Notes On Linear Algebra

Spencer T. Parkin

August 14, 2012

Abstract

This paper is a formal compilation of all my notes on linear algebra. It proceeds as a treatment of linear algebra guided by the material given in the reference section, but here in some places I go into greater detail so as to understand more fully what is being communicated in those sources. The paper [1] is studied in particular.

1 Enter Linear Functions

Linear algebra is the study of linear functions defined on linear spaces. Linear spaces are more often refered to as vector spaces, suggesting to the mind a geometric interpretation of the elements of such spaces. In fact, what we'll find is that geometric algebra facilitates the study of linear algebra.

Holding the definitions of a vector space and linear independence as already known, we begin with a formal definition of a linear function. We arbitrarily define all vector spaces over the field of real numbers $\mathbb{R}$.

Definition 1.1. *A function $f : \mathbb{A} \to \mathbb{B}$ is a mapping from a vector space $\mathbb{A}$ to a vector space $\mathbb{B}$ with the property of preserving both scalar-vector multiplication and vector addition. That is, for any scalar $\lambda \in \mathbb{R}$, and any two vectors $x, y \in \mathbb{A}$, we have $f(\lambda x) = \lambda f(x)$ and $f(x + y) = f(x) + f(y)$.*

It is not enitrely clear to me how much, if any, loss in generality we incur by restricting our study of such functions to those that map to and from the same vector space. Nevertheless, since this is the class of linear functions for which I am most interested, we will proceed with this restriction.

That said, let $f : \mathbb{V}^n \to \mathbb{V}^n$ be the linear function we will study, where $\mathbb{V}^n$ denotes an n-dimensional vector space. Then, for any $x \in \mathbb{V}^n$, right away we learn two interesting things about linear functions. Letting $\{e_i\}_{i=1}^n \subset \mathbb{V}^n$ be any set of n linearly independent vectors taken from $\mathbb{V}^n$, we have

$$f(x) = f\left(\sum_{k=1}^{n} x_i e_i\right) = \sum_{k=1}^{n} x_i f(e_i), \tag{1}$$

where $x_i = x \cdot e_i$. The first thing this shows is that any linear function is determined entirely by how it transforms the set of basis vectors $\{e_i\}_{i=1}^n$, so that when we're faced with formulating a linear transformation, we need only consider how it transforms a basis of $\mathbb{V}^n$. The second thing this shows is that every linear function is a change of basis transformation. That is, the set of coordinates $\{x_i\}_{i=1}^n$ for a vector x are preserved while the set of basis vectors $\{e_i\}_{i=1}^n$ are replaced with a new set of vectors $\{f(e_i)\}_{i=1}^n$.

2 Existance and Uniqueness of Linear Function Inverses

Equation (1) also brings to bair immediate implications on the invertibility of f. That is, we have enough to prove at this point that f is invertible if and only if f preserves linear independence in the sense that if $\{e_i\}_{i=1}^n$ is a linearly independent set, then so is $\{f(e_i)\}_{i=1}^n$. Many graphics transformations can be easily formulated this way.

Recall that one direction of the statement $x = y \iff f(x) = f(y)$ is the requirement of a well defined function, while the other direction is optional, and is the condition upon which f^{-1} exists. Specifically, if there exist distinct vectors $x, y \in \mathbb{V}^n$ such that $f(x) = f(y)$, then f^{-1} does not exist. If $z = f(x) = f(y)$, then do we let $f^{-1}(z) = x$ or $f^{-1}(z) = y$?

Suppose for the moment that the set $\{f(e_i)\}_{i=1}^n$ is linearly dependent. Then, without loss of generality, we can write $f(e_n)$ as $\sum_{i=1}^{n-1} \lambda_i f(e_i)$, where each $\lambda_i \in \mathbb{R}$. Now let $x_n = 0$ and let $y \in \mathbb{V}^n$ be $\sum_{i=1}^n y_i e_i$, where $y_n \neq 0$, and for all $i < n$, let $y_i = x_i - \lambda_i y_n$. Clearly $x \neq y$, and we see that

$$\begin{aligned} f(x) &= \sum_{i=1}^{n-1} x_i f(e_i) = \sum_{i=1}^{n-1} (y_i + \lambda_i y_n) f(e_i) \\ &= \sum_{i=1}^{n-1} y_i f(e_i) + y_n \sum_{i=1}^{n-1} \lambda_i f(e_i) = \sum_{i=1}^{n-1} y_i f(e_i) + y_n f(e_n) = f(y), \end{aligned}$$

showing that f is non-invertible.

