

9.1 Invariant Subspaces

Chapter 4 raised a question it could not answer: given an operator T on a finite-dimensional space, how simple can its matrix $[T]_{\mathcal{B}}^{\mathcal{B}}$ be made by choosing the basis $\mathcal{B}$ well? Chapter 8 found the first piece of an answer. A zero block in the lower-left corner of $[T]_{\mathcal{B}}^{\mathcal{B}}$ appears exactly when the first few basis vectors span a subspace that T maps into itself. This section gives that property a name, collects the standard examples, and shows that it can fail badly: some operators map almost no subspace into itself. It ends by recording how such subspaces cut an operator, and its characteristic polynomial, into smaller pieces. The next section looks at the smallest nonzero case, a line, and finds eigenvectors there.

9.1.1 Subspaces an operator maps into itself

After a vector space we studied its subspaces, and after linear maps their kernels and images. For an operator $T \in \mathcal{L}(V)$, the domain and the codomain are the same space, so a new question makes sense: which subspaces of V are compatible with T ?

Here is the result that makes the question worth asking. Let V be finite-dimensional, $T \in \mathcal{L}(V)$, and U a subspace with $\{0\} \neq U \neq V$. Take a basis $(u_1, \dots, u_k)$ of U and extend it to a basis $\mathcal{B} = (u_1, \dots, u_k, w_1, \dots, w_l)$ of V . By Theorem 8.1.7, the lower-left $l \times k$ block of $[T]_{\mathcal{B}}^{\mathcal{B}}$ is zero **if and only if** $Tu \in U$ for every $u \in U$. For example, the cyclic shift $T(x, y, z) = (y, z, x)$ on $\mathbb{R}^3$ keeps the plane $x + y + z = 0$ in place, and in Example (A Plane Mapped into Itself) this produced a matrix with a zero corner. The property “ T maps U into U ” is what we need to control, so it gets a name.

A subspace is invariant under T when T never moves a vector of the subspace out of it.

Definition 9.1.1. Invariant Subspace

Let V be a vector space over F and $T \in \mathcal{L}(V)$. A subspace U of V is **invariant under T** , or **T -invariant**, if

$$Tu \in U \quad \text{for every } u \in U,$$

that is, if $T(U) \subseteq U$.

In words: we start from a **subspace** U , not an arbitrary subset. We apply T to **every** vector of U , and each image must land back **in** U . Nothing is required of vectors outside U , and nothing says that a vector of U stays where it is: Tu may be a different vector of U . The condition is an inclusion $T(U) \subseteq U$, not an equality.

Checking “for every $u \in U$ ” looks like infinitely many checks. Linearity reduces it to finitely many when U has a finite spanning list.

Lemma 9.1.1. Invariance Can Be Checked on a Spanning List

Let $T \in \mathcal{L}(V)$ and $U = \text{span}(u_1, \dots, u_k)$. Then U is T -invariant if and only if $Tu_i \in U$ for each $i = 1, \dots, k$.

Proof. $(\Rightarrow)$ Each u_i lies in U . $(\Leftarrow)$ Let $u = c_1u_1 + \dots + c_ku_k \in U$. By Theorem 2.1.4, $Tu = c_1Tu_1 + \dots + c_kTu_k$, a linear combination of vectors of U , which lies in U because U is a subspace (Theorem 1.2.1).

Examples.

- **The trivial subspaces.** For every $T \in \mathcal{L}(V)$, both $\{0\}$ and V are T -invariant: $T0 = 0$ (Theorem 2.1.1), and $Tv \in V$ always. These are the degenerate cases, and they matter as a warning about what the word does **not** tell us: invariance carries information only for subspaces strictly between $\{0\}$ and V . At the other extreme, for $T = c \operatorname{id}_V$ **every** subspace is invariant, since $cu \in U$ whenever $u \in U$.
- **Kernel and image.** $\ker T$ is invariant: if $u \in \ker T$, then $Tu = 0 \in \ker T$. So is $\operatorname{im} T$: if $w \in \operatorname{im} T$, then Tw is T of something, so $Tw \in \operatorname{im} T$. More generally, for every polynomial $p \in F[x]$, the subspaces $\ker p(T)$ and $\operatorname{im} p(T)$ are T -invariant, by Theorem 6.7.4 (b). For instance $\ker(T - c \operatorname{id}_V)$ is invariant for every scalar c .
- **A line in the plane.** Let $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$ act on $\mathbb{R}^2$. The line $\operatorname{span}(\mathbf{e}_1)$ is invariant, by Lemma 9.1.1, because $\mathbf{A}\mathbf{e}_1 = (2, 0) = 2\mathbf{e}_1 \in \operatorname{span}(\mathbf{e}_1)$. The line $\operatorname{span}((1, 1))$ is invariant as well, since $\mathbf{A}(1, 1) = (3, 3) = 3(1, 1)$.
- **Polynomials of bounded degree.** Let $D: F[x] \rightarrow F[x]$ be differentiation (Example (Differentiation)). For each $k \in \mathbb{N}$, the subspace $F[x]_{\leq k}$ is D -invariant, because differentiating does not raise the degree. It is spanned by $1, x, \dots, x^k$, and $D(x^j) = jx^{j-1} \in F[x]_{\leq k}$.

