CISC 271 Class 3

Vector Spaces

Text Correspondence: §3.1

Main Concepts:

- *Algebraic properties of a vector space*
- *Size: number of entries in a vector*
- *Dimension: number of basis vectors of a vector space*

Sample Problem, Data Analytics: What space do vectors "live" in?

The concepts of vector spaces and linear transformations should be familiar to the student from prerequisite material, so this is a brief summary that re-phrases the ideas in terms of matrices.

3.1 Vector Space: Properties

Recall, from prerequisite courses, the definition of a vector space $\mathbb{V}$. This space is a set of vectors that meet 8 criteria. These criteria, also called axioms, are usually written in terms of vectors $\vec{u} \in \mathbb{V}$, $\vec{v} \in \mathbb{V}$, $\vec{w} \in \mathbb{V}$, and real numbers a and b. They can be summarized as:

Addition is associative	$\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$
Addition is commutative	$\vec{u} + \vec{v} = \vec{v} + \vec{u}$
Addition has an identity	$\exists\, \vec{0} \in \mathbb{V}$ such that $\forall\, \vec{v} \in \mathbb{V}$, $\vec{v} + \vec{0} = \vec{v}$
Addition has an inverse	$\forall\, \vec{v} \in \mathbb{V}$, $\exists(-\vec{v}) \in \mathbb{V}$ such that $\vec{v} + (-\vec{v}) = \vec{0}$
Addition distributes over multiplication	$(a + b)\vec{v} = a\vec{v} + b\vec{v}$
Multiplication is compatible	$a(b\vec{v}) = (ab)\vec{v}$
Multiplication has an identity	$1\vec{v} = \vec{v}$
Multiplication distributes over addition	$a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$

We will use a typical abbreviation for a *real vector space*, the symbol $\mathbb{R}$. In this course, we will use distinct terminology for concepts that other texts or sources may use differently. Here, a vector $\vec{v} \in \mathbb{R}^m$ will be referred to as having a *size* of m, meaning that it takes m real numbers to specify the vector.

We will reserve the term *dimension* for the number k of basis vectors needed to describe the vector space $\vec{v}_m \in \mathbb{V}^k$ under consideration. Where it is clear that the vector space that is being studied is also $\mathbb{V}^m = \mathbb{R}^m$, the size m and the dimension m may be used interchangeably.

A critically important concept for this course is:

The dimension k of a vector $\vec{v}$ in a vector space $\mathbb{V}$ can be less than the size m of the vector $\vec{v}$

To understand this concept better, consider the plane $\mathbb{R}^2$. A vector $\vec{v} \in \mathbb{R}^2$ is written as

$$\vec{v} = \begin{bmatrix} v_x \\ v_y \end{bmatrix}$$

What we ordinarily think of as the X axis is a 1-dimensional space $\mathbb{X}^1 \subset \mathbb{R}^2$. A vector $\vec{x}$ in this space can be written as

$$\vec{x} = \begin{bmatrix} x \\ 0 \end{bmatrix}$$

3.1.1 Vector Space: Interpretation

We will interpret the axioms of a vector space computationally. Unless otherwise specified, a *vector* is a finite column of real numbers.

We interpret vector addition as the entry-wise addition, so that we compute $\vec{u}+\vec{v}$ by finding the i^{th} entry of each vector and adding these real numbers. We interpret multiplication as the entry-wise product of the real number with the entries of the vector. Although the axioms allow for other alternatives, our interpretation requires that vectors that are summed must have the same number of entries.

The simplest, and perhaps most often used, vector space is the *coordinate space*. For vectors that have m entries, this space is $\mathbb{R}^m$. When in doubt, this is the vector space that we will refer to when we work with vectors.

With our interpretation, there is an easy way to check that a proposed set of vectors is a vector space. For the set $\mathbb{V}$ to be a vector space we must have:

For any $\vec{u} \in \mathbb{V}$, any $\vec{v} \in \mathbb{V}$, any $a \in \mathbb{R}$, and any $b \in \mathbb{R}$,

$$[a\vec{u} + b\vec{v}] \in \mathbb{V}$$

3.1.2 Vector Space: Examples

A careful student might verify that a vector $\vec{x} \in \mathbb{X}^1$ satisfies the above 8 axioms. A student might similarly verify that vectors of the form

$$\vec{y} = \begin{bmatrix} 0 \\ y \end{bmatrix}$$

form a 1-dimensional space in $\mathbb{R}^2$.

