RIGID MOTIONS AND ORTHOGONAL MATRICES

I am including the pages from Artin's *Algebra* (1st edition, pages 126–128) where he shows that a rigid motion fixing the origin is given by multiplication by an orthogonal matrix. An orthogonal matrix is a matrix A satisfying $A^T A = I$. The proof in the second edition (pages 156–157) is brilliant, short, but also tricky, so I felt that you might benefit from another proof.

(5.13) **Proposition.** The following conditions on a real $n \times n$ matrix A are equivalent:

(a) A is orthogonal.
(b) Multiplication by A preserves dot product, that is, $(AX \cdot AY) = (X \cdot Y)$ for all column vectors X, Y.
(c) The columns of A are mutually orthogonal unit vectors.

A basis consisting of mutually orthogonal unit vectors is called an *orthonormal* basis. An orthogonal matrix is one whose columns form an orthonormal basis.

Left multiplication by an orthogonal matrix is also called an *orthogonal operator*. Thus the orthogonal operators on $\mathbb{R}^n$ are the ones which preserve dot product.

Proof of Proposition (5.13). We write $(X \cdot Y) = X^t Y$. If A is orthogonal, then $A^t A = I$, so

$$(X \cdot Y) = X^t Y = X^t A^t A Y = (AX)^t (AY) = (AX \cdot AY).$$

Conversely, suppose that $X^t Y = X^t A^t A Y$ for all X and Y. We rewrite this equality as $X^t B Y = 0$, where $B = I - A^t A$. For any matrix B,

$$e_i{}^t B e_j = b_{ij}. \tag{5.14}$$

So if $X^t B Y = 0$ for all X, Y, then $e_i{}^t B e_j = b_{ij} = 0$ for all i, j, and $B = 0$. Therefore $I = A^t A$. This proves the equivalence of (a) and (b). To prove that (a) and (c) are equivalent, let A_j denote the jth column of the matrix A. The (i, j) entry of the product matrix $A^t A$ is $(A_i \cdot A_j)$. Thus $A^t A = I$ if and only if $(A_i \cdot A_i) = 1$ for all i, and $(A_i \cdot A_j) = 0$ for all $i \neq j$, which is to say that the columns have length 1 and are orthogonal. □

The geometric meaning of multiplication by an orthogonal matrix can be explained in terms of rigid motions. A *rigid motion* or *isometry* of $\mathbb{R}^n$ is a map $m: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ which is distance preserving; that is, it is a map satisfying the following condition: If X, Y are points of $\mathbb{R}^n$, then the distance from X to Y is equal to the distance from $m(X)$ to $m(Y)$:

$$|m(X) - m(Y)| = |X - Y|. \tag{5.15}$$

Such a rigid motion carries a triangle to a congruent triangle, and therefore it preserves angles and shapes in general.

Note that the composition of two rigid motions is a rigid motion, and that the inverse of a rigid motion is a rigid motion. Therefore the rigid motions of $\mathbb{R}^n$ form a group M_n, with composition of operations as its law of composition. This group is called the *group of motions*.

(5.16) **Proposition.** Let m be a map $\mathbb{R}^n \longrightarrow \mathbb{R}^n$. The following conditions on m are equivalent:

(a) m is a rigid motion which fixes the origin.
(b) m preserves dot product; that is, for all $X, Y \in \mathbb{R}^n$, $(m(X) \cdot m(Y)) = (X \cdot Y)$.
(c) m is left multiplication by an orthogonal matrix.

(5.17) **Corollary.** A rigid motion which fixes the origin is a linear operator.

This follows from the equivalence of (a) and (c).

Proof of Proposition (5.16). We will use the shorthand $'$ to denote the map m, writing $m(X) = X'$. Suppose that m is a rigid motion fixing 0. With the shorthand notation, the statement (5.15) that m preserves distance reads

$$(X' - Y' \cdot X' - Y') = (X - Y \cdot X - Y) \tag{5.18}$$

for all vectors X, Y. Setting $Y = 0$ shows that $(X' \cdot X') = (X \cdot X)$ for all X. We expand both sides of (5.18) and cancel $(X \cdot X)$ and $(Y \cdot Y)$, obtaining $(X' \cdot Y') = (X \cdot Y)$. This shows that m preserves dot product, hence that (a) implies (b).

To prove that (b) implies (c), we note that the only map which preserves dot product and which also fixes each of the basis vectors e_i is the identity. For, if m preserves dot product, then $(X \cdot e_j) = (X' \cdot e_j')$ for any X. If $e_j' = e_j$ as well, then

$$x_j = (X \cdot e_j) = (X' \cdot e_j') = (X' \cdot e_j) = x_j'$$

for all j. Hence $X = X'$, and m is the identity.

Now suppose that m preserves dot product. Then the images $e_1', \ldots, e_n'$ of the standard basis vectors are orthonormal: $(e_i' \cdot e_i') = 1$ and $(e_i' \cdot e_j') = 0$ if $i \neq j$. Let $\mathbf{B}' = (e_1', \ldots, e_n')$, and let $A = [\mathbf{B}']$. According to Proposition (5.13), A is an orthogonal matrix. Since the orthogonal matrices form a group, A^{-1} is also orthogonal. This being so, multiplication by A^{-1} preserves dot product too. So the composed motion $A^{-1}m$ preserves dot product, and it fixes each of the basis vectors e_i. Therefore $A^{-1}m$ is the identity map. This shows that m is left multiplication by A, as required.

Finally, if m is a linear operator whose matrix A is orthogonal, then $X' - Y' = (X - Y)'$ because m is linear, and $|X' - Y'| = |(X - Y)'| = |X - Y|$ by (5.13b). So m is a rigid motion. Since a linear operator also fixes 0, this shows that (c) implies (a). □

One class of rigid motions which do not fix the origin, and which are therefore not linear operators, is the translations. Given any fixed vector $b = (b_1, \ldots, b_n)^t$ in $\mathbb{R}^n$, *translation* by b is the map

(5.19)
$$t_b(X) = X + b = \begin{bmatrix} x_1 + b_1 \\ \vdots \\ x_n + b_n \end{bmatrix}.$$

This map is a rigid motion because $t_b(X) - t_b(Y) = (X + b) - (Y + b) = X - Y$, and hence $|t_b(X) - t_b(Y)| = |X - Y|$.

(5.20) **Proposition.** Every rigid motion m is the composition of an orthogonal linear operator and a translation. In other words, it has the form $m(X) = AX + b$ for some orthogonal matrix A and some vector b.

Proof. Let $b = m(0)$. Then $t_{-b}(b) = 0$, so the composed operation $t_{-b}m$ is a rigid motion which fixes the origin: $t_{-b}(m(0)) = 0$. According to Proposition (5.16), $t_{-b}m$ is left multiplication by an orthogonal matrix A: $t_{-b}m(X) = AX$. Applying t_b to both sides of this equation, we find $m(X) = AX + b$.

Note that both the vector b and the matrix A are uniquely determined by m, because $b = m(0)$ and A is the operator $t_{-b}m$. □