This discussion is about the space $\mathbb{R}^d$ (vectors and matrices have real entries), but most readily generalizes to finite dimensional vector spaces (and matrices) over arbitrary fields F.

1. A function $T : \mathbb{R}^d \to \mathbb{R}^d$ is **linear** if it satisfies the following two conditions: T is *additive* and T is *homogeneous.*

The first condition is that T is a homomorphism from $\mathbb{R}^d$ to $\mathbb{R}^d$ (here the word homomorphism only expresses that the structure of $\mathbb{R}^d$ *as an additive group* is preserved):

$$(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^d)\, T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}) \quad \text{(additivity).}$$

The second condition is

$$(\forall \mathbf{x} \in \mathbb{R}^d)\, (\forall c \in \mathbb{R}) T(c\mathbf{x}) = cT(\mathbf{x}) \quad \text{(homogeneity).}$$

Sometimes we combine these two conditions into one:

$$(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^d)(\forall c_1, c_2 \in \mathbb{R}) T(c_1\mathbf{x} + c_2\mathbf{y}) = c_1 T(\mathbf{x}) + c_2 T(\mathbf{y})$$

— the function T *preserves linear combinations.*

This combined property may also be phrased as "T is a homomorphism," but this time the word homomorphism expresses the stronger demand that the structure of $\mathbb{R}^d$ *as a vector space* be preserved (both vector addition and scalar multiplication are preserved). A *linear function* on $\mathbb{R}^d$ is also called a *linear transformation* on $\mathbb{R}^d$ or a *linear operator* on $\mathbb{R}^d$.

2. The unit vectors $\mathbf{e}^1 = (1, 0, 0, \ldots, 0)$, $\mathbf{e}^2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $\mathbf{e}^d = (0, 0, 0, \ldots, 1)$ form a basis in $\mathbb{R}^d$: every vector can be expressed in a unique way as a linear combination of the unit vectors — indeed, $(x_1, x_2, \ldots, x_d) = x_1\mathbf{e}^1 + x_2\mathbf{e}^2 + \ldots + x_d\mathbf{e}^d$.

Hence, if a linear operator (on $\mathbb{R}^d$) is zero on all unit vectors, then it's zero everywhere. Consequently, if two linear operators agree on all unit vectors, then they agree everywhere.

3. Given a linear operator T on $\mathbb{R}^d$, we construct a $d \times d$ matrix A whose jth column contains $T(\mathbf{e}^j)$ $(j = 1, 2, \ldots, d)$. (When an ordered basis $B = (\mathbf{b}_1, \ldots, \mathbf{b}_d)$ other than the standard basis is used, one should phrase it as "the jth column contains the coordinates of the image $T(\mathbf{b}_j)$ of the jth basis vector – expressed in B itself.) Thus, by the rules of matrix multiplication, $T(\mathbf{e}^j) = A\mathbf{e}^j$, $(j = 1, 2, \ldots, d)$, and hence $T(\mathbf{x}) = A\mathbf{x}$ for all vectors $\mathbf{x} \in \mathbb{R}^d$. This matrix A is the "coordinatized form" of the operator T (in the standard basis).

Examples

$$I\text{: identity on } \mathbb{R}^2 \quad \leftrightarrow \quad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \leftrightarrow \quad I(x, y) = (x, y)$$

$$T_1\text{: rotation of } \mathbb{R}^2 \text{ around } \mathbf{0} \text{ by } \pi/2 \quad \leftrightarrow \quad A_1 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \leftrightarrow \quad T_1(x, y) = (-y, x)$$

$$T_2\text{: reflection about } y = x \quad \leftrightarrow \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \leftrightarrow \quad T_2(x, y) = (y, x)$$

$$T_3\text{: orthogonal projection onto the } x\text{-axis} \quad \leftrightarrow \quad A_3 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \leftrightarrow \quad T_3(x, y) = (x, 0)$$

The fourth operator is not an isometry, but the first three are; they satisfy the orthogonality condition (see next page): $A^{-1} = A^t$, that is, $AA^t = A^tA = I$. The first two are direct isometries (determinant $= 1$), and the third one is an inverse isometry (determinant $= -1$).

Orthogonal matrices

Definition. *A $d \times d$ matrix A is* **orthogonal** *if the length of every column vector is 1, and any two column vectors have dot product 0.*

A more elegant form: *The columns of A form an orthonormal basis of the vector space $\mathbb{R}^d$.* Writing A^t for the transpose of A, the orthogonality condition can be written as $A^tA = I$.

Theorem 1. *A matrix A is orthogonal if and only if $A^{-1} = A^t$, that is, the three conditions $A^tA = I$, $AA^t = I$, and $A^{-1} = A^t$ are equivalent.*
The determinant of an orthogonal matrix is 1 or -1.

Proof. It is enough to show that A is invertible, since then multiplying by A^{-1} (from left and from right) leads to the equivalence. Now, invertibility of an orthogonal matrix A (as well as $\det(A) = \pm 1$) follows from the following simple lemmas of linear algebra.

