

11.7 Isometries and Unitary Maps

Chapter 2 asked which linear maps are invertible. Now that we can measure, we can ask a sharper question: which linear maps leave the measurements alone? A map that never changes a length never changes an angle either, and the two facts turn out to be the same fact. This section identifies these maps in five equivalent ways, collects them into the orthogonal and unitary groups, and classifies them completely in dimensions two and three, where they are exactly the rotations and reflections that the words suggest. At the end we drop linearity from the hypothesis and get it back as a conclusion.

As throughout the chapter, F is $\mathbb{R}$ or $\mathbb{C}$, inner products are linear in the first slot, and spaces are finite-dimensional unless stated otherwise.

11.7.1 Maps that do not change length

Take the plane with its usual inner product and ask which linear maps move the picture rigidly. Turning the plane about the origin clearly does not change any length. Flipping it across a line through the origin does not either. Scaling by 2 obviously does. Shearing, $(x_1, x_2) \mapsto (x_1 + x_2, x_2)$, also does: it fixes $\mathbf{e}_1$ but sends $\mathbf{e}_2$ to $(1, 1)$, of length $\sqrt{2}$. So “does not change length” is a genuine restriction, and it is worth a name.

An isometry is a linear map that every ruler agrees with: the image of a vector is exactly as long as the vector.

Definition 11.7.1. Isometry

Let V and W be inner product spaces over F . A linear map $T \in \mathcal{L}(V, W)$ is an **isometry** if

$$\|Tv\| = \|v\| \quad \text{for every } v \in V.$$

In words: the condition is required for **every** vector, not merely for the vectors of a basis, and it compares a norm in W with a norm in V . The word comes from *iso* (equal) and *metron* (measure). Some authors reserve “isometry” for a bijective such map; we do not, and [Definition 11.7.2](#) below names the bijective case separately.

Two consequences are immediate. An isometry is injective, since $Tv = 0$ forces $\|v\| = \|Tv\| = 0$ and hence $v = 0$ (Theorem 2.2.3). And if $Tv = \lambda v$ with $v \neq 0$, then $\|v\| = \|\lambda v\| = |\lambda| \|v\|$ by Theorem 11.1.2 (b), so $|\lambda| = 1$. We record the second observation, which the classification below will use twice.

Proposition 11.7.1. Eigenvalues of an Isometry Have Modulus 1

Let $T \in \mathcal{L}(V)$ be an isometry of an inner product space over F , and let $\lambda \in F$ be an eigenvalue of T . Then $|\lambda| = 1$. In particular, if $F = \mathbb{R}$ then $\lambda = 1$ or $\lambda = -1$.

Proof. Let $v \neq 0$ satisfy $Tv = \lambda v$ (Definition 9.2.1). Then $\|v\| = \|Tv\| = \|\lambda v\| = |\lambda| \|v\|$ by Theorem 11.1.2 (b), and $\|v\| \neq 0$, so dividing gives $|\lambda| = 1$. The only real numbers of modulus 1 are ± 1 . ■

Example. Isometries and Near-Isometries

Decide which of the following linear maps are isometries, with the standard inner products.

- (a) $T: \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $T(x_1, x_2) = (x_1, x_2, 0)$.
- (b) $T_{\mathbf{P}}: F^n \rightarrow F^n$ for a permutation matrix $\mathbf{P}$ (Definition 3.4.2).
- (c) $T: \mathbb{C} \rightarrow \mathbb{C}$, $T(z) = \alpha z$, where $|\alpha| = 1$.
- (d) $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $T(x_1, x_2) = (x_1 + x_2, x_2)$.
- (e) The orthogonal projection P_U onto a subspace U with $\{0\} \neq U \neq V$.

Solution.

- (a) Yes: $\|T\mathbf{x}\|^2 = x_1^2 + x_2^2 + 0 = \|\mathbf{x}\|^2$. Note that T is not surjective, so an isometry need not be invertible.
- (b) Yes. A permutation matrix has exactly one 1 in each row and column, so $\mathbf{P}\mathbf{x}$ is the list $x_1, \dots, x_n$ in a different order, and $\|\mathbf{P}\mathbf{x}\|^2 = \sum_i |x_{\sigma(i)}|^2 = \sum_i |x_i|^2$, the same sum reordered.
- (c) Yes: $\|\alpha z\| = |\alpha z| = |\alpha||z| = |z|$.
- (d) No: $\|T\mathbf{e}_2\| = \|(1, 1)\| = \sqrt{2} \neq 1 = \|\mathbf{e}_2\|$. One vector is enough to refute.
- (e) No: pick $v \in U^\perp$ with $v \neq 0$, which exists because $U \neq V$ (Theorem 11.3.2 (a)). Then $P_U v = 0$ (Theorem 11.3.3 (a)), so $\|P_U v\| = 0 \neq \|v\|$. A projection destroys length by design.

11.7.2 Five ways to say the same thing

The definition mentions only norms, but norms and inner products determine each other. In one direction this is Definition 11.1.3. In the other it is the following identity, which reconstructs the inner product from lengths alone; it is the reason that preserving length automatically preserves angle.

Lemma 11.7.2. Polarization Identities

Let V be an inner product space over F and let $u, v \in V$.

- (a) If $F = \mathbb{R}$, then $\langle u, v \rangle = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2)$.
- (b) If $F = \mathbb{C}$, then $\langle u, v \rangle = \frac{1}{4} \sum_{k=0}^3 i^k \|u + i^k v\|^2$.

Idea. Expand every norm squared with $\|u + w\|^2 = \|u\|^2 + 2\operatorname{Re}\langle u, w \rangle + \|w\|^2$ (Theorem 11.1.2 (c)). In the real case the two expansions differ only in the middle term, which survives doubled. In the complex case the four terms $\|u\|^2 + \|w\|^2$ are the same in all four summands and are killed by $\sum_k i^k = 0$; what is left is a sum in which the $\langle u, v \rangle$ parts reinforce and the $\overline{\langle u, v \rangle}$ parts cancel.

Proof.

- (a) By Theorem 11.1.2 (c), $\|u \pm v\|^2 = \|u\|^2 \pm 2\langle u, v \rangle + \|v\|^2$, since over $\mathbb{R}$ the inner product is its own real part. Subtracting, $\|u + v\|^2 - \|u - v\|^2 = 4\langle u, v \rangle$.
- (b) Write $z = \langle u, v \rangle$. Since $|i^k| = 1$, Theorem 11.1.2 (b) and (c) give $\|i^k v\| = \|v\|$ and

$$\|u + i^k v\|^2 = \|u\|^2 + \|v\|^2 + 2\operatorname{Re}\langle u, i^k v \rangle = \|u\|^2 + \|v\|^2 + 2\operatorname{Re}(i^{-k} z),$$

where $\langle u, i^k v \rangle = \overline{i^k} z = i^{-k} z$ by conjugate-linearity in the second slot. Multiply by i^k and sum over $k = 0, 1, 2, 3$. The first two terms contribute $(\|u\|^2 + \|v\|^2) \sum_k i^k = 0$, because $1 + i - 1 - i = 0$. For the remaining terms use $2\operatorname{Re}(w) = w + \bar{w}$:

$$\sum_{k=0}^3 i^k \cdot 2\operatorname{Re}(i^{-k} z) = \sum_{k=0}^3 i^k (i^{-k} z + i^k \bar{z}) = \sum_{k=0}^3 z + \bar{z} \sum_{k=0}^3 i^{2k} = 4z + 0,$$

since $\sum_k i^{2k} = 1 - 1 + 1 - 1 = 0$. Dividing by 4 gives the identity. ■

Now the main theorem. It says that length, angle, the adjoint, and orthonormal bases all detect the same class of maps, and it is the reason isometries are easy to work with: whichever of the five descriptions a problem hands you, you may switch to whichever one you want.

