Physics 342 Lecture 3

Welcome back to more linear algebra! By the end of last lecture we had identified some features of linear operators and how to access their information. For a matrix (i.e., linear operator) $\mathbb{M}$, we can find its element in the i^{th} row, j^{th} column, M_{ij}, by multiplying by vectors $\vec{v}_i$ and $\vec{v}_j$ such that

$$\vec{v}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \} i \qquad \vec{v}_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \} j \quad , \quad \vec{v}_i \cdot \vec{v}_j = \delta_{ij}$$

where

$$M_{ij} = \vec{v}_i^{\,T} \mathbb{M} \vec{v}_j .$$

We we able to genealize this for general linear operators that act on continuous functions. For some linear operator $\hat{\mathcal{O}}$, such as the derivative $\partial/\partial x$, its "$(i,j)^{th}$" element is defined by the integral:

$$\mathcal{O}_{ij} = \int_{x_0}^{x_1} dx\, f_i(x)\, \hat{\mathcal{O}}\, f_j(x) , \quad \text{where the operator is defined on the domain } x \in [x_0, x_1],$$

and the functions $f_i(x)$, $f_j(x)$ are elements of L^2-norm:

$$1 = \int_{x_0}^{x_1} dx \left(f_i(x)\right)^2 = \int_{x_0}^{x_1} dx \left(f_j(x)\right)^2 , \quad \delta_{ij} = \int_{x_0}^{x_1} dx\, f_i(x) f_j(x) .$$

Before continuing, we need to address one more aspect of this method for identification of matrix elements, $\mathcal{O}_{ij}$. To know everything about a linear operator $\hat{\mathcal{O}}$, we need to know <u>all</u> of its matrix elements $\mathcal{O}_{ij}$, where i and j run over <u>all</u> possible values. For example, if we consider a 2x2 matrix $\mathbb{M}$, then i and j run over 1,2:

$i, j \in \{1,2\}$. For this matrix, we need to know the four matrix elements: M_{11}, M_{12}, M_{21}, and M_{22} to completely specify the matrix. Requiring complete knowledge of a linear operator enforces properties on the set of vectors $\{\vec{v}_i\}$ that are used to define and isolate the matrix elements.

One constraint on the vectors $\{\vec{v}_i\}$ is simple. If $\mathbb{M}$ is an $N \times N$ matrix, then to identify all $N \times N$ matrix elements, there need to be (at least) N vectors in the set $\{\vec{v}_i\}$. We have already said that all vectors in this set are normalized:

$\vec{v}_i \cdot \vec{v}_i = 1$, and orthogonal $\vec{v}_i \cdot \vec{v}_j = 0$, for $i \neq j$.

An $N \times N$ matrix acts on N dimensional vectors; that is, it manipulates some N dimensional space. If we have identified N orthonormal N-dimensional vectors $\{\vec{v}_i\}_{i=1}^{N}$, then these vectors necessarily span the whole space of N dimensional vectors. That is, some N dimensional vector $\vec{u}$ can be expressed as a linear combination of the $\vec{v}_i$s:

$$\vec{u} = \sum_{i=1}^{N} \alpha_i \vec{v}_i, \quad \text{where } \alpha_i \text{ is some number.}$$

If these properties are true, then we state that the set of vectors $\{\vec{v}_i\}_{i=1}^{N}$ are orthonormal and <u>complete</u>. Given a complete set of orthonormal vectors, we can find out <u>anything</u> we want about the matrix $\mathbb{M}$.

