Notes on Linear Algebra 2

David L. Duncan

This is a continuation of the previous notes [1] I posted on Sakai. The purpose of these notes is to give you some insight into the various formulas present in the theory of Fourier series by giving you much of the underlying linear algebra. As a device to assist you in the reading, I have established the following convention: Any material not directly related to this purpose (e.g. material that has been included *only* to help illuminate concepts from linear algebra) has been indented and demoted to an 'Example' or a 'Remark'. In particular, this material is not necessary to understand the logical flow of the exposition, however, it may help with understanding the various definitions.

Furthermore, I will use the following conventions (mostly established in [1]):

- V, W will be used to denote arbitrary vector spaces, and u, v will denote vectors in these spaces.
- When dealing with the specific vector space $\mathbb{R}^n$ (or any of its subspaces) I will use the following notation to denote the vectors:
  - $\mathbf{x}, \mathbf{y}$; or
  - $\langle x_1, \ldots, x_n \rangle, \langle y_1, \ldots, y_n \rangle$; or
  - $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$

  depending on the context (i.e. depending on what is convenient).
- We will use $F(a, b)$ to denote the vector space of functions from $[a, b]$ to $\mathbb{R}$. Similarly, $F(-\infty, \infty)$ denotes the vector space of functions from $(-\infty, \infty)$ to $\mathbb{R}$. The vectors in these vector spaces (and their subspaces) will be denoted by f, g.
- Arbitrary scalars will be denoted by k (which is a real number).

1 Linear Transformations

In mathematics there are two central objects: sets (collections of objects), and functions (maps taking the objects in one set to the objects in another). When considering sets with additional structure, it is natural to consider functions that preserve that structure. For example, vector spaces are sets with the extra structure of vector addition and scalar multiplication. The functions that preserve this structure are called *linear transformation*, and these are what we will study in this section. Here is the formal definition:

Definition 1. *Let V, W be vector spaces and $T : V \to W$ a function. Then T is called a* **linear transformation** *if each of the following conditions are satisfied for all $u, v \in V$ and every $k \in \mathbb{R}$:*

1. $T(u + v) = T(u) + T(v)$

2. $T(ku) = kT(u)$

The next theorem gives two examples of linear transformations that exist for any vector space.

Theorem 2. *Let V be a vector space. Then the following are linear transformations:*

- *The* **zero function** $T_0 : V \to \{0\}$. *This is the function that sends v to 0 for every $v \in V$.*
- *The* **identity function** $T_{\text{Id}} : V \to V$. *This is the function that sends v to v for every $v \in V$.*

Proof. We will prove that the identity function is a linear transformation, and leave the proof of the analogous statement for the zero function as an exercise. Let $u, v \in V$ and $k \in \mathbb{R}$. Then $T_{\text{Id}}(u) = u$ and $T_{\text{Id}}(v) = v$, so

$$T_{\text{Id}}(u + v) = u + v = T_{\text{Id}}(u) + T_{\text{Id}}(v).$$

Similarly,

$$T_{\text{Id}}(ku) = ku = kT_{\text{Id}}(u).$$

This is exactly what we needed to demonstrate in order to show that T_{Id} is a linear transformation, so we are done.

□

Exercise 1. *Complete the proof by showing that the zero function T_0 is a linear transformation.*

1.1 Linear Transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$

In this subsection we will begin with some examples of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$, for specific values of n and m. Then we will classify, in a certain sense, all linear transformations of this type.

When dealing with linear transformation, it is convenient to write the vectors in $\mathbb{R}^n$ as $n \times 1$ matrices:

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Note that this means we will be writing elements of $\mathbb{R}^1$ as $[x]$.

1.1.1 The Case $n = m = 1$

Here we are studying linear transformations of the form

$$T : \mathbb{R}^1 \longrightarrow \mathbb{R}^1$$

Example 1. Define

$$T_2 : \mathbb{R}^1 \longrightarrow \mathbb{R}^1$$

by

$$T_2([x]) := 2\,[x]\,.$$

Then T_2 is a linear transformation because of the distributive property:

$$\begin{aligned} T_2([x]+[y]) &= 2\,([x]+[y]) \\ &= 2\,[x] + 2\,[y] \\ &= T_2\,([x]) + T_2\,([y])\,, \end{aligned}$$

and because of the commutative property of multiplication in $\mathbb{R}$:

$$\begin{aligned} T_2(k\,[x]) &= 2(k\,[x]) \\ &= k(2\,[x]) \\ &= kT_2([x]). \end{aligned}$$

As you might guess, there is nothing special about the number '2' in the previous example. In fact, given any real number $a \in \mathbb{R}$, we can define a function

$$T_a : \mathbb{R}^1 \longrightarrow \mathbb{R}^1$$

by the formula

$$T_a([x]) := a\,[x]\,,$$

and essentially the same argument as the one given for T_2 shows that T_a is a linear transformation. As we will see in Theorem 4 below, every linear transformation from $\mathbb{R}^1$ to itself is of the form T_a for some a.

Example 2. Here is an example of a function that is *not* a linear transformation: Consider the map $S : \mathbb{R}^1 \to \mathbb{R}^1$ defined by $S([x]) := \left[x^2\right]$. Then S is not a linear transformation because

$$S(2\,[1]) = S([2]) = 2^2 \neq 2 = 2S([1]).$$

1.1.2 General n, m

In this subsection we are studying linear transformations of the form

$$\mathbb{R}^n \longrightarrow \mathbb{R}^m.$$

The next example illustrates how an $m \times n$ matrix determines a *function* of this form. The theorem that follows shows that this function is in fact a linear transformation.