Nonexample by minimal change. Keep $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$ and replace the line $\operatorname{span}(\mathbf{e}_1)$ by the line $\operatorname{span}(\mathbf{e}_2)$. It is still a subspace, and $\mathbf{A}\mathbf{e}_2 = (1, 3)$ is still some vector of $\mathbb{R}^2$. What fails is the clause “ $T\mathbf{u} \in U$ for every $\mathbf{u} \in U$ ” at $\mathbf{u} = \mathbf{e}_2$: the vector $(1, 3)$ is not a multiple of $\mathbf{e}_2$. In the same way, $\operatorname{span}(x, x^2)$ inside $\mathbb{R}[x]_{\leq 3}$ is not D -invariant, because $D(x) = 1 \notin \operatorname{span}(x, x^2)$.

The examples with lines all have the same shape, which is worth recording now, because the whole next section grows out of it.

Proposition 9.1.2. Invariant Lines

Let $T \in \mathcal{L}(V)$ and $v \in V$ with $v \neq 0$. Then $\operatorname{span}(v)$ is T -invariant if and only if $Tv = \lambda v$ for some $\lambda \in F$.

Proof. By Lemma 9.1.1, $\operatorname{span}(v)$ is invariant if and only if $Tv \in \operatorname{span}(v)$, and $\operatorname{span}(v) = \{\lambda v : \lambda \in F\}$ by Definition 1.3.2. ■

So an invariant line is a line on which T acts by stretching with a single factor λ . Does every operator have one? Not over $\mathbb{R}$.

Example. A Rotation Has No Invariant Lines

Let $R: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $R(x, y) = (-y, x)$, be the rotation by a right angle. Find all R -invariant subspaces of $\mathbb{R}^2$.

Solution. A subspace U of $\mathbb{R}^2$ has $\dim U \in \{0, 1, 2\}$ by Theorem 1.6.7. If $\dim U = 0$, then $U = \{\mathbf{0}\}$; if $\dim U = 2$, then $U = \mathbb{R}^2$ by Theorem 1.6.8. Both are invariant.

Suppose $\dim U = 1$. A basis of U is a single vector $\mathbf{v} = (a, b) \neq \mathbf{0}$, and $U = \text{span}(\mathbf{v})$. If U were invariant, then by Proposition 9.1.2 there would be $\lambda \in \mathbb{R}$ with $R\mathbf{v} = \lambda\mathbf{v}$, that is,

$$-b = \lambda a, \quad a = \lambda b.$$

Substituting the first equation into the second gives $a = \lambda b = \lambda(-\lambda a) = -\lambda^2 a$, so $(1 + \lambda^2)a = 0$. Since λ is real, $1 + \lambda^2 \geq 1 > 0$, so $a = 0$. Then $b = -\lambda a = 0$ by the first equation, and $\mathbf{v} = \mathbf{0}$, a contradiction.

Hence the only R -invariant subspaces of $\mathbb{R}^2$ are $\{\mathbf{0}\}$ and $\mathbb{R}^2$. Geometrically, a quarter-turn moves every line through the origin onto the perpendicular line.

The same computation over $\mathbb{C}$ does not end in a contradiction: $1 + \lambda^2 = 0$ has the solutions $\lambda = \pm i$, and we will see in the next section that $\mathbb{C}^2$ does contain R -invariant lines. Whether invariant lines exist depends on the field.

Why this definition. Two variations suggest themselves, and both are worse.

- *Demand* $T(U) = U$. Then $\ker T$ would fail whenever T is not injective, since then $T(\ker T) = \{\mathbf{0}\} \neq \ker T$. But the kernel is exactly the kind of subspace that produces a zero block, and Theorem 8.1.7 needs only the inclusion.
- *Demand* $Tu = u$ for every $u \in U$. That says T acts as the identity on U , a far stronger condition. The line $\text{span}((1, 1))$ above is invariant under $\mathbf{A}$, but $\mathbf{A}$ triples its vectors.

