6
Diagonalization

Because diagonal matrices are especially easy to handle, we want to pinpoint when a linear operator is represented by a diagonal matrix.

6.0 Eigenbasis

When is the matrix of a linear operator diagonal? To find more tractable matrices associated to a linear operator, we focus on bases that yield diagonal matrices.

6.0.0 Problem. Find a diagonal matrix $\mathbf{\Lambda} := \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ that is similar to the matrix $\mathbf{A} := \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix}$.

Solution. Suppose that there exists an invertible matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P} = \mathbf{\Lambda}$. Let $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ in $\mathbb{Q}^2$ denote the columns of the matrix $\mathbf{P}$. Hence, we obtain

$$\begin{bmatrix} \mathbf{A}\,\boldsymbol{p}_1 & \mathbf{A}\,\boldsymbol{p}_2 \end{bmatrix} = \mathbf{A}\,\mathbf{P} = \mathbf{P}\,\mathbf{\Lambda} = \begin{bmatrix} \boldsymbol{p}_1 & \boldsymbol{p}_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} \lambda_1\,\boldsymbol{p}_1 & \lambda_2\,\boldsymbol{p}_2 \end{bmatrix} .$$

By comparing columns, we see that $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are eigenvectors for $\mathbf{A}$ and the diagonal entries in $\mathbf{\Lambda}$ are eigenvalues of $\mathbf{A}$; see [5.0.0]. By definition [5.1.2], the characteristic polynomial of the matrix $\mathbf{A}$ is

$$\begin{aligned} \mathrm{p}_{\mathbf{A}}(t) &= \det(t\,\mathbf{I} - \mathbf{A}) = \det\left(\begin{bmatrix} t & -1 \\ 2 & t-3 \end{bmatrix}\right) \\ &= t(t-3) + 2 = t^2 - 3\,t + 2 = (t-1)(t-2)\,, \end{aligned}$$

so the eigenvalues of $\mathbf{A}$ are 1 and 2; see [5.1.5]. By determining the reduced row echelon form of the relevant matrices, we find the corresponding eigenvectors:

$$\begin{aligned} \mathbf{I} - \mathbf{A} &= \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix} \xrightarrow[\sim]{r_2 \mapsto r_2 - 2\,r_1} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \quad \Rightarrow \quad \operatorname{Ker}(\mathbf{I} - \mathbf{A}) = \operatorname{Span}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right), \\ 2\,\mathbf{I} - \mathbf{A} &= \begin{bmatrix} 2 & -1 \\ 2 & -1 \end{bmatrix} \xrightarrow[\sim]{r_2 \mapsto r_2 - \ r_1} \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} \quad \Rightarrow \quad \operatorname{Ker}(2\,\mathbf{I} - \mathbf{A}) = \operatorname{Span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right). \end{aligned}$$

Hence, we obtain

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},$$

and we conclude that $\mathbf{A} \approx \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$. $\square$

6.0.1 Definition. A linear operator $T\colon V \to V$ is ***diagonalizable*** if there exists an ordered basis $\mathcal{B}$ for the vector space V such that the associated matrix $(T)_{\mathcal{B}}^{\mathcal{B}}$ of T relative to $\mathcal{B}$ is diagonal.

A square matrix is diagonalizable if it is similar to a diagonal matrix.

6.0.2 Theorem (Diagonalizability criterion)**.** *A linear operator T on a finite-dimensional $\mathbb{K}$-vector space V is diagonalizable if and only if T has* $\dim(V)$ *linearly independent eigenvectors.*

An ***eigenbasis*** is a basis consisting of eigenvectors. With this terminology, a linear operator on a finite-dimensional vector space is diagonalizable if and only if it has an eigenbasis.

Proof. Let $n := \dim(V)$. Choose an ordered basis $\mathcal{B}$ for the vector space V and consider the matrix $\mathbf{A} := (T)_{\mathcal{B}}^{\mathcal{B}}$. Since similar matrices represent the same linear operator relative to different ordered basis [4.2.7], it is enough to show that the matrix $\mathbf{A}$ is diagonalizable if and only if the matrix $\mathbf{A}$ has n linearly independent eigenvectors.

