Eigenvalues and eigenvectors in the presence of an inner product

We’re still going to be considering linear maps from a vector space to itself, i.e., \( T: V \to V \). But now we will assume that \( V \) has a real or Hermitian inner product (or scalar product or dot product) \( \langle \mathbf{v}, \mathbf{w} \rangle \) which satisfies the usual axioms:

\[
\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle} \quad \text{or} \quad \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle \quad \text{if } V \text{ is real}
\]

\[
\langle \mathbf{v}_1 + \mathbf{v}_2, \mathbf{w} \rangle = \langle \mathbf{v}_1, \mathbf{w} \rangle + \langle \mathbf{v}_2, \mathbf{w} \rangle
\]

\[
\langle \alpha \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle
\]

\[
\langle \mathbf{v}, \mathbf{v} \rangle > 0 \quad \text{for } \mathbf{v} \neq \mathbf{0}
\]

1. Show that these axioms imply

(a) \( \langle \mathbf{v}, \alpha \mathbf{w} \rangle = \overline{\alpha} \langle \mathbf{v}, \mathbf{w} \rangle \)

(b) \( \langle \mathbf{v}, \mathbf{w}_1 + \mathbf{w}_2 \rangle = \langle \mathbf{v}, \mathbf{w}_1 \rangle + \langle \mathbf{v}, \mathbf{w}_2 \rangle \)

(c) If \( T: V \to V \) and \( \lambda \) is an eigenvalue of \( T \) with eigenvector \( \mathbf{v} \) then

\[
\lambda = \frac{\langle T(\mathbf{v}), \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}.
\]

(d) If \( T: V \to V \) and \( \lambda \) is an eigenvalue of \( T \) with eigenvector \( \mathbf{v} \) then

\[
\overline{\lambda} = \frac{\langle \mathbf{v}, T(\mathbf{v}) \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}.
\]

(e) If \( T: V \to V \) and \( \lambda \) is a real eigenvalue of \( T \) with eigenvector \( \mathbf{v} \) then

\[
\langle T(\mathbf{v}), \mathbf{v} \rangle = \langle \mathbf{v}, T(\mathbf{v}) \rangle.
\]

(f) If \( T: V \to V \) and \( \lambda \) is a purely imaginary eigenvalue of \( T \) with eigenvector \( \mathbf{v} \) then

\[
\langle T(\mathbf{v}), \mathbf{v} \rangle = -\langle \mathbf{v}, T(\mathbf{v}) \rangle.
\]

Now we’ll consider two important classes of linear maps that interact in a particularly nice way with the inner product:

**Definition:** A linear map \( T: V \to V \) on an inner-product space \( V \) is called Hermitian (or symmetric in the real case) if

\[
\langle T(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, T(\mathbf{w}) \rangle \quad \text{for all } \mathbf{v} \text{ and } \mathbf{w} \text{ in } V
\]

and \( T \) is called skew-Hermitian (or skew-symmetric in the real case) if

\[
\langle T(\mathbf{v}), \mathbf{w} \rangle = -\langle \mathbf{v}, T(\mathbf{w}) \rangle \quad \text{for all } \mathbf{v} \text{ and } \mathbf{w} \text{ in } V
2. Let $C^\infty([a, b])$ be the space of infinitely-differentiable real-valued functions on the closed interval $[a, b] \subset \mathbb{R}$. Equip $C^\infty([a, b])$ with the (real) inner product

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

(a) Show that multiplication by a fixed (infinitely-differentiable) function $p(x)$ (so $T(f)(x) = p(x)f(x)$) is a symmetric operator.

(b) Show that differentiation is a skew-symmetric operator on the subspace of $C^\infty([a, b])$ of functions for which $f(a) = f(b)$.

(c) Suppose $p(x)$ and $q(x)$ are fixed infinitely-differentiable functions. Show that the Sturm-Liouville operator

$$T(f) = (pf')' + qf$$

is a symmetric operator on the subspace of $C^\infty([a, b])$ of functions for which $f(a) = f(b) = 0$.

(d) Show that the second derivative operator $T(f) = f''$ is a Sturm-Liouville operator on the space of functions $C^\infty([0, \pi])$ satisfying $f(0) = f(\pi) = 0$. Find its eigenvalues and eigenvectors (eigenfunctions), i.e., the values of $\lambda$ for which there are non-zero solutions of

$$f'' = \lambda f, \quad f(0) = 0, \quad f(\pi) = 0$$

together with the corresponding functions.

3. Show that

(a) The eigenvalues of a Hermitian (symmetric) operator are all real, and eigenvectors corresponding to distinct eigenvalues are orthogonal to one another.

(b) The eigenvalues of a skew-Hermitian (skew-symmetric) operator are all purely imaginary, and eigenvectors corresponding to distinct eigenvalues are orthogonal to each other.

(c) What does this imply about the functions you found in the answer to problem 2(d) above?

