8. VECTORS, MATRICES AND TENSORS

Matthew Baring — Lecture Notes for PHYS 516, Fall 2022

1 Linear Vector Spaces: Review

Reading Review: *Content of this section is posted on-line.*

2 Linear Operators and Matrices: Review

Reading Review: *Content of this section is posted on-line.*

3 Eigenvalue Problems

Now we explore special cases where linear operators leave select vectors invariant in direction under the prescribed transformation. These are termed **eigenvalue problems**. Define a column vector $\vec{x}$ in an n-dimensional vector space via

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{1}$$

Then, for a linear operator

$$\mathcal{A} \equiv \begin{pmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} & \cdots & \mathcal{A}_{1n} \\ \mathcal{A}_{21} & \mathcal{A}_{22} & \cdots & \mathcal{A}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{A}_{n1} & \mathcal{A}_{n2} & \cdots & \mathcal{A}_{nn} \end{pmatrix} \tag{2}$$

we can form another vector in the same space

$$\vec{y} = \mathcal{A}\,\vec{x} = \begin{pmatrix} \mathcal{A}_{11}x_1 + \mathcal{A}_{12}x_2 + \cdots + \mathcal{A}_{1n}x_n \\ \mathcal{A}_{21}x_1 + \mathcal{A}_{22}x_2 + \cdots + \mathcal{A}_{2n}x_n \\ \vdots \\ \mathcal{A}_{n1}x_1 + \mathcal{A}_{n2}x_2 + \cdots + \mathcal{A}_{nn}x_n \end{pmatrix} \,. \tag{3}$$

In general, $\vec{y}$ is linearly independent of $\vec{x}$, i.e. $\vec{y} \neq \lambda\vec{x}$ for constant λ. However, in special cases, $\vec{y} = \mathcal{A}\,\vec{x} = \lambda\vec{x}$. Then the scalar λ is termed an **eigenvalue** of the operator $\mathcal{A}$, and $\vec{x}$ is said to be the corresponding **eigenvector** (characteristic vector) of this operator.

• Such eigenvectors are preferred vectors: they do not change their components (up to the scaling of the eigenvalues) under the application of the operator. This desirable character connects to natural modes of a system if the operator is the differential operator characterizing the equation of motion.

∗ Examples abound in quantum mechanics, for example where kinetic energy, angular momentum and spin operators appear. Eigenvalues then constitute quantized physical quantities.

We can recast the eigenvector/eigenvalue equation in matrix language as $(\mathcal{A} - \lambda \mathcal{I})\, \vec{x} = \vec{0}$, where

$$\mathcal{I} \equiv \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \tag{4}$$

is the diagonal identity matrix. Note that $\mathcal{I}$ commutes with everything. The eigenvalue equation $(\mathcal{A} - \lambda \mathcal{I})\, \vec{x} = \vec{0}$ has non-trivial solutions if and only if the determinant of the corresponding matrix is zero:

$$\left| \mathcal{A} - \lambda \mathcal{I} \right| = 0 \quad . \tag{5}$$

This is known as the **secular or characteristic equation**. In general, this is an nth order algebraic equation for λ with n roots in the complex plane. The roots are *not necessarily distinct.*

• To solve a given **eigenvalue problem**, first solve the characteristic equation for the eigenvalues λ_i, $i = 1, 2, \ldots n$, and then plug each λ_i into the eigenvalue equation to find the corresponding eigenvectors.

  ∗ Observe that this gives $n-1$ ratios of components of each eigenvector, i.e. it determines their "directions" but not their magnitudes. This is due to the linearity of the eigenvalue equation: if $\vec{x}$ is an eigenvector, then so also is any scalar multiple of $\vec{x}$.