$\Rightarrow$: Suppose that there is an invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{\Lambda}$ such that $\mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P} = \mathbf{\Lambda}$ or equivalently $\mathbf{A}\,\mathbf{P} = \mathbf{P}\,\mathbf{\Lambda}$. It follows that, for each $1 \leqslant k \leqslant n$, the matrix $\mathbf{A}$ times the k-th column of $\mathbf{P}$ is the k-th diagonal entry of $\mathbf{\Lambda}$ times the k-th column of $\mathbf{P}$. By definition [5.0.0], the k-th column of $\mathbf{P}$ is an eigenvector of $\mathbf{A}$ with eigenvalue equal to the k-th diagonal entry of $\mathbf{\Lambda}$. Since $\mathbf{P}$ is invertible, the characterization of invertible matrices implies that these are n eigenvectors are linearly independent.

$\Leftarrow$: Suppose that the vectors $\boldsymbol{p}_1, \boldsymbol{p}_2 \dots, \boldsymbol{p}_n$ are linearly independent eigenvectors of the matrix $\mathbf{A}$. The characterization of invertible matrices establishes that the matrix $\mathbf{P}$, having these eigenvectors as its columns, is invertible. For all $1 \leqslant k \leqslant n$, let the scalar λ_k in $\mathbb{K}$ be the eigenvalue associated to the eigenvector $\boldsymbol{p}_k$. It follow that

$$\begin{aligned}
\mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P} &= \mathbf{P}^{-1}\begin{bmatrix} \mathbf{A}\,\boldsymbol{p}_1 & \mathbf{A}\,\boldsymbol{p}_2 & \cdots & \mathbf{A}\,\boldsymbol{p}_n \end{bmatrix} \\
&= \mathbf{P}^{-1}\begin{bmatrix} \lambda_1\,\boldsymbol{p}_1 & \lambda_2\,\boldsymbol{p}_2 & \cdots & \lambda_n\,\boldsymbol{p}_n \end{bmatrix} \\
&= \mathbf{P}^{-1}\begin{bmatrix} \boldsymbol{p}_1 & \boldsymbol{p}_2 & \cdots & \boldsymbol{p}_n \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \\
&= \mathbf{P}^{-1}\,\mathbf{P}\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.
\end{aligned}$$

□

6.0.3 Corollary (Sufficient condition for diagonalizable)**.** *Let V be a finite-dimensional $\mathbb{K}$-vector space. Any linear operator $T\colon V \to V$ with* $\dim(V)$ *distinct eigenvalues is diagonalizable.*

This condition is not necessary!

Proof. The eigenvectors corresponding the distinct eigenvalues are linearly independent [5.0.6], so the diagonalizablility criterion [6.0.2] implies that the linear operator T is diagonalizable. □

We next demonstrate that this sufficient condition is not necessary.

6.0.4 Problem. Is the matrix $\mathbf{B} := \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 2 \end{bmatrix}$ diagonalizable?

Solution. The characteristic polynomial of the matrix $\mathbf{B}$ is

$$\mathrm{p}_{\mathbf{B}}(t) = \det(t\,\mathbf{I} - \mathbf{B}) = \det\left(\begin{bmatrix} t-1 & -2 & 0 & 0 \\ 0 & t+1 & 0 & 0 \\ 0 & 0 & t-1 & 0 \\ 0 & 0 & 1 & t-2 \end{bmatrix}\right) = (t+1)(t-1)^2(t-2)\,,$$

so this (4×4)-matrix has only 3 distinct eigenvalues. We find the corresponding eigenvectors as follows:

$$-1\,\mathbf{I} - \mathbf{B} = \begin{bmatrix} -2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 1 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \operatorname{Ker}(-1\,\mathbf{I} - \mathbf{B}) = \operatorname{Span}\left(\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}\right),$$

$$1\,\mathbf{I} - \mathbf{B} = \begin{bmatrix} 0 & -2 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \operatorname{Ker}(1\,\mathbf{I} - \mathbf{B}) = \operatorname{Span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}\right),$$

$$2\,\mathbf{I} - \mathbf{B} = \begin{bmatrix} -3 & -2 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \operatorname{Ker}(2\,\mathbf{I} - \mathbf{B}) = \operatorname{Span}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right).$$

