Orthogonal and Unitary transformatons; Rigid Motions

☐ Definition: A matrix U such that $U^* = U^{-1}$ is said to be *unitary* if its entries are complex, and *orthogonal* if they are real.

PROPOSITION:

1. $U^*U = I \Rightarrow UU^* = I$.
2. The rows of a unitary/orthogonal matrix written out in the standard basis form an orthonormal basis for $\mathbb{C}^n(\mathbb{R}^n)$. So do the columns.
3. If H is self-adjoint, then $\exp(iH)$ is unitary.

PROPOSITION: If U is unitary, then for all $\mathbf{x}, \mathbf{y} \in V$, $||U\mathbf{x} - U\mathbf{y})|| = ||\mathbf{x} - \mathbf{y}||$. That is, the linear transformation $\mathbf{x} \to U\mathbf{x}$ preserves lengths.

PROOF: We compute

$$\begin{aligned} ||U\mathbf{x} - U\mathbf{y})||^2 &= < U\mathbf{x} - U\mathbf{y}, U\mathbf{x} - U\mathbf{y} > \\ &= < U(\mathbf{x} - \mathbf{y}), U(\mathbf{x} - \mathbf{y}) > \\ &= < \mathbf{x} - \mathbf{y}, U^*U(\mathbf{x} - \mathbf{y}) > \\ &= < \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} > = ||\mathbf{x} - \mathbf{y}||^2 \end{aligned}$$

A unitary transformation also preserves the scalar product between two vectors:

$$\begin{aligned} < U\mathbf{x}, U\mathbf{y} > &= < \mathbf{x}, U^*U\mathbf{y} > \\ &= < \mathbf{x}, \mathbf{y} > \end{aligned}$$

So if U is orthogonal, it preserves angles as well as distances, since

$$\frac{< U\mathbf{x}, U\mathbf{y} >}{||U\mathbf{x}||\ ||U\mathbf{y}||} = \frac{< \mathbf{x}, \mathbf{y} >}{||\mathbf{x}||\ ||\mathbf{y}||} = \cos(\theta),$$

where θ is the angle between the two vectors.

☐ Definition: Let $F : V \to V$ be a function with the property $||F(\mathbf{x}) - F(\mathbf{y})|| = ||\mathbf{x} - \mathbf{y}||$, $\forall \mathbf{x}, \mathbf{y} \in V$. Then F is said to be a *rigid motion* of V.

We already know that unitary transformations are rigid motions. They are also linear transformations.

☐ Definition: A transformation of the form $F_{\mathbf{a}}(\mathbf{x}) = \mathbf{x} - \mathbf{a}$, where $\mathbf{a} \neq \mathbf{0}$ is called a *translation by* $\mathbf{a}$.

Translations are not linear transformations, since $T_{\mathbf{a}}(\mathbf{0}) \neq \mathbf{0}$. But translations are rigid motions: $||T_{\mathbf{a}}(\mathbf{x}) - T_{\mathbf{a}}(\mathbf{y})|| = ||\mathbf{x} - \mathbf{a} - (\mathbf{y} - \mathbf{a})|| = ||\mathbf{x} - \mathbf{y}||$.

THEOREM 1: Let V be an n-dimensional i.p.s. over $\mathbb{R}$. Then there is a distance-preserving isomorphism $\phi : V \to \mathbb{R}^n$, where the scalar product in $\mathbb{R}^n$ is given by the dot product.

PROOF: Choose an orthonormal basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ for V. Then for $\mathbf{x}, \mathbf{y} \in V$, we have

$$\begin{aligned} < \mathbf{x}, \mathbf{y} > &= < \sum_{i=1}^{n} < \mathbf{x}, \mathbf{e}_i > \mathbf{e}_i, \sum_{j=1}^{n} < \mathbf{y}, \mathbf{e}_j > \mathbf{e}_j > \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j < \mathbf{e}_i, \mathbf{e}_j > \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \delta_{ij} \\ &= \sum_{i=1}^{n} x_i y_i, \end{aligned}$$

where x_i, y_i are the coordinates of $\mathbf{x}, \mathbf{y}$ in the given orthonormal basis. Now define

$$\phi(\mathbf{x}) = \sum_{i=1}^{n} x_i E_i \in \mathbb{R}^n,$$

where $\{E_i : 1 \leq i \leq n\}$ is the standard orthonormal basis in $\mathbb{R}^n$. Then clearly (a) ϕ is linear, and (b) $\phi(\mathbf{x}) \bullet \phi(\mathbf{y}) = \sum_{i=1}^{n} x_i y_i = < \mathbf{x}, \mathbf{y} >$. It follows from this that $||\phi(\mathbf{x}) - \phi(\mathbf{y})|| = ||\mathbf{s} - \mathbf{y}||$ (why?) so the proof is complete ■

□ **Definition:** A vector space isomorphism which preserves the distances between vectors is called an *isometry*.

Note:

- The isometry is far from unique. We can use the same construction with any o.n. basis in V and any o.n. basis in $\mathbb{R}^n$.
- As a consequence of the theorem, any time we want to demonstrate something about a finite-dimensional real inner product space, it suffices to demonstrate it on $\mathbb{E}^n$, since the result can be transferred to V via the isometry ϕ^{-1}. So we don't give up any generality by working in $\mathbb{E}^n$.
- The theorem as stated is also true for finite-dimensional complex inner product spaces. There are isometries between them and $\mathbb{C}^n$ with the standard scalar product $< \mathbf{z}, \mathbf{w} >= \mathbf{w}^* \mathbf{z}$.

