COMS10003 : Linear Algebra

Vector Spaces, Span and Basis

Andrew Calway

March 16, 2015

Introduction

In this section we carry on our look at linear algebra by considering how we can *represent* sets of vectors in terms of *linear combinations* of other vectors. We are interested in particular sets of vectors - known as vector *spaces and subspaces* - and what sets of vectors we can use to represent them in terms of linear combinations - these are called *spanning sets* - and also spanning sets with the smallest number of vectors - known as *bases*. From the latter we will also generlise what we mean by the coordinates of a vector. Finally, we'll look at the concept of projecting vectors onto a subspace of vectors, which will prepare us for looking at solving sets of linear equations in the next section. As previously, I have made use of several textbooks from my bookshelves and these are listed below.

Theory and problems of linear algebra by Seymour Lipschutz, McGraw-Hill,1981.

Linear Algebra and Probability for CS Applications by Ernest Davis, CRC Press, 2012.

Coding the Matrix by Philip N Klein, Newtonian Press, 2013.

Vector Spaces and Subspaces

We start by looking at two important concepts - the vector space and the vector subspace. As with vectors, these are sometimes considered in abstract terms, but we are going to look at them in concrete terms based on our vectors with real number components. We start with some definitions:

Linear Combination

Let $\mathcal{V}$ be a set of vectors. Then the vector $\mathbf{u}$ is a **linear combination** over $\mathcal{V}$ if there exists vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\ldots$, $\mathbf{v}_m$ in V and scalars a_1, a_2, $\ldots$, a_m such that $\mathbf{u} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m = \sum_{i=1}^{m} a_i\mathbf{v}_i$. In other words, we can write the vector $\mathbf{u}$ as a weighted sum of vectors from $\mathcal{V}$ where the weights are the scalars a_i.

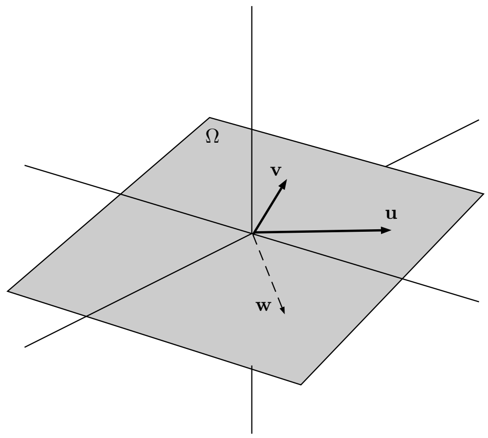

Figure 1: The vectors **u** and **v** span a subspace containing all vectors such as **w** which are in the plane Ω.

Span

Let $\mathcal{S}$ be a set of vectors. The **span** of $\mathcal{S}$, denoted $Span(\mathcal{S})$, is the set of linear combinations over $\mathcal{S}$. In other words, it is the set of all vectors that are 'generated' by computing all the linear combinations of the vectors in $\mathcal{S}$. Thus we say that the vectors in $\mathcal{S}$ span the vectors in the set $Span(\mathcal{S})$.

Example 1

Let $\mathcal{S} = \{\mathbf{s}_1, \mathbf{s}_2\}$ where $\mathbf{s}_1 = (1, 2)$ and $\mathbf{s}_2 = (-1, 1)$. Then the vector $\mathbf{v} = (5, 1)$ is a linear combination over $\mathcal{S}$ since $(5, 1) = 2(1, 2) - 3(-1, 1)$. In fact, all 2-D vectors in $\mathcal{R}^2$ are spanned by $\mathcal{S}$, since $Span(\mathcal{S}) = \{(a - b, 2a + b) : a, b \in \mathcal{R}\}$ and suitable choices of a and b enable us to generate any vector in $\mathcal{R}^2$.

Vector Spaces and Subspaces

A set of vectors $\mathcal{V}$ is a **vector space** if there exists a set of vectors $\mathcal{S}$ for which $\mathcal{V} = Span(\mathcal{S})$. In other words, a vector space is the set of all the possible linear combinations of a set of vectors. This is not a complete and strict definition of a vector space, but it suffices for our purposes. You should take some time to look up the formal definition and also convince yourself that $\mathcal{R}^n$ is a vector space. We can also define a **vector subspace** : if $\mathcal{V}$ and $\mathcal{W}$ are both vector spaces and $\mathcal{W} \subset \mathcal{V}$, then $\mathcal{W}$ is a subspace of $\mathcal{V}$.

Example 2

Let $\mathcal{S} = \{(-1, 0, 1), (1, 1, 1)\}$ then all vectors $(b - a, b, a + b)$ for $a, b \in \mathcal{R}$ are spanned by $\mathcal{S}$ and hence form a subspace of $\mathcal{R}^3$. The vector $(1, 2, 3)$ is in the subspace but the vector $(-2, 1, -1)$ is not in the subspace (you should convince yourself of this). It should be clear that in terms of geometry, all the vectors in the subspace are in a plane, similar to that shown in Fig. 1. Note also that both $Span((-1,0,1))$ and $Span((1,1,1))$ are also subspaces - they contain all the scalar multiples of $(-1,0,1)$ and $(1,1,1)$, respectively, and correspond to lines in 3-D.

Basis and Dimension

Having defined the notions of vector spaces and subspaces as sets of vectors which can be generated by another set of vectors, it is natural to ask how many sets of vectors exist that will generate a given subspace, how many vectors are in them and whether there are some that are 'better' than others.