Example 3. Let M denote the matrix

$$M := \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 0 \end{bmatrix},$$

and define the function

$$T_M : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$$

by

$$T_M\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) := \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

For example,

$$T_M\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix}.$$

Then we will see in Theorem 3 below that T_M is a linear transformation.

Exercise 2. *Multiply the matrices in the definition of T_M to compute an explicit formula for*

$$T_M\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right).$$

More generally, suppose we have an $m \times n$ matrix M. Then M determines a function

$$T_M : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

defined by matrix multiplication

$$T_M\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) := M \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

This is always a linear transformation:

Theorem 3. *For any matrix M, the function $T_M : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation.*

Proof. Since the components are not relevant for this computation, I will use $\mathbf{x}, \mathbf{y}$ to denote vectors in $\mathbb{R}^n$, instead of $n \times 1$ matrices. So let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $k \in \mathbb{R}$. Then

$$\begin{aligned} T_M(\mathbf{x} + \mathbf{y}) &= M(\mathbf{x} + \mathbf{y}) \\ &= M\mathbf{x} + M\mathbf{y} \\ &= T_M(\mathbf{x}) + T_M(\mathbf{y}) \end{aligned}$$

$$\begin{aligned} T_M(k\mathbf{x}) &= M(k\mathbf{x}) \\ &= kM\mathbf{x} \\ &= kT_M(\mathbf{x}) \end{aligned}$$

These computations prove that T_M is a linear transformation. □

It turns out that every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be expressed as multiplication by an $m \times n$ matrix:

Theorem 4. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there is an $m \times n$ matrix M for which $T = T_M$.*

I will not prove this here, but if you are feeling ambitious you can try to prove it as an extra credit assignment.

1.2 Linear Transformations of Functions Spaces

We saw in the last subsection that linear transformations of $\mathbb{R}^n$ are, in the sense of Theorem 4, just matrices. In this section, we will see that if we study more general spaces than $\mathbb{R}^n$ (e.g. certain function spaces) then we get many more linear transformations than just matrices.[1]

Example 4. Let $a, b \in \mathbb{R}$ with $a < b$, and consider the vector space

$$C(a,b) := \{f \in F(a,b) \mid f \text{ is continuous}\}.$$

(See the previous set of notes [1] for more details on this vector space.) Given a vector $f \in C(a,b)$ (which is a function in this case), we can integrate it to obtain a real number:

$$\int_a^b f(x)\ dx \in \mathbb{R}.$$

This defines a map, which we denote by

$$T_{\int_a^b} : C(a,b) \longrightarrow \mathbb{R}.$$

Explicitly, it is given by

$$T_{\int_a^b}(f) := \int_a^b f(x)\ dx.$$

Then $T_{\int_a^b}$ is a linear transformation because

$$\begin{aligned} T_{\int_a^b}(f+g) &= \int_a^b f(x) + g(x)\ dx \\ &= \int_a^b f(x)\ dx + \int_a^b g(x)\ dx \\ &= T_{\int_a^b}(f) + T_{\int_a^b}(g) \end{aligned}$$

$$\begin{aligned} T_{\int_a^b}(kf) &= \int_a^b kf(x)\ dx \\ &= k\int_a^b f(x)\ dx \\ &= kT_{\int_a^b}(f) \end{aligned}$$

for $f, g \in C(a,b)$ and $k \in \mathbb{R}$.

[1] This is because these function space examples are *infinite dimensional*, whereas the $\mathbb{R}^n$ examples are all *finite dimensional*. However, we will not deal directly with the notion of *dimension* in these notes.

Example 5. Let c be a real number between a and b, and define

$$T_{\mathrm{ev}_c} : F(a,b) \longrightarrow \mathbb{R}$$

by

$$T_{\mathrm{ev}_c}(f) := f(c).$$

Exercise 3. *Show that T_{ev_c} is a linear transformation.*

Example 6. Define

$$C^1(a,b) := \{f \in F(a,b) \mid f, f' \text{ both exist and are continuous}\}.$$

Then if $f \in C^1(a,b)$ we have that

$$\frac{d}{dx} f = f' \in C(a,b)$$

is continuous, by definition of $C^1(a,b)$. So differentiation gives a map

$$T_{\frac{d}{dx}} : C^1(a,b) \longrightarrow C(a,b).$$

defined by

$$T_{\frac{d}{dx}}(f) := f'$$

Exercise 4. *Show that $T_{\frac{d}{dx}}$ is a linear transformation*

Example 7. The Laplace transform is a linear transformation defined on the vector space consisting of piecewise continuous functions that are exponentially bounded.

Remark. When dealing with a linear transformation on a function space, it is important that the domain be suitably chosen so that everything makes sense. For example, we couldn't choose the domain of $T_{\frac{d}{dx}}$ to be the vector space of continuous functions because there are continuous functions f that are not differentiable, so the notation f' doesn't even make sense. This is where many of these crazy function spaces are coming from: We want to study a linear transformation, so we have to find the correct domain, and sometimes the correct domain is a little weird.

