5.1: Eigenvalues and Eigenvectors.

Definition: A linear operator T on a finite-dimensional vector space V is called diagonalizable if $\exists$ an ordered basis β for V such that $[T]_\beta$ is a diagonal matrix.

A square matrix A is called diagonalizable if L_A is diagonalizable

Recap: matrix representation $[T]_\beta$, where $\beta = \{1, x, x^2\}$ of $P_2(\mathbb{R})$

Definition: Let $\beta = \{v_1, v_2, \dots, v_n\}$ be an ordered basis for a finite dimensional vector space V. For $x \in V$, let $a_1, a_2, \dots, a_n$ be the unique scalars s.t

$$x = \sum_{i=1}^{n} a_i v_i$$

we define the coordinate vector of x relative to β, denoted $[x]_\beta$, by

$$[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

Example.

Let $V = P_2(\mathbb{R})$, $\beta = \{\underset{v_1}{1}, \underset{v_2}{x}, \underset{v_3}{x^2}\}$: standard ordered basis, $f(x) = 4 + 6x - 7x^2$

$$[f]_\beta = \begin{array}{c|c} \begin{pmatrix} 4 \\ 6 \\ -7 \end{pmatrix} & \begin{matrix} a_1 \\ a_2 \\ a_3 \end{matrix} \end{array} \quad f(x) = \sum a_i v_i = 4(1) + 6(x) - 7(x^2)$$

Recap: L_A: left multiplication transformation L_A, where A is an $m\times n$ matrix. L_A is probably the most important tool for transfering properties about transformation to analogous properties about matrices and vice versa. For example, we use it to prove that matrix multiplication is associative

Definition: Let A be $m\times n$ matrix with entries from a field F. We denote by L_A the mapping $L_A: F^n \to F^m$ defined by $L_A(x) = Ax$ for each column vector $x \in F^n$.

Ex) Let $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$, then $A \in M_{2\times 3}(\mathbb{R})$ and $L_A: \mathbb{R}^3 \to \mathbb{R}^2$. If $x = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}$,

then,

$$L_A(x) = Ax = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \end{pmatrix}$$

Theorem 2.15: Let A be an $m\times n$ matrix with entries from F. Then $L_A: F^n \to F^m$ is linear. Furthermore, if B is any other $m\times n$ matrix and β and γ are the standard ordered bases for F^n and F^m, respectively, then we have the following properties

(a) $[L_A]_\beta^\gamma = A$

(b) $L_A = L_B$ iff $A = B$

(c) $L_{A+B} = L_A + L_B$ and $L_{aA} = aL_A \;\; \forall a \in F$

Computing $[L_A]_\beta$. Consider $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, $\beta = \left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\}$

$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3\cdot\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 0\cdot\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 1\begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Hence $[L_A]_\beta = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$

If $\beta = \{v_1, v_2, \dots, v_n\}$ is an ordered basis for V such that $T(v_j) = \lambda_j v_j$ for some scalars $\lambda_1, \lambda_2, \dots, \lambda_n$, then clearly,

$$[T]_\beta = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Definition: Let T be a linear operator on a vector space V. A nonzero vector $v \in V$ is called eigenvector of T if $\exists$ a scalar λ s.t $T(v) = \lambda v$. The scalar λ is called eigenvalue corresponding to the eigenvector v.

Theorem 5.1: A linear operator T on a finite-dimensional vector space V is diagonalizable iff $\exists$ an ordered basis β for V consisting of eigenvectors of T. Furthermore, if T is diagonalizable, $\beta = \{v_1, v_2, \dots, v_n\}$ is an ordered basis of eigenvectors of T, and $D = [T]_\beta$ is a diagonal matrix and D_{jj} is the eigenvalue corresponding to v_j for $1 \le j \le n$.

Ex) Let $A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$, $v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$

$L_A(v_1) = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \end{pmatrix} = -2\begin{pmatrix} 1 \\ -1 \end{pmatrix} = -2v_1$, hence v_1 is an eigenvector of L_A, and hence of A. Here $\lambda_1 = -2$ corresponding to v_1.

Furthermore,

$L_A(v_2) = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 15 \\ 20 \end{pmatrix} = 5\begin{pmatrix} 3 \\ 4 \end{pmatrix} = 5v_2$, hence v_2 is an eigenvector of L_A and A, and $\lambda_2 = 5$.

Note that $\beta = \{v_1, v_2\}$ is an ordered basis for $\mathbb{R}^2$ consisting of eigenvectors of both A and L_A, and therefore A and L_A are diagonalizable.

