

9.3 Algebraic and Geometric Multiplicity

An eigenvalue can be counted in two ways. It is a root of the characteristic polynomial, and a root has a multiplicity. It also has an eigenspace, and a subspace has a dimension. The two numbers need not agree, and the gap between them is exactly what can stop an operator from having a diagonal matrix. This section defines both counts, proves that the dimension never exceeds the root multiplicity, and shows that the root multiplicities package the trace and the determinant. The next section turns the comparison into a test for diagonalizability.

9.3.1 Two ways to count an eigenvalue

Compare two matrices over any field F :

$$2\mathbf{I}_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

Both are triangular with diagonal 2, 2, so both have characteristic polynomial $(x-2)^2$ and spectrum $\{2\}$. By the only measure we have so far, the eigenvalue data agree. Yet the eigenvectors are very different. For $2\mathbf{I}_2$, every nonzero vector is an eigenvector and $E_2 = F^2$. For $\mathbf{J}$, the matrix $\mathbf{J} - 2\mathbf{I} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ kills exactly the multiples of $\mathbf{e}_1$, so $E_2(\mathbf{J}) = \text{span}(\mathbf{e}_1)$ is a line. The root 2 “counts twice” in both characteristic polynomials, but it supplies two independent eigenvectors for $2\mathbf{I}_2$ and only one for $\mathbf{J}$. Both numbers deserve names.

The algebraic multiplicity counts an eigenvalue as a root; the geometric multiplicity counts the independent eigenvectors it supplies.

Definition 9.3.1. Algebraic Multiplicity

Let V be a finite-dimensional vector space over F with $\dim V \geq 1$, let $T \in \mathcal{L}(V)$ and $\lambda \in F$. The **algebraic multiplicity** of λ for T is

$$a_T(\lambda) := \text{mult}_\lambda(p_T),$$

the multiplicity of λ as a root of the characteristic polynomial (Definition 6.4.1). For $\mathbf{A} \in M_n(F)$, $a_{\mathbf{A}}(\lambda) := \text{mult}_\lambda(p_{\mathbf{A}})$.

Definition 9.3.2. Geometric Multiplicity

Let V be a finite-dimensional vector space over F , $T \in \mathcal{L}(V)$, and $\lambda \in F$. The **geometric multiplicity** of λ for T is

$$g_T(\lambda) := \dim E_\lambda(T) = \dim \ker(T - \lambda \text{id}_V).$$

For $\mathbf{A} \in M_n(F)$, $g_{\mathbf{A}}(\lambda) := \dim E_\lambda(\mathbf{A}) = n - \text{rank}(\mathbf{A} - \lambda \mathbf{I}_n)$.

When T or $\mathbf{A}$ is clear from context we write simply $a(\lambda)$ and $g(\lambda)$.

In words: $a_T(\lambda)$ is the **largest** m such that $(x - \lambda)^m$ divides p_T . It is a statement about a polynomial and can be computed by factoring. $g_T(\lambda)$ is the dimension of the subspace of solutions of $Tv = \lambda v$, including 0. It is the largest number of **linearly independent** eigenvectors for λ , not the number of eigenvectors, which is infinite over an infinite field. For matrices the second

formula comes from Theorem 3.5.7 applied to $\mathbf{A} - \lambda \mathbf{I}_n$, and it is the one we compute with: find the rank of $\mathbf{A} - \lambda \mathbf{I}$ by elimination and subtract from n .

Well-definedness. Both numbers are defined for every $\lambda \in F$, eigenvalue or not. By Theorem 9.2.1, λ is an eigenvalue if and only if $p_T(\lambda) = 0$, if and only if $a_T(\lambda) \geq 1$; and by Definition 9.2.1, if and only if $E_\lambda(T) \neq \{0\}$, that is, $g_T(\lambda) \geq 1$. So for a scalar that is not an eigenvalue both multiplicities are 0, and for an eigenvalue both are at least 1. Neither depends on a basis: p_T does not (Definition 7.10.2), and $E_\lambda(T)$ is defined without one. If $\mathcal{B}$ is a basis of V , then $a_T(\lambda) = a_{[T]_{\mathcal{B}}}(\lambda)$ because $p_T = p_{[T]_{\mathcal{B}}}$, and $g_T(\lambda) = g_{[T]_{\mathcal{B}}}(\lambda)$ because the coordinate map is an isomorphism from $\ker(T - \lambda \text{id}_V)$ onto $\text{nul}([T]_{\mathcal{B}} - \lambda \mathbf{I})$ (Theorem 4.1.5 (a) and Corollary 4.1.3).