Next, consider a vector of the form

$$\vec{w} = \begin{bmatrix} x \\ 2x \end{bmatrix}$$

Such a vector needs 2 real numbers to be specified, so $\vec{w} \in \mathbb{R}^2$. A careful student might verify that vectors of this form satisfy the 8 axioms, so they are a vector space. For the purposes of immediate discussion, we will call this space $\mathbb{W}$.

Because the numbers needed to specify any $\vec{w} \in \mathbb{W}$ have a strict relation, there is only 1 dimension to the space $\mathbb{W}$. Geometrically, the space $\mathbb{W}$ is the line in the plane that passes through the origin and has a slope of 2. We can generalize this notion to observe that, for any real numbers c and d, vectors of the form

$$\vec{w} = \begin{bmatrix} cx \\ dx \end{bmatrix}$$

are in a vector space $\mathbb{W}$ of dimension less than 2. (What is the dimension of the space when $c = d = 0$?)

It is not as easy to visualize the real space $\mathbb{R}^3$, but it is not too difficult and it provides rich examples of vector spaces. As before, we can see that the coordinate axes are 1-D vector spaces; we can write these as

$$\vec{x} = \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} \qquad \vec{y} = \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} \qquad \vec{z} = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} \tag{3.1}$$

For any real numbers x, d, and e, with at least one not equal to zero, a vector in a 1-D space of $\mathbb{R}^3$ can be written as

$$\vec{v} = \begin{bmatrix} cx \\ dx \\ ex \end{bmatrix} \tag{3.2}$$

As before, A careful student might verify that this vector is in a vector space. An additional property, to be verified, is that the sum of two vectors that have the form of Equation 3.2 must also have the same form; this means that the space is closed and is truly of just 1 dimension. Geometrically, the space is a line that passes through the origin.

The ordinary plane is a 2-D space of $\mathbb{R}^3$; a vector in it can be written as

$$\vec{v} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \tag{3.3}$$

This vector $\vec{v}$ can be written in another way: we could also have written it as

$$\vec{v} = \vec{x} + \vec{y} \tag{3.4}$$

where $\vec{x}$ and $\vec{y}$ are of the form in Equation 3.1. These vector *span* the XY plane and act as *basis* vectors. We can easily reason that the XZ and YZ planes are 2-D spaces of $\mathbb{R}^3$, being different linear combinations of the elementary vectors of Equation 3.1.

From these examples, we can better recall definitions from prerequisite material.

A *subspace* is a subset of a vector space that is also a vector space. For example, in the 3-D real vector space $\mathbb{R}^3$, the subspaces of 1 or 2 dimensions can be specified as:

1-D: Any scalar multiple of a specific vector $\vec{v} \in \mathbb{R}^3$, that is, $a\vec{v}$ for any real number a

2-D: Any linear combination of two specific vectors $\vec{v} \in \mathbb{R}^3$ and $\vec{w} \in \mathbb{R}^3$, that is, $a\vec{v} + b\vec{w}$

There are various mathematical definitions of, or criteria for, a subspace. Suppose that two vectors $\vec{v} \in \mathbb{V}$ and $\vec{w} \in \mathbb{V}$ are in a vector space $\mathbb{V}$. They are also in a subspace $\mathbb{W} \subseteq \mathbb{V}$ if and only if:

$\vec{v} \in \mathbb{W}$ and $\vec{w} \in \mathbb{W}$ implies that $(a\vec{v} + b\vec{w}) \in \mathbb{W}$

A *linear span*, or simply a *span*, of a vector space $\mathbb{V}$ is a set of vectors that can be linearly combined to produce any vector in the space $\mathbb{V}$. This can be defined in two complementary ways that amount to the same thing: we can start with the universal vector space and narrow it down to the vector space $\mathbb{V}$, or we can begin with the set of vectors and describe the vector space $\mathbb{V}$ that they span. Here, we will use the second method.

