infty)$ with the weight function $x^\alpha e^{-x}$. Their norms squared,*

$$||L_n^\alpha||^2 = \frac{\Gamma(n + \alpha + 1)}{n!}.$$

They satisfy the Laguerre equation

$$[x^{\alpha+1} e^{-x} (L_n^\alpha)']' + n x^\alpha e^{-x} L_n^\alpha = 0.$$

For $x > 0$ and $|z| < 1$,

$$\sum_{n=0}^{\infty} L_n^\alpha(x) z^n = \frac{e^{-xz/(1-z)}}{(1-z)^{\alpha+1}}.$$

[1] Alex Jones does get into it: `https://www.youtube.com/watch?v=i91XV07Vsc0`. Check out the Alex Jones Prison Planet `https://www.youtube.com/watch?v=kn_dHspHd8M`. Turns out that Alex Jones's crazy ranting makes for decent death metal vocals. The gay frogs and America first remix are pretty decent too.

Exercise 3. *Find the polynomial of at most degree 7 which minimizes*

$$\int_0^\infty |f(x) - P(x)|^2 x^\alpha e^{-x} dx.$$

Since the Laguerre polynomials are an orthogonal basis for $\mathcal{L}^2(0, \infty)$ with weight function $x^\alpha e^{-x}$, we define

$$c_n = \frac{\int_0^\infty f(x) L_n^\alpha(x) x^\alpha e^{-x} dx}{||L_n^\alpha||^2}.$$

The polynomial we seek is:

$$P(x) = \sum_{n=0}^{7} c_n L_n^\alpha(x).$$

3. Best approximation summary

Assume that based on theoretical considerations we know that a certain collection of functions

$$e^{inx}, \quad \cos, \quad \sin, \quad \text{orthogonal polynomials}, \quad \text{Bessel functions}, \quad \text{weird SLP functions},$$

are an orthogonal basis on a bounded interval. In the case of SLP functions, *do not forget the weight function* in case the weight function is not simply 1. Let us call such function ϕ_n. Then the best approximation to any f in $\mathcal{L}^2$ of the bounded interval under consideration is its Fourier-ϕ_n expansion, which is

$$\sum \frac{\langle f, \phi_n \rangle}{||\phi_n||^2} \phi_n(x).$$

Recall

$$\langle f, \phi_n \rangle = \int f(y) \overline{\phi_n(y)} w(y) dy, \quad \text{if the weight function is } w(y),$$

and

$$||\phi_n||^2 = \langle \phi_n, \phi_n \rangle.$$

One can also do best approximations using Hermite and Laguerre polynomials on $\mathbb{R}$ and $(0, \infty)$, respectively, with the weight functions e^{-x^2} and $x^\alpha e^{-x}$, respectively. It works in very much the same way in all these cases.

4. Distributions done the right way

The mathematical concept of a *distribution,* or, as they are sometimes called, *generalized function,* has been badly abused not only by physicists but also by mathematicians. You may have already heard about the so-called "delta function." It's not really a function. It's not a 'generalized function.' It has its very own terminology, and that is that it is a distribution. Now, distributions are not as mysterious and weird as the mystique in which they are often shrouded.

Distributions are functions which themselves take as input a function. A particularly nice class of distributions are the *tempered distributions.* These distributions take in a Schwarz class function and spit out a number.

Definition 4. *Assume that f is a smooth function on $\mathbb{R}$. Then, we say that $f \in \mathcal{S}$ if for all k and for all n,*

$$\lim_{|x| \to \infty} x^n f^{(k)}(x) = 0.$$

In other words, f and all its derivatives decay rapidly at $\pm\infty$. There are quite a few functions which satisfy this. For example, all smooth functions which live on a bounded interval (compactly supported) satisfy this property.

Exercise 4. *Show that if $f \in \mathcal{S}$ then all of the derivatives of f are in $\mathcal{S}$. Show that if $f \in \mathcal{S}$ then its Fourier transform is also in $\mathcal{S}$.*

Definition 5. *A* tempered distribution *is a function which maps $\mathcal{S}$ to $\mathbb{C}$, which satisfies the following conditions:*

- *It is linear, so for a distribution denoted by L, we have*
$$L(\alpha f + \beta g) = \alpha L(f) + \beta L(g),$$
for all f and g in $\mathcal{C}_c^\infty(\mathbb{R})$ and for all complex numbers α and β.
- *There is a non-negative integer N and a constant $C \geq 0$ such that for all $f \in \mathcal{S}$*
$$|L(f)| \leq C \sum_{j+k \leq N} \sup_{x \in \mathbb{R}} |x^j f^{(k)}(x)|.$$

Let's do an example. We define a distribution in the following way. For $f \in \mathcal{C}_c^\infty(\mathbb{R})$,
$$L(f) := f(0).$$
That is, the distribution takes in the function, f, and spits out the value of f at the point $0 \in \mathbb{R}$. This distribution satisfies for any f and g in $\mathcal{C}_c^\infty(\mathbb{R})$ and for any α and $\beta \in \mathbb{C}$,
$$L(\alpha f + \beta g) = \alpha L(f) + \beta L(g).$$
Moreover, we have the estimate that
$$|L(f)| \leq |f(0)| \leq \sup_{x \in \mathbb{R}} |f(x)|.$$
So the estimate required is satisfied with $N = 0$ and $C = 1$. This distribution has a name. It is called the *delta* distribution. It is usually written with the letter δ. It is nothing other than a function which takes a function as its input and spits out a number as its output.

