MATH 217 - LINEAR ALGEBRA
HOMEWORK 11, SOLUTIONS

Part A (10 points)

Solve the following problems from the book:
Section 7.2: 12.

Solution. The characteristic polynomial is

$$\det\begin{bmatrix} 2-x & -2 & 0 & 0 \\ 1 & -1-x & 0 & 0 \\ 0 & 0 & 3-x & -4 \\ 0 & 0 & 2 & -3-x \end{bmatrix} = (x^2-x)(x^2-1) = x(x-1)^2(x+1).$$

Thus eigenvalues are 0 with algebraic multiplicity 1, 1 with algebraic multiplicity 2, and -1 with algebraic multiplicity 1.

Section 7.3: 20.

Solution. The characteristic polynomial is $\chi_A(x) = (x-1)((x-1)^2-1) = x(x-1)(x-2)$. Eigenvalues are 0, 1, and 2.

Eigenspaces are

- $E_0 = \ker(A) = \mathrm{Span}(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix})$. A basis of E_0: $\{\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\}$.
- $E_1 = \ker(\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}) = \mathrm{Span}(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix})$. A basis of E_1: $\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\}$.
- $E_2 = \ker(\begin{bmatrix} -1 & 0 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}) = \mathrm{Span}(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix})$. A basis of E_2: $\{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\}$.

Diagonalizing A:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix}^{-1}.$$

Section 7.3: 24.

Solution. For example, we can choose a matrix A to be $\begin{bmatrix} 0 & 2 \\ 0.5 & 2 \end{bmatrix}$.

A "complete" way to think about this is by problem 1c of part B below, $A = S\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}S^{-1}$.
Any invertible matrix A with the first column is a nonzero vector in Span($\begin{bmatrix}2\\1\end{bmatrix}$) will work.

Section 7.3: 36.

Solution. The first matrix has characteristic polynomial is x^2-3x-5 while the second matrix has characteristic polynomial x^2-4x-5. Since similar matrices have the same characteristic polynomial, the two given matrix are not similar.

Section 7.5: 14.

Solution. The matrix S will have a form $\begin{bmatrix}\vec{w} & \vec{v}\end{bmatrix}$ where $\vec{v}+i\vec{w}$ is an eigenvector of the matrix (Theorem 7.5.3 textbook or Problem 1d part B below). Thus we want to find a complex eigenvector of the matrix/ For this, the characteristic polynomial is x^2+1, and thus has two root i and $-i$. The eigenspace

$$E_i = \ker(\begin{bmatrix}1-i & -2\\ 1 & -1-i\end{bmatrix}) = \mathrm{Span}(\begin{bmatrix}2\\ 1-i\end{bmatrix}).$$

Therefore we can choose $S = \begin{bmatrix}0 & 2\\ -1 & 1\end{bmatrix}$.

Section 8.1: 24.

Solution. The characteristic polynomial is $\xi_A(x) = -x\det\begin{bmatrix}-x & 1 & 0\\ 1 & -x & 0\\ 0 & 0 & -x\end{bmatrix} - 1\det\begin{bmatrix}0 & -x & 1\\ 0 & 1 & -x\\ 1 & 0 & 0\end{bmatrix} = x^2(x^2-1)-(x^2-1) = (x-1)^2(x+1)^2$. Eigenspaces

$$E_1 = \mathrm{Span}(\vec{e}_1+\vec{e}_4, \vec{e}_2+\vec{e}_3),$$

and

$$E_{-1} = \mathrm{Span}(\vec{e}_1-\vec{e}_4, \vec{e}_2-\vec{e}_3).$$

E_1 has an orthonormal basis $(\frac{1}{\sqrt{2}}(\vec{e}_1+\vec{e}_4), \frac{1}{\sqrt{2}}(\vec{e}_2+\vec{e}_3))$ and E_{-1} has an orthonormal basis $(\frac{1}{\sqrt{2}}(\vec{e}_1-\vec{e}_4), \frac{1}{\sqrt{2}}(\vec{e}_2-\vec{e}_3))$. Thus $(\frac{1}{\sqrt{2}}(\vec{e}_1+\vec{e}_4), \frac{1}{\sqrt{2}}(\vec{e}_2+\vec{e}_3), \frac{1}{\sqrt{2}}(\vec{e}_1-\vec{e}_4), \frac{1}{\sqrt{2}}(\vec{e}_2-\vec{e}_3)$ is an orthonormal basis for A.