1.2.1 The Transformation $T_{\frac{d^2}{dx^2}-\lambda \mathrm{Id}}$

We will see later that this example is of central interest to our discussion of Fourier series. Let $p > 0$ be a real number, and denote the space of twice differentiable $2p$-periodic functions by

$$\mathrm{PerFun}^2(2p) :=$$

$$\left\{ f \in F(-\infty, \infty) \;\middle|\; \begin{array}{c} f, f', f'' \text{ are defined on } [a,b], \\ f \text{ is } 2p\text{-periodic} \end{array} \right\}.$$

Exercise 5. *Show that $\mathit{PerFun}^2(2p)$ is a subspace of $F(-\infty, \infty)$.*

Let $\lambda \in \mathbb{R}$, and define

$$T_{\frac{d^2}{dx^2}-\lambda \mathrm{Id}} : \mathrm{PerFun}^2(2p) \longrightarrow F(-\infty, \infty).$$

by the formula

$$\left(T_{\frac{d^2}{dx^2}-\lambda \mathrm{Id}}\right)(f) := \frac{d^2}{dx^2} f - \lambda f.$$

Example 8. The function $f(x) = \sin(x) + \cos(3x) \in \mathrm{PerFun}^2(2\pi)$, and we have

$$\left(T_{\frac{d^2}{dx^2}-\lambda \mathrm{Id}}\right)(f) = -(1+\lambda)\sin(x) - (9+\lambda)\cos(3x).$$

Exercise 6. *Show that $T_{\frac{d^2}{dx^2}-\lambda \mathrm{Id}}$ is a linear transformation.*

Remark. Notice that

$$T_{\frac{d^2}{dx^2}-\lambda\mathrm{Id}} = T_{\frac{d^2}{dx^2}} - \lambda T_{\mathrm{Id}},$$

which is to say $T_{\frac{d^2}{dx^2}-\lambda\mathrm{Id}}$ is a linear combination of other linear transformations. More generally, any linear combination of linear transformations from V to W is again a linear transformation from V to W. in other words, the set of linear transformations from V to W is a vector space.

1.3 The Kernel

I will end the discussion of linear transformations with a few words about the *kernel* of a linear transformation. This is the set of vectors that are mapped to the zero vector under a linear transformation. We will see that this set is always a subspace, and so provides a very easy way of proving that a subset is a subspace, which we will illustrate through some examples.

Definition 5. *Let V, W be vector spaces and $T : V \to W$ a linear transformation. Then define the* **kernel**[2] *of T to be the set*

$$\ker(T) := \{v \in V \mid T(v) = 0\}.$$

The reason this definition is useful to us is the following theorem:

Theorem 6. *Let $T : V \to W$ be a linear transformation. Then $\ker(T)$ is a subspace of V.*

Proof. I will only prove half of what needs to be shown, and leave the other half for the reader as an exercise. Let $u, v \in \ker(T)$, so $T(u) = 0$ and $T(v) = 0$. We want to show $u + v \in \ker(T)$, which follows directly from this computation:

$$\begin{aligned} T(u+v) &= T(u) + T(v) \\ &= 0 + 0 \\ &= 0 \end{aligned}$$

□

Exercise 7. *Finish the proof by showing that if $u \in \ker(T)$ and $k \in \mathbb{R}$, then $ku \in \ker(T)$.*

Example 9. Let $T_2 : \mathbb{R}^1 \to \mathbb{R}^1$ be the linear transformation from Example 1, given by multiplication by 2. Then

[2] Some authors call this the **null space**.

$$\begin{aligned}
\ker(T_2) &= \left\{ [x] \in \mathbb{R}^1 \,\middle|\, T_2([x]) = [0] \right\} \\
&= \left\{ [x] \in \mathbb{R}^1 \,\middle|\, 2\,[x] = [0] \right\} \\
&= \left\{ [x] \in \mathbb{R}^1 \,\middle|\, [2x] = [0] \right\} \\
&= \{[0]\}
\end{aligned}$$

So the kernel of T_2 is the set containing only the zero vector. In [1] we saw that this is a subspace, thereby confirming Theorem 6.

Example 10. Consider the 2×2 matrix

$$N = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}.$$

By Theorem 3 multiplication by this matrix determines a linear transformation

$$T_N : \mathbb{R}^2 \longrightarrow \mathbb{R}^2.$$

Its kernel is given by

$$\begin{aligned}
\ker(T_N) &= \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \,\middle|\, \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \\
&= \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \,\middle|\, x_1 + 2x_2 = 0, \text{ and } 2x_1 + 4x_2 = 0 \right\} \\
&= \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \,\middle|\, x_1 + 2x_2 = 0 \right\} \\
&= \left\{ \begin{bmatrix} x_1 \\ -(1/2)x_1 \end{bmatrix} \in \mathbb{R}^2 \right\} \\
&= \left\{ x \begin{bmatrix} 1 \\ -1/2 \end{bmatrix} \,\middle|\, x \in \mathbb{R} \right\}
\end{aligned}$$

Note that this is a line (with slope -1/2) through the origin of $\mathbb{R}^2$, and recall from [1] that the non-trivial subspaces of $\mathbb{R}^2$ are precisely the lines through the origin. So we have again confirmed Theorem 6.

Exercise 8. *Let $T_M : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation from Section 1.1.2, given by multiplication by the matrix*

$$M = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Compute $\ker(T_M)$.

Example 11. Consider the linear operator given by differentiation

$$T_{\frac{d}{dx}} : C^1(a,b) \longrightarrow C(a,b).$$

It has an interesting kernel

$$\begin{aligned} \ker\left(\tfrac{d}{dx}\right) &= \left\{f \in C^1(a,b) \mid f' = 0\right\} \\ &= \left\{f \in C^1(a,b) \mid f \text{ is constant}\right\} \end{aligned}$$

So the kernel of $T_{\frac{d}{dx}}$ is the set of constant functions that are defined on $[a,b]$.

Exercise 9. *Show directly (that is, without appealing to Theorem 6) that the set of constant functions is a subspace of $F(a,b)$.*

1.3.1 $\ker\left(T_{\frac{d^2}{dx^2}-\lambda \mathrm{Id}}\right)$

In this subsection we compute the kernel of the linear transformation

$$T_{\frac{d^2}{dx^2}-\lambda \mathrm{Id}} : \mathrm{PerFun}^2(2p) \longrightarrow F(-\infty,\infty)$$

for all values of $\lambda \in \mathbb{R}$ (see Section 1.2.1).

