3 - Vector Spaces

Vectors in $\mathbf{R}^2$ and $\mathbf{R}^3$ are essentially matrices. They can be viewed either as column vectors (matrices of size 2×1 and 3×1, respectively) or row vectors (1×2 and 1×3 matrices). The addition and scalar multiplication defined on real vectors are precisely the corresponding operations on matrices. Thus the matrix definitions provide a structure on $\mathbf{R}^n$, real n-space or n-dimensional space, whose vectors can be thought of as either row or column matrices with n elements for any n = 1, 2, 3, ... This is the central idea in linear algebra: the notion of vector space which we now define.

__Definition__

Let V be a set and K be either the real, **R**, or the complex numbers, **C**. We call V a ***vector space*** (or ***linear space***) over the field of scalars K provided that there are two operations, vector addition and scalar multiplication, such that for any vectors **u, v,** and **w** in V and for any scalars " and $ in K:

1. (*Closure)* **v** + **w** and "**v** are in V,
2. (*Associativity*) **u** + (**v** + **w**) = (**u** + **v**) + **w**
3. (*Commutativity*) **v** + **w** = **w** + **v**
4. There exists some element **0** in V with **v** + **0** = **v** (**0** which is independent of **v** is called the *identity* of V.)
5. (*Inverses*) There exists an element **-v** in V with **v** + **-v** = **0**
6. (*Distribution*) "(**v** + **w**) = "**v** + "**w**
7. (" + $) **v** = "**v** + $**v**
8. "($**v**) = ("$)**v**
9. 1**v** = **v**

The process of abstracting the primary structure of a system into a general definition is extremely common in mathematics and serves several purposes. Proving results about the generalized object yields corresponding information about any specific case which satisfies the axioms of the system. In addition it is sometimes actually easier to see (and prove) what is going on in the general object which does not have extra features which might otherwise obscure its essence. The examples below are to testify to the wide range of vector spaces.

Examples

1. For any positive integers m and n, $M_{m\times n}(\mathbf{R})$, the set of m by n matrices with real entries, is a vector space over $\mathbf{R}$ with componentwise addition and scalar multiplication.
2. We use $M_{m\times n}(\mathbf{C})$ to denote the set of m by n matrices whose entries are complex numbers. This forms a vector space over either the reals or the complexes which is to say, we may consider the scalars here to come from either $\mathbf{R}$ or $\mathbf{C}$.
3. $\mathbf{R}^n$, as mentioned above, is a vector space over the reals.
4. $\mathbf{C}^n$ considered as either $M_{1\times n}(\mathbf{C})$ or $M_{n\times 1}(\mathbf{C})$ is a vector space with its field of scalars being either $\mathbf{R}$ or $\mathbf{C}$.
5. The set of all real valued functions, F, on $\mathbf{R}$ with the usual function addition and scalar multiplication is a vector space over $\mathbf{R}$.
6. The set of all polynomials with coefficients in $\mathbf{R}$ and having degree less than or equal to n, denoted P_n, is a vector space over $\mathbf{R}$.

Theorem

Suppose that $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are elements of some vector space. Then

1. If $\mathbf{u} + \mathbf{v} = \mathbf{w} + \mathbf{v}$, then $\mathbf{u} = \mathbf{w}$. (The cancellation property holds.)
2. The inverse of $\mathbf{v}$, $-\mathbf{v}$, is unique.
3. The identity element $\mathbf{0}$ is unique.
4. $0\mathbf{v} = \mathbf{0}$
5. $(-1)\mathbf{v} = -\mathbf{v}$

Proof

Suppose that $\mathbf{u} + \mathbf{v} = \mathbf{w} + \mathbf{v}$. Adding $-\mathbf{v}$ to both sides of the equation yields,

$$(\mathbf{u} + \mathbf{v}) + -\mathbf{v} = (\mathbf{w} + \mathbf{v}) + -\mathbf{v}$$

So $$\mathbf{u} + (\mathbf{v} + -\mathbf{v}) = \mathbf{w} + (\mathbf{v} + -\mathbf{v})$$

Thus, $$\mathbf{u} + \mathbf{0} = \mathbf{w} + \mathbf{0}$$

which means we get $\mathbf{u} = \mathbf{w}$ which proves property 1.