Part B (25 points)

Problem 1. *(Classifying non-diagonalizable*[1] 2×2 *matrices.)* Let $A\in\mathbb{R}^{2\times 2}$ be a 2×2 matrix.

[1] We work over $\mathbb{R}$ throughout this problem. So"eigenvalue" means *real eigenvalue*, "diagonalizable" means *diagonalizable over* $\mathbb{R}$, and "similar" means *similar over* $\mathbb{R}$.

(a) Suppose that A has eigenvalue 0 but is not diagonalizable. Prove that[2] $\operatorname{im}(A) = E_0$, and conclude from this that $A^2 = 0$.

(b) Let $\lambda \in \mathbb{R}$ and suppose that A has eigenvalue λ but is not diagonalizable. Prove that we have $(A - \lambda I_2)^2 = 0$, and deduce from this that $A\vec{v} - \lambda\vec{v} \in E_\lambda$ for every $\vec{v} \in \mathbb{R}^2$.

[Hint: apply part (a) to the matrix $A - \lambda I_2$.]

(c) Prove that if A has eigenvalue λ but is not diagonalizable, then A is similar to $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$.

[Hint: consider the basis $\mathcal{B} = (A\vec{v} - \lambda\vec{v}, \vec{v})$ where $\vec{v} \notin E_\lambda$.]

(d) Prove that if A does not have any real eigenvalues, then A is similar to a matrix of the form λQ where Q is an orthogonal matrix and $\lambda > 0$.

Solution.

(a) Since A has an eigenvalue 0, A is not invertible. On the other hand, as A is not diagonalizable, the only eigenvalue of A is 0 and A is not the zero matrix. Therefore, $\operatorname{im}(A)$ has 1 dimension. For every nonzero vector $\vec{v}$ in $\operatorname{im}(A)$, the vector Av is in $\operatorname{im}(A)$, thus a multiple of v. Hence v must be an eigenvector. Combining with 0 is the only eigenvalue, it follows that $\operatorname{im}(A) = E_0$. Hence, $A^2\vec{x} = A(A\vec{x}) = 0$ for every $x \in \mathbb{R}^2$. By picking x to be $\vec{e}_1$ and $\vec{e}_2$, we obtain that A^2 is the zero matrix.

(b) We first prove the following claim:

Claim: A is diagonalizable with eigenvalue λ_1, λ_2 (not necessarily to be different) if and only if $A - \lambda I_2$ is diagonalizable with eigenvalues $\lambda_1 - \lambda, \lambda_2 - \lambda$.

Proof of Claim: A is diagonalizable with eigenvalue λ_1, λ_2 iff there exists an invertible matrix S such that $A = S \begin{bmatrix} \lambda_1\vec{e}_1 & \lambda_2\vec{e}_2 \end{bmatrix} S^{-1}$. This is equivalent with

$$A - \lambda I_2 = S \begin{bmatrix} \lambda_1\vec{e}_1 & \lambda_2\vec{e}_2 \end{bmatrix} S^{-1} - S \begin{bmatrix} \lambda\vec{e}_1 & \lambda\vec{e}_2 \end{bmatrix} S^{-1} = S \begin{bmatrix} \lambda_1 - \lambda & 0 \\ 0 & \lambda_2 - \lambda \end{bmatrix} S^{-1}.$$

Thus the claim follows.

Back to the problem, by Claim, the matrix $A - \lambda I_2$ has eigenvalue 0 but is not diagonalizable. Hence, by part (a), $(A - \lambda I)^2 = 0$. Therefore, for every $\vec{v} \in \mathbb{R}^2$,

$$(A - \lambda I_2)(A\vec{v} - \lambda\vec{v}) = (A - \lambda I_2)^2\vec{v} = 0.$$

This implies that $A\vec{v} - \lambda\vec{v} \in \ker(A - \lambda I_2) = E_\lambda$.