• **Example 1:** For the Hermitian operator

$$\mathcal{A} \equiv \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{6}$$

the secular equation is

$$\begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix} = 0 \quad , \tag{7}$$

which reduces to $\lambda(\lambda^2 - 1) = 0$, i.e. yields eigenvalues $\lambda = -1, 0, 1$. For $\lambda = -1$ we have

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} . \tag{8}$$

This yields $x_{11} + x_{12} = 0$ and $x_{13} = 0$. The result is

$$\vec{x}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad \text{for} \quad \lambda = -1 . \tag{9}$$

Observe that this can be multiplied by any constant to also yield a "parallel" eigenvector; the present choice is of unit normalization. Likewise for $\lambda = 0$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_{21} \\ x_{22} \\ x_{23} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} , \tag{10}$$

yielding $x_{21} = x_{22} = 0$ and $x_{23} = 1$, or an arbitrary constant. Then

$$\vec{x}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{for} \quad \lambda = 0 . \tag{11}$$

Finally, for $\lambda = +1$,

$$\begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_{31} \\ x_{32} \\ x_{33} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} . \tag{12}$$

This yields $x_{31} - x_{32} = 0$ and $x_{33} = 0$. The result is

$$\vec{x}_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{for} \quad \lambda = 1 . \tag{13}$$

Observe that $\vec{x}_i \cdot \vec{x}_j = \delta_{ij}$, i.e. *these are orthonormal eigenvectors*.

• In the above example, all the eigenvalues are real and distinct: this convenient circumstance is not guaranteed and is often not realized.

3.1 Gram-Schmidt Orthogonalization

For cases with degenerate eigenvalues, we develop an efficient protocol for isolating two linearly independent eigenvectors for the same eigenvalue.

- **Example 2:** Consider the Hermitian operator

$$\mathcal{A} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad . \tag{14}$$

Its secular equation is

$$\begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = 0 \quad , \tag{15}$$

which solves to $(1-\lambda)(\lambda^2 - 1) = 0$, i.e. yields eigenvalues $\lambda = -1, 1, 1$. For $\lambda = -1$ we have

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad . \tag{16}$$

This yields $x_{12} + x_{13} = 0$ and $x_{11} = 0$. The result is

$$\vec{x}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \quad \text{for} \quad \lambda = -1 \quad . \tag{17}$$

For $\lambda = 1$ we repeat the process:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_{21} \\ x_{22} \\ x_{23} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad . \tag{18}$$

This yields $x_{22} - x_{23} = 0$ and $x_{21} = 0$ (arbitrary choice). The result is

$$\vec{x}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \text{for} \quad \lambda = 1 \quad . \tag{19}$$

Observe that $\vec{x}_1 \cdot \vec{x}_2 = 0$. To arrive at a third, mutually orthogonal eigenvector try $\vec{x}_3 = \vec{x}_1 \times \vec{x}_2 = (1,0,0)$. This does satisfy the eigenvalue equation for $\lambda = 1$, and automatically is orthogonal to the other two eigenvectors.

An alternative and more general way to find $\vec{x}_3$ is to choose a **trial vector** $\vec{x}_{3T}$ that is independent of the two known eigenvectors, for example

$$\vec{x}_{3T} = a \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} . \tag{20}$$

and define (for the convenient choice of $a = 1/\sqrt{3}$)

$$\vec{y}_3 = \vec{x}_{3T} - (\vec{x}_{3T} \cdot \vec{x}_2)\,\vec{x}_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{3}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} . \tag{21}$$

This is just $\vec{x}_{3T}$ minus its projection onto $\vec{x}_2$ in the direction of the latter vector; since

$$\vec{y}_3 \cdot \vec{x}_2 = \left\{ 1 - |\vec{x}_2|^2 \right\} (\vec{x}_{3T} \cdot \vec{x}_2) , \tag{22}$$

it is necessarily orthogonal to $\vec{x}_2$ if $\vec{x}_2$ is of unit normalization. Now form the vector $\vec{x}_3 = \vec{y}_3/|\vec{y}_3|$ of unit normalization. The result is (as above)

$$\vec{x}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} . \tag{23}$$

Observe that this protocol could adopt $\vec{x}_1$ to define $\vec{x}_{3T}$, instead of $\vec{x}_2$, and would then still end up with the same result.