Consider a set of m vectors, each being $\vec{w}_j \in \mathbb{R}^n$. They span the vector space $\mathbb{W}$ of all vectors that are linear combinations of the original set. Another way to say this is that, from the given vectors $\vec{w}_j$, the vector space $\mathbb{W}$ can be constructed by summation. Formally, for any real numbers $x_1, x_2, \cdots, x_m$ the vector space $\mathbb{W}$ can be constructed as the linear combination

$$\left[\sum_{j=1}^{m} x_j \vec{w}_j\right] \in \mathbb{W}$$

3.2 Matrices and Vector Spaces

A matrix represents a linear transformation between vector spaces. A matrix $A \in \mathbb{R}^{m \times n}$ is one way of representing the linear map

$$A : \vec{v}_n \mapsto \vec{w}_m$$

where $\vec{v}_n$ is a vector of size n (a column with n entries) and $\vec{w}_m$ is a vector of size m (a column with m entries). Important qualifications to keep in mind are that, for a given matrix A:

- It is possible that not every vector $\vec{v}_n$ maps to a unique $\vec{w}_m$, that is, there may be distinct vectors $\vec{u}_n \neq \vec{v}_n$ such that $A\vec{u} = M\vec{v}$; and
- It is possible that not every vector $\vec{w}_m$ can be "mapped to", that is, there may not exist a vector $\vec{v}_n$ such that $A\vec{v} = \vec{w}$.

For a matrix, or for the linear transformation that it represents, these are equivalent concepts:

- The column space of A;
- The image of the map $A : \vec{v}_n \mapsto \vec{w}_m$;
- The range of the map $A : \vec{v}_n \mapsto \vec{w}_m$.

and these are also equivalent concepts:

- The pre-image of the map $A : \vec{v}_n \mapsto \vec{w}_m$;
- The domain of the map $A : \vec{v}_n \mapsto \vec{w}_m$.

In this course, the space that we will examine in greatest detail is the column space. This is the space that a matrix A *maps to*, which is the range of the map; it is sometimes referred to as the "valid" image of A (but correctly, it is simply the image of A). It has a dimension that is no greater than n and may be much less. For example, a small rectangular matrix $A \in \mathbb{R}^{3\times 2}$ has a column space that may be of dimension 2, or 1, or even 0 (if A is 0×0) but cannot be of dimension 3.

The columns of a matrix can be considered to be an ordered set of vectors, so it is natural to inquire about their meaning.

The *column space* of a matrix $A \in \mathbb{R}^{m\times n}$ is the set of all vectors $\vec{w} \in \mathbb{R}^m$ that are linear combinations of the column vectors. This is not necessarily a minimal set; for example, the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$$

has two column vectors in its span but, because the second column is a scalar multiple (2) of the first column, the space is 1-D.

A *column basis* is a minimal set of vectors than span the column space. By "minimal" we mean that each vector in the basis is linearly independent of the other vectors. For example, a randomly generated small rectangular matrix $A \in \mathbb{R}^{3\times 2}$ will almost certainly have 2 vectors in its column span; this happens if and only if the column space is 2-D. In such a case, the column space is of size 3, because each vector has 3 entries, but is 2-D because $A\vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2$ so there are only 2 independent ways of specifying the range of A.

3.3 Echelon Forms of a Matrix $A \in \mathbb{R}^{m\times n}$

We previously considered the pivoted LR decomposition of a matrix in the context of a linear equation. Here, we will look at the matrix only; this is justified because any linear equation can be expressed and manipulated as an augmented matrix.

The *leading coefficient* of a row is defined as the first non-zero entry of the row. This is also called the *pivot* entry of the row, so the column of the pivot entry is the *pivot column*.