(c) Consider $\mathcal{B} = (A\vec{v} - \lambda\vec{v}, \vec{v})$, where $\vec{v} \notin E_\lambda$. We note that $A\vec{v} - \lambda\vec{v} \neq 0$ and belongs to E_λ by part (b), while $\vec{v}$ is a nonzero vector not in E_λ. Therefore they are not multiple of each other, equivalently, they are linearly independent. Thus $\mathcal{B}$ is a basis of $\mathbb{R}^2$. By part (b), $A(A\vec{v} - \lambda\vec{v}) = \lambda(A\vec{v} - \lambda\vec{v})$. It follows that $[A(A\vec{v} - \lambda\vec{v})]_{\mathcal{B}} = \begin{bmatrix} \lambda \\ 0 \end{bmatrix}$. On the other hand, $[A\vec{v}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ \lambda \end{bmatrix}$. Therefore,

$$[T_A]_{\mathcal{B}} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix},$$

where T_A is the linear map with A as its standard matrix. Now the claim follows as $A = [T_A]_{\mathcal{E}}$ is similar to $[T_A]_{\mathcal{B}}$.

[2] Recall that for each $\lambda \in \mathbb{R}$, $E_\lambda = \{\vec{v} \in \mathbb{R}^2 \ : \ A\vec{v} = \lambda\vec{v}\}$.

(d) Let $p(x)$ be the characteristic polynomial of A. Suppose that $a + bi$ is a root of p. Then $p(a - bi) = \overline{p(a + bi)} = 0$, where the first equality follows from the fact that all coefficients of p are real. Thus, if A does not have real eigenvalue, we can assume that A has two distinct complex eigenvalues $a + bi$ and $a - bi$.

Suppose $A(\vec{v} + i\vec{w}) = (a + bi)(\vec{v} + i\vec{w})$, so also $A(\vec{v} - i\vec{w}) = (a - bi)(\vec{v} - i\vec{w})$. Then, diagonalizing A over $\mathbb{C}$, we have

$$\begin{bmatrix} | & | \\ \vec{v} + i\vec{w} & \vec{v} - i\vec{w} \\ | & | \end{bmatrix}^{-1} A \begin{bmatrix} | & | \\ \vec{v} + i\vec{w} & \vec{v} - i\vec{w} \\ | & | \end{bmatrix} = \begin{bmatrix} a + bi & \\ & a - bi \end{bmatrix},$$

Multiply both sides by $\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$ (on left) and $\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}^{-1}$ (on right), we get

$$\begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} | & | \\ \vec{v} + i\vec{w} & \vec{v} - i\vec{w} \\ | & | \end{bmatrix}^{-1} A \begin{bmatrix} | & | \\ \vec{v} + i\vec{w} & \vec{v} - i\vec{w} \\ | & | \end{bmatrix} \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

By direct computation we have $\begin{bmatrix} | & | \\ \vec{v} + i\vec{w} & \vec{v} - i\vec{w} \\ | & | \end{bmatrix} \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \vec{w} & \vec{v} \end{bmatrix}$, so

$$\begin{bmatrix} \vec{w} & \vec{v} \end{bmatrix}^{-1} A \begin{bmatrix} \vec{w} & \vec{v} \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

The claim now follows from the fact that $\frac{1}{\sqrt{a^2+b^2}} \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ is an orthogonal matrix.

Problem 2. *("Spectral theorem" for skew-symmetric matrices.)* Let A be an $n \times n$ matrix such that $A^\top = -A$. Write T_A for the linear map induced by A, so $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is defined by $T_A(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^n$.

