MATH 5510 LECTURE NOTES
PRELIMINARIES:
LINEAR ALGEBRA AND L^2 FUNCTIONS

A. Kaffel

Topics covered

- Linear algebra (vectors and matrices)
  - Inner products, orthogonality and adjoints
  - Solving linear systems using an eigenvector basis
- Theory for L^2 functions
  - Definition (inner product, norm) and weighted L^2 spaces
  - Convergence for L^2 functions
  - Preview of the framework for solving DEs

Main goals

- Start building the main framework for solving linear DEs, using familiar linear algebra
- Recall why representation by an orthogonal basis of functions are useful (using Fourier series as an example), and what it means to 'converge' for such a series
- A preview of the main ideas for the next few weeks

1. The framework - for vectors and matrices

Here a 'review' of linear algebra introduces the framework that will be used to solve differential equations. The structure we review for **vectors** and **matrices** in the space $\mathbb{R}^n$ to solve **linear systems** $Ax = b$ will be adapted to solve differential equations.

1.1. The space.
In linear algebra, you studied the space $\mathbb{R}^n$. To review:

Definitions (linear algebra in $\mathbb{R}^n$):

- The space of n-dimensional real vectors:
$$\mathbb{R}^n = \{\mathbf{x} = (x_1, x_2, \cdots, x_n), \quad x_j \in \mathbb{R}\}$$
- We can define an **inner product** (the 'dot product') on this space by
$$\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^T \mathbf{y} = \sum_{j=1}^{n} x_j y_j.$$
- This also defines a norm (the 'ℓ^2 norm')
$$\|\mathbf{x}\|_2 := \Big(\sum_{j=1}^{n} |x_j|^2 \Big)^{1/2} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

- Two vectors $\mathbf{x}, \mathbf{y}$ are called **orthogonal** if

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0.$$

Note that two vectors are orthogonal, geometrically, if they are **perpendicular**. Some properties of the inner product and norm are worth highlighting:

- **Norm property:** A vector $\mathbf{x}$ has norm zero if and only if it is the zero vector:

$$\|\mathbf{x}\| = 0 \iff \mathbf{x} \equiv 0.$$

- **Linearity:** The inner product is **linear** in each argument, e.g. for the first argument,

$$\langle c_1 \mathbf{u} + c_2 \mathbf{v}, \mathbf{y} \rangle = c_1 \langle \mathbf{u}, \mathbf{y} \rangle + c_2 \langle \mathbf{v}, \mathbf{y} \rangle \text{ for all } c_1, c_2 \in \mathbb{R} \text{ and } \mathbf{u}, \mathbf{v} \in \mathbb{R}^n.$$

We can define **linear operators** L on $\mathbb{R}^n$, which are functions

$$L : \mathbb{R}^n \to \mathbb{R}^n$$

that are **linear** as defined above:

$$L(c_1 \mathbf{x} + c_2 \mathbf{y}) = c_1 L\mathbf{x} + c_2 L\mathbf{y} \text{ for all} c_1, c_2 \in \mathbb{R} \text{ and } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

In $\mathbb{R}^n$, linear operators are equivalent to $n \times n$ matrices:

$$L \text{ is a linear operator } \iff \text{ there is an } n \times n \text{ matrix } A \text{ s.t. } L\mathbf{x} = A\mathbf{x}.$$

Linear operators L can have eigenvalues and eigenvectors, i.e. $\lambda \in \mathbb{C}$ and $\phi \in \mathbb{R}^n$ such that

$$L\phi = \lambda \phi.$$

See the review document for further details.

1.2. **Adjoints.** Consider a linear operator L on $\mathbb{R}^n$.

Definition (Adjoint): The **adjoint** L^* of a linear operator L is the operator such that

$$\langle L\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, L^* \mathbf{y} \rangle \text{ for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

An operator is **self-adjoint** if it is equal to its adjoint ($L = L^*$).

In $\mathbb{R}^n$, the adjoint to L is the transpose: $L^* = L^T$ (viewing L as a matrix) since

$$\langle L\mathbf{x}, \mathbf{y} \rangle = (L\mathbf{x})^T \mathbf{y} = \mathbf{x}^T L^T \mathbf{y} = \langle \mathbf{x}, L^T \mathbf{y} \rangle$$

and an operator on $\mathbb{R}^n$ is self-adjoint if and only if the matrix is symmetric.

