Gaussian Sobolev Spaces and Polynomial Approximation

February 26, 2025

So far, we've worked with the Sobolev model on even 2-periodic functions.

$$\mathcal{M} = \{\text{even 2-periodic } m(x) : \rho(m) \le B\} \quad \text{for} \quad \rho(m) = \left\| \frac{d}{dx} m \right\|_{L_2} \tag{1}$$

Because curves on $[0,1]$ are in one-to-one correspondence with even periodic functions,[1] this is a natural model for functions on the unit interval $[0,1]$. Because $\|u\|_{L_2}^2 = \mathrm{E}\, u(X)^2$ for X uniformly distribution on $[0,1]$, our constraint requires $|m'(X)|$ to be small most of the time for X with this distribution.

In this homework, we'll see what happens when we impose the same constraint for X with another distribution: the standard normal distribution. This is a bit more natural if our covariates X_i tend to be near zero but aren't bounded. And it saves us a little trouble, as we'll be modeling functions on all of $\mathbb{R}$ directly, so we won't need to think about periodic extension. Here's the model.

$$\begin{aligned} \mathcal{M} &= \left\{ m : \lim_{x \to \pm\infty} m(x)^2 f(x) = 0 \ \text{ and } \ \rho(m) \le B \right\} \\ &\text{for} \quad \rho(m) = \sqrt{\int_{-\infty}^{+\infty} \left\{ \frac{d}{dx} m(x) \right\}^2 f(x)\ dx} \end{aligned} \tag{2}$$

where $f(x) = (1/\sqrt{2\pi})e^{-x^2/2}$ is the probability density for the standard normal distribution. This is a space of smooth functions m from $\mathbb{R} \to \mathbb{R}$ that don't grow too fast as x approaches $\pm\infty$.

We'll find a Fourier-series characterization of this model, show that the basis functions $\phi_0, \phi_1, \ldots$ involved in this characterization are the polynomials of order $0, 1, 2, \ldots$, and find a uniform bound on the error we get when we use polynomials of order $N-1$ to approximate the functions in this model. We can use this bound to implement an approximate least squares estimator using this model using finite-dimensional approximation, much like we did with the Sobolev model (1). But, least squares in polynomial models being fairly popular, it's also useful interpretively. E.g., when somebody says fitting a cubic polynomial model is 'good enough', we can look at the approximation error we get using cubic polynomials to approximate possible signals μ in the model $\mathcal{M}$ to think about whether we agree.

Throughout, we'll use the gaussian inner product and associated norm.

$$\langle u, v\rangle = \int_{-\infty}^{+\infty} u(x)v(x)f(x)dx \quad \text{and} \quad \|v\| = \sqrt{\langle v, v\rangle} \quad \text{for} \quad f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \tag{3}$$

If we prefer, we can think of these in terms of expectations involving a standard normal random variable.

$$\langle u, v\rangle = \mathrm{E}\, u(X)v(X) \quad \text{and} \quad \|v\| = \sqrt{\mathrm{E}\, v(X)^2} \quad \text{for} \quad X \sim N(0,1). \tag{4}$$

1 Fourier Series

Our first step is characterizing the adjoint of the differential operator $\frac{d}{dx}$. This isn't $-\frac{d}{dx}$, as it was when we were talking about the model (1), because we're working with a different vector space of functions with a different inner product.

Exercise 1 *Show that if we're using the gaussian inner product, the adjoint of the first derivative operator $Lv(x) = \frac{d}{dx}v(x)$ is $L^\star u(x) = xu(x) - \frac{d}{dx}u(x)$.*

Hint. *You can prove that the adjoint $L^\star$ satisfies $\langle L^\star u,\ v\rangle = \langle u,\ Lv\rangle$ using integration by parts on $\int\{f(x)u(x)\}v'(x)dx$. Why is it important that the functions and their derivatives don't grow too fast as $x \to \pm\infty$? How does that relate to the periodicity restriction we use with* (1)*?*

Using this, we can characterize the *self-adjoint operator* $L^\star L$. One advantage of this alternate definition is that it allows us to define a *family of models*, $\mathcal{M}^p$ for positive integers p, based on powers of the operator.[2]

$$\begin{aligned} \mathcal{M}^p = &\left\{ m \ :\ \lim_{x\to\pm\infty} m(x)^2\phi(x) = 0 \ \text{ and } \ \rho_p(m) \le B \right\} \\ &\text{for} \quad \rho_p(m) = \sqrt{\langle (L^\star L)^p\ m,\ m\rangle} \end{aligned} \tag{5}$$

