Worksheet 27: The Spectral Theorem (§8.1)

Definition. A square matrix A is **orthogonally diagonalizable** if EITHER of the following two EQUIVALENT conditions holds:

- There exists an *orthogonal* matrix S such that $S^{-1}AS$ (equivalently, $S^{\top}AS$) is diagonal;
- A has an orthonormal eigenbasis.

Spectral Theorem: *A matrix is* **orthogonally diagonalizable** *if and only if it is symmetric.*

Problem 1. Warm-up.

(a) Use the spectral theorem to prove that every diagonal matrix is orthogonally diagonalizable. Now give a direct proof, by showing that the standard basis is an orthonormal eigenbasis.

(b) Use the Spectral theorem to find an example of a 2×2 diagonalizable matrix that does not have an eigenbasis consisting of vectors that are perpendicular to eachother.

Solution: A diagonal matrix is symmetric, so the Spectral theorem tells us it has an orthonormal eigenbasis. But also, if A is diagonal, then $A\vec{e}_j = a_{jj}\vec{e}_j$. So the standard basis is an eigenbasis, and since it is also orthonormal, the standard basis is an orthonormal eigenbasis.

For (b), any diagonalizable matrix that is not symmetric will do. The matrix $\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ is one such. It is obviously not symmetric, and its eigenvalues are 1 and 3. Since there are 2 different eigenvalues, their eigenvectors are linearly independent, hence must be a basis (since the dimension of the source/target space is 2).

Problem 2. Consider the linear transformation T of $\mathbb{R}^2$ given by left mutliplication by $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

(a) Use the Spectral Theorem to decide whether or not T has an *orthonormal eigenbasis*.

(b) Describe T geometrically. Find an orthonormal eigenbasis, using purely geometric thinking.

(c) Write down an *orthogonal* S and *diagonal* D such that $D = S^{-1}AS$. This is called "orthogonally diagonalizing A". Find a matrix S such that $S^{\top}AS$ is diagonal.

Solution:

(a) The matrix is symmetric, so the Spectral theorem tells us it has an eigenbasis consisting of orthonormal eigenvectors.

(b) The map is reflection over the line $y = x$. The vectors on this line (for example $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$) are eigenvectors with eigenvalue 1 (since the map takes them to themselves). The vectors

$\vec{v}$ perpendicular to this line are reflected to $-\vec{v}$, so they are eigenvectors with eigenvalue -1. So $(\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix})$ is an orthogonal eigenbasis. If we want an **orthonormal** eigenbasis, we need to scale down each by its length. An **orthonormal** eigenbasis is $\mathcal{B} = (\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}, \begin{bmatrix}1/\sqrt{2}\\-1/\sqrt{2}\end{bmatrix})$.

(c) $S = \begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}$ is orthogonal (since its columns are **orthonormal**). It is the change of basis matrix from $\mathcal{B}$ to the standard basis. Since $S^\top = S^{-1}$ we can take the same S.

Problem 3. Give quick proofs of the following corollaries of the Spectral theorem.

(a) Every symmetric matrix is similar to a diagonal matrix.

(b) If A is symmetric, then there exists an S such that $S^\top AS$ is diagonal.

(c) If A is an $n \times n$ symmetric matrix, then all eigenvalues of A are *real*; that is, there are no non-real complex eigenvalues of A.

(d) The standard matrix of orthogonal projection onto a two dimensional subspace V of $\mathbb{R}^3$ is symmetric.

Solution:

(a) If A is symmetric, the Spectral theorem tells us it has an eigenbasis $\mathcal{B}$ (even more: it tells us it has an orthonormal eigenbasis). So $[T_A]_\mathcal{B}$ is diagonal, and $A = S[T_A]_\mathcal{B}S^{-1}$ where S is the change of basis matrix from $\mathcal{B}$ to standard. So A is similar to a diagonal matrix.