Suppose that vectors $\mathbf{u}$ and $\mathbf{w}$ are both inverses of $\mathbf{v}$. Then $\mathbf{u} + \mathbf{v} = \mathbf{0} = \mathbf{w} + \mathbf{v}$. Applying the cancellation we just proved gives us $\mathbf{u} = \mathbf{w}$, so inverses are unique. Even more readily, if $\mathbf{0}$ and $\mathbf{0}'$ both will serve as the identity, then $\mathbf{0} = \mathbf{0} + \mathbf{0}' = \mathbf{0}'$. Thus a vector space has only one identity. From this it follows that, since, $\mathbf{v} = (1 + 0)\mathbf{v} = 1\mathbf{v} + 0\mathbf{v} = \mathbf{v} + 0\mathbf{v}$ implies that $0\mathbf{v}$ is an identity, $0\mathbf{v} = \mathbf{0}$. Finally, $0\mathbf{v} = (1 + -1)\mathbf{v} = 1\mathbf{v} + (-1)\mathbf{v} = \mathbf{v} + (-1)\mathbf{v}$ and so, by the uniqueness of inverses, $-\mathbf{v} = (-1)\mathbf{v}$.

Numerous important examples of vector spaces are subsets of other vector spaces.

<u>Definition</u>

Let S be a subset of a vector space V over K. S is a ***subspace*** of V if S is itself a vector space over K under the addition and scalar multiplication of V.

<u>Theorem</u>

Suppose that S is a nonempty subset of V, a vector space over K. The following are equivalent:

1. S is a subspace of V.
2. S is closed under vector addition and scalar multiplication.
3. S is closed under the process of taking linear combinations, i.e., if $\mathbf{v}$ and $\mathbf{w}$ are in S and α and β are in K, then $\alpha\mathbf{v} + \beta\mathbf{w}$ is in S.

<u>Proof</u>

Suppose that S is a subspace of V. Then S is a vector space and so is closed with respect to addition and scalar multiplication. Thus, 1 implies 2. Also, if α and β are scalars and $\mathbf{v}$ and $\mathbf{w}$ are vectors in S, then $\alpha\mathbf{v}$ and $\beta\mathbf{w}$ are in S, so the linear combination $\alpha\mathbf{v} + \beta\mathbf{w}$ is also in S. Hence, 1 implies 3.

Now since S is (simply) a subset of V, it satisfies properties 2, 3, and 6 - 9. For example, suppose $\mathbf{v}$ and $\mathbf{w}$ are in S. Then $\mathbf{v}$ and $\mathbf{w}$ are also in V so $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$, i.e., S satisfies property 3, commutativity of addition. If we assume that S is closed under vector addition and scalar multiplication (condition 2 above), then this is precisely property 1 in the definition of vector space. Also since S is not empty there is some $\mathbf{v}$ in S. Closure under scalar multiplication then implies that $0\mathbf{v} = \mathbf{0}$ is in S. Thus, S includes the identity as required by property 4. Then if $\mathbf{v}$ is any vector in S, $(-1)\mathbf{v} = -\mathbf{v}$ is also in S. So S possesses inverses and so satisfies property 5.

Hence we have shown that S is a vector space and, consequently, condition 2 implies 1.

Finally, if S satisfies condition 3 and so is closed under linear combinations, then assume that **v** and **w** are in S and (is any scalar. Taking " = 1 = $, we see that "**v** + $**w** = **v** + **w** is in S and S is closed under addition. Letting " = 0, the fact that "**v** + (**w** = 0**v** + (**w** = **0** + (**w** = (**w** shows that S is closed under scalar multiplication. Thus, 3 implies 2 and the theorem is established.