So we want to understand the set

$$\ker\left(T_{\frac{d^2}{dx^2}-\lambda \mathrm{Id}}\right) = \left\{f \in \mathrm{PerFun}^2(2p) \mid f'' - \lambda f = 0\right\}$$

This is equivalent to finding the solutions of the following boundary value problem

$$f'' - \lambda f = 0, \qquad f(x) = f(x+2p)$$

Case 1. $\lambda = \alpha^2 > 0$: In this case the differential equation becomes

$$f'' - \alpha^2 f = 0.$$

The solutions of this are of the form

$$f(x) = A\cosh(\alpha x) + B\sinh(\alpha x)$$

The hyperbolic trig functions are not periodic, so the only way this function can satisfy the $2p$-periodicity condition is for $A = B = 0$. That is

$$\ker\left(T_{\frac{d^2}{dx^2}-\lambda \mathrm{Id}}\right) = \{0\} \qquad \text{for } \lambda > 0$$

Case 2. $\lambda = 0$: The differential equation becomes $f'' = 0$ and the solutions have the form

$$f(x) = Ax + B.$$

Again, the periodicity condition $f(x) = f(x + 2p)$ forces $A = 0$. However, the constant solution

$$f(x) = B$$

is $2p$-periodic. So we get

$$\ker\left(T_{\frac{d^2}{dx^2} - \lambda \mathrm{Id}}\right) = \{B \mid B \in \mathbb{R}\} \qquad \text{for } \lambda = 0.$$

Case 3. $\lambda = -\alpha^2 < 0$: The differential equation becomes $f'' + \alpha^2 f = 0$, and the solutions are

$$f(x) = A\cos(\alpha x) + B\sin(\alpha x).$$

Imposing the periodicity condition forces $\alpha = n\pi/p$ for some $n = 1, 2, 3, \ldots$, a positive integer. So in this case we have

$$f(x) = A\cos\left(\frac{n\pi}{p}x\right) + B\sin\left(\frac{n\pi}{p}x\right).$$

In conclusion, our case analysis yields the following result

$$\ker\left(T_{\frac{d^2}{dx^2} - \lambda \mathrm{Id}}\right) = \begin{cases} \left\{A\cos\left(\frac{n\pi}{p}x\right) + B\sin\left(\frac{n\pi}{p}x\right) \middle| A, B \in \mathbb{R}\right\} & \text{if } \lambda = (n\pi/p)^2 \\ & \text{for } n = 1, 2, 3, \ldots \\ \{B \mid B \in \mathbb{R}\} & \lambda = 0 \\ \{0\} & \text{otherwise} \end{cases} \tag{1}$$

1.3.2 Using the Kernel

Here we work through two examples of how you can prove that a subset is a sub*space* by realizing it as the kernel of a linear transformation.

Example 12. Consider the set

$$S := \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \,\middle|\, \begin{array}{r} x_1 + x_2 = 0 \\ x_2 - 3x_1 = 0 \end{array} \right\}.$$

By writing the defining equations in terms of matrices, we get

$$
\begin{aligned}
S &= \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \middle| \begin{bmatrix} 1 & 1 & 0 \\ -3 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}. \\
&= \ker(T_M),
\end{aligned}
$$

where

$$
M = \begin{bmatrix} 1 & 1 & 0 \\ -3 & 1 & 0 \end{bmatrix}.
$$

So, by Theorem 6, S is automatically a subspace of $\mathbb{R}^3$.

Example 13. Our second example involves the derivative linear transformation. Define the set

$$
S' := \left\{ f \in F(-\pi, \pi) \middle| f \text{ is continuous and } \int_{-\pi}^{\pi} f(x)\ dx = f(1) \right\}
$$

Suppose we want to show that S' is a vector space. We can do this by realizing it as the kernel of a linear transformation, as follows:

$$
\begin{aligned}
S' &= \left\{ f \in F(-\pi, \pi) \mid f \text{ is continuous and } \int_{-\pi}^{\pi} f(x)\ dx = f(1) \right\} \\
&= \left\{ f \in C(-\pi, \pi) \middle| \int_{-\pi}^{\pi} f(x)\ dx - f(1) = 0 \right\} \\
&= \ker\left(T_{\int_{-\pi}^{\pi}} - T_{\mathrm{ev}_1} \right)
\end{aligned}
$$

where

$$
T_{\int_{-\pi}^{\pi}} - T_{\mathrm{ev}_1} : C(-\pi, \pi) \longrightarrow \mathbb{R}
$$

is the linear transformation given by

$$
\left(T_{\int_{-\pi}^{\pi}} - T_{\mathrm{ev}_1} \right)(f) = \int_{-\pi}^{\pi} f(x)\ dx - f(1).
$$

Here we are using the fact that any linear combination of linear transformations is again a linear transformation. See the Remark at the end of Subsection 1.2.1

2 The Eigenvalue Problem

Let V be a vector space, and $T : V \to V$ a linear transformation. The **eigenvalue problem** for T is the following: Find all of the numbers $\lambda \in \mathbb{C}$ for which the equation

$$T(v) = \lambda v \tag{2}$$

has a non-zero solution $v \in V$. These values λ are called the **eigenvalues**, and any non-zero solution v of equation (2) is called an **eigenvector** for λ.

> **Remark.** If V is a function space (i.e. a vector space whose elements are functions), then the eigenvectors are often called **eigenfunctions**.

The textbook [2] has a pretty good discussion for the case when $V = \mathbb{R}^n$, in which case $T = T_M$ is given by matrix multiplication M. The reason I have included this section in these notes is to work out the following function space example.