(b) Following up from (b), since we can assume the eigenbasis $\mathcal{B}$ is orthonormal, the change of basis matrix $S = S_{\mathcal{B}\to\mathcal{E}}$ will have columns that are orthonormal, which means the matrix S is orthogonal. In particular, $S^{-1} = S^\top$. So that $S^{-1}AS = S^\top AS$ is diagonal.

(c) If A is symmetric, the Spectral theorem tells us A has an eigenbasis with real eigenvalues. So the characteristic polynomial (up to sign) is the product of the n factors $(x - \lambda_i)$ where the λ_i range through the n real eigenvalues corresponding to the eigenbasis elements. There are no others.

(c) If we let $(\vec{u}_1, \vec{v}_2)$ be an orthonormal basis for V, and $\vec{n}$ be a unit normal vector to V, then $(\vec{u}_1, \vec{v}_2, \vec{n})$ is an orthonormal eigenbasis for the projection. Thus the projection has an orthonormal eigenbasis. By the spectral theorem, its standard matrix must be symmetric.

Problem 4. Let $A = \begin{bmatrix}3 & 1\\1 & 3\end{bmatrix}$.

(a) Without computing anything, is A diagonalizable? What can you say about its eigenspaces?

(b) Verify that $\begin{bmatrix}1\\1\end{bmatrix}$ is an eigenvector of A. Use the spectral theorem to find another eigenvector.

(c) Without computing the characteristic polynomial, find the eigenvalues for the eigenvectors found in (b). Find a diagonal matrix similar to A.

(d) Find an *orthonormal eigenbasis* $(\vec{u}_1, \vec{u}_2)$ for A.

(e) Find orthogonal S and a diagonal D such that $S^{-1}AS = D$. Is there an S such that $S^\top AS$ is diagonal?

(f) Consider the unit square $\mathcal{Q} = \{s\vec{u}_1 + t\vec{u}_2 \mid 0 \le s, t \le 1\}$ formed by the vectors of your eigenbasis $(\vec{u}_1, \vec{u}_2)$. Explain why the image of $\mathcal{Q}$ under A is a *rectangle* and sketch it. What are the lengths of the sides of $A[\mathcal{Q}]$?

Solution:

(a) The Spectral theorem tells us that because A is symmetric, it has an eigenbasis, so it is diagonalizable. Indeed, it has eigenvectors that form an orthonormal set spanning $\mathbb{R}^2$. There are two eigenspaces and they are orthogonal to each other.

(b) Compute $A\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}4\\4\end{bmatrix} = 4\begin{bmatrix}1\\1\end{bmatrix}$. Since A has an orthonormal eigenbasis, we know that it has eigenvector perpendicular to $\begin{bmatrix}1\\1\end{bmatrix}$, so $\begin{bmatrix}1\\-1\end{bmatrix}$ must be an eigenvector. We check: $\begin{bmatrix}3 & 1\\1 & 3\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix} = \begin{bmatrix}2\\-2\end{bmatrix} = 2\begin{bmatrix}1\\-1\end{bmatrix}$.

(c) The computations above tell us 4 and 2 are eigenevalues. Since the space is two-dimensional, there are no others. A is similar to a diagonal matrix with 2 and 4 on the diagonal (in either order).

(d) The two eigenvectors we found are orthogonal, so get an orthonormal basis we need to scale them: $\mathcal{U} = (\begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}, \begin{bmatrix}1/\sqrt{2}\\-1/\sqrt{2}\end{bmatrix})$.

(e) We can take $S = \begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}$, which is orthogonal (since its columns are **orthonormal**). It is the change of basis matrix from $\mathcal{U}$ to the standard basis. Since $S^{-1} = S^\top$ for orthogonal matrices, the same matrix satisfies $S^\top AS$ is diagonal. That diagonal matrix is the $\mathcal{U}$-matrix for A, so it is $D = \begin{bmatrix}4 & 0\\0 & 2\end{bmatrix}$.

