

9.4 Diagonalization

We can now answer the question of Chapter 4 in the best case. An operator has a diagonal matrix in some basis exactly when there are enough eigenvectors to form a basis, and the previous two sections give precise ways to count “enough”: by dimensions of eigenspaces, or by comparing geometric with algebraic multiplicities. This section proves that these tests agree, shows how to find the diagonalizing basis for a concrete matrix, and uses it to compute powers. The payoff is a closed formula for the Fibonacci numbers. We also meet the operators for which the method fails, and see that the field is part of the answer.

9.4.1 Diagonalizable operators

Chapter 2 introduced similarity to compare the matrices of one operator in different bases (Theorem 4.2.4), and asked for the simplest representative. Section 2 unwound the wish “ $[T]_{\mathcal{B}}^{\mathcal{B}}$ is diagonal” and found that it says exactly that every vector of $\mathcal{B}$ is an eigenvector. Diagonal matrices are the matrices we can compute with most easily: they add, multiply, power, and invert entry by entry. So the operators that admit such a basis deserve a name.

An operator is diagonalizable when some basis consists of vectors that the operator only stretches.

Definition 9.4.1. Diagonalizable

Let V be a finite-dimensional vector space over F . An operator $T \in \mathcal{L}(V)$ is **diagonalizable** if there **exists** a basis $\mathcal{B}$ of V such that $[T]_{\mathcal{B}}^{\mathcal{B}}$ is a diagonal matrix.

A matrix $\mathbf{A} \in M_n(F)$ is **diagonalizable over F** if it is similar to a diagonal matrix: there **exist** an **invertible** $\mathbf{P} \in M_n(F)$ and a diagonal $\mathbf{D} \in M_n(F)$ with $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$, equivalently $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$.

In words: for an operator, we are free to choose **any** basis, and we ask whether **one** of them makes the matrix diagonal; most bases will not. For a matrix, the change-of-basis matrix $\mathbf{P}$ and the diagonal entries of $\mathbf{D}$ must both have entries **in F** . The two notions match: $\mathbf{A}$ is diagonalizable over F if and only if $T_{\mathbf{A}} \in \mathcal{L}(F^n)$ is diagonalizable, by Theorem 4.2.4 (c), because the matrices of $T_{\mathbf{A}}$ in the bases of F^n are exactly the matrices similar to $\mathbf{A}$. Likewise, T is diagonalizable if and only if $[T]_{\mathcal{C}}^{\mathcal{C}}$ is diagonalizable for one, equivalently every, basis $\mathcal{C}$ of V , by Theorem 4.2.4 (a) and (b).

Examples.

- **Diagonal matrices** are diagonalizable, with $\mathbf{P} = \mathbf{I}$. So is every scalar operator $c \operatorname{id}_V$, whose matrix is $c\mathbf{I}$ in every basis.
- **The degenerate cases.** Every 1×1 matrix is diagonal. On $V = \{0\}$, the empty basis gives the empty matrix, which counts as diagonal. These cases are harmless, but they remind us that diagonalizability is a question about repeated eigenvalues and missing eigenvectors, which need room to occur.
- **A 2×2 matrix.** In Example (Eigenvalues and Eigenspaces of a 2×2 Matrix), $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ has eigenvectors $(2, 3)$ for 4 and $(1, -1)$ for -1 . They form a basis of $\mathbb{R}^2$, so $\mathbf{A}$ is diagonalizable, with $\mathbf{P} = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$ and $\mathbf{D} = \operatorname{diag}(4, -1)$; we check this below.
- **Projections and involutions.** A projection P satisfies $V = \operatorname{im} P \oplus \ker P$ with P the identity on $\operatorname{im} P$ and zero on $\ker P$ (Theorem 4.5.1). A basis of $\operatorname{im} P$ followed by a basis of

$\ker P$ is a basis $\mathcal{B}$ with $[P]_{\mathcal{B}}^{\mathcal{B}} = \text{diag}(1, \dots, 1, 0, \dots, 0)$. In characteristic not 2, an involution splits $V = E_+ \oplus E_-$ (Theorem 4.5.5), and gives $\text{diag}(1, \dots, 1, -1, \dots, -1)$ in the same way.

Nonexample by minimal change. Change the $(1, 2)$ -entry of $\mathbf{I}_2$ from 0 to 1, giving $\mathbf{J} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ over any field F . It is still triangular with the single eigenvalue 1, and still invertible.

What fails is the existence of a basis of eigenvectors: $E_1(\mathbf{J}) = \ker \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \text{span}(\mathbf{e}_1)$ is a line, so every eigenvector of $\mathbf{J}$ is a multiple of $\mathbf{e}_1$, and no two eigenvectors are independent. By the unwinding in Section 2, no basis $\mathcal{B}$ makes $[T_{\mathbf{J}}]_{\mathcal{B}}^{\mathcal{B}}$ diagonal. (Directly: if $\mathbf{P}^{-1}\mathbf{J}\mathbf{P} = \mathbf{D}$ were diagonal, its diagonal entries would be eigenvalues of $\mathbf{J}$, so $\mathbf{D} = \mathbf{I}$, and then $\mathbf{J} = \mathbf{P}\mathbf{I}\mathbf{P}^{-1} = \mathbf{I}$, which is false.)

Why this definition. We ask for similarity to a diagonal matrix, not equality, because the basis is ours to choose; a matrix like $\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ has nothing diagonal about it until the right basis is found. We insist that $\mathbf{P}$ and $\mathbf{D}$ have entries in F , because otherwise the word would not depend on the field and would hide a real phenomenon: the rotation $\mathbf{R} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has no real eigenvector, so it is not diagonalizable over $\mathbb{R}$, while over $\mathbb{C}$ it becomes $\text{diag}(i, -i)$ (Section 2). The name says what the definition promises: the operator can be **made** diagonal.

Warning. **Diagonalizability depends on the field, and even on its characteristic.**

The rotation $\mathbf{R}$ is diagonalizable over $\mathbb{C}$ but not over $\mathbb{R}$. A subtler case: $\mathbf{S} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is diagonalizable over $\mathbb{Q}$, with eigenvectors $(1, 1)$ for 1 and $(1, -1)$ for -1 . Over $\mathbb{F}_2$ the same matrix has $p_{\mathbf{S}} = x^2 - 1 = (x - 1)^2$, since $-1 = 1$, and $\mathbf{S} - \mathbf{I} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ has rank 1, so its only eigenvectors are the multiples of $(1, 1)$, and $\mathbf{S}$ is **not** diagonalizable over $\mathbb{F}_2$. Always say over which field.

