Cayley's parameterization of orthogonal matrices

Canonical name	CayleysParameterizationOfOrthogonalMatrices
Date of creation	2013-03-22 14:51:38
Last modified on	2013-03-22 14:51:38
Owner	rspuzio (6075)
Last modified by	rspuzio (6075)
Numerical id	20
Author	rspuzio (6075)
Entry type	Theorem
Classification	msc 22E70
Classification	msc 15A57
Defines	Cayley transform

Any orthogonal matrix O which does not have -1 as an eigenvalue can be expressed as

$$O = (I + A)(I - A)^{-1}$$

for some suitable skew-symmetric matrix A. Conversely, any skew-symmetric matrix A can be expressed in terms of a suitable orthogonal matrix O by a similar formula,

$$A = (O + I)^{-1}(O - I).$$

These two formulae are each other's inverses and set up a one-to-one correspondence between orthogonal and skew-symmetric matrices.

0.0.1 Proof

The restriction on the eigenvalues of O is necessary in order for $I + O$ to be invertible.

It is a matter of simple computation why these formulae are correct. Suppose that A is skew-symmetric. Then

$$O^T O = \left((I + A)(I - A)^{-1}\right)^T (I + A)(I - A)^{-1}$$

Using the fact that the transpose of a product is the product of the transposes in the opposite order,

$$= ((I - A)^{-1})^T (I + A)^T (I + A)(I - A)^{-1}$$

Using the fact that the transpose of a sum is the sum of transposes and the transpose of an inverse is the inverse of the transpose,

$$= (I^T - A^T)^{-1}(I^T + A^T)(I + A)(I - A)^{-1}$$

By the definition of skew-symmetry, $A^T = -A$ and $I^T = I$,

$$= (I + A)^{-1}(I - A)(I + A)(I - A)^{-1}$$

Finally, since $I + A$ and $I - A$ commute, we may switch the order of the second and third factors:

$$= (I + A)^{-1}(I + A)(I - A)(I - A)^{-1}$$

Then the first two factors and the last two factors cancel, showing that $O^T O = I$.

Next, we verify that the second formula is indeed the inverse of the first formula. Multiplying by $I - A$ on both sides,

$$O(I - A) = I + A$$

Expanding this and moving terms from one side of the equation to the other,

$$O - I = A + OA$$

Factoring,

$$O - I = (I + O)A$$

Multiplying both sides by $(I + O)^{-1}$, we obtain the desired formula:

$$(O + I)^{-1}(O - I) = A$$

Finally, one can show that, if O is orthogonal, then A is skew-symmetric using the same sort of computation that was used to show the converse:

$$A^T = \left((O + I)^{-1}(O - I)\right)^T$$

Using the facts about transposes of sums, products, and inverses,

$$= (O^T - I^T)(O^T + I^T)^{-1}$$

Since O is orthogonal, $O^T = O^{-1}$. As usual $I^T = I$.

$$= (O^{-1} - I)(O^{-1} + I)^{-1}$$

Insert an identity matrix between the two factors like so:

$$= (O^{-1} - I)I(O^{-1} + I)^{-1}$$

Replace the identity matrix with OO^{-1}:

$$= (O^{-1} - I)OO^{-1}(O^{-1} + I)^{-1}$$

Absorbing the O and the O^{-1} into the factors,

$$= (I - O)(I + O)^{-1} = -(O - I)(O + I)^{-1}.$$

Since

$$O(O + I)^{-1} = ((O + I)O^T)^{-1} = (I + O^T)^{-1} = (O^T(O + I))^{-1} = (O + I)^{-1}O,$$

$(O - I)$ and $(O + I)^{-1}$ commute, and consequently

$$= -(O + I)^{-1}(O - I) = -A.$$

0.0.2 Cayley Transform

The relation between A and O which is set up by the formulas

$$O = (I + A)(I - A)^{-1}$$

and

$$A = (O + I)^{-1}(O - I)$$

is sometimes known as the *Cayley transform*. Note that the proof that these two formulas are each other's inverses did not require A to be skew-symmetric or O to be orthogonal. Hence, the Cayley transform is defined for all matrices such that -1 is not an eigenvalue of O. (Recall that this condition is necessary to insure that $O + I$ is invertible.

0.0.3 Generalizations

The Cayley parameterization can be generalized to unitary transforms. Namely, if U is a unitary matrix, then U is the Cayley transform of a skew-Hermitean matrix A. Since a skew-Hermitean matrix can be written as i times a Hermitean matrix, the Cayley transform is often written as follows when dealing with unitary matrices:

$$U = (iI + H)(iI - H)^{-1}$$

$$iH = (U + I)^{-1}(U - I)$$

where H is Hermitean. The proof in this case is substantially the same as was presented above; all one has to do is replace matrix transposition with Hermitean conjugation.

A special case of this worth pointing out is the case of one-dimensional unitary matrices. The sole entry of a one dimensional unitary matrix must have modulus 1 and the sole entry of a one-dimensional Hermitean matrix must be real. In that case, the Cayley transform reduces to

$$u = \frac{i + h}{i - h}$$

$$ih = \frac{u + i}{u - i},$$

which is a fractional linear transform that maps the unit circle to the real axis.

The Cayley parameterization can be generalized to the case of a general inner product with arbitrary signature (see Sylvester's law for the definition of signature — Cayley and Sylvester were the best of friends). We simply need to define the transpose of a matrix M by the condition $(M^T u) \cdot v = u \cdot (Mv)$ for all vectors u and v. In particular, this allows one to parameterize pseudo-orthogonal matrices such as Lorentz transformations using a Cayley parameterization. Likewise, given a conjugate linear inner product on a complex vector space, one has a Cayley parameterization of the unitary (or pseudo-unitary) transforms which preserve the product.

In conclusion, it might be worth pointing out that the Cayley transform generalizes to the case of infinite dimensions, if one replaces matrices with operators on a Hilbert space. In particular, it is useful because unitary and orthogonal operators are bounded whereas Hermitean and skew-symmetric operators may or may not be bounded. For instance, it is often easier to obtain the spectral decomposition of a Hermitean operator or study symmetric extensions of a symmetric operator by first performing a Cayley transform and dealing with the resulting bounded operator.