Math 416 Lecture Note: Week 10

Daesung Kim

1 Diagonalization and eigenvectors

Definition 1.1. A matrix $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ is a diagonal matrix if $A_{ij} = 0$ for all $i \neq j$. In this case, we use the notation

$$A = \mathrm{diag}(A_{11}, A_{22}, \cdots, A_{nn}).$$

Example 1.2. A matrix $\begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$ is not diagonal but $\begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}$ is.

Definition 1.3. Let V be a finite dimensional vector space over $\mathbb{R}$. A linear transformation $T : V \to V$ is diagonalizable if there exists a basis β for V such that $[T]_\beta$ is diagonal.

Example 1.4. Let $A = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$ and consider $L_A : \mathbb{R}^2 \to \mathbb{R}^2$. Let $\beta = \{v_1 = (1, 1), v_2 = (-1, 1)\}$, then β is a basis for $\mathbb{R}^2$, $Av_1 = 8v_1$, and $Av_2 = 2v_2$. Thus, T is diagonalizable and

$$[L_A]_\beta = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}.$$

If T is diagonalizable and β is a basis for V such that

$$[T]_\beta = \mathrm{diag}(\lambda_1, \cdots, \lambda_n).$$

Then, we have $T(v_i) = \lambda_i v_i$ for all $v_i \in \beta$.

Definition 1.5. An eigenvector for a linear transformation $T : V \to V$ is a nonzero vector $v \in V$ such that $T(v) = \lambda v$ for some $\lambda \in \mathbb{R}$. The scalar λ is called the eigenvalue associated to v.

Remark 1.6. Note that the zero vector 0 is not a eigenvector for all λ.

Example 1.7. Let $A = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$ and consider $T = L_A$. One can see that $v = (1, 1)$ is an eigenvector and $T(v) = 2v$. Thus the eigenvalue associated to v is 2. If $w = (1, 2)$, then

$$T(w) = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 11 \\ 13 \end{pmatrix}.$$

Since the last vector is not a scalar multiple of w, w is not an eigenvector.

Theorem 1.8. *Let V be a finite dimensional vector space over $\mathbb{R}$. A linear transformation $T : V \to V$ is diagonalizable if and only if there exists a basis for V consisting of eigenvectors of V.*

Definition 1.9. A matrix $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ is diagonalizable if L_A is diagonalizable. An eigenvector for A is the one for L_A. (That is, a nonzero vector v is an eigenvalue for A if $Av = \lambda v$ for some $\lambda \in \mathbb{R}$.)

Remark 1.10. A matrix $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ is diagonalizable if and only if there exists an invertible matrix $Q \in \mathcal{M}_{n\times n}(\mathbb{R})$ such that $Q^{-1}AQ = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$. (That is, A is similar to a diagonal matrix.)

Remark 1.11. Why are we interested in the diagonalization of a matrix? This is because a diagonal matrix is very convenient for multiplication. For example, if we want to compute $A^k = A \cdots A$ (k times), it is usually very hard. But, if A is diagonal, say, $A = \mathrm{diag}(a_1, \cdots, a_n)$, then it is easy to see that

$$A^k = \mathrm{diag}(a_1^k, \cdots, a_n^k)$$

(HW). Suppose A is diagonalizable so that $Q^{-1}AQ = D = \mathrm{diag}(d_1, \cdots, d_n)$. In other words, $A = QDQ^{-1}$. One can see that

$$A^k = (QDQ^{-1})^k = QD^kQ^{-1}$$

(HW), which makes the computation a lot easier!

Example 1.12. Let $A = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$ and consider $L_A : \mathbb{R}^2 \to \mathbb{R}^2$. Let $\beta = \{e_1, e_2\}$ be the standard basis for $\mathbb{R}^2$ and $\beta' = \{v_1 = (1,1), v_2 = (-1,1)\}$. We have seen that

$$A = [L_A]_\beta = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}, \qquad [L_A]_{\beta'} = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}.$$

Thus, we have

$$[I_{\mathbb{R}^2}]_{\beta'}^{\beta} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} =: Q, \qquad [I_{\mathbb{R}^2}]_{\beta}^{\beta'} = Q^{-1}$$

and

$$[L_A]_{\beta'} = \mathrm{diag}(8,2) = [I_{\mathbb{R}^2}]_{\beta}^{\beta'} [L_A]_\beta [I_{\mathbb{R}^2}]_{\beta'}^{\beta} = Q^{-1}AQ.$$