Since

$$\det\left(\begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\right) = -1 \neq 0\,,$$

the characterizations of determinants and invertible matrices show that the matrix $\mathbf{B}$ has 4 linearly independent eigenvectors. Therefore, the criterion for diagonalizability [6.0.2] proves that the matrix $\mathbf{B}$ is diagonalizable. □

6.0.5 Problem. Let V be the $\mathbb{R}$-vector space of trigonometric polynomials of degree at most 1. Is the linear operator $S \colon V \to V$ defined, for all f in V, by $S[f] := f(x + \pi/4)$ diagonalizable?

Solution. Let $\mathcal{T} := \big(1, \cos(x), \sin(x)\big)$ be canonical ordered basis of the $\mathbb{R}$-vector space V. The special values $\cos(\pi/4) = \sin(\pi/4) = 1/\sqrt{2}$ together with the angle sum formulae give

The angle sum formulae assert that $\sin(\theta+\phi) = \sin(\theta)\cos(\phi) + \sin(\phi)\cos(\theta)$, $\cos(\theta+\phi) = \cos(\theta)\cos(\phi) - \sin(\phi)\sin(\theta)$.

$$S[1] = 1 \qquad \Rightarrow \quad (1)_{\mathcal{T}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

$$S[\cos] = \cos\big(x + \tfrac{\pi}{4}\big) = \tfrac{1}{\sqrt{2}}\cos(x) - \tfrac{1}{\sqrt{2}}\sin(x) \;\Rightarrow\; (\cos)_{\mathcal{T}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

$$S[\sin] = \sin\big(x + \tfrac{\pi}{4}\big) = \tfrac{1}{\sqrt{2}}\cos(x) + \tfrac{1}{\sqrt{2}}\sin(x) \;\Rightarrow\; (\sin)_{\mathcal{T}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Since $(S)_{\mathcal{J}}^{\mathcal{T}} = \dfrac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}$, the characteristic polynomial of S is

$$\begin{aligned} \mathrm{p}_S(t) = \det((t\,\mathrm{id}_V - S)_{\mathcal{J}}^{\mathcal{T}}) &= \det \left(\begin{bmatrix} t-1 & 0 & 0 \\ 0 & t - \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & t - \frac{1}{\sqrt{2}} \end{bmatrix} \right) \\ &= (t-1)\left(\left(t - \tfrac{1}{\sqrt{2}}\right)^2 + \tfrac{1}{2} \right) \\ &= (t-1)\left(t - \tfrac{1}{\sqrt{2}}(1+\mathrm{i})\right)\left(t - \tfrac{1}{\sqrt{2}}(1-\mathrm{i})\right). \end{aligned}$$

Since $\dim(V) = 3$ and there are 3 distinct eigenvalues, we conclude that the linear operator S is diagonalizable [6.0.3]. □

Exercises

6.0.6 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. The identity and zero operators on any finite-dimensional vector space are both diagonalizable.

ii. Every linear operator on a finite-dimensional vector space is diagonalizable.

iii. A linear operator is diagonalizable if and only if its underlying vector space has a basis consisting of eigenvectors.

iv. A linear operator is diagonalizable if and only if the number distinct eigenvalues equals the dimension of the vector space.

v. When a linear operator is diagonalizable, the associated matrix relative to any ordered basis is diagonal.