Exercise 5. *Assume that $f \in \mathcal{C}_c^\infty(\mathbb{R})$. Show that by defining*
$$L_f(g) = \int_{\mathbb{R}} f(x) g(x) dx, \quad g \in \mathcal{C}_c^\infty(\mathbb{R}),$$
L_f is a tempered distribution.

In fact, the assumption that $f \in \mathcal{C}_c^\infty(\mathbb{R})$ wasn't even necessary. You can show that for $f \in \mathcal{L}^2(\mathbb{R})$ or $f \in \mathcal{L}^1(\mathbb{R})$, the distribution, L_f defined above (it takes in a function $g \in \mathcal{C}_c^\infty(\mathbb{R})$ and integrates the product with f over $\mathbb{R}$), is a distribution. So, here's something which is rather cool. The elements in $\mathcal{L}^2(\mathbb{R})$ and $\mathcal{L}^1(\mathbb{R})$ are in general *not* differentiable at all. However, the *distributions* we can make out of them *are* differentiable. Here's how we do that.

Definition 6. *The derivative of a tempered distribution, L is another tempered distribution, denoted by $L' \in \mathcal{D}(\mathbb{R})$, which is defined by*
$$L'(g) = -L(g'), \quad g \in \mathcal{S}.$$

To see that this definition makes sense, we think about the special case where $L = L_f$, and $f \in \mathcal{S}$. Then, we *can* take the derivative of f, and it is also an element of $\mathcal{S}$. So, we can define $L_{f'}$ in the analogous way. Let's write it down when it takes in $g \in \mathcal{S}$,

$$L_{f'}(g) = \int_{\mathbb{R}} f'(x)g(x)dx.$$

We can do integration by parts. The boundary terms vanish, so we get

$$L_{f'}(g) = \int_{\mathbb{R}} f'(x)g(x)dx = -\int_{\mathbb{R}} f(x)g'(x)dx.$$

So,

$$L_{f'}(g) = -L_f(g') = (L_f)'(g).$$

This is why it makes a lot of sense to define the derivative of a distribution in this way. For the heavyside function, we define

$$L_H, \quad L_H(g) = \int_0^\infty g(x)dx.$$

Then, we compute that

$$L_H'(g) = -L_H(g') = -\int_0^\infty g'(x)dx.$$

Due to the fact that $g \in \mathcal{S}$,

$$\lim_{x\to\infty} g(x) = 0.$$

Hence, we have

$$-\int_0^\infty g'(x)dx = -(0 - g(0)) = g(0) = \delta(g).$$

So, we see that the derivative of L_H is the δ distribution! Pretty neat!

In this way, distributions can solve differential equations! For example, we'd say that a distribution L satisfies the equation

$$L'' + \lambda L = 0$$

if, for every $g \in \mathcal{S}$ we have

$$L''(g) + \lambda L(g) = 0.$$

This turns out to be incredibly useful and important in the theory of partial differential equations. However, the way it usually works is that instead of actually finding a distribution which solves the PDE, one shows by abstract mathematics that there *exists* a distribution which solves the PDE. Then, one can use clever methods to show that the mere existence of a distribution solving the PDE, which is called a *weak solution,* actually implies that there exists a genuinely differentiable solution to the PDE. We don't want to get ahead of ourselves here, so conclude with one last exercise, which proves that you can differentiate distributions as many times as you like!

Exercise 6. *Use induction to show that you can differentiate a distribution as many times as you like, by defining*

$$L^{(k)}(g) := (-1)^k L(g^{(k)}).$$

In a similar way, we can define the Fourier transform of a distribution.

Definition 7. *Assume that L is a tempered distribution. The Fourier transform of L is the distribution, $\hat{L}$ which for $f \in \mathcal{S}$ acts as follows*

$$\hat{L}(f) := L(\hat{f}).$$

In this way, we can compute the Fourier transform of our favorite distribution, δ.

$$\hat{\delta}(f) := \delta(\hat{f}) = \hat{f}(0) = \int_{\mathbb{R}} f(x)dx.$$

So, we could think of the Fourier transform of δ as the distribution which acts by

$$\hat{\delta} : f \in \mathcal{S} \mapsto \int_{\mathbb{R}} f(x)dx.$$

On the other hand, by the FIT,

$$\delta(f) = f(0) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\xi)d\xi = \frac{1}{2\pi}\hat{\delta}(\hat{f}) = \frac{1}{2\pi}\hat{\hat{\delta}}(f).$$

So that's kind of cute. It says that

$$\delta = \frac{1}{2\pi}\hat{\hat{\delta}}.$$