Examples.

- **The hook.** For $2\mathbf{I}_2$: $a(2) = 2$ and $g(2) = 2 - \text{rank } \mathbf{O} = 2$. For $\mathbf{J}$: $a(2) = 2$ and $g(2) = 2 - \text{rank} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1$.
- **A 1×1 matrix.** For $\mathbf{A} = (c)$, $p_{\mathbf{A}} = x - c$, so $a(c) = 1$, and $\mathbf{A} - c\mathbf{I} = \mathbf{O}_1$ has rank 0, so $g(c) = 1$. This degenerate case already shows the pattern of the theorem below: the two counts can only differ when there is room, a repeated root.
- **Differentiation.** On $V = \mathbb{R}[x]_{\leq n}$, the matrix of D in the basis $(1, x, \dots, x^n)$ is strictly upper triangular, so $p_D = x^{n+1}$ and $a_D(0) = n + 1$. But $E_0(D)$ is the line of constant polynomials (the examples after Definition 9.2.2), so $g_D(0) = 1$. The gap can be as large as n .

- **A matrix from Chapter 7.** For $\mathbf{A} = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 3 & 1 \\ 1 & 0 & 2 \end{pmatrix}$, Example (A 3×3 Characteristic Polynomial) found $p_{\mathbf{A}} = (x - 1)(x - 3)^2$, so $a(1) = 1$ and $a(3) = 2$. The matrix $\mathbf{A} - 3\mathbf{I} = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix}$ has a zero column, and its first two rows are independent while the third

is minus the first, so its rank is 2 and $g(3) = 3 - 2 = 1$. Similarly $\mathbf{A} - \mathbf{I} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ has third row equal to the first and first two rows independent, so its rank is 2 and $g(1) = 1$.

Minimal change. Take the triangular matrices

$$\mathbf{B} = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{B}' = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which differ in one entry. Both have $p = (x - 2)^2(x - 1)$, so $a(2) = 2$ for both: the algebraic multiplicity does not notice the change. But $\mathbf{B} - 2\mathbf{I}$ has nonzero rows $(0, 0, 1)$ and $(0, 0, -1)$, rank 1, so $g_{\mathbf{B}}(2) = 2$; while $\mathbf{B}' - 2\mathbf{I}$ has rows $(0, 1, 1)$ and $(0, 0, -1)$, rank 2, so $g_{\mathbf{B}'}(2) = 1$. One off-diagonal entry, invisible to p , removed an eigenvector.

Why these definitions. The algebraic multiplicity is the natural count for a polynomial: with it, the roots of p_T add up to $\deg p_T = \dim V$ whenever p_T splits (Theorem 6.4.3), and the trace and determinant come out right (Theorem 9.3.3 below). The geometric multiplicity is the natural count for eigenvectors: it is the number of basis vectors we can take from $E_\lambda(T)$ when we try to build a basis of eigenvectors. The names record where each number lives, in algebra (a polynomial) or in geometry (a subspace).

Warning. The geometric multiplicity is not read off from p_T . Two operators with the same characteristic polynomial can have different geometric multiplicities: $2\mathbf{I}_2$ and $\mathbf{J}$ above, or $\mathbf{B}$ and $\mathbf{B}'$. To find $g(\lambda)$ you must compute $\text{rank}(\mathbf{A} - \lambda\mathbf{I})$. Also, $g(\lambda)$ is **not** “how many eigenvectors” λ has: over $\mathbb{R}$, $E_\lambda \neq \{0\}$ always contains infinitely many eigenvectors.

Quick check. Find $a(0)$ and $g(0)$ for $\mathbf{N} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in M_3(F)$.

Solution. $\mathbf{N}$ is triangular with zero diagonal, so $p_{\mathbf{N}} = x^3$ and $a(0) = 3$. The rank of $\mathbf{N} - 0\mathbf{I} = \mathbf{N}$ is 1, so $g(0) = 3 - 1 = 2$, with $E_0(\mathbf{N}) = \text{span}(\mathbf{e}_1, \mathbf{e}_3)$.

9.3.2 The geometric multiplicity never exceeds the algebraic

In every example so far, $g(\lambda) \leq a(\lambda)$. That is always the case.