Now suppose that $\{f(e_i)\}_{i=1}^n$ is linearly independent. We must show that for any $x, y \in \mathbb{V}^n$, if $f(x) = f(y)$, then $x = y$. This follows immediately from the equation

$$0 = f(x) - f(y) = \sum_{i=1}^{n} (x_i - y_i) f(e_i),$$

because we must have for all integers $i \in [1, n]$, $x_i = y_i$ on the grounds that $\{f(e_i)\}_{i=1}^n$ is a linearly independent set.

Having now established the conditions upon which f^{-1} exists, let's quickly prove the uniqueness of f^{-1}. Suppose the functions g and h are distinct inverses of f. By distinct, this must mean that there exists $y \in f(\mathbb{V}^n)$ such that $g(y) \neq h(y)$. Let $x \in \mathbb{V}^n$ be such that $f(x) = y$. We then have $x = g(f(x)) = g(y) \neq h(y) = h(f(x)) = x$, which is a contradiction. Inverses of functions in general are therefore unique.

3 Enter Outermorphic Functions

To this point we have satisfied the basic questions of existance and uniqueness for inverses of linear functions. Given a linear function f, what we would now hope to be able to do is find f^{-1}. This is where geometric algebra comes into play.

Definition 3.1. *A function f is called an outermorphism if it preserves the outer product. That is, for any two vectors $x, y \in \mathbb{V}^n$, we have $f(x \wedge y) = f(x) \wedge f(y)$.*

Every linear function f can be extended to an outermorphism which we denote by $\underline{f}$. If f does not already possess the defining characteristic of an outermorphism, (i.e., $f \neq \underline{f}$), then it can be extended to an outermorphism by simply defining

$$\underline{f}(B) = \bigwedge_{i=1}^{s} f(b_i),$$

where B is the s-blade $\bigwedge_{i=1}^{s} b_i$. Clearly this extension of f preserves the original function and is therefore both linear and outermorphic. Furthermore, if we find $\underline{f}^{-1}$, then we have also found f^{-1}.

Taking zero as being a blade of any grade, it is clear that outermorphic functions are grade preserving. From what we already know so far, it is clear that if a linear outermorphism also preserves non-zero blades as non-zero blades, then it must be invertible.

4 The Adjoint Outermorphism

Given a linear function f, of particular interest to us will be the function $\bar{f}$ implicity defined as

$$x \cdot f(y) = \bar{f}(x) \cdot y,$$

where $x, y \in \mathbb{V}^n$. I believe we call this the adjoint of f. An explicit definition of $\bar{f}$ is therefore given by

$$\bar{f}(x) = \sum_{i=1}^{n} (\bar{f}(x) \cdot e_i) e_i = \sum_{i=1}^{n} (x \cdot f(e_i)) e_i,$$

which deserves careful comparison to equation (1). Can we relate the concept of reciprical basis here?

As one can check, $\bar{f}$ is clearly a linear function. The extension of this function to an outermorphism, which we'll also denote by $\bar{f}$, is what I believe we call the adjoint outermorphism of f. An interesting reformulation of the $\bar{f}$ is given by

$$\begin{aligned}
\bar{f}(x) &= \sum_{i=1}^{n} e_i (x \cdot f(e_i)) \\
&= \sum_{i=1}^{n} e_i (x \cdot \partial_i f(y)) \\
&= \sum_{i=1}^{n} e_i \partial_i (x \cdot f(y)) \\
&= \nabla_y (x \cdot f(y)),
\end{aligned}$$

where here we must be careful to note that ∇_y denotes the gradient of its operand with respect to y, not the directional derivative of its operand in the direction y.