(e) The square $\mathcal{Q}$ has sides of lengths one and vertices $0, \begin{bmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}, \begin{bmatrix}1/\sqrt{2}\\-1/\sqrt{2}\end{bmatrix}, \begin{bmatrix}2/\sqrt{2}\\0\end{bmatrix}$. Applying A, the vector $\vec{u}_1$ which makes up one edge of the square, is stretched to 4 times its length, and the vector $\vec{u}_2$ which makes up another edge of the square, is stretched to twice its length. So $\mathcal{Q}$ is stretched into a rectangle with side lengths 4 and 2 which lies right in the corner formed by the two (perpendicular) eigenspaces of A.

Problem 5. Prove or disprove.

(a) An orthogonal matrix is orthogonally diagonalizable.

(b) For any $d \times n$ matrix B, the matrix $B^\top B$ has an orthonormal eigenbasis.

(c) If A is symmetric, then every eigenbasis is orthonormal.

(d) There exists a 2×2 matrix Q that has no real eigenvalues.

(e) There exists a 2×2 matrix Q such that $QQ^\top$ has no real eigenvalues.

(f) There exists a symmetric 3×3 matrix with an eigenvalue whose geometric multiplicity is one and whose algebraic multiplicity two.

Solution:

(a) False! For example, $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is orthogonal (as it is rotation 90 degrees counterclockwise) but not orthogonally diagonalizable, as it has no real eigenvalues (or because it is not symmetric).

(b) True! This follows from the spectral theorem since $B^\top B$ is symmetric: $(B^\top B)^\top = B^\top B$.

(c) False! The identity matrix is symmetric, but every basis for $\mathbb{R}^n$ is an eigenbasis even though not all bases are orthonormal.

(d) True. See (a).

(e) False: $QQ^\top$ is symmetric, so all eigenvalues are real.

(f) False: since a symmetric matrix is diagonalizable, all geometric multiplicities must equal the corresponding algebraic multiplicity.

The Proof of the Spectral Theorem.

Problem 6. Some easier bits.

(a) Explain why the two bulleted statements in the Spectral Theorem are *equivalent*.

(b) Prove that if A is orthogonally diagonalizable, then A is symmetric.
[Hint: Write A as a product of matrices that are easy to transpose.]

Solution:

(a) If A has an orthonormal eigenbasis $\mathcal{U}$, then the $\mathcal{U}$-matrix of A will be diagonal. Using the change of basis formula:

$$[A]_{\mathcal{U}} = S^{-1}[A]_{\mathcal{E}} S$$

where $S = S_{\mathcal{U}\to\mathcal{E}}$ is the change of basis matrix from $\mathcal{U}$ to the standard basis. Note that $[A]_{\mathcal{U}}$ is diagonal (since $\mathcal{U}$ is an eigenbasis), $[A]_{\mathcal{E}}$ is just A and $S_{\mathcal{U}\to\mathcal{E}}$ has columns consisting of the vectors in the orthonormal basis $\mathcal{U}$. So S is orthogonal (its columns form an orthonormal basis). This shows that if A has an orthonormal eigenbasis, then A is orthogonally diagonalizable. Conversely, if $S^{-1}AS$ is diagonal where S is orthogonal, then the columns of S are an orthonormal basis for $\mathbb{R}^n$ and, since $S^{-1}AS = [A]_{\mathcal{U}}$ is diagonal, we see that $\mathcal{U}$ is also an eigenbasis.

(b) Suppose that A is orthogonally diagonalizable. This means that there is an orthogonal matrix S such that $S^{-1}AS = D$ where D is diagonal. Equivalently, $A = SDS^{-1}$. By definition of orthogonal matrix, $S^{-1} = S^\top$, so also $A = SDS^\top$. Transpose both sides: $A^\top = (SDS^\top)^\top = (S^\top)^\top D^\top S^\top = SD^\top S^\top = SDS^\top$ (using that $D = D^\top$ for a diagonal matrix). So $A^\top = A$, and A is symmetric.