Here, we are concerned with a special matrix decomposition, the *reduced row echelon form* or *RREF*. It is part of a simple hierarchy of decompositions, each with useful properties:

Echelon Form: Result of "ordinary" Gaussian elimination. For A, doing elimination $A = LR$ without pivoting, matrix R is in echelon form. This form has the properties:

1. The leading coefficient of a given row is strictly to the right of the leading coefficient of every row above it
2. Every entry below a leading coefficient is zero

Row Echelon Form: Permutation of echelon form so that all non-zero rows are "above" all zero rows. For A, doing elimination $PA = LR$ using pivoting, matrix R is in row echelon form. This form has the properties of echelon form, plus:

3. All non-zero rows are "above" all zero rows

A special case is when the leading coefficient of each row of R is 1. This occurs if and only if R is *unit upper triangular*, or *upper unitriangular*, or *strictly upper triangular*.

We can write this case as $PA = LDR$ where D is a diagonal matrix. By convention, the diagonal entries of D are all 1 unless the corresponding row needs scaling.

Reduced Row Echelon Form, RREF: Row echelon form, with every row is scaled so that its leading coefficient is 1, and "upwards eliminated" so that every entry above a leading coefficient is 0. This decomposition has the properties of row echelon form, plus:

4. Every leading coefficient is 1
5. Every entry above the leading coefficient is zero

The upwards elimination can be formulated as a matrix decomposition. Because it works oppositely to Gaussian elimination, it uses an upper-triangular matrix U that operates just as the lower-triangular L operates. The RREF of a matrix A can be written as the decomposition $PA = LDUR$.

Permuted RREF: A special case of the reduced row echelon form. In addition to the five requirements of RREF, such a matrix also has the property:

6. The upper left block is the identity matrix

The permuted RREF can also be defined in terms of operations. Beginning with the reduced row echelon form, columns are permuted so that the leading entry of each row (which must be 1) has the same column index as row index. The result has a special block partitioning:

$$R = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}$$

The process of finding the RREF of a matrix is straightforward, but we must pay attention when pivoting is required.

These examples are suggested for self-study before the class is presented. The instructor will work through these examples in class. After the class, sample computations will be made available so that a student can check their own results.

Consider these 2×3 examples:

$$
\begin{aligned}
A_1 &= \begin{bmatrix} 1 & -4 & 3 \\ 1 & -2 & 4 \end{bmatrix} \\
A_2 &= \begin{bmatrix} 1 & -4 & 3 \\ 1 & -4 & 2 \end{bmatrix} \\
A_3 &= \begin{bmatrix} 2 & -8 & 6 \\ 1 & -4 & 3 \end{bmatrix}
\end{aligned}
$$

Consider this 3×2 example:

$$
A_4 = \begin{bmatrix} 1 & 1 \\ -4 & 2 \\ 3 & 4 \end{bmatrix}
$$

3.4 Null Space and the RREF

We will use the RREF often, because it describes two important vector spaces related the the matrix A. The first important space is the *null space* of a matrix. The null space is defined as:

> *The null space of a matrix A, often written as $N(A)$, is the set of weight vectors $\vec{d}$ such that $A\vec{d} = \vec{0}$*

To see how the null space works, consider one matrix from the textbook:

$$
A_5 = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 3 & 8 & 6 & 16 \end{bmatrix}
$$

The RREF decomposition produces the upper unitriangular matrix

$$
R_5 = \begin{bmatrix} \mathbf{1} & \mathbf{0} & 2 & 0 \\ \mathbf{0} & \mathbf{1} & 0 & 2 \end{bmatrix}
$$

The RREF columns in bold font, column 1 and column 2, are linearly independent; this means that remaining two columns of R_5 can be computed as linear combinations of the first two columns. (In this case the computation is trivial, but in general the computation is less obvious.)

This linear dependence produces a peculiar and important consequence. Inspecting the first row of R_5, we see that if it operates on the weight vector

$$\vec{d_1} = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

then the result is zero. The second row also produces a zero value, so we can conclude that

$$R_5\vec{d_1} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{3.5}$$

Inspecting the second row of R_5, we see that if it operates on the weight vector

$$\vec{d_2} = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 1 \end{bmatrix}$$

then the result is also zero. We can conclude that

$$R_5\vec{d_2} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{3.6}$$

Putting together Equation 3.5 and Equation 3.6, we can conclude that any linear combination of the weight vectors $\vec{d_1}$ and $\vec{d_2}$ produce a zero vector. This is equivalent to saying that $\vec{d_1}$ and $\vec{d_2}$ ***span*** the null space of R_5.