Theorem 1.13. *Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, then the following are equivalent.*

(i) *$\lambda \in \mathbb{R}$ is an eigenvalue for A.*

(ii) *$\mathcal{N}(A - \lambda I_n) \neq \{0\}$.*

(iii) *$\det(A - \lambda I_n) = 0$.*

Example 1.14. Let $A = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$, then

$$A - 2I_n = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}, \qquad A - 8I_n = \begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix}.$$

Proof. By definition, $\lambda \in \mathbb{R}$ is an eigenvalue for A if and only if there exists a nonzero $v \in \mathbb{R}^n$ such that

$$\begin{aligned} Av = \lambda v &\Leftrightarrow (A - \lambda I_n)v = 0 \\ &\Leftrightarrow \mathcal{N}(A - \lambda I_n) \neq \{0\} \\ &\Leftrightarrow A - \lambda I_n \text{ is not one-to-one.} \\ &\Leftrightarrow A - \lambda I_n \text{ is not invertible. (By Dimension Theorem.)} \\ &\Leftrightarrow \det(A - \lambda I_n) = 0. \end{aligned}$$

□

Definition 1.15. The characteristic polynomial of $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ is $f(t) = \det(A - tI_n)$.

Example 1.16. Let $A = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$, then

$$f(t) = \det \begin{pmatrix} 5-t & 3 \\ 3 & 5-t \end{pmatrix} = (5-t)^2 - 9 = t^2 - 10t + 16 = (t-2)(t-8).$$

Theorem 1.17. *The eigenvalue of A is the root of its characteristic polynomial.*

2 Finding eigenvalues and eigenvectors

Question 2.1. How can we find the eigenvalues of a matrix A?

To answer the question, we recall the following.

(i) A scalar $\lambda \in \mathbb{R}$ is an eigenvalue for A if and only if $\det(A - \lambda I_n) = 0$.

(ii) The characteristic polynomial is $f(t) = \det(A - tI_n)$. So the roots of $f(t)$ are exactly the eigenvalues of A.

Thus, it suffices to find the roots of the characteristic polynomial of a matrix.

Example 2.2. Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 4 \end{pmatrix}$, then

$$f(t) = \det(A - tI_n) = \det \begin{pmatrix} 1-t & 2 & 0 \\ 0 & 2-t & -1 \\ 0 & 0 & 4-t \end{pmatrix} = (1-t)(2-t)(4-t).$$

Thus, the eigenvalues of A are 1,2, and 4.

Example 2.3. Let $A = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix}$, then

$$\begin{aligned} f(t) = \det(A - tI_n) &= \det \begin{pmatrix} 4-t & 0 & 1 \\ 2 & 3-t & 2 \\ 1 & 0 & 4-t \end{pmatrix} \\ &= (4-t)\det \begin{pmatrix} 3-t & 2 \\ 0 & 4-t \end{pmatrix} + \det \begin{pmatrix} 2 & 3-t \\ 1 & 0 \end{pmatrix} \\ &= (4-t)^2(3-t) - (3-t) \\ &= (3-t)((4-t)^2 - 1) \\ &= (3-t)^2(5-t). \end{aligned}$$

Thus, the eigenvalues of A are 3 and 5.

Example 2.4. Let $A = \begin{pmatrix} 0 & 3 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, then

$$f(t) = \det(A - tI_n) = \det \begin{pmatrix} -t & 3 & 1 \\ 0 & -t & 2 \\ 0 & 0 & -t \end{pmatrix} = -t^3.$$

Thus, 0 is the only eigenvalue of A.

It is important to note that $f(t)$ has degree n and the leading coefficient is $(-1)^n$. Thus, the characteristic polynomial has at most n real roots.

Question 2.5. How can we find the corresponding eigenvectors?

Last time, we have seen that a nonzero vector v is an eigenvector for a matrix A associated to λ if and only if v belongs the null space $\mathcal{N}(A - \lambda I_n)$. Thus, it suffice to find the null space $\mathcal{N}(A - \lambda I_n)$ to answer the question.

Example 2.6. Let $A = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix}$, then the eigenvalues of A are 3 and 5. By row operations, we have

$$A - 3I_n = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Thus, $\mathcal{N}(A - 3I_n) = \{(t, s, -t) : t, s \in \mathbb{R}\}$. That is, every nonzero vector $(t, s, -t)$ for $t, s \in \mathbb{R}$ is the eigenvector for A associated to 3. Similarly,

$$A - 5I_n = \begin{pmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Thus, $\mathcal{N}(A - 5I_n) = \{(t, 2t, t) : t \in \mathbb{R}\}$. That is, every nonzero vector $t(1, 2, 1)$ for $t \in \mathbb{R} \setminus \{0\}$ is the eigenvector for A associated to 5.

