Linear Algebra: Characteristics of Spaces

Daven Farnham

August 2015

Linear Transformations

I want to try to connect the definition of a linear transformation with the definition of a homomorphism in group theory.

If I have two groups $(G, \circ)$ and $(H, \star)$ and some homomorphism between them $\varphi(x)$, then $\forall g_1, g_2 \in G$, $\varphi(g_1 g_2) = \varphi(g_1)\varphi(g_2)$ where the operation on the left comes from G and the operation on the right comes from H.

A homomorphism **preserves structural properties** s.t. the identity element in G, e_G, maps to the identity element in H, e_H, the inverse in G maps to the inverse in H, if G is abellian then H is abellian, etc.

Identity:

$$\varphi(e_G) = \varphi(e_G \circ e_G) = \varphi(e_G) \star \varphi(e_G)$$
$$\varphi(e_G) \star e_H = \varphi(e_G) \star \varphi(e_G)$$
$$e_H = \varphi(e_G)$$

Inverse:

$$\varphi(g_1 g_1^{-1}) = \varphi(g_1) \star \varphi(g_1^{-1})$$
$$\varphi(g_1 g_1^{-1}) = \varphi(e_G)$$
$$\varphi(e_G) = \varphi(g_1) \star \varphi(g_1^{-1}) = \varphi(g_1) \star \varphi(g_1)^{-1} \text{ (by the definition of inverses)}$$
$$\varphi(g_1) \star \varphi(g_1^{-1}) = \varphi(g_1) \star \varphi(g_1)^{-1}$$
$$\varphi(g_1^{-1}) = \varphi(g_1)^{-1}$$

This should look very much like the definition of a linear transformation, which says for a transformation $R^m \to R^n$:

- **(1)** T(0) = 0
- **(2)** T(v + w) = T(v) + T(w) for all $v, w \in R^m$

- **(3)** T(kx) = kT(x) for all $x \in R^m$

This is a transformation going from the reals to the reals. The identity element, in both cases, should be 0 since $0 + x = x$, which is the case in (1). This is really just a corner case of (2) (which can also be used to prove (3)).

The second property here is called **additivity** and I think you should look at it in the context of vectors. If I transform one vector, that should be the same as transforming its component vectors. Any vector will be a line, so a linear transformation is one where transforming components is the same as simply transforming the entire vector.

Basis, Image, Kernel

Section 3 in Bretscher is probably the most important in the book. In it, he discusses 3 key characteristics of matrices: basis, image, and kernel.

0.1 Basis

Given some space $V = \{v_1, v_2, ..., v_n\}$ where each v_n is a column vector, the basis of space V is the minimum number of vectors in V that completely describes the image of the matrix.

Put in other words, if I have a matrix A and I find its basis B, then the Image(A) = Image(B).

Furthermore, $\{\forall\ v_i \in$ the basis of $V\}$, v_i cannot be created through a **linear combination** of the other vectors $v_n | n \neq i$. This means all vectors in the basis are **non-redundant** and the only **linear relation** between them is the trivial one.

0.1.1 Linear Combination

A linear combination looks like this:

$$a_1v_1 + a_2v_2 +, ..., +a_nv_n = c$$

In a basis, since no vectors are redundant, given a vector v_i there's no way to create v_i from the remaining vectors. So, regardless what coefficients a_i you pick:

$$a_1v_1 + a_2v_2 +, ..., +a_{i-1}v_{i-1} + a_{i+1}v_{i+1} +, ..., +a_nv_n \neq v_i$$

A linear relation then is just an expansion on the idea of a linear combination. A linear relation is a linear combination that equals 0. So, taking the above example:

$$a_1v_1 + a_2v_2 +, ..., +a_nv_n = 0$$

is an example of a linear relation. If all vectors are non-redundant, meaning no vector can be created from the others, then it's impossible to get 0 unless all coefficients are 0 (which is the trivial relation). A super simple example involves the standard vectors in R^3, e_1, e_2, e_3. Regardless what non-zero coefficients you pick, no summation of the vectors will ever equal zero

$$a_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + a_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \neq 0 \tag{1}$$

which should be obvious from the fact that for each non-zero element in a particular vector, all the other vectors have a zero in that position; so there's no way to eliminate the value.

0.1.2 Unique Representation

On page 120 Bretscher gives a cool, alternative definition of a basis. If I have a basis of some space V in R^n, then how many ways are there to represent a vector $v \in V$?