Theorem 11.7.3. Characterizations of Isometries

Let V and W be finite-dimensional inner product spaces over F , and let $T \in \mathcal{L}(V, W)$. The following are equivalent.

- (a) T is an isometry: $\|Tv\| = \|v\|$ for every $v \in V$.
- (b) T preserves inner products: $\langle Tu, Tv \rangle = \langle u, v \rangle$ for all $u, v \in V$.
- (c) $T^*T = \operatorname{id}_V$.
- (d) T sends **every** orthonormal list in V to an orthonormal list in W .
- (e) T sends **some** orthonormal basis of V to an orthonormal list in W .
- (f) For one (equivalently every) pair of orthonormal bases $\mathcal{B}$ of V and $\mathcal{C}$ of W , the matrix $\mathbf{A} = [T]_{\mathcal{B}}^{\mathcal{C}} \in M_{m \times n}(F)$, where $n = \dim V$ and $m = \dim W$, satisfies $\mathbf{A}^* \mathbf{A} = \mathbf{I}_n$; that is, the n columns of $\mathbf{A}$ are orthonormal vectors of F^m .

If in addition $\dim W = \dim V$, then (a)–(f) are further equivalent to each of

- (g1) T is invertible with $T^{-1} = T^*$;
- (g2) $TT^* = \operatorname{id}_W$.

Idea. Run the cycle (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (a). The first step is polarization: lengths are all we are given, and [Lemma 11.7.2](#) converts them into an inner product. The second is the adjoint doing its job, $\langle Tu, Tv \rangle = \langle u, T^*Tv \rangle$, plus nondegeneracy. The third is the same identity read backwards with $u = v$. Then (b) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (b) is a second, shorter cycle, and (c) $\Leftrightarrow$ (f) is Theorem 11.6.3. The last two need a dimension count: $T^*T = \operatorname{id}$ makes T injective, and in equal dimensions injective means invertible.

Proof.

- (a) $\Rightarrow$ (b). Suppose $\|Tv\| = \|v\|$ for all v . If $F = \mathbb{R}$, then by Lemma 11.7.2 (a) applied in W and then in V , and by the linearity of T ,

$$\begin{aligned}\langle Tu, Tv \rangle &= \frac{1}{4}(\|T(u+v)\|^2 - \|T(u-v)\|^2) \\ &= \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2) = \langle u, v \rangle.\end{aligned}$$

If $F = \mathbb{C}$, the same two lines work with Lemma 11.7.2 (b), since $T(u+iv) = Tu+ivTv$.

- (b) $\Rightarrow$ (c). For all $u, v \in V$, Definition 11.6.1 gives $\langle u, T^*Tv \rangle = \langle Tu, Tv \rangle = \langle u, v \rangle$, so $\langle u, T^*Tv - v \rangle = 0$ for every u . Taking $u = T^*Tv - v$ makes its norm zero, so $T^*Tv = v$ for every v .
- (c) $\Rightarrow$ (a). $\|Tv\|^2 = \langle Tv, Tv \rangle = \langle v, T^*Tv \rangle = \langle v, v \rangle = \|v\|^2$, and norms are nonnegative, so $\|Tv\| = \|v\|$.
- (d) $\Rightarrow$ (d). Let $(u_1, \dots, u_k)$ be orthonormal in V , so $\langle u_i, u_j \rangle = \delta_{ij}$. Then $\langle Tu_i, Tu_j \rangle = \langle u_i, u_j \rangle = \delta_{ij}$, which is exactly what it means for $(Tu_1, \dots, Tu_k)$ to be orthonormal (Definition 11.2.1).
- (e) $\Rightarrow$ (e). V is finite-dimensional, so it has an orthonormal basis by Theorem 11.2.2, and (d) applies to it.
- (f) $\Rightarrow$ (b). Let $\mathcal{B} = (v_1, \dots, v_n)$ be an orthonormal basis of V with $(Tv_1, \dots, Tv_n)$ orthonormal. Given $u, v \in V$, expand them in $\mathcal{B}$ as $u = \sum_i a_i v_i$ and $v = \sum_j b_j v_j$. Expanding both slots and using orthonormality twice,

$$\langle Tu, Tv \rangle = \sum_{i,j} a_i \bar{b}_j \langle Tv_i, Tv_j \rangle = \sum_i a_i \bar{b}_i = \sum_{i,j} a_i \bar{b}_j \langle v_i, v_j \rangle = \langle u, v \rangle,$$

since in both double sums only the terms with $i = j$ survive.

- (g) $\Leftrightarrow$ (f). Fix orthonormal bases $\mathcal{B}$ and $\mathcal{C}$ and put $\mathbf{A} = [T]_{\mathcal{B}}^{\mathcal{C}}$. By Theorem 11.6.3, $[T^*]_{\mathcal{C}}^{\mathcal{B}} = \mathbf{A}^*$, and by Theorem 4.1.2, $[T^*T]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{A}^* \mathbf{A}$. Since a map is determined by its matrix in fixed bases, $T^*T = \text{id}_V$ if and only if $\mathbf{A}^* \mathbf{A} = \mathbf{I}_n$. The (i, j) entry of $\mathbf{A}^* \mathbf{A}$ is the standard inner product of column j with column i , so this says the columns are orthonormal. As (c) does not mention the bases, the condition holds for one pair of orthonormal bases if and only if it holds for every pair.

Now assume $\dim W = \dim V$.

- (c) $\Rightarrow$ (g1). By (c) and Theorem 2.2.3, T is injective, so $\text{rank } T = \dim V = \dim W$ by Theorem 2.3.1, and T is surjective. Hence T is invertible, and multiplying $T^*T = \text{id}_V$ on the right by T^{-1} gives $T^* = T^{-1}$.

(g1) $\Rightarrow$ (g2). $TT^* = TT^{-1} = \text{id}_W$.

(g2) $\Rightarrow$ (c). $TT^* = \text{id}_W$ makes T surjective, hence invertible by the equal dimensions and Theorem 2.3.1, and multiplying on the left by T^{-1} gives $T^* = T^{-1}$, so $T^*T = \text{id}_V$. This proves the theorem. ■

Condition (f) is the one to use on a concrete matrix: **check that the columns are orthonormal**. Condition (c) is the one to use in a proof. Condition (e) is the cheapest to verify on an abstract map, since one basis suffices, and (d) is what you get out afterwards.

Warning. $T^*T = \text{id}$ **does not give** $TT^* = \text{id}$ **when the dimensions differ.** Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ be $T(x_1, x_2) = (x_1, x_2, 0)$, with matrix $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then $\mathbf{A}^\top \mathbf{A} = \mathbf{I}_2$, so T is an isometry; but

$$\mathbf{A}\mathbf{A}^\top = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq \mathbf{I}_3.$$

Ask *identity on what?* The composite TT^* is the orthogonal projection of $\mathbb{R}^3$ onto $\text{im } T$, the x_1x_2 -plane, and it is the identity there and nowhere else. Only when T is surjective can both products be identities.

Quick check. Is $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ the matrix of an isometry of $\mathbb{R}^2$? If not, is some scalar multiple of it?

Solution. No: the first column has norm $\sqrt{2} \neq 1$, so the columns are not orthonormal and condition (f) of [Theorem 11.7.3](#) fails. The columns are, however, orthogonal, so dividing by $\sqrt{2}$ normalizes both and $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ is an isometry.

11.7.3 Unitary and orthogonal maps

Conditions (g1) and (g2) describe the case that matters most, an isometry of a space with itself. It is the case with a name.

Definition 11.7.2. Unitary and Orthogonal

Let V be a finite-dimensional inner product space over F . An operator $T \in \mathcal{L}(V)$ is **unitary** if

$$T^*T = TT^* = \text{id}_V,$$

equivalently if T is an invertible isometry. When $F = \mathbb{R}$, such a T is also called **orthogonal**.

