LINEAR TRANSFORMATIONS DEFINED GEOMETRICALLY

1. Introduction. Any $n \times n$ matrix A defines a linear transformation:

$$A : \mathbb{R}^n \to \mathbb{R}^n;$$

that is, matrices 'act on column vectors by left-multiplication' (move them around). Now we turn things around and ask: suppose we can describe a transformation T of $\mathbb{R}^n$ by the geometry of its effect on vectors: T 'rotates vectors by 30 degrees counterclockwise about the axis with direction $(1, 1, 1)$', or 'projects vectors onto the plane (subspace) $x + y + z = 0$, parallel to the vector $(2, 1, 1)$', etc. Suppose we know, in addition, that the transformation T is *linear*, that is, respects linear combinations:

$$T(c_1v_1 + \ldots + c_rv_r) = c_1T(v_1) + \ldots c_rT(v_r).$$

Problem: Can we find an $n \times n$ matrix A_0 having the same geometric action on vectors as T?

If we find such a matrix, then we can easily compute the effect of T on *any* vector in $\mathbb{R}^n$; with the geometric description alone, we can compute easily the action on only a small number of vectors (to be precise: only on special subspaces).

We already know one basic fact to help solve this problem: the columns of a matrix correspond to its action on the vectors e_i of the 'standard basis' of $\mathbb{R}^n$:

$$A_0 = [A_0e_1|A_0e_2|\ldots|A_0e_n] \quad \text{(by columns)}$$

Thus we can find A_0 if we can compute the action of T on the e_i.

Example 1. $T : \mathbb{R}^2 \to \mathbb{R}^2$ projects vectors onto the one-dimensional subspace E spanned by $v_1 = (2, 1)$, parallel to the line K (subspace) spanned by $v_2 = (1, 1)$. Thus we know:

$$T(2, 1) = (2, 1), \quad T(1, 1) = 0.$$

To compute the effect of T on the standard basis of $\mathbb{R}^2$, we note that:

$$e_1 = (1, 0) = (2, 1)-(1, 1) = v_1-v_2; \quad e_2 = (0, 1) = 2(1, 1)-(2, 1) = 2v_2-v_1.$$

Thus, by linearity:

$$Te_1 = Tv_1 - Tv_2 = (2, 1); \quad Te_2 = 2Tv_2 - Tv_1 = (-2, -1).$$

So A_0 is the matrix having $(2,1)$ and $(-2,-1)$ as column vectors:

$$A_0 = \begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix}.$$

From this we can easily compute the action of T on an arbitrary vector, say: $T(3,5) = (-2,-2)$.

(*Note:* There is a subtle abuse of notation here- we are identifying the geometrically defined transformation T with a 2×2 matrix A_0 that has the same action as T on any vector. Soon we'll see that there are many matrices associated with the same (geometrically defined) T; but since we always identify a vector in $\mathbb{R}^n$ with its 'coordinates in the standard basis' (see below), identifying T with its 'matrix in the standard basis' A_0 leads to correct results. This note will become clearer once Definitions 1,2 below have been absorbed.)

Assorted remarks on example 1. (1) Note that the column space of A_0 is exactly the subspace E, corresponding to the fact that the *image* of T is E (that is, T maps all of $\mathbb{R}^2$ to E); (2) The *nullspace* (or *kernel*) of A_0 has defining equation $x_1 = x_2$, so it coincides with the subspace K; this corresponds to the fact that T maps all of K to the zero vector. (3)Note that, in this case, $N(A_0)$ and $Col(A_0)$ intersect only at 0, and together span $\mathbb{R}^2$. This *not* true for a general T (although the *dimensions* of $N(T)$ and $Ran(T)$ always add up to n), but it is always true for projections (as we'll see later). (4) Once you've projected a vector onto E, projecting again won't do anything; that is, it is clear geometrically that $T^2(v) = T(v)$, for any $v \in \mathbb{R}^2$. On the matrix side, we would expect the corresponding identity: $A_0^2 = A_0$, and this is indeed true (check!)

Example 2. Let R_θ be the linear transformation of $\mathbb{R}^2$ that rotates vectors by θ radians, counterclockwise. From basic trigonometry:

$$R_\theta(e_1) = (\cos\theta, \sin\theta), \quad R_\theta(e_2) = (-\sin\theta, \cos\theta),$$

so we associate to R_θ the 'rotation matrix':

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

For example, to compute the effect of rotating $(4,2)$ by 60 degrees (counterclockwise) we use $R_{\pi/3}$:

$$\begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 2-\sqrt{3} \\ 2\sqrt{3}+1 \end{bmatrix}.$$

