Worksheet 27: The Spectral Theorem (§8.1)

Problem 1. Find the eigenvalues of the *symmetric* 2×2 matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$. Under what conditions on the (real) entries a, b, c are these eigenvalues real?

Solution: The characteristic polynomial is $t^2 - (a+c)t + ac - b^2$, so the eigenvalues are

$$t \;=\; \frac{(a+c) \pm \sqrt{(a+c)^2 - 4ac + 4b^2}}{2} \;=\; \frac{1}{2}\left(a + c \pm \sqrt{(a-c)^2 + 4b^2}\right) \;\in\; \mathbb{R}.$$

So the eigenvalues are real no matter what a, b, c are.

Problem 2. Extend your observation in (1) and prove that for any symmetric $n \times n$ matrix A with real entries, A has n real eigenvalues (counting multiplicities). Here are some suggestions on how to proceed:

- Since A is $n \times n$, its characteristic polynomial has degree n, and thus A has n *complex* eigenvalues (counting multiplicities) by the Fundamental Theorem of Algebra. So all you need to show is that every complex eigenvalue of A is actually real.
- To show this, let λ be a (complex) eigenvalue of A with corresponding (complex) eigenvector $\vec{z}$. Then $\overline{\vec{z}}$ is also an eigenvector of A, with corresponding eigenvalue $\overline{\lambda}$, as you showed on the previous worksheet. Starting with the equation
$$A\overline{\vec{z}} \;=\; \overline{\lambda}\,\overline{\vec{z}},$$
take the transpose of both sides and then multiply on the right by $\vec{z}$.

Solution: Let $\lambda \in \mathbb{C}$ be an eigenvalue of A, with corresponding (complex) eigenvector z. Then as we showed on the previous worksheet, $A\overline{z} = \overline{\lambda}\,\overline{z}$. Therefore $\overline{z}^\top A^\top = \overline{\lambda}\overline{z}^\top$, so

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

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

* * *

Now you are ready to prove the ...

Spectral Theorem: An $n \times n$ matrix with real entries is symmetric if and only if it is *orthogonally diagonalizable*, meaning that there is an $n \times n$ orthogonal matrix Q and a diagonal matrix D such that $A = QDQ^\top$.

(Remember that if Q is orthogonal, then $Q^\top = Q^{-1}$)

Problem 3. First prove the backward direction: show that if A is orthogonally diagonalizable, then A must be symmetric.

Solution: Suppose A is orthogonally diagonalizable, say $A = QDQ^\top$ where D is diagonal and Q is orthogonal, so $D^\top = D$ and $Q^\top = Q^{-1}$. Then

$$A^\top \;=\; (QDQ^\top)^\top \;=\; (Q^\top)^\top D^\top Q^\top \;=\; QDQ^\top \;=\; A,$$

so A is symmetric.

Problem 4. The forward direction is harder, and we will prove it by induction on n.

(a) State and prove the base case.

(b) State the induction hypothesis.

Now, let A be an $(n+1)\times(n+1)$ real symmetric matrix. By Problem (2), A has a real eigenvalue, say λ, with corresponding unit eigenvector $\vec{u}$. Let's complete $\vec{u}$ to an orthonormal basis $\{\vec{u}, \vec{u}_1, \ldots, \vec{u}_n\}$ of $\mathbb{R}^{n+1}$, and let $Q = [\vec{u}\ \ \vec{u}_1\ \ \cdots\ \ \vec{u}_n]$.

(c) 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 $A \in \mathbb{R}^{1\times1}$, if A is symmetric then there is orthogonal $Q \in \mathbb{R}^{1\times1}$ such that $Q^\top AQ$ is diagonal." Since every 1×1 matrix is both symmetric and diagonal, the base case holds trivially.

(b) "Let $n \geq 1$, and suppose that for every $n \times n$ matrix A, if A is symmetric then A 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 5. Prove, without using the Spectral Theorem, that if λ and μ are distinct eigenvalues of the symmetric matrix A with corresponding eigenvectors $\vec{v}$ and $\vec{w}$, then $\vec{v} \cdot \vec{w} = 0$.