9.4.2 The Diagonalization Theorem

How do we decide whether a basis of eigenvectors exists without searching for one? Eigenvectors for different eigenvalues are automatically independent (Theorem 9.2.2), so the only question is whether each eigenspace supplies enough of them. The theorem below lists five equivalent ways to say “enough”.

Theorem 9.4.1. Diagonalization Theorem

Let V be a finite-dimensional vector space over F with $\dim V = n \geq 1$, let $T \in \mathcal{L}(V)$, and let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of T (possibly $k = 0$). The following are equivalent:

- (a) T is diagonalizable;
- (b) V has a basis consisting of eigenvectors of T ;
- (c) $V = E_{\lambda_1}(T) \oplus \dots \oplus E_{\lambda_k}(T)$;
- (d) $\dim E_{\lambda_1}(T) + \dots + \dim E_{\lambda_k}(T) = n$;
- (e) p_T splits over F , and $g_T(\lambda_i) = a_T(\lambda_i)$ for every $i = 1, \dots, k$.

In that case, if $\mathcal{B}$ is a basis of eigenvectors, then $[T]_{\mathcal{B}}^{\mathcal{B}}$ is diagonal, its diagonal entries are the eigenvalues of the corresponding basis vectors, and each λ_i appears on the diagonal exactly $a_T(\lambda_i) = g_T(\lambda_i)$ times.

(For $k = 0$ the sum in (c) is $\{0\}$ and the sum in (d) is 0; all five conditions are then false, since $n \geq 1$.)

Idea. Plan: prove (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (b) $\Rightarrow$ (a).

① (a) $\Leftrightarrow$ (b) is the unwinding of columns from Section 2. ② From a basis of eigenvectors, the eigenspaces add up to V , and Corollary 9.2.3 makes the sum direct. ③ Dimensions add in a direct sum. ④ (d) $\Rightarrow$ (e) is a sandwich: $n = \sum g \leq \sum a \leq n$, so every inequality is an equality. ⑤ (e) $\Rightarrow$ (b) runs the sandwich backwards and glues bases of the eigenspaces together.

The hard work was done in Sections 2 and 3; this proof only assembles it.

Proof.

- (a) $\Leftrightarrow$ (b). Let $\mathcal{B} = (v_1, \dots, v_n)$ be a basis of V . By Definition 4.1.1, column j of $[T]_{\mathcal{B}}^{\mathcal{B}}$ is $[Tv_j]_{\mathcal{B}}$, and this column is $d_j \mathbf{e}_j$ if and only if $Tv_j = d_j v_j$. Hence $[T]_{\mathcal{B}}^{\mathcal{B}} = \text{diag}(d_1, \dots, d_n)$ if and only if $Tv_j = d_j v_j$ for every j . Since basis vectors are nonzero, this says that every v_j is an eigenvector of T , for the eigenvalue d_j . So a basis makes $[T]_{\mathcal{B}}^{\mathcal{B}}$ diagonal exactly when it consists of eigenvectors, which also proves the final statement about the diagonal entries, apart from the count.
- (b) $\Rightarrow$ (c). Let $\mathcal{B}$ be a basis of eigenvectors. Each vector of $\mathcal{B}$ lies in $E_{\lambda_i}(T)$ for some i , so $\mathcal{B} \subseteq E_{\lambda_1}(T) + \dots + E_{\lambda_k}(T)$. This sum is a subspace (Theorem 1.7.1) containing a spanning list of V , so it equals V . By Corollary 9.2.3 the sum is direct.
- (c) $\Rightarrow$ (d). By Theorem 1.7.5 ((a) $\Rightarrow$ (e)), $n = \dim V = \sum_i \dim E_{\lambda_i}(T)$.
- (d) $\Rightarrow$ (e). By Theorem 9.3.1, $g_T(\lambda_i) \leq a_T(\lambda_i)$ for each i . The eigenvalues are exactly the roots of p_T in F (Theorem 9.2.1), so by Theorem 6.4.3, $\sum_i a_T(\lambda_i) \leq \deg p_T = n$ (Theorem 7.10.2). With (d),

$$n = \sum_{i=1}^k g_T(\lambda_i) \leq \sum_{i=1}^k a_T(\lambda_i) \leq n.$$

Hence both inequalities are equalities. The second equality, $\sum_i a_T(\lambda_i) = \deg p_T$, means that p_T splits over F (Theorem 6.4.3). The first says $\sum_i (a_T(\lambda_i) - g_T(\lambda_i)) = 0$, a sum of nonnegative integers, so every term is 0 and $g_T(\lambda_i) = a_T(\lambda_i)$ for each i .

- (e) $\Rightarrow$ (b). Since p_T splits, Theorem 6.4.3 gives $\sum_i a_T(\lambda_i) = n$, and so $\sum_i g_T(\lambda_i) = n$ by (e). For each i let $\mathcal{B}_i$ be a basis of $E_{\lambda_i}(T)$, of length $g_T(\lambda_i)$, and let $\mathcal{B}$ be the list $\mathcal{B}_1, \dots, \mathcal{B}_k$. The sum $W = E_{\lambda_1}(T) + \dots + E_{\lambda_k}(T)$ is direct (Corollary 9.2.3), so by Theorem 1.7.5 ((a) $\Rightarrow$ (d)), $\mathcal{B}$ is a basis of W , and $\dim W = n = \dim V$. By Theorem 1.6.8, $W = V$. So $\mathcal{B}$ is a basis of V , and its vectors are eigenvectors, being nonzero vectors of eigenspaces.

Finally, suppose $\mathcal{B}$ is a basis of eigenvectors, and let m_i be the number of its vectors with eigenvalue λ_i . By (a) $\Leftrightarrow$ (b), $[T]_{\mathcal{B}}^{\mathcal{B}}$ is diagonal with λ_i appearing m_i times, so $p_T = \prod_i (x - \lambda_i)^{m_i}$ (the triangular case after Definition 7.10.1). Since the λ_i are distinct, $(x - \lambda_i)^{m_i}$ divides

p_T and the remaining factors do not vanish at λ_i , so $m_i = a_T(\lambda_i)$ by Lemma 6.4.1; and $a_T(\lambda_i) = g_T(\lambda_i)$ by (e). This proves the theorem. ■

This is the theorem promised in Chapter 1, §7: **an operator is diagonalizable exactly when its eigenspaces split V as a direct sum**, and then T acts on the i -th piece as the scalar λ_i . In the language of Theorem 9.1.5, the pieces are invariant, the blocks are $\lambda_i \mathbf{I}_{g(\lambda_i)}$, and the matrix is block diagonal with scalar blocks.