(a) Prove that $\ker(A) = \operatorname{im}(A)^\perp$, and if n is odd then $\ker(A) \neq \{\vec{0}\}$.
(b) Prove that for every $\vec{v} \in \mathbb{R}^n$, $(A\vec{v}) \cdot \vec{v} = 0$.
(c) Prove that if $\vec{v}$ is an eigenvector of A, then $\vec{v} \in \ker(A)$.
(d) Let $T : \operatorname{im}(A) \to \operatorname{im}(A)$ be the restriction of T_A to $\operatorname{im}(A)$, i.e., the linear transformation from $\operatorname{im}(A)$ to $\operatorname{im}(A)$ defined by $T(x) = A\vec{x}$ for every $\vec{x} \in \operatorname{im}(A)$. Prove that T is invertible.
(e) Prove that the dimension of $\operatorname{im}(A)$ is even.
(f) Prove that A^2 is symmetric. Thus, by the Spectral Theorem, A^2 is diagonalizable, and in particular every eigenvalue of A^2 is real. Prove that every eigenvalue of A^2 is non-positive. [Hint: if $A^2\vec{v} = \lambda\vec{v}$ with $\vec{v} \neq \vec{0}$, compute $\lambda\vec{v} \cdot \vec{v}$.]
(g) Prove that if $\vec{v} \neq 0$ and $A^2\vec{v} = -\lambda\vec{v}$ for some $\lambda > 0$, then $(\vec{v}, A\vec{v})$ is linearly independent and $A\vec{v}$ is also an eigenvector of A^2 with eigenvalue $-\lambda$.
(h) Prove that if $\vec{v} \neq 0$ and $A^2\vec{v} = -\lambda\vec{v}$ for some $\lambda > 0$, then[3] $T_A[V] = V$ and $T_A[V^\perp \cap \operatorname{im}(A)] \subseteq V^\perp \cap \operatorname{im}(A)$, where $V = \operatorname{Span}(\vec{v}, A\vec{v})$.

[3]Recall from the *More Joy of Sets* handout that for any function $f : X \to Y$ and subset $V \subseteq X$, we define $f[V] = \{f(x) : x \in V\}$.

(i) Suppose now that $n = 3$ and that A is nonzero (and skew-symmetric). Prove that there is an orthonormal basis $\mathcal{U}$ of $\mathbb{R}^3$ and a scalar $\lambda > 0$ such that

$$[T_A]_{\mathcal{U}} = \begin{bmatrix} 0 & \sqrt{\lambda} & 0 \\ -\sqrt{\lambda} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

(j) **(Recreational.)** Prove that the previous part holds for any n, in the following sense: for every skew-symmetric matrix $A \in \mathbb{R}^{n\times n}$, there is an orthonormal basis $\mathcal{U}$ of $\mathbb{R}^n$ such that $[T_A]_{\mathcal{U}}$ is a block-diagonal matrix whose diagonal blocks are either zero or have the form $\begin{bmatrix} 0 & \sqrt{\lambda} \\ -\sqrt{\lambda} & 0 \end{bmatrix}$ for some $\lambda > 0$.

Solution.

(a) We have that $\ker(A) = \operatorname{im}(A^T)^\perp = \operatorname{im}(-A)^\perp = \operatorname{im}(A)^\perp$. For the second claim, when n is odd then $\det(A) = \det(A^T) = \det(-A) = (-1)^n \det(A) = -\det(A)$. Hence $\det(A) = 0$, and A is not invertible. Thus $ker(A) \neq \{\vec{0}\}$.

(b) For every $\vec{v} \in \mathbb{R}^n$,
$$\vec{v} \cdot (A\vec{v}) = \vec{v}^T A\vec{v} = -\vec{v}^T A^T \vec{v} = -(A\vec{v}) \cdot \vec{v} = -\vec{v} \cdot (A\vec{v}).$$
It follows that $(A\vec{v}) \cdot \vec{v} = \vec{v} \cdot (A\vec{v}) = 0$.

(c) Suppose $\vec{v}$ is a μ-eigenvector of A. Then $Av = \mu\vec{v}$. By part (b), $\mu\vec{v} \cdot \vec{v} = (A\vec{v}) \cdot \vec{v} = 0$. Since $\vec{v} \neq \vec{0}$, we must have $\mu = 0$. Therefore $v \in E_0 = \ker(A)$.

