1 Function Spaces, Norms, Metrics

1.1 Overview

Function spaces can typically be thought of as a linear space. For example, since polynomials of degree less than or equal to three can be uniquely identified by the coefficients to ax^3+bx^2+cx+d, we can think of polynomials of degree less than or equal to three as having the same structure as the linear space $\mathbb{R}^4$ (some work is needed to show that polynomials of degree less than or equal to three satisfies the axioms of linear algebra).

Collocation and FDM are entirely driven by the theory of Banach spaces. The class will be primarily about convergence theorems; for it is convergence theorems that give us guarantees on how the discrete technique approximates the true answer to a problem.

The structure of the ideas is that inner products → norms → metrics → convergence statements. The axioms for each of inner products, norms, and metrics are easily referenced online or in a textbook. The spaces we will talk about are correspondingly inner product spaces, norm spaces, and metric spaces.

If our inner product is into the complex number field, one significant generalization is that the inner product symmetry is now conjugated: $\sigma(u, v) = \overline{\sigma(v, u)}$.

1.2 Cauchy-Schwarz Inequality

If the norm space is complete we call it a Banach space. If the norm arises from an inner product then we refer to it as a Hilbert space (that is, all Hilbert spaces are Banach spaces, but not necessarily vice versa). The link between Banach and Hilbert spaces is primarily established with the Cauchy-Schwarz inequality:

$$\sigma(u, v) \leq \|u\|\|v\|$$

Where $\sigma(u, v)$ is the inner product and u, v are elements of the underlying space.

Proof.

$$\begin{aligned} 0 &\leq \sigma(au - bv, au - bv) \\ &\leq a^2\|u\|^2 + b^2\|v\|^2 - 2ab\sigma(u, v) \end{aligned}$$

Take $a = \|v\|$ and $b = \|u\|$.

$$\begin{aligned} 0 &\leq 2\|u\|^2\|v\|^2 - 2\|u\|\|v\|\sigma(u, v) \\ \sigma(u, v) &\leq \|u\|\|v\| \end{aligned}$$

□

From here, we produce a version of the triangle inequality:
Claim: $(\|u\| + \|v\|)^2 \geq \|u+v\|^2$

Proof.

$$\begin{aligned}(\|u\| + \|v\|)^2 &\geq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &\geq \|u\|^2 + 2\sigma(u,v) + \|v\|^2 \text{ from Cauchy-Schwarz} \\ &= \sigma(u,u) + \sigma(u,v) + \sigma(u,v) + \sigma(v,v) \\ &= \sigma(u,u+v) + \sigma(u+v,v) \\ &= \sigma(u+v,u+v) \\ &= \|u+v\|^2\end{aligned}$$

□

1.3 L^p Spaces

If a function f is in the function space L^2 (that is, $\int_{-\infty}^{\infty} |f(x)|^2 \; dx < \infty$) it is said to be "*square integrable*" or "*square summable*". Similarly, if f is in L^1 (that is, $\int_{-\infty}^{\infty} |f(x)| \; dx < \infty$) it is said to be "*absolutely integrable*" or "*absolutely summable*". If a function is in L^∞ this means that we are using the sup. norm under the integral sign; as a side note, all convergence in the L^∞ space is uniform. For example, consider the function $f(x) = \lim_{n\to\infty} x^n$ for $x \in [0,1]$. The statement "the limit does not converge uniformly to f" is the same as saying "f is not in L^∞".

1.4 Bounded Functions

If L is a linear transformation from space V to space W then we can define M to be the smallest number such that $\|Lu\| \leq M\|u\|$ for all $u \in V$. We can thus identify any linear function L with a functional norm M. L is uniformly continuous if and only if M exists. If V is finite dimensional, then all linear transformations L will have a norm. Finally, any bounded linear transformation will remain so under completion of the space. In other words: If V is a normed linear space then it is a dense subset of a Hilbert space (and thus permits a completion $\overline{V}$ which is a Hilbert space). L extends into $\overline{V}$ similarly; so $\overline{V}$ has the linear transformation $\overline{L}$ which is bounded as well due to L being bounded.

2 Differential Equations

Partial differential equations can be thought of as linear transformations on their inputs. For example:

$$\alpha\nabla^2 u + \beta\nabla u + \gamma u + \delta = 0$$

the linear operator in this case is $\delta \circ L \equiv \{L + \delta = 0\}$, an affine map where $L = \alpha\nabla^2 + \beta\nabla + \gamma$. In this specific example, there are names for the constant terms: α i sknown as the diffusion term, β the transport term, and γ the decay term. δ is sometimes referred to as the source term, or forcing term.

A generic second order PDE looks like

$$A\frac{\partial^2 u}{\partial x^2} + 2B\frac{\partial^2 u}{\partial x \partial y} + C\frac{\partial^2 u}{\partial y^2} + D\frac{\partial u}{\partial x} + E\frac{\partial u}{\partial y} + Fu + G = 0$$

There are several spaces here: there is the domain of the PDE, i.e. the domain $\mathcal{D}$ of the x and y arguments. Then there is the domain $\mathcal{F}$ of the functions that are operated on under the PDE. Keep these separate in your mind, because they are not at all the same.

