Chapter 10

Tuesday, July 26, 2011

10.1 Subspaces and Linear Transformations

To understand the deep theory behind how Linear Transformations act on the vector spaces $\mathbb{R}^n$, we must have the notion of a subspace, a subset of $\mathbb{R}^n$ that is closed under scalar multiplication and addition.

10.1.1 Subspaces of $\mathbb{R}^n$

A *subspace* V of $\mathbb{R}^n$ is a subset of $\mathbb{R}^n$ in which any two vectors $\mathbf{x}, \mathbf{y} \in V$ and $\alpha \in \mathbb{R}$, the sum $\mathbf{x} + \mathbf{y}$ is also in V and $\alpha \cdot \mathbf{x}$ is also in V. Thus, a subspace of a vector space is a vector space in its own right. It is closed under scalar multiplication and addition.

10.1.2 Important Examples of Subspaces of $\mathbb{R}^n$

Example - The zero subspace.

The subset consisting of the single element $(0, 0, \ldots, 0)$ (called the zero-vector) is a subspace of $\mathbb{R}^n$. Notice that

$$\alpha \cdot (0, 0, \ldots, 0) = (\alpha \cdot 0, \alpha \cdot 0, \ldots, \alpha \cdot 0) = (0, 0, \ldots, 0),$$

thus it is closed under scalar multiplication. Further, since there is only one vector in this space, it is clear that it is closed under addition:

$$(0, 0, \ldots, 0) + (0, 0, \ldots, 0) = (0, 0, \ldots, 0).$$

Thus, both requirements to be a subspace are satisfied and the zero subset is indeed a subspace.

Example - Subspaces of $\mathbb{R}^3$.

Consider the subset V of all vectors of the form $(x_1, x_2, 0)$ with $x_i \in \mathbb{R}$; thus, this is the subset with the third entry zero. Given any $\alpha \in \mathbb{R}$,

$$\alpha \cdot (x_1, x_2, 0) = (\alpha \cdot x_1, \alpha \cdot x_2, \alpha \cdot 0) = (\alpha \cdot x_1, \alpha \cdot x_2, 0);$$

thus, for any $\mathbf{x} \in V$, $\alpha \cdot \mathbf{x} \in V$ and V is closed under scalar multiplication. Further, if $\mathbf{x} = (x_1, x_2, 0) \in V$ and $\mathbf{y} = (y_1, y_2, 0) \in V$, then

$$\mathbf{x} + \mathbf{y} = (x_1, x_2, 0) + (y_1, y_2, 0) = (x_1 + y_1, x_2 + y_2, 0).$$

Since the third term of the sum is zero, this vector sum is still in V, and thus this subset is closed under vector addition. Therefore, both the requirements for being a subspace are satisfied and V is indeed a subspace.

Notice that this construction will also work if the first or second (or both) entries are required to be 0.

Example - The Kernel of the Linear Transformation L_A.

Consider an $n \times m$ matrix A; of course, this gives rise to a linear transformation

$$L_A : \mathbb{R}^m \to \mathbb{R}^n$$

given by left-multiplication $\mathbf{x} \mapsto A \cdot \mathbf{x}$. Consider the subset of $\mathbb{R}^m$ given by

$$\mathrm{Ker}(A) = \{\mathbf{x} \in \mathbb{R}^m | A \cdot \mathbf{x} = \mathbf{0}\},$$

where $\mathbf{0}$ is the zero element $(0, 0, \ldots, 0) \in \mathbb{R}^n$. This subset is called the *kernel* of the map L_A (also known as the kernel of A). We will prove that this subset of $\mathbb{R}^m$ is a subspace. First, we prove that the kernel of L_A is closed under scalar multiplication. If $\mathbf{x} \in \mathrm{Ker}(A)$ and $\alpha \in \mathbb{R}$ is a scalar, then we must show that $\alpha \cdot \mathbf{x} \in \mathrm{Ker}(A)$. Since $\mathbf{x} \in \mathrm{Ker}(A)$, then

$$L_A(\mathbf{x}) = A \cdot \mathbf{x} = \mathbf{0}.$$

We see that

$$L_A(\alpha \cdot \mathbf{x}) = A(\alpha \cdot \mathbf{x}) = \alpha \cdot A\mathbf{x} = \alpha\mathbf{0} = \mathbf{0}.$$

Thus, $\alpha\mathbf{x}$ is also in $\mathrm{Ker}(A)$. Further, if $\mathbf{x}, \mathbf{y} \in \mathrm{Ker}(A)$, then $A\mathbf{x} = A\mathbf{y} = \mathbf{0}$. We see that their sum $\mathbf{x} + \mathbf{y}$ is also in $\mathrm{Ker}(A)$:

$$L_A(\mathbf{x} + \mathbf{y}) = A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$

Since both conditions are met, $\mathrm{Ker}(A)$ is a subspace of $\mathbb{R}^m$.

