Congruence and rigid motions

This note is meant to explain the physical meaning of the concept of strong congruence. Some of the points described below were mentioned informally in the lectures. The basic reference is Section II.4 of the lecture notes.

If X and Y are subsets of $\mathbf{R}^n$, then X and Y are weakly congruent if there is an isometry F from $\mathbf{R}^n$ to itself such that Y is the image of X. The results in Section II.4 imply that every isometry F is given by a first degree vector equation $F(\mathbf{v}) = A\mathbf{v} + \mathbf{c}$, where A is an orthogonal matrix and $\mathbf{c} \in \mathbf{R}^n$. A matrix P is orthogonal if and only if its inverse is equal to its transpose, so if P is orthogonal then

$$\det P \quad = \quad \det {}^{\mathbf{T}}P \quad = \quad (\det P)^{-1} \ .$$

This implies that $\det P = \pm 1$, and we have defined X and Y to be strongly congruent if one can choose the isometry F such that $\det A = +1$.

Rigid motions

The following definition formalizes the intuitive idea of taking an object and moving it without changing its size or shape.

Definition. A *rigid motion* of $\mathbf{R}^n$ is a continuous function H from $\mathbf{R}^n \times [a, b]$ to $\mathbf{R}^n$, where $[a, b]$ is a closed interval in the real line, such that the following hold:

(i) For all $\mathbf{v} \in \mathbf{R}^n$ we have $H(\mathbf{v}, a) = \mathbf{v}$.

(ii) For every $t \in [a, b]$ the map F_t from $\mathbf{R}^n$ to itself defined by $F_t(\mathbf{v}) - H(\mathbf{v}, t)$ is an isometry.

(iii) If A_t is the orthogonal matrix such that $A_t = DF_t(\mathbf{u})$ for some (equivalently, for all) $\mathbf{u} \in \mathbf{R}^n$, then the matrix valued function A_t is a continuous function of t (*i.e.*, its coordinates are continuous functions of t).

Notes. **1.** The set $\mathbf{R}^n \times [a, b]$ has a natural interpretation as a subspace of $\mathbf{R}^{n+1}$, and using this interpretation it is meaningful to discuss continuity of a function defined on the given set.

2. We know that if G is an isometry of $\mathbf{R}^n$ and $G(\mathbf{v}) = B\mathbf{v} + \mathbf{w}$, then the derivative of G at every point is equal to B, so the derivative matrix $DG(\mathbf{u})$ is a constant function of $\mathbf{u}$.

Definition. If $X \subset \mathbf{R}^n$, then for each $t \in [a, b]$ the image set $X_t = F_t(X)$ is weakly congruent to $X = X_a$, and we say that each such set is obtained from X by a **continuous rigid motion** or that there is a continuous rigid motion taking X to each set X_t.

In fact we can draw the following stronger conclusion about each of the sets X_t:

PROPOSITION. *If X_t and the other notation is given as in the preceding discussion, then X_t is strongly congruent to X for all t.*

Proof. As usual write $F_t(\mathbf{v}) = A_t\mathbf{v} + \mathbf{c}_t$ where A_t is orthogonal. By construction we know that A_0 is the identity matrix. Consider the function $q(t) = \det A_t$. The hypotheses on rigid motions, and the expressibility of a determinant as a polynomial in the entries of a matrix, imply that $q(t)$ is continuous and $q(0) = 1$. Since each A_t is orthogonal we know that $q(t) = \pm 1$ for all t. The only way one can have a continuous function of this sort is if $q(t) = 1$ for all t. Therefore $\det A_t = +1$, which implies that for each t the set X_t is strongly congruent to X.■

Strong congruence implies continuous rigid motion

The rest of the discussion is aimed at proving a converse to the preceding result:

THEOREM. *Suppose that X and Y are subsets of $\mathbf{R}^n$ and X is strongly congruent to Y. Then there is a continuous rigid motion taking X to Y.*