A matrix $\mathbf{A} \in M_n(\mathbb{C})$ with $\mathbf{A}^* \mathbf{A} = \mathbf{I}_n$ is a **unitary matrix**; a matrix $\mathbf{A} \in M_n(\mathbb{R})$ with $\mathbf{A}^\top \mathbf{A} = \mathbf{I}_n$ is an **orthogonal matrix**.

For a **square** matrix the one-sided condition $\mathbf{A}^* \mathbf{A} = \mathbf{I}$ already forces $\mathbf{A} \mathbf{A}^* = \mathbf{I}$, by part (g2) of [Theorem 11.7.3](#), or directly by Theorem 3.4.5. The warning above shows that squareness is not a formality.

Definition 11.7.3. The Unitary and Orthogonal Groups

For $n \geq 1$ set

$$U(n) := \{\mathbf{A} \in M_n(\mathbb{C}) : \mathbf{A}^* \mathbf{A} = \mathbf{I}_n\}, \quad O(n) := \{\mathbf{A} \in M_n(\mathbb{R}) : \mathbf{A}^\top \mathbf{A} = \mathbf{I}_n\},$$

and $SO(n) := \{\mathbf{A} \in O(n) : \det \mathbf{A} = 1\}$, the **special orthogonal group**.

Proposition 11.7.4. The Groups and Their Determinants

- (a) $U(n)$ and $O(n)$ are groups under matrix multiplication, and $O(n) = U(n) \cap M_n(\mathbb{R})$.
- (b) If $\mathbf{A} \in U(n)$ then $|\det \mathbf{A}| = 1$. If $\mathbf{A} \in O(n)$ then $\det \mathbf{A} = 1$ or $\det \mathbf{A} = -1$.
- (c) $SO(n)$ is a subgroup of $O(n)$.

Proof.

- (a) $\mathbf{I}^* \mathbf{I} = \mathbf{I}$, so $\mathbf{I} \in U(n)$. If $\mathbf{A}, \mathbf{B} \in U(n)$ then $(\mathbf{AB})^*(\mathbf{AB}) = \mathbf{B}^* \mathbf{A}^* \mathbf{AB} = \mathbf{B}^* \mathbf{B} = \mathbf{I}$, so $\mathbf{AB} \in U(n)$. Each $\mathbf{A} \in U(n)$ is invertible with $\mathbf{A}^{-1} = \mathbf{A}^*$, and $(\mathbf{A}^*)^* \mathbf{A}^* = \mathbf{AA}^* = \mathbf{I}$, so $\mathbf{A}^{-1} \in U(n)$. Associativity is inherited from matrix multiplication. The same three lines prove the claim for $O(n)$, with $\mathbf{A}^* = \mathbf{A}^\top$ for a real matrix, which is also the last assertion.
- (b) The determinant is the sum $\sum_{\sigma} \text{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}$ of Definition 7.4.1, and conjugation of complex numbers respects sums and products, so $\det \overline{\mathbf{A}} = \overline{\det \mathbf{A}}$. With Theorem 7.4.5 this gives $\det \mathbf{A}^* = \det(\overline{\mathbf{A}}^\top) = \det \overline{\mathbf{A}} = \overline{\det \mathbf{A}}$. Hence, by Theorem 7.5.4,

$$1 = \det \mathbf{I}_n = \det(\mathbf{A}^* \mathbf{A}) = \overline{\det \mathbf{A}} \det \mathbf{A} = |\det \mathbf{A}|^2.$$

For real $\mathbf{A}$ the determinant is real, and the real numbers of modulus 1 are ± 1 .

- (c) $\det \mathbf{I} = 1$, and $\det$ is multiplicative (Theorem 7.5.4), so $\det(\mathbf{AB}) = 1$ and $\det(\mathbf{A}^{-1}) = (\det \mathbf{A})^{-1} = 1$ whenever $\det \mathbf{A} = \det \mathbf{B} = 1$. ■

Warning. $|\lambda| = 1$ **does not mean** $\lambda = \pm 1$, and $\det \mathbf{A} = 1$ **does not mean** “rotation”. The matrix $-\mathbf{I}_4 \in SO(4)$ has determinant $(-1)^4 = 1$, but it fixes no nonzero vector, so it is not a rotation about an axis in any sense. Likewise $-\mathbf{I}_3 \in O(3)$ has determinant -1 but is not a reflection in a plane: a reflection fixes a plane pointwise, and $-\mathbf{I}_3$ fixes only $\mathbf{0}$. The clean classification below is a feature of dimensions 2 and 3, not a general theorem.

11.7.4 The Frobenius norm and unitary factors

Section 1 made $M_{m \times n}(F)$ into an inner product space by declaring

$$\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr}(\mathbf{B}^* \mathbf{A}),$$

the **Frobenius inner product**, and observing that it is nothing but the standard inner product of F^{mn} after the entries are listed in a single column. Its induced norm (Definition 11.1.3) was named there but never given a symbol. This is the place to give it one, because the single property that makes it useful is a statement about unitary factors, and unitary factors have only just arrived.

For $\mathbf{A} \in M_{m \times n}(\mathbb{C})$, the **Frobenius norm** of $\mathbf{A}$ is

$$\|\mathbf{A}\|_F := \sqrt{\langle \mathbf{A}, \mathbf{A} \rangle} = \left(\sum_{i,j} |a_{ij}|^2 \right)^{1/2}.$$

The subscript F is there because $\|\cdot\|$ with no subscript already means the norm of a *vector*, and a matrix will later be asked to be both. Two remarks on the formula. First, it treats $\mathbf{A}$ as a list of mn numbers and forgets that they are arranged in a rectangle: it is blind to the difference between $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Second, it is computable with no eigenvalues, no row reduction, and no choices: add up mn squared moduli.

The trace form of the same quantity is the one that does the work.

Quick check. Why is $\|\mathbf{A}\|_F^2 = \text{tr}(\mathbf{A}^* \mathbf{A})$?

Solution. The (j, j) entry of $\mathbf{A}^* \mathbf{A}$ is $\sum_i \overline{a_{ij}} a_{ij} = \sum_i |a_{ij}|^2$, the squared length of the j -th column. Summing over j sweeps every entry of $\mathbf{A}$ exactly once, so $\text{tr}(\mathbf{A}^* \mathbf{A}) = \sum_{i,j} |a_{ij}|^2$. Equivalently, it is $\langle \mathbf{A}, \mathbf{A} \rangle$ read off the definition of the Frobenius inner product.

Now the property that makes this the right yardstick whenever orthonormal bases are in play: unitary factors are invisible to it.

Lemma 11.7.5. Unitary Invariance of the Frobenius Norm

Let $\mathbf{A} \in M_{m \times n}(\mathbb{C})$, let $\mathbf{U} \in U(m)$ and let $\mathbf{V} \in U(n)$. Then

$$\|\mathbf{U}\mathbf{A}\|_F = \|\mathbf{A}\|_F = \|\mathbf{A}\mathbf{V}\|_F.$$

In particular, if $\mathbf{A}$ is square and $\mathbf{U}$ is unitary then $\|\mathbf{U}^* \mathbf{A} \mathbf{U}\|_F = \|\mathbf{A}\|_F$.