Exercise 2 *What is the self-adjoint operator $L^\star L$ in this case? In terms of it, write an alternate definition of the seminorm ρ in* (2)*. Something like the definition of ρ_p in from* (5) *for $p = 1$, but more concrete. Then, if you like, write one for the seminorm ρ_2 from* (5) *too.*

Having this definition (5) also reduces Fourier series representation to the calculation of eigenvalues and eigenvectors of $L^\star L$. We'll show that these eigenvectors are polynomials and work out an efficient way of computing them.

We have at least two polynomial eigenvectors: the functions 1 and x are eigenvectors with corresponding eigenvalues 0 and 1. You can check. To find more, we'll start by working out a recursive formula.

Exercise 3 *Show that if ϕ is an eigenvector of $L^\star L$ with eigenvalue λ, then ϕ' is an eigenvector of $L^\star L$ with eigenvalue $\lambda - 1$.*

Hint. *Differentiate both sides of the identity $L^\star L\ \phi(x) = \lambda\phi(x)$.*

Let's interpret this. Suppose that $L^\star L$ has an eigenvector ϕ_j that is a polynomial of order j with corresponding eigenvalue is λ_j. Because the derivative of a jth order polynomial is a $(j-1)$th order polynomial, Exercise 3 would imply that we would also have an eigenvector ϕ_j' that's a polynomial of order $j-1$ with corresponding eigenvalue $\lambda_j - 1$. And therefore, applying the same reasoning to ϕ_j', that we'd also have an eigenvector ϕ_j'' with eigenvalue $\lambda_j - 2$, etc. This would be consistent with what we know already: we do have polynomial eigenvectors ϕ_0 and ϕ_1 of orders 0 and 1 and the corresponding eigenvalues $\lambda_0 = 0$ and $\lambda_1 = 1$ do satisfy the recursive formula $\lambda_{j-1} = \lambda_j - 1$. But none of that tells us that our premise is true, i.e. that $L^\star L$ actually does have an eigenvector ϕ_j that is a polynomial of order j.

What we do know is that it has eigenvectors ϕ_j that are polynomials of order j, with corresponding eigenvalues $\lambda_j = j$, for $j < 2$. If we want to show this is true for all j, we need to go 'up' from $j = 0$ to $j = 1$ to $j = 2$ etc. instead of going 'down' from j to $j-1$ to $j-2$ etc. To *increase* the order of a polynomial eigenvector $\phi_{j-1}(x)$, we can multiply it by x. The result, $x\phi_j(x)$, won't be an eigenvector. But it's a place to start. What we'll do is write out our $j+1$st eigenvector as a 'corrected version' of this guess, i.e. $\phi_{j+1}(x) = x\phi_j(x) - u(x)$, then work out what 'correction' $u(x)$ we need to use to make it an eigenvector.

Exercise 4 *Suppose that ϕ_{j-1} satisfies $L^\star L\phi_{j-1}(x) = (\lambda_j - 1)\phi_{j-1}(x)$. Find a function $u_j(x)$ so that $\phi_j(x) = x\phi_{j-1}(x) - u_j(x)$ satisfies $L^\star L\ \phi_j(x) = \lambda_j\phi_j(x)$.*

Tip. *You could solve for $u_j(x)$ so that $L^\star L\{x\phi_{j-1}(x) - u_j(x)\} = \lambda_j\{x\phi_j(x) - u_j(x)\}$, but that's working harder than you need to. Exercise 3 suggests an easier approach. We know that if ϕ_j is an eigenvector, then so is ϕ_j' and therefore $\frac{1}{c}\phi_j'$ for any constant c. So we can try to find $u_j(x)$ by solving the simpler differential equation $\phi_j' = \{x\phi_{j-1}(x) - u_j(x)\}' = c\phi_{j-1}(x)$, then* check *that it satisfies $L^\star L\{x\phi_{j-1}(x) - u_j(x)\} = \lambda_j\{x\phi_{j-1}(x) - u_j(x)\}$.*