Condition (e) is the practical test. It separates two different ways to fail, and they have different remedies. If p_T does not split, the field is too small, and enlarging it (from $\mathbb{R}$ to $\mathbb{C}$) may help. If p_T splits but some $g(\lambda) < a(\lambda)$, there are genuinely too few eigenvectors, and no change of field helps; Chapter 10 explains what the best matrix is then.

The easiest case of (e) needs no rank computation at all.

Corollary 9.4.2. Distinct Eigenvalues Imply Diagonalizable

Let V be a finite-dimensional vector space over F with $\dim V = n \geq 1$, and let $T \in \mathcal{L}(V)$ have n distinct eigenvalues. Then T is diagonalizable. The same holds for $\mathbf{A} \in M_n(F)$ with n distinct eigenvalues in F .

Proof. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues. Each λ_i is an eigenvalue, so $E_{\lambda_i}(T) \neq \{0\}$ (Definition 9.2.2) and $\dim E_{\lambda_i}(T) \geq 1$, so $\sum_i \dim E_{\lambda_i}(T) \geq n$, and by Corollary 9.2.3 the sum is at most n . Hence it equals n , and T is diagonalizable by Theorem 9.4.1 ((d) $\Rightarrow$ (a)). ■

So a matrix whose characteristic polynomial has n distinct roots in F is diagonalizable over F ; for example $\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$ is, because its eigenvalues 2 and 3 are distinct. This is the generic situation, but it is only a sufficient condition.

Warning. The converse fails: diagonalizable matrices may have repeated eigenvalues. $\mathbf{I}_n$ is diagonal, and its only eigenvalue is 1. Less trivially, the matrix of Example (Diagonalizing a 3×3 Matrix with a Repeated Eigenvalue) below has eigenvalues 1, 1, 3 and is diagonalizable. A repeated eigenvalue is a reason to **check** $g(\lambda) = a(\lambda)$, not a verdict.

Quick check. Which of $\begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$, $\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$, $\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$ are diagonalizable over $\mathbb{R}$?

Solution. The first has distinct eigenvalues 3, 5, so it is diagonalizable by Corollary 9.4.2. The

second has $a(3) = 2$ but $\text{rank} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1$, so $g(3) = 1 < a(3)$, and it is not diagonalizable by Theorem 9.4.1 (e). The third is already diagonal.

9.4.3 Computing $\mathbf{P}$ and $\mathbf{D}$

For a matrix $\mathbf{A} \in M_n(F)$, a basis of eigenvectors turns into the matrices $\mathbf{P}$ and $\mathbf{D}$ of Definition 9.4.1 directly. Let $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ be a basis of F^n with $\mathbf{A}\mathbf{v}_j = d_j\mathbf{v}_j$, let $\mathbf{P}$ be the matrix with columns $\mathbf{v}_1, \dots, \mathbf{v}_n$, and $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$. By the column view of matrix multiplication (Theorem 0.9.2), column j of $\mathbf{AP}$ is $\mathbf{A}\mathbf{v}_j = d_j\mathbf{v}_j$, and column j of $\mathbf{PD}$ is $\mathbf{P}(d_j\mathbf{e}_j) = d_j\mathbf{v}_j$. So

$$\mathbf{AP} = \mathbf{PD}.$$

The columns of $\mathbf{P}$ form a basis, so $\mathbf{P}$ is invertible (Theorem 3.4.4), and multiplying on the left by $\mathbf{P}^{-1}$ gives $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$. The order matters in one way only: the j -th column of $\mathbf{P}$ must be an eigenvector for the j -th diagonal entry of $\mathbf{D}$. Reordering the eigenvectors reorders the diagonal of $\mathbf{D}$ in the same way, and any order works.

For $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$, with $\mathbf{P} = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$ and $\mathbf{D} = \text{diag}(4, -1)$: $\mathbf{AP} = \begin{pmatrix} 8 & -1 \\ 12 & 1 \end{pmatrix}$ and $\mathbf{PD} = \begin{pmatrix} 8 & -1 \\ 12 & 1 \end{pmatrix}$. Checking $\mathbf{AP} = \mathbf{PD}$ avoids computing $\mathbf{P}^{-1}$, and together with $\det \mathbf{P} = -5 \neq 0$ it proves $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$.

The full procedure, then: compute $p_{\mathbf{A}}$ and its roots; if it does not split over F , stop (not diagonalizable over F); for each eigenvalue compute a basis of $E_{\lambda}(\mathbf{A})$; if some $g(\lambda) < a(\lambda)$, stop; otherwise put the basis vectors as columns of $\mathbf{P}$.

Example. Diagonalizing a 3×3 Matrix with a Repeated Eigenvalue

Let

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & -2 \\ -2 & -1 & 2 \\ -2 & -2 & 3 \end{pmatrix} \in M_3(\mathbb{R}).$$

Decide whether $\mathbf{A}$ is diagonalizable over $\mathbb{R}$. If it is, find an invertible $\mathbf{P}$ and a diagonal $\mathbf{D}$ with $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$.

Solution. Characteristic polynomial. Expanding along the first row (Theorem 7.6.1 over $\mathbb{R}[x]$),

$$\begin{aligned} p_{\mathbf{A}}(x) &= \det \begin{pmatrix} x-3 & -2 & 2 \\ 2 & x+1 & -2 \\ 2 & 2 & x-3 \end{pmatrix} \\ &= (x-3)[(x+1)(x-3) + 4] + 2[2(x-3) + 4] \\ &\quad + 2[4 - 2(x+1)]. \end{aligned}$$

The brackets are $x^2 - 2x + 1 = (x-1)^2$, $2(x-1)$, and $-2(x-1)$, so the last two terms cancel and

$$p_{\mathbf{A}}(x) = (x-1)^2(x-3).$$

It splits over $\mathbb{R}$, with $a(1) = 2$ and $a(3) = 1$. As a check, $\text{tr } \mathbf{A} = 5 = 1 + 1 + 3$ (Theorem 9.3.3).