Proof. Suppose that A is an arbitrary $n \times n$ matrix. Then standard results from linear algebra imply the existence of an orthonormal basis for $\mathbf{R}^n$ of the form

$$\{\, \mathbf{u}_1, \mathbf{v}_1, \ \cdots, \mathbf{u}_k, \mathbf{v}_k, \ \mathbf{w}_1 \ \cdots \ , \mathbf{w}_\ell \,\}$$

such that for each j between 1 and k we have

$$\begin{aligned} A(\mathbf{u}_j) &= \cos\theta_j\, \mathbf{u}_j + \sin\theta_j\, \mathbf{v}_j \\ A(\mathbf{v}_j) &= -\sin\theta_j\, \mathbf{u}_j + \cos\theta_j\, \mathbf{v}_j \end{aligned}$$

for some real numbers θ_j, and for each i between 1 and ℓ we have

$$A(\mathbf{w}_i) = \varepsilon_i\, \mathbf{w}_i$$

where $\varepsilon_i = \pm 1$. We allow cases where $k = 0$ or $\ell = 0$, the only restriction being that $2k + \ell$ must be equal to n. A proof of this result, which only uses material from Mathematics 131 and 132a, appears in Appendix D of the following online notes:

`http://math.ucr.edu/~res/math205A.gentopnotes2005.*`

We shall need the following slight refinement of this normal form for orthogonal matrices.

CLAIM. *It is possible to choose k and ℓ such that $\ell \leq 2$, with $\varepsilon_i = (-1)^i$ if $\ell = 2$.* — Here is a PROOF: First of all, we can rearrange the $\mathbf{w}$'s so that the vectors with negative eigenvalues precede the vectors with positive eigenvalues. If we have a pair of eigenvectors $\mathbf{w}_1$ and $\mathbf{w}_2$ with eigenvalue -1 we can move them to the first part of the orthonormal basis, renaming them as $\mathbf{u}_{k+1}$ and $\mathbf{v}_{k+1}$ and taking $\theta_{k+1} = \pi$. Repeating this if necessary, we can modify the original form so that the list of $\mathbf{w}$'s with eigenvalue equal to -1 has length zero or one. Likewise, if we have a pair of eigenvectors $\mathbf{w}_m$ and $\mathbf{w}_{m+1}$ with eigenvalue $+1$ we can move them to the first part of the orthonormal basis, renaming them as $\mathbf{u}_{k+1}$ and $\mathbf{v}_{k+1}$ and taking $\theta_{k+1} = 2\pi$. Once again, we can repeat this if necessary so that the list of $\mathbf{w}$'s with eigenvalue equal to $+1$ has length zero or one.

Note that the sign of $\det A$ is negative if there is an eigenvector with eigenvalue equal to -1 and the determinant of A is positive otherwise.

Suppose now that $\det A = +1$ so that there are no eigenvectors for the eigenvalue -1 and $\ell \leq 1$; in fact, $\ell = 0$ if n is even and $\ell = 1$ if n is odd. We may now define a continuous matrix valued function on the unit interval [0,1] by means of the following parametrized family of linear transformations:

$$\begin{aligned} L_t(\mathbf{u}_j) &= \cos(t\theta_j)\, \mathbf{u}_j + \sin(t\theta_j)\, \mathbf{v}_j \\ L_t(\mathbf{v}_j) &= -\sin(t\theta_j)\, \mathbf{u}_j + \cos(t\theta_j)\, \mathbf{v}_j \\ L_t(\mathbf{w}_i) &= \mathbf{w}_i \end{aligned}$$

Specifically, let B_t denote the matrices of the linear transformations L_t with respect to the given orthonormal basis. Clearly B_t is continuous in t. The result on normal forms implies that $A = P^{-1} B\, P$ for some orthogonal matrix P, and if we take $A_t = P^{-1} B_t\, P$ it will follow that A_t is also continuous in t. By construction both B_0 and A_0 are equal to the identity matrix, and we also have $A_1 = A$.

We now have enough data to define a rigid motion H such that F_0 is the identity and F_1 is the original isometry F. It is a routine exercise to check that this mapping has all the required properties.∎