The name records that the subspace, as a whole, is unchanged in the sense that it is carried into itself; its individual vectors may move.

Warning. Invariant does not mean fixed. The vectors of a T -invariant subspace may move, as long as they stay inside it. For the swap $T(x, y) = (y, x)$ on $\mathbb{R}^2$, the whole plane is invariant, yet $T(1, 0) = (0, 1) \neq (1, 0)$. Conversely, containing a fixed vector does not make a subspace invariant. For $T(x, y, z) = (x, z, y + z)$ on $\mathbb{R}^3$, the plane $z = 0$ contains $\mathbf{e}_1$, and $T\mathbf{e}_1 = \mathbf{e}_1$; but it also contains $\mathbf{e}_2$, and $T\mathbf{e}_2 = \mathbf{e}_3$ leaves the plane.

Quick check. Let $T(x, y) = (x + y, y)$ on $\mathbb{R}^2$. Is the x -axis T -invariant? Is the y -axis?

Solution. By Lemma 9.1.1 it suffices to test one spanning vector. $T\mathbf{e}_1 = (1, 0) = \mathbf{e}_1$, which lies on the x -axis, so the x -axis is invariant. $T\mathbf{e}_2 = (1, 1)$, which is not a multiple of $\mathbf{e}_2$, so the y -axis is **not** invariant.

9.1.2 Restriction and the induced operator

An invariant subspace U lets us split T into two smaller operators. One lives on U : since T sends U into U , we may simply forget the rest of V . The other lives on the quotient V/U , which records what U cannot see. Chapter 8 met both in the proof of Theorem 8.1.7, and Chapter

2 constructed the second in Exercise (C2: The operator induced on a quotient). We now give them names.

Definition 9.1.2. Restriction and Induced Operator

Let $T \in \mathcal{L}(V)$ and let U be a T -invariant subspace of V .

- (a) The **restriction** of T to U is the operator $T|_U: U \rightarrow U$, $u \mapsto Tu$.
- (b) The **induced operator** (or quotient operator) on V/U is $\bar{T}: V/U \rightarrow V/U$, $v + U \mapsto Tv + U$.

Well-definedness. For (a), the rule makes sense as a map **into** U only because U is invariant; it is linear because T is, and U carries the operations of V . For (b), the rule is phrased in terms of a representative v of the coset, so we must check that it does not depend on that choice. Suppose $v + U = v' + U$. Then $v - v' \in U$ (Lemma 4.4.3), so $Tv - Tv' = T(v - v') \in U$ by invariance, and hence $Tv + U = Tv' + U$. Linearity follows from the coset operations (Theorem 4.4.4): for instance

$$\begin{aligned}\bar{T}((v + U) + (w + U)) &= \bar{T}(v + w + U) = Tv + Tw + U \\ &= \bar{T}(v + U) + \bar{T}(w + U),\end{aligned}$$

and scalars work the same way. Invariance was used exactly once in each part, which is where the hypothesis earns its place.

Examples.

- **Differentiation.** Let D act on $V = \mathbb{R}[x]_{\leq 2}$ and $U = \mathbb{R}[x]_{\leq 1}$, which is D -invariant. The restriction $D|_U$ is differentiation on $\mathbb{R}[x]_{\leq 1}$. The quotient V/U has dimension 1 (Theorem 4.4.5), with basis $x^2 + U$, and $\bar{D}(x^2 + U) = 2x + U = 0 + U$, because $2x \in U$. So $\bar{D} = 0$.
- **The cyclic shift.** For $T(x, y, z) = (y, z, x)$ and the plane U where $x + y + z = 0$, Example (A Plane Mapped into Itself) showed that $\bar{T}$ is the identity of the line $\mathbb{R}^3/U$: the shift permutes coordinates and so never changes their sum.
- **Degenerate cases.** For $U = V$, the restriction is T itself and V/V is the zero space. For $U = \{0\}$, the restriction is the operator on the zero space, and $\bar{T}$ is T in disguise, since $v + \{0\}$ is just v .

With these names, the matrix theorem of Chapter 8 reads as follows. We add the consequence for determinants and characteristic polynomials, which is what the rest of this chapter uses.