The most important fact about Hermitian and skew-Hermitian linear transformations of finite-dimensional inner product spaces is they can always be diagonalized. And the change of basis that accomplishes this can always be chosen in a special way. The same is true for symmetric matrices. Skew-symmetric matrices cannot be diagonalized by a real change of basis, but they can be put into a special form as well. We explore this next.

4. Let $T: V \to V$ be a Hermitian transformation of the finite-dimensional inner product space $V$.

(a) Explain how we know that $T$ has at least one eigenvalue $\lambda_1$, and hence at least one eigenvector $\mathbf{v}_1$. Further, explain why we can choose $\mathbf{v}_1$ so that $|\mathbf{v}_1| = 1$.

(b) Let $V_1 \subset V$ be the set of vectors in $V$ that are orthogonal to $\mathbf{v}_1$. Explain why $V_1$ is a linear subspace of $V$. 
(c) Explain why \( T(v_1) \in V_1 \) for any vector \( v_1 \in V_1 \). Therefore we can consider the restriction of \( T \) to \( V_1 \). Explain why \( T: V_1 \to V_1 \) is a Hermitian linear transformation of \( V_1 \).

(d) Show that now we can repeat parts (a), (b) and (c) for the restriction of \( T \) to \( V_1 \), and obtain a second eigenvector \( v_2 \), such that \( |v_2| = 1 \) and \( \langle v_1, v_2 \rangle = 0 \), and \( T \) maps the subspace \( V_2 \subset V \) consisting of all vectors in \( V \) orthogonal to both \( v_1 \) and \( v_2 \) to itself.

(e) Now use induction to show that there is an orthonormal basis of \( V \) consisting of eigenvectors of \( T \).

5. Repeat exercise 4 but this time assume \( T \) is skew-hermitian.

6. Suppose \( \{e_1, e_2, \ldots, e_n\} \) is any orthonormal basis of \( V \), and \( T: V \to V \).

   (a) If \( V \) is complex and \( T \) is Hermitian, show that the entries \( a_{ij} \) of the matrix \( A \) of \( T \) with respect to this orthonormal basis satisfy \( a_{ji} = \overline{a_{ij}} \). (We then say that \( A \) is a Hermitian matrix.)

   (b) If \( V \) is complex and \( T \) is skew-Hermitian, show that \( a_{ji} = -\overline{a_{ij}} \). (We then say that \( A \) is a skew-Hermitian matrix.)

   (c) If \( V \) is real and \( T \) is symmetric, show that \( a_{ji} = a_{ij} \). (So \( A \) is a symmetric matrix.)

   (d) If \( V \) is real and \( T \) is skew-symmetric, show that \( a_{ji} = -a_{ij} \). (So \( A \) is a skew-symmetric matrix.)

7. Show that every Hermitian or skew-Hermitian matrix \( A \) is similar to a diagonal matrix, \( D = C^{-1}AC \), where the diagonal entries of \( D \) are real if \( A \) is Hermitian and purely imaginary if \( A \) is skew-Hermitian. Furthermore, show that \( C \) can be chosen so that the \( C^{-1} = CT \) (such a matrix \( C \) is called a unitary matrix).

8. If \( A \) is a real symmetric matrix, show that \( A \) can be diagonalized, so \( D = C^{-1}AC \) and in this case \( C \) can be chosen so that \( C^{-1} = CT \) (such a matrix is called an orthogonal matrix).

9. What can you say about the determinant of an orthogonal matrix? ... about its eigenvalues? What about those of a unitary matrix?

10. Find an orthogonal matrix \( C \) so that \( C^{-1}AC \) is diagonal if

\[
A = \begin{bmatrix}
16 & 12 \\
12 & 9
\end{bmatrix} \quad A = \begin{bmatrix}
2 & 3 & 0 \\
3 & 1 & -1 \\
0 & -1 & 2
\end{bmatrix}.
\]

11. Show that the set of \( n \)-by-\( n \) unitary matrices is closed under matrix multiplication. Likewise for the set of \( n \)-by-\( n \) orthogonal matrices.

12. Referring back to problem 2(d), find an orthonormal basis of the space of \( C^\alpha([0, \pi]) \)-functions with \( f(0) = f(\pi) = 0 \) with respect to which the second derivative operator is diagonal.

13. Find all 2-by-2 orthogonal matrices (don’t forget the ones with determinant \(-1\)).
14. (a) Show that a “quadratic form” in two variables (i.e., a polynomial in two variables with only terms of degree two) can be represented in the form

\[ p(x, y) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \]

(b) Why is diagonalization of a 2-by-2 quadratic form referred to as “rotation of axes”? How might that help us classify or graph quadratic equations in two variables?

(c) For practice, diagonalize and graph the following quadratic equations:

\[ 5x^2 + 4xy + 2y^2 + 2x - 3y + 5 = 0 \]
\[ x^2 + 4xy - 2y^2 - 6x + 4y - 13 = 0 \]