Proof. Using $\|\mathbf{B}\|_F^2 = \text{tr}(\mathbf{B}^* \mathbf{B})$, then $(\mathbf{U}\mathbf{A})^* = \mathbf{A}^* \mathbf{U}^*$ and $\mathbf{U}^* \mathbf{U} = \mathbf{I}_m$,

$$\|\mathbf{U}\mathbf{A}\|_F^2 = \text{tr}(\mathbf{A}^* \mathbf{U}^* \mathbf{U} \mathbf{A}) = \text{tr}(\mathbf{A}^* \mathbf{A}) = \|\mathbf{A}\|_F^2.$$

For the right-hand factor, Theorem 0.9.6 (3) lets the $\mathbf{V}$ travel around the trace, and $\mathbf{V}\mathbf{V}^* = \mathbf{I}_n$:

$$\|\mathbf{A}\mathbf{V}\|_F^2 = \text{tr}(\mathbf{V}^* \mathbf{A}^* \mathbf{A} \mathbf{V}) = \text{tr}(\mathbf{A}^* \mathbf{A} \mathbf{V} \mathbf{V}^*) = \|\mathbf{A}\|_F^2.$$

Both norms are nonnegative reals, so equality of the squares gives equality of the norms. For the last statement, $\mathbf{U}^*$ is unitary too, since $(\mathbf{U}^*)^* \mathbf{U}^* = \mathbf{U} \mathbf{U}^* = \mathbf{I}$; apply the first equality to the left factor $\mathbf{U}^*$ and the second to the right factor $\mathbf{U}$. ■

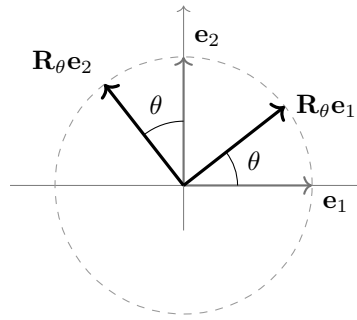
The consequence to keep in mind is this. If $T \in \mathcal{L}(V)$ is an operator on a finite-dimensional complex inner product space and we compute $\|[T]_{\mathcal{B}}\|_F$ in an **orthonormal** basis $\mathcal{B}$, the answer does not depend on which orthonormal basis we chose, because two such matrices differ by a unitary similarity (Theorem 4.2.3). So the Frobenius norm is a property of T , not of our

coordinates — exactly like the trace and the determinant, and unlike, say, the largest entry.

11.7.5 Rotations and reflections of the plane

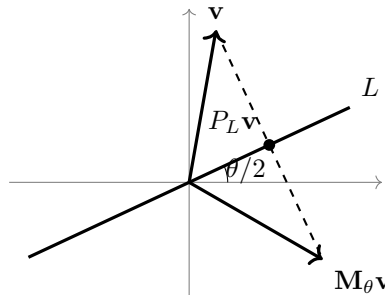
In dimension 2 the group $O(2)$ can be written down completely, and the two determinant values split it into exactly the two families the words suggest. For $\theta \in \mathbb{R}$ put

$$\mathbf{R}_\theta := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad \mathbf{M}_\theta := \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}.$$



the columns of $\mathbf{R}_\theta$ are the rotated axes

The first matrix rotates the plane by θ : its columns are $\mathbf{e}_1$ and $\mathbf{e}_2$ turned through θ . The second is a reflection, and the picture says in which mirror.



$\mathbf{M}_\theta$ reflects in the line L at angle $\theta/2$

Theorem 11.7.6. Classification of the Plane Isometries

Let $\mathbf{A} \in O(2)$.

- (a) If $\det \mathbf{A} = 1$, then $\mathbf{A} = \mathbf{R}_\theta$ for a unique $\theta \in [0, 2\pi)$.
- (b) If $\det \mathbf{A} = -1$, then $\mathbf{A} = \mathbf{M}_\theta$ for a unique $\theta \in [0, 2\pi)$. In this case $\mathbf{A}^2 = \mathbf{I}$, the eigenvalues of $\mathbf{A}$ are 1 and -1 , and $\mathbf{A}$ is the reflection in the line $L = \text{span}((\cos \frac{\theta}{2}, \sin \frac{\theta}{2}))$: it fixes L pointwise and negates $L^\perp$.

Consequently $O(2) = \{\mathbf{R}_\theta\} \cup \{\mathbf{M}_\theta\}$, and $SO(2) = \{\mathbf{R}_\theta : \theta \in [0, 2\pi)\}$.

Idea. Condition (f) of Theorem 11.7.3 says the two columns are orthonormal. A unit vector of $\mathbb{R}^2$ is $(\cos \theta, \sin \theta)$ for exactly one $\theta \in [0, 2\pi)$, and there are exactly two unit vectors orthogonal to it. So there are only two candidates once the first column is chosen, and the determinant tells them apart. The geometric description of $\mathbf{M}_\theta$ is then a two-line verification:

guess the fixed vector from the picture and check it.

Proof. Write $\mathbf{A} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in O(2)$. By [Theorem 11.7.3](#) (f), the columns are orthonormal: $a^2 + b^2 = 1$, $c^2 + d^2 = 1$, and $ac + bd = 0$. Since (a, b) is a unit vector of $\mathbb{R}^2$, there is a unique $\theta \in [0, 2\pi)$ with $a = \cos \theta$ and $b = \sin \theta$. The vectors orthogonal to (a, b) form a line, spanned by $(-b, a)$, which is already a unit vector; so the unit vectors on that line are $\pm(-b, a)$, and $(c, d) = \pm(-\sin \theta, \cos \theta)$. The two choices give

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \mathbf{R}_\theta \quad \text{and} \quad \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} = \mathbf{M}_\theta,$$

with determinants $\cos^2 \theta + \sin^2 \theta = 1$ and $-\cos^2 \theta - \sin^2 \theta = -1$. This proves (a) and the first sentence of (b), and uniqueness of θ in both cases, since θ is already determined by the first column.

For the rest of (b), put $\mathbf{u} = (\cos \frac{\theta}{2}, \sin \frac{\theta}{2})$ and $\mathbf{w} = (-\sin \frac{\theta}{2}, \cos \frac{\theta}{2})$, two orthogonal unit vectors. Writing $\theta = \frac{\theta}{2} + \frac{\theta}{2}$ and using the subtraction formulas for cosine and sine,

$$\mathbf{M}_\theta \mathbf{u} = \begin{pmatrix} \cos \theta \cos \frac{\theta}{2} + \sin \theta \sin \frac{\theta}{2} \\ \sin \theta \cos \frac{\theta}{2} - \cos \theta \sin \frac{\theta}{2} \end{pmatrix} = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} \end{pmatrix} = \mathbf{u},$$

$$\mathbf{M}_\theta \mathbf{w} = \begin{pmatrix} -\cos \theta \sin \frac{\theta}{2} + \sin \theta \cos \frac{\theta}{2} \\ -\sin \theta \sin \frac{\theta}{2} - \cos \theta \cos \frac{\theta}{2} \end{pmatrix} = \begin{pmatrix} \sin \frac{\theta}{2} \\ -\cos \frac{\theta}{2} \end{pmatrix} = -\mathbf{w}.$$

So $(\mathbf{u}, \mathbf{w})$ is an orthonormal basis of eigenvectors with eigenvalues 1 and -1 , and $\mathbf{M}_\theta^2$ fixes both of them, hence $\mathbf{M}_\theta^2 = \mathbf{I}$. Since $L = \text{span}(\mathbf{u})$, we have $L^\perp = \text{span}(\mathbf{w})$, because $\mathbf{w} \perp \mathbf{u}$ and $\dim L^\perp = 1$ ([Theorem 11.3.2](#) (c)), so $\mathbf{M}_\theta$ fixes L pointwise and negates $L^\perp$, which is the description of a reflection. This proves the theorem. ■

So $SO(2)$ is the circle of rotations, and the other half of $O(2)$ is the circle of reflections. Multiplying two rotations adds the angles, $\mathbf{R}_\alpha \mathbf{R}_\beta = \mathbf{R}_{\alpha+\beta}$, so $SO(2)$ is abelian; [Exercise \(C3: Two reflections make a rotation\)](#) shows what happens when two reflections are multiplied.

Example. Naming Two Plane Isometries

Classify each of the following as a rotation or a reflection, and give the angle or the mirror line.

(a) $\mathbf{A} = \frac{1}{2} \begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}.$

(b) $\mathbf{B} = \frac{1}{5} \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}.$

Solution.