Eigenspace for 1.

$$\mathbf{A} - \mathbf{I} = \begin{pmatrix} 2 & 2 & -2 \\ -2 & -2 & 2 \\ -2 & -2 & 2 \end{pmatrix}.$$

Every row is a multiple of $(1, 1, -1)$, so the rank is 1 and $g(1) = 3 - 1 = 2 = a(1)$. The system reduces to $x + y - z = 0$; with the free variables y and z , a basis of $E_1(\mathbf{A})$ is $(-1, 1, 0)$, $(1, 0, 1)$ (Theorem 3.5.4).

Eigenspace for 3. Here $g(3) = 1 = a(3)$ by Theorem 9.3.1, so condition (e) of Theorem 9.4.1 holds and $\mathbf{A}$ is diagonalizable. To find the eigenvector,

$$\mathbf{A} - 3\mathbf{I} = \begin{pmatrix} 0 & 2 & -2 \\ -2 & -4 & 2 \\ -2 & -2 & 0 \end{pmatrix}.$$

Row 1 gives $y = z$, and row 3 gives $x = -y$; row 2 then reads $2y - 4y + 2y = 0$. So $E_3(\mathbf{A}) = \text{span}((1, -1, -1))$.

The matrices. Put the eigenvectors as columns, in the order of the eigenvalues we want on the diagonal:

$$\mathbf{P} = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}, \quad \mathbf{D} = \text{diag}(1, 1, 3).$$

Check. $\mathbf{AP}$ has columns $\mathbf{A}(-1, 1, 0) = (-1, 1, 0)$, $\mathbf{A}(1, 0, 1) = (1, 0, 1)$, and $\mathbf{A}(1, -1, -1) = (3, -3, -3)$, which are the columns of $\mathbf{PD}$. Expanding along the first column, $\det \mathbf{P} = -1 \cdot (0 + 1) - 1 \cdot (-1 - 1) = 1 \neq 0$, so $\mathbf{P}$ is invertible, and in fact

$$\mathbf{P}^{-1} = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix},$$

as multiplying out $\mathbf{PP}^{-1} = \mathbf{I}$ confirms. Hence $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$.

Had we listed the eigenvectors in the order $(1, -1, -1), (-1, 1, 0), (1, 0, 1)$, we would get $\mathbf{D} = \text{diag}(3, 1, 1)$ instead. Both are correct.

9.4.4 Powers and polynomials

Why diagonalize? Powers of diagonal matrices are free: $\text{diag}(d_1, \dots, d_n)^k = \text{diag}(d_1^k, \dots, d_n^k)$. Similarity carries this over.

Theorem 9.4.3. Powers and Polynomials of a Diagonalizable Matrix

Let $\mathbf{A} = \mathbf{PDP}^{-1}$ with $\mathbf{P} \in M_n(F)$ invertible and $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$. Then for every $k \in \mathbb{N}$ and every $p \in F[x]$,

$$\mathbf{A}^k = \mathbf{P} \text{diag}(d_1^k, \dots, d_n^k) \mathbf{P}^{-1}, \quad p(\mathbf{A}) = \mathbf{P} \text{diag}(p(d_1), \dots, p(d_n)) \mathbf{P}^{-1}.$$

If moreover every $d_i \neq 0$, then $\mathbf{A}$ is invertible and $\mathbf{A}^{-1} = \mathbf{P} \text{diag}(d_1^{-1}, \dots, d_n^{-1}) \mathbf{P}^{-1}$.

Proof. Since $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, Proposition 4.2.6 (c) gives $p(\mathbf{D}) = \mathbf{P}^{-1}p(\mathbf{A})\mathbf{P}$, so $p(\mathbf{A}) = \mathbf{P}p(\mathbf{D})\mathbf{P}^{-1}$. A diagonal matrix is the direct sum $(d_1) \oplus \cdots \oplus (d_n)$ of 1×1 blocks, so by Theorem 8.1.4 (b), extended to n blocks by induction on the number of blocks, $p(\mathbf{D}) = (p(d_1)) \oplus \cdots \oplus (p(d_n)) = \text{diag}(p(d_1), \dots, p(d_n))$. Taking $p = x^k$ gives the formula for $\mathbf{A}^k$. If every $d_i \neq 0$, put $\mathbf{B} = \mathbf{P} \text{diag}(d_1^{-1}, \dots, d_n^{-1}) \mathbf{P}^{-1}$; then $\mathbf{A}\mathbf{B} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}\mathbf{P} \text{diag}(d_i^{-1}) \mathbf{P}^{-1} = \mathbf{P} \text{diag}(d_i d_i^{-1}) \mathbf{P}^{-1} = \mathbf{P}\mathbf{P}^{-1} = \mathbf{I}$, and $\mathbf{B} = \mathbf{A}^{-1}$ by Theorem 3.4.5. ■

In words: to apply a polynomial to a diagonalizable matrix, apply it to the eigenvalues and change basis back. The work is done once, when $\mathbf{P}$ is found; every power after that costs a single product of three matrices. For [Example \(Diagonalizing a \$3 \times 3\$ Matrix with a Repeated Eigenvalue\)](#),

$$\mathbf{A}^k = \mathbf{P} \text{diag}(1, 1, 3^k) \mathbf{P}^{-1} = \begin{pmatrix} 3^k & 3^k - 1 & 1 - 3^k \\ 1 - 3^k & 2 - 3^k & 3^k - 1 \\ 1 - 3^k & 1 - 3^k & 3^k \end{pmatrix},$$

which is $\mathbf{I}$ for $k = 0$ and $\mathbf{A}$ for $k = 1$, as it must be.

The most famous application is to a recurrence. A linear recurrence of order two becomes a first-order recurrence for vectors, whose solution is a matrix power.

Example. A Closed Formula for the Fibonacci Numbers

The Fibonacci numbers are defined by $F_0 = 0$, $F_1 = 1$, and $F_{k+2} = F_{k+1} + F_k$ for $k \geq 0$. (Here F_k is a number, not a field; we work over $\mathbb{R}$.) Find a formula for F_k by diagonalizing a 2×2 matrix.