Theorem 9.3.1. Geometric Multiplicity Is at Most Algebraic Multiplicity

Let V be a finite-dimensional vector space over F with $\dim V = n \geq 1$, let $T \in \mathcal{L}(V)$, and let $\lambda \in F$ be an eigenvalue of T . Then

$$1 \leq g_T(\lambda) \leq a_T(\lambda).$$

Idea. Start from the smallest subspace in sight, $E_\lambda(T)$. Take a basis of it and extend it to a basis of V . In that basis, the first g columns of the matrix are $\lambda\mathbf{e}_1, \dots, \lambda\mathbf{e}_g$, so the matrix is block upper triangular with top-left block $\lambda\mathbf{I}_g$. The block triangular determinant then factors $(x - \lambda)^g$ out of p_T , and a root with a factor $(x - \lambda)^g$ has multiplicity at least g . The inequality can be strict because the other block may contribute further factors $x - \lambda$, as it does for $\mathbf{J}$.

Proof. Since λ is an eigenvalue, $E_\lambda(T) \neq \{0\}$, so $g = g_T(\lambda) \geq 1$. Let $(v_1, \dots, v_g)$ be a basis of $E_\lambda(T)$. By Theorem 1.6.5, it extends to a basis $\mathcal{B} = (v_1, \dots, v_g, w_1, \dots, w_{n-g})$ of V .

Case $g = n$. Then $Tv_i = \lambda v_i$ for every basis vector, so $[T]_{\mathcal{B}}^{\mathcal{B}} = \lambda\mathbf{I}_n$ by Definition 4.1.1, and $p_T = (x - \lambda)^n$ by the triangular case of Definition 7.10.1. Hence $a_T(\lambda) = n = g$.

Case $g < n$. For $i \leq g$, $Tv_i = \lambda v_i$, so column i of $[T]_{\mathcal{B}}^{\mathcal{B}}$ is $\lambda\mathbf{e}_i$. Hence

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} \lambda\mathbf{I}_g & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{pmatrix}, \quad x\mathbf{I}_n - [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} (x - \lambda)\mathbf{I}_g & -\mathbf{B} \\ \mathbf{O} & x\mathbf{I}_{n-g} - \mathbf{D} \end{pmatrix},$$

for some $\mathbf{B} \in M_{g \times (n-g)}(F)$ and $\mathbf{D} \in M_{n-g}(F)$. The second matrix is block upper triangular

over $F[x]$ with square diagonal blocks. By Theorem 7.8.4, which holds over the commutative ring $F[x]$ (the remark after it), and Definition 7.10.2,

$$p_T = \det((x - \lambda)\mathbf{I}_g) \det(x\mathbf{I}_{n-g} - \mathbf{D}) = (x - \lambda)^g p_{\mathbf{D}},$$

where $\det((x - \lambda)\mathbf{I}_g) = (x - \lambda)^g$ because the matrix is diagonal. So $(x - \lambda)^g \mid p_T$. By Definition 6.4.1, $a_T(\lambda)$, the **largest** m with $(x - \lambda)^m \mid p_T$, satisfies $a_T(\lambda) \geq g$. This proves the theorem. ■

The inequality is strict exactly when λ is also a root of $p_{\mathbf{D}}$, the characteristic polynomial of the “rest” of the operator. For $\mathbf{J}$, with $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2)$, the block $\mathbf{D}$ is (2) , and $p_{\mathbf{D}} = x - 2$ contributes the second factor.

A useful consequence: if $a(\lambda) = 1$, then $g(\lambda) = 1$. A **simple** root of p_T always has a one-dimensional eigenspace, and so the only eigenvalues whose eigenspaces need checking are the repeated ones.

Both multiplicities are properties of the operator, not of the basis, so they cannot distinguish similar matrices. That makes them tools for proving that two matrices are **not** similar.

Proposition 9.3.2. Similar Matrices Have the Same Multiplicities

Let $\mathbf{A}, \mathbf{B} \in M_n(F)$ with $\mathbf{A} \sim \mathbf{B}$, and let $\lambda \in F$. Then $a_{\mathbf{A}}(\lambda) = a_{\mathbf{B}}(\lambda)$ and $g_{\mathbf{A}}(\lambda) = g_{\mathbf{B}}(\lambda)$.

Proof. Let $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. By Theorem 7.10.3, $p_{\mathbf{A}} = p_{\mathbf{B}}$, so the algebraic multiplicities agree. Also $\mathbf{B} - \lambda\mathbf{I} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} - \lambda\mathbf{P}^{-1}\mathbf{P} = \mathbf{P}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{P}$, so $\mathbf{B} - \lambda\mathbf{I} \sim \mathbf{A} - \lambda\mathbf{I}$, and these have the same rank by Proposition 4.2.6 (a). Hence $g_{\mathbf{B}}(\lambda) = n - \text{rank}(\mathbf{B} - \lambda\mathbf{I}) = n - \text{rank}(\mathbf{A} - \lambda\mathbf{I}) = g_{\mathbf{A}}(\lambda)$. ■

For example, $2\mathbf{I}_2$ and $\mathbf{J}$ have the same characteristic polynomial but $g(2) = 2$ versus $g(2) = 1$, so they are not similar. The same test shows $\mathbf{B} \not\sim \mathbf{B}'$ in the minimal change above.