However, as hinted at in the last lecture, individual matrix elements are not well-defined. For example, let's take our 2×2 matrix $\mathbb{M}$. To identify its matrix elements M_{ij}, we need to specify a complete basis of orthonormal vectors. This specification is not unique. For example, let's take

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{so clearly } \vec{v}_1 \cdot \vec{v}_1 = \vec{v}_2 \cdot \vec{v}_2 = 1$$

and $\vec{v}_1 \cdot \vec{v}_2 = 0$. Then, we find, for example that

$M_{12} = \vec{v}_1^{\,T} \mathbb{M} \vec{v}_2$. In this basis, we can express the matrix $\mathbb{M}$ as

$$\mathbb{M} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$$

Your friend, however, defines as their complete basis:

$$\vec{u}_1 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}, \quad \text{for some (any!) angle } \theta.$$

Note that $\vec{u}_1 \cdot \vec{u}_1 = (\cos\theta \ \ \sin\theta) \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = \cos^2\theta + \sin^2\theta = 1$

and similarly, $\vec{u}_2 \cdot \vec{u}_2 = 1$. Also, you can show that $\vec{u}_1 \cdot \vec{u}_2 = 0$, and so we have two, orthonormal vectors $\vec{u}_1, \vec{u}_2$ which therefore define another complete basis in which to study $\mathbb{M}$. However in this basis, the 12 element of $\mathbb{M}$ is

$$\vec{u}_1^{\,T} \mathbb{M} \vec{u}_2 = (\cos\theta \ \ \sin\theta) \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} =$$

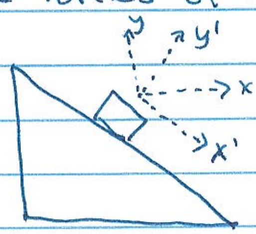
$$= \begin{pmatrix} M_{11}\cos\theta + M_{21}\sin\theta & M_{12}\cos\theta + M_{22}\sin\theta \end{pmatrix} \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

$$= -M_{11}\cos\theta\sin\theta - M_{21}\sin^2\theta + M_{12}\cos^2\theta + M_{22}\cos\theta\sin\theta.$$

This is clearly different than what we found using the $\vec{v}$-vector basis! Who is right and who is wrong?

Neither are wrong: they have just expressed the elements of $\mathbb{M}$ in a different basis, and so find a different explicit representation of $\mathbb{M}$. However, the matrix $\mathbb{M}$ is independent of any particular basis, as the choice of basis was something that we did arbitrarily. $\mathbb{M}$ doesn't change to our whims; it is what it is.

As an example of something familiar that you've seen that manifests this property, consider analyzing the forces of a block on a ramp:



The first step in such a problem, before even drawing a free-body diagram, is to determine the coordinate system you will use. Do you align the x-axis with the ramp, or do you align it horizontally? Or, do you point it some other direction entirely? You are allowed to orient your coordinates in <u>any</u> way you want: the physics is independent of your particular, idiosyncratic, choice. A pithy way to express this is that life (= Nature) cannot imitate art (= our coordinate system). Whatever coordinates we use, we must find the same acceleration of the block, for example.

Going back to our matrix example, what is the analogy of the coordinate-basis invariant physics? What properties of a matrix $\mathbb{M}$ are independent of basis in which we choose to express its matrix elements? Well if $\mathbb{M}$ is an $N \times N$ matrix, then clearly, and perhaps trivially (!), the dimension of $\mathbb{M}$, N, is basis-independent. Every orthonormal, complete basis for $\mathbb{M}$ must consist of N vectors.

Further, even though in general there are a continuous infinity of possible orthonormal, complete bases in which to express $\mathbb{M}$, there is only one for which $\mathbb{M}$ is expressed as a diagonal matrix, with only non-zero entries on the diagonal. Let's see why this basis is unique. Let's assume the basis in which $\mathbb{M}$ is diagonal is $\{\vec{t}_i\}_{i=1}^{N}$. "Diagonal" means that the matrix elements M_{ij} are

$$M_{ij} = \lambda_i \delta_{ij} = \vec{t}_i^{\,T} \mathbb{M} \vec{t}_j, \text{ for some number } \lambda_i.$$

However, if $\mathbb{M}$ is diagonal in this basis:

$$\mathbb{M} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{pmatrix},$$ then the matrix product is simply:

$$\mathbb{M}\vec{t}_j = \lambda_j \vec{t}_j,$$ so that the matrix element M_{ij} is

$$\vec{t}_i^{\,T} \mathbb{M} \vec{t}_j = \lambda_j \vec{t}_i^{\,T} \vec{t}_j = \lambda_j \vec{t}_i \cdot \vec{t}_j = \lambda_j \delta_{ij}.$$

On the right, we used orthonormality of the basis vectors.