- (a) The columns $\frac{1}{2}(1, \sqrt{3})$ and $\frac{1}{2}(-\sqrt{3}, 1)$ have norm $\frac{1}{2}\sqrt{1+3} = 1$ and inner product $\frac{1}{4}(-\sqrt{3} + \sqrt{3}) = 0$, so $\mathbf{A} \in \text{O}(2)$. Its determinant is $\frac{1}{4}(1+3) = 1$, so $\mathbf{A} = \mathbf{R}_\theta$ with $\cos \theta = \frac{1}{2}$ and $\sin \theta = \frac{\sqrt{3}}{2}$: rotation by $\theta = \pi/3$.
- (b) The columns $\frac{1}{5}(3, 4)$ and $\frac{1}{5}(4, -3)$ are orthonormal, and $\det \mathbf{B} = \frac{1}{25}(-9 - 16) = -1$, so $\mathbf{B} = \mathbf{M}_\theta$ with $\cos \theta = \frac{3}{5}$ and $\sin \theta = \frac{4}{5}$. The mirror is $L = \text{span}((\cos \frac{\theta}{2}, \sin \frac{\theta}{2}))$, and rather than halve the angle we find the fixed line directly: $\mathbf{B}\mathbf{x} = \mathbf{x}$ reads $3x_1 + 4x_2 = 5x_1$, that is $x_1 = 2x_2$. So $L = \text{span}((2, 1))$. Check: $\mathbf{B}(2, 1) = \frac{1}{5}(6 + 4, 8 - 3) = (2, 1)$, and $\mathbf{B}(-1, 2) = \frac{1}{5}(-3 + 8, -4 - 6) = (1, -2) = -(-1, 2)$, as a reflection must do to $L^\perp$.

11.7.6 Rotations of space

In dimension 3 the answer is the one physics assumes: every orientation-preserving isometry of $\mathbb{R}^3$ is a rotation about a line. The proof is where Chapter 9 is needed, for one reason only: to produce an eigenvector.

Theorem 11.7.7. Every Element of $\text{SO}(3)$ Is a Rotation

Let $\mathbf{A} \in \text{SO}(3)$. Then there is a unit vector $\mathbf{u} \in \mathbb{R}^3$ with $\mathbf{A}\mathbf{u} = \mathbf{u}$, and an orthonormal basis $\mathcal{B} = (\mathbf{u}, \mathbf{w}_1, \mathbf{w}_2)$ of $\mathbb{R}^3$ in which

$$[T_{\mathbf{A}}]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

for some θ . That is, $\mathbf{A}$ fixes the **axis** $\text{span}(\mathbf{u})$ and rotates the plane $\text{span}(\mathbf{u})^\perp$ by θ . Moreover $\text{tr } \mathbf{A} = 1 + 2 \cos \theta$.

Idea. Three steps. ① Get a real eigenvalue: the characteristic polynomial is real of degree 3, and an odd-degree real polynomial has a real root. ② That eigenvalue is ± 1 , since $T_{\mathbf{A}}$ is an isometry. ③ Split off the eigenline: its orthogonal complement is invariant, the restriction there lies in $\text{O}(2)$, and [Theorem 11.7.6](#) finishes. If the eigenvalue was -1 the restriction has determinant -1 , so it is a reflection and supplies a $+1$ eigenvector inside the plane; feed that back into step ③.

Proof. Write $T = T_{\mathbf{A}}$, an isometry of $\mathbb{R}^3$ by [Theorem 11.7.3](#).

Step 1: a real eigenvalue exists. By Theorem 7.10.2, $p_{\mathbf{A}} \in \mathbb{R}[x]$ is monic of degree 3, which is odd, so $p_{\mathbf{A}}$ has a real root λ by Corollary 6.5.9, and λ is an eigenvalue of $\mathbf{A}$ by Theorem 9.2.1. By [Proposition 11.7.1](#), $\lambda = 1$ or $\lambda = -1$. Fix a unit eigenvector $\mathbf{u}_0$ for λ , possible after dividing by its norm.

Step 2: the orthogonal complement of an invariant line is invariant.

Claim. If $U \subseteq \mathbb{R}^3$ is a subspace with $T(U) \subseteq U$, then $T(U^\perp) \subseteq U^\perp$.

Proof. $T|_U: U \rightarrow U$ is injective, because T is, so it is surjective by Theorem 2.5.7 applied on the finite-dimensional space U . Let $\mathbf{w} \in U^\perp$ and $\mathbf{u} \in U$. Write $\mathbf{u} = T\mathbf{u}'$ with $\mathbf{u}' \in U$. Then $\langle T\mathbf{w}, \mathbf{u} \rangle = \langle T\mathbf{w}, T\mathbf{u}' \rangle = \langle \mathbf{w}, \mathbf{u}' \rangle = 0$, the middle step by [Theorem 11.7.3](#) (b) and the

last because $\mathbf{u}' \in U$. So $T\mathbf{w} \in U^\perp$. ■

Step 3: reduce to the plane. Put $U = \text{span}(\mathbf{u}_0)$, which is T -invariant. By the claim, $P := U^\perp$ is T -invariant, and $\dim P = 2$ by Theorem 11.3.2 (c). Choose an orthonormal basis $(\mathbf{w}_1, \mathbf{w}_2)$ of P (Theorem 11.2.2); then $\mathcal{B}_0 = (\mathbf{u}_0, \mathbf{w}_1, \mathbf{w}_2)$ is an orthonormal basis of $\mathbb{R}^3$, and in it

$$[T]_{\mathcal{B}_0}^{\mathcal{B}_0} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & c_{11} & c_{12} \\ 0 & c_{21} & c_{22} \end{pmatrix}, \quad \mathbf{C} := \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix},$$

the first column because $T\mathbf{u}_0 = \lambda\mathbf{u}_0$, and the two zeros in the first row because $T\mathbf{w}_1, T\mathbf{w}_2 \in P$ have no $\mathbf{u}_0$ -component. This matrix is orthogonal, being the matrix of an isometry in orthonormal bases (Theorem 11.7.3 (f)), so its second and third columns are orthonormal, which says $\mathbf{C}^\top \mathbf{C} = \mathbf{I}_2$ and $\mathbf{C} \in \text{O}(2)$; and $\mathbf{C}$ is the matrix of $T|_P$ in the basis $(\mathbf{w}_1, \mathbf{w}_2)$. Expanding the determinant along the first column, and using that similar matrices have equal determinants (Corollary 7.5.7, with Theorem 4.2.4 (a)), $1 = \det \mathbf{A} = \lambda \det \mathbf{C}$, so $\det \mathbf{C} = \lambda$.

Case $\lambda = 1$. Then $\det \mathbf{C} = 1$, so $\mathbf{C} = \mathbf{R}_\theta$ for some θ by Theorem 11.7.6 (a), and $\mathcal{B} = \mathcal{B}_0$ and $\mathbf{u} = \mathbf{u}_0$ are as required.

Case $\lambda = -1$. Then $\det \mathbf{C} = -1$, so by Theorem 11.7.6 (b) the restriction $T|_P$ is a reflection of the plane P and has 1 as an eigenvalue. Let $\mathbf{u} \in P$ be a unit eigenvector for 1, so $T\mathbf{u} = \mathbf{u}$. Now rerun Steps 2 and 3 with $\mathbf{u}$ in place of $\mathbf{u}_0$. Because $\mathbf{u}$ was chosen with $T\mathbf{u} = \mathbf{u}$, the eigenvalue this time is 1, so the rerun lands in the previous case and cannot return here.

Finally, $\mathbf{A}$ and $[T]_{\mathcal{B}}^{\mathcal{B}}$ are matrices of the same operator, hence similar (Theorem 4.2.4 (a)), and similar matrices have equal traces (Theorem 4.2.5), so $\text{tr } \mathbf{A} = 1 + 2 \cos \theta$. This proves the theorem. ■

The angle is determined by $\mathbf{A}$ up to sign, since reversing the order of $\mathbf{w}_1, \mathbf{w}_2$ replaces θ by $-\theta$; the trace formula $\cos \theta = \frac{1}{2}(\text{tr } \mathbf{A} - 1)$ gives $|\theta|$ with no computation at all. This settles O(3) as well: if $\det \mathbf{A} = -1$ then $-\mathbf{A} \in \text{SO}(3)$, so $\mathbf{A}$ is a rotation composed with the antipodal map.

