MATB24 TUTORIAL PROBLEMS 8, SOLUTIONS

KEY WORDS: Eigenvalue and eigenvectors, diagonalization

RELEVANT SECTIONS IN THE TEXTBOOK: RELEVANT SECTIONS IN THE TEXTBOOK: 7.2 FB or 5.C SA

WARM-UP: As usual, write down a complete definition or a complete mathematical characterization for the following terms. You may need to consult the material for MATA22 to these definitions.

- Eigenvector of a matrix.
- Eigenvalues of a matrix.
- A diagonalizable matrix.
- Characteristic polynomial of a matrix.
- Orthogonal complement

Recall from MATA22: An $n \times n$ matrix is diagonalizable if it is similar to a diagonal matrix. That is if there is an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$. Here is how you decide whether a given A is diagonalizable with your MATA22 knowledge:

(1) Find all eigenvalues of A by finding the roots of the characteristic polynomial of A, that is $\det(A - xI) = 0$. For each root, λ, the algebraic multiplicity of λ is the largest number n $(x - \lambda)^n$ appears in factorization of $\det(A - xI)$.

(2) For each eigenvalue λ, find the eigenspace $E_\lambda = Nul(A - \lambda I)$. Then geometric multiplicity of λ is $\dim(E_\lambda)$.

(3) A is diagonalizable if 1) for each λ, algebraic multiplicity of λ is equal to the geometric multiplicity of λ 2) the sum of all geometric multiplicities is equal to n.

(4) The columns of P are basis of E_λ's, and the diagonal entries of D are eigenvalues of A (put in a compatible order)

In this worksheet you develop a new point of view to the concept of diagonalizable matrices using the notion of change of basis, which sheds light on the above procedure. First, you need to generalize the concepts of eigenvalue, eigenvectors and characteristic polynomial from matrices to linear transformations.

A:

DEFINITION 0.1. *Let V be a vector space of dimension n, and let $T : V \to V$ be a linear transformation. A non-zero vector $\vec{v} \in V$ is said to be an eigenvector of T if*

$$T(\vec{v}) = \lambda \vec{v}$$

for some scalar λ. The scalar λ is called the eigenvalue corresponding to $\vec{v}$.

(1) Show that for any scalar λ, the set $V_\lambda = \{\vec{v} \in V : T(\vec{v}) = \lambda \vec{v}\}$ is a subspace of V. When is this subspace nontrivial (i.e., not equal to $\{\vec{0}\}$)?

(2) For the following transformations, find an eigenvector using any methods you can think of, including basic geometry, if this is possible. What are the corresponding eigenvalues?

  (a) $V = \mathbb{R}^2$, $T =$ reflection over the x-axis.

(b) $V = \mathbb{R}^2$, $T =$ reflection over the line $x = y$.
(c) $V = \mathbb{R}^2$, $T =$ rotation by $90°$.
(d) $V = \mathbb{R}^2$, $T =$ left multiplication by $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$
(e) $V = \mathbb{P}_3$ the space of polynomials of degree less than or equal 3 in the variable t, $T(f) = f'$.

Solution.

(1) $T(\vec{0}) = \vec{0} = \lambda\vec{0}$, so $\vec{0} \in V_\lambda$. For $\vec{v}, \vec{w} \in V_\lambda$ and a scalar r, $T(\vec{v}+r\vec{w}) = \lambda\vec{v}+r(\lambda\vec{w}) = \lambda(\vec{v} + r\vec{w})$. The subspace is non-trivial when λ is an eigenvalue of T.

(2) (a) Have $T([x, y]) = [x, -y]$, so $T([1, 0]) = [1, -0] = [1, 0]$, so $[1, 0]$ is an eigenvector with eigenvalue 1.

(b) $T([x, y]) = [y, x]$, so $T([1, 1]) = [1, 1]$ and thus $[1, 1]$ is an eigenvector with eigenvalue 1.

(c) $T([x, y]) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} [x, y]^t = [-y, x]$, so there are no eigenvectors: $[-y, x] = \lambda[x, y]$ implies $-y = \lambda x = \lambda^2 y$, so if $y \neq 0$, then $-1 = \lambda^2$ which is impossible, so $y = 0$, in which case $x = \lambda y = 0$.