2.1 The Eigenvalue Problem for $\frac{d^2}{dx^2}$

Let $p > 0$ and consider the linear transformation:

$$T_{\frac{d^2}{dx^2}} : \mathrm{PerFun}^2(2p) \longrightarrow F(-\infty, \infty).$$

defined by differentiating twice

$$\left(T_{\frac{d^2}{dx^2}}\right)(f) := \frac{d^2}{dx^2} f.$$

We want to solve the eigenvalue problem for this operator. So we want to find the values of $\lambda \in \mathbb{C}$ for which the following subset contains more than just the zero function:

$$\left\{ f \in \mathrm{PerFun}^2(2p) \,\middle|\, \frac{d^2}{dx^2} f = \lambda f \right\}. \tag{3}$$

Observe that

$$\left\{ f \in \mathrm{PerFun}^2(2p) \,\middle|\, \frac{d^2}{dx^2} f = \lambda f \right\} = \ker\left(T_{\frac{d^2}{dx^2} - \lambda \mathrm{Id}}\right),$$

the latter of which we studied in Section 1.3.1 for real values of λ.

Fact: It can be shown that $T_{\frac{d^2}{dx^2}}$ only has real eigenvalues (I will not prove this, but if you are daring you can try to prove it for extra credit).

So the analysis in Section 1.3.1 is indeed sufficient to solve the eigenvalue problem. The conclusion of this analysis is Equation (1), where we see that the eigenvalues of $T_{\frac{d^2}{dx^2}}$ are

$$\left\{0, (n\pi/p)^2 \mid n = 1, 2, 3, \ldots\right\},$$

and the eigenvectors associated to $(\pi n/p)^2$ have the form

$$A \cos\left(\frac{n\pi}{p}x\right) + B \sin\left(\frac{n\pi}{p}x\right)$$

where A, B are real numbers with $(A, B) \neq (0, 0)$, and the eigenvectors associated to 0 are the nonzero constant functions (recall that eigenvectors cannot be zero, by definition).

In particular, taking all coefficients to be 1 and separating the sines and cosines, we have that the set

$$\left\{1, \cos\left(\frac{\pi}{p}x\right), \cos\left(\frac{2\pi}{p}x\right), \ldots, \sin\left(\frac{\pi}{p}x\right), \sin\left(\frac{2\pi}{p}x\right), \ldots\right\} \tag{4}$$

represents the eigenvectors of $T_{\frac{d^2}{dx^2}}$ in the following sense: Any eigenvector with eigenvalue λ can be written as a linear combination of elements from the set (4). Note that the functions appearing in this linear combination will necessarily have the same eigenvalue λ.

Example 14. The function

$$\cos\left(\frac{13\pi}{p}x\right) - 2\sin\left(\frac{13\pi}{p}x\right)$$

is an eigenvector of $T_{\frac{d^2}{dx^2}}$ with eigenvalue $(13\pi/p)^2$. Clearly it can be written as a linear combination of $\cos\left(\frac{13\pi}{p}x\right)$ and $\sin\left(\frac{13\pi}{p}x\right)$, which are elements of the set (4).

Exercise 10. *Let $p > 0$ and define $C_0^2(0, p)$ to be the subspace of $F(0, p)$ consisting of functions f for which f, f', f'' are continuous, and $f(0) = f(p) = 0$. Solve the eigenvalue problem for the linear transformation*

$$T_{\frac{d^2}{dx^2}} : C_0^2(0, p) \longrightarrow C(0, p)$$

to show that the eigenvalues are $\left\{(n\pi/p)^2 \mid n = 1, 2, 3, \ldots\right\}$ and the eigenvectors associated to the eigenvalue $(n\pi/p)^2$ have the form

$$B \sin\left(\frac{n\pi}{p}x\right)$$

for $B \in \mathbb{R}$, and $B \neq 0$.

3 Inner Products

Let V be a vector space. An *inner product* on V is a mathematical device that allows one to measure lengths of vectors and angles between vectors. We will not define these in general, instead we will only discuss the two inner products that will arise in this class.

3.1 The Standard Inner Product on $\mathbb{R}^n$

The **standard inner product** on $\mathbb{R}^n$ is defined by the dot product

$$\langle x_1, \ldots, x_n \rangle \bullet \langle y_1, \ldots, y_n \rangle := x_1 y_1 + \ldots + x_n y_n.$$

In an attempt to simplify notation, when the exact components of the vector are not important, I will use $\mathbf{x}$ rather than $\langle x_1, \ldots, x_n \rangle$.

The key properties that the standard inner product satisfies are the following: For all $\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in \mathbb{R}^n$, and every $k \in \mathbb{R}$ we have

1. $\mathbf{x} \bullet \mathbf{y} = \mathbf{y} \bullet \mathbf{x}$;
2. $\mathbf{x} \bullet \mathbf{x} \geq 0$;
3. $\mathbf{x} \bullet \mathbf{x} = 0$ if and only if $\mathbf{x} = 0$;
4. $(\mathbf{x}_1 + \mathbf{x}_2) \bullet \mathbf{y} = \mathbf{x}_1 \bullet \mathbf{y} + \mathbf{x}_2 \bullet \mathbf{y}$;
5. $(k\mathbf{x}) \bullet \mathbf{y} = k\,(\mathbf{x} \bullet \mathbf{y})$.

Exercise 11. *Use Properties 1, 4, and 5 to show*

$$\mathbf{y} \bullet (\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{y} \bullet \mathbf{x}_1 + \mathbf{y} \bullet \mathbf{x}_2$$

and

$$\mathbf{x} \bullet (k\mathbf{y}) = k\,(\mathbf{x} \bullet \mathbf{y})\,.$$

As I mentioned, inner products can be used to measure lengths of vectors. We use the term *norm* to refer to the length of a vector. The **norm** of $\mathbf{x} \in \mathbb{R}^n$ is defined by

$$\|\mathbf{x}\| := \sqrt{\mathbf{x} \bullet \mathbf{x}}.$$

Notice that by Property 2 the norm of a vector is always a non-negative real number.