6.0.7 Problem. Consider the linear operator $T \colon \mathbb{R}[t]_{\leqslant 2} \to \mathbb{R}[t]_{\leqslant 2}$ defined, for all p in $\mathbb{R}[t]_{\leqslant 2}$, by $T[p] := (1-t^2)p''(t) - t\,p'(t) + 2\,p(t)$. Show that T is diagonalizable and find an eigenbasis.

6.1 Eigenspaces

HOW DO WE FIND A LARGEST POSSIBLE COLLECTION OF LINEARLY INDEPENDENT EIGENVECTORS? Motivated by our diagonalizability criterion [6.0.2], we seek a maximal set of linearly independent eigenvectors for a given linear operator. To accomplish this task, we introduce the linear subspace associated to an eigenvalue.

6.1.0 Definition. For any scalar λ, the ***λ-eigenspace*** of linear operator is the span of all eigenvectors with eigenvalue λ. The dimension of the λ-eigenspace called the ***geometric multiplicity*** of λ.

When the scalar λ is not a eigenvalue, the λ-eigenspace is the zero linear subspace and its geometric multiplicity is 0.

6.1.1 Problem. Find a basis for the 2-eigenspace of $\mathbf{A} := \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$.

Solution. Since

$$2\,\mathbf{I} - \mathbf{A} = \begin{bmatrix} -2 & 1 & -6 \\ -2 & 1 & -6 \\ -2 & 1 & -6 \end{bmatrix} \sim \begin{bmatrix} -2 & 1 & -6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \Rightarrow \quad \mathrm{Ker}(2\,\mathbf{I} - \mathbf{A}) = \mathrm{Span}\left(\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 6 \\ 1 \end{bmatrix} \right),$$

so $\begin{bmatrix} 1 & 2 & 0 \end{bmatrix}^{\mathsf{T}}$ and $\begin{bmatrix} 0 & 6 & 1 \end{bmatrix}^{\mathsf{T}}$ are a basis for the 2-eigenspace. □

6.1.2 Lemma. *Each vector in the λ-eigenspace of a linear operator is either an eigenvector of the linear operator with eigenvalue λ or the zero vector.*

Proof. It suffices to prove that any nonzero linear combination of eigenvectors with eigenvalue λ is also an eigenvector with eigenvalue λ. Suppose that $\boldsymbol{v}$ and $\boldsymbol{w}$ are eigenvectors of a linear operator $T \colon V \to V$ both with eigenvalue λ. For any scalars c and d, we have $T[c\,\boldsymbol{v} + d\,\boldsymbol{w}] = c\,T[\boldsymbol{v}] + d\,T[\boldsymbol{w}] = \lambda(c\,\boldsymbol{v} + d\,\boldsymbol{w})$. When $c\,\boldsymbol{v} + d\,\boldsymbol{w} \neq \boldsymbol{0}$, this linear combination is an eigenvector with eigenvalue λ. □

The next proposition establishes that choosing a basis for each eigenspace produces a linearly independent set of eigenvectors.

6.1.3 Proposition. *For any linear operator on a finite-dimensional vector space, the union of any bases for its eigenspaces is linearly independent.*

Proof. Let V be an finite-dimensional $\mathbb{K}$-vector space. Consider the distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_\ell$ of a linear operator $T \colon V \to V$. For each $1 \leqslant j \leqslant \ell$, choose an ordered basis $\mathcal{B}_j := (\boldsymbol{v}_{j,1}, \boldsymbol{v}_{j,2}, \dots, \boldsymbol{v}_{j,d_j})$ for the λ_j-eigenspace of T. By definition, the eigenvalue λ_j has geometric multiplicity d_j. It suffices to prove that the union

$$\bigcup_{j=1}^{\ell} \mathcal{B}_j = \{\boldsymbol{v}_{1,1}, \boldsymbol{v}_{1,2}, \dots, \boldsymbol{v}_{1,d_1}, \boldsymbol{v}_{2,1}, \boldsymbol{v}_{2,2}, \dots, \boldsymbol{v}_{2,d_2}, \ \dots, \ \boldsymbol{v}_{\ell,1}, \boldsymbol{v}_{\ell,2}, \dots, \boldsymbol{v}_{\ell,d_\ell}\}$$

is a linearly independent set of vectors. Suppose that, for all $1 \leqslant j \leqslant \ell$ and all $1 \leqslant k \leqslant d_j$, there are scalars $c_{j,k}$ in $\mathbb{K}$ such that

$$\sum_{j=1}^{\ell} \sum_{k=1}^{d_j} c_{j,k}\, \boldsymbol{v}_{j,k} = \boldsymbol{0}\,.$$