THEOREM: Suppose that G is a linear transformation on $\mathbb{E}^n$ which preserves the scalar product. Then G is orthogonal.

PROOF: We're given that $G\mathbf{x}\bullet G\mathbf{y} = \mathbf{x}\bullet\mathbf{y}$, $\forall \mathbf{x}, \mathbf{y}$. Then $\mathbf{x}\bullet G^*Gy = \mathbf{x}\bullet\mathbf{y}$. Moving everything to the left hand side, we get

$$\mathbf{x}\bullet G^*G\mathbf{y} - \mathbf{x}\bullet\mathbf{y} = \mathbf{x}\bullet\big(G^*G\mathbf{y} - \mathbf{y}\big) = \mathbf{0}.$$

For any given $\mathbf{y}$, this equality holds for all $\mathbf{x} \in \mathbb{E}^n$. This means that for any given $\mathbf{y}$, we must have $G^*G\mathbf{y} - \mathbf{y} = \mathbf{0}$, or $G^*G\mathbf{y} = \mathbf{y}$ for any $\mathbf{y}$. And this means that $G{*}^G = I$, so G is orthogonal ■

THEOREM: Any rigid motion F in $\mathbb{E}^n$ may be written in the form $F(\mathbf{x}) = (T_{\mathbf{a}} \circ U)(\mathbf{x})$, where U is orthogonal and $T_{\mathbf{a}}$ is a translation.

PROOF: Given F, we define G by $G(\mathbf{x}) = F(\mathbf{x}) - F(\mathbf{0})$. Then since $||G(\mathbf{x}) - G(\mathbf{y})|| = ||F(\mathbf{x}) - F(\mathbf{y})|| = ||\mathbf{x} - \mathbf{y}||$, G is a rigid motion that satisfies $G(\mathbf{0}) = \mathbf{0}$. We'll first show that G preserves the scalar product, and then that G is linear (which by the theorem above means that G is orthogonal:

(a) Since G is a rigid motion, $\big(G(\mathbf{x}) - G(\mathbf{y})\big)\bullet\big(G(\mathbf{x}) - G(\mathbf{y})\big) = (\mathbf{x} - \mathbf{y})\bullet(\mathbf{x} - \mathbf{y})$. On the left hand side of this, we have

$$\begin{aligned}\big(G(\mathbf{x}) - G(\mathbf{y})\big)\bullet\big(G(\mathbf{x}) - G(\mathbf{y})\big) &= ||G(\mathbf{x})||^2 + ||G(\mathbf{y})||^2 - 2G(\mathbf{x})\bullet G(\mathbf{y}) \\ &= ||\mathbf{x}||^2 + ||\mathbf{y}||^2 - 2G(\mathbf{x})\bullet G(\mathbf{y}) \text{ (because } G \text{ is a rigid motion)}\end{aligned}$$

On the right hand side,

$$||\mathbf{x} - \mathbf{y}||^2 = ||\mathbf{x}||^2 + ||\mathbf{y}||^2 - 2\mathbf{x}\bullet\mathbf{y},$$

from which we conclude that $G(\mathbf{x})\bullet G(\mathbf{y}) = \mathbf{x}\bullet\mathbf{y}$, so G preserves the scalar product.

(b) We now show that G is linear:

$$\begin{aligned}||G(\mathbf{x} + c\mathbf{y}) - G(\mathbf{x}) - cG(\mathbf{y})||^2 &= ||[G(\mathbf{x} + c\mathbf{y}) - G(\mathbf{x})] - cG(\mathbf{y})||^2 \\ &= ||G(\mathbf{x} + c\mathbf{y}) - G(\mathbf{x})||^2 + c^2||G(\mathbf{y})||^2 - 2c\big(G(\mathbf{x} + c\mathbf{y}) - G(\mathbf{x})\big)\bullet G(\mathbf{y}) \\ &= ||(\mathbf{x} + c\mathbf{y}) - \mathbf{x}||^2 + c^2||\mathbf{y}||^2 - 2cG(\mathbf{x} + c\mathbf{y})\bullet G(\mathbf{y}) + 2cG(\mathbf{x})\bullet G(\mathbf{y}) \\ &= 2c^2||\mathbf{y}||^2 - 2c(\mathbf{x} + c\mathbf{y})\bullet\mathbf{y} + 2c\mathbf{x}\bullet\mathbf{y} \\ &= 2c^2||\mathbf{y}||^2 - 2c\mathbf{x}\bullet\mathbf{y} - 2c^2||\mathbf{y}||^2 + 2c\mathbf{x}\bullet\mathbf{y} \\ &= 0,\end{aligned}$$

where we've used the facts that G is a rigid motion and that G preserves the scalar product.

COMMENT: This last theorem gives another way to think about Euclidean geometry. If you examine your high school geometry text carefully, you might conclude (depending on the text of course), that the important term *congruent* is not really well-defined. For instance, Wikipedia (not exactly the basic reference for mathematics) says that 2 figures are congruent "if they have the same shape and size", whatever that means. What they should have said is that 2 figures A and B are congruent if there's a rigid motion F such that $F(A) = B$.