The reason you should think of this as the length of a vector is because it is just the Pythagorean Theorem in n dimensions. The next exercise illustrates this in the two dimensional case.

Exercise 12. *Plot the point $(3, 4)$ on the plane, and find the distance from it to the origin. Now find the norm of the vector $\langle 3, 4 \rangle$, and compare your answers.*

We can use the following formula from vector calculus to measure the angle θ between two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$:

$$\cos(\theta) = \frac{\mathbf{x} \bullet \mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}. \tag{5}$$

We say that two vectors $\mathbf{x}$ and $\mathbf{y}$ are **orthogonal** if $\mathbf{x} \bullet \mathbf{y} = 0$. In light of Formula (5), two vectors are orthogonal if and only if they are perpendicular (meet at a right angle). More generally, we say that a set of vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ in $\mathbb{R}^n$ is **orthogonal** if every pair is orthogonal: $\mathbf{x}_i \bullet \mathbf{x}_j = 0$ for all $i \neq j$.

Example 15. The set

$$\{\langle 2,0,0\rangle, \langle 0,1,0\rangle, \langle 0,0,1\rangle\}$$

is orthogonal, but the set

$$\{\langle 0,2,0\rangle, \langle 0,1,0\rangle, \langle 0,0,1\rangle\}$$

is not.

The utility of 'large' orthogonal sets is expressed in the following theorem:

Theorem 7. *If* $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subseteq \mathbb{R}^n$ *is an orthonormal set of size* n*, and* $\mathbf{y} \in \mathbb{R}^n$*, then*

$$\mathbf{y} = \frac{\mathbf{y} \bullet \mathbf{x}_1}{\|\mathbf{x}_1\|^2}\ \mathbf{x}_1 + \ldots + \frac{\mathbf{y} \bullet \mathbf{x}_n}{\|\mathbf{x}_n\|^2}\ \mathbf{x}_n.$$

This is a slightly more general statement than Theorem 7.7.1 in [2], which only applies to the case where $\mathbf{x}_1, \ldots, \mathbf{x}_n$ all have norm 1.

Example 16. Consider the case where $n = 1$. Then take $\mathbf{x}_1 \neq 0 \in \mathbb{R}^1$. The set $\{\mathbf{x}_1\}$ is obviously orthogonal and of size 1, so it satisfies the hypotheses of the theorem. On $\mathbb{R}^1$ the dot product is just usual multiplication: $\mathbf{x} \bullet \mathbf{y} = \mathbf{x}\mathbf{y}$, and so we obviously have

$$\mathbf{y} = \frac{\mathbf{y}\mathbf{x}_1}{\mathbf{x}_1\mathbf{x}_1}\mathbf{x}_1 = \frac{\mathbf{y} \bullet \mathbf{x}_1}{\|\mathbf{x}_1\|^2}\mathbf{x}_1$$

for all $\mathbf{y} \in \mathbb{R}^1$.

Exercise 13. *Prove Theorem 7. Feel free to refer to the proof of Theorem 7.7.1 in [2] for help.*

The real numbers

$$\frac{\mathbf{y} \bullet \mathbf{x}_1}{\|\mathbf{x}_1\|^2}, \ldots, \frac{\mathbf{y} \bullet \mathbf{x}_n}{\|\mathbf{x}_n\|^2}$$

in the conclusion of Theorem 7 are called the **coefficients of** y **in the basis** $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$.

An orthogonal set in $\mathbb{R}^n$ of size n is often called an **orthogonal basis**. On the other hand, an orthogonal set that satisfies the conclusion of Theorem 7 is called **complete**.[3] Theorem 7 says that for $\mathbb{R}^n$, the notions of orthogonal bases, and completeness coincide.

[3] The term 'complete' is used because such a set contains enough vectors to 'get to any other vector' in $\mathbb{R}^n$ by taking linear combinations – if one of the vectors were removed from this set then this would not be possible, and the set would be 'incomplete'.

Exercise 14. *Show that*

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

are each eigenvectors for the linear transformation determined by the matrix

$$\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}.$$

Show that

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

is an orthogonal basis, and find the coefficients for the vector

$$\begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

in this basis.

3.2 The L^2 Inner Product on $C(a,b)$

Let $a < b$ be real numbers, and $V \subseteq C(a,b)$ a subspace. For $f, g \in V$, define

$$(f,g) := \int_a^b f(x)g(x)\ dx.$$

This is an inner product on V, called the L^2 **inner product**.[4]

> **Remark.** We have restricted to vector spaces consisting of continuous functions defined on closed bounded intervals in order to guarantee that the integral in the definition of the inner product is well-defined. However, the results of this subsection carry over to many more vector spaces than just those appearing as subspaces of $C(a,b)$. For example, they hold for the vector space of piecewise continuous functions defined on $[a,b]$.

There is a similar inner product on $\mathrm{PerFun}^2(2p)$ defined by

$$(f,g) := \int_{-p}^{p} f(x)g(x)\ dx.$$

> **Remark.** The functions in $\mathrm{PerFun}^2(2p)$ are $2p$-periodic. This means that essentially all of their information is contained in any interval of length $2p$. In particular, there is nothing special about integrating over the interval from $-p$ to p in the definition of this inner product. Indeed, any interval of length $2p$ would yield exactly the same inner product. This is the content of the next exercise.

[4]'L' stands for Henri Lebesgue, an early 20th century mathematician, and the '2' is for the 2 that appears in the norm below.