Theorem 9.1.3. Invariant Subspaces and Block Triangular Matrices

Let V be finite-dimensional, $T \in \mathcal{L}(V)$, and U a subspace with $\{0\} \neq U \neq V$. Let $\mathcal{B}_U = (u_1, \dots, u_k)$ be a basis of U , extended to a basis $\mathcal{B} = (u_1, \dots, u_k, w_1, \dots, w_l)$ of V , and put $\bar{\mathcal{B}} = (w_1 + U, \dots, w_l + U)$.

- (a) U is T -invariant if and only if $[T]_{\mathcal{B}}^{\mathcal{B}}$ has zero lower-left $l \times k$ block.
- (b) In that case

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} [T|_U]_{\mathcal{B}_U}^{\mathcal{B}_U} & \mathbf{B} \\ \mathbf{0} & [\bar{T}]_{\bar{\mathcal{B}}}^{\bar{\mathcal{B}}} \end{pmatrix}$$

for some $\mathbf{B} \in M_{k \times l}(F)$, and

$$\det T = \det(T|_U) \det \bar{T}, \quad p_T = p_{T|_U} p_{\bar{T}}.$$

Proof. Part (a), and the block form in (b), are Theorem 8.1.7 in the language of [Definition 9.1.1](#) and [Definition 9.1.2](#); $\tilde{\mathcal{B}}$ is a basis of V/U by Theorem 4.4.5.

Write $\mathbf{A} = [T|_U]_{\mathcal{B}_U}^{\mathcal{B}_U} \in M_k(F)$ and $\mathbf{D} = [\bar{T}]_{\tilde{\mathcal{B}}}^{\tilde{\mathcal{B}}} \in M_l(F)$. By Definition 7.5.1 and Theorem 7.8.4, $\det T = \det [T]_{\mathcal{B}}^{\mathcal{B}} = \det \mathbf{A} \det \mathbf{D} = \det (T|_U) \det \bar{T}$. For the characteristic polynomials,

$$x\mathbf{I}_{k+l} - [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} x\mathbf{I}_k - \mathbf{A} & -\mathbf{B} \\ \mathbf{O} & x\mathbf{I}_l - \mathbf{D} \end{pmatrix}$$

is block upper triangular with square diagonal blocks, as a matrix over $F[x]$. The block triangular determinant holds over the commutative ring $F[x]$ (the remark after Theorem 7.8.4), so $\det(x\mathbf{I} - [T]_{\mathcal{B}}^{\mathcal{B}}) = \det(x\mathbf{I}_k - \mathbf{A}) \det(x\mathbf{I}_l - \mathbf{D})$. By Definition 7.10.2, this says $p_T = p_{T|_U} p_{\bar{T}}$. ■

The theorem is the reason invariant subspaces matter. It splits the study of T into the study of two operators on smaller spaces, and every quantity that sees only the diagonal blocks splits with it. For differentiation on $\mathbb{R}[x]_{\leq 2}$, with $U = \mathbb{R}[x]_{\leq 1}$ and $\mathcal{B} = (1, x, x^2)$,

$$[D]_{\mathcal{B}}^{\mathcal{B}} = \left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{array} \right),$$

with upper-left block $[D|_U]_{\mathcal{B}_U}^{\mathcal{B}_U} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, lower-right block $[\bar{D}]_{\tilde{\mathcal{B}}}^{\tilde{\mathcal{B}}} = \mathbf{O}_1$, and $p_D = x^3 = x^2 \cdot x$. Section 7 of this chapter will run this splitting repeatedly, by induction on the dimension.

9.1.3 Sums, intersections, and direct sums

Invariant subspaces can be combined. Sums and intersections of invariant subspaces are invariant, in contrast to many other properties, and the proof is one line each.

Proposition 9.1.4. Sums and Intersections of Invariant Subspaces

Let $T \in \mathcal{L}(V)$ and let $U_1, \dots, U_k$ be T -invariant subspaces of V . Then $U_1 + \dots + U_k$ and $U_1 \cap \dots \cap U_k$ are T -invariant.

Proof. Both are subspaces, by Theorem 1.7.1 and Theorem 1.2.2. Let $u = u_1 + \dots + u_k$ with $u_i \in U_i$. Then $Tu = Tu_1 + \dots + Tu_k$ with $Tu_i \in U_i$ by invariance, so $Tu \in U_1 + \dots + U_k$. If $u \in U_1 \cap \dots \cap U_k$, then $Tu \in U_i$ for each i , by invariance of each U_i , so $Tu \in U_1 \cap \dots \cap U_k$. ■

When V is a **direct** sum of invariant subspaces, all the off-diagonal blocks vanish, not just one corner. Chapter 8 proved this for two summands (Theorem 8.1.6). The diagonalization theorem of Section 4 needs any number of summands.