Solution. A vector recurrence. Put $\mathbf{x}_k = (F_{k+1}, F_k) \in \mathbb{R}^2$. The recurrence says

$$\mathbf{x}_{k+1} = \begin{pmatrix} F_{k+2} \\ F_{k+1} \end{pmatrix} = \begin{pmatrix} F_{k+1} + F_k \\ F_{k+1} \end{pmatrix} = \mathbf{A}\mathbf{x}_k, \quad \mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

By induction on k , $\mathbf{x}_k = \mathbf{A}^k \mathbf{x}_0$ with $\mathbf{x}_0 = (1, 0)$.

Eigenvalues. $\text{tr } \mathbf{A} = 1$ and $\det \mathbf{A} = -1$, so $p_{\mathbf{A}}(x) = x^2 - x - 1$, with the two real roots

$$\varphi = \frac{1 + \sqrt{5}}{2}, \quad \psi = \frac{1 - \sqrt{5}}{2}.$$

They are distinct, so $\mathbf{A}$ is diagonalizable over $\mathbb{R}$ ([Corollary 9.4.2](#)). Each root satisfies $\lambda^2 = \lambda + 1$, and we will use $\varphi + \psi = 1$, $\varphi\psi = -1$, and $\varphi - \psi = \sqrt{5}$.

Eigenvectors. For a root λ , the vector $(\lambda, 1)$ is an eigenvector: $\mathbf{A}(\lambda, 1) = (\lambda + 1, \lambda) = (\lambda^2, \lambda) = \lambda(\lambda, 1)$. So

$$\mathbf{P} = \begin{pmatrix} \varphi & \psi \\ 1 & 1 \end{pmatrix}, \quad \mathbf{D} = \text{diag}(\varphi, \psi), \quad \mathbf{P}^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -\psi \\ -1 & \varphi \end{pmatrix},$$

where $\det \mathbf{P} = \varphi - \psi = \sqrt{5} \neq 0$ and $\mathbf{P}^{-1}$ comes from Theorem 0.9.8.

The power. By Theorem 9.4.3, $\mathbf{x}_k = \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1}\mathbf{x}_0$. Working from the right:

$$\begin{aligned}\mathbf{P}^{-1}\mathbf{x}_0 &= \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \\ \mathbf{D}^k\mathbf{P}^{-1}\mathbf{x}_0 &= \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi^k \\ -\psi^k \end{pmatrix}, \\ \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1}\mathbf{x}_0 &= \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi^{k+1} - \psi^{k+1} \\ \varphi^k - \psi^k \end{pmatrix}.\end{aligned}$$

The second entry of $\mathbf{x}_k$ is F_k , so

$$F_k = \frac{\varphi^k - \psi^k}{\sqrt{5}} = \frac{1}{\sqrt{5}} \left[\left(\frac{1 + \sqrt{5}}{2} \right)^k - \left(\frac{1 - \sqrt{5}}{2} \right)^k \right].$$

Check. $k = 0$: 0. $k = 1$: $(\varphi - \psi)/\sqrt{5} = 1$. $k = 2$: $(\varphi^2 - \psi^2)/\sqrt{5} = (\varphi - \psi)(\varphi + \psi)/\sqrt{5} = 1$. And the right side satisfies the recurrence, because $\varphi^{k+2} = \varphi^{k+1} + \varphi^k$ (multiply $\varphi^2 = \varphi + 1$ by φ^k), and likewise for ψ .

Since $|\psi| < 1$, the term $\psi^k/\sqrt{5}$ has absolute value less than $1/2$, so F_k is the integer nearest to $\varphi^k/\sqrt{5}$; the Fibonacci numbers grow like powers of $\varphi \approx 1.618$, the larger eigenvalue.

The irrational numbers in the formula are forced. Over $\mathbb{Q}$, the polynomial $x^2 - x - 1$ has no root (by Theorem 6.3.6, the only candidates are ± 1 , and neither works), so $\mathbf{A}$ has no eigenvalue in $\mathbb{Q}$ and is **not** diagonalizable over $\mathbb{Q}$, although all its entries and powers are integers. Chapter 7 met the same numbers as determinants of tridiagonal matrices (Example (Fibonacci Numbers as Determinants)); here the recurrence is solved rather than just recognized. Section 11 of this chapter treats general linear recurrences and Markov chains the same way.

9.4.5 When diagonalization fails

Two examples collect the ways Theorem 9.4.1 (e) can fail.

Too few eigenvectors. $\mathbf{J} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable over **any** field, because $p_{\mathbf{J}} = (x - 1)^2$ splits but $g(1) = 1 < 2 = a(1)$. Enlarging the field changes nothing: the rank of $\mathbf{J} - \mathbf{I}$ is 1 over every field. Its powers are still easy, $\mathbf{J}^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ by induction, but they are not of the form $\mathbf{P} \operatorname{diag}(1, 1) \mathbf{P}^{-1} = \mathbf{I}$; the entry k , growing linearly rather than like a power, is the footprint of the missing eigenvector. Chapter 10 shows that such blocks are the only additional ingredient.

Too small a field. The rotation $\mathbf{R} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has $p_{\mathbf{R}} = x^2 + 1$, which does not split over $\mathbb{R}$, so $\mathbf{R}$ is not diagonalizable over $\mathbb{R}$. Over $\mathbb{C}$ it has the distinct eigenvalues $\pm i$ with eigenvectors $(1, -i)$ and $(1, i)$, so with $\mathbf{P} = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$,

$$\mathbf{P}^{-1}\mathbf{R}\mathbf{P} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Then $\mathbf{R}^k = \mathbf{P} \operatorname{diag}(i^k, (-i)^k) \mathbf{P}^{-1}$, and since $i^4 = 1$, $\mathbf{R}^4 = \mathbf{I}$: four quarter-turns make a full turn, as they should.

When a matrix depends on a parameter, the same test decides diagonalizability for every value at once.

Example. Diagonalizability with a Parameter

For which $t \in \mathbb{R}$ is $\mathbf{A}_t = \begin{pmatrix} 2 & t & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ diagonalizable over $\mathbb{R}$?