Exercise 15. *Let $a \in \mathbb{R}$, and f a $2p$-periodic function defined on $\mathbb{R}$. Show that*

$$\int_{-p}^{p} f(x)\ dx = \int_{a}^{a+2p} f(x)\ dx.$$

Conclude that the inner product on $\mathrm{PerFun}^2(2p)$ *is equivalently defined by*

$$(f,g) = \int_{a}^{a+2p} f(x)g(x)\ dx.$$

These inner products satisfy the following key properties (compare with the standard inner product on $\mathbb{R}^n$): For all $f, f_1, f_2, g \in V$ and every $k \in \mathbb{R}$

1. $(f,g) = (g,f)$;
2. $(f,f) \geq 0$;
3. $(f,f) = 0$ if and only if $f = 0$;
4. $(f_1 + f_2, g) = (f_1, g) + (f_2, g)$;
5. $(kf, g) = k(f,g)$.

As in the case of $\mathbb{R}^n$, we define the **norm** of $f \in V$ by

$$\|f\| := \sqrt{(f,f)} = \sqrt{\int_a^b f^2\ dx}.$$

This should be thought of as the 'length' of f.

We say that a set $\{f_1, \ldots, f_n, \ldots\} \subseteq V$ is **orthogonal** if $(f_i, f_j) = 0$ for all $i \neq j$. (This set is allowed to be infinite.)

3.3 The Eigenvectors of $T_{\frac{d^2}{dx^2}}$ are Orthogonal

In Subsection 2.1 we found that each element of the set

$$\left\{1, \cos\left(\frac{\pi}{p}x\right), \cos\left(\frac{2\pi}{p}x\right), \ldots, \sin\left(\frac{\pi}{p}x\right), \sin\left(\frac{2\pi}{p}x\right), \ldots\right\}$$

is an eigenvector for the linear transformation

$$T_{\frac{d^2}{dx^2}} : \mathrm{PerFun}^2(2p) \longrightarrow F(-\infty, \infty).$$

We will show that this set is orthogonal in $\mathrm{PerFun}^2(2p)$.

Let $n, m > 0$ be integers. It suffices to show

1. $\left(\cos\left(\frac{m\pi}{p}x\right), \sin\left(\frac{n\pi}{p}x\right)\right) = 0$ for all integers n, m;
2. $\left(\cos\left(\frac{m\pi}{p}x\right), \cos\left(\frac{n\pi}{p}x\right)\right) = 0$ for all integers with $n \neq m$;

3. $\left(\sin\left(\frac{m\pi}{p}x\right), \sin\left(\frac{n\pi}{p}x\right)\right) = 0$ for all integers with $n \neq m$;

Proof of 1: Our goal is to prove

$$\int_{-p}^{p} \cos\left(\frac{m\pi}{p}x\right) \sin\left(\frac{n\pi}{p}x\right)\ dx = 0.$$

However, this is immediate because the integrand is an odd function and the integral is over a symmetric interval.

Proof of 2: Let $n \neq m$ be integers. To show

$$\int_{-p}^{p} \cos\left(\frac{m\pi}{p}x\right) \cos\left(\frac{n\pi}{p}x\right)\ dx = 0$$

we use the identity

$$\cos(A)\cos(B) = \frac{1}{2}\left(\cos(A+B) + \cos(A-B)\right):$$

$$\begin{aligned}
\int_{-p}^{p} \cos\left(\frac{m\pi}{p}x\right) \cos\left(\frac{n\pi}{p}x\right)\ dx &= \frac{1}{2}\int_{-p}^{p} \cos\left(\frac{(m+n)\pi}{p}x\right) + \cos\left(\frac{(m-n)\pi}{p}x\right)\ dx \\
&= \frac{1}{2}\left[\frac{p}{(m+n)\pi}\sin\left(\frac{(m+n)\pi}{p}x\right)\right. \\
&\qquad \left. + \frac{p}{(m-n)\pi}\sin\left(\frac{(m-n)\pi}{p}x\right)\right]\Bigg|_{-p}^{p} \\
&= 0.
\end{aligned}$$

Proof of 3: This is the same as in the proof of 2, except you should use the identity

$$\sin(A)\sin(B) = \frac{1}{2}\left(\cos(A-B) - \cos(A+B)\right).$$

I will leave this computation for the interested reader.

While we are at it, we may as well compute the (square of the) norm of each element in this orthogonal set. First the easy one:

$$\|1\|^2 = \int_{-p}^{p} 1\ dx = 2p$$

For the trig functions, take $m > 0$ an integer. Then, using the trig identities again we get

$$\left\|\cos\left(\frac{m\pi}{p}x\right)\right\|^2 = \left\|\sin\left(\frac{m\pi}{p}x\right)\right\|^2 = p.$$

Exercise 16. *Show that* $\left\{\sin\left(\frac{\pi}{p}x\right), \sin\left(\frac{2\pi}{p}x\right), \sin\left(\frac{3\pi}{p}x\right), \ldots\right\}$ *is an orthogonal set for* $C_0^2(0,p)$.

3.4 Completeness

It turns out the notion of 'orthogonal basis' does not carry over well to $C(a,b)$ or $\text{PerFun}^2(2p)$ (because they are infinite dimensional), however the notion of an orthogonal set being 'complete' does. Theorem 7 provides inspiration for the next definition:

Definition 8. *Let* V *be a subspace of* $C(a,b)$ *or* $\text{PerFun}^2(2p)$. *An orthogonal set* $\{f_1, \ldots, f_n, \ldots\} \subseteq V$ *is* **complete** *if for every* $g \in V$ *we have*

$$g = \frac{(g, f_1)}{\|f_1\|^2} f_1 + \frac{(g, f_2)}{\|f_2\|^2} f_2 + \ldots + \frac{(g, f_n)}{\|f_n\|^2} f_n + \ldots$$

> **Remark.** Since this sum may be infinite, care must be taken due to convergence/divergence issues. For the most part, we will ignore these issues in this class.