Problem 7. Proof of Spectral Theorem for Diagonalizable Matrices. Let A bs a symmetric matrix.

(a) Let λ_1 and λ_2 be *distinct* eigenvalues of A. Show that the corresponding eigenspaces are *orthogonal.* [Hint: Remember the Lemma: $B\vec{v}\cdot\vec{w} = \vec{v}\cdot B^\top\vec{w}$ for any $n\times n$ matrix B and $\vec{v},\vec{w}\in\mathbb{R}^n$.]

(b) Prove that if A is diagonalizable, then A is orthogonally diagonalizable. [Hint: Use Gram-Schmidt on each eigenspace.]

(This is not a complete proof of the Spectral Theorem—we still need to see why a symmetric matrix is diagonalizable. It is not even clear it has real eigenvalues!).

Solution:

(a) Compute:

$$\lambda_1\vec{v_1}\cdot\vec{v_2} = A\vec{v_1}\cdot\vec{v_2} = (A\vec{v_1})^\top\vec{v_2} = \vec{v_1}^\top A^\top\vec{v_2} = \vec{v_1}^\top A\vec{v_2} = \vec{v_1}^\top\lambda_2\vec{v_2} = \lambda_2\vec{v_1}^\top\vec{v_2} = \lambda_2\vec{v_1}\cdot\vec{v_2}.$$

Since $\lambda_1\neq\lambda_2$, this implies $\vec{v_1}\cdot\vec{v_2} = 0$.

(b) If A is diagonalizable, then it has an eigenbasis. We can find one by taking a basis for each eigenspace E_λ and taking their union. With the basis for each E_λ, we can first apply Gram-Schmidt to it to make it orthonormal. Then the union of the resulting orthonormal sets is orthonormal and since we have the correct number of them, it is an orthonormal basis.

Problem 8. Symmetric Matrices have real eigenvalues. Let $A\in\mathbb{R}^{2\times 2}$ be symmetric. Our goal is to show all eigenvalues of A are real.

(a) Explain why A has exactly n complex eigenvalues, counting (algebraic) multiplicities. Explain why to achieve our goal, it is enough to show that $\overline{\lambda} = \lambda$ for each of them.

(b) Let λ be a (complex) eigenvalue of A with corresponding (complex) eigenvector $\vec{z}$. Explain why $\overline{\vec{z}}$ is also an eigenvector of A, with corresponding eigenvalue $\overline{\lambda}$.

(c) Show that $\lambda = \overline{\lambda}$ by starting with the equation

$$A\overline{\vec{z}} = \overline{\lambda}\,\overline{\vec{z}},$$

transposing both sides, and then multiplying on the right by $\vec{z}$.

Solution:

(a) The characteristic polynomial has exactly n roots, counted with multiplicity. These are the complex eigenvalues of A. A complex number λ is real if $\lambda = \overline{\lambda}$.

(b) We showed this on the previous worksheet.

(c) Given $A\overline{z} = \overline{\lambda}\,\overline{z}$, we transpose to get $\overline{z}^\top A^\top = \overline{\lambda}\overline{z}^\top$. Multiplying on the right by $\vec{z}$, we have

$$\lambda\overline{z}^\top z \;=\; \overline{z}^\top \lambda z \;=\; \overline{z}^\top A z \;=\; \overline{z}^\top A^\top z \;=\; \overline{\lambda}\overline{z}^\top z.$$

Since $\overline{z}^\top z = \sum \overline{z}_i z = \sum |z_i|^2 > 0$, this implies $\lambda = \overline{\lambda}$, ie, $\lambda \in \mathbb{R}$.

Problem 9. Proof of the Spectral Theorem. To prove the Spectral Theorem, it remains to show a symmetric matrix is diagonalizable. We will induce on n.

(a) State and prove the base case.

(b) State the induction hypothesis.

