Normal forms for orthogonal transformations

The Spectral Theorem in linear algebra implies that a normal linear transformation on a complex inner product space (one that commutes with its adjoint) has an orthonormal basis of eigenvectors. In particular, since the adjoint of a unitary transformation is its inverse, the result implies that every unitary transformation has an orthonormal basis of eigenvectors.

It is clear that one cannot have a direct generalization of the preceding result to orthogonal transformations on real inner product spaces. In particular, plane rotations given by matrices of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

do not have such a basis over the reals except in the relatively trivial cases when θ is an integral multiple of π and the matrices reduce to $(-I)^k$ for some integer k. However, if one takes these into account it is possible to prove the following strong result on the existence of a "good" orthonormal basis for a given orthogonal transformation.

NORMAL FORM. *Let V be a finite-dimensional real inner product space, and let $T : V \to V$ be an orthogonal transformation of V. Then there is an orthonormal direct sum decomposition of V into T-invariant subspaces W_i such that the dimension of each W_i is either* 1 *or* 2*.*

In particular, this result implies that there is an ordered orthonormal basis for V such that the matrix of T with respect to this ordered orthonormal basis is a block sum of 2×2 and 1×1 orthogonal matrices.

It is beyond the scope of these notes to go into detail about the results from a standard linear algebra course that we use in the proof of the result on normal forms. Virtually all of the background information can be found in nearly any linear algebra text that includes a proof of the Spectral Theorem.

Small dimensional orthogonal transformations

Since orthogonal transformations preserve the lengths of vectors it is clear that a 1-dimensional orthogonal transformation is just multiplication by ± 1. It is also not difficult to describe 2-dimensional orthogonal transformations completely using the fact that their columns must be orthonormal. In particular, it follows that the matrices representing 2-dimensional orthogonal transformations of $\mathbf{R}^2$ with respect to the standard inner product have the form

$$\begin{pmatrix} \cos\theta & \mp\sin\theta \\ \sin\theta & \pm\cos\theta \end{pmatrix}$$

for some real number θ. We have already noted that one class of cases corresponds to rotation through θ, and it is an exercise in linear algebra to check that each of the remaining matrices

$$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

has an orthonormal basis of eigenvectors, with one vector in the basis having eigenvalue -1 and the other having eigenvalue $+1$. The details of verifying this are left to the reader as an exercise (*Hint:* First verify that the characteristic polynomial is $t^2 - 1$).

One can use this to prove a geometrically sharper version of the result on normal form.

STRONG NORMAL FORM. *Let V be a finite-dimensional real inner product space, and let $T : V \to V$ be an orthogonal transformation of V. Then there is an orthonormal direct sum decomposition of V into T-invariant subspaces W_i such that the dimension of each W_i is either* 1 *or* 2 *and T operates by a plane rotation on each* 2*-dimensional summand.*

This is an immediate consequence of the preceding discussion, for if T operates on an invariant 2-dimensional subspace with by a map that is not a rotation, then we can split the subspace into two 1-dimensional eigenspaces.

Complexification

If $V = \mathbf{R}^n$ and T is represented by the orthogonal matrix A, it is clear how one can extend T to a unitary transformation on $\mathbf{C}^n$. We shall need a version of this principle that works for an arbitrary real inner product space V. The idea is simple but a little inelegant; it can be done better if one uses tensor products, but we want to prove the result without introducing them.

One defines the complexification of V formally to be very much as one defines the complex numbers. The underlying set of complex vectors $V_{\mathbf{C}}$ is given by $V \times V$, addition is defined in a coordinatewise fashion, multiplication by a complex scalar $a + bi$ is given by the formula

$$[a + bi](v, w) = (av - bw, bv + aw)$$

and the complexified inner product is defined by

$$\langle\, (v, w), (v', w')\, \rangle_{\text{complex}} =$$

$$(\langle v, v' \rangle + \langle w, w' \rangle) + i \cdot (\langle w, v' \rangle - \langle v, w' \rangle) \ .$$

Both intuitively and formally the pair (v, w) can be viewed as $v + i\, w$ if one identifies a vector v in the original space with $(v, 0)$ in the complexification. Verification that the structure defined above is actually a complex inner product space is essentially an exercise in bookkeeping and will not be carried out here.

If we are given a linear transformation $T : V \to W$ of real inner product spaces, then

$$T \times T : V \times V \to W \times W$$

defines a complex linear transformation $T_{\mathbf{C}}$ on the complexification. Moreover, this construction is compatible with taking adjoints and composites:

$$(T^*)_{\mathbf{C}} = (T_{\mathbf{C}})^*$$

$$(S \circ T)_{\mathbf{C}} = S_{\mathbf{C}} \circ T_{\mathbf{C}}$$

In particular, *if T is orthogonal then $T_{\mathbf{C}}$ is unitary.*

The main ideas behind the proof are

(1) to extract as much information as possible using the Spectral Theorem for the unitary transformation $T_{\mathbf{C}}$,

(2) prove the result by induction on the dimension of V.

In order to do the latter we need the following observation that mirrors one step in the proof of the Spectral Theorem.

PROPOSITION. *Let T be as above, and assume that $W \subset V$ is T-invariant. Then the orthogonal complement $W^\perp$ is also T-invariant.*∎

In the Spectral Theorem the induction proof begins by noting that there is an invariant 1-dimensional subspace corresponding to an eigenvector for T. One then obtains a transformation on the orthogonal complement of this subspace and can apply the inductive hypothesis to the associated transformation on this invariant subspace. We would like to do something similar here, and in order to begin the induction we need the following result.

KEY LEMMA. *If V is a finite dimensional real inner product space and $T : V \to V$ is an orthogonal transformation, then there is a subspace $W \subset V$ of dimension 1 or 2 that is T-invariant.*

If we have this subpsace, then we can proceed as before, using the induction hypothesis to split $W^\perp$ into an orthogonal direct sum of T-invariant subspaces of dimension 1 or 2, and this will complete the derivation of the normal form.∎

Proof of the key lemma. Suppose that λ is an eigenvalue of $T_{\mathbf{C}}$ and x is a nonzero eigenvector; since $T_{\mathbf{C}}$ is unitary we know that $|\lambda| = 1$. Express both λ and x in terms of their real and imaginary components, using the fact that $|\lambda| = 1$:

$$\lambda \quad = \quad \cos\theta + i\,\sin\theta, \qquad x \quad = \quad (v, w) \quad = \quad v + i\,w$$

Then the eigenvalue equation $T_{\mathbf{C}}(x) = \lambda\,x$ may be rewritten in the form

$$T(v) + i\,T(w) \quad = \quad T_{\mathbf{C}}(x) \quad = \quad \lambda\,x \quad = \quad (\cos\theta + i\,\sin\theta)\cdot(v + i\,w) \quad =$$

$$\Big(\cos\theta\,v - \sin\theta\,w\Big) + i\,\Big(\sin\theta\,v + \cos\theta\,w\Big)$$

which yields the following pair of equations:

$$T(v) \quad = \quad \cos\theta\,v - \sin\theta\,w$$

$$T(w) \quad = \quad \sin\theta\,v + \cos\theta\,w$$

It follows that the subspace W spanned by v and w is a T-invariant subspace, and since it is spanned by two vectors its dimension is at most 2. On the other hand, since $x \neq 0$ we also know that at least one of v and w is nonzero and therefore the dimension of W is at least 1.∎