Examples

1. Consider S = {(0, y, z): y and z are any real numbers}. S is a subset of $\mathbf{R}^3$**.** S is also a subspace since addition and scalar multiplication is by components so the 0 in the first component will be preserved and we get that S is closed under both operations. Note that S is essentially $\mathbf{R}^2$.
2. Letting $D_{m\times n}$ be the set of all m×n diagonal matrices it is easy to see that $D_{m\times n}$ is a subspace of $M_{m\times n}$.
3. The smallest subspace of any vector space is {**0**}, the set consisting solely of the zero vector.
4. For any n the set of lower triangular n×n matrices is a subspace of $M_{n\times n} = M_n$.
5. The set of all n×n symmetric matrices is a subspace of M_n. What properties of the transpose are used to show this?
6. Let S = { f : f is in F and f(2) = 0}. This is a subspace of F since if f and g are in S and c is in **R**, f(2) = 0 = g(2) so (f + g) (2) = f(2) + g(2) = 0 (so f + g is in S) and (cf)(2) = c0 = 0 (so cf is in S.)
7. Define a subset of $\mathbf{R}^3$ by setting S = { (x, y, z) ; 5x - y + 7z = 0}. Suppose that **v** = (x, y, z) and **w** = (s, t, u) are in S. Then, by definition of S, 5x - y + 7z = 0 = 5s - t + 7u so that 5(x + s) - (y + t) + 7(z + u) = 0. Thus **v** + **w** is in S. Additionally, for any scalar a, 5(ax) - ay + 7az = a(5x - y + 7z) = 0, so S is closed under scalar multiplication. S is, therefore, a subspace of $\mathbf{R}^3$.

Problems

1. Is the set of all upper triangular matrices a vector space?

2. Is the set of all unit lower triangular matrices of size n×n a subspace of M_n?

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n$ be vectors in some vector space *V*. Define the ***subspace generated by these vectors*** by S = $<\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n> = \{"_1\mathbf{v}_1 + "_2\mathbf{v}_2 + \ldots + "_n\mathbf{v}_n$: each $"_k$ is in K}. Then ***S*** is nothing more that all possible linear combinations of these vectors. To see that ***S*** is indeed a subspace of ***V*** use the linear combination characterization of subspace (condition 3) along with the commutative and associative properties and the “fake” distributive and associative properties (properties 7 and 8 in the definition of vector space.) This is the *smallest* subspace of ***V*** containing the original vectors.

Example

Geometrically in $\mathbf{R}^2$ and $\mathbf{R}^3$ (and the same is true in $\mathbf{R}^n$ for any n = 2, 3, 4, ...) we have seen that scalar multiples of a given vector **v** result in a vector in the same or opposite direction as **v**. The set of all such vectors, the subspace generated by **v**, <**v**>, is a line through the origin provided **v** ... **0.** (What if **v** = **0**?) Assuming that **v** and **w** are vectors in $\mathbf{R}^n$ that are not parallel (and so not scalar multiples of each other), consider P = <**v**, **w**>. The subspace P includes the lines through the origin “containing” **v** and **w** as well as all the linear combinations $"\mathbf{v} + \$\mathbf{w}$. Each vector $"\mathbf{v}$ or $\$\mathbf{w}$ corresponds to a point on one of the lines through the origin and **v** or **w**, respectively. By the parallelogram rule for visualizing vector addition, the vector sum $"\mathbf{v} + \$\mathbf{w}$ is the diagonal (or vertex) of the parallelogram whose other vertices are the origin, $"\mathbf{v}$, and $\$\mathbf{w}$. This a point on the plane determined by the origin and the points **v** and **w**. Thus, *P* fills up this plane and is a plane containing the origin. In particular, if $v = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ & $w = \begin{pmatrix} 2 \\ 7 \\ 0 \end{pmatrix}$, then the associated plane is given by $P = \{ (x, y z) \mid 7x - 2y + 7z = 0 \}$ which is precisely the plane containing the three distinct points (0,0,0) , (1,0,-1), and (2,7,0).