Example - The Image of the Linear Transformation L_A

Given the same $n \times m$-matrix A, we can also consider the subset of $\mathbb{R}^n$ which is the image of the map L_A. This subset is called the *column space of* A and is given by

$$\mathrm{Col}(A) = \{\mathbf{y} \in \mathbb{R}^n | A \cdot \mathbf{x} = \mathbf{y} \text{ for some } \mathbf{x} \in \mathbb{R}^m\}.$$

Thus, $\mathrm{Col}(A)$ are all the vectors in $\mathbb{R}^n$ that are mapped onto by some vector in $\mathbb{R}^m$. To show that this subset is closed under scalar multiplication, we note that if $\mathbf{y} \in \mathbb{R}^n$ such that there exists some $\mathbf{x} \in \mathbb{R}^n$ such that $A \cdot \mathbf{x} = \mathbf{y}$. So, to show that $\alpha \cdot \mathbf{y}$ is in $\mathrm{Col}(A)$, we must find a pre-image that maps to $\alpha \cdot \mathbf{y}$. Consider $\alpha \cdot \mathbf{x}$. If we plug this into our map L_A we obtain

$$L_A(\alpha \cdot \mathbf{x}) = A(\alpha \cdot \mathbf{x}) = \alpha \cdot A\mathbf{x} = \alpha \cdot \mathbf{y}.$$

Thus, $\alpha \cdot \mathbf{y}$ is also in the image and thus $\alpha \cdot \mathbf{y} \in \mathrm{Col}(A)$. To show that $\mathrm{Col}(A)$ is closed under addition, take $\mathbf{y}, \mathbf{z} \in \mathrm{Col}(A)$; thus, there are pre-images $\mathbf{x}, \mathbf{w} \in \mathbb{R}^m$ such that $A \cdot \mathbf{x} = \mathbf{y}$ and $A \cdot \mathbf{w} = \mathbf{z}$. To show that $\mathbf{y} + \mathbf{z} \in \mathrm{Col}(A)$, we consider the vector $\mathbf{x} + \mathbf{w} \in \mathbb{R}^n$:

$$L_A(\mathbf{x} + \mathbf{w}) = A(\mathbf{x} + \mathbf{w}) = A\mathbf{x} + A\mathbf{w} = \mathbf{y} + \mathbf{z}.$$

Thus, $\mathbf{y} + \mathbf{z}$ is in the image $\mathrm{Col}(A)$. Thus, the two requirements are satisfied and $\mathrm{Col}(A)$ is a subspace of $\mathbb{R}^n$.

10.2 Vector Spaces

In the previous section, we investigated the notion of a subspace of $\mathbb{R}^n$; these are precisely the subsets of $\mathbb{R}^n$ that are closed under matrix addition and scalar multiplication. These two properties ensure that we can simply focus on these subsets; these spaces, known as *vector spaces*, act very much like $\mathbb{R}^k$ (for perhaps a different value of k).

10.2.1 Definitions and Examples

A key property of an n-vector $\mathbf{v}$ in $\mathbb{R}^n$ is that if we multiply it by a real number $\alpha \in \mathbb{R}$ or add it another n-vector $\mathbf{w} \in \mathbb{R}^n$, then we obtain other n-vectors $\alpha\mathbf{v}$ and $\mathbf{v} + \mathbf{w}$. Thus, $\mathbb{R}^n$ has the property that it is ***closed under scalar multiplication*** and ***closed under vector addition***. Of course, we could say the same thing about any *subspace* of $\mathbb{R}^n$ (like the *kernel* or *column space* of a linear transformation). If we look beyond sets of n-vectors to general sets with a notion of scalar multiplication and vector addition which is closed under these two operations, we obtain the definition of a ***vector space***.

To solidify the concept of a vector space, let us investigate some important examples.

Example - Polynomials of degree n. The set $\mathbb{P}_n$ consists of all polynomials $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ of degree at most n. Vector addition is given by regular polynomial addition, and scalar multiplication is regular multiplication of a polynomial by a real number. Notice that if we take two polynomials of degree at most n, then their degree is also at most n. Furthermore, when we scale a polynomial of degree n by some real number, it also remains of degree at most n. Thus, $\mathbb{P}_n$ is a vector space.

Example - All polynomials. A larger space $\mathbb{P}$ is the set of all polynomials (with no restriction on the degree). Using the same operations as with $\mathbb{P}_n$, $\mathbb{P}$ also becomes a vector space since the sum or scalar multiple of a polynomial is once again a polynomial.