Warning. A repeated eigenvalue does not force a shortage of eigenvectors, and equal eigenvalues do not force similarity. The identity $\mathbf{I}_2$ has the repeated eigenvalue 1, yet $g(1) = a(1) = 2$ and its matrix is as simple as possible. In the other direction, $\mathbf{I}_2$ and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ have the same eigenvalues with the same algebraic multiplicities, and they are not similar. Even agreement of **both** multiplicities for every eigenvalue does not force similarity; [Exercise \(C2: Multiplicities do not determine similarity\)](#) gives two 4×4 matrices showing this.

9.3.3 Multiplicities, trace, and determinant

By Theorem 6.4.3, the algebraic multiplicities of the eigenvalues of T add up to at most $\deg p_T = \dim V$, with equality exactly when p_T splits over F . Over $\mathbb{C}$ this always happens (Corollary 6.5.5). When it does, p_T is determined by the eigenvalues and their algebraic multi-

plicities,

$$p_T = (x - \lambda_1)^{a(\lambda_1)} \cdots (x - \lambda_k)^{a(\lambda_k)},$$

and two of its coefficients are familiar: by Theorem 7.10.2, the coefficient of x^{n-1} is $-\operatorname{tr} T$ and the constant term is $(-1)^n \det T$. Comparing the two descriptions expresses trace and determinant through eigenvalues.

Theorem 9.3.3. Trace and Determinant from Eigenvalues

Let V be a finite-dimensional vector space over F with $\dim V = n \geq 1$, and let $T \in \mathcal{L}(V)$. Suppose p_T splits over F , and let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of T . Then

$$a(\lambda_1) + \cdots + a(\lambda_k) = n, \quad \operatorname{tr} T = \sum_{i=1}^k a(\lambda_i) \lambda_i, \quad \det T = \prod_{i=1}^k \lambda_i^{a(\lambda_i)}.$$

In words: listing each eigenvalue as often as its algebraic multiplicity, the trace is the sum of the list and the determinant is its product. In particular this holds for every operator on a nonzero finite-dimensional complex space, and for every $\mathbf{A} \in M_n(\mathbb{C})$.

Idea. Rather than expanding a product of linear factors by hand, notice that the product is itself a characteristic polynomial: that of the diagonal matrix with the eigenvalues on its diagonal. Chapter 7 already knows the x^{n-1} coefficient and the constant term of any characteristic polynomial, so we read them off twice and compare.

Proof. Since p_T splits and is monic of degree n (Theorem 7.10.2), Definition 6.4.2 gives $p_T = (x - c_1) \cdots (x - c_n)$ for some $c_1, \dots, c_n \in F$. For $c, \lambda \in F$, the polynomial $x - c$ has $\operatorname{mult}_\lambda(x - c) = 1$ if $c = \lambda$ and 0 otherwise, since λ is a root of $x - c$ only when $c = \lambda$ (Theorem 6.1.3). So by Theorem 6.4.2, $\operatorname{mult}_\lambda(p_T) = \sum_j \operatorname{mult}_\lambda(x - c_j)$ is the number of indices j with $c_j = \lambda$. In particular the values occurring among $c_1, \dots, c_n$ are exactly the roots of p_T in F , which are $\lambda_1, \dots, \lambda_k$ by Theorem 9.2.1, and λ_i occurs exactly $a(\lambda_i)$ times. Counting the n factors gives $\sum_i a(\lambda_i) = n$, and

$$c_1 + \cdots + c_n = \sum_i a(\lambda_i) \lambda_i, \quad c_1 \cdots c_n = \prod_i \lambda_i^{a(\lambda_i)}.$$

Let $\mathbf{C} = \operatorname{diag}(c_1, \dots, c_n)$. It is triangular, so $p_{\mathbf{C}} = (x - c_1) \cdots (x - c_n) = p_T$ (the examples after Definition 7.10.1). Apply Theorem 7.10.2 to both $\mathbf{C}$ and T : the coefficient of x^{n-1} in this one polynomial is $-\operatorname{tr} \mathbf{C} = -(c_1 + \cdots + c_n)$ and also $-\operatorname{tr} T$; the constant term is $(-1)^n \det \mathbf{C} = (-1)^n c_1 \cdots c_n$ (Theorem 7.4.6) and also $(-1)^n \det T$. Hence $\operatorname{tr} T = c_1 + \cdots + c_n$ and $\det T = c_1 \cdots c_n$, which are the displayed formulas. The complex case follows from Corollary 6.5.5. ■

Examples.