Example. Axis and Angle of a Rotation

Let $\mathbf{A} = \frac{1}{3} \begin{pmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ -2 & -1 & 2 \end{pmatrix}$. Verify that $\mathbf{A} \in \text{SO}(3)$ and find its axis and angle.

Solution. Orthogonality. The three columns are $\frac{1}{3}(2, 1, -2)$, $\frac{1}{3}(-2, 2, -1)$ and $\frac{1}{3}(1, 2, 2)$. Their squared norms are $\frac{1}{9}(4 + 1 + 4) = 1$, $\frac{1}{9}(4 + 4 + 1) = 1$ and $\frac{1}{9}(1 + 4 + 4) = 1$, and their pairwise inner products are $\frac{1}{9}(-4 + 2 + 2) = 0$, $\frac{1}{9}(2 + 2 - 4) = 0$ and $\frac{1}{9}(-2 + 4 - 2) = 0$. So the columns are orthonormal and $\mathbf{A} \in \text{O}(3)$ by Theorem 11.7.3 (f).

Determinant. Expanding along the first row, $27 \det \mathbf{A} = 2(4 + 2) - (-2)(2 + 4) + 1(-1 + 4) = 12 + 12 + 3 = 27$, so $\det \mathbf{A} = 1$ and $\mathbf{A} \in \text{SO}(3)$.

Axis. Solve $\mathbf{A}\mathbf{u} = \mathbf{u}$, that is $(3\mathbf{I} - \mathbf{A})\mathbf{u} = \mathbf{0}$:

$$\begin{pmatrix} 1 & 2 & -1 \\ -1 & 1 & -2 \\ 2 & 1 & 1 \end{pmatrix} \mathbf{u} = \mathbf{0}.$$

Adding row one to row two gives $3u_2 - 3u_3 = 0$, so $u_2 = u_3$; then row one gives $u_1 = -u_2$. The axis is $\text{span}((-1, 1, 1))$, and indeed $\mathbf{A}(-1, 1, 1) = \frac{1}{3}(-2-2+1, -1+2+2, 2-1+2) = (-1, 1, 1)$.

Angle. $\text{tr } \mathbf{A} = \frac{1}{3}(2+2+2) = 2$, so $\cos \theta = \frac{1}{2}(2-1) = \frac{1}{2}$ and $|\theta| = \pi/3$. The map is a rotation by 60° about the line through $(-1, 1, 1)$.

Quick check. A matrix $\mathbf{A} \in \text{SO}(3)$ has $\text{tr } \mathbf{A} = -1$. What is its angle, and what is $\mathbf{A}^2$?

Solution. $\cos \theta = \frac{1}{2}(-1-1) = -1$, so $\theta = \pi$: a half-turn about the axis. Then $\mathbf{A}^2$ is the rotation by 2π about the same axis, which is $\mathbf{I}_3$. (In the basis of [Theorem 11.7.7](#) the matrix is $\text{diag}(1, -1, -1)$, whose square is $\mathbf{I}_3$.)

11.7.7 Rigid motions

We assumed linearity from the start. It is worth knowing that we did not have to: for maps of $\mathbb{R}^n$ that fix the origin, preserving distance forces linearity.

Theorem 11.7.8. Distance-Preserving Maps Are Affine

Let $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be any function, not assumed linear, with

$$\|f(\mathbf{x}) - f(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\| \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

- (a) If $f(\mathbf{0}) = \mathbf{0}$, then f is linear and orthogonal.
- (b) In general, $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ for a unique $\mathbf{A} \in \text{O}(n)$ and a unique $\mathbf{b} \in \mathbb{R}^n$.

Idea. Setting $\mathbf{y} = \mathbf{0}$ turns the hypothesis into “preserves norms”, and expanding the squared distance then turns it into “preserves inner products” without any polarization at all. Once inner products are preserved, f sends the standard basis to an orthonormal basis, and [Theorem 11.2.1](#) writes $f(\mathbf{x})$ as an explicit linear expression in the coordinates of $\mathbf{x}$. Part (b) is part (a) applied to $\mathbf{x} \mapsto f(\mathbf{x}) - f(\mathbf{0})$.

Proof.

- (a) Assume $f(\mathbf{0}) = \mathbf{0}$. Taking $\mathbf{y} = \mathbf{0}$ gives $\|f(\mathbf{x})\| = \|\mathbf{x}\|$ for all $\mathbf{x}$. Expanding both sides of the hypothesis with [Theorem 11.1.2](#) (c),

$$\|f(\mathbf{x})\|^2 - 2\langle f(\mathbf{x}), f(\mathbf{y}) \rangle + \|f(\mathbf{y})\|^2 = \|\mathbf{x}\|^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2,$$

and the four outer terms cancel in pairs, leaving

$$\langle f(\mathbf{x}), f(\mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle \quad \text{for all } \mathbf{x}, \mathbf{y}. \quad (*)$$

Put $\mathbf{f}_i := f(\mathbf{e}_i)$. By (*), $\langle \mathbf{f}_i, \mathbf{f}_j \rangle = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$, so $(\mathbf{f}_1, \dots, \mathbf{f}_n)$ is an orthonormal list of n vectors in $\mathbb{R}^n$, hence an orthonormal basis (it is independent by Theorem 11.1.4 and has $n = \dim \mathbb{R}^n$ elements). Now let $\mathbf{x} = (x_1, \dots, x_n)$ be arbitrary. By Theorem 11.2.1 applied to the basis $(\mathbf{f}_1, \dots, \mathbf{f}_n)$, and then by (*),

$$f(\mathbf{x}) = \sum_{i=1}^n \langle f(\mathbf{x}), \mathbf{f}_i \rangle \mathbf{f}_i = \sum_{i=1}^n \langle f(\mathbf{x}), f(\mathbf{e}_i) \rangle \mathbf{f}_i = \sum_{i=1}^n \langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{f}_i = \sum_{i=1}^n x_i \mathbf{f}_i.$$

The right-hand side is linear in $\mathbf{x}$, so f is linear; and it preserves norms, so it is an isometry, hence orthogonal by Definition 11.7.2, since it maps $\mathbb{R}^n$ to $\mathbb{R}^n$.

- (b) Put $\mathbf{b} := f(\mathbf{0})$ and $g(\mathbf{x}) := f(\mathbf{x}) - \mathbf{b}$. Then g preserves distances, since the translation cancels, and $g(\mathbf{0}) = \mathbf{0}$; by (a), g is orthogonal, say $g(\mathbf{x}) = \mathbf{A}\mathbf{x}$ with $\mathbf{A} \in O(n)$, and $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$. For uniqueness, evaluating at $\mathbf{0}$ recovers $\mathbf{b}$, and then $\mathbf{A}\mathbf{x} = f(\mathbf{x}) - \mathbf{b}$ determines $\mathbf{A}$. This proves the theorem. ■

So the rigid motions of $\mathbb{R}^n$ are exactly the maps “rotate or reflect, then translate”, and the linear ones are exactly the orthogonal maps. That is the sense in which $O(n)$ is the symmetry group of Euclidean geometry, and it is why the same group reappears whenever a problem has a distance in it.

11.7.8 Exercises

11.7.8.1 A. Check your understanding

Exercise. A1

- Define an isometry $T \in \mathcal{L}(V, W)$, and state two conditions equivalent to it.
- What must be checked about a square matrix $\mathbf{A}$ to decide whether it is unitary? State the check in terms of its columns.
- True or false: every isometry is invertible. Justify your answer.
- True or false: if $\det \mathbf{A} = 1$ for $\mathbf{A} \in O(n)$, then $\mathbf{A}$ fixes a nonzero vector. Justify your answer.
- Give the two families of matrices making up $O(2)$, and say how the determinant distinguishes them.
- An operator T on a complex inner product space is unitary. Why must every eigenvalue of T lie on the unit circle?

