Extra: More details on the spectral theorem

Let us add $\langle \boldsymbol{v}, \boldsymbol{w} \rangle$ to our notations for the dot product: $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \boldsymbol{v}^T \boldsymbol{w} = \boldsymbol{v} \cdot \boldsymbol{w}$.

- In our story of orthogonality, the important player has been the dot product. However, one could argue that the fundamental quantity is actually the norm:
  $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \frac{1}{4}(\|\boldsymbol{v} + \boldsymbol{w}\|^2 - \|\boldsymbol{v} - \boldsymbol{w}\|^2)$. See Example 28.

- Accepting the dot product as immensely important, we see that symmetric matrices (i.e. matrices A such that $A = A^T$) are of interest.
  For every matrix A, $\langle A\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, A^T \boldsymbol{w} \rangle$.
  It follows that, a matrix A is symmetric if and only if $\langle A\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, A\boldsymbol{w} \rangle$ for all vectors $\boldsymbol{v}, \boldsymbol{w}$.

- Similarly, let Q be an orthogonal matrix (i.e. Q is a square matrix with $Q^T Q = I$).
  Then, $\langle Q\boldsymbol{v}, Q\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$.
  In fact, a matrix A is orthogonal if and only if $\langle A\boldsymbol{v}, A\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ for all vectors $\boldsymbol{v}, \boldsymbol{w}$.
  Comment. We observed in Example 149 that orthogonal matrices Q correspond to rotations ($\det Q = 1$) or reflections ($\det Q = -1$) [or products thereof]. The equality $\langle Q\boldsymbol{v}, Q\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ encodes the fact that these types (and only these!) of geometric transformations preserve angles and lengths.

(spectral theorem)

A $n \times n$ matrix A is symmetric if and only if it can be decomposed as $A = PDP^T$, where

- D is a diagonal matrix, $(n \times n)$

The diagonal entries λ_i are the **eigenvalues** of A.

- P is orthogonal. $(n \times n)$

The columns of P are **eigenvectors** of A.

Note that, in particular, A is always diagonalizable, the eigenvalues (and hence, the eigenvectors) are all real, and, most importantly, the eigenspaces of A are orthogonal.
The "only if" part says that, if A is symmetric, then we get a diagonalization $A = PDP^T$. The "if" part says that, if $A = PDP^T$, then A is symmetric (which follows from $A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$).

Let us prove the following important parts of the spectral theorem.

Theorem 156.

(a) If A is symmetric, then the eigenspaces of A are orthogonal.

(b) If A is real and symmetric, then the eigenvalues of A are real.

Proof.

(a) We need to show that, if $\boldsymbol{v}$ and $\boldsymbol{w}$ are eigenvectors of A with different eigenvalues, then $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \boldsymbol{0}$.
Suppose that $A\boldsymbol{v} = \lambda \boldsymbol{v}$ and $A\boldsymbol{w} = \mu \boldsymbol{w}$ with $\lambda \neq \mu$.
Then, $\lambda \langle \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \lambda \boldsymbol{v}, \boldsymbol{w} \rangle = \langle A\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, A^T \boldsymbol{w} \rangle = \langle \boldsymbol{v}, A\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \mu \boldsymbol{w} \rangle = \mu \langle \boldsymbol{v}, \boldsymbol{w} \rangle$.
However, since $\lambda \neq \mu$, $\lambda \langle \boldsymbol{v}, \boldsymbol{w} \rangle = \mu \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ is only possible if $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0$.

(b) Suppose λ is a nonreal eigenvalue with nonzero eigenvector $\boldsymbol{v}$. Then, $\bar{\boldsymbol{v}}$ is a $\bar{\lambda}$-eigenvector and, since $\lambda \neq \bar{\lambda}$, we have two eigenvectors with different eigenvalues. By the first part, these two eigenvectors must be orthogonal in the sense that $\bar{\boldsymbol{v}}^T \boldsymbol{v} = 0$. But $\bar{\boldsymbol{v}}^T \boldsymbol{v} = \boldsymbol{v}^* \boldsymbol{v} = \|\boldsymbol{v}\|^2 \neq 0$. This shows that it is impossible to have a nonzero eigenvector for a nonreal eigenvalue. □

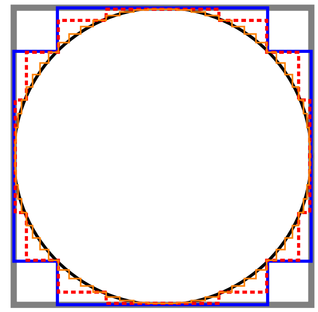
Armin Straub
straub@southalabama.edu

Let us highlight the following point we used in our proof:

Let A be a real matrix. If $\boldsymbol{v}$ is a λ-eigenvector, then $\bar{\boldsymbol{v}}$ is a $\bar{\lambda}$-eigenvector.

See, for instance, Example 83. This is just a consequence of the basic fact that we cannot algebraically distinguish between $+i$ and $-i$.

Remark 157. (Pre-April Fools' Day!) π is the perimeter of a circle enclosed in a square with edge length 1. The perimeter of the square is 4, which approximates π. To get a better approximation, we "fold" the vertices of the square towards the circle (and get the blue polygon). This construction can be repeated for even better approximations and, in the limit, our shape will converge to the true circle. At each step, the perimeter is 4, so we conclude that $\pi = 4$, contrary to popular belief.



Can you pin-point the fallacy in this argument?

Comment. We'll actually come back to this. It's related to linear algebra in infinite dimensions.

Armin Straub
straub@southalabama.edu