Self-adjoint operators on $\mathbb{R}^n$ (symmetric matrices) have an important structure (in fact, the same structure we'll exploit for functions in solving differential equations!). The main theorem goes as follows:

Theorem ('spectral' theorem for real matrices) If L is a self-adjoint operator on $\mathbb{R}^n$ (i.e. a real symmetric matrix), then:

- There are n eigenvectors $\phi_1, \cdots, \phi_n$ with distinct (real) eigenvalues
- The eigenvectors are an orthogonal basis for $\mathbb{R}^n$

That is, every $\mathbf{x} \in \mathbb{R}^n$ has a unique representation in the form

$$\mathbf{x} = \sum_{i=1}^{n} c_i \phi_i$$

for coefficients $c_j \in \mathbb{R}$ (the basis part) and (the orthogonality part)

$$\langle \phi_i, \phi_j \rangle = 0 \text{ for } i \neq j. \tag{1}$$

Adjoint eigenvalues: The eigenvalues of L^* are the same as for L since the equations for the eigenvalues of L and for L^* are the same:

$$0 = \det(L - \lambda I) = \det((L - \lambda I)^T) = \det(L^T - \lambda I).$$

The *eigenvalues* of L and L^* are the same, but the *eigenvectors* are different.

Bi-orthogonality: Let $\{\phi_j\}$ and $\{\psi_j\}$ be the eigenvectors for L and L^*. The two sets of eigenvectors are **bi-orthogonal**, which means that

$$\langle \phi_j, \psi_k \rangle = 0 \text{ if } \lambda_j \neq \lambda_k.$$

Contrast with 'self-orthogonality' of the single set $\{\phi_j\}$ for a self adjoint operator (eq. (1)). The proof is useful, as it illustrates a typical manipulation of the adjoint.

Proof. Let λ_j, ϕ_j and λ_k, ϕ_k be eigenvalue/vector pairs for L and L^*, respectively.

$$\begin{aligned}
\lambda_j \langle \phi_j, \psi_k \rangle &= \langle \lambda_j \phi_j, \psi_k \rangle \\
&= \langle L\phi_j, \psi_k \rangle \quad (\phi_j \text{ is an eigenvector of } L) \\
&= \langle \phi_j, L^* \psi_k \rangle \quad (\text{adjoint property}) \\
&= \lambda_k \langle \phi_j, \psi_k \rangle \quad (\psi_k \text{ is an eigenvector of } L^*)
\end{aligned}$$

Subtracting the RHS from the LHS, we get

$$(\lambda_j - \lambda_k)\langle \phi_j, \psi_k \rangle = 0 \implies \langle \phi_j, \psi_k \rangle = 0 \text{ if } \lambda_j \neq \lambda_k.$$

□

1.3. **Projection.** Let $\{\phi_i\}$ be the eigenvectors of a self-adjoint operator on $\mathbb{R}^n$. We know that for any $\mathbf{x} \in \mathbb{R}^n$,

$$\mathbf{x} = \sum_{i=1}^{n} c_i \phi_i$$

for some coefficients c_i. We find a coefficient c_j by **projecting** onto the j-th component. Take the inner product with ϕ_j to get

$$\langle \mathbf{x}, \phi_j \rangle = \sum_{i=1}^{n} c_i \langle \phi_i, \phi_j \rangle = c_j \langle \phi_j, \phi_j \rangle \implies \boxed{c_j = \frac{\langle \mathbf{x}, \phi_j \rangle}{\langle \phi_j, \phi_j \rangle}.}$$

The map $\mathbf{x} \to \langle \mathbf{x}, \phi_j \rangle$ projects $\mathbf{x}$ onto its ϕ_j component and returns the coefficient.[1]

[1] Note that the actual projection operator (returning a vector) is $P_j \mathbf{x} := \langle \mathbf{x}, \phi_j \rangle \phi_j$, which projects $\mathbf{x}$ onto the subspace spanned by ϕ_j. Since this space is one dimensional, it is simpler to just take the inner product and return the coefficient of ϕ_j.

Now suppose the $\{\phi_i\}$'s are eigenvectors of an operator L that form a basis for $\mathbb{R}^n$. Even if they are not orthogonal, we also have eigenvectors $\{\psi_i\}$'s for the adjoint. Then, by bi-orthogonality,

$$\langle \mathbf{x}, \psi_j \rangle = \sum_{i=1}^{n} c_i \langle \phi_i, \psi_j \rangle = c_j \langle \phi_j, \psi_j \rangle \implies \boxed{c_j = \frac{\langle \mathbf{x}, \psi_j \rangle}{\langle \phi_j, \psi_j \rangle}.}$$

assuming all the eigenvalues are distinct (see HW for an example).