Lemma. *The identity I satisfies $\det(I) = 1$.*
For any square matrix A, $\det(A^t) = \det(A)$.

Lemma. *Determinant is multiplicative: For any two $d \times d$ matrices A and B,*

$$\det(A \cdot B) = \det(A) \cdot \det(B).$$

In other words, determinant is a homomorphism from the set of all $d \times d$ matrices to the set of real numbers (both considered only as multiplicative structures).

Lemma. *A square matrix is invertible if and only if its determinant is not 0.*

Theorem 2. *The product of two orthogonal matrices is orthogonal.*

Proof.

$$(AB)^{-1} = B^{-1}A^{-1} = B^tA^t = (AB)^t.$$

Remark. Note the important rule we used: When either inverting or transposing a product, one **must** reverse order!

Isometries and linear algebra

We will use the (somewhat ridiculous) abbreviation **ifo** for an **i**_sometry_ **f**_ixing the_ **o**_rigin._

Short summary

The Following Are Equivalent

- f is an **ifo**.
- f preserves dot products.
- f is linear with an orthogonal matrix.

Detailed statements

Theorem 3. _A function_ $f : \mathbb{R}^d \to \mathbb{R}^d$ _is an_ **ifo** _if and only if it preserves the dot product (and hence the lengths) of vectors:_

$$(\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^d)\; f(\mathbf{x}) \cdot f(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$$

Theorem 4. _An_ **ifo** _is linear._

Theorem 5. _If_ f _is linear with matrix_ A_, then_ f _is an_ **ifo** _if and only if_ A _is orthogonal_ $(A^{-1} = A^t)$_._

Note that so far everything was about arbitrary dimensions and isometries (direct or inverse).

Theorem 6. _Let_ $f : \mathbb{R}^3 \to \mathbb{R}^3$ _be an_ **ifo**. _Then,_ f _is a direct isometry (the determinant of the matrix of_ f _is 1) if and only if_ f _is a rotation._

Definition. _The set of all_ $d \times d$ _orthogonal matrices is denoted by_ $O(d)$_. It is a group with respect to matrix multiplication, called the_ **orthogonal group.**

The set of all matrices in $O(d)$ _with determinant 1 clearly form a subgroup. It is denoted by_ $SO(d)$ _and is called the_ **special orthogonal group.**

Because of the equivalences above, we often use $O(d)$ $[SO(d)]$ _to also denote the set of all_ [_direct_] _isometries of_ $\mathbb{R}^d$ _fixing the origin._

Thus, Theorem 6 can be restated as follows:

Theorem 7. _Let_ A _be a_ 3×3 _real matrix. Then,_ $A \in SO(3)$ _if and only if the function_ $f : \mathbb{R}^3 \to \mathbb{R}^3 : \mathbf{x} \mapsto A\mathbf{x}$ _is a rotation (about some axis through the origin)._

Using Theorem 2, we obtain the following non-trivial corollary.

Corollary 8. _The product of two rotations fixing the origin is again a rotation._

("Product," of course, means composition here.)

Proofs

To simplify notation, we will write a' for $f(a)$, x' for $f(x)$, etc.

Proof of Theorem 3. Let f be an **ifo** on $\mathbb{R}^d$. Since the origin is fixed by f, the length of every vector is also preserved (since length is distance from the origin).
Now let $x, y \in \mathbb{R}^d$, and write $x' = f(x)$ and $y' = f(y)$. Since

$$\|\mathbf{x}-\mathbf{y}\|^2 = (\mathbf{x}-\mathbf{y})\cdot(\mathbf{x}-\mathbf{y}) = \mathbf{x}\cdot\mathbf{x}+\mathbf{y}\cdot\mathbf{y}-2\mathbf{x}\cdot\mathbf{y} = \|\mathbf{x}\|^2+\|\mathbf{y}\|^2-2\mathbf{x}\cdot\mathbf{y},$$
$$\|\mathbf{x}'-\mathbf{y}'\|^2 = (\mathbf{x}'-\mathbf{y}')\cdot(\mathbf{x}'-\mathbf{y}') = \mathbf{x}'\cdot\mathbf{x}'+\mathbf{y}'\cdot\mathbf{y}'-2\mathbf{x}'\cdot\mathbf{y}' = \|\mathbf{x}'\|^2+\|\mathbf{y}'\|^2-2\mathbf{x}'\cdot\mathbf{y}',$$

and $\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}'-\mathbf{y}'\|^2$, $\|\mathbf{x}\|^2 = \|\mathbf{x}'\|^2$, $\|\mathbf{y}\|^2 = \|\mathbf{y}'\|^2$, so $\mathbf{x}\cdot\mathbf{y} = \mathbf{x}'\cdot\mathbf{y}'$.