Theorem 9.1.5. Invariant Direct Sums and Block Diagonal Matrices

Let V be finite-dimensional with $V = U_1 \oplus \cdots \oplus U_k$, where each $U_i \neq \{0\}$ has a basis $\mathcal{B}_i$ of length n_i . Let $\mathcal{B}$ be the list $\mathcal{B}_1, \dots, \mathcal{B}_k$ written one after another, and $T \in \mathcal{L}(V)$. Then $\mathcal{B}$ is a basis of V , and the following are equivalent:

- (a) each U_i is T -invariant;
- (b) $[T]_{\mathcal{B}}^{\mathcal{B}}$ is block diagonal with respect to the partition $n_1, \dots, n_k$.

In that case

$$[T]_{\mathcal{B}}^{\mathcal{B}} = [T|_{U_1}]_{\mathcal{B}_1}^{\mathcal{B}_1} \oplus \cdots \oplus [T|_{U_k}]_{\mathcal{B}_k}^{\mathcal{B}_k}, \quad p_T = p_{T|_{U_1}} \cdots p_{T|_{U_k}}.$$

Idea. Read the matrix one block of columns at a time. The columns of block j are the coordinates of Tb for the vectors b of $\mathcal{B}_j$. Such a Tb lies in U_j exactly when all its coordinates outside block j vanish, because a vector of U_j is already a combination of $\mathcal{B}_j$, and coordinates are unique.

Proof. By Theorem 1.7.5 ((a) $\Rightarrow$ (d)), $\mathcal{B}$ is a basis of V . Fix j and a vector b of $\mathcal{B}_j$; the column of $[T]_{\mathcal{B}}^{\mathcal{B}}$ belonging to b is $[Tb]_{\mathcal{B}}$ (Definition 4.1.1). We claim: $Tb \in U_j$ if and only if the entries of this column outside the rows of block j are all 0. If $Tb \in U_j = \text{span}(\mathcal{B}_j)$, then Tb is a combination of $\mathcal{B}_j$ alone, which is a representation in the basis $\mathcal{B}$ with zero coefficients outside block j ; by Theorem 1.5.1, it is **the** coordinate vector. Conversely, if those entries vanish, then Tb is a combination of $\mathcal{B}_j$, so it lies in U_j .

- (a) $\Rightarrow$ (b). If U_j is invariant, then $Tb \in U_j$ for each b in $\mathcal{B}_j$, so by the claim every block (i, j) with $i \neq j$ is zero. As j was arbitrary, $[T]_{\mathcal{B}}^{\mathcal{B}}$ is block diagonal. Moreover the entries of column b in the rows of block j are the coordinates of $Tb = T|_{U_j} b$ in $\mathcal{B}_j$, so the (j, j) block is $[T|_{U_j}]_{\mathcal{B}_j}^{\mathcal{B}_j}$.
- (b) $\Rightarrow$ (a). If the blocks (i, j) with $i \neq j$ vanish, the claim gives $Tb \in U_j$ for every b in $\mathcal{B}_j$, and U_j is invariant by [Lemma 9.1.1](#).

Finally, $x\mathbf{I} - [T]_{\mathcal{B}}^{\mathcal{B}}$ is block diagonal with diagonal blocks $x\mathbf{I}_{n_j} - [T|_{U_j}]_{\mathcal{B}_j}^{\mathcal{B}_j}$. Its determinant over $F[x]$ is the product of the determinants of these blocks (the remark after Theorem 7.8.4, applied by induction on the number of blocks), which is $p_T = \prod_j p_{T|_{U_j}}$ by Definition 7.10.2. ■

So an operator that preserves each piece of a direct sum decomposition is, in a suitable basis, a list of independent smaller operators. The best possible case is a decomposition into invariant **lines**: then every block is 1×1 , and $[T]_{\mathcal{B}}^{\mathcal{B}}$ is diagonal. By [Proposition 9.1.2](#), that happens exactly when V has a basis of vectors v with $Tv = \lambda v$. This observation is the starting point of the next section.

A tempting shortcut is to assume that once one invariant subspace is found, a complement of it can be chosen invariant too, and the matrix becomes block diagonal. This fails.