What does this say about the null space of the original matrix A_5? Because the RREF decomposition of A_5 can be written as

$$A_5 = [P_5^{-1} L_5 D_5 U_5] R_5 = M_5 R_5$$

any weight vector $\vec{d}$ that is mapped to zero by R_5 must also be mapped to zero by A_5. So the null space of R_5 and of A_5 are the same.

Another example from the textbook is the matrix

$$A_6 = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 2 & 2 & 8 & 10 \\ 3 & 3 & 10 & 13 \end{bmatrix}$$

The RREF decomposition produces the upper unitriangular matrix

$$R_6 = \begin{bmatrix} \mathbf{1} & 1 & \mathbf{0} & 1 \\ \mathbf{0} & 0 & \mathbf{1} & 1 \\ \mathbf{0} & 0 & \mathbf{0} & 0 \end{bmatrix}$$

The pivot columns of R_6, columns 1 and 3 shown in bold face, are linearly independent. The remaining columns, 2 and 4, are linear combinations of the pivot columns.

The null space of R_6 is less obvious. In a later class, we will develop a simple algorithm for finding the null space from the permuted RREF; for now, a careful student might verify that the vectors

$$\vec{d}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \vec{d}_2 = \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

are in the null space of R_6 and that they are also in the null space of A_6.

It is common usage to assemble a set of independent null vectors into a matrix. Using this convention, we could write the null spaces of A_5 and A_6 as

$$\begin{aligned} N(A_5) &= \begin{bmatrix} -2 & 0 \\ 0 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \\ N(A_6) &= \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

A careful student might verify that, for the other example matrices studied in the class, the null spaces are:

$$
\begin{aligned}
N(A_1) &= \begin{bmatrix} -5 \\ -0.5 \\ 1 \end{bmatrix} \\
N(A_2) &= \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix} \\
N(A_3) &= \begin{bmatrix} 4 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}
$$

For the matrix A_4, there are no non-empty 3-vectors that map to the zero vector. We say that the null space for A_4 is trivial.

Extra Notes

These are the instructor's computations for the examples above.

$$
\begin{aligned}
A_1 &= \begin{bmatrix} 1 & -4 & 3 \\ 1 & -2 & 4 \end{bmatrix} \\
\text{do LR:} &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -4 & 3 \\ 0 & 2 & 1 \end{bmatrix} \\
\text{scale:} &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{2} \end{bmatrix} \begin{bmatrix} 1 & -4 & 3 \\ 0 & 1 & 0.5 \end{bmatrix} \\
\text{up-elim:} &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{-4} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 0.5 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
A_2 &= \begin{bmatrix} 1 & -4 & 3 \\ 1 & -4 & 2 \end{bmatrix} \\
\text{do LR:} &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -4 & 3 \\ 0 & 0 & -1 \end{bmatrix} \\
\text{scale:} &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{-1} \end{bmatrix} \begin{bmatrix} 1 & -4 & 3 \\ 0 & 0 & 1 \end{bmatrix} \\
\text{up-elim:} &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{3} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} 1 & -4 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
A_3 &= \begin{bmatrix} 2 & -8 & 6 \\ 1 & -4 & 3 \end{bmatrix} \\
\text{do LR:} &= \begin{bmatrix} 1 & 0 \\ 0.5 & 1 \end{bmatrix} \begin{bmatrix} 2 & -8 & 6 \\ 0 & 0 & 0 \end{bmatrix} \\
\text{scale:} &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{2} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} 1 & -4 & 3 \\ 0 & 0 & 0 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
A_4 &= \begin{bmatrix} 1 & 1 \\ -4 & 2 \\ 3 & 4 \end{bmatrix} \\
\text{do LR:} &= \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 0 & 1 \end{bmatrix} \\
\text{up-elim:} &= \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{1} \\ \mathbf{0} & \mathbf{1} & \mathbf{2} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}
$$

End of Extra Notes