- For $\mathbf{A} = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 3 & 1 \\ 1 & 0 & 2 \end{pmatrix}$, with $a(1) = 1$ and $a(3) = 2$: $\operatorname{tr} \mathbf{A} = 2 + 3 + 2 = 7 = 1 + 3 + 3$, and $\det \mathbf{A} = 9 = 1 \cdot 3 \cdot 3$.

- For the real matrix $\begin{pmatrix} 3 & -2 \\ 4 & -1 \end{pmatrix}$ of Example (A Real Matrix with Non-Real Eigenvalues), p does not split over $\mathbb{R}$, but it splits over $\mathbb{C}$ with roots $1 \pm 2i$. The theorem over $\mathbb{C}$ gives $(1 + 2i) + (1 - 2i) = 2 = \operatorname{tr} \mathbf{A}$ and $(1 + 2i)(1 - 2i) = 1 + 4 = 5 = \det \mathbf{A}$. The trace and determinant are real, while the eigenvalues are not.
- For a 2×2 matrix whose characteristic polynomial splits, the theorem says the eigenvalues λ, μ (listed with multiplicity) satisfy $\lambda + \mu = \operatorname{tr} \mathbf{A}$ and $\lambda\mu = \det \mathbf{A}$. This is a quick way to guess eigenvalues: for $\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$, we want two numbers with sum 3 and product -4 , namely 4 and -1 .

Without splitting, the theorem says nothing. If p_T does not split, the eigenvalues in F , counted with algebraic multiplicity, do not account for all of p_T , and summing or multiplying them gives no information. For $\mathbf{A} = \begin{pmatrix} 3 & -2 \\ 4 & -1 \end{pmatrix}$ over $\mathbb{R}$ there are no eigenvalues: the multiplicities add up to $0 \neq 2$, the empty sum 0 is not $\operatorname{tr} \mathbf{A} = 2$, and the empty product 1 is not $\det \mathbf{A} = 5$. For the real rotation $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, with trace 0 and determinant 1, the empty sum and product happen to give the right values, but only by coincidence. For real matrices the remedy is to pass to $\mathbb{C}$, where $p_{\mathbf{A}}$ always splits, as in the second example above.

9.3.4 Exercises

9.3.4.1 A. Check your understanding

Exercise. A1

- Define the algebraic and the geometric multiplicity of $\lambda \in F$ for $\mathbf{A} \in M_n(F)$.
- State the inequality between them for an eigenvalue λ , and name the tool its proof uses.
- True or false: if $a(\lambda) = 1$, then $g(\lambda) = 1$. Justify your answer.
- True or false: for every $\mathbf{A} \in M_n(\mathbb{C})$, the geometric multiplicities of the eigenvalues add up to n . Justify your answer.
- True or false: if $\mathbf{A}, \mathbf{B} \in M_n(F)$ have the same characteristic polynomial, then $g_{\mathbf{A}}(\lambda) = g_{\mathbf{B}}(\lambda)$ for every λ . Justify your answer.
- For $\mathbf{A} \in M_3(\mathbb{C})$ with eigenvalues 2 (algebraic multiplicity 2) and -1 , what are $\operatorname{tr} \mathbf{A}$ and $\det \mathbf{A}$?

Solution.

- $a_{\mathbf{A}}(\lambda) = \operatorname{mult}_{\lambda}(p_{\mathbf{A}})$, the largest m with $(x - \lambda)^m \mid p_{\mathbf{A}}$; $g_{\mathbf{A}}(\lambda) = \dim E_{\lambda}(\mathbf{A}) = n - \operatorname{rank}(\mathbf{A} - \lambda \mathbf{I})$ (Definition 9.3.1, Definition 9.3.2).
- $1 \leq g(\lambda) \leq a(\lambda)$ (Theorem 9.3.1). The proof extends a basis of E_{λ} to a basis of the space and uses the block triangular determinant (Theorem 7.8.4) over $F[x]$.
- True. An eigenvalue has $g(\lambda) \geq 1$, and $g(\lambda) \leq a(\lambda) = 1$.
- False. For $\mathbf{J} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, the only eigenvalue is 2, with $g(2) = 1 \neq 2$. (The **algebraic** multiplicities do add up to n over $\mathbb{C}$.)