Solution.

- $\|Tv\| = \|v\|$ for every $v \in V$ (Definition 11.7.1). Equivalent: $\langle Tu, Tv \rangle = \langle u, v \rangle$ for all u, v ; and $T^*T = \text{id}_V$ (Theorem 11.7.3).
- That its columns are orthonormal in F^n , that is $\mathbf{A}^*\mathbf{A} = \mathbf{I}_n$ (Theorem 11.7.3 (f)). For a square matrix this already gives $\mathbf{A}\mathbf{A}^* = \mathbf{I}_n$.

- (c) False. $T: \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $T(x_1, x_2) = (x_1, x_2, 0)$, is an isometry and is not surjective ([Example \(Isometries and Near-Isometries\)](#) (a)). An isometry is always injective, and it is invertible exactly when $\dim W = \dim V$.
- (d) False for even n : $-\mathbf{I}_4 \in \text{SO}(4)$ has determinant 1 and fixes only $\mathbf{0}$. It is true for odd n , though not by the argument of [Theorem 11.7.7](#), which uses the $\text{O}(2)$ classification of a 2-dimensional complement. The general reason is a count: the nonreal eigenvalues of a real $\mathbf{A}$ come in conjugate pairs $\lambda, \bar{\lambda}$ with $\lambda\bar{\lambda} = |\lambda|^2 = 1$ for an isometry, so they contribute 1 to the determinant and an even number to the count. With n odd, an odd number of eigenvalues are real, each ± 1 , and their product is $\det \mathbf{A} = 1$; an odd number of factors ± 1 with product 1 cannot all be -1 , so 1 is an eigenvalue and some nonzero vector is fixed.
- (e) The rotations $\mathbf{R}_\theta$, with determinant 1, and the reflections $\mathbf{M}_\theta$, with determinant -1 ([Theorem 11.7.6](#)).
- (f) A unitary operator is an isometry, so [Proposition 11.7.1](#) applies: $|\lambda| = 1$.

11.7.8.2 B. Practice

Exercise. B1: Determine which are unitary

Determine which of the following matrices are orthogonal or unitary. Justify your answer; for those that are not, name the failing condition.

- (a) $\frac{1}{3} \begin{pmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ -2 & -1 & 2 \end{pmatrix}$ over $\mathbb{R}$.
- (b) $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$ over $\mathbb{C}$.
- (c) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ over $\mathbb{R}$.
- (d) $\frac{1}{5} \begin{pmatrix} 3 & 4 \\ 4 & 3 \end{pmatrix}$ over $\mathbb{R}$.

Solution.

- (a) Orthogonal, as computed in [Example \(Axis and Angle of a Rotation\)](#): the three columns are orthonormal.
- (b) Unitary. The columns are $\frac{1}{\sqrt{2}}(1, i)$ and $\frac{1}{\sqrt{2}}(i, 1)$, of norm $\frac{1}{\sqrt{2}}\sqrt{1+1} = 1$, and their inner product is

$$\frac{1}{2}(1 \cdot \bar{i} + i \cdot \bar{1}) = \frac{1}{2}(-i + i) = 0.$$

So $\mathbf{A}^* \mathbf{A} = \mathbf{I}_2$. (It is unitary but not orthogonal, since it is not real.)

- (c) Not orthogonal. The second column $(1, 1)$ has norm $\sqrt{2} \neq 1$, so the columns are not orthonormal. Equivalently, the map stretches $\mathbf{e}_2$.
- (d) Not orthogonal. Both columns are unit vectors, $\frac{1}{5}(3, 4)$ and $\frac{1}{5}(4, 3)$, but their inner product is $\frac{1}{25}(12 + 12) = \frac{24}{25} \neq 0$. Unit columns are not enough; they must also be orthogonal. (Its determinant is $\frac{1}{25}(9 - 16) = -\frac{7}{25}$, neither 1 nor -1 , which confirms it by [Proposition 11.7.4](#).)

Exercise. B2: Name the plane isometry

Each of the following lies in $O(2)$. Decide whether it is a rotation or a reflection, and give the angle of rotation or the mirror line.

- (a) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
- (b) $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$.
- (c) $\frac{1}{13} \begin{pmatrix} 5 & 12 \\ 12 & -5 \end{pmatrix}$.

Solution.

- (a) Determinant -1 , so a reflection. $\mathbf{Ax} = \mathbf{x}$ reads $x_2 = x_1$, so the mirror is $L = \text{span}((1, 1))$, the line $x_2 = x_1$. Check: $\mathbf{A}(1, -1) = (-1, 1) = -(1, -1)$.
- (b) Determinant $\frac{1}{2}(1 + 1) = 1$, so a rotation $\mathbf{R}_\theta$. Comparing with $\mathbf{R}_\theta$, $\cos \theta = \frac{1}{\sqrt{2}}$ and $\sin \theta = -\frac{1}{\sqrt{2}}$, so $\theta = -\pi/4$: a clockwise half of a right angle.
- (c) Determinant $\frac{1}{169}(-25 - 144) = -1$, so a reflection. $\mathbf{Ax} = \mathbf{x}$ reads $5x_1 + 12x_2 = 13x_1$, that is $3x_2 = 2x_1$, so $L = \text{span}((3, 2))$. Check: $\mathbf{A}(3, 2) = \frac{1}{13}(15 + 24, 36 - 10) = (3, 2)$.

Exercise. B3: Permutations are orthogonal

Let $\sigma \in S_n$ and let $\mathbf{P}_\sigma \in M_n(\mathbb{R})$ be its permutation matrix (Definition 3.4.2), whose j -th column is $\mathbf{e}_{\sigma(j)}$.

- (a) Prove that $\mathbf{P}_\sigma \in O(n)$.
- (b) Deduce that $\det \mathbf{P}_\sigma = \pm 1$, and identify the sign.
- (c) Show that $\mathbf{P}_\sigma^{-1} = \mathbf{P}_\sigma^\top = \mathbf{P}_{\sigma^{-1}}$.

Solution.

- (a) Column j of $\mathbf{P}_\sigma$ is $\mathbf{e}_{\sigma(j)}$, by Definition 3.4.2. So each column is a unit vector, and for $i \neq j$ the columns $\mathbf{e}_{\sigma(i)}$ and $\mathbf{e}_{\sigma(j)}$ are distinct standard basis vectors, because σ is injective, hence orthogonal. By Theorem 11.7.3 (f), $\mathbf{P}_\sigma \in O(n)$.
- (b) By Proposition 11.7.4 (b), $\det \mathbf{P}_\sigma = \pm 1$. The sign is $\text{sgn}(\sigma)$: permuting the columns of $\mathbf{I}_n$ by σ multiplies the determinant by $\text{sgn}(\sigma)$ (Theorem 7.2.1), and $\det \mathbf{I}_n = 1$.
- (c) $\mathbf{P}_\sigma^{-1} = \mathbf{P}_\sigma^\top$ holds because $\mathbf{P}_\sigma$ is orthogonal, by (a). For the second equality, $(\mathbf{P}_\sigma)_{ij} = 1$ exactly when $i = \sigma(j)$, so $(\mathbf{P}_\sigma^\top)_{ij} = (\mathbf{P}_\sigma)_{ji} = 1$ exactly when $j = \sigma(i)$, that is when $i = \sigma^{-1}(j)$; and that is the defining condition for $(\mathbf{P}_{\sigma^{-1}})_{ij} = 1$.