Example - Continuous functions on $\mathbb{R}$. Let $C(\mathbb{R})$ be the set of all continuous functions on $\mathbb{R}$, with addition and scalar multiplication being function addition and scalar multiplication. Since the sum of two continuous functions is continuous and the scalar multiple of a continuous function is continuous, $C(\mathbb{R})$ is a vector space that includes both $\mathbb{P}$ and $\mathbb{P}_n$ as subsets (in fact, subspaces). We may generalize this set to the $C^k(\mathbb{R})$, the set of all functions on $\mathbb{R}$ which are k times differentiable.

Example - Convergent Sequences. Consider the set of all sequences that converge to some real number (this limit may be different for the individual sequences). We may add the sequences x_n and y_n by having its k-th term be the sum of $x_k + y_k$. Furthermore, we may scale a sequence x_n by $\alpha \in \mathbb{R}$ to produce a new sequence where the k-th term is αx_k. Since the sum of two convergent sequences is convergent and the scalar multiple of a convergent sequence is convergent, we see that this set does indeed form a vector space.

10.3 Bases and Dimensions of Vector Spaces

10.3.1 Bases, Span, and Linear Independence

Notice that if we have the 3-vector $\mathbf{x} = (4, -\pi, \sqrt{2})$, we can write it as a sum of three scaled vectors:

$$\mathbf{x} = (4, -\pi, \sqrt{2}) = 4 \cdot (1, 0, 0) - \pi \cdot (0, 1, 0) + \sqrt{2} \cdot (0, 0, 1).$$

Also, notice that this is actually the unique way to write this particular vector. These are examples of the notion of *span* and *linear independence*. Any subset of $\mathbb{R}^3$ that have these two properties (in this example, it's the vectors $\mathbf{e_1} = (1, 0, 0), \mathbf{e_2} = (0, 1, 0)$, and $\mathbf{e_3} = (0, 0, 1)$) is called a *basis* for $\mathbb{R}^3$.

Formally, given any set of n vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ in a vector space V, we say that this set ***spans*** the vector space V if every vector $\mathbf{x} \in V$ can be written as a ***linear combination*** of the $\mathbf{v}_i$'s. That is, there are scalars $a_1, \ldots, a_n \in \mathbb{R}$ such that

$$\mathbf{x} = a_1\mathbf{v_1} + a_2\mathbf{v_2} + \cdots + a_n\mathbf{v_n}.$$

Thus, a subset of vectors *spans* a vector space if there are enough of them so that any other vector in V is just a linear combination of them.

We say that a subset of vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ is a ***linearly independent*** subset of a vector space V if whenever we can write the zero vector as a linear combination

$$\mathbf{0} = a_1\mathbf{v_1} + a_2\mathbf{v_2} + \cdots + a_n\mathbf{v_n},$$

then we must have that

$$a_1 = a_2 = \cdots = a_n = 0.$$

Thus, a set of vectors is linearly independent if the only way to write the zero vector is the trivial way (with all the scalars being zero). Intuitively, this means that the subset of vectors $\mathbf{v_1}, \ldots, \mathbf{v_n}$ is small enough that we can write each vector uniquely.

If we are lucky enough to a subset of vectors $v_1, v_2, \ldots, v_n$ that spans V and is linearly independent, then this set is known as a ***basis***. Thus, a basis is big enough to be able to write any vector in V as a linear combination of them, but small enough that there is only one way to do it. It is not difficult to prove that one natural choice of basis for $\mathbb{R}^n$ is

$$\mathbf{e_1} = (1, 0, 0, \ldots, 0), \mathbf{e_2} = (0, 1, 0, \ldots, 0), \ldots \mathbf{e_n} = (0, 0, 0, \ldots, 1);$$

actually, this is referred to as the *canonical basis* for $\mathbb{R}^n$. Of course, there are many other *bases* for $\mathbb{R}^n$. We will shortly give a criterion for deciding if a collection of vectors is a basis for $\mathbb{R}^n$.

10.3.2 Dimension of a Vector Space

Given an arbitrary vector space V with a finite basis $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$, we say that the dimension of V is n (the number of vectors in the basis). This concept is well-defined because of the following theorem.

Theorem. Let V be a vector space with a basis of n vectors. Then, any other basis for V will also have n vectors.

For example, $\mathbf{e_1}, \mathbf{e_2}, \ldots \mathbf{e_n}$ is a basis for $\mathbb{R}^n$ and thus its dimension is n (as expected). Note that the above theorem tells us that any other basis for $\mathbb{R}^n$ will also have n vectors.

Another somewhat obvious theorems is that subspaces of a vector space have smaller dimensions.

Theorem. If V is a subspace of a finite-dimensional vector space W, then

$$\dim(V) \leq \dim(W).$$