$[L_A]_\beta = \begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix}$

Moreover, by Theorem 5.1 and its corollary, if

$Q = \begin{array}{c} \begin{pmatrix} 1 & 3 \\ -1 & 4 \end{pmatrix} \\ \;\; v_1 \;\; v_2 \end{array}$, then $Q^{-1}AQ = [L_A]_\beta = \begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix}$

Q: Q is the $n \times n$ matrix whose jth column is v_j for $j = 1, 2, \dots, n$.

D: $D = Q^{-1}AQ$ is a diagonal matrix such that D_{jj} is the eigenvalue of A corresponding to v_j.

Theorem 5.2: Let $A \in M_{n\times n}(F)$. Then a scalar λ is an eigenvalue of A iff $\det(A - \lambda I) = 0$

+ Definition: Let $A \in M_{n\times n}(F)$. The polynomial $f(t) = \det(A - tI_n)$ is called the characteristic polynomial of A.

Example: To find the eigenvalues of

$A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$,

we compute its characteristic polynomial:

$\det(A - tI_2) = \det\begin{pmatrix} 1-t & 1 \\ 4 & 1-t \end{pmatrix} = t^2 - 2t - 3 = (t-3)(t+1)$. Hence, only eigenvalues of A are 3 and -1.

Theorem 5.4 : Let $A \in M_{n\times n}(F)$, and let λ be an eigenvalue of A. Vector $v \in F^n$ is an eigenvector of A corresponding to λ iff $v \neq 0$ and $(A-\lambda I)v = 0$.

Example:

To find all the eigenvalues of the matrix $A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$, A has two eigenvalues, $\lambda_1 = 3$, and $\lambda_2 = -1$.

Let $B_1 = A - \lambda_1 I = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} - 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix}$.

Then, $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ is an eigenvector corresponding to $\lambda_1 = 3$ iff $x \neq 0$ and $(A - \lambda_1 I)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$.

$$\begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2x_1 + x_2 \\ 4x_1 - 2x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Clearly the set of all solutions to this equation is

$\left\{ t\begin{pmatrix} 1 \\ 2 \end{pmatrix} : t \in \mathbb{R} \right\}$.

Hence, x is an eigenvector corresponding to $\lambda_1 = 3$ iff $x = t\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ for some $t \neq 0$.

Now, suppose x is an eigenvector of A corresponding to $\lambda_2 = -1$.

Let $B_2 = A - \lambda_2 I = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} + 1\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}$

$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$.

$(A - \lambda_2 I)x = \begin{pmatrix} 2x_1 + x_2 \\ 4x_1 + 2x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$N(L_{B_2}) = \left\{ t\begin{pmatrix} 1 \\ -2 \end{pmatrix} : t \in \mathbb{R} \right\}$.

Hence, x is an eigenvector corresponding to $\lambda_2 = -1$ iff $x = t\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ for some $t \neq 0$.

$\left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right\}$ is basis for $\mathbb{R}^2$ consisting of eigenvectors of A. Thus, A is diagonalizable.

5.2: Diagonalizability

: We need a simple test to determine whether an operator or a matrix can be diagonalized, as well as a method for actually finding a basis of eigenvectors.

Definition: A polynomial $f(t)$ in $P(F)$ splits over F if $\exists$ scalars $c, a_1, \ldots, a_n$ (not necessarily distinct) in F such that
$$f(t) = c(t-a_1)(t-a_2)\cdots(t-a_n)$$

Example:

$t^2-1 = (t+1)(t-1)$ splits over $\mathbb{R}$

but $(t^2+1)(t-2)$ does not split over $\mathbb{R}$ because t^2+1 cannot be factored into a product of linear factors.

However, $(t^2+1)(t-2)$ does split over $\mathbb{C}$ because it factors into the product $(t+i)(t-i)(t-2)$.

Theorem 5.6: The characteristic polynomial of any diagonalizable linear operator on a vector space V over a field F splits over F.

Proof: Let T be a diagonalizable linear operator on the n-dimensional vector space V, and β be an ordered basis for V, such that $[T]_\beta = D$ is a diagonal matrix.

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

and let $f(t)$ be the characteristic polynomial of T. Then

$$f(t) = \det(D - tI) = \det\begin{pmatrix} \lambda_1 - t & 0 & \cdots & 0 \\ 0 & \lambda_2 - t & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n - t \end{pmatrix} = (\lambda_1 - t)(\lambda_2 - t)\cdots(\lambda_n - t)$$

☆ The converse of Theorem 5.6 is false; that is, the characteristic polynomial of T may split, but T need not to be diagonalizable

Definition: Let λ be an eigenvalue of a linear operator or matrix with characteristic polynomial $f(t)$. The multiplicity (sometimes called the algebraic multiplicity) of λ is the largest positive integer k for which $(t-\lambda)^k$ is a factor of $f(t)$.