(c) Now, let A be an $(n+1)\times(n+1)$ real symmetric matrix. By Problem (8), A has a real eigenvalue, say λ, with corresponding unit eigenvector $\vec{u}$. Complete $\vec{u}$ to an orthonormal basis $\{\vec{u}, \vec{u}_1, \ldots, \vec{u}_n\}$ of $\mathbb{R}^{n+1}$, and let $Q = [\vec{u}\ \ \vec{u}_1\ \ \cdots\ \ \vec{u}_n]$.

What is the first column of $Q^\top AQ$? How about the first row?

(d) Let B be the $n \times n$ submatrix of $Q^\top AQ$ obtained by deleting the first row and first column of $Q^\top AQ$. Is B symmetric?

(e) Apply the induction hypothesis and try to finish the proof by showing that A is orthogonally diagonalizable. (This may take some work! Write down carefully what the matrix $Q^\top AQ$ looks like after you apply your induction hypothesis).

Solution:

(a) The base case $n = 1$ is: "For any $A \in \mathbb{R}^{1\times 1}$, if A is symmetric then there is orthogonal $Q \in \mathbb{R}^{1\times 1}$ such that $Q^\top AQ$ is diagonal." Since every 1×1 matrix is both symmetric and diagonal, the base case holds trivially.

(b) Fix n. "Assume that every $n \times n$ symmetric matrix is orthogonally diagonalizable."

(c) Let $\mathcal{U} = (\vec{u}, \vec{u}_1, \ldots, \vec{u}_n)$, so $Q^\top = Q^{-1} = S_{\mathcal{E}\to\mathcal{U}}$. Then

$$Q^\top AQ\vec{e}_1 \;=\; Q^\top A\vec{u} \;=\; Q^\top \lambda\vec{u} \;=\; \lambda S_{\mathcal{E}\to\mathcal{U}}[\vec{u}]_{\mathcal{E}} \;=\; \lambda[\vec{u}]_{\mathcal{U}} \;=\; \lambda\vec{e}_1.$$

Since $(Q^\top AQ)^\top = Q^\top AQ$, we also have $\vec{e}_1^\top Q^\top AQ = \lambda\vec{e}_1$. Thus $Q^\top AQ$ has the form

$$\begin{bmatrix} \lambda & 0 \\ 0 & B \end{bmatrix}$$

where B is $n \times n$.

(d) Since $Q^\top AQ$ is symmetric (as shown above), so is B.

(e) Since B is an $n \times n$ symmetric matrix, we know by the induction hypothesis that B is orthogonally diagonalizable, say $B = RDR^\top$ where D is diagonal and R is orthogonal. Then

$$Q^\top AQ = \begin{bmatrix} \lambda & 0 \\ 0 & B \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & RDR^\top \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} \lambda & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & R^\top \end{bmatrix},$$

and therefore

$$\begin{bmatrix} 1 & 0 \\ 0 & R^\top \end{bmatrix} Q^\top A\, Q \begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & D \end{bmatrix}.$$

Since $\begin{bmatrix} \lambda & 0 \\ 0 & D \end{bmatrix}$ is diagonal and $\begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix}$ is orthogonal, and a product of orthogonal matrices is orthogonal, this shows that A is orthogonally diagonalizable, completing the induction.

Problem 10. Consider the matrix $A = \begin{bmatrix} 1 & -2 & 2 \\ -2 & 4 & -4 \\ 2 & -4 & 4 \end{bmatrix}$.

(a) Is A orthogonally diagonalizable? (Do not actually try to diagonalize it yet – just answer yes or no).

(b) Without finding the characteristic polynomial of A, find one eigenvalue. What is its geometric multiplicity?

(c) Again without using the characteristic polynomial of A, find the remaining eigenvalue(s).

(d) Describe the transformation $T(\vec{x}) = A\vec{x}$ geometrically by thinking about what effect it has on vectors in each eigenspace. (Hint: what is the relationship between the eigenspaces?)

(e) Show, geometrically, that $A^2 = 9A$.

(f) Orthogonally diagonalize A.