Let $\gamma(s) = (x(s), y(s))$ be a rectifiable curve on $\mathcal{D}$. Thus $u \circ \gamma$ is a function $\mathbb{R} \to \mathbb{R}$. Applying the chain rule we have that the tangent vectors are $(u \circ \gamma)_x = \frac{\partial u}{\partial x}\frac{\partial x}{\partial s} = \frac{\partial u}{\partial x}x'$.

Repeating the argument for the other component we have the following

$$\begin{bmatrix} (u \circ \gamma)_x \\ (u \circ \gamma)_y \end{bmatrix} = \begin{bmatrix} \frac{\partial u}{\partial x}x' \\ \frac{\partial u}{\partial y}y' \end{bmatrix}$$

Repeating the differentiation once more to get the second partials (assuming smoothness for mixed partials being equivalent), we arrive at

$$\begin{bmatrix} \frac{\partial^2 u}{\partial x^2} & \frac{\partial^2 u}{\partial x \partial y} \\ \frac{\partial^2 u}{\partial x \partial y} & \frac{\partial^2 u}{\partial y^2} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

Now if we let $H = -\left(D\frac{\partial u}{\partial x} + E\frac{\partial u}{\partial y} + Fu + G\right)$ we can rewrite the entire PDE and establish the following linear system:

$$\begin{bmatrix} A & 2B & C \\ x' & y' & 0 \\ 0 & x' & y' \end{bmatrix} \begin{bmatrix} \frac{\partial^2 u}{\partial x^2} \\ \frac{\partial^2 u}{\partial x \partial y} \\ \frac{\partial^2 u}{\partial y^2} \end{bmatrix} = \begin{bmatrix} H \\ (u(\gamma))_{xx} \\ (u(\gamma))_{yy} \end{bmatrix}$$

What have we done? Given the general second-order PDE above, we can use the parameters $A, B, \ldots$ to deduce certain facts about the behavior of u on γ. For example, if the 3×3 matrix on the left is singular, then there must be a jump discontinuty in the solution (and vice versa). To get singularities of the solution, we notice that the determinant of the matrix can be rewritten as $A\left(\frac{y'}{x'}\right)^2 - 2B\left(\frac{y'}{x'}\right) + C = 0$ and then noticing that this is a quadratic equation on $\frac{y'}{x'}$. Using our familiar theory from pre-calculus, we can categorize the types of problems as follows:

If $B^2 - 4AC < 0$ then we have an *elliptical* problem; the determinant has no real roots and therefore the whole matrix is nonsingular.

If $B^2 - 4AC = 0$ then we have a *parabolic* problem: one real root, thus one possible singularity.

If $B^2 - 4AC > 0$ then we have a *hyperbolic* problem: two independent roots exist, and are thus our singularities.

3 Generalized Polynomial Interpolation

Polynomial interpolation can be accomplished through the use of a Vandermonde matrix. The degree n Vandermonde matrix for one variable is as follows (where x_i are all distinct points):

$$V = \begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{bmatrix}$$

Solving for the polynomial is equivalent to solving the following linear system: $V\vec{a} = \vec{y} = p(\vec{x})$ where $\vec{a}$ are the undetermined coefficients of the polynomial, and $\vec{y}$ is the vector containing, in order, $p(x_0), p(x_1), ...$

In order to solve this linear system it is necessary to show V is nonsingular. This can be done in a variety of ways, but the method we choose to favor is one of deduction.

3.1 Determinant Finding Example - Degree 2

Our task is to show that the matrix $V = \begin{bmatrix} 1 & x_0 & x_0^2 \\ 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \end{bmatrix}$ is nonsingular. Let $R_1 = (1, x_0, x_0^2)$, $R_2 = (1, x_1, x_1^2)$, $R_3 = (1, x_2, x_2^2)$ be the rows. We know that the determinant of V is invariant under the row operation $R_i \to R_i + cR_j$ where c is any scalar. Several facts:

- The determinant for this matrix must be an expression of degree 3 ("degree" here meaning the combined power of any monomial expression). The proof of this is done inductively on the bottom right corner of a general $n \times n$ Vandermonde matrix.
- Any factor of a row must be a factor of the determinant. This is because the row operation $R_i \to cR_i$ changes the determinant by a factor of c.

From this we can perform row operations to try to establish factors of rows, effectively "filling in" our knowledge of what the determinant must be. For instance:

$$\begin{aligned} R_1 &\to R_1 - R_2 \\ &= (0, x_0 - x_1, x_0^2 - x_1^2) \\ &= (x_0 - x_1) \cdot (0, 1, x_0 + x_1) \end{aligned}$$

Thus we conclude $(x_0 - x_1)$ must be a factor of the determinant. Working with the other rows, we arrive at $c(x_0 - x_1)(x_0 - x_2)(x_1 - x_2)$ as the expression for the determinant. The c is necessary because we have determined this polynomial exclusively through a factorizing technique on expressions of the form $(x_i - x_j)$— that is to say, we have no information at all about scalar factors. However, this can be resolved: in this instance, we know that if we evaluate the determinant "the usual way" (that is, expansion by minors) then the first term of the expression is $x_1 x_2^2 - ...$; looking at our factored form, we must conclude that in this situation $c = -1$. Hence the determinant of V can be written as $-(x_0 - x_1)(x_0 - x_2)(x_1 - x_2)$; c is nonzero, and if the points are all distinct then V must be nonsingular.

The procedure generalizes to arbitrary n.