Warning. The complement of an invariant subspace need not be invariant. Let

$\mathbf{N} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ act on F^2 . The line $U = \text{span}(\mathbf{e}_1)$ is invariant, since $\mathbf{N}\mathbf{e}_1 = \mathbf{0}$. Every complement of U is a line $W = \text{span}((a, 1))$ for some $a \in F$ (a line not equal to U contains a vector with second entry 1). But $\mathbf{N}(a, 1) = (1, 0)$, which is not a multiple of $(a, 1)$. So **no** complement of U is $\mathbf{N}$ -invariant, and no basis $\mathcal{B}$ makes $[T_{\mathbf{N}}]_{\mathcal{B}}^{\mathcal{B}}$ block diagonal with two 1×1 blocks.

Quick check. Let U_1, U_2 be T -invariant. Must the union $U_1 \cup U_2$ be T -invariant? Must it even be a subspace?

Solution. The question has a trap: “invariant subspace” requires a subspace, and a union of two subspaces is usually not one (Proposition 1.2.3). For $T = \text{id}_{\mathbb{R}^2}$, the two axes are invariant, but their union is not a subspace, so the word does not apply. The subspace that plays the role of the union is the sum $U_1 + U_2$, which is invariant by Proposition 9.1.4.

9.1.4 Exercises

9.1.4.1 A. Check your understanding

Exercise. A1

- Define what it means for a subspace U of V to be invariant under $T \in \mathcal{L}(V)$.
- True or false: for every $T \in \mathcal{L}(V)$, the subspaces $\ker T$ and $\text{im } T$ are T -invariant. Justify your answer.
- True or false: if U is T -invariant, then $Tu = u$ for every $u \in U$. Justify your answer.
- True or false: if U is invariant under S and under T , then U is invariant under $S + T$ and under ST . Justify your answer.
- For $v \neq 0$, what condition on Tv makes $\text{span}(v)$ invariant?
- In a basis whose first k vectors form a basis of a T -invariant subspace U , which block of $[T]_{\mathcal{B}}^{\mathcal{B}}$ is zero, and what are the two diagonal blocks?

Solution.

- U is T -invariant if $Tu \in U$ for every $u \in U$ (Definition 9.1.1).
- True. If $u \in \ker T$, then $Tu = 0 \in \ker T$. If $w \in \text{im } T$, then $Tw \in \text{im } T$ by definition of the image.
- False. For $T = 2 \text{id}_{\mathbb{R}^2}$, the subspace $\mathbb{R}^2$ is invariant, but $T\mathbf{e}_1 = 2\mathbf{e}_1 \neq \mathbf{e}_1$.
- True. For $u \in U$, both Su and Tu lie in U , so $(S + T)u = Su + Tu \in U$ since U is a subspace. Also $Tu \in U$, so $(ST)u = S(Tu) \in U$ by invariance under S .
- $Tv = \lambda v$ for some scalar λ (Proposition 9.1.2).
- The lower-left block is zero. The upper-left block is $[T|_U]_{\mathcal{B}_U^U}^{\mathcal{B}_U^U}$ and the lower-right block is the matrix of the induced operator $\bar{T}$ on V/U (Theorem 9.1.3).

9.1.4.2 B. Practice

Exercise. B1: All invariant subspaces of a diagonal matrix

Let $\mathbf{A} = \text{diag}(1, 2)$ act on $\mathbb{R}^2$. Find all $\mathbf{A}$ -invariant subspaces of $\mathbb{R}^2$.

Solution. As in [Example \(A Rotation Has No Invariant Lines\)](#), a subspace of $\mathbb{R}^2$ is $\{\mathbf{0}\}$, $\mathbb{R}^2$, or a line $\text{span}(\mathbf{v})$ with $\mathbf{v} = (a, b) \neq \mathbf{0}$. The first two are invariant. By [Proposition 9.1.2](#), the line is invariant if and only if $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ for some $\lambda \in \mathbb{R}$, that is,

$$a = \lambda a, \quad 2b = \lambda b.$$

If $a \neq 0$, the first equation gives $\lambda = 1$, and then $2b = b$ gives $b = 0$; so $\mathbf{v}$ is a multiple of $\mathbf{e}_1$. If $a = 0$, then $b \neq 0$, the second equation gives $\lambda = 2$, and $\mathbf{v}$ is a multiple of $\mathbf{e}_2$. Conversely $\mathbf{A}\mathbf{e}_1 = \mathbf{e}_1$ and $\mathbf{A}\mathbf{e}_2 = 2\mathbf{e}_2$. Hence the invariant subspaces are exactly

$$\{\mathbf{0}\}, \quad \text{span}(\mathbf{e}_1), \quad \text{span}(\mathbf{e}_2), \quad \mathbb{R}^2.$$

Compare with the identity $\mathbf{I}_2$, for which every line is invariant: changing one diagonal entry from 1 to 2 destroys all lines except the two axes.