Example:

Let $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 4 \end{pmatrix}$, which has characteristic polynomial $f(t) = -(t-3)^2(t-4)$. Hence $\lambda = 3$ is an eigenvalue of A with multiplicity 2, and $\lambda = 4$ is an eigenvalue of A with multiplicity 1.

Definition: Let T be a linear operator on a vector space V, and let λ be an eigenvalue of T.

Define $E_\lambda = \{x \in V : T(x) = \lambda x\} = N(T - \lambda I_V)$. The set E_λ is called the eigenspace of T corresponding to the eigenvalue λ.

Analogously, we define the eigenspace of a square matrix A corresponding to the eigenvalue λ to be the eigenspace of L_A corresponding to λ.

Clearly, E_λ is a subspace of V consisting of the zero vector and the eigenvectors of T corresponding to the eigenvalue λ.

Theorem 5.7: Let T be a linear operator on a finite-dimensional vector space V, and let λ be an eigenvalue of T having multiplicity m. Then $1 \le \dim(E_\lambda) \le m$.

Example: Let T be the linear operator on $\mathbb{R}^3$ defined by

$$T\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 4a_1 + a_3 \\ 2a_1 + 3a_2 + 2a_3 \\ a_1 + 4a_3 \end{pmatrix}.$$

Let β be the standard ordered basis for $\mathbb{R}^3$, then

$$[T]_\beta = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix},$$ and hence the characteristic polynomial T is

$$\det([T]_\beta - tI) = \det\begin{pmatrix} 4-t & 0 & 1 \\ 2 & 3-t & 2 \\ 1 & 0 & 4-t \end{pmatrix} = -(t-5)(t-3)^2$$

$\lambda_1 = 5$ with multiplicity 1

$\lambda_2 = 3$ with multiplicity 2.

$$\begin{pmatrix} 4-\lambda_1 & 0 & 1 \\ 2 & 3-\lambda_1 & 2 \\ 1 & 0 & 4-\lambda_1 \end{pmatrix}$$

$$E_{\lambda_1} = N(T - \lambda_1 I) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : \begin{pmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

E_{λ_1} is the solution space of the system of linear equations,

$$\begin{aligned} -x_1 \quad\quad + x_3 &= 0 \\ 2x_1 - 2x_2 + 2x_3 &= 0 \\ x_1 \quad\quad - x_3 &= 0 \end{aligned} \Rightarrow \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\},$$ is a basis for E_{λ_1}. Hence, $\dim(E_{\lambda_1}) = 1$

$$E_{\lambda_2} = N(T - \lambda_2 I) \quad \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$\begin{aligned} x_1 \quad + x_3 &= 0 \\ 2x_1 \quad + 2x_3 &= 0 \\ x_1 \quad + x_3 &= 0, \end{aligned} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = s\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \text{ for } s, t \in \mathbb{R}$$

It follows $\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$ is a basis for E_{λ_2}, and $\dim(E_{\lambda_2}) = 2 = m$

In this case, the multiplicity of each eigenvalue λ_i is equal to the dimension of the corresponding eigenspace E_{λ_i}.

Theorem 5.8: Let T be a linear operator on a finite-dimensional vector space V such that the characteristic of T splits. Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of T. Then

(a) T is diagonalizable iff the multiplicity of λ_i is equal to $\dim(E_{\lambda_i})$ $\forall i$

(b) If T is diagonalizable and β_i is an ordered basis for E_{λ_i} for each i, then $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ is an ordered basis for V consisting of eigenvectors of T.

Test for Diagonalizability

: Let T be a linear operator on an n-dimensional vector space V. Then T is diagonalizable iff both of the following conditions hold.

1. The characteristic polynomial of T splits.
2. For each eigenvalue λ of T, the multiplicity of λ equals $\text{nullity}(T-\lambda I)$, that is, the multiplicity of λ equals $n - \text{rank}(T-\lambda I)$.

Example:

Let $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} \in M_{3\times 3}(\mathbb{R})$, we test A for diagonalizability.

The characteristic polynomial of A is $\det(A-\lambda I) = -(t-4)(t-3)^2$, which splits, and so condition 1 is satisfied.

$\lambda_1 = 4$ with multiplicity 1.

$\lambda_2 = 3$ with multiplicity 2.

$A - \lambda_1 I = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, and $\text{rank}(A-\lambda_1 I) = 2$. Since $\dim(A) - \text{rank}(A-\lambda_1 I) = 1$ = multiplicity of λ_1, satisfied.

$A - \lambda_2 I = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, and $\text{rank}(A-\lambda_2 I) = 2$. $\dim(A) - \text{rank}(A-\lambda_2 I) = 1 \neq$ multiplicity of λ_2.

Thus, condition 2 fails for λ_2, and A is therefore not diagonalizable.