Definition 2.7. Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ and $\lambda \in \mathbb{R}$ be an eigenvalue for A. The eigenspace of A corresponding to λ is defined by

$$E_\lambda = \mathcal{N}(A - \lambda I_n).$$

Remark 2.8. Note that $E_\lambda = \mathcal{N}(A - \lambda I_n)$ is a subspace of $\mathbb{R}^n$. If $\lambda_1 \neq \lambda_2$, then $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$. (Homework.)

Example 2.9. Let $A = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix}$, then the eigenvalues of A are 3 and 5 and

$$E_3 = \mathcal{N}(A - 3I_n) = \{(t, s, -t) : t, s \in \mathbb{R}\}, \qquad E_5 = \mathcal{N}(A - 5I_n) = \{(t, 2t, t) : t \in \mathbb{R}\}.$$

Let $v_1 = (1, 0, -1) \in E_3$, $v_2 = (0, 1, 0) \in E_3$, and $v_3 = (1, 2, 1) \in E_5$. One can see that $\beta = \{v_1, v_2, v_3\}$ is a basis for $\mathbb{R}^3$. Then, A is diagonalizable and

$$[L_A]_\beta = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

This leads to

$$A = [I_{\mathbb{R}^3}]_\beta^\gamma [L_A]_\beta [I_{\mathbb{R}^3}]_\gamma^\beta = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix}^{-1}$$

where γ is the standard basis for $\mathbb{R}^3$.

Definition 2.10. Let V be a finite dimensional vector space over $\mathbb{R}$ and $T : V \to V$ linear. Let $\lambda \in \mathbb{R}$ be an eigenvalue for T. The eigenspace of T corresponding to λ is defined by

$$E_\lambda = \mathcal{N}(T - \lambda I_V) = \{v \in V : T(v) = \lambda v\}.$$

Theorem 2.11. *Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ and $v_1, \cdots, v_k$ be eigenvectors for A corresponding to eigenvalues $\lambda_1, \cdots, \lambda_k$. If λ_i are distinct, then $\{v_1, \cdots, v_k\}$ is linearly independent.*

Proof. Use an induction on k. If $k = 1$, then $\{v_1\}$ is linearly independent because every eigenvector is not zero. Suppose $k \geq 2$ and the theorem is true for $k - 1$. Let $v_1, \cdots, v_k$ be eigenvectors for A corresponding to disting eigenvalues $\lambda_1, \cdots, \lambda_k$. To see $\{v_1, \cdots, v_k\}$ is linearly independent, we consider

$$a_1 v_1 + \cdots + a_k v_k = 0.$$

Let $T = L_A$, then we have

$$\begin{aligned}(T - \lambda_k I_n)(a_1 v_1 + \cdots + a_k v_k) &= a_1(T - \lambda_k I_n)(v_1) + \cdots + a_k(T - \lambda_k I_n)(v_k) \\ &= a_1(\lambda_1 - \lambda_k)v_1 + \cdots + a_k(\lambda_k - \lambda_k)v_k \\ &= a_1(\lambda_1 - \lambda_k)v_1 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} \\ &= 0.\end{aligned}$$

Since $\{v_1, \cdots, v_{k-1}\}$ is linearly independent by the induction hypothesis, we have

$$a_1(\lambda_1 - \lambda_k) = \cdots = a_{k-1}(\lambda_{k-1} - \lambda_k) = 0.$$

Since $\lambda_1, \cdots, \lambda_k$ are distinct, we conclude $a_1 = \cdots = a_{k-1} = 0$. Since $v_k \neq 0$, we get $a_k = 0$ as desired. $\square$

3 Diagonalizability

Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ and $\lambda_1, \cdots, \lambda_k$ be distinct eigenvalues for A. Let $v_i \in E_{\lambda_i}$ for each $i = 1, 2, \cdots, k$, then we have seen last time that $\{v_1, \cdots, v_k\}$ is linearly independent. This leads to the following.