Exercise. B2: Kernels of $T - c \text{id}$

Let $T \in \mathcal{L}(V)$ and $c \in F$.

- (a) Prove directly from the definition that $\ker(T - c \text{id}_V)$ is T -invariant.
- (b) Hence show that **every** subspace of $\ker(T - c \text{id}_V)$ is T -invariant.

Solution.

- (a) Let $K = \ker(T - c \text{id}_V)$, a subspace by Theorem 2.2.1. For $u \in K$ we have $Tu - cu = 0$, that is, $Tu = cu$. Since K is a subspace, $cu \in K$. Hence $Tu \in K$, and K is T -invariant.
- (b) Let W be a subspace of K and $w \in W$. Since $w \in K$, part (a) gives $Tw = cw$, and $cw \in W$ because W is a subspace. Hence W is T -invariant. (On K , T acts as $c \text{id}$, and for a scalar operator every subspace is invariant.)

Exercise. B3: Invariance under differentiation

Let D be differentiation on $V = \mathbb{R}[x]_{\leq 3}$. Determine which of the following subspaces are D -invariant. Justify your answer.

- (a) $U_1 = \text{span}(1, x)$.
- (b) $U_2 = \text{span}(x, x^2)$.
- (c) $U_3 = \{p \in V : p(0) = 0\}$.
- (d) $U_4 = \text{span}(x^2 + 2x, 2x + 2, 2)$.
- (e) $U_5 = \text{span}(1, x^2)$, the even polynomials in V .

Solution. We use [Lemma 9.1.1](#) throughout.

- (a) Invariant: $D(1) = 0 \in U_1$ and $D(x) = 1 \in U_1$.
- (b) Not invariant: $D(x) = 1$, and $1 \notin \text{span}(x, x^2)$ since every element of U_2 has constant term 0.
- (c) Not invariant: $x \in U_3$, but $D(x) = 1$ and 1 evaluated at 0 is $1 \neq 0$.
- (d) Invariant: $D(x^2 + 2x) = 2x + 2$, $D(2x + 2) = 2$, and $D(2) = 0$, all in U_4 . (In fact $U_4 = \mathbb{R}[x]_{\leq 2}$, since the three spanning polynomials have degrees 2, 1, 0 and so are independent by Theorem 1.4.4.)
- (e) Not invariant: $D(x^2) = 2x$, which is odd and nonzero, so it is not in $\text{span}(1, x^2)$.

9.1.4.3 C. Going deeper

Exercise. C1: A chain of invariant subspaces

Let F be any field and

$$\mathbf{N} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in M_3(F),$$

acting on F^3 , so $\mathbf{N}\mathbf{e}_1 = \mathbf{0}$, $\mathbf{N}\mathbf{e}_2 = \mathbf{e}_1$, $\mathbf{N}\mathbf{e}_3 = \mathbf{e}_2$. Put $W_0 = \{\mathbf{0}\}$, $W_1 = \text{span}(\mathbf{e}_1)$, $W_2 = \text{span}(\mathbf{e}_1, \mathbf{e}_2)$, $W_3 = F^3$.

- (a) Show that each W_j is $\mathbf{N}$ -invariant.
- (b) Let U be an $\mathbf{N}$ -invariant subspace, and let j be the largest index such that some vector of U has a nonzero j -th entry (put $j = 0$ if $U = \{\mathbf{0}\}$). Prove that $U = W_j$.
- (c) Deduce that the $\mathbf{N}$ -invariant subspaces are exactly $W_0 \subset W_1 \subset W_2 \subset W_3$, and that any two of them are comparable under inclusion.

Hint: for (b), apply $\mathbf{N}$ and $\mathbf{N}^2$ to a vector with nonzero j -th entry.

Solution.