Conversely, if f preserves dot product, then, in particular, it preserves length, and thus it fixes the origin (the only vector of zero length). Also, for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$,

$$d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|^2 = \mathbf{x}\cdot\mathbf{x}+\mathbf{y}\cdot\mathbf{y}-2\mathbf{x}\cdot\mathbf{y},$$

so distance is also preserved. Thus, f is an **ifo**. $\square$

Proof of Theorem 4. (We'll use the same 'prime' notations as in the previous proof.)
First, let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ and let $\mathbf{z} = \mathbf{x}+\mathbf{y}$. We need to show that $\mathbf{z}' = \mathbf{x}'+\mathbf{y}'$. Now,

$$0 = [\mathbf{z}-(\mathbf{x}+\mathbf{y})]\cdot[\mathbf{z}-(\mathbf{x}+\mathbf{y})] = \mathbf{x}\cdot\mathbf{x}+\mathbf{y}\cdot\mathbf{y}+\mathbf{z}\cdot\mathbf{z}+2\mathbf{x}\cdot\mathbf{y}-2\mathbf{x}\cdot\mathbf{z}-2\mathbf{y}\cdot\mathbf{z}.$$

Since f preserves dot products (Theorem 3), the right-hand side in the last equation equals

$$= \mathbf{x}'\cdot\mathbf{x}'+\mathbf{y}'\cdot\mathbf{y}'+\mathbf{z}'\cdot\mathbf{z}'+2\mathbf{x}'\cdot\mathbf{y}'-2\mathbf{x}'\cdot\mathbf{z}'-2\mathbf{y}'\cdot\mathbf{z}' = [\mathbf{z}'-(\mathbf{x}'+\mathbf{y}')]\cdot[\mathbf{z}'-(\mathbf{x}'+\mathbf{y}')].$$

Thus, $[\mathbf{z}'-(\mathbf{x}'+\mathbf{y}')]\cdot[\mathbf{z}'-(\mathbf{x}'+\mathbf{y}')] = 0$, that is, $\mathbf{z}' = (\mathbf{x}'+\mathbf{y}')$.

Similarly, let $\mathbf{x} \in \mathbb{R}^d$ and $c \in \mathbb{R}$. Write $\mathbf{z} = c\mathbf{x}$. We need to show that $\mathbf{z}' = c\mathbf{x}'$. Now,

$$0 = [\mathbf{z}-c\mathbf{x}]\cdot[\mathbf{z}-c\mathbf{x}] = c^2\mathbf{x}\cdot\mathbf{x}-2c\mathbf{x}\cdot\mathbf{z}+\mathbf{z}\cdot\mathbf{z} = c^2\mathbf{x}'\cdot\mathbf{x}'-2c\mathbf{x}'\cdot\mathbf{z}'+\mathbf{z}'\cdot\mathbf{z}' = [\mathbf{z}'-c\mathbf{x}']\cdot[\mathbf{z}'-c\mathbf{x}'] = 0$$

proving $\mathbf{z}' = c\mathbf{x}'$. $\square$

Proof of Theorem 5.

Part I. Assume f is an **ifo**. Then f preserves length and dot product (by Theorem 3). Thus, in particular, f preserves the unit length of the standard basis vectors $\mathbf{e}_i$, $i = 1, \ldots, d$, so the diagonal elements in the matrix A^tA are all 1. Similarly, for $i \neq j$, the zero dot product $\mathbf{e}_i \cdot \mathbf{e}_j$ is preserved, so the rest of the matrix A^tA is all zero. Thus, $A^tA = I$.

Part II. Assume A is orthogonal. We will show that f preserves dot product (and hence it's an **ifo**). Indeed, let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. Using the matrix-multiplication form $\mathbf{x}\cdot\mathbf{y} = \mathbf{x}^t\mathbf{y}$ (we indicate matrix multiplication by dropping the dot product symbol $\cdot$), we get

$$f(\mathbf{x})\cdot f(\mathbf{y}) = A\mathbf{x}\cdot A\mathbf{y} = (A\mathbf{x})^tA\mathbf{y} = (\mathbf{x}^tA^t)A\mathbf{y} = \mathbf{x}^t(A^tA)\mathbf{y} = x^tI\mathbf{y} = \mathbf{x}^t\mathbf{y} = \mathbf{x}\cdot\mathbf{y} \quad \square$$

Remark.

Elements of $SO(d)$, that is, orthogonal $d \times d$ matrices with determinant 1, are called "rotations" in d-space. It is easy to see that for $d = 2$ they indeed correspond to rotations about the origin. We show now that for $d = 3$ they also correspond to rotations in the geometric sense.

Let A be a 3×3 orthogonal matrix with determinant 1. We claim that it correspond to a linear operator T which is a rotation, that is, T fixes the points of a line through the origin, and rotates the rest of the 3-space about that line.

Indeed,

$$\det(A-I) = \det(A-I)\det(A^t) = \det[(A-I)A^t] = \det(I-A^t) = \det(I-A) = (-1)^3 \det(A-I).$$

Hence, $\det(A - I) = 0$, that is, 1 is an eigenvalue of A.

Thus, A indeed fixes the points of a line ℓ through the origin. If we choose ℓ to be the third axis of a new Cartesian coordinate system, the new matrix B of T will contain the vector $(0, 0, 1)$ as its last column, and since B is also orthogonal (Why?), its last row is also $(0, 0, 1)$. The remaining 2×2 submatrix is in $SO(2)$ (Why?) and hence a rotation.