There are several important spaces associated with any $m \times n$ matrix A. The ***column space** of*

A, denoted $C(A) = C_A$, is the subspace of $\mathbb{R}^m$ generated by the n columns of A. (Why is $C(A) = C_A$ a subspace of $\mathbb{R}^m$?) That is, if $\{c_1, c_2, c_3, \ldots, c_n\}$ are the n columns of A, then

$$C(A) = C_A = \{ \alpha_1 c_1 + \ldots + \alpha_n c_n \mid \alpha_1, \ldots, \alpha_n \in \mathbb{R} \}.$$

We have seen that the matrix equation $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$ has a solution if and only if **b** is a linear combination of the columns of $\boldsymbol{A}$ (see the chapter 2 - Solving Linear Equations.) That is, $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$ is consistent if and only

$$x_1\,\boldsymbol{c_1} + x_2\,\boldsymbol{c_2} + \ldots + x_n\,\boldsymbol{c_n} = \boldsymbol{b}$$

for some $x_1, \ldots, x_n \in \mathbb{R}$ where $\{c_1, c_2, c_3, \ldots, c_n\}$ are the n columns of A.

In the present terminology we can state this fact as: A**x** = **b** has a solution if and only if **b** is in C(A).

Problems

Set A = $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -2 \end{pmatrix}$, B = $\begin{pmatrix} 1 & -3 \\ -5 & 15 \\ 4 & -12 \end{pmatrix}$, and C = $\begin{pmatrix} 1 & 6 & 2 & 0 \\ 0 & 0 & 7 & 7 \\ -1 & -6 & 0 & 2 \end{pmatrix}$.

Determine the column spaces of A, B, and C.

The ***null space** of A* is defined by N(A) = {**x** : A**x** = **0**}. It consists precisely of those vectors which when multiplied by A yield the zero vector. If A is an m×n matrix, then N(A) is a subset of $\mathbf{R}^n$, that is, it consists of vectors which can be multiplied on the left by A.

Example

Suppose that $A = \begin{pmatrix} 1 & 1 & 3 & 1 \\ 2 & 1 & 5 & 4 \\ 1 & 2 & 4 & -1 \end{pmatrix}$. Then the vectors $\begin{pmatrix} -17 \\ 9 \\ 1 \\ 5 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 8 \\ -4 \\ 2 \end{pmatrix}$ belong to $N(A)$

while $\begin{pmatrix} -17 \\ 9 \\ 1 \\ 0 \end{pmatrix}$ fails to belong to the null space. (How does one verify that statement?)

In fact, N(A) is a *subspace* of $\mathbf{R}^n$ as we now show. Since **0** is in N(A) (the homogeneous system always has the trivial solution) N(A) is nonempty. Suppose that **x** and **y** are in N(A) and that " is a real number. Then, by definition, A**x** = **0** = A**y**. Therefore, A(**x** + **y**) = A**x** + A**y** = **0** + **0** = **0.** Thus, **x** + **y** is in N(A), so N(A) is closed under addition. Also the fact that A("**x**) = "(A**x**) = "**0** = **0** tells us that "**x** is in N(A). Hence, N(A) is closed under scalar multiplication. We, therefore, have that the null space of ***A*** is a subspace.

An immediate application of null spaces is as solution sets for systems of linear equations. If ***A*** is the coefficient matrix for a system and the constants of it are all zeroes (so the system is homogeneous), then is corresponds to the matrix equation ***A*** **x** = **0** and solving the system consists of describing the null space of ***A***. When the system is not homogeneous (the matrix equation is ***A*x** = **b** and **b** … **0**), the solution set is not a subspace since **0** is not a solution. Still N(A) is useful in describing the solution set as we shall see later. For now we consider the homogeneous case.

To solve a linear system of equations we have used elimination to simplify the equations or its augmented matrix to a "staircase" form. We make precise the forms that we use in the following

Definition.

A matrix is in ***echelon*** (or ***row echelon***) form if

1. Any rows consisting entirely of zeros are grouped at the bottom of the matrix.
2. The first nonzero element of any row is called a *pivot*. The pivot in any row is located to the right of the pivot in the row directly above it.