The collection of λ_i numbers on the diagonal are of course nothing but the eigenvalues of the matrix $\mathbb{M}$ which can be found by solving the characteristic polynomial equation:

$$\det(\mathbb{M} - \lambda \mathbb{I}) = 0$$

where $\mathbb{I}$ is the $N \times N$ identity matrix. For such a matrix the determinant of $\mathbb{M} - \lambda \mathbb{I}$ is in general an N^{th} degree polynomial equation for λ, for which there are ingeneral N solutions, corresponding to the $\lambda_1, \lambda_2, \ldots, \lambda_N$ values.

This "diagonal" basis for $\mathbb{M}$ is analogous to the proper framein special relativity. Though energies, momenta, lengths, times, etc., are different in every inertial reference frame, every frame can agree what would be seen in the special frame which is at rest with respect to the object of interest. Diagonalizing a matrix $\mathbb{M}$ (that is, finding its eigenvalues) is like boosting to this proper frame.

Thus, the only basis-independent quantities of a linear operator $\mathbb{M}$ are its dimension and collection of eigenvalues, $\{\lambda_i\}$. You know the eigenvalues of a matrix, you know everything about that matrix.

Our discussion in most of this lecture has been limited to familiar finite, discrete matrices, but the structure applies to all linear operators, $\hat{O}$. In the last few minutes, let's consider its consequences for our good friend the derivative operator $\partial/\partial x$. Extrapolating from our matrix analysis, the eigenvalue equation for the derivative would be:

$$\frac{\partial}{\partial x} f_\lambda(x) = \lambda f_\lambda(x), \text{ for some eigenvalue } \lambda \text{ and "eigenfunction" } f_\lambda(x).$$

This eigenvalue equation, for some λ, is a simple differential equation for which the solution is just an exponential function:

$$f_\lambda(x) = e^{\lambda x}.$$

In our analogy, these $f_\lambda(x)$ functions should themselves be a complete, orthonormal basis. However, this form may suggest some problems with these requirements. First, the position x can range over all values: $x \in (-\infty, \infty)$, and for real λ, the function $e^{\lambda x}$ is poorly-behaved as $x \to \pm\infty$. Orthogonality of these functions would further mean, for two eigenvalues λ_1, λ_2 such that $\lambda_1 \neq \lambda_2$:

$$0 = \int_{-\infty}^{\infty} dx\, f_{\lambda_1}(x)\, f_{\lambda_2}(x) = \int_{-\infty}^{\infty} dx\, e^{(\lambda_1+\lambda_2)x},$$

but this is clearly crazy. The assumption we made that was too restricting was that λ is <u>real</u>: if λ is imaginary

$\lambda = ik$, for some real number k, then these eigenfunctions

are very nice, if <u>complex</u> functions. Now, the eigenfunctions are

$f_k(x) = e^{ikx}$, which has absolute value/magnitude of 1; it never diverges for all x.

This imaginary exponential is familiar from your study of Fourier transforms for which you know that

$$\int_{-\infty}^{\infty} dx\, e^{i(k_1-k_2)x} = 0, \text{ if } k_1 \neq k_2,$$ a manifestation of orthogonality.

This all seems very good so far, other than the strange fact that we had to introduce the imaginary number to do it.

As a teaser for next week, let's re-write the derivative's eigenvalue equation with this new insight:

$$\frac{\partial}{\partial x} f_k(x) = ik\, f_k(x).$$

This is all well and good, but the "ik" on the right is a bit aesthetically displeasing, so let's move the i to the left:

$$\left(-i\frac{\partial}{\partial x}\right) f_k(x) = k\, f_k(x)$$

Recall that k is a real number. This is the eigenvalue equation for the modified derivative operator:

$$\mathcal{O} = -i\frac{\partial}{\partial x},$$

which, apparently exclusively has real eigenvalues, even though (or because of) the explicit factor of i in its expression.

What does this mean? More next week...