Corollary 3.1. *If $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ has n distinct eigenvalues for A, then A is diagonalizable.*

Example 3.2. Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 4 \end{pmatrix}$, then

$$f(t) = \det(A - tI_n) = \det \begin{pmatrix} 1-t & 2 & 0 \\ 0 & 2-t & -1 \\ 0 & 0 & 4-t \end{pmatrix} = (1-t)(2-t)(4-t).$$

Thus, the eigenvalues of A are 1,2, and 4. Since the eigenvalues are distinct, A is diagonalizable.

Indeed, we can generalize it in the following way.

Theorem 3.3. *Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ and $\lambda_1, \cdots, \lambda_k$ be distinct eigenvalues for A. For each $i = 1, 2, \cdots, k$, let β_i be a linearly independent subset of E_{λ_i} and $\beta = \beta_1 \cup \cdots \cup \beta_k$. Then, β is linearly independent.*

Proof. Let $\beta_i = \{v_1^i, \cdots, v_{d_i}^i\}$ for each $i = 1, 2, \cdots, k$. Note that $\beta_i \cap \beta_j = \varnothing$ for $i \neq j$ because $E_{\lambda_i} \cap E_{\lambda_j} = \{0\}$. Suppose

$$(a_1^1 v_1^1 + \cdots + a_{d_1}^1 v_{d_1}^1) + (a_1^2 v_1^2 + \cdots + a_{d_2}^2 v_{d_2}^2) + \cdots + (a_1^k v_1^k + \cdots + a_{d_k}^k v_{d_k}^k) = 0.$$

We note that if $w_i = a_1^i v_1^i + \cdots + a_{d_i}^i v_{d_i}^i \neq 0$ then w_i is an eigenvector of A associated to λ_i. Since $\{w_i : w_i \neq 0\}$ is linearly independent, $w_1 + \cdots + w_k = 0$ implies $w_1 = \cdots = w_k = 0$. Since β_i is linearly independent, $w_i = 0$ implies $a_1^i = \cdots = a_{d_i}^i = 0$ for all i. Thus, β is linearly independent. □

Corollary 3.4. *Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ and $\lambda_1, \cdots, \lambda_k$ be distinct eigenvalues for A. Let $d_i = \dim(E_{\lambda_i})$ for each $i = 1, 2, \cdots, k$. Then A is diagonalizable if and only if $n = d_1 + \cdots + d_k$.*

We will now consider when $n = \dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k})$ holds. To this end, we study the roots of the characteristic polynomial.

Definition 3.5. Let F be a field. (For example, $F = \mathbb{R}$ or $F = \mathbb{C}$.) A polynomial $f(t)$ in $\mathcal{P}(F)$ (the set of all polynomial where the coefficients are in $\mathbb{F}$) splits over F if

$$f(t) = c(t - a_1) \cdots (t - a_n)$$

where $c, a_1, \cdots, a_n \in F$.

Example 3.6. (i) $t^2 - 1$ splits over $\mathbb{R}$ because $t^2 - 1 = (t-1)(t+1)$.

(ii) $t^2 + 1$ does not split over $\mathbb{R}$, but splits over $\mathbb{C}$ because $t^2 + 1 = (t+i)(t-i)$.

Theorem 3.7 (Fundamental theorem of algebra)**.** *Every polynomial in $\mathcal{P}(\mathbb{C})$ splits over $\mathbb{C}$.*

Theorem 3.8. *If a matrix $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ is diagonalizable, then the characteristic polynomial splits over $\mathbb{R}$.*

Proof. If $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ is diagonalizable, then there exists an invertible matrix Q such that $A = QDQ^{-1}$ where $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$. Then, we have

$$(A - tI_n) = QDQ^{-1} - tQQ^{-1} = Q(D - tI_n)Q^{-1}$$

and so

$$\begin{aligned} f(t) = \det(A - tI_n) &= \det(Q(D - tI_n)Q^{-1}) = \det(D - tI_n) \\ &= (\lambda_1 - t) \cdots (\lambda_n - t) = (-1)^n (t - \lambda_1) \cdots (t - \lambda_n). \end{aligned}$$

□

Remark 3.9. Similar matrices have the same characteristic polynomial.

Example 3.10. Let $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, then

$$f(t) = \det \begin{pmatrix} t & 1 \\ -1 & t \end{pmatrix} = t^2 + 1.$$

Since $f(t)$ does not split over $\mathbb{R}$, A is not diagonalizable.

The converse is not true in general. That is, even though the characteristic polynomial splits over $\mathbb{R}$, the corresponding matrix may not be diagonalizable.