(Hint: compare $\lambda(\vec{v} \cdot w)$ and $\mu(\vec{v} \cdot \vec{w})$)

Solution: Suppose λ and μ are distinct eigenvalues of the symmetric matrix A, with corresponding eigenvalues $\vec{v}$ and $\vec{w}$. Then

$$\lambda\vec{v} \cdot \vec{w} \;=\; A\vec{v} \cdot \vec{w} \;=\; (A\vec{v})^\top \vec{w} \;=\; \vec{v}^\top A^\top \vec{w} \;=\; \vec{v}^\top A\vec{w} \;=\; \vec{v}^\top \mu\vec{w} \;=\; \mu\vec{v}^\top \vec{w} \;=\; \mu\vec{v} \cdot \vec{w}.$$

Since $\lambda \neq \mu$, this implies $\vec{v} \cdot \vec{w} = 0$.

Can you give another proof using the Spectral Theorem?

Solution: Let A be a symmetric matrix. Using the Spectral Theorem, write $A = QDQ^\top$ where D is diagonal and Q is orthogonal. Then the diagonal entries of D are the eigenvalues of A, and for each eigenvalue λ of A, the columns of Q that correspond to the occurrences of λ in D form a basis of E_λ, the eigenspace corresponding to λ. Suppose now that λ and μ are distinct eigenvalues of A, and let $(\vec{u}_1, \ldots, \vec{u}_r)$ and $(\vec{v}_1, \ldots, \vec{v}_s)$ be the columns of Q that correspond to λ and μ, respectively, in D, so that the former is a basis of E_λ and the latter is a basis of E_μ. Let $\vec{0} \neq \vec{w}_1 \in E_\lambda$ and $\vec{0} \neq \vec{w}_2 \in E_\mu$, and write $\vec{w}_1 = \sum_{i=1}^{m} c_i\vec{u}_i$ and $\vec{w}_2 = \sum_{j=1}^{n} d_j\vec{v}_j$. Then

$$\vec{w}_1 \cdot \vec{w}_2 \;=\; \left(\sum_{i=1}^{m} c_i\vec{u}_i \right) \cdot \left(\sum_{j=1}^{n} d_j\vec{v}_j \right) \;=\; \sum_{i=1}^{m} \sum_{j=1}^{n} c_i d_j (\vec{u}_i \cdot \vec{v}_j) \;=\; 0$$

since the columns of Q are orthonormal and therefore $\vec{u}_i \cdot \vec{v}_j = 0$ for all i, j.

Problem 6. 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 7. 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. Fix an eigenvalue λ of A, along with an eigenvector $\vec{v}$ corresponding to λ. Let $\mathcal{B} = (\vec{v}, \vec{v}_1, \ldots, \vec{v}_n)$ be a basis of $\mathbb{R}^{n+1}$, and let P be the matrix whose columns are the basis vectors in $\mathcal{B}$. Then $P^{-1}AP$ has the form

$$P^{-1}AP \;=\; \begin{bmatrix} \lambda & C \\ 0 & B \end{bmatrix}$$

where B is an $n \times n$ matrix. Since A has $n + 1$ real eigenvalues and the characteristic polynomial of A is $(t - \lambda) \cdot \det(tI_n - B)$, B has n real eigenvalues. Therefore, by the inductive hypothesis we can write $B = SRS^{-1}$ where R is upper triangular. Then

$$A \;=\; P \begin{bmatrix} \lambda & C \\ 0 & B \end{bmatrix} P^{-1} \;=\; P \begin{bmatrix} \lambda & C \\ 0 & SRS^{-1} \end{bmatrix} P^{-1} \;=\; P \begin{bmatrix} 1 & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} \lambda & CS \\ 0 & R \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & S^{-1} \end{bmatrix} P^{-1},$$

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.