Example 161. (extra) We can identify complex numbers $x+iy$ with vectors $\begin{bmatrix} x \\ y \end{bmatrix}$ in $\mathbb{R}^2$. Then, what is the geometric effect of multiplying with i?

Solution. Algebraically, the effect of multiplying $x+iy$ with i obviously is $i(x+iy) = -y+ix$.
Since multiplication with i is obviously linear, we can represent it using a 2×2 matrix J acting on vectors $\begin{bmatrix} x \\ y \end{bmatrix}$.
$J\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ (this is the same as saying $i\cdot 1 = i$) and $J\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ (this is the same as saying $i\cdot i = -1$).

Hence, $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. This is precisely the rotation matrix for a rotation by $90°$.

In other words, multiplication with i has the geometric effect of rotating complex numbers by $90°$.

Comment. The relation $i^2 = -1$ translates to $J^2 = -I$.

Complex numbers as 2×2 matrices. In light of the above, we can express complex numbers $x+iy$ as the 2×2 matrix $xI + yJ = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}$. Adding and multiplying these matrices behaves exactly the same way as adding or multiplying the complex numbers directly.

For instance, $(2+3i)(4-i) = 8+10i-3i^2 = 11+10i$ versus $\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 4 & 1 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} 11 & -10 \\ 10 & 11 \end{bmatrix}$.

Likewise for inverses: $\frac{1}{2+3i} = \frac{2-3i}{(2+3i)(2-3i)} = \frac{2-3i}{13}$ versus $\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}^{-1} = \frac{1}{13}\begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix}$

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 $Q^TQ = I$).
  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 155 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.

Armin Straub
straub@southalabama.edu

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.

We already proved the first part in Theorem 94 using the same argument and only slightly different notation.

Theorem 162.

(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.

Alternative proof. Note that a complex number λ is real if and only if $\bar{\lambda} = \lambda$. Suppose that λ is an eigenvalue with nonzero eigenvector $\boldsymbol{v}$ so that $A\boldsymbol{v} = \lambda \boldsymbol{v}$. We now observe that $\lambda \boldsymbol{v}^* \boldsymbol{v} = \boldsymbol{v}^* (\lambda \boldsymbol{v}) = \boldsymbol{v}^* A \boldsymbol{v} = \boldsymbol{v}^* A^* \boldsymbol{v} = (A\boldsymbol{v})^* \boldsymbol{v} = (\lambda \boldsymbol{v})^* \boldsymbol{v} = \bar{\lambda} \boldsymbol{v}^* \boldsymbol{v}$. Dividing by $\|\boldsymbol{v}\|^2 = \boldsymbol{v}^* \boldsymbol{v}$ (which is not zero!) we find $\lambda = \bar{\lambda}$ from which we conclude that λ is real. □

Advanced comment. Note that the alternative proof of the second part shows that any Hermitian matrix A (that is, a complex matrix A such that $A^* = A$) has only real eigenvalues. If A is Hermitian, what can we conclude about the eigenspaces if we follow the argument in the first part?

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 89. This is just a consequence of the basic fact that we cannot algebraically distinguish between $+i$ and $-i$.