Set $\boldsymbol{w}_j := \sum_{k=1}^{d_j} c_{j,k}\, \boldsymbol{v}_{j,k}$ for all $1 \leqslant j \leqslant \ell$, so $\boldsymbol{w}_1 + \boldsymbol{w}_2 + \cdots + \boldsymbol{w}_\ell = \boldsymbol{0}$. Since the vector $\boldsymbol{w}_j$ lies in the λ_j-eigenspace of the linear operator T, Lemma 6.1.2 establishes that $\boldsymbol{w}_j$ is either an eigenvector of T with eigenvalue λ_j or the zero vector. Since the $\lambda_1, \lambda_2, \dots, \lambda_\ell$ are distinct, the corresponding eigenvectors are linearly independent [5.0.6]. Hence, the equation $\boldsymbol{w}_1 + \boldsymbol{w}_2 + \cdots + \boldsymbol{w}_\ell = \boldsymbol{0}$ implies that, for all $1 \leqslant j \leqslant \ell$, we have $\boldsymbol{0} = \boldsymbol{w}_j = c_{j,1}\, \boldsymbol{v}_{j,1} + c_{j,2}\, \boldsymbol{v}_{j,2} + \cdots + c_{j,d_j}\, \boldsymbol{v}_{j,d_j}$. For all $1 \leqslant j \leqslant \ell$, we deduce that $c_{j,1} = c_{j,2} = \cdots = c_{j,d_j} = 0$ because the set $\mathcal{B}_j$ is linearly independent. We conclude that the union $\bigcup_{j=1}^{\ell} \mathcal{B}_j$ is also linearly independent. □

Although Proposition 6.1.3 creates sets of linearly independent eigenvectors, these collections can be small.

6.1.4 Problem. Fix a positive integer n and a scalar λ. Compute the geometric and algebraic multiplicity for the unique eigenvalue of the $(n \times n)$-matrix

$$\mathbf{J} := \begin{bmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}.$$

Solution. Since the determinant of a triangular matrix is the product of the entries on its diagonal, we have

$$\mathrm{p}_{\mathbf{J}}(t) = \det\left(\begin{bmatrix} t-\lambda & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & t-\lambda & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & t-\lambda & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & t-\lambda & -1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & t-\lambda \end{bmatrix}\right) = (t-\lambda)^n ,$$

so λ is the only eigenvalue and it has algebraic multiplicity n. Since

$$\lambda\,\mathbf{I} - \mathbf{J} = \begin{bmatrix} 0 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

and $\mathrm{Ker}(\lambda\,\mathbf{I} - \mathbf{J}) = \mathrm{Span}(\boldsymbol{e}_1)$, the geometric multiplicity of λ is 1. □

Remarkably, the eigenvalue of the matrix $\mathbf{J}$ has the maximal possible algebraic multiplicity and the minimal possible geometric multiplicity.

Geometric and algebraic multiplicity are related by an inequality.