To go further, consider Example 1 again: there is a basis of $\mathbb{R}^2$ for which we can easily write down the action of T, namely: $\mathcal{B} = \{v_1, v_2\}$. Indeed,

$$Tv_1 = v_1 = 1v_1 + 0v_2, \quad Tv_2 = 0 = 0v_1 + 0v_2.$$

That is, the 'coordinate vector of Tv_1 in the basis $\mathcal{B}$' is $[Tv_1]_{\mathcal{B}} = (1, 0)$, and the 'coordinate vector of Tv_2 in the basis $\mathcal{B}$' is $[Tv_2]_{\mathcal{B}} = (0, 0)$. Using these 'coordinate vectors' as columns, we define the *matrix of T in the basis $\mathcal{B}$* as:

$$[T]_{\mathcal{B}} = [[Tv_1]_{\mathcal{B}} \quad |[Tv_2]_{\mathcal{B}}] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

This is by analogy with the matrix of T in the standard basis $\mathcal{B}_0 = \{e_1, e_2\}$. Above we computed $A_0 = [T]_{\mathcal{B}_0}$, the 'matrix of T in the standard basis'. The point of these contortions is that any projection in the Universe will have the very simple form of $[T]_{\mathcal{B}}$ above if we choose the basis $\mathcal{B}$ appropriately, and then there is a simple formula (given below) to get the matrix of T in the standard basis (the one we really use). Before introducing formal definitions, here is another simple example.

Example 3. T contracts vectors in E (spanned by $v_1 = (1, 1)$) by a factor $1/2$, expands vectors in $E^\perp$ (spanned by $v_2 = (1, -1)$) by a factor 3. We want the matrices of T in the basis $\mathcal{B} = \{(1, 1), (1, -1)\}$ and in the standard basis. Since, by the definition of T:

$$Tv_1 = (1/2)v_1, \quad Tv_2 = 3v_2,$$

we have immediately:

$$[T]_{\mathcal{B}} = \begin{bmatrix} 1/2 & 0 \\ 0 & 3 \end{bmatrix}.$$

We easily find the coordinates of e_1, e_2 in the basis $\mathcal{B}$:

$$(1, 0) = (1/2)(1, 1) + (1/2)(1, -1), \quad (0, 1) = (1/2)(1, 1) - (1/2)(1, -1),$$

so that:

$$T(1, 0) = (1/2)(1/2)v_1 + (1/2)3v_2 = (7/4, -5/4),$$
$$T(0, 1) = (1/2)(1/2)v_1 - (1/2)3v_2 = (-5/4, 7/4).$$

The matrix of T in the standard basis has these two vectors as columns:

$$A_0 = [T]_{\mathcal{B}_0} = \begin{bmatrix} 7/4 & -5/4 \\ -5/4 & 7/4 \end{bmatrix}.$$

Remark: A subspace of $\mathbb{R}^n$ in which T acts as a pure expansion/contraction (as E and $E^\perp$ in this example) is called an 'eigenspace' for T, with 'eigenvalue' given by the expansion/contraction factor (1/2 and 3 in this example). Precise definitions will be given soon.

2. Coordinate changes.

Any basis $\mathcal{B} = \{v_1, \ldots v_n\}$ of $\mathbb{R}^n$ determines a coordinate system. From the definition of basis, given any $v \in \mathbb{R}^n$ there is a unique expansion as a linear combination:

$$v = y_1 v_1 + y_2 v_2 + \ldots + y_n v_n.$$

Definition 1. The *coordinate vector* of v in the basis $\mathcal{B}$ (denoted $[v]_\mathcal{B}$) is:

$$[v]_\mathcal{B} = (y_1, \ldots, y_n), \text{ where } v = y_1 v_1 + \ldots + y_n v_n.$$

Given a basis $\mathcal{B}$ and a vector $v = (x_1, \ldots, x_n) \in \mathbb{R}^n$ (recall vectors are identified with their 'coordinate vectors' in the standard basis), how do we find the coordinates of v in the basis $\mathcal{B}$? Define an $n \times n$ matrix B by:

$$B = [v_1 | v_2 | \ldots | v_n] \quad \text{(by columns)}.$$

Then B is invertible (why?) and $Be_i = v_i$ for $i = 1, \ldots n$, where $\mathcal{B}_0 = \{e_1, \ldots, e_n\}$ is the standard basis. Thus if $v = y_1 v_1 + y_2 v_2 + \ldots y_n v_n$, we find:

$$B^{-1} v = y_1 B^{-1} v_1 + \ldots + y_n B^{-1} v_n = y_1 e_1 + \ldots y_n e_n = (y_1, \ldots y_n).$$

This means we have the *change of coordinates formula (for vectors):*

$$[v]_\mathcal{B} = B^{-1}[v]_{\mathcal{B}_0}, \quad [v]_{\mathcal{B}_0} = B[v]_\mathcal{B}$$