Solution. The matrix is upper triangular, so $p_{\mathbf{A}_t}(x) = (x - 2)^2(x - 3)$ for every t , by Theorem 7.4.6. This splits, with $a(2) = 2$ and $a(3) = 1$. Since $1 \leq g(3) \leq a(3) = 1$ by Theorem 9.3.1, only $\lambda = 2$ needs checking. Here

$$\mathbf{A}_t - 2\mathbf{I} = \begin{pmatrix} 0 & t & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

whose rank is 1 if $t = 0$ and 2 if $t \neq 0$. By rank-nullity for matrices (Theorem 3.5.7), $g(2) = 3 - \text{rank}(\mathbf{A}_t - 2\mathbf{I})$ is 2 when $t = 0$ and 1 when $t \neq 0$. By Theorem 9.4.1 (e), $\mathbf{A}_t$ is diagonalizable exactly when $t = 0$.

Notice what did not change: the eigenvalues and their algebraic multiplicities are the same for every t . Only the geometric multiplicity, a rank, detects the difference.

9.4.6 Exercises

9.4.6.1 A. Check your understanding

Exercise. A1

- Define what it means for $\mathbf{A} \in M_n(F)$ to be diagonalizable over F .
- State the Diagonalization Theorem, with at least four equivalent conditions.
- True or false: a diagonalizable $n \times n$ matrix has n distinct eigenvalues. Justify your answer.
- True or false: every invertible matrix is diagonalizable. Justify your answer.
- True or false: every diagonalizable matrix is invertible. Justify your answer.
- If $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$ is diagonal, what are the columns of $\mathbf{P}$, and how are they matched with the diagonal of $\mathbf{D}$?

Solution.

- There are an invertible $\mathbf{P} \in M_n(F)$ and a diagonal $\mathbf{D} \in M_n(F)$ with $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$ (Definition 9.4.1).
- See Theorem 9.4.1: for $T \in \mathcal{L}(V)$, $\dim V = n \geq 1$, with distinct eigenvalues $\lambda_1, \dots, \lambda_k$, the following are equivalent: T is diagonalizable; V has a basis of eigenvectors; $V = \bigoplus_i E_{\lambda_i}(T)$; $\sum_i \dim E_{\lambda_i}(T) = n$; p_T splits and $g(\lambda_i) = a(\lambda_i)$ for all i .
- False. $\mathbf{I}_n$ is diagonal, with the single eigenvalue 1.

- (d) False. $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has determinant 1 but is not diagonalizable.
- (e) False. The zero matrix is diagonal and not invertible.
- (f) The columns of $\mathbf{P}$ form a basis of eigenvectors of $\mathbf{A}$, and column j is an eigenvector for the j -th diagonal entry of $\mathbf{D}$, since $\mathbf{AP} = \mathbf{PD}$.

9.4.6.2 B. Practice

Exercise. B1: Diagonalize or show impossible

For each real matrix, either find an invertible $\mathbf{P}$ and a diagonal $\mathbf{D}$ with $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$, or show that none exist.

- (a) $\mathbf{A}_1 = \begin{pmatrix} 0 & 1 \\ -6 & 5 \end{pmatrix}$.
- (b) $\mathbf{A}_2 = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ -3 & 3 & 2 \end{pmatrix}$.
- (c) $\mathbf{A}_3 = \begin{pmatrix} -1 & 2 & -2 \\ -2 & 3 & -2 \\ 2 & -2 & 3 \end{pmatrix}$.

Solution.

- (a) $\text{tr } \mathbf{A}_1 = 5$, $\det \mathbf{A}_1 = 6$, so $p = x^2 - 5x + 6 = (x - 2)(x - 3)$. The eigenvalues are distinct, so $\mathbf{A}_1$ is diagonalizable (Corollary 9.4.2). $\mathbf{A}_1 - 2\mathbf{I} = \begin{pmatrix} -2 & 1 \\ -6 & 3 \end{pmatrix}$ gives $y = 2x$, eigenvector $(1, 2)$; $\mathbf{A}_1 - 3\mathbf{I} = \begin{pmatrix} -3 & 1 \\ -6 & 2 \end{pmatrix}$ gives $y = 3x$, eigenvector $(1, 3)$. So

$$\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}, \quad \mathbf{D} = \text{diag}(2, 3), \quad \mathbf{P}^{-1} = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix},$$

$$\text{and indeed } \mathbf{A}_1\mathbf{P} = \begin{pmatrix} 2 & 3 \\ 4 & 9 \end{pmatrix} = \mathbf{PD}.$$

- (b) Expanding along the first row,

$$\begin{aligned} p(x) &= \det \begin{pmatrix} x-1 & -1 & -1 \\ -2 & x & -1 \\ 3 & -3 & x-2 \end{pmatrix} \\ &= (x-1)[x(x-2)-3] + [-2(x-2)+3] - [6-3x]. \end{aligned}$$

The first bracket is $(x-3)(x+1)$, and the last two terms add up to $-2x+7-6+3x =$

$x + 1$. So $p(x) = (x + 1)[(x - 1)(x - 3) + 1] = (x + 1)(x - 2)^2$, and $a(2) = 2$. Now

$$\mathbf{A}_2 - 2\mathbf{I} = \begin{pmatrix} -1 & 1 & 1 \\ 2 & -2 & 1 \\ -3 & 3 & 0 \end{pmatrix}.$$

Row 3 equals row 1 minus row 2, and rows 1 and 2 are not proportional, so the rank is 2 and $g(2) = 1 < 2 = a(2)$. By [Theorem 9.4.1](#) (e), $\mathbf{A}_2$ is **not** diagonalizable, over $\mathbb{R}$ or over $\mathbb{C}$.

(c) Expanding along the first row,

$$\begin{aligned} p(x) &= \det \begin{pmatrix} x+1 & -2 & 2 \\ 2 & x-3 & 2 \\ -2 & 2 & x-3 \end{pmatrix} \\ &= (x+1)[(x-3)^2 - 4] + 2[2(x-3) + 4] \\ &\quad + 2[4 + 2(x-3)]. \end{aligned}$$

The first bracket is $(x-5)(x-1)$ and the other two are $2(x-1)$ each, so $p(x) = (x-1)[(x+1)(x-5) + 8] = (x-1)(x^2 - 4x + 3) = (x-1)^2(x-3)$. For $\lambda = 1$,

$\mathbf{A}_3 - \mathbf{I} = \begin{pmatrix} -2 & 2 & -2 \\ -2 & 2 & -2 \\ 2 & -2 & 2 \end{pmatrix}$ has rank 1, so $g(1) = 2 = a(1)$, and E_1 is the plane $x - y + z = 0$, with basis $(1, 1, 0), (-1, 0, 1)$. Also $g(3) = 1 = a(3)$, so $\mathbf{A}_3$ is diagonalizable.