We will now proceed to uncover an important property of the outermorphic extenstion of f and its adjoint outermorphism. Letting B be the s-blade

$\bigwedge_{i=1}^{s} b_i$, we see that

$$\begin{aligned} x \cdot \underline{f}(B) &= -\sum_{i=1}^{s}(-1)^i (x \cdot \underline{f}(b_i)) \bigwedge_{j=1, j\neq i}^{s} \underline{f}(b_j) \\ &= -\sum_{i=1}^{s}(-1)^i (\bar{f}(x) \cdot b_i) \bigwedge_{j=1, j\neq i}^{s} \underline{f}(b_j) \\ &= \underline{f}\left(-\sum_{i=1}^{s}(-1)^i (\bar{f}(x) \cdot b_i) \bigwedge_{j=1, j\neq i}^{s} b_j \right) \\ &= \underline{f}(\bar{f}(x) \cdot B). \end{aligned}$$

Now let A be the r-blade $\bigwedge_{i=1}^{r} a_i$, where $r \leq s$, and recall that

$$A \cdot B = \left(\bigwedge_{i=1}^{r-1} a_i \right) \cdot (a_r \cdot B). \tag{2}$$

By repeated application of this identity, we get

$$A \cdot B = a_1 \cdot \cdots \cdot a_r \cdot B,$$

where here it is understood that the inner product is meant to be right-to-left associative. Similarly, we may write

$$A \cdot \underline{f}(B) = \left(\bigwedge_{i=1}^{r-1} \bar{f}(a_i) \right) \cdot \underline{f}(\bar{f}(a_r) \cdot B),$$

and again repeatedly apply the identity (2) recursively to obtain

$$A \cdot \underline{f}(B) = \underline{f}(\bar{f}(a_1) \cdot \cdots \cdot \bar{f}(a_r) \cdot B) = \underline{f}(\bar{f}(A) \cdot B),$$

where again the inner product here is understood to be right-to-left associative.

5 The Inverse Outermorphism

We're now ready to find the inverse of an outermorphism. We begin by making the observation that

$$\lambda A I = A \underline{f}(I) = \underline{f}(\bar{f}(A) I),$$

where here $\lambda \in \mathbb{R}$ and I is the unit psuedo-scalar of $\mathbb{G}(\mathbb{V}^n)$. In fact, we define $\det f = \lambda = \underline{f}(I)I^{-1}$. It follows that

$$(\det f)\underline{f}^{-1}(AI) = \bar{f}(A)I.$$

Solving $\underline{f}^{-1}(A)$, we get

$$\underline{f}^{-1}(A) = \frac{\bar{f}(AI^{-1})I}{\det f} = \frac{\bar{f}(AI)I}{\det f}(-1)^{n(n-1)/2} = \frac{\bar{f}(AI)I^{-1}}{\det f}.$$

When $A \in \mathbb{V}^n$, this gives us the inverse function f^{-1}! That is, for all $x \in \mathbb{V}^n$, we have

$$f^{-1}(x) = \frac{\bar{f}(xI)I^{-1}}{\det f}. \tag{3}$$

It can be shown that f^{-1} is outermorphic. A proof of this is given in [1] which siffices me, so I will not repeat it here. The implication of this is that the inverse of the outermorphic extension of f is the outermorphic extension of the inverse of f.

6 A Small Comparison To Matrix Algebra

It is common to represent an invertible linear transformation as a set of basis vectors $\{a_i\}_{i=1}^n$ of $\mathbb{V}^n$. Matrices, for example, represent linear transformations this way, the basis vectors taking up the rows or columns of the matrix, and matrix multiplication is, by definition, a change of basis transformation. Using what we now know about linear algebra, we see that the set of basis vectors $\{f^{-1}(e_i)\}_{i=1}^n$ must represent the inverse change of basis transformation where $f(x) = \sum_{i=1}^n (x \cdot e_i)a_i$. Using equation (3), we can find a formula for $f^{-1}(e_i)$. Attempting to do so, we get

$$f^{-1}(e_i) = \frac{-(-1)^i}{\det f}\left(\bigwedge_{j=1,j\neq i}^{n} \sum_{k=1}^{n}(e_j \cdot a_k)e_k\right) I^{-1}.$$

Admittedly, this isn't as reduced as I would like to see it, but it is interesting to me none-the-less. Geometric algebra has given us a formula for the inverse basis in terms of the original basis.

7 Eigen Blades

References

[1] David Hestenes. The design of linear algebra and geometry. *Acta Applicandae Mathematicae*, 1991.

[2] Alan Macdonald. A survey of geometric algebra and geometric calculus, 2012.