• This is an example of the **Gram-Schmidt orthogonalization procedure**, which works for dimensions higher than $n = 3$ also. For an m-fold degenerate set of eigenvalues, we can exploit the arbitrariness of the eigenvectors to create a set of m linearly independent eigenvectors, then use the Gram-Schmidt procedure to convert these into m orthogonal eigenvectors.

• Comparison of the Gram-Schmidt technique to using the cross product approach in this case indicates comparable computational involvement. However, for vectors of much larger dimensions, the cross product approach requires of the order of n^2 computations, and the Gram-Schmidt method is of the order of $3n$ tasks so that it becomes much more efficient.

4 Hermitian Matrices and Diagonalization

The next goal is to determine when the n eigenvalues of an operator are real. This is an important case since it establishes conditions under which we can define physically meaningful quantities such as quantum numbers.

Theorem: If $\mathcal{H}$ is a Hermitian operator, i.e., $\mathcal{H}^\dagger = \mathcal{H}$ (real-symmetric), then (a) its eigenvalues are real, and (b) the eigenvectors associated with different eigenvalues are orthogonal.

Proof: For two eigenvector/eigenvalue pairs we have the identities

$$
\begin{aligned}
\mathcal{H}\,\vec{x}_i &= \lambda_i \vec{x}_i \quad, \quad i = 1,2 \\
(\vec{x}_j)^\dagger \mathcal{H}\,\vec{x}_i &= \lambda_i\,(\vec{x}_j)^\dagger \vec{x}_i \quad, \quad i \neq j \quad.
\end{aligned}
\tag{24}
$$

We then form the complex conjugate of the second of these:

$$
\begin{aligned}
\left\{ (\vec{x}_2)^\dagger \mathcal{H}\,\vec{x}_1 \right\}^* &= \sum_{k,l} \{x_{2k}^* \mathcal{H}_{kl} x_{1l}\}^* = \sum_{k,l} x_{1l}^* \mathcal{H}_{kl}^* x_{2k} \\
&= \sum_{k,l} x_{1l}^* \mathcal{H}_{lk} x_{2k} = (\vec{x}_1)^\dagger \mathcal{H}\,\vec{x}_2 \quad.
\end{aligned}
\tag{25}
$$

We then apply the first equation in Eq. (24) to this to get

$$
\lambda_1^* \left\{ (\vec{x}_2)^\dagger \vec{x}_1 \right\}^* = \lambda_2\,(\vec{x}_1)^\dagger \vec{x}_2 \quad,
\tag{26}
$$

from which it follows that

$$
\{\lambda_1^* - \lambda_2\}\,(\vec{x}_1)^\dagger \vec{x}_2 = 0 \quad \text{since} \quad \left\{ (\vec{x}_2)^\dagger \vec{x}_1 \right\}^* = (\vec{x}_1)^\dagger \vec{x}_2 \quad.
\tag{27}
$$

∗ It immediately follows that (a) if $\lambda_1 = \lambda_2$ and the corresponding eigenvectors are identical, then since $(\vec{x}_1)^\dagger \vec{x}_1 > 0$ and $\lambda_1^* = \lambda_1$, i.e. λ_1 is real.

∗ It also can be concluded that (b) if $\lambda_1 \neq \lambda_2$, then $(\vec{x}_1)^\dagger \vec{x}_2 = 0$, i.e. the two eigenvectors are orthogonal. *QED*

● Illustrations of these properties of Hermitian matrices are contained in the two previous examples.