- (e) False. $2\mathbf{I}_2$ and $\mathbf{J}$ both have $p = (x - 2)^2$, but $g(2) = 2$ and $g(2) = 1$.
 (f) By [Theorem 9.3.3](#), listing eigenvalues with algebraic multiplicity as $2, 2, -1$: $\operatorname{tr} \mathbf{A} = 3$ and $\det \mathbf{A} = -4$.

9.3.4.2 B. Practice

Exercise. B1: Computing multiplicities

For each real matrix, find every eigenvalue together with its algebraic and geometric multiplicity.

- (a) $\mathbf{A}_1 = \begin{pmatrix} 4 & -1 \\ 1 & 2 \end{pmatrix}$.
 (b) $\mathbf{A}_2 = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$.
 (c) $\mathbf{A}_3 = \begin{pmatrix} 2 & 1 & 1 \\ -1 & 0 & -1 \\ 1 & 1 & 2 \end{pmatrix}$, given that $p_{\mathbf{A}_3} = p_{\mathbf{A}_2}$.
 (d) The matrix of the operator $S(q)(x) = q(x + 1)$ on $\mathbb{R}[x]_{\leq 2}$, in the basis $(1, x, x^2)$.

Solution.

- (a) $\operatorname{tr} \mathbf{A}_1 = 6$, $\det \mathbf{A}_1 = 8 + 1 = 9$, so $p = x^2 - 6x + 9 = (x - 3)^2$: the only eigenvalue is 3, with $a(3) = 2$. $\mathbf{A}_1 - 3\mathbf{I} = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$ has rank 1, so $g(3) = 1$.
 (b) Expand $\det(x\mathbf{I} - \mathbf{A}_2)$ along the first column:

$$\det \begin{pmatrix} x-2 & -1 & -1 \\ 0 & x-1 & 1 \\ 1 & 1 & x-1 \end{pmatrix} = (x-2)[(x-1)^2 - 1] + 1 \cdot \det \begin{pmatrix} -1 & -1 \\ x-1 & 1 \end{pmatrix}.$$

The first bracket is $x^2 - 2x = x(x - 2)$ and the last determinant is $-1 + (x - 1) = x - 2$, so $p_{\mathbf{A}_2} = (x - 2)(x^2 - 2x + 1) = (x - 1)^2(x - 2)$. Thus $a(1) = 2$ and $a(2) = 1$. By [Theorem 9.3.1](#), $g(2) = 1$. For $\lambda = 1$,

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

Row 3 is $-(\text{row } 1) - (\text{row } 2)$, and rows 1 and 2 are independent, so the rank is 2 and $g(1) = 1$.

- (c) Here $a(1) = 2$ and $a(2) = 1$, as for $\mathbf{A}_2$, so $g(2) = 1$. Now

$$\mathbf{A}_3 - \mathbf{I} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

has all rows multiples of $(1, 1, 1)$, so its rank is 1 and $g(1) = 2$. By [Proposition 9.3.2](#), $\mathbf{A}_2$ and $\mathbf{A}_3$ are not similar, although they have the same characteristic polynomial.

- (d) $S(1) = 1$, $S(x) = 1 + x$, $S(x^2) = 1 + 2x + x^2$, so the matrix is $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ (as in Example (Characteristic Polynomials of Two Operators)). It is triangular, so $p = (x - 1)^3$ and $a(1) = 3$. The matrix minus $\mathbf{I}$ is $\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, of rank 2, so $g(1) = 1$: the only polynomials with $q(x + 1) = q(x)$ in $\mathbb{R}[x]_{\leq 2}$ are the constants.

Exercise. B2: A parameter that merges eigenvalues

For $t \in \mathbb{R}$, let $\mathbf{A}_t = \begin{pmatrix} t & 1 \\ -1 & 3 \end{pmatrix}$. Find all t for which $\mathbf{A}_t$ has a real eigenvalue λ with $g(\lambda) < a(\lambda)$.

Solution. $\text{tr } \mathbf{A}_t = t + 3$ and $\det \mathbf{A}_t = 3t + 1$, so $p(x) = x^2 - (t + 3)x + (3t + 1)$. For a 2×2 matrix, $g(\lambda) < a(\lambda)$ is possible only if $a(\lambda) = 2$, because $a(\lambda) = 1$ forces $g(\lambda) = 1$ ([Theorem 9.3.1](#)). So p must have a repeated real root, which happens exactly when its discriminant vanishes:

$$(t + 3)^2 - 4(3t + 1) = t^2 - 6t + 5 = (t - 1)(t - 5) = 0.$$

If $t = 1$, then $p = x^2 - 4x + 4 = (x - 2)^2$ and $\mathbf{A}_1 - 2\mathbf{I} = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}$, of rank 1, so $g(2) = 1 < 2 = a(2)$. If $t = 5$, then $p = (x - 4)^2$ and $\mathbf{A}_5 - 4\mathbf{I} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$, of rank 1, so $g(4) = 1 < 2$. Hence the answer is $t \in \{1, 5\}$. (In both cases $g = a$ was impossible anyway: $g(\lambda) = 2$ would mean $\mathbf{A}_t - \lambda\mathbf{I} = \mathbf{O}$, but the off-diagonal entries of $\mathbf{A}_t$ are nonzero.)