Example 3.11. Let $A = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 1 \end{pmatrix}$, then

$$f(t) = \det(A - tI_n) = \det \begin{pmatrix} -t & 1 & 2 \\ 0 & -t & 3 \\ 0 & 0 & 1-t \end{pmatrix} = -t^2(t-1).$$

Thus, $f(t)$ splits over $\mathbb{R}$ and 0 and 1 are the eigenvalues of A. But, it turns out that A is not diagonalizable. This is because we have

$$E_0 = \mathcal{N}(A) = \{t(1,0,0) : t \in \mathbb{R}\}, \qquad E_1 = \mathcal{N}(A - I_n) = \{t(5,3,1) : t \in \mathbb{R}\}.$$

and so the eigenvectors are not enough to form a basis for $\mathbb{R}^3$ (we need three linearly independent eigenvectors but $\dim(E_0) = \dim(E_1) = 1$).

Definition 3.12. Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ and λ be an eigenvalue for A. The algebraic multiplicity of λ is the largest positive integer k for which $(t-\lambda)^k$ is a factor of the characteristic polynomial $f(t)$. We denote by $\mathrm{m}_{\mathrm{alg}}(\lambda)$.

Definition 3.13. Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ and λ be an eigenvalue for A. The geometric multiplicity of λ is $\dim(E_\lambda)$. We denote by $\mathrm{m}_{\mathrm{geo}}(\lambda)$.

Lemma 3.14. *Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ and λ be an eigenvalue for A, then $1 \le \mathrm{m}_{\mathrm{geo}}(\lambda) \le \mathrm{m}_{\mathrm{alg}}(\lambda)$.*

Proof. Let $d = \mathrm{m}_{\mathrm{geo}}(\lambda) = \dim(E_\lambda)$ and $\{v_1, \cdots, v_d\}$ be a basis for E_λ. We extend it to a basis $\beta = \{v_1, \cdots, v_d, \cdots, v_n\}$ for $\mathbb{R}^n$, then

$$[L_A]_\beta = \begin{pmatrix} \lambda I_d & B \\ O & C \end{pmatrix}.$$

Thus, we get

$$\begin{aligned} f(t) &= \det(A - tI_n) = \det([L_A]_\beta - tI_n) \\ &= \det \begin{pmatrix} (\lambda - t)I_d & B \\ O & C - tI_{n-d} \end{pmatrix} \\ &= \det((\lambda - t)I_d)\det(C - tI_{n-d}) \\ &= (\lambda - t)^d \det(C - tI_{n-d}) \end{aligned}$$

(by HW). Since $(\lambda - t)^d$ divides $f(t)$, we obtain $\mathrm{m}_{\mathrm{geo}}(\lambda) \le \mathrm{m}_{\mathrm{alg}}(\lambda)$. $\square$

Theorem 3.15. *A matrix $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ is diagonalizable if and only if*

(i) *the characteristic polynomial of A splits over $\mathbb{R}$, and*

(ii) *for every eigenvalue λ, $\mathrm{m}_{\mathrm{geo}}(\lambda) = \mathrm{m}_{\mathrm{alg}}(\lambda)$.*

Proof. Suppose that A is diagonalizable, then we have seen that the characteristic polynomial of A splits over $\mathbb{R}$. Let β be a basis consisting of eigenvectors of A and b_i the number of vectors in β and E_{λ_i}. Then, $\sum_i b_i = n$. Since $\sum_i \mathrm{m_{alg}}(\lambda_i) = n$ and

$$b_i \leq \mathrm{m_{geo}}(\lambda_i) \leq \mathrm{m_{alg}}(\lambda_i),$$

we get $\mathrm{m_{geo}}(\lambda_i) = \mathrm{m_{alg}}(\lambda_i)$ for all i.

Suppose that the characteristic polynomial of A splits over $\mathbb{R}$ and $\mathrm{m_{geo}}(\lambda_i) = \mathrm{m_{alg}}(\lambda_i)$ for each $i = 1, 2, \cdots, k$. Then,

$$n = \sum_i \mathrm{m_{alg}}(\lambda_i) = \sum_i \mathrm{m_{geo}}(\lambda_i) = \dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k}),$$

which implies that A is diagonalizable. $\square$

References

[FIS] Freidberg, Insel, and Spence, *Linear Algebra*, 4th edition, 2002.

[Bee] Beezer, *A First Course in Linear Algebra*, Version 3.5, 2015.

Department of Mathematics, University of Illinois at Urbana-Champaign
E-mail address: daesungk@illinois.edu