1.4. **Solving linear equations.** Consider the linear equation

$$L\mathbf{x} = \mathbf{y}, \qquad L \text{ self-adjoint.}$$

Let $\{\phi_j\}$ be the eigenvectors. We will solve this equation using projection, to illustrate how it works (for a non-self-adjoint example, see HW).

Version 1 (Separation): First, we decompose $\mathbf{x}$ and $\mathbf{y}$ into the eigenvector basis:

$$\mathbf{x} = \sum_{j=1}^{n} x_j \phi_j, \quad \mathbf{y} = \sum_{j=1} y_j \phi_j.$$

Plugging into the equation and using that $L\phi_j = \lambda_j \phi_j$,

$$\sum_{j=1}^{n} \lambda_j x_j \phi_j = \sum_{j=1}^{n} y_j \phi_j.$$

Since the ϕ_j's are linearly independent (or you could project onto the j-th component), it must be that the LHS and RHS are equal term-by-term:

$$\lambda_j x_j = y_j \implies x_j = \frac{y_j}{\lambda_j} = \frac{1}{\lambda_j} \frac{\langle \mathbf{y}, \phi_j \rangle}{\langle \phi_j, \phi_j \rangle}.$$

Version 2 (adjoints): Start by projecting the equation onto ϕ_j:

$$\langle L\mathbf{x}, \phi_j \rangle = \langle \mathbf{y}, \phi_j \rangle, \qquad j = 1, \cdots, n.$$

These n-equations will determine $\mathbf{x}$. However, the $\mathbf{x}$ is stuck inside the operator.[2] To solve the problem, we use the adjoint property. In complete detail:

$$\begin{aligned} \langle \mathbf{y}, \phi_j \rangle &= \langle L\mathbf{x}, \phi_j \rangle \\ &= \langle \mathbf{x}, L^* \phi_j \rangle \quad \text{(def'n of the adjoint)} \\ &= \langle \mathbf{x}, L\phi_j \rangle \quad (L \text{ is self-adjoint}) \\ &= \langle \mathbf{x}, \lambda_j \phi_j \rangle \quad (\phi_j \text{ is an eigenvector}) \\ &= \lambda_j \langle \mathbf{x}, \phi_j \rangle \quad \text{(linearity of the inner product)} \end{aligned}$$

which gives us the information needed to get the coefficient of ϕ_j:

$$\langle \mathbf{x}, \phi_j \rangle = \frac{1}{\lambda_j} \langle \mathbf{y}, \phi_j \rangle \implies \mathbf{x} = \sum_{j=1}^{n} \frac{1}{\lambda_j} \frac{\langle \mathbf{y}, \phi_j \rangle}{\langle \phi_j, \phi_j \rangle} \phi_j.$$

[2] We can just expand $\mathbf{x}$ and use the eigenvector property as in Version 1 here; the point of the exercise is that for DE's, the trick in Method 2 will be relevant as opposed to using Version 1.

Highlight: There are two key ideas at play here. First, the basis allows use to decompose the vector into simple parts, and projection allows us to deal with simple equations for each component. Second, the eigenvector property means that the components do not interact - this **decouples** the structure into independent one-dimensional pieces.

If the operator is not self-adjoint, we can use bi-orthogonality instead (see HW).

2. Functional analysis review

Here we start to build up the function version of this framework, starting with familiar structure from Fourier series. Note that this will be done with the goal of practical use in mind, so some technical details will be elided.

2.1. **The space.** How do we generalize inner products in $\mathbb{R}^n$ to functions? Given a function $f(x)$ from an interval $[a,b]$ to $\mathbb{R}$, we define an **inner product** and the 'L^2 norm'

$$\langle f, g\rangle = \int_a^b f(x)g(x)\,dx, \qquad \|f\|_2 = \left(\int_a^b |f(x)|^2\,dx\right)^{1/2}. \tag{2}$$

The space in which our functions reside is the space

$$L^2[a,b] = \{f : [a,b] \to \mathbb{R} : \int_a^b |f(x)|^2\,dx < \infty\} \tag{3}$$

The space is pronounced 'L^2'; a function in this space is an 'L^2 function' or 'square-integrable' function). The inner product is well-defined and finite for any two functions f, g in $L^2[a,b]$.

Similarly, we may define the **weighted space** where the integral is weighted by a non-negative function $w(x)$:

$$L^2_w[a,b] = \{f : [a,b] \to \mathbb{R} : \int_a^b |f(x)|^2 w(x)\,dx < \infty\} \tag{4}$$

along with the inner product and norm

$$\langle f, g\rangle_w = \int_a^b f(x)g(x)w(x)\,dx, \qquad \|f\|_{2,w} = \left(\int_a^b |f(x)|^2 w(x)\,dx\right)^{1/2}. \tag{5}$$

Typically, the w is omitted when the space is implied (which is usually the case).