- (a) W_0 and W_3 are invariant for every operator. By [Lemma 9.1.1](#), W_1 is invariant since $\mathbf{N}\mathbf{e}_1 = \mathbf{0}$, and W_2 is invariant since $\mathbf{N}\mathbf{e}_1 = \mathbf{0}$ and $\mathbf{N}\mathbf{e}_2 = \mathbf{e}_1$ lie in W_2 .
- (b) If $U = \{\mathbf{0}\}$, then $j = 0$ and $U = W_0$. Otherwise $j \geq 1$. By the choice of j , every vector of U has entries 0 beyond position j , so $U \subseteq W_j$. Pick $\mathbf{v} = (a, b, c) \in U$ with nonzero j -th entry. We have $\mathbf{N}\mathbf{v} = (b, c, 0)$ and $\mathbf{N}^2\mathbf{v} = (c, 0, 0)$, and all of these lie in U by invariance.
 - Case $j = 3$. Here $c \neq 0$. Then $\mathbf{e}_1 = c^{-1}\mathbf{N}^2\mathbf{v} \in U$, $\mathbf{e}_2 = c^{-1}(\mathbf{N}\mathbf{v} - b\mathbf{e}_1) \in U$, and $\mathbf{e}_3 = c^{-1}(\mathbf{v} - a\mathbf{e}_1 - b\mathbf{e}_2) \in U$. So $W_3 \subseteq U$.
 - Case $j = 2$. Here $c = 0$ and $b \neq 0$. Then $\mathbf{e}_1 = b^{-1}\mathbf{N}\mathbf{v} \in U$ and $\mathbf{e}_2 = b^{-1}(\mathbf{v} - a\mathbf{e}_1) \in U$, so $W_2 \subseteq U$.
 - Case $j = 1$. Here $b = c = 0$ and $a \neq 0$, so $\mathbf{e}_1 = a^{-1}\mathbf{v} \in U$ and $W_1 \subseteq U$.
 In each case $W_j \subseteq U \subseteq W_j$, so $U = W_j$.
- (c) By (a) each W_j is invariant, and by (b) every invariant subspace is one of them. Since $W_0 \subset W_1 \subset W_2 \subset W_3$, any two are comparable. In particular $\mathbf{N}$ has exactly one invariant line, W_1 , and exactly one invariant plane, W_2 . Contrast [Exercise \(B1: All invariant subspaces of a diagonal matrix\)](#), where $\text{diag}(1, 2)$ has two invariant lines (in

the sense of Proposition 9.1.2), neither containing the other.

Exercise. C2: Invariance under the inverse

- (a) Let V be finite-dimensional, $T \in \mathcal{L}(V)$ invertible, and U a T -invariant subspace. Prove that $T(U) = U$, and deduce that U is T^{-1} -invariant.
- (b) Show that (a) fails without finite dimension. On $V = F^{\mathbb{Z}}$, the space of all functions $\mathbb{Z} \rightarrow F$ written as two-sided sequences $(s_n)_{n \in \mathbb{Z}}$, let $(Ss)_n = s_{n-1}$. Show that S is invertible, that $U = \{s : s_n = 0 \text{ for all } n < 0\}$ is S -invariant, and that U is **not** S^{-1} -invariant.

Hint: for (a), consider the restriction $T|_U$ and its kernel.

Solution.

- (a) The restriction $T|_U \in \mathcal{L}(U)$ exists by invariance. Its kernel is $\ker T \cap U = \{0\}$, since T is injective, so $T|_U$ is injective (Theorem 2.2.3). The space U is finite-dimensional (Theorem 1.6.7), so by Theorem 2.5.7, $T|_U$ is surjective onto U . That is, $T(U) = U$. Now let $u \in U$. By surjectivity of $T|_U$, there is $u' \in U$ with $Tu' = u$. Applying T^{-1} gives $T^{-1}u = u' \in U$. Hence U is T^{-1} -invariant.
- (b) $V = F^{\mathbb{Z}}$ is a vector space with pointwise operations (Example (The Standard Vector Spaces)), and S is linear because it only re-indexes entries. Define $(S's)_n = s_{n+1}$. Then $(S'Ss)_n = (Ss)_{n+1} = s_n$ and $(SS's)_n = (S's)_{n-1} = s_n$, so S is invertible with $S^{-1} = S'$.

U is a subspace (the conditions $s_n = 0$ are preserved by sums and scalar multiples). If $s \in U$ and $n < 0$, then $n - 1 < 0$, so $(Ss)_n = s_{n-1} = 0$; thus $Ss \in U$, and U is S -invariant.

Let $\delta \in U$ be the sequence with $\delta_0 = 1$ and $\delta_n = 0$ for $n \neq 0$. Then $(S^{-1}\delta)_{-1} = \delta_0 = 1 \neq 0$, so $S^{-1}\delta \notin U$. Hence U is not S^{-1} -invariant. In (a), the step that breaks is “ $T|_U$ injective, hence surjective”, which needs finite dimension: here $S|_U$ is injective but its image misses δ .