(d) We first note that T is well-defined as for every $\vec{x} \in \operatorname{im}(A)$, $T(\vec{x}) = A\vec{x} \in im(A)$. To prove T is invertible, we let $\vec{x} \in \ker(T) \subset \operatorname{im}(A)$. Then $A\vec{x} = 0$, which implies that $\vec{x} \in \ker(A)$. Hence, $\vec{x} \in \ker(A) \cap \operatorname{im}(A)$. By part (a), $\vec{x} = \vec{0}$. Therefore, $\ker(T) = \{\vec{0}\}$. Since $\operatorname{im}(A)$ has finite dimension, T is invertible.

(e) Suppose, for the sake of contradiction, that $\operatorname{im}(A)$ has odd dimension. Then T has an eigenvector. But eigenvector of T is also an eigenvector of A. By part (c), the eigenvalue associated to this eigenvector is 0. This contradicts to the invertibility of T by part (d). Therefore $\operatorname{im}(A)$ has an even dimension.

(f) The claim A^2 is symmetric follows from $(A^2)^T = (A^T)^2 = (-A)^2 = A^2$. For the second claim, suppose that $\vec{v}$ is a η-eigenvector for A^2. Then
$$\eta\vec{v} \cdot \vec{v} = \vec{v}^T A^2 \vec{v} = -\vec{v}^T A^T A\vec{v} = -(A\vec{v}) \cdot (A\vec{v}).$$
As $\vec{v} \cdot \vec{v} > 0$ and $(A\vec{v}) \cdot (A\vec{v}) \geq 0$, it follows that $\eta \leq 0$.

(g) We first show that $(\vec{v}, A\vec{v})$ is linearly independent. Indeed, let $c, d \in \mathbb{R}^2$ such that $c\vec{v} + dA\vec{v} = \vec{0}$. Apply A to both sides, we get $cA\vec{v} - d\lambda\vec{v} = \vec{0}$. We thus obtain
$$\lambda dc\vec{v} + \lambda d^2 A\vec{v} = \vec{0},$$
and
$$c^2 A\vec{v} - \lambda dc\vec{v} = \vec{0}.$$
Adding two equations, we obtain $(c^2 + \lambda d^2)A\vec{v} = \vec{0}$. Since $A^2\vec{v} \neq \vec{0}$, we must have $A\vec{v} \neq \vec{0}$. it follows that $c^2 + \lambda d^2 = 0$, which implies $c = d = 0$ as $\lambda > 0$ by assumption.
For the second claim, $A^2(A\vec{v}) = A(A^2\vec{v}) = A(-\lambda\vec{v}) = -\lambda A\vec{v}$.

(h) If we set $V = \mathrm{Span}(\vec{v}, A\vec{v})$ then $T_A[V] = \mathrm{Span}(A\vec{v}, A^2\vec{v}) = \mathrm{Span}(A\vec{v}, -\lambda\vec{v}) = V$.
For the second claim, if $\vec{w} \in V^\perp \cap \mathrm{im}(A)$ then $\vec{w} \cdot \vec{v} = 0 = \vec{w} \cdot (A\vec{v})$. Thus

$$(A\vec{w}) \cdot \vec{v} = -\vec{w} \cdot (A\vec{v}) = 0,$$

and

$$(A\vec{w}) \cdot (A\vec{v}) = -\vec{w} \cdot (A^2\vec{v}) = 0.$$

This implies $A\vec{w} \in V^\perp$. Therefore $T_A[V^\perp \cap \mathrm{im}(A)] \subset V^\perp \cap \mathrm{im}(A)$.

(i) By part (a), A is not invertible. By part (e) and the assumption A is not the zero matrix, $\mathrm{im}(A)$ has two dimension, and thus $\ker(A)$ has 1 dimension. The matrix A^2 is symmetric. A^2 must have a different eigenvalue from 0, follows from part (d). Let $\lambda > 0$ such that $-\lambda$ is an eigenvalue of A^2 (part f). By part (g), $(-\lambda)$-eigenspace of A^2 must have at least two dimensions, and hence have exactly two dimension. Since A^2 is orthogonally diagonalizable, the $(-\lambda)$-eigenspace of A^2 is othogonal to $\ker(A)$ (the 0-eigenspace). Thus $(-\lambda)$-eigenspace of A^2 is a subspace of $\mathrm{im}(A)$, having the same dimension as $\dim(\mathrm{im}(A))$, hence conincides with $\mathrm{im}(A)$.