(d) $T([x, y]) = [2x + y, 2y]$, so $T([1, 0]) = [2, 0]$ and hence $[1, 0]$ is an eigenvector with eigenvalue 2.

(e) $T(a + bt + ct^2 + dt^3) = b + 2ct + 3dt^2$, so $T(1) = 0 = 0 \cdot 1$ and hence the constant polynomial 1 is an eigenvector with eigenvalue 0.

B: A SYSTEMATIC WAY TO FIND EIGENVALUES

DEFINITION 0.2. *Let V be a vector space of dimension n, and let $T : V \to V$ be a linear transformation. Determinant of T, denoted by $\det(T)$ is defined to be $\det([T]_{\mathcal{B}})$, where $\mathcal{B}$ is any basis for B.*

(1) Is the definition of $\det(T)$ well defined? What may go wrong? Prove that this definition is well-defined.
(2) Let λ be an eigenvalue of $T : V \to V$ and let $\vec{v}$ be an eigenvector with eigenvalue λ. Then
$$(T - \lambda\mathrm{id})\vec{v} = \vec{0}.$$
This means that the linear transformation $T - \lambda\mathrm{id}$ has a nontrivial kernel. Suppose $\mathcal{B}$ is a basis for V. Discuss with your group why we can deduce
$$\det([T - \lambda\mathrm{id}]_{\mathcal{B}}) = 0.$$ [1]
(3) Conversely, reversing the argument shows that if $\det([T - \lambda\mathrm{id}]_{\mathcal{B}}) = 0$, then λ is an eigenvalue of T.
(4) Show that $\det([T - \lambda\mathrm{id}]_{\mathcal{B}})$ does not depend on our choice of basis $\mathcal{B}$. That is choose a different basis $\mathcal{A}$ and show that $\det([T - \lambda\mathrm{id}]_{\mathcal{B}}) = \det([T - \lambda\mathrm{id}]_{\mathcal{A}})$.[2]
(5) Discuss with your group why the following definition is well defined:

DEFINITION 0.3. *(Characteristic polynomial) The characteristic polynomial of the linear transformation T is the polynomial in the variable λ given by $\det([T-\lambda\mathrm{id}]_{\mathcal{B}})$, where $\mathcal{B}$ is any basis of V.*

---

[1]Recall from MATA22 that a matrix have non-zero determinant if and only if it is invertible.
[2]Remember from MATA22 that $\det(AB) = \det(A)\det(B)$.

(6) Let A and B be similar matrices and let $r \in F$. Prove that $A - rI$ and $B - rI$ are similar.
(7) Let $\mathcal{B}$ be a basis of V. Show that
$$([T - \lambda \mathrm{id}]_{\mathcal{B}}) = [T]_{\mathcal{B}} - \lambda I$$
Conclude the characteristic polynomial of T is equal to $\det([T]_{\mathcal{B}} - \lambda I)$.
(8) To find all the eigenvalues λ of T, we find roots of $\det([T]_{\mathcal{B}} - \lambda I) = 0$. Use the previous parts to justify this. Discuss with your groups what happens if we choose a different basis.
(9) Choose two transformations in problem A part 2 and find their eigenvalues using this method.
(10) Write down the characteristic polynomials for the linear transformations you chose in previous part.
(11) Suppose $\vec{v}_1$ and $\vec{v}_2$ are eigenvectors corresponding to *distinct* eigenvalues λ_1 and λ_2. Show that $\vec{v}_1$ and $\vec{v}_2$ are linearly independent.
(12) Extend the statement and proof of the previous problem to r eigenvectors.

Solution.