Solution:

(a) Yes! A is symmetric, so it is (orthogonally) diagonalizable by the Spectral Theorem.

(b) By inspection one sees that the columns of A are linearly dependent, so one eigenvalue of A is 0. Since the second and third columns of A are scalar multiples of the first (which is nonzero), $\dim(\ker(A)) = \text{gemu}(0) = 2$.

(c) Since A is 3×3 and $\text{gemu}(0) = \text{almu}(0) = 2$, we are looking for one more eigenvalue. As every eigenvector of A belongs to $\ker(A)$ or $\text{im}(A)$ and we have already considered $\ker(A)$, our final eigenspace is the 1-dimensional $\text{im}(A)$. This is spanned by the eigenvector $A\vec{e}_1$, with corresponding eigenvalue 9. Note that $\text{almu}(9) = \text{gemu}(9) = 1$.

(d) For each $\vec{v} \in \mathbb{R}^3$, $A\vec{v}$ is 9 times the projection of $\vec{v}$ onto the line generated by $A\vec{e}_1$.

(e) The fact that $A^2 = 9A$ is clear from the geometric description in (d). Alternatively, it is clear that $A^2\vec{v} = 9A\vec{v}$ for each vector $\vec{v}$ in any orthonormal eigenbasis of $\mathbb{R}^3$ for A, and such a basis must exist by the Spectral Theorem.

(f) By inspection we can complete $\begin{bmatrix}1\\-2\\2\end{bmatrix}$ to the orthogonal basis $\left(\begin{bmatrix}1\\-2\\2\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}-4\\-1\\1\end{bmatrix}\right)$. Normalizing, we obtain $A = PDP^\top$ where

$$D = \begin{bmatrix}9 & & \\ & 0 & \\ & & 0\end{bmatrix} \qquad \text{and} \qquad P = \begin{bmatrix}\frac{1}{3} & 0 & \frac{-4}{\sqrt{18}} \\ \frac{-2}{3} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{18}} \\ \frac{2}{3} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{18}}\end{bmatrix}.$$

Problem 11. Recall that by the Fundamental Theorem of Algebra, every $n \times n$ matrix has exactly n complex eigenvalues, counting multiplicities.

(a) Prove that if the $n \times n$ matrix A has n real eigenvalues (counting multiplicities), then A is similar (over $\mathbb{R}$) to an upper triangular matrix.

[HINT: mimic the proof of the Spectral Theorem.]

Solution: We argue by induction on n, the claim being obvious for the base case $n = 1$. For the inductive step, let $n \geq 1$, suppose every $n \times n$ matrix with n real eigenvalues is similar over $\mathbb{R}$ to an upper triangular matrix, and let A be an $(n+1) \times (n+1)$ matrix with $n+1$ real eigenvalues. Abusing notation as the book does, we think of A as a linear transformation from $\mathbb{R}^{n+1}$ to itself with standard matrix A. That is, letting $\mathcal{E}$ denote the standard basis of $\mathbb{R}^{n+1}$, we have $[A]_\mathcal{E} = A$.

Fix an eigenvalue λ of A, along with an eigenvector $\vec{v}$ corresponding to λ. Let $\mathcal{B} = (\vec{v}, \vec{v}_1, \dots, \vec{v}_n)$ be a basis of $\mathbb{R}^{n+1}$. Then the matrix of the tranformation A in the basis $\mathcal{B}$ has a block form

$$[A]_\mathcal{B} = \begin{bmatrix}\lambda & \vec{c}^\top \\ 0 & B\end{bmatrix}$$

where $B \in \mathbb{R}^{n\times n}$ and $\vec{c}^\top \in \mathbb{R}^{1\times n}$. Of course the matrix $[A]_\mathcal{B}$ is similar to A, as they represent the same transformation in different bases; in particular the matrix $[A]_\mathcal{B}$ has the same characteristic polynomial as A. Computing the determinant of