4 Wrap-up: Fourier Series

Let me recap what we have done: In Subsection 2.1 we studied the eigenvalue problem for the second derivative linear transformation

$$T_{\frac{d^2}{dx^2}} : \text{PerFun}^2(2p) \longrightarrow F(-\infty, \infty)$$

and found that the set

$$\left\{1, \cos\left(\frac{\pi}{p}x\right), \cos\left(\frac{2\pi}{p}x\right), \ldots, \sin\left(\frac{\pi}{p}x\right), \sin\left(\frac{2\pi}{p}x\right), \ldots\right\} \tag{6}$$

represents the eigenvectors for this operator. In Subsection 3.3 we showed that this set is orthogonal, and

$$\|1\|^2 = 2p, \qquad \left\|\cos\left(\frac{n\pi}{p}x\right)\right\|^2 = p, \qquad \left\|\sin\left(\frac{n\pi}{p}x\right)\right\|^2 = p. \tag{7}$$

The following theorem is of fundamental importance to the study of Fourier series. I will not include its proof, which relies on some technical details involving convergence that are beyond the scope of this course.

Theorem 9. *The set (6) is complete in* $\mathrm{PerFun}^2(2p)$.

Translating to a (possibly) more familiar language gives:

Corollary 10. *(Fourier Series) Let* $f \in \mathrm{PerFun}^2(2p)$. *Then*

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{p}x\right) + b_n \sin\left(\frac{n\pi}{p}x\right),$$

where

$$\begin{aligned} a_0 &= \frac{1}{p}\int_{-p}^{p} f(x)\ dx \\ a_n &= \frac{1}{p}\int_{-p}^{p} f(x)\cos\left(\frac{n\pi}{p}x\right)\ dx \\ b_n &= \frac{1}{p}\int_{-p}^{p} f(x)\sin\left(\frac{n\pi}{p}x\right)\ dx \end{aligned}$$

Proof. This is just unravelling the definitions. By Theorem 9 together with the definition of the term 'complete', we have

$$\begin{aligned} f(x) &= \tfrac{(f,1)}{\|1\|^2}1 + \tfrac{\left(f,\cos\left(\frac{\pi}{p}x\right)\right)}{\left\|\cos\left(\frac{\pi}{p}x\right)\right\|^2}\cos\left(\tfrac{\pi}{p}x\right) + \ldots + \tfrac{\left(f,\sin\left(\frac{\pi}{p}x\right)\right)}{\left\|\sin\left(\frac{\pi}{p}x\right)\right\|^2}\sin\left(\tfrac{\pi}{p}x\right) + \ldots \\ &= \tfrac{(f,1)}{\|1\|^2} + \sum_{n=1}^{\infty}\tfrac{\left(f,\cos\left(\frac{\pi}{p}x\right)\right)}{\left\|\cos\left(\frac{\pi}{p}x\right)\right\|^2}\cos\left(\tfrac{\pi}{p}x\right) + \tfrac{\left(f,\sin\left(\frac{\pi}{p}x\right)\right)}{\left\|\sin\left(\frac{\pi}{p}x\right)\right\|^2}\sin\left(\tfrac{\pi}{p}x\right) \end{aligned}$$

Now use (7):

$$\begin{aligned} f(x) &= \tfrac{(f,1)}{2p} + \sum_{n=1}^{\infty}\tfrac{\left(f,\cos\left(\frac{\pi}{p}x\right)\right)}{p}\cos\left(\tfrac{\pi}{p}x\right) + \tfrac{\left(f,\sin\left(\frac{\pi}{p}x\right)\right)}{p}\sin\left(\tfrac{\pi}{p}x\right) \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos\left(\tfrac{n\pi}{p}x\right) + b_n\sin\left(\tfrac{n\pi}{p}x\right), \end{aligned}$$

where

$$\begin{aligned} a_0 &= \frac{1}{p}(f, 1) \\ &= \frac{1}{p}\int_{-p}^{p} f(x)\ dx \end{aligned}$$

$$\begin{aligned} a_n &= \frac{1}{p}\left(f, \cos\left(\tfrac{n\pi}{p}x\right)\right) \\ &= \frac{1}{p}\int_{-p}^{p} f(x)\cos\left(\frac{n\pi}{p}x\right)\ dx \end{aligned}$$

$$\begin{aligned} b_n &= \frac{1}{p}\left(f, \sin\left(\tfrac{n\pi}{p}x\right)\right) \\ &= \frac{1}{p}\int_{-p}^{p} f(x)\sin\left(\frac{n\pi}{p}x\right)\ dx. \end{aligned}$$

This is exactly what we needed to show. □

Exercise 17. *(Sine Series) In Exercise 10 and 15 we saw that the set*

$$\left\{\sin\left(\frac{\pi}{p}x\right), \sin\left(\frac{2\pi}{p}x\right), \sin\left(\frac{3\pi}{p}x\right), \ldots\right\} \subseteq C_0^2(0,p)$$

consists of orthogonal eigenvectors for

$$T_{\frac{d^2}{dx^2}} : C_0^2(0,p) \to C(0,p).$$

It turns out that this set is complete in $C_0^2(0,p)$ *(you may assume this statement without proof). Use these facts to find an analogue of Corollary 10 for this set-up. (Hint: You will arrive at the sine-series for a function.)*

References

[1] D. Duncan. *Notes on Linear Algebra*. Unpublished. 2011.

[2] W. Wright, D. Zill. *Advanved Engineering Mathematics*. 4 ed. Jones and Bartlett. 2011.