(1) It is well defined, we need to show that if we have two basis $\mathcal{A}, \mathcal{B}$, we want to show $\det([T]_{\mathcal{A}}) = \det([T]_{\mathcal{B}})$. Assume C is the change of basis matrix from basis $\mathcal{A}$ to $\mathcal{B}$, C is invertible. We know $[T]_{\mathcal{A}} = C^{-1}[T]_{\mathcal{B}}C$, then $\det([T]_{\mathcal{A}}) = \det(C)^{-1}\det([T]_{\mathcal{B}})\det(C) = \det([T]_{\mathcal{B}})$, thus it is well defined.
(2) Since $T - \lambda \mathrm{id}$ has non-trivial kernel, therefore it is not injective and hence not invertible. Therefore its matrix representation is not invertible and hence it has determinant 0.
(3) It follows that $T - \lambda \mathrm{id}$ is not invertible. Since $T : V \to V$, therefore being not invertible means being neither injective nor surjective (e.g. rank-nullity theorem). So in particular T has non-trivial kernel, i.e. there exists $\vec{v} \neq \vec{0}$ such that $(T - \lambda \mathrm{id})(\vec{v}) = \vec{0}$, i.e. $T(\vec{v}) = \lambda \vec{v}$.
(4) There exists an invertible matrix C such that $[T - \lambda \mathrm{id}]_{\mathcal{B}} = C[T - \lambda \mathrm{id}]_{\mathcal{A}}C^{-1}$, and hence $\det[T - \lambda \mathrm{id}]_{\mathcal{B}} = \det(C[T - \lambda \mathrm{id}]_{\mathcal{A}}C^{-1}) = \det C \det[T - \lambda \mathrm{id}]_{\mathcal{A}} \det C^{-1} = \det C \det C^{-1} \det[T - \lambda \mathrm{id}]_{\mathcal{A}} = 1 \cdot \det[T - \lambda \mathrm{id}]_{\mathcal{A}}$.
(5) Previously it was shown that the choice of basis does not matter.
(6) There is some invertible matrix C such that $A = CBC^{-1}$, so $C(B - rI)C^{-1} = CBC^{-1} + rCIC^{-1} = A + rI$.
(7) Let $\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_n)$. Recall $[T - \lambda \mathrm{id}]_{\mathcal{B}}$ is the matrix with columns $[(T - \lambda \mathrm{id})(\vec{b}_i)]_{\mathcal{B}}$. Using linearity of taking basis representations, $[(T - \lambda \mathrm{id})(\vec{b}_i)]_{\mathcal{B}} = [T(\vec{b}_i) + \lambda \vec{b}_i]_{\mathcal{B}} = [T(\vec{b}_i)]_{\mathcal{B}} + \lambda[\vec{b}_i]_{\mathcal{B}} = [T(\vec{b}_i)]_{\mathcal{B}} + \lambda \vec{e}_i$ where $\vec{e}_i$ is the standard basis for $\mathbb{R}^n$. It is seen that $[T]_{\mathcal{B}} - \lambda I$ has the same columns and hence they are the same matrix. The equality of the characteristic polynomial follows from the definition and the equality of the matrices implying equality of their determinants.
(8) T has an eigenvalue λ_0 iff there is some eigenvector $\vec{v} \neq \vec{0}$ such that $T(\vec{v}) = \lambda_0 \vec{v}$. (1), (2) show this is equivalent to $\det[T - \lambda \mathrm{id}]_{\mathcal{B}} = 0$ having solution $\lambda = \lambda_0$. In terms of (4) this means the characteristic polynomial has solution $\lambda = \lambda_0$. (3), (6) shows this is equivalent to $\lambda = \lambda_0$ being a solution of $\det([T]_{\mathcal{B}} - \lambda I) = 0$.
(9) Consider (d). Get $\det \begin{bmatrix} 2 - \lambda & 1 \\ 0 & 2 - \lambda \end{bmatrix} = (2 - \lambda)^2 = 0$ with solutions $\lambda = 2$. So (d) has eigenvalue 2.