Pick a nonzero $\vec{v} \in \mathrm{im}(A)$. Then by part (g), $(\vec{v}_1, \vec{v}_2) = (\frac{\vec{v}}{\|\vec{v}\|}, -\frac{A\vec{v}}{\|A\vec{v}\|})$ is an orthonormal basis of $\mathrm{im}(A)$. Pick $\vec{v}_3$ to be a unit vector in $\ker(A)$. Then $\mathcal{B} = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$ is an orthonormal basis of $\mathbb{R}^3$. With respect to this basis, we must have And,

$$[T_A]_{\mathcal{B}} = \begin{bmatrix} 0 & \lambda_1 & 0 \\ \lambda_2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

for some $\lambda_1, \lambda_2 \in \mathbb{R}$. To complete the proof, we prove that $\lambda_1 = \sqrt{\lambda}$ and $\lambda_2 = -\sqrt{\lambda}$. For this we note that $A\vec{v}_1 = \lambda_2\vec{v}_2$, $A\vec{v}_2 = \lambda_1\vec{v}_1$, and $A^2\vec{v}_1 = -\lambda\vec{v}_1$. Therefore $\lambda_1\lambda_2 = -\lambda$. On the other hand, let $S = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix}$. Then S is orthogonal, and thus $[T_A]_{\mathcal{B}} = S^TAS$ is skew symmetric. This implies $\lambda_1 = -\lambda_2$. As λ_2 must be negative, we must have $\lambda_1 = \sqrt{\lambda}$ and $\lambda_2 = -\sqrt{\lambda}$. *Alternatively (thanks to Yuanjun), we have that*

$$A\vec{v}_1 = \frac{A\vec{v}}{\|\vec{v}\|} = \frac{A\vec{v}}{\sqrt{\frac{-1}{\lambda}\vec{v} \cdot (A^2\vec{v})}} = \frac{A\vec{v}}{\sqrt{\frac{1}{\lambda}\vec{v}^TA^TA\vec{v}}} = \sqrt{\lambda}\frac{A\vec{v}}{\|A\vec{v}\|} = -\sqrt{\lambda}\vec{v}_2.$$

And similarly $A\vec{v}_2 = \sqrt{\lambda}\vec{v}_1$.

In conclusion,

$$[T_A]_{\mathcal{B}} = \begin{bmatrix} 0 & \sqrt{\lambda} & 0 \\ -\sqrt{\lambda} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

(j) (Sketch) First we apply Spectral theorem to A^2 to obtain that there exist $\lambda_1, \lambda_2, \ldots, \lambda_k > 0$ such that $\lambda_1, \lambda_2, \ldots, \lambda_k$ and possibly $\lambda_0 = 0$ are eigenvalue of A^2. Moreover $E_{\lambda_i}^{(A^2)}$ (the superscript indicates the eigenspace for A^2) is orthogonal to $E_{\lambda_j}^{(A^2)}$ if $i \neq j$. For each $i = 1, \ldots, k$ we are going to choose a "good" orthogonal basis for $E_{\lambda_i}^{(A^2)}$. The procedure is as follows

- First pick a nonzero $\vec{v}_1 \in E_{\lambda_i}^{(A^2)}$, by part (b) and (g), $A\vec{v}_1 E_{\lambda_i}^{(A^2)}$ and is orthogonal to $\vec{v}_1$.
- If $\mathrm{Span}(\vec{v}_1, A\vec{v}_1) \neq E_{\lambda_i}^{(A^2)}$, then pick $\vec{v}_2 \in E_{\lambda_i}^{(A^2)} \cap \mathrm{Span}(\vec{v}_1, A\vec{v}_1)^\perp$. By part (b), (g), and (h), $A\vec{v}_2 \in E_{\lambda_i}^{(A^2)}$ and $(\vec{v}_1, A\vec{v}_1, \vec{v}_2, A\vec{v}_2)$ is orthogonal.