Example 1: L^2 functions **do not** need to be continuous or bounded. Consider

$$f(x) = \frac{1}{x^p}, \quad x \in [0,1].$$

for a real number p. Computing the square of the norm, we get

$$\|f\|_2^2 = \int_0^1 \frac{1}{x^{2p}}\,dx = \begin{cases} \infty & p \geq 1/2 \\ \dfrac{1}{2p-1}\dfrac{1}{x^{2p-1}}\Big|_{x=0}^{x=1} & p < 1/2 \end{cases}$$

so f is in L^2 if and only if $p < 1/2$ - that is, if it does not diverge to ∞ too quickly around $x = 0$. It is, however, allowed to be unbounded!

Example 2: It is not quite true that zero norm implies the function is zero. Precisely,

$$\|f\|_2 = 0 \iff f = 0 \text{ a.e.}$$

where 'a.e.' means 'almost everywhere'. For our purposes, a.e. can be interpreted to mean 'except at an isolated set of points'. For example, let $f \in L^2[0,1]$ be defined by

$$f = \begin{cases} 0 & x \neq 1/2 \\ 1 & x = 1/2 \end{cases}.$$

Since it is non-zero only at one point,

$$\int_0^1 |f|^2\,dx = 0.$$

However, f is not identically zero. We say that it is 'zero a.e.' or just $f \equiv 0$ for short, since it is effectively zero for most purposes.

2.2. **Convergence: definitions.** For a sequence of functions $\{f_n\}$ in $L^2[a,b]$ there are three important notions of convergence to a limiting function f. We'll need all of them to make sense of expressions like

$$f = \sum_{j=1}^{\infty} c_j \phi_j$$

where the sequence of functions are the 'partial sums'

$$f_n = \sum_{j=1}^{n} c_j \phi_j.$$

Convergence (definitions): Let f_n be the sequence of functions in $L^2[a,b]$.

- The sequence is said to converge **in norm** (or 'in L^2') to a limit f if
$$\|f_n - f\|_2 \to 0 \quad \text{as } n \to \infty \tag{6}$$
That is,
$$\int_a^b |f_n(x) - f(x)|^2\,dx \to 0 \text{ as } n \to \infty.$$
- The sequence converges **pointwise** to f if
$$f_n(x) \to f(x) \text{ as } n \to \infty \text{ for all } x \in [a,b]. \tag{7}$$
- The sequence converges **uniformly** to f if
$$\max_{x \in [a,b]} |f_n(x) - f(x)| \to 0 \text{ as } n \to \infty. \tag{8}$$

The only general relation we have is that uniform convergence is the strongest:

$$\text{uniform} \implies \text{pointwise, norm.}$$

It is not quite true that pointwise convergence implies norm convergence or vice versa.

Example (not uniform): Consider

$$f_n = x^n \text{ on } [0,1], \quad n = 1, 2, \cdots$$

which converges pointwise to the function

$$f(x) = \begin{cases} 0 & \text{if } 0 \le x < 1 \\ 1 & \text{if } x = 1 \end{cases}.$$

To see this, note that $f_n(1) = 1$ for all n (so $f_n(1) \to 1$) and

$$\lim_{n\to\infty} x^n = 0 \text{ if } 0 < x < 1.$$

It is easy to check that it converges in norm to f as well:

$$\int_0^1 |f_n - f|^2 \, dx = \int_0^1 x^{2n} \, dx = \frac{1}{2n+1} \to 0 \text{ as } n \to \infty$$

since the value at $x = 1$ does not contribute to the integral.

However, the convergence is **not uniform**. For $x < 1$,

$$|f_n(x) - f(x)| = x^n$$

and x^n can be made arbitrarily close to 1 by taking x close enough to 1 (for any n). Thus the maximum error is always 1. The interval where the error is near 1 shrinks in size as n increases, but the max. error never decreases.