Analogous formulas hold for linear transformations and matrices. Given a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$, and a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$, we have:

Definition 2. The *matrix of T in the basis $\mathcal{B}$* (denoted $[T]_\mathcal{B}$) is defined by the relation:

$$[Tv]_\mathcal{B} = [T]_\mathcal{B}[v]_\mathcal{B}.$$

(As with vectors, we identify $[T]_{\mathcal{B}_0}$ with T.) To compute what this means, consider the expansions of Tv_i in the basis $\mathcal{B}$, for each of the basis vectors $v_1, \ldots, v_n$:

$$Tv_1 = a_{11} v_1 + a_{21} v_2 + \ldots + a_{n1} v_n, \ldots$$

$$Tv_j = a_{1j}v_1 + a_{2j}v_2 + \ldots + a_{nj}v_n, \ldots$$

$$Tv_n = a_{1n}v_1 + a_{2n}v_2 + \ldots + a_{nn}v_n.$$

That is, with summation notation:

$$Tv_j = \sum_{i=1}^{n} a_{ij}v_i.$$

Then if $[v]_{\mathcal{B}} = (y_1, \ldots, y_n)$, by linearity of T:

$$Tv = \sum_{j=1}^{n} y_j Tv_j = \sum_{i=1}^{n}(\sum_{j=1}^{n} a_{ij}y_j)v_i.$$

By Definition 1, this means:

$$[Tv]_{\mathcal{B}} = (\sum_{j=1}^{n} a_{1j}y_j, \ldots, \sum_{j=1}^{n} a_{nj}).$$

The right-hand side of this equality is exactly the result of multiplying the matrix $A = (a_{ij})$ by the (column) vector $(y_1, \ldots, y_n) = [v]_{\mathcal{B}}$. We conclude:

$$[T]_{\mathcal{B}} = A,$$

where the entries (a_{ij}) of A are defined above. Note that A is the $n \times n$ matrix whose columns are given by the coordinate vectors of Tv_i in the basis $\mathcal{B}$:

$$A = [T]_{\mathcal{B}} \Leftrightarrow A = [[Tv_1]_{\mathcal{B}}| \ldots |[Tv_n]_{\mathcal{B}}] \text{ (by columns).}$$

From the change of coordinates formula for vectors follows one for matrices. Since, for any $v \in \mathbb{R}^n$:

$$[v]_{\mathcal{B}} = B^{-1}[v]_{\mathcal{B}_0} \quad \text{ and } [Tv]_{\mathcal{B}} = B^{-1}[Tv]_{\mathcal{B}_0} = B^{-1}[T]_{\mathcal{B}_0}[v]_{\mathcal{B}_0},$$

it follows that:

$$[Tv]_{\mathcal{B}} = B^{-1}[T]_{\mathcal{B}_0}B[v]_{\mathcal{B}}.$$

Using Definition 2, we obtain the *change of coordinates formula for matrices:*

$$[T]_{\mathcal{B}} = B^{-1}[T]_{\mathcal{B}_0}B, \quad [T]_{\mathcal{B}_0} = B[T]_{\mathcal{B}}B^{-1}.$$

(This is one of the most useful formulas in Mathematics.)

Remark. Although this formula was derived for $\mathcal{B}_0$ thought of as the 'standard' basis, the same formula relates the matrices of a linear transformation T in two *arbitrary* bases $\mathcal{B}, \mathcal{B}_0$, as long as the 'change of basis matrix' B has as its column vectors the $[v_i]_{\mathcal{B}_0}$, where $\mathcal{B} = \{v_1, v_2, \ldots v_n\}$:

$$B = [[v_1]_{\mathcal{B}_0} | \ldots [v_n]_{\mathcal{B}_0}].$$

Example 3.-Projections. Let $P : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear operator that projects vectors onto the plane (subspace) $E = \{x; x_1 + x_2 + x_3 = 0\}$, parallel to the subspace F spanned by $(1, 2, 2)$ (a line). This means $v - Pv = c(1, 1, 2)$, for some constant c.