Tip. *To find $u_j(x)$, compare $c\phi_{j-1}(x) = \{x\phi_{j-1}(x) - u_j(x)\}'$ to $(\lambda_j - 1)\phi_{j-1}(x) = L^\star L\phi_{j-1}(x) = x\phi_{j-1}'(x) - \phi_{j-1}''(x)$.*

What does this tell us about whether the eigenvectors of $L^\star L$ are polynomials? If ϕ_{j-1} and u_j are both polynomials, then so is $\phi_j = x\phi_{j-1} - u_j$. Using the 'base case' that $L^\star L$ has polynomial eigenvectors ϕ_j of order j and corresponding eigenvalues $\lambda_j = j$ for all natural numbers $j < K$ when $K = 2$, we can prove by induction on K that the same is true for all natural numbers j.

Because $L^\star L$ is a self-adjoint linear operator, we know that these polynomials ϕ_j are orthogonal. In fact, we know that $\phi_0 \ldots \phi_j$ is an orthogonal basis for the space of all polynomials of order j, i.e., for the vector space spanned by the vectors $1, x, x^2, \ldots, x^j$. Why? Because it's a set of $j+1$ orthogonal, and therefore linearly independent, vectors in a $j+1$ dimensional vector space. If we're willing to buy the premise that we can write every function in our model as a *power series*, i.e. that our model is spanned by the basis functions $1, x, x^2, \ldots$, it follows that it's also spanned by the eigenvectors $\phi_0, \phi_1, \phi_2, \ldots$.

That means we've got an orthogonal basis we can use to write a Fourier series representation of the model. However, when we do that, we will want to scale our eigenvectors so that $\|\phi_j\| = 1$.[1]

Exercise 5 *Suppose that ϕ_j with $\|\phi_j\| = 1$ is an eigenvector of $L^\star L$ with eigenvalue $\lambda_j = j$ for $j \geq 1$. Show that $\|\phi_j'\| = \sqrt{\lambda_j}$, so Exercise 3 implies that $\phi_{j-1} = \phi_j'/\sqrt{j}$ is an eigenvector of $L^\star L$ with corresponding eigenvalue $\lambda_{j-1} = j-1$ and norm $\|\phi_{j-1}\| = 1$.*

Hint. *$\|\phi'\|^2 = \langle L\phi',\ L\phi' \rangle = \langle L^\star L\phi,\ \phi \rangle$ for $L = \frac{d}{dx}$.*

We'll also want an efficient way of computing these eigenvectors. Something like the recursive formula in Exercise 4, which we can use to compute the eigenvector ϕ_j from eigenvectors $\phi_0 \ldots \phi_{j-1}$, but ideally without needing to compute any derivatives.

Exercise 6 *Suppose that, for all natural numbers j, ϕ_j with $\|\phi_j\| = 1$ is an eigenvector of $L^\star L$ with eigenvalue $\lambda_j = j$. Furthermore, suppose that this sequence satisfies $\phi_{j-1} = \phi_j'/\sqrt{j}$ for all $j \geq 1$. Prove that $\phi_j = x\phi_{j-1}/\sqrt{j} - \phi_{j-2}\sqrt{(j-1)/j}$ for $j \geq 2$.*

Hint. *$j\phi_j = x\phi_j' - \phi_j''$. How are ϕ_j' and ϕ_j'' related to ϕ_{j-1} and ϕ_{j-2}?*

2 Polynomial Approximation

Given the Fourier series representation of the model, we can use results from the Sobolev models homework to get a uniform bound on the error we get using polynomials of order $N-1$ to approximate functions in the model.

Exercise 7 *What is the maximal error of polynomial approximations of order $N-1$ for functions m in $\mathcal{M}^p$, i.e. what is*

$$\max_{m \in \mathcal{M}^p} \min_{a_0 \ldots a_{N-1}} \left\| m - \sum_{j=0}^{N-1} a_j x^j \right\| \quad \text{for} \quad \|u\| = \sqrt{\int_{-\infty}^{+\infty} u(x)^2 \phi(x) dx}?$$

If we want to ensure this is no more than ϵ, how large do we need to make N?

With reference to the Sobolev models homework, briefly explain how you know that you cannot do better than this using a different N-dimensional basis. A sentence should do.

[1] Often, e.g. in this Wikipedia article, the eigenvectors are 'normalized' so the coefficient on x^j in the polynomial ϕ_j is 1. This is a different thing.