- If $\mathrm{Span}(\vec{v}_1, A\vec{v}_1, \vec{v}_2, A\vec{v}_2)$ is not $E^{(A^2)}_{\lambda_i}$, the we pick $\vec{v}_3 \in E^{(A^2)}_{\lambda_i} \cap \mathrm{Span}(\vec{v}_1, A\vec{v}_1, \vec{v}_2, A\vec{v}_2)^\perp$. Continue this process until we absorb all $E^{(A^2)}_{\lambda_i}$. This process also shows that dimension of each $E^{(A^2)}_{\lambda_i}$ is even.
- The result of this process is that there exist nonzero vectors $\vec{v}_1, \ldots, \vec{v}_m$ such that $E^{(A^2)}_{\lambda_i} = \mathrm{Span}(\vec{v}_1, A\vec{v}_1, \vec{v}_2, A\vec{v}_2, \ldots, \vec{v}_m, A\vec{v}_m)$ and $(\vec{v}_1, A\vec{v}_1, \vec{v}_2, A\vec{v}_2, \ldots, \vec{v}_m, A\vec{v}_m)$ is orthogonal.
- We pick a basis $\mathcal{B}_i = (\frac{\vec{v}_1}{\|\vec{v}_1\|}, -\frac{A\vec{v}_1}{\|A\vec{v}_1\|}, \frac{\vec{v}_2}{\|\vec{v}_2\|}, -\frac{A\vec{v}_2}{\|A\vec{v}_2\|}, \ldots, \frac{\vec{v}_m}{\|\vec{v}_m\|}, -\frac{A\vec{v}_m}{\|A\vec{v}_m\|})$. This is an orthogonal basis of $E^{(A^2)}_{\lambda_i}$.

Do the same procedure for other eigenspace of A^2, to obtain orthogonal bases $\mathcal{B}_1, \ldots, \mathcal{B}_k$ of eigenspaces. Note that $\ker(A) = \ker(A^2)$ (by part (d)). We let $\mathcal{B}_0$ be an orthogonal basis for $\ker(A)$. Since A^2 is orthogonally diagonalizable, the union $\mathcal{B} = \mathcal{B}_0 \cup \mathcal{B}_1 \cup \ldots \mathcal{B}_k$ is an orthogonal basis of $\mathcal{R}^n$. Now easily check that $[T_A]_\mathcal{B}$ is a block-diagonal matrix with diagonal blocks are either zero or of the form $\begin{bmatrix} 0 & \sqrt{\lambda_i} \\ -\sqrt{\lambda_i} & 0 \end{bmatrix}$.

Problem 3. Let $(V, \langle \cdot, \cdot \rangle)$ be a finite dimensional inner product vector space, and let $T : V \to V$ be a linear transformation.

(a) Prove that there is a unique transformation $T^* : V \to V$ such that $\langle T(x), y\rangle = \langle x, T^*(y)\rangle$ for all $x, y \in V$.
(b) Prove that if $T = T^*$, then there is a basis $\mathcal{B}$ of V such that $[T]_\mathcal{B}$ is symmetric.
(c) Prove that if $T = T^*$, then T is orthogonally diagonalizable.

Solution.

(a) If V were $\mathbb{R}^n$ and $\langle \cdot, \cdot \rangle$ the dot product, we could represent T by its standard matrix A so that for all $x, y \in \mathbb{R}^n$,

$$\langle T(x), y\rangle \;=\; Ax \cdot y \;=\; x^\top A^\top y \;=\; x \cdot A^\top y,$$

and thus T^* would be the linear map induced by $A^\top$. With this in mind, let us fix an orthonormal basis $\mathcal{U}$ of V, so we have $\langle x, y\rangle = [x]_\mathcal{U} \cdot [y]_\mathcal{U}$ for all $x, y \in V$, and define T^* to be the linear map from V to V whose $\mathcal{U}$-matrix is $[T]^\top_\mathcal{U}$. Thus, formally, T^* is defined by the rule

$$T^*(y) \;=\; L^{-1}_\mathcal{U}\big([T]^\top_\mathcal{U}\, [y]_\mathcal{U}\big) \quad \text{for all } y \in V.$$