6.1.5 Proposition (Multiplicity inequality). *For any linear operator on a finite-dimensional vector space, the geometric multiplicity of an eigenvalue is less than or equal to the algebraic multiplicity of the same eigenvalue.*

Proof. Fix a positive integer n. Let V be an n-dimensional $\mathbb{K}$-vector space. Consider a linear operator $T \colon V \to V$ and a scalar λ in $\mathbb{K}$. Choose an ordered basis $(\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_d)$ for the λ-eigenspace of the linear operator T. Extend this list of linearly independent vectors to an ordered basis $\mathcal{B} := (\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_d, \boldsymbol{v}_{d+1}, \boldsymbol{v}_{d+2}, \dots, \boldsymbol{v}_n)$ of V. For all $1 \leqslant k \leqslant d$, we have $T[\boldsymbol{v}_k] = \lambda\,\boldsymbol{v}_k$; for all $d+1 \leqslant k \leqslant n$, there are scalars $a_{1,k}, a_{2,k}, \dots, a_{n,k}$ such that $T[\boldsymbol{v}_k] = a_{1,k}\,\boldsymbol{v}_1 + a_{2,k}\,\boldsymbol{v}_2 + \cdots + a_{n,k}\,\boldsymbol{v}_n$. It follows that

$$(t\ \mathrm{id}_V - T)^{\mathcal{B}}_{\mathcal{B}} = \begin{bmatrix} t-\lambda & 0 & 0 & \cdots & 0 & -a_{1,d+1} & -a_{1,d+2} & \cdots & -a_{1,n} \\ 0 & t-\lambda & 0 & \cdots & 0 & -a_{2,d+1} & -a_{2,d+2} & \cdots & -a_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & t-\lambda & -a_{d,d+1} & -a_{d,d+2} & \cdots & -a_{d+2,n} \\ 0 & 0 & 0 & \cdots & 0 & t-a_{d+1,d+1} & -a_{d+1,d+2} & \cdots & -a_{d+1,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -a_{n,d+1} & -a_{n,d+2} & \cdots & t-a_{n,n} \end{bmatrix},$$

so $\det\big((t\ \mathrm{id}_V - T)^{\mathcal{B}}_{\mathcal{B}}\big) = (t-\lambda)^d\, q(t)$ where q is a polynomial of degree $n-d$. Therefore, the algebraic multiplicity of λ is at least d. □

Exercises

6.1.6 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. The 1-eigenspace of the identity operator on a vector space V is always equal to V.

ii. Every vector $\boldsymbol{v}$ in the λ-eigenspace of a linear operator T satisfies the equation $T[\boldsymbol{v}] = \lambda\,\boldsymbol{v}$.

iii. The geometric multiplicity of a scalar λ in $\mathbb{K}$ is positive if and only if λ is an eigenvalue.

iv. When the algebraic multiplicity of an eigenvalue is 1, its geometric multiplicity must also be 1.

6.1.7 Problem. Let V denote the $\mathbb{R}$-vector space of trigonometric polynomials having degree at most 1. Consider the linear operator $J \colon V \to V$ defined, for all f in V, by $\big(J[f]\big)(x) := \int_0^{\pi} f(x-t)\,dt$. Show that J is diagonalizable and find an eigenbasis.

6.2 Diagonalizability

How do we characterize diagonalizable linear operators? By consolidating our knowledge about eigenbases and eigenspaces, we completely describe diagonalizable linear operators.

6.2.0 Theorem (Characterization of diagonalizable operators)**.** *Let V be a finite-dimensional $\mathbb{K}$-vector space. For any linear operator $T \colon V \to V$, the following are equivalent:*

a. *the linear operator T is diagonalizable,*

b. *the union of any bases for the eigenspaces of T contains* $\dim(V)$ *vectors,*

c. *the sum of the algebraic multiplicities of all eigenvalues equals* $\dim V$ *and, for each eigenvalue, the algebraic and geometric multiplicities are equal.*

Proof. Set $n := \dim V$ and let the scalars $\lambda_1, \lambda_2, \dots, \lambda_\ell$ in $\mathbb{K}$ denote the distinct eigenvalues of the linear operator T. For all $1 \leqslant k \leqslant \ell$, we write d_k and m_k for the geometric and algebraic multiplicity of λ_k respectively.