11.7.8.3 C. Going deeper**Exercise. C1: Determinant as a homomorphism**

- (a) Prove that $\det: O(n) \rightarrow \{1, -1\}$ is a surjective group homomorphism with kernel $SO(n)$, and deduce that exactly half of $O(n)$ has determinant -1 , in the sense that $O(n)$ is the disjoint union of $SO(n)$ and one coset $\mathbf{A}_0 SO(n)$.
- (b) Prove that $\det: U(n) \rightarrow \{z \in \mathbb{C} : |z| = 1\}$ is a surjective group homomorphism.
- (c) Describe $O(1)$ and $U(1)$ completely.

Solution.

- (a) $\det$ is multiplicative (Theorem 7.5.4) and takes only the values ± 1 on $O(n)$ (Proposition 11.7.4 (b)), so it is a homomorphism into $\{1, -1\}$. It is surjective: $\det \mathbf{I}_n = 1$, and $\mathbf{A}_0 := \text{diag}(-1, 1, \dots, 1)$ is orthogonal, since its columns are $\pm \mathbf{e}_i$, with $\det \mathbf{A}_0 = -1$. Its kernel is $\{\mathbf{A} : \det \mathbf{A} = 1\} = SO(n)$ by definition. If $\det \mathbf{A} = -1$ then $\det(\mathbf{A}_0^{-1} \mathbf{A}) = 1$, so $\mathbf{A} \in \mathbf{A}_0 SO(n)$; and no matrix has both determinants, so the union is disjoint.
- (b) Multiplicativity again, and $|\det \mathbf{A}| = 1$ on $U(n)$ by Proposition 11.7.4 (b). For surjectivity, given z with $|z| = 1$, the matrix $\text{diag}(z, 1, \dots, 1)$ has orthonormal columns, hence lies in $U(n)$, and its determinant is z (Theorem 7.4.6).
- (c) A 1×1 matrix (a) satisfies $\bar{a}a = 1$ exactly when $|a| = 1$. So $U(1)$ is the unit circle of $\mathbb{C}$ and $O(1) = \{(1), (-1)\}$. Consistently, $SO(1) = \{(1)\}$ is trivial.

Exercise. C2: Linearity is a hypothesis, not a consequence

- (a) Let $c: \mathbb{C} \rightarrow \mathbb{C}$ be $c(z) = \bar{z}$, where $\mathbb{C}$ carries $\langle z, w \rangle = z\bar{w}$. Show that $\|c(z)\| = \|z\|$ for every z , that c is not linear over $\mathbb{C}$, and that c does not preserve the inner product. Which step of Theorem 11.7.3 fails?
- (b) Let V be a **complex** inner product space and let $T \in \mathcal{L}(V)$ satisfy $\langle Tv, v \rangle = 0$ for every $v \in V$. Prove that $T = 0$.
- (c) Show that (b) fails over $\mathbb{R}$ by exhibiting a nonzero $T \in \mathcal{L}(\mathbb{R}^2)$ with $\langle T\mathbf{v}, \mathbf{v} \rangle = 0$ for all $\mathbf{v}$.

Hint for (b): apply the hypothesis to $v = u + w$ and to $v = u + iw$.

Solution.

- (a) $\|c(z)\| = |\bar{z}| = |z| = \|z\|$. But $c(iz) = \overline{iz} = -i\bar{z} = -ic(z) \neq ic(z)$ unless $z = 0$, so c is not $\mathbb{C}$ -linear. And $\langle c(1), c(i) \rangle = \langle 1, -i \rangle = 1 \cdot \overline{-i} = i$, while $\langle 1, i \rangle = -i$. The proof of (a) $\Rightarrow$ (b) in Theorem 11.7.3 uses $T(u + i^k v) = Tu + i^k Tv$, which is exactly what c does not satisfy.
- (b) Let $u, w \in V$. Applying the hypothesis to $u + w$ and expanding both slots,

$$\begin{aligned} 0 &= \langle T(u + w), u + w \rangle = \langle Tu, u \rangle + \langle Tu, w \rangle + \langle Tw, u \rangle + \langle Tw, w \rangle \\ &= \langle Tu, w \rangle + \langle Tw, u \rangle. \end{aligned} \tag{1}$$

Applying it to $u + iw$, and using $\langle Tu, iw \rangle = -i\langle Tu, w \rangle$, $\langle T(iw), u \rangle = i\langle Tw, u \rangle$ and $\langle T(iw), iw \rangle = i\bar{i}\langle Tw, w \rangle = 0$,

$$0 = -i\langle Tu, w \rangle + i\langle Tw, u \rangle.$$

Multiplying by i gives

$$\langle Tu, w \rangle - \langle Tw, u \rangle = 0. \quad (2)$$

Adding (1) and (2), $2\langle Tu, w \rangle = 0$, so $\langle Tu, w \rangle = 0$ for all u, w . Taking $w = Tu$ gives $\|Tu\|^2 = 0$, so $Tu = 0$ for every u , that is $T = 0$.

- (c) Let T be the rotation $\mathbf{R}_{\pi/2}$, that is $T(x_1, x_2) = (-x_2, x_1)$. Then $\langle T\mathbf{x}, \mathbf{x} \rangle = -x_2x_1 + x_1x_2 = 0$ for every $\mathbf{x}$, and $T \neq 0$. The complex proof breaks down because the second equation is unavailable: there is no i to substitute.

Exercise. C3: Two reflections make a rotation

Let $\alpha, \beta \in \mathbb{R}$ and let $\mathbf{M}_\alpha, \mathbf{M}_\beta \in \text{O}(2)$ be the reflections of [Theorem 11.7.6](#), in the lines L_α and L_β at angles $\alpha/2$ and $\beta/2$.

- (a) Prove that $\mathbf{M}_\alpha \mathbf{M}_\beta = \mathbf{R}_{\alpha-\beta}$.
- (b) Deduce that the composite of the reflections in two lines through the origin is the rotation by **twice** the angle from the second line to the first, and that every rotation of the plane is a product of two reflections.
- (c) Deduce that $\text{O}(2)$ is not abelian, by computing $\mathbf{M}_\beta \mathbf{M}_\alpha$.

Solution.

- (a) Multiplying out,

$$\mathbf{M}_\alpha \mathbf{M}_\beta = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix} \begin{pmatrix} \cos \beta & \sin \beta \\ \sin \beta & -\cos \beta \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix},$$

where $c = \cos \alpha \cos \beta + \sin \alpha \sin \beta = \cos(\alpha - \beta)$ and $s = \sin \alpha \cos \beta - \cos \alpha \sin \beta = \sin(\alpha - \beta)$, by the subtraction formulas. The top-right entry is $\cos \alpha \sin \beta - \sin \alpha \cos \beta = -\sin(\alpha - \beta)$, as displayed. So $\mathbf{M}_\alpha \mathbf{M}_\beta = \mathbf{R}_{\alpha-\beta}$.

- (b) The angle from L_β to L_α is $\frac{\alpha}{2} - \frac{\beta}{2}$, and (a) says the composite is the rotation by $\alpha - \beta$, twice that. Conversely, given θ , choose $\alpha = \theta$ and $\beta = 0$; then $\mathbf{R}_\theta = \mathbf{M}_\theta \mathbf{M}_0$, a product of two reflections.
- (c) By (a) with the roles exchanged, $\mathbf{M}_\beta \mathbf{M}_\alpha = \mathbf{R}_{\beta-\alpha} = \mathbf{R}_{\alpha-\beta}^{-1}$. These two agree only when $\mathbf{R}_{\alpha-\beta}^2 = \mathbf{I}$, that is when $\alpha - \beta$ is a multiple of π . Taking $\alpha = \pi/2$ and $\beta = 0$ gives $\mathbf{M}_{\pi/2} \mathbf{M}_0 = \mathbf{R}_{\pi/2} \neq \mathbf{R}_{-\pi/2} = \mathbf{M}_0 \mathbf{M}_{\pi/2}$, so $\text{O}(2)$ is not abelian.