Note that this definition is independent of our choice of orthonormal basis $\mathcal{U}$ of V, since if $\mathcal{U}'$ is another orthonormal basis of V then the change-of-coordinates matrix $S = S_{\mathcal{U}\to\mathcal{U}'}$ is orthogonal, so $[T]^\top_{\mathcal{U}'} = \big(S[T]_\mathcal{U} S^\top\big)^\top = S[T]^\top_\mathcal{U} S^\top$, which implies

$$[T^*]_\mathcal{U} = [T]^\top_\mathcal{U} \quad \text{if and only if} \quad [T^*]_{\mathcal{U}'} = [T]^\top_{\mathcal{U}'}.$$

Then for all $x, y \in V$, we have

$$\begin{aligned} \langle T(x), y\rangle \;&=\; [T(x)]_\mathcal{U} \cdot [y]_\mathcal{U} \;=\; [T]_\mathcal{U}[x]_\mathcal{U} \cdot [y]_\mathcal{U} \;=\; [x]^\top_\mathcal{U}[T]^\top_\mathcal{U}[y]_\mathcal{U} \\ &=\; [x]^\top_\mathcal{U}[T^*]_\mathcal{U}[y]_\mathcal{U} \;=\; [x]_\mathcal{U} \cdot [T^*(y)]_\mathcal{U} \;=\; \langle x, T^*(y)\rangle. \end{aligned}$$

This establishes existence of T^*, and for uniqueness suppose $S : V \to V$ is another linear map satisfying $\langle T(x), y\rangle = \langle x, S(y)\rangle$ for all $x, y \in V$. Let $y \in V$ be arbitrary, and note that for all $x \in V$ we have

$$\langle x, S(y) - T^*(y)\rangle \;=\; \langle x, S(y)\rangle - \langle x, T^*(y)\rangle \;=\; \langle T(x), y\rangle - \langle T(x), y\rangle \;=\; 0.$$

In particular, $\langle S(y) - T^*(y), S(y) - T^*(y)\rangle = 0$, which implies $S(y) = T^*(y)$ by positive-definiteness of the inner product. This implies $S = T^*$, establishing uniqueness.

(b) In part (a) we saw that $[T^*]_{\mathcal{U}} = [T]_{\mathcal{U}}^\top$ for any orthonormal basis $\mathcal{U}$ of V. Thus if $T = T^*$, we have that $[T]_{\mathcal{U}}$ is symmetric for any orthonormal basis $\mathcal{U}$ of V.

(c) Suppose $T = T^*$, and fix an orthonormal basis $\mathcal{U} = (u_1, \ldots, u_n)$ of V. Then $[T]_{\mathcal{U}}$ is symmetric, hence orthogonally diagonalizable by the Spectral Theorem, so we can fix an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ such that $Q^\top [T]_{\mathcal{U}} Q$ is diagonal. For each $1 \le i \le n$, let $b_i = L_{\mathcal{U}}^{-1}(Q\vec{e}_i)$, and let $\mathcal{B} = (b_1, \ldots, b_n)$, so $Q = S_{\mathcal{B}\to\mathcal{U}}$. Then

$$[T]_{\mathcal{B}} \;=\; S_{\mathcal{U}\to\mathcal{B}}[T]_{\mathcal{U}} S_{\mathcal{B}\to\mathcal{U}} \;=\; Q^\top [T]_{\mathcal{U}} Q,$$

so $[T]_{\mathcal{B}}$ is diagonal. Furthermore, since $\mathcal{U}$ is orthonormal and $Q = S_{\mathcal{B}\to\mathcal{U}}$ is orthogonal,

$$\langle b_i, b_j\rangle \;=\; [b_i]_{\mathcal{U}} \cdot [b_j]_{\mathcal{U}} \;=\; Q[b_i]_{\mathcal{B}} \cdot Q[b_j]_{\mathcal{B}} \;=\; e_i^\top Q^\top Q e_j \;=\; \vec{e}_i \cdot \vec{e}_j \;=\; \delta_{ij}$$

for all $1 \le i, j \le n$, so $\mathcal{B}$ is orthonormal and T is indeed orthogonally diagonalizable.