a $\Rightarrow$ *b*: The diagonalizability criterion [6.0.2] implies that the linear operator T has n linearly independent eigenvectors. If d_k' of these eigenvectors correspond to the eigenvalue λ_k for all $1 \leqslant k \leqslant \ell$, then any basis for the λ_k-eigenspace contains at least d_k' vectors. Hence, the union of the bases of the eigenspaces contains at least n vectors. Since the Proposition 6.1.3 shows that this union is linearly independent, the union contains at most n vectors.

b $\Rightarrow$ *c*: Suppose that $n = d_1 + d_2 + \cdots + d_\ell$. For all $1 \leqslant k \leqslant \ell$, the characteristic polynomial $\mathrm{p}_T(t)$ is divisible by $(t-\lambda_k)^{m_k}$ and the degree of $\mathrm{p}_T(t)$ is n. It follows that $m_1 + m_2 + \cdots + m_\ell \leqslant n$

and $(m_1 - d_1) + (m_2 - d_2) + \cdots + (m_\ell - d_\ell) \leqslant 0$. Since the geometric multiplicity of a eigenvalue is at most its algebraic multiplicity [6.1.5], we have $m_k - d_k \geqslant 0$ for all $1 \leqslant k \leqslant \ell$. We deduce that $m_k = d_k$ for all $1 \leqslant k \leqslant \ell$ and $m_1 + m_2 + \cdots + m_\ell = n$.

c ⇒ *a*: The characteristic polynomial is a product of linear factors if and only if $m_1 + m_2 + \cdots + m_\ell = n$. Under the additional hypothesis that $d_k = m_k$ for all $1 \leqslant k \leqslant \ell$, the linear operator T has $d_1 + d_2 + \cdots + d_\ell = m_1 + m_2 + \cdots + m_\ell = n$ linearly independent eigenvectors [6.1.3]. Thus, the diagonalizability criterion [6.0.2] shows that the linear operator T is diagonalizable. □

6.2.1 Problem. Let $T\colon \mathbb{R}[x]_{\leqslant 2} \to \mathbb{R}[x]_{\leqslant 2}$ be the linear operator defined, for all p in $\mathbb{R}[x]_{\leqslant 2}$, by

$$T[p] := \tfrac{3}{8} \int_{-1}^{1} \big(1 + 4(x+t) + 5(x^2+t^2) - 15x^2t^2\big)\, p(t)\, dt\,.$$

Show that T is diagonalizable, find an eigenbasis, and describe T^{-1}.

Solution. Fix the monomial basis $\mathcal{M} := (1, x, x^2)$ for $\mathbb{R}[x]_{\leqslant 2}$. Since

$$\begin{aligned}
T[1] &= \tfrac{3}{8}\int_{-1}^{1} 1+4(x+t)+5(x^2+t^2)-15x^2t^2\, dt = \tfrac{3}{8}\Big[t+4xt+2t^2+5x^2t+\tfrac{5}{3}t^3-5x^2t^3\Big]_{t=-1}^{t=1} = 2+3x\,,\\
T[x] &= \tfrac{3}{8}\int_{-1}^{1} t+4xt+4t^2+5x^2t+5t^3-15x^2t^3\, dt = \tfrac{3}{8}\Big[\tfrac{1}{2}t^2+2xt^2+\tfrac{4}{3}t^3+\tfrac{5}{2}x^2t^2+\tfrac{5}{4}t^4-\tfrac{15}{4}x^2t^4\Big]_{t=-1}^{t=1} = 1\,,\\
T[x^2] &= \tfrac{3}{8}\int_{-1}^{1} t^2+4xt^2+4t^3+5x^2t^2+5t^4-15x^2t^4\, dt = \tfrac{3}{8}\Big[\tfrac{1}{3}t^3+\tfrac{4}{3}xt^3+t^4+\tfrac{5}{3}x^2t^3+t^5-3x^2t^5\Big]_{t=-1}^{t=1} = 1+x-x^2\,,
\end{aligned}$$

we have $(T)_{\mathcal{M}}^{\mathcal{M}} = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$. Hence, the characteristic polynomial is

$$\mathrm{p}_T(t) = \det\left(\begin{bmatrix} t-2 & -1 & -1 \\ -3 & t & -1 \\ 0 & 0 & t+1 \end{bmatrix}\right) = (t+1)\big((t-2)t-3\big) = (t+1)^2(t-3)\,.$$

