Orthogonal Transformations

In the previous section we studied orthogonal matrices G. These are $n \times n$ matrices $G = [v_1, \ldots, v_n]$ of column matrices such that v_i are mutually orthogonal (i.e. $v_i^t v_j = 0$ for $i \neq j$) unit vectors (i.e. $v_i^t v_i = 1$); in other words, the basis $v_1, \ldots, v_n$ is an *orthonormal* basis. It follows easily that $G^t G = 1$ so that $G^t = G^{-1}$. Moreover, for any pair v, w of column vectors, we have $(Gw)^t(Gv) = w^t(G^tG)v = w^tv$. Thus, we can geometrically describe them by saying that the preserve lengths of vectors and angles between vectors.

From this description it is clear that a product of orthogonal matrix is an orthogonal matrix. The inverse of an orthogonal matrix is orthogonal and the identity matrix is orthogonal. In other words, orthogonal matrices form a group.

Conversely, suppose given a linear transformation $v \mapsto Gv$ which satisfies the property that $(Gw)^t(Gv) = w^tv$ for all vectors v and w. It follows that if $v_i = Ge_i$ (where e_i is the standard column vector) then the vectors v_i form an orthonormal basis. Thus G is given by an orthogonal matrix. Such a G is called an orthogonal transformation. We will use the term orthogonal matrix and orthogonal transformation interchangeably.

In this section, we will study orthogonal matrices (or the corresponding transformations) in various ways.

Products of Reflections

We work in a field of characteristic different from 2. In other words, 1/2 lies in our field.

Given a non-zero (column) vector v, we define an orthogonal transformation R_v which sends v to $-v$ and is identity in the plane perpendicular to v; due to similarity with reflection in a mirror, such a transformation is called a *reflection* and is given by the formula

$$R_v(w) = w - 2\frac{v^tw}{v^tv}v$$

We note that $R_v \cdot R_v = 1$. Given two linearly independent vectors v and w such that $v^tv = w^tw$ we note that $(v-w)^t(v+w) = 0$. We can form the reflection R_{v-w}. We see that $R_{v-w}(v+w) = v+w$. On the other hand $R_{v-w}(v-w) = w-v$. It follows easily that $R_{v-w}(v) = w$ and $R_{v-w}(w) = v$.

Given an orthogonal matrix G. Suppose that G is different from identity, then there is a vector v for which $w = Gv \neq v$. Note that $w^tw = v^tv$ by the orthogonality of G. We can then form the orthogonal transformation $G' = R_{v-w}G$. From the above calculation, we see that $G'v = v$. We can now restrict G' to the vector space $v^\perp = \{w | w^tv = 0\}$ and repeat this process. It follows that any orthogonal $n \times n$ matrix can be written as a product of at most n reflections.

Rotations

Given v and w two linearly independent vectors consider the orthgonal transformation $R_v R_w$. If we restrict our attention to $< v, w >^{\perp} = \{u | u^t v = u^t w = 0\}$, then $R_v R_w$ acts as identity. Hence, we only need to understand its action on the plane spanned by v and w.

Exercise: Calculate the matrix of the linear transformation $R_v R_w$ on the plane $< v, w >$ spanned by v and w in a suitable orthonormal basis, and show that it has the form $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ where $a^2 + b^2 = 1$ and (a, b) can be calculated in terms of $v^t w$.

In other words, this is a plane rotation. Combining this with the arguments given above, we can show that any orthogonal matrix is a product of at most $n/2$ rotations and possibly one reflection. (We can combine the reflections in adjacent pairs to give rotations.)

However, there is nothing "canonical" about such an expression. Moreover, the rotations are not in planes perpendicular to each other.

Canonical Form of an Orthogonal matrix

Given an orthogonal matrix G. We can collect all eigenvectors of G with eigenvalue ± 1 into a subspace V and restrict our attention to the subspace $W = V^{\perp}$. Hence, we can assume that ± 1 are not eigenvalues of G.

The Cayley transform of G is defined as $H = (G+1)(G-1)^{-1}$. We see easily that H is skew-symmetric and invertible, so that it does not have 0 as an eigenvalue.

Given a skew-symmetric matrix H with the additional assumption that 0 is not an eigen value of H. Consider the matrix $A = \begin{pmatrix} 0 & H \\ -H & 0 \end{pmatrix}$. We see that A is a symmetric matrix and so, as seen above, it has an eigenvector $(p, q)^t$. It follows that $Hp = aq$ and $Hq = -ap$ for a suitable real number a. Since 0 is not an eigenvalue, we see that p and q are linearly independent.

One then checks that $Gp = bp + cq$ and $Gq = -cp + bq$ where $b = (1-a^2)/(1+a^2)$ and $c = 2a/(1 + a^2$ so that $b^2 + c^2 = 1$. In other words, G is a rotation restricted to the plane $< p, q >$ spanned by p and q.

Repeating the above with the space perpendicular to the plane $< p, q >$ allows us to write G in block diagonal form with blocks consisting of 1's, -1's and planar rotations. (Note that this proof did not use the fundamental theorem of algebra.) These rotations are on planes that given an orthogonal decomposition of W.

Over the field of complex numbers $p + \sqrt{-1}q$ and $p - \sqrt{-1}q$ can be seen to be eigenvectors of G with eigenvalues $b + \sqrt{-1}c$ and $b - \sqrt{-1}c$ respectively. Doing this for every plane as above shows that G is diagonalisable over the field of complex numbers. In particular, this means that the minimal polynomial of G has distinct roots.