(10) $p(\lambda) = (2-\lambda)^2$.
(11) Suppose $r\vec{v}_1 + t\vec{v}_2 = \vec{0}$. Suppose for a contradiction that at least one of r, t are nonzero. Without loss of generality assume $r \neq 0$. Applying T to the dependence relation yields $r\lambda_1\vec{v}_1 + t\lambda_2\vec{v}_2 = \vec{0}$. Multiplying the original dependence relation by λ_2 yields $r\lambda_2\vec{v}_1 + t\lambda_2\vec{v}_2 = \vec{0}$. Subtracting these last two equations yields $\vec{0} = r(\lambda_1 - \lambda_2)\vec{v}_1$. Since $r \neq 0$ and $\vec{v}_1 \neq \vec{0}$ this implies that $\lambda_1 = \lambda_2$, a contradiction.
(12) Statement: Given r eigenvectors $\vec{v}_1, \ldots, \vec{v}_r$ corresponding to r distinct eigenvalues $\lambda_1, \ldots, \lambda_r$, it follows that $\vec{v}_1, \ldots, \vec{v}_r$ are linearly independent. Proof: Proceed by induction. For the induction step, assume the claim holds for r and now consider it for $r+1$. Without loss of generality there is a dependence relation $\sum_1^r t_i\vec{v}_i = \vec{v}_{r+1}$. Applying T yields $\sum_1^r t_i\lambda_i\vec{v}_i = \lambda_{r+1}\vec{v}_{r+1} = \sum_1^r t_i\lambda_{r+1}\vec{v}_i$. Subtracting these yields $\sum_1^r t_i(\lambda_i - \lambda_{r+1})\vec{v}_i = \vec{0}$. Since $\lambda_i \neq \lambda_{r+1}$ for all i by assumption, and at least one t_i is nonzero, therefore this is a dependence relation for $\vec{v}_1, \ldots, \vec{v}_r$. But this is impossible by the inductive step.

B: EIGENBASIS

DEFINITION 0.4. *Let V be an n dimensional vector space and let $T : V \to V$ be a linear transformation. T is diagonalizable if there exists a basis $\mathcal{B}$ for V such that $[T]_{\mathcal{B}}$ is diagonal.*

DEFINITION 0.5. *Let V be an n dimensional vector space and let $T : V \to V$ be a linear transformation.. A basis $\mathcal{B} = \{b_1, \cdots, b_n\}$ of V is called an eigenbasis if all b_i's are eigenvectors of T.*

(1) Suppose $\mathcal{B} = \{b_1, \cdots, b_n\}$ is an eigenbasis for T with the corresponding eigenvalues $\lambda_1, \cdots, \lambda)n$. Write down $[T]_{\mathcal{B}}$.
(2) Suppose the matrix if T with respect to some basis $\mathcal{A}$ is diagonal. What can we say about the vectors in $\mathcal{A}$?
(3) Prove: T is diagonalizable if and only if V has an eigenbasis for T.
(4) Now read the process of diagonalizing matrices again. Think of A as the standard matrix of a linear transformation from F^n to itself. Explain how the steps 1,2 and 3 in diagonalization of matrices at the top of the worksheet related to finding an eigenbasis? What is D in the context of change of basis? What is another name for matrix P and P^{-1}?

Solution.
(1) Since $[T(b_i)]_{\mathcal{B}} = [\lambda_i b_i]_{\mathcal{B}} = \lambda_i e_i$, so $[T]_{\mathcal{B}}$ is the diagonal matrix with λ_i on the diagonal.
(2) If T with respect to some basis $\mathcal{A}$ is also diagonal, then $\mathcal{A}$ is just the rescale version of $\mathcal{B}$, which means every elements of $\mathcal{A}$ is a multiple of some element in $\mathcal{B}$
(3) If V has eigenbasis, from question(a) we see that the matrix representation with respect to $\mathcal{B}$ is diagonal. If T is diagonal, then by definition there are basis $\mathcal{B} = \{b_1, \ldots, b_n\}$ such that matrix representation with respect to $\mathcal{B}$ is diagonal.Then $[T(b_i)]_{\mathcal{B}} = \lambda_i[b_i]_{\mathcal{B}}$, so $\mathcal{B}$ is an eigenbasis.
(4) If we have $A = P^{-1}DP$, where D is a diagonal matrix, A is the standard matrix representation of a linear transformation, $\mathcal{B}$ is the corresponding eigenbasis. Then D

is the matrix representation of the transformation with respect to $\mathcal{B}$, P is the change of basis matrix from standard basis to eigenbasis $\mathcal{B}$.