Since

$$\begin{aligned}
(-\mathrm{id} + T)_{\mathcal{M}}^{\mathcal{M}} &= \begin{bmatrix} -3 & -1 & -1 \\ -3 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 3 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \Rightarrow \quad \operatorname{Ker}(-\mathrm{id} + T) = \operatorname{Span}(3x-1, 3x^2-1)\,,\\
(3\,\mathrm{id} + T)_{\mathcal{M}}^{\mathcal{M}} &= \begin{bmatrix} 1 & -1 & -1 \\ -3 & 3 & -1 \\ 0 & 0 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \Rightarrow \quad \operatorname{Ker}(3\,\mathrm{id} - T) = \operatorname{Span}(x+1)\,,
\end{aligned}$$

we see that $\mathcal{C} := (3x-1, 3x^2-1, x+1)$ forms an eigenbasis,

$$(T)_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad \text{and} \qquad (T^{-1})_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \tfrac{1}{3} \end{bmatrix}.$$

When $f := c_1(3x-1) + c_2(3x^2-1) + c_3(x+1)$, it follows that

$$\begin{aligned}
T[f] &= -c_1(3x-1) - c_2(3x^2-1) + (3)c_3(x+1)\\
T^{-1}[f] &= -c_1(3x-1) - c_2(3x^2-1) + \big(\tfrac{1}{3}\big)c_3(x+1)\,.
\end{aligned}$$

□

6.2.2 Problem. Fix two positive integers m and n such that $m \leqslant n$. Let $\mathbf{A}$ be an $(m \times n)$-matrix and let $\mathbf{B}$ be an $(n \times m)$-matrix. Prove the $(n \times n)$-matrix $\mathbf{B}\,\mathbf{A}$ has the same eigenvalues with the same algebraic multiplicity as the $(m \times m)$-matrix $\mathbf{A}\,\mathbf{B}$ together with an additional $n - m$ eigenvalues equal to 0.

Solution. Consider the following two block-matrix identities:

$$\begin{bmatrix} \mathbf{A}\,\mathbf{B} & \mathbf{0} \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{A}\,\mathbf{B} & \mathbf{A}\,\mathbf{B}\,\mathbf{A} \\ \mathbf{B} & \mathbf{B}\,\mathbf{A} \end{bmatrix} \qquad \begin{bmatrix} \mathbf{I} & \mathbf{A} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{B} & \mathbf{B}\,\mathbf{A} \end{bmatrix} = \begin{bmatrix} \mathbf{A}\,\mathbf{B} & \mathbf{A}\,\mathbf{B}\,\mathbf{A} \\ \mathbf{B} & \mathbf{B}\,\mathbf{A} \end{bmatrix}.$$

Since the matrix $\begin{bmatrix} \mathbf{I} & \mathbf{A} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$ is invertible, it follows that

$$\begin{bmatrix} \mathbf{I} & \mathbf{A} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}\,\mathbf{B} & \mathbf{0} \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{B} & \mathbf{B}\,\mathbf{A} \end{bmatrix}$$

Since the eigenvalues of these similar matrices are the same, the assertion follows. □

Exercises

6.2.3 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. A linear operator over a complex vector space is diagonalizable.
ii. A linear operator is diagonalizable if and only if the algebraic multiplicity of each eigenvalue equals its geometric multiplicity.
iii. A linear operator over a complex vector space is diagonalizable if and only if the algebraic and geometric multiplicity are equal for each eigenvalue.

6.2.4 Problem. Find all values of k for which the matrix $\mathbf{A} := \begin{bmatrix} 1 & 1 & k \\ 1 & 1 & k \\ 1 & 1 & k \end{bmatrix}$ is diagonalizable.