From $\mathbf{A}_3 - 3\mathbf{I} = \begin{pmatrix} -4 & 2 & -2 \\ -2 & 0 & -2 \\ 2 & -2 & 0 \end{pmatrix}$, row 3 gives $y = x$ and row 2 gives $z = -x$, so $E_3 = \text{span}((1, 1, -1))$. Hence

$$\mathbf{P} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}, \quad \mathbf{D} = \text{diag}(1, 1, 3).$$

Check: $\mathbf{A}_3(1, 1, 0) = (1, 1, 0)$, $\mathbf{A}_3(-1, 0, 1) = (-1, 0, 1)$, $\mathbf{A}_3(1, 1, -1) = (3, 3, -3)$, so $\mathbf{A}_3\mathbf{P} = \mathbf{PD}$; and $\det \mathbf{P} = 1(0 - 1) + 1(-1 - 0) + 1(1 - 0) = -1 \neq 0$.

Exercise. B2: A tenth power

Let $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \in M_2(\mathbb{R})$.

- (a) Diagonalize $\mathbf{A}$.
- (b) Hence find a formula for $\mathbf{A}^k$, $k \in \mathbb{N}$, and compute $\mathbf{A}^{10}$.

Solution.

(a) $\text{tr } \mathbf{A} = 2$, $\det \mathbf{A} = -3$, so $p = x^2 - 2x - 3 = (x - 3)(x + 1)$. $\mathbf{A} - 3\mathbf{I} = \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix}$

gives the eigenvector $(1, 1)$, and $\mathbf{A} + \mathbf{I} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$ gives $(1, -1)$. So $\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$,

$\mathbf{D} = \text{diag}(3, -1)$, and $\mathbf{P}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ by Theorem 0.9.8.

(b) By Theorem 9.4.3,

$$\mathbf{A}^k = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 3^k & 0 \\ 0 & (-1)^k \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 3^k + (-1)^k & 3^k - (-1)^k \\ 3^k - (-1)^k & 3^k + (-1)^k \end{pmatrix}.$$

For $k = 1$ this gives $\frac{1}{2} \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix} = \mathbf{A}$. With $3^{10} = 59049$ and $(-1)^{10} = 1$,

$$\mathbf{A}^{10} = \begin{pmatrix} 29525 & 29524 \\ 29524 & 29525 \end{pmatrix}.$$

Exercise. B3: A matrix with a parameter

For $t \in \mathbb{R}$, let $\mathbf{A}_t = \begin{pmatrix} 1 & 1 & 0 \\ 0 & t & 0 \\ 0 & 0 & 2 \end{pmatrix}$. Determine all t for which $\mathbf{A}_t$ is diagonalizable over $\mathbb{R}$. Justify your answer.

Solution. $\mathbf{A}_t$ is upper triangular, so $p(x) = (x - 1)(x - t)(x - 2)$, which splits over $\mathbb{R}$.

Case $t \notin \{1, 2\}$. The eigenvalues $1, t, 2$ are distinct, so $\mathbf{A}_t$ is diagonalizable (Corollary 9.4.2).

Case $t = 1$. Then $a(1) = 2$, and $\mathbf{A}_1 - \mathbf{I} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ has rank 2, so $g(1) = 1 < 2$. Not diagonalizable.

Case $t = 2$. Then $a(2) = 2$, and $\mathbf{A}_2 - 2\mathbf{I} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ has rank 1, so $g(2) = 2 = a(2)$; also $g(1) = a(1) = 1$. Diagonalizable, with eigenvectors $\mathbf{e}_1$ for 1 and $(1, 1, 0), \mathbf{e}_3$ for 2.

Hence $\mathbf{A}_t$ is diagonalizable over $\mathbb{R}$ if and only if $t \neq 1$. In both exceptional cases a diagonal entry is repeated, but only for $t = 1$ is an eigenvector lost: the entry 1 in position $(1, 2)$ couples $\mathbf{e}_1$ and $\mathbf{e}_2$, and it matters only when their diagonal entries agree.

9.4.6.3 C. Going deeper

Exercise. C1: Matrices with $\mathbf{A}^k = \mathbf{I}$

Let $\mathbf{A} \in M_n(\mathbb{C})$, $n \geq 1$, and $k \geq 1$.

- Prove that if $\mathbf{A}^k = \mathbf{I}$, then every eigenvalue λ of $\mathbf{A}$ satisfies $\lambda^k = 1$.
- Prove that if $\mathbf{A}$ is diagonalizable and every eigenvalue satisfies $\lambda^k = 1$, then $\mathbf{A}^k = \mathbf{I}$.

(c) Show that the hypothesis “diagonalizable” in (b) cannot be dropped.

Solution.

- (a) Let $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$. By Exercise (B3: Eigenvalues of powers and inverses) (a), $\mathbf{v} = \mathbf{I}\mathbf{v} = \mathbf{A}^k\mathbf{v} = \lambda^k\mathbf{v}$, so $(\lambda^k - 1)\mathbf{v} = \mathbf{0}$ and $\lambda^k = 1$ by Theorem 1.1.9.
- (b) Write $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$ with $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$ (Definition 9.4.1). The d_i are eigenvalues of $\mathbf{A}$: $\mathbf{D}$ and $\mathbf{A}$ are similar, so $p_{\mathbf{D}} = p_{\mathbf{A}}$ (Theorem 7.10.3), and each d_i is a root of $p_{\mathbf{D}} = \prod_j (x - d_j)$. So $d_i^k = 1$ for all i , and by Theorem 9.4.3, $\mathbf{A}^k = \mathbf{P} \text{diag}(d_1^k, \dots, d_n^k) \mathbf{P}^{-1} = \mathbf{P}\mathbf{I}\mathbf{P}^{-1} = \mathbf{I}$.
- (c) Let $\mathbf{J} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Its only eigenvalue is 1, and $1^k = 1$. But $\mathbf{J}^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ by induction ($\mathbf{J}^{k+1} = \mathbf{J}^k\mathbf{J}$), which is not $\mathbf{I}$ since $k \neq 0$ in $\mathbb{C}$. (Section 8 proves the converse of (b) in a strong form: over $\mathbb{C}$, $\mathbf{A}^k = \mathbf{I}$ forces $\mathbf{A}$ to be diagonalizable.)