Exercise. B3: Eigenvalues from trace and determinant

- (a) Let $\mathbf{A} \in M_3(\mathbb{R})$ have the eigenvalue $1 + i$ (as a complex matrix) and $\det \mathbf{A} = 6$. Find all eigenvalues of $\mathbf{A}$ over $\mathbb{C}$, and $\text{tr } \mathbf{A}$.
- (b) Let $\mathbf{B} \in M_2(\mathbb{R})$ with $\text{tr } \mathbf{B} = 6$ and $\det \mathbf{B} = 9$. Show that $\text{spec}(\mathbf{B}) = \{3\}$ with $a(3) = 2$. Give one such $\mathbf{B}$ with $g(3) = 2$ and one with $g(3) = 1$.

Solution.

- (a) By [Theorem 9.2.5](#), $1 - i$ is also an eigenvalue. By [Corollary 6.5.5](#), $p_{\mathbf{A}}$ splits over $\mathbb{C}$ into three linear factors, and two of them are $x - (1 + i)$ and $x - (1 - i)$; let μ be the third root. By [Theorem 9.3.3](#), $6 = \det \mathbf{A} = (1 + i)(1 - i)\mu = 2\mu$, so $\mu = 3$. Hence the eigenvalues are $1 + i, 1 - i, 3$, and $\text{tr } \mathbf{A} = (1 + i) + (1 - i) + 3 = 5$.
- (b) By [Theorem 7.10.2](#), $p_{\mathbf{B}} = x^2 - 6x + 9 = (x - 3)^2$. So $\text{spec}(\mathbf{B}) = \{3\}$ and $a(3) = 2$. For

$\mathbf{B} = 3\mathbf{I}_2$, $g(3) = 2$. For $\mathbf{B} = \begin{pmatrix} 4 & -1 \\ 1 & 2 \end{pmatrix}$ of [Exercise \(B1: Computing multiplicities\)](#) (a), $\text{tr } \mathbf{B} = 6$, $\det \mathbf{B} = 9$, and $g(3) = 1$.

9.3.4.3 C. Going deeper

Exercise. C1: Nilpotent matrices

Let $\mathbf{A} \in M_n(\mathbb{C})$, $n \geq 1$, be nilpotent ($\mathbf{A}^k = \mathbf{O}$ for some $k \geq 1$).

- (a) Prove that $a(0) = n$ and $g(0) = n - \text{rank } \mathbf{A}$.
- (b) Deduce that $g(0) = a(0)$ if and only if $\mathbf{A} = \mathbf{O}$.
- (c) For $\mathbf{N} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, check that $\mathbf{N}^4 = \mathbf{O}$ and compute $g(0)$. How large can $a(0) - g(0)$ be for a nilpotent $n \times n$ matrix?

Solution.

- (a) By [Exercise \(C2: Nilpotent operators\)](#) (b), $p_{\mathbf{A}} = x^n$, so $a(0) = \text{mult}_0(x^n) = n$. By [Definition 9.3.2](#), $g(0) = n - \text{rank}(\mathbf{A} - 0\mathbf{I}) = n - \text{rank } \mathbf{A}$.
- (b) By (a), $g(0) = a(0)$ if and only if $n - \text{rank } \mathbf{A} = n$, that is, $\text{rank } \mathbf{A} = 0$, that is, $\mathbf{A} = \mathbf{O}$ (a matrix of rank 0 has no nonzero column).
- (c) $\mathbf{N}\mathbf{e}_1 = \mathbf{0}$ and $\mathbf{N}\mathbf{e}_j = \mathbf{e}_{j-1}$ for $j \geq 2$. So $\mathbf{N}^j\mathbf{e}_i = \mathbf{0}$ once $j \geq i$, and $\mathbf{N}^4\mathbf{e}_i = \mathbf{0}$ for $i \leq 4$; hence $\mathbf{N}^4 = \mathbf{O}$. The columns of $\mathbf{N}$ are $\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, so $\text{rank } \mathbf{N} = 3$ and $g(0) = 1$, while $a(0) = 4$. In general, a nilpotent $\mathbf{A}$ has $0 \in \text{spec}(\mathbf{A})$ ([Exercise \(C2: Nilpotent operators\)](#) (a)), so $g(0) \geq 1$ and $a(0) - g(0) \leq n - 1$; the $n \times n$ analogue of $\mathbf{N}$, with ones just above the diagonal, has rank $n - 1$ and attains $a(0) - g(0) = n - 1$.