To compute the matrix of P in the standard basis, we first find the matrix in a basis $\mathcal{B} = \{v_1, v_2, v_3\}$ adapted to the situation: $v_1 = (1, 1, 2)$ and $\{v_2, v_3\}$ is a basis of E. For example, we may take $v_2 = (1, 0, -1), v_3 = (0, 1, -1)$. We clearly have (since v_2 and v_3 are in E):

$$Pv_1 = 0, \quad Pv_2 = v_2 = 0v_1 + 1v_2 + 0v_3, \quad Pv_3 = v_3 = 0v_1 + 0v_2 + 1v_3,$$

which gives the columns of the matrix of P in the basis $\mathcal{B}$:

$$[P]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The 'change of basis' matrix B and its inverse are:

$$B = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & -1 & -1 \end{bmatrix}, \quad B^{-1} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & -1 \\ -1 & 3 & -1 \end{bmatrix},$$

and the change of basis formula gives the matrix of P in the standard basis:

$$[P]_{\mathcal{B}_0} = B[P]_{\mathcal{B}}B^{-1} = \frac{1}{4} \begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -2 & -2 & 2 \end{bmatrix}.$$

Example 4.-Eigenvalues Using the same line F and plane E as in the previous example, let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation which contracts vectors by $1/2$ in E, and expands vectors by 5 in F. Formally:

$$Tv = (1/2)v, \quad v \in E; \quad Tv = 5v, \quad v \in F.$$

This means 5 and 1/2 are eigenvalues of T, with 'eigenspaces' F and E (respectively).

Using the same basis $\mathcal{B}$ as in example 3, we clearly have:

$$[T]_{\mathcal{B}} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}.$$

Using B and B^{-1} given above, we find the matrix of T in the standard basis:

$$[T]_{\mathcal{B}_0} = B[T]_{\mathcal{B}}B^{-1} = \frac{1}{8}\begin{bmatrix} 13 & 9 & 9 \\ 9 & 13 & 11 \\ 21 & 18 & 20 \end{bmatrix}.$$

3. Direct sums and projections. In all the examples above we have two subspaces E and F which together span all of $\mathbb{R}^n$, and in each of which a linear transformation T acts in a simple way. This is a common situation. In fact, given two subspaces E, F of $\mathbb{R}^n$, the following two conditions are equivalent:

1- Any $v \in \mathbb{R}^n$ can be written (in a unique way) as a sum $v = u + w$, where $u \in E$, $w \in F$;

2-$dimE + dimF = n$ and $E \cap F = \{0\}$ (that is, E and F intersect only at the origin.)

Whenever this happens, we say '$\mathbb{R}^n$ is the direct sum of subspaces E and F', denoted:

$$\mathbb{R}^n = E \oplus F.$$

(Note: in general E and F need not be orthogonal). In this situation, we may define $P = P_{E,F}$, the linear operator 'projection on E along F', by the rule:

$$Pv = u, \text{ if } v = u + w \text{ as in (1)},$$

or equivalently:

$$Pv = v \text{ if } v \in E; \quad Pv = 0 \text{ if } v \in F.$$

Of course, $P_{F,E}$ is similarly defined- it projects onto F along E- and, from (1), for all $v \in \mathbb{R}^n$:

$$v = P_{E,F}v + P_{F,E}v,$$

that is to say, we have a relation between these two operators:

$$P_{E,F} + P_{F,E} = Id_n,$$

the identity operator in $\mathbb{R}^n$. As seen in the example below, this is more useful than it sounds.

Let $P = P_{E,F}$. The following facts are completely obvious: (i) $P^2 = P$ (projecting again doesn't do anything new); (ii) Ran(P)=E; (iii)Kernel(P)=F. Conversely, given any linear operator $P : \mathbb{R}^n \to \mathbb{R}^n$, if $P^2 = P$ it is not hard to show that $Kernel(P) \cap Ran(P) = \{0\}$, and their dimensions add up to n (good exercise for the theoretically minded). That is, any operator satisfying $P^2 = P$ is automatically the projection onto its range, parallel to its kernel (as constructed above).

Definition 3. A linear operator in $\mathbb{R}^n$ is a *projection* if $P^2 = P$.

Example 5. Consider again the subspaces E and F of Example 3 above. Let $v = (1, 2, 3)$. Find the projections of v onto E (along F) and onto F (along E).

First solution: using the matrix of $P_{E,F}$ already computed, we find $u = P_{E,F}v$:

$$u = \frac{1}{4}\begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -2 & -2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 0 \end{bmatrix}.$$

And then $w = P_{F,E}v$ is obtained by taking the difference:

$$w = v - u = (3/2, 3/2, 3).$$

Second solution: Without using the change of coordinates formula, note that all we need to find are numbers c,w_1 and w_2 so that:

$$(1, 2, 3) = c(1, 1, 2) + u_1(1, 0, -1) + u_2(0, 1, -1).$$

That is, we need to solve the linear system:

$$c + u_1 = 1, \quad c + u_2 = 2, \quad 2c - u_1 - u_2 = 3.$$

Proceeding in the usual way, we find the unique solution:

$$c = 3/2, \quad u_1 = -1/2, \quad u_2 = 1/2,$$

which gives:

$$u = u_1(1, 0, -1)+u_2(0, 1, -1) = (-1/2, 1/2, 0), \quad w = c(1, 1, 2) = (3/2, 3/2, 3).$$