A matrix is in ***reduced echelon*** form if it is in echelon form, its pivots are all ones, and any column containing a pivot consists entirely of zeros in its remaining entries.

Example

1. $\boldsymbol{A} = \begin{pmatrix} 1 & -2 & 3 & 9 \\ -1 & 3 & 0 & -4 \\ 2 & -5 & 5 & 17 \end{pmatrix}$ is not in *row echelon form*.

2. $\boldsymbol{U} = \begin{pmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 2 & 4 \end{pmatrix}$ is in *row echelon form*.

3. $\boldsymbol{R} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$ is in both *row echelon form* and *reduced row echelon form*.

The target of the elimination that we did previously was an upper triangular matrix U, the echelon form. The process made use of elementary equation operations though working with the augmented matrix using row operations accomplishes the same end (in fact, recall that we proceeded to frame the process in terms of multiplications by elementary matrices.) To achieve *reduced* echelon form requires one additional operation. Here, for convenience and reference, are all three operations.

Elementary Row Operations

1. Multiplication of any row by a nonzero constant.
2. Interchanging any two rows.
3. Addition of a multiple of one row to another.

Directly corresponding to systems of equations being equivalent if and only if they have the same solution set is the definition that two matrices are *row equivalent* if one can be transformed into the other by a finite sequence of elementary row operations. In our previous work we did not need to use the first operation, multiplication of a row by a nonzero constant, because we were not transforming to reduced form. To change pivot values to ones typically requires this operation. An advantage of reduced echelon form is the fact that any given matrix is row equivalent to a unique matrix in reduced echelon form.

To solve a system in the spirit in which we have worked requires that we first row reduce the augmented matrix to upper triangular or echelon form. Note that in the present discussion of homogeneous systems reducing the coefficient matrix suffices. Scalar multiplication of rows will then make all the pivots into ones and adding multiples of lower rows to the upper rows (what amounts to back substitution) will complete the transformation into reduced echelon form.

Example

Solve the following system expressing the solution in terms of vectors in the null space.

$$\begin{aligned} 2x + 6y + 8z + 6w &= 0 \\ -4x - 12y - 20z - 20w &= 0 \\ 4x + 12y + 14z + 8w &= 0 \end{aligned}$$

The coefficient matrix is

$$\begin{pmatrix} 2 & 6 & 8 & 6 \\ -4 & -12 & -20 & -20 \\ 4 & 12 & 14 & 8 \end{pmatrix}$$

Adding 2 times row one to row two and -2 times row one to row three gives the matrix

$$\begin{pmatrix} 2 & 6 & 8 & 6 \\ 0 & 0 & -4 & -8 \\ 0 & 0 & -2 & -4 \end{pmatrix}$$

Adding -1/2 times row two to row three yields the echelon form

$$\begin{pmatrix} 2 & 6 & 8 & 6 \\ 0 & 0 & -4 & -8 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Multiply row two by -1/4 and row one by ½ to make the pivots one.

$$\begin{pmatrix} 1 & 3 & 4 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Then finally adding - 4 times row two to row one results in the desired reduced echelon form.

$$\begin{pmatrix} 1 & 3 & 0 & -5 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The solution can be read as $x + 3y \quad -5w = 0$

$z + 2w = 0$

or in vector form as

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} -3y+5w \\ y \\ -2w \\ w \end{pmatrix} = y\begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + w\begin{pmatrix} 5 \\ 0 \\ -2 \\ 1 \end{pmatrix}$$

Notice that y and w can be any value. We call these ***free variables*** and point out that they came from columns that do not have pivots. The variables x and z are called ***pivot variables***. They were associated with the columns that had pivots. The solution consists of all linear combinations of the vectors $\begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} and \begin{pmatrix} 5 \\ 0 \\ -2 \\ 1 \end{pmatrix}$. That is, the null space of this coefficient matrix is

$$\left\langle \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 \\ 0 \\ -2 \\ 1 \end{pmatrix} \right\rangle$$