———

• If we transform to a new coordinate system with basis vectors $\vec{e}\,' = \gamma\,\vec{e}$, then any eigenvectors $\vec{x}_i$ of an operator $\mathcal{A}$ transform via $\vec{x}_i' = \gamma^{-1}\vec{x}_i$. Then,

$$\mathcal{A}\,\vec{x}_i \;=\; \lambda_i\vec{x}_i \;\;\Rightarrow\;\; \left\{\gamma^{-1}\mathcal{A}\,\gamma\right\}\gamma^{-1}\vec{x}_i \;=\; \lambda_i\gamma^{-1}\vec{x}_i \;\;\Rightarrow\;\; \mathcal{A}'\,\vec{x}_i' \;=\; \lambda_i\vec{x}_i' \quad, \quad (28)$$

since $\mathcal{A}' = \gamma^{-1}\mathcal{A}\,\gamma$ is our similarity transformation. Hence *the eigenvalues and eigenvectors of $\mathcal{A}$ are <u>independent of the coordinate system</u>.*

∗ An important consequence is that key physical properties such as conservation of angular momentum are not dependent on the choice of coordinates.

• If we choose $\vec{e}_i'$ to be the orthogonal eigenvectors of $\mathcal{A}$, then $\mathcal{A}'\vec{e}_i' = \lambda_i\vec{e}_i'$ so that the orthogonality implies $\mathcal{A}'_{ij} = \lambda_j\delta_{ij}$. Accordingly, $\mathcal{A}' = \gamma^{-1}\mathcal{A}\,\gamma$ is diagonal, and the diagonal elements are the eigenvalues. The transformation matrix is

$$\gamma \;=\; \begin{pmatrix} \vec{e}_1' & \vec{e}_2' & \cdots & \vec{e}_n' \\ \downarrow & \downarrow & & \downarrow \end{pmatrix} \quad, \quad (29)$$

i.e. its columns are the eigenvectors of $\mathcal{A}$. Therefore, *the eigenvalues of an operator are invariant under a similarity transformation.*

• **A Hermitian operator $\mathcal{H}$ can always be diagonalized by the unitary transformation $\mathcal{H}' = \gamma^{-1}\mathcal{H}\,\gamma$, where the columns of γ are the eigenvectors of $\mathcal{H}$.** This is an important result. A major motivation for eigenvalue/eigenvector determination for Hermitian operators becomes apparent – we can form γ, diagonalize $\mathcal{H}$ and then invert to solve the system.

• **Example 3:** The solution of simultaneous equations can now be formulated. These can be written in the matrix form

$$\vec{y} \;=\; \mathcal{A}\,\vec{x} \quad. \quad (30)$$

Now define a coordinate transformation with matrix γ given by the eigenvectors of $\mathcal{A}$. Since $\vec{x} = \gamma\vec{x}\,'$ and $\vec{y} = \gamma\vec{y}\,'$, it follows that

$$\mathcal{A}\gamma\vec{x}\,' \;=\; \vec{y} \;=\; \gamma\vec{y}\,' \;\;\Rightarrow\;\; \left\{\gamma^{-1}\mathcal{A}\,\gamma\right\}\vec{x}\,' \;=\; \gamma^{-1}\vec{y} \;=\; \vec{y}\,' \quad. \quad (31)$$

The similarity transformation yields a diagonal $\mathcal{A}' \equiv \gamma^{-1}\mathcal{A}\,\gamma$, so that the values for $\vec{x}\,'$ can immediately be read off. It then follows that the solution $\vec{x} = \gamma\vec{x}\,'$ can be quickly obtained. *Once the eigenvectors are obtained, only matrix multiplications are involved in this process.*

● **Example 4:** As an example of diagonalization, consider the operator in Example 2:

$$\mathcal{A} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad . \tag{32}$$

The orthonormal eigenvalues are already assembled there, so that the transformation matrix can be written in the form

$$\gamma \equiv \begin{pmatrix} 0 & 0 & 1 \\ \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} & 0 \\ -\dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} & 0 \end{pmatrix} \quad . \tag{33}$$

This has an inverse

$$\gamma^{-1} \equiv \begin{pmatrix} 0 & \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \\ 0 & \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{pmatrix} \quad . \tag{34}$$

Routine matrix algebra then yields a transformed operator

$$\mathcal{A}' \equiv \gamma^{-1}\mathcal{A}\gamma = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad , \tag{35}$$

wherein the diagonal elements are the eigenvalues of both $\mathcal{A}$ and $\mathcal{A}'$.