Classification of linear transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$

In mathematics, one way we "understand" mathematical objects is to classify them (when we can). For this, we have some definition of the objects as being isomorphic (essentially the same), and then understand when two objects are isomorphic. If we're really lucky, we have a list of clear examples such that every object is isomorphic to one of them.

Here we'll do this for linear transformations T from $\mathbb{R}^2$ to $\mathbb{R}^2$. What can they look like? Each T has the form T_A, defined by $T_A(x) = Ax$, for some 2×2 matrix A. We define T_A and T_B to be isomorphic if there is an invertible linear transformation S from $\mathbb{R}^2$ to $\mathbb{R}^2$ such that $S^{-1}T_AS = T_B$.

We think of S as changing coordinates, just "renaming" points in $\mathbb{R}^2$ It is like: you can give a point an English name or a French name, and then go about the vector space business. Whatever names are given to describe points, the basic nature of addition and scalar multiplication is the same. S just translates from English to French, say.

Then, we can think of T_A as describing the linear transformation in English, and $S^{-1}T_AS = T_B$ as describing the linear transformation in French. Fundamentally the same thing is being done by T_A and T_B.

Above, $S = T_U$ for some invertible 2 matrix U. For T_A and T_B to be isomorphic is then equivalent to A and B being similar matrices: there exists U such that $U^{-1}AU = B$. So our classification problem is equivalent to classifying 2×2 matrices up to similarity. (Next page.)

Classification of 2×2 matrices up to similarity

Suppose A is a 2×2 matrix with real number entries. Then A will be similar to matrices of just a few types. Let χ_A denote the characteristic polynomial of A. Similar matrices have the same characteristic polynomial.

1. **Suppose χ_A has a root which is not a real number.**
   Let the root be $\lambda = a + ib$, where a and b are real and b is not zero. Then $a - ib$ must be the other root, because χ_A has real coefficients. The matrix A is (by Theorem 9, Sec. 5.5 of Lay) similar to the matrix

   $$C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} .$$

   We can understand the linear transformation T_C better geometrically by writing $a + ib$ in the form $r(\cos\theta + i\sin\theta)$, with $r > 0$. Then we have

   $$C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = r \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} .$$

   The linear transformation T_C, sending x to Cx, is a composition of two geometrically described actions. First, T_C rotates x counterclockwise around the origin through angle θ. Then it dilates by the factor r.

2. **Suppose χ_A has two distinct real roots, α and β.**
   Then A is similar to the diagonal matrix $\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$. $\mathbb{R}^2$ has a basis of eigenvectors of A.

3. **Suppose χ_A has one, repeated real root α and the dimension of the eigespace is 2.**
   This means the null space of $A - \alpha I$ has dimension 2.
   So, the null space of $A - \alpha I$ is $\mathbb{R}^2$, and this means $A - \alpha I = 0$.
   Therefore $A = \alpha I = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}$.

(Next page for the last case.)

4. Suppose χ_A has one, repeated real root α and the dimension of the eigenspace is 1.

Then A is similar to the matrix $= \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix}$.

(Such a matrix A will not be similar to αI as in Case 3.)

This case is the trickiest to prove. Let B denote the matrix $A - \alpha I$. Then B has rank 1 (because rank + dimension of null space = 2).
For characteristic polynomials, we have $\chi_A(t) = (t-\alpha)^2$ $\chi_B(t) = t^2$.
The characteristic polynomial of B has a repeated root, zero.

First, pick a nonzero vector w which is not in the null space of B. Then Bw is not zero.

CLAIM: Bw is in the null space of B.
PROOF OF CLAIM: Note, the range of B^2 is contained in the range of B, since for any x, $B^2x = B(Bx)$. Also, the dimension of the range of B is 1. Thus, since Bw is not zero, the vector B^2w must be a multiple of Bw. But the only eigenvalue of B is zero. Therefore Bw is in the null space of B. This proves the claim.

Now define v to be Bw. Then $0 = Bv = (A - \alpha I)v = Av - \alpha v$. So, $Av = \alpha v$. Also, $v = Bw = (A - \alpha I)w = Aw - \alpha w$, so $Aw = \alpha w + v$. Now define a matrix $U = \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix}$ which has v as its first column and w as its second column. It follows that

$$A \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} = \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix}$$

because the equality of first columns is $Av = \alpha v$ and the equality of second columns is $Aw = \alpha w + v$. Writing this matrix equality in terms of U, we get

$$AU = U \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix}$$

and therefore

$$U^{-1}AU = \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix} .$$