It turns out there is only **one**. Since I have a basis, there'll be at least one way to represent v since that's the definition of a basis. Let's now use linear combinations to rewrite v and *assume* there are multiple ways to represent it:

$$v = a_1v_1 + a_2v_2 + ... + a_nv_n$$

$$v = b_1v_1 + b_2v_2 + ... + b_nv_n$$

$$0 = (a_1 - b_1)v_1 + (a_2 - b_2)v_2 + ... + (a_n - b_n)v_n$$

Since the only linear relation among basis vectors is the trivial one, $a_n = b_n$ for all coefficients in the above combinations, meaning there can be only a single representation for a vector $v \in V$.

0.2 Image

The image of a linear transformation is everything in the target space the transformation maps to. Specifically, it's a linear combination of matrix A's column vectors. Consider a transformation $T(Ax)$:

$$\begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{2}$$

In matrix multiplication, this simplifies to:

$$x_1 \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} \tag{3}$$

which means it maps $R^2 \to R^3$, going from a two dimensional domain to a three dimensional target space where every vector in the target space can be represented by a linear combination of the two vectors [**1 4 7**] and [**2 5 8**]. These two vectors **span** the target space, creating a plane (since there are only two vectors it would be impossible to span all of R^3).

To succinctly describe the image, it's good to eliminate any redundant vectors in A; that is, make sure all vectors in A are **linearly independent**. Otherwise, there's redundancy.

For example, in the above matrix A if there was another vector [**3 12 21**], well that's just 3 * [**1 4 7**]. You could describe

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 12 \\ 7 & 8 & 21 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{4}$$

like this:

$$x_1 \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 12 \\ 21 \end{bmatrix} \tag{5}$$

but [**1 4 7**] and [**3 12 21**] are multiples of each other and lie along the same line. Therefore, as x_1 varies from $-\infty$ to ∞, it will consume [**3 12 21**] whatever it is.

Eliminating this redundant vector, we get the **basis**.

0.3 Kernel

The kernel of an n by m matrix A are all vectors x in m that map to zero in the range. So the kernel is the set of all vectors in the domain that map to zero; put more succinctly:

$$\ker(\mathrm{A}) = \{x \in R^m \mid Ax = 0\}$$

Something that should be remembered is that **invertibility** implies bijectivity. If I have an invertible matrix A from $R^n \to R^m$, then I can pick any vector $y \in R^m$ and $\{\exists x \in R^n \mid A^{-1}[y] = x\}$. When it comes to the kernel, this means the kernel of an invertible matrix $= \{0\}$.

This should make sense, since an invertible matrix can be row reduced into the identity matrix. The identity matrix, then, is just the standard vectors and above I showed that the standard vectors are linearly independent, meaning there's no way to eliminate the value from e_1 with any combination of the other vectors $e_2, e_3, ..., e_n$.

The dimension of some subspace V is the number of vectors in its basis.

The **rank** of a matrix is the number of columns in the matrix with a leading variable. This matrix:

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{6}$$

then, has a rank of 3. You can also think of this as the number of linearly independent vectors in a matrix.

If the rank of a n by m matrix A equals m, then there are leading variables in each column of the matrix and thus, the kernel of the matrix has to be zero. Note, this doesn't immediately imply bijectivity though, if $m \neq n$ then the matrix isn't invertible.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} x \tag{7}$$

$$x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{8}$$

which means this is a transformation from $R^2 \to R^4$, so I can't pick any vector in R^4 and have it map back to a vector in R^2. For example, if I picked vector $[\mathbf{0}\ \mathbf{0}\ \mathbf{1}\ \mathbf{1}]$ there's absolutely no vector $x \in R^2$ that maps to it. Therefore, although the above transformation is **injective**, it is not **surjective** and thus not bijective.

A couple things to keep in mind:

- if $m > n$ and every column vector in m is non-trivial (i.e. not the zero vector) then there has to be free variables and thus some vector in the kernel.

- You can think of the above, also, in the frame of going from a higher dimension to a lower dimension, and thus you have to *collapse* the vectors of the domain into a smaller dimension. Therefore, some of the vectors from the higher dimension have to be relegated to the kernel and map to zero.
- If you go in the other direction, like the above example when I went from $R^2 \to R^4$, the kernel won't always be zero (like if you have redundant vectors) but there's the possibility the kernel will equal zero. In this case, the matrix is kind of **expanding** upon the dimension of the domain and thus there's no need to cram multiple vector types into the kernel.

0.4 Rank-Nullity

Given some matrix A in R^n, the number of vectors that make up its image (dim(Im(A)) or rank(A)) + the number of vectors in its kernel (Nullity(A)) = n.

Thinking about this in terms of the vectors in the matrix A, all column vectors with a leading 1 end up in the image whereas all column vectors with free variables are in the kernel. Number of Leading variables + number of Free variables = total number of columns, so the Dim(Im(A)) + Nullity(A) = n.