Exercise. C2: Polynomials of diagonalizable matrices

- (a) Prove that if $\mathbf{A} \in M_n(F)$ is diagonalizable over F , then so is $p(\mathbf{A})$ for every $p \in F[x]$, and that every eigenvalue of $p(\mathbf{A})$ has the form $p(\lambda)$ with $\lambda \in \text{spec}(\mathbf{A})$.
- (b) Show that the converse of the first statement fails: find a nondiagonalizable $\mathbf{A}$ and a polynomial p with $p(\mathbf{A})$ diagonalizable.

Solution.

- (a) Let $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$, $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$. By Theorem 9.4.3,

$$p(\mathbf{A}) = \mathbf{P} \text{diag}(p(d_1), \dots, p(d_n)) \mathbf{P}^{-1},$$

so $p(\mathbf{A})$ is similar to a diagonal matrix, that is, diagonalizable. By Theorem 7.10.3, $p_{p(\mathbf{A})} = \prod_i (x - p(d_i))$, so by Theorem 9.2.1 every eigenvalue of $p(\mathbf{A})$ is some $p(d_i)$. As in Exercise (C1: Matrices with $\mathbf{A}^k = \mathbf{I}$) (b), each d_i is an eigenvalue of $\mathbf{A}$.

- (b) Let $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $p = x^2$. Then $\mathbf{A}$ is not diagonalizable: $p_{\mathbf{A}} = x^2$ gives $a(0) = 2$, while $\text{rank } \mathbf{A} = 1$ gives $g(0) = 1$ (Theorem 9.4.1 (e)). But $p(\mathbf{A}) = \mathbf{A}^2 = \mathbf{O}$ is diagonal.

Exercise. C3: The transpose map on $M_n(F)$

Let $n \geq 1$ and $\tau: M_n(F) \rightarrow M_n(F)$, $\tau(\mathbf{X}) = \mathbf{X}^\top$.

- (a) Suppose the characteristic of F is not 2. Prove that τ is diagonalizable, and find $\dim E_1(\tau)$ and $\dim E_{-1}(\tau)$.
- (b) Under the same hypothesis, deduce p_τ , $\text{tr } \tau$, and $\det \tau$.
- (c) Now let $F = \mathbb{F}_2$ and $n \geq 2$. Prove that τ is **not** diagonalizable.

Hint: for the dimensions, use the matrices $\mathbf{E}_{ij} + \mathbf{E}_{ji}$ and $\mathbf{E}_{ij} - \mathbf{E}_{ji}$.

Solution.

- (a) Since $(\mathbf{X}^\top)^\top = \mathbf{X}$ (Theorem 0.9.5), $\tau^2 = \text{id}$. By Theorem 4.5.5, $M_n(F) = E_+ \oplus E_-$, where $E_+ = E_1(\tau)$ is the space of symmetric matrices and $E_- = E_{-1}(\tau)$ the space of skew-symmetric matrices.

A symmetric $\mathbf{X} = (x_{ij})$ is determined by its entries on and above the diagonal, and $\mathbf{X} = \sum_i x_{ii} \mathbf{E}_{ii} + \sum_{i < j} x_{ij} (\mathbf{E}_{ij} + \mathbf{E}_{ji})$. These $n + \binom{n}{2}$ matrices are symmetric and independent, since their nonzero entries sit in disjoint positions. So $\dim E_1(\tau) = n(n+1)/2$. A skew-symmetric $\mathbf{X}$ has $x_{ii} = -x_{ii}$, so $2x_{ii} = 0$ and $x_{ii} = 0$ (as $2 \neq 0$); and $\mathbf{X} = \sum_{i < j} x_{ij} (\mathbf{E}_{ij} - \mathbf{E}_{ji})$, with these $\binom{n}{2}$ matrices independent for the same reason. So $\dim E_{-1}(\tau) = n(n-1)/2$.

The eigenvalues of τ lie among ± 1 (as in Exercise (B2: The transpose map)), and $n(n+1)/2 + n(n-1)/2 = n^2 = \dim M_n(F)$. If $n \geq 2$, both eigenspaces are nonzero, and the dimensions of the eigenspaces add up to n^2 ; if $n = 1$, $\tau = \text{id}$ and $E_1(\tau)$ is everything. Either way Theorem 9.4.1 ((d) $\Rightarrow$ (a)) shows that τ is diagonalizable.

- (b) In a basis $\mathcal{B}$ of eigenvectors, $[\tau]_{\mathcal{B}}^{\mathcal{B}} = \text{diag}(1, \dots, 1, -1, \dots, -1)$ with $n(n+1)/2$ ones and $n(n-1)/2$ entries -1 (Theorem 9.4.1). This matrix is triangular, so

$$\begin{aligned} p_\tau &= (x-1)^{n(n+1)/2} (x+1)^{n(n-1)/2}, \\ \text{tr } \tau &= \frac{n(n+1)}{2} - \frac{n(n-1)}{2} = n, \\ \det \tau &= (-1)^{n(n-1)/2}. \end{aligned}$$

As a check on the trace: in the basis $\mathcal{E} = (\mathbf{E}_{ij})$, $\tau(\mathbf{E}_{ij}) = \mathbf{E}_{ji}$, so the diagonal entry of $[\tau]_{\mathcal{E}}^{\mathcal{E}}$ at $\mathbf{E}_{ij}$ is 1 if $i = j$ and 0 otherwise, and the trace is n .

- (c) Over $\mathbb{F}_2$, $-1 = 1$. If $\tau(\mathbf{X}) = \lambda \mathbf{X}$ with $\mathbf{X} \neq \mathbf{O}$, then as in Exercise (B2: The transpose map) (a), $\lambda^2 = 1$, so $(\lambda - 1)^2 = \lambda^2 - 2\lambda + 1 = \lambda^2 + 1 = 0$ and $\lambda = 1$. So $\text{spec}(\tau) = \{1\}$ and $E_1(\tau)$ is the space of symmetric matrices. The spanning argument in (a) did not divide by 2, so $\dim E_1(\tau) = n(n+1)/2$. For $n \geq 2$, $n(n+1)/2 < n^2$, so condition (d) of Theorem 9.4.1 fails, and τ is not diagonalizable. For instance, $\mathbf{E}_{12}$ is not a sum of eigenvectors.