2.3. **Orthogonal bases.** An **basis**[3] for this L^2 space is a set of functions $\{\phi_j\}$ (for $j = 1, 2, \cdots$) such that every f in $L^2([a,b])$ can be written as

$$f = \sum_{i=1}^{\infty} c_i \phi_i$$

for unique coefficients $\{c_j\}$. Here the 'equals' sign means that the series converges in norm to f; that is, the **partial sums**

$$f_n = \sum_{i=1}^{n} c_i \phi_i$$

converge to f in norm ($\|f_n - f\|_2 \to 0$). The set of functions is **orthogonal** if

$$\langle \phi_i, \phi_j \rangle = 0 \text{ for } i \neq j$$

Projection works just as it did for vectors. Let $\{\phi_j\}$ be an orthogonal basis for $L^2[a,b]$. If f is an L^2 function then there are coefficients c_i such that

$$f = \sum_{i=1}^{\infty} c_i \phi_i.$$

[3]Technically, this is a 'Hilbert basis', but the distinction is not relevant. We'll refer to it as a 'basis'.

Take the inner product of this with ϕ_j to get

$$\langle f, \phi_j \rangle = c_j \langle \phi_j, \phi_j \rangle \implies \boxed{c_j = \frac{\langle f, \phi_j \rangle}{\langle \phi_j, \phi_j \rangle}}.$$

Highlight: Each projection onto a basis function is one-dimensional as in $\mathbb{R}^n$, but there are an **infinite** number of basis functions in total. This is the main source of similarities (projections are nice) and differences (the infinite sum is not nice).

2.4. **Linear operators.** The **linear operator** of interest are operators that map a function on some domain to another such function. For example,

$$f \to \frac{df}{dx}$$

is a linear operator. Note that is not defined for all L^2 functions, but we'll get to the subtleties there later. Another example, is

$$Lf \to \int_a^x f(\xi)\, d\xi \text{ for } f \in L^2([a,b]).$$

Linear operators can have eigenvalues λ and **eigenfunctions** ϕ satisfying

$$L\phi = \lambda\phi.$$

Since we have an inner product, the operator L should have an adjoint L^*, and we might expect to have similar eigenfunctions and so on as in $\mathbb{R}^n$. This turns out to be more or less true, with some complications - our primary goal is to explore this generalization in depth.

3. Where are we going with all this?

Ignoring all the details, let's see the general idea. This example is intended to motivate some of the key questions. Consider a **partial** differential equation

$$\frac{\partial u}{\partial t} = Lu, \qquad x \in (a,b), \quad t > 0$$

for a function $u(x,t)$ and a linear operator L with only x-derivatives like

$$Lu = \frac{\partial^2 u}{\partial x^2}.$$

Suppose L and L^* have eigenfunctions ϕ_j and ψ_j, and that the ϕ_j's form a basis for $L^2(a,b)$. Then we have that

$$u(x,t) = \sum_{j=1}^{\infty} c_j(t)\phi_j(x)$$

for coefficients $c_j(t)$. Plug into the PDE:

$$\sum_{j=1}^{\infty} c_j'(t)\phi_j(x) = \sum_{j=1}^{\infty} \lambda_j c_j(t)\phi_j(x)$$

from which it (should) follow that

$$c_j' = \lambda_j c_j.$$

The orthogonal basis of eigenfunctions allows us to convert the **PDE** into a set of one dimensional **ODEs** for the coefficients.

The Big Picture: The sketch above and the contrast between linear algebra in $\mathbb{R}^n$ and functions in L^2 raises some key questions that will motivate the topics to come. There are some equivalences, and many questions left to answer:

vectors in $\mathbb{R}^n$(or $\mathbb{C}^n$)	$\Longleftrightarrow$	functions in $L^2[a,b]$
linear systems $Ax = b$	$\Longleftrightarrow$	differential equations?
$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i$	$\Longleftrightarrow$	$\langle f, g \rangle = \int_a^b f(x)g(x)\,dx$
$n \times n$ matrices	$\Longleftrightarrow$	linear operators L (e.g. $d/dx, ??$)
symmetric matrices	$\Longleftrightarrow$	self-adjoint operators?
spectral thm: $L = L^* \to \{\phi_j\}$	$\Longleftrightarrow$	????

- **What is the operator?** We want an orthogonal basis of eigenvectors for some linear operator L. This means identifying the right operator and understanding when it will do what we want.

- **When does an operator have nice eigenfunctions?** Self-adjoint matrices give a basis of eigenvectors. When is this true of operators in L^2? This is a key question that will require some work to sort out.

- **Infinite dimensions?** The basis for the function space is infinite dimensional - this has consequences that make the story more complicated than linear systems in $\mathbb{R}^n$. We'll also need to get around some problems that arise with infinite series - finite sums can be added and differentiated and so on freely, but infinite series take more care.

- **What are the eigenfunctions?** We will need to compute the eigenfunctions to have practical solutions, or at least understand their properties.