$$[A]_\mathcal{B} - xI_{n+1}$$

using a Laplace expansion along the first column, we see that the characteristic polynomial of $[A]_\mathcal{B}$ and hence of A is

$$(\lambda - x)\ \det(B - xI_n).$$

Since all $n+1$ roots of this characteristic polynomial are real, we conclude that all roots of $\det(B - xI_n)$ must be real. In other words, the $n \times n$ matrix B has all of its eigenvalues real. By induction, B is similar to an upper triangular matrix R—that is, $B = SRS^{-1}$ for some invertible $n \times n$ matrix S. In other words,

$$[A]_\mathcal{B} = \begin{bmatrix}\lambda & \vec{c}^\top \\ 0 & SRS^{-1}\end{bmatrix}, \tag{1}$$

where $R \in \mathbb{R}^{n\times n}$ is upper triangular. So if we let $\tilde{S}$ be the $(n+1)\times(n+1)$ block diagonal matrix

$$\tilde{S} = \begin{bmatrix}1 & \vec{0}^\top \\ \vec{0} & S\end{bmatrix}$$

where $\vec{0} \in \mathbb{R}^{n\times 1}$, then using block matrix multiplication it's easy to see that

$$\begin{bmatrix} 1 & \vec{0}^\top \\ \vec{0} & S \end{bmatrix} \begin{bmatrix} \lambda & \vec{c}^\top S \\ 0 & R \end{bmatrix} \begin{bmatrix} 1 & \vec{0}^\top \\ \vec{0} & S^{-1} \end{bmatrix} = \begin{bmatrix} \lambda & \vec{c}^\top SS^{-1} \\ 0 & SRS^{-1} \end{bmatrix} = \begin{bmatrix} \lambda & \vec{c}^\top \\ 0 & B \end{bmatrix} = [A]_{\mathcal{B}}.$$

Define $\tilde{R}$ to be the matrix $\begin{bmatrix} \lambda & \vec{c}^\top S \\ 0 & R \end{bmatrix}$, and note that $\tilde{R}$ is upper triangular. So $[A]_{\mathcal{B}}$ is similar to the upper triangular matrix $\tilde{R}$. So also $A = [A]_{\mathcal{E}}$ is similar to the same upper triangular matrix $\tilde{R}$. [In general, matrices similar to the same matrix are similar to each other. In this case, letting $P = S_{\mathcal{B}\to\mathcal{E}}$, we have

$$A = [A]_{\mathcal{E}} = P[A]_{\mathcal{B}}P^{-1} = P\tilde{S}\tilde{R}\tilde{S}^{-1}P^{-1} = (P\tilde{S})\tilde{R}(P\tilde{S})^{-1},$$

showing that A is indeed similar (over $\mathbb{R}$) to an upper triangular matrix.] This completes the induction, which completes the proof.

(b) Explain why your proof in part (a) shows that *every* square matrix is similar over $\mathbb{C}$ to an upper triangular matrix (possibly with complex entries).

Solution: Let A be a $n \times n$ matrix. Then A has n complex eigenvalues (counting multiplicities) by the Fundamental Theorem of Algebra. Thus we can run the same argument used in part (a), with the only difference being that we allow λ and the entries in the matrices P, C, B, R, and S to be complex. This argument then shows that A is similar over $\mathbb{C}$ to an upper triangular matrix.

(c) Explain why part (b) shows that the trace and determinant of a square matrix A are the sum and product, respectively, of all the (complex) eigenvalues of A.

Solution: Let A be an $n \times n$ matrix. By part (b), we know that A is similar (over $\mathbb{C}$) to an upper triangular matrix, call it U. Then the (complex) eigenvalues $\lambda_1, \ldots, \lambda_n$ of U are just the diagonal entries of U. But since U is triangular, its trace and determinant are just the sum and product, respectively, of its diagonal entries, i.e., its eigenvalues. The corresponding claim for A now follows from the fact that A and U are similar, so they have the same eigenvalues, trace, and determinant.