Exercise. C2: Multiplicities do not determine similarity

Let

$$\mathbf{M}_1 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{M}_2 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in M_4(\mathbb{R}).$$

- (a) Show that $\mathbf{M}_1$ and $\mathbf{M}_2$ have the same eigenvalues with the same algebraic **and** geometric multiplicities.
- (b) Compute $\text{rank}(\mathbf{M}_1 - \mathbf{I})^2$ and $\text{rank}(\mathbf{M}_2 - \mathbf{I})^2$. Deduce that $\mathbf{M}_1 \not\sim \mathbf{M}_2$.
- (c) Construct a matrix $\mathbf{M} \in M_4(\mathbb{R})$ with $\text{spec}(\mathbf{M}) = \{2, 5\}$, $a(2) = 3$, $g(2) = 1$, and $a(5) = g(5) = 1$.

Hint: for (b), if $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, relate $(\mathbf{B} - \mathbf{I})^2$ to $(\mathbf{A} - \mathbf{I})^2$.

Solution.

(a) Both matrices are upper triangular with diagonal 1, 1, 1, 1, so both have $p = (x - 1)^4$, spectrum $\{1\}$, and $a(1) = 4$. The matrix $\mathbf{M}_1 - \mathbf{I}$ has columns $\mathbf{0}, \mathbf{e}_1, \mathbf{0}, \mathbf{e}_3$, of rank 2; $\mathbf{M}_2 - \mathbf{I}$ has columns $\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{0}$, also of rank 2. So $g(1) = 4 - 2 = 2$ for both.

(b) Let $\mathbf{N}_1 = \mathbf{M}_1 - \mathbf{I}$, so $\mathbf{N}_1 \mathbf{e}_2 = \mathbf{e}_1$, $\mathbf{N}_1 \mathbf{e}_4 = \mathbf{e}_3$, and $\mathbf{N}_1 \mathbf{e}_1 = \mathbf{N}_1 \mathbf{e}_3 = \mathbf{0}$. Then $\mathbf{N}_1^2$ kills every $\mathbf{e}_j$, so $\mathbf{N}_1^2 = \mathbf{0}$ and $\text{rank } \mathbf{N}_1^2 = 0$. Let $\mathbf{N}_2 = \mathbf{M}_2 - \mathbf{I}$, so $\mathbf{N}_2 \mathbf{e}_2 = \mathbf{e}_1$, $\mathbf{N}_2 \mathbf{e}_3 = \mathbf{e}_2$, and $\mathbf{N}_2 \mathbf{e}_1 = \mathbf{N}_2 \mathbf{e}_4 = \mathbf{0}$. Then $\mathbf{N}_2^2 \mathbf{e}_3 = \mathbf{e}_1 \neq \mathbf{0}$ and $\mathbf{N}_2^2$ kills $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4$, so $\mathbf{N}_2^2 = \mathbf{E}_{13}$ has rank 1.

Suppose $\mathbf{M}_2 = \mathbf{P}^{-1} \mathbf{M}_1 \mathbf{P}$. With $p(x) = (x - 1)^2$, Proposition 4.2.6 (c) gives $(\mathbf{M}_2 - \mathbf{I})^2 = \mathbf{P}^{-1} (\mathbf{M}_1 - \mathbf{I})^2 \mathbf{P}$, and then (a) of the same proposition gives equal ranks, $1 = 0$. This contradiction shows $\mathbf{M}_1 \not\sim \mathbf{M}_2$. So eigenvalues together with both multiplicities do not determine a matrix up to similarity; Chapter 10 finds invariants that do.

(c) Take

$$\mathbf{M} = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

It is triangular, so $p_{\mathbf{M}} = (x - 2)^3(x - 5)$, $a(2) = 3$, and $a(5) = 1$, hence $g(5) = 1$ by Theorem 9.3.1. The matrix $\mathbf{M} - 2\mathbf{I}$ has columns $\mathbf{0}, \mathbf{e}_1, \mathbf{e}_2, 3\mathbf{e}_4$, which has rank 3, so $g(2) = 4 - 3 = 1$.