Of course, we need to be clear what we mean by 'better', but a good choice would seem to be those containing the smallest number of vectors. Why? Well it means that we would then be able to represent any vector in the space using just those common set of vectors plus the same number of scalars. Thus it would be a *minimal representation* requiring the least amount of storage - which makes us happy as computer scientists.

Also, we would likely be interested in how easy it is to generate such representations and then prefer those sets for which generation required the least amount of computation. We will look at that issue in the section once we have determined the minimum number of generating vectors we need. To start, we need some more definitions.

Linear (In)Dependence

A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ are said to be a **linearly dependent** set if there exists scalars a_i such that $a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \ldots a_n\mathbf{u}_n = 0$. Another way of putting this is that the set is dependent if we can write any one of the vectors as a linear combination of the others. The counter case is a **linearly independent** set - no vector can be written as a linear comniation of the others.

For example, the vectors $(-1,0,1)$, $(1,1,1)$ and $(0,1,2)$ are dependent since $(0,1,2) = (-1,0,1) + (1,1,1)$, whereas $(-1,0,1)$, $(1,1,1)$ and $(1,0,1)$ are independent - convince yourself of this.

Basis and Dimension

We are now in a position to determine the minimal number of vectors that need to be in a generating set for a given vector space. Consider a generating set $\mathcal{S}$ for a vector space $\mathcal{V}$. Assume first that $\mathcal{S}$ is a dependent set. This means that one or more vectors in $\mathcal{S}$ can be written as a linear combination of the others. Thus if we remove them from the set then we can still generate the vector space - they are redundent; if we need them then we just generate them from the reduced set.

It follows that the minimum number of vectors we need to generate the space is the number that is left once we have removed all the dependent vectors, i.e. until we are left with an independent set. Thus - the smallest set that we can have is one that both generates the vector space, i.e. its vectors span the space, and is a linearly independent set.

Such a set of vectors is known as a **basis** for the space - think of it as 'the base' of the space from which all others are built - and the number of vectors in the set is known as the **dimension** of the space - think of it as the number of degrees of freedom in the space.

Example 3

The set $\mathcal{S} = \{(-1,0,1),(1,1,1)\}$ are linearly independent and thus form a basis for the subspace $Span(\mathcal{S})$ corresponding geometrically to a plane in 3-D. The dimension of the subspace is 2. The set $\mathcal{S} = \{(-1,0,1),(1,1,1),(0,1,2)\}$ are linearly dependent and so although it spans the subspace it is not a basis for it.

Example 4

The 5-D vectors $\mathbf{v}_1 = (1,0,2,-1,1)$, $\mathbf{v}_2 = (-2,1,0,1,-1)$, $\mathbf{v}_3 = (0,-1,1,-2,0)$ and $\mathbf{v}_4 = (4,2,1,3,3)$ are linearly dependent since $\mathbf{v}_4 = 2\mathbf{v}_1 - \mathbf{v}_2 - 3\mathbf{v}_3$. The set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ are independent and are therefore a basis for a 3-D subspace of $\mathcal{R}^5$.

Coordinates and Orthonormal Bases

So now we know the minimum number of vectors we need to generate a vector space - we need a basis - the next question to ask is how we determine the representation for a given vector in our space in terms of our basis. Or more specifically, if $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m\}$ for a subspace of dimension m, how do we find the scalars a_i such that a vector $\mathbf{v}$ in the subspace is given by $\mathbf{v} = \sum_{i=1}^{m} a_i \mathbf{b}_i$?

The scalars a_i are known as the **coordinates** of the vector $\mathbf{v}$ with respect to the basis set $\mathcal{B}$. Actually, we have been using coordinates since we started - the components of our vectors are coordinates w.r.t to what we call the **standard basis**. For example, in $\mathcal{R}^3$, the standard basis is the set $\{(1,0,0),(0,1,0),(0,0,1)\}$ and it is easy to see that the weights required to build any vector from a linear combination of these vectors correspond to the components of the vector.

But what about the coordinates w.r.t any another basis? In such cases we have solve for the a_i in the linear combination $\mathbf{v} = \sum_{i=1}^{m} a_i \mathbf{b}_i$. For example, taking the vectors in *Example 4*, if we assume that $(4,2,1,3,3) = a_1(1,0,2,-1,1) + a_2(-2,1,0,1,-1) + a_3(0,-1,1,-2,0)$

then we get

$$4 = a_1 - 2a_2 \quad 2 = a_2 - a_3 \quad 1 = 2a_1 + a_3 \quad 3 = -a_1 + a_2 - 2a_3 \quad 3 = a_1 - a_2$$

By substitution, we then get $a_1 = 2$, $a_2 = -1$ and $a_3 = -3$, as in *Example 4*. Note that these values satisfy all five of the above equations. What would it mean if we couldn't find values to satisfy all five? It would mean that $(4, 2, 1, 3, 3)$ couldn't be represented as a linear combination of the other vectors and thus was outside of the subspace.

It turns out however that there is a quicker way of determining the coordinates w.r.t a given basis if we choose the latter carefully. If as well as being **linearly independent** the basis vectors are also **unit vectors** and **orthogonal** to each other, then we can show (you will do this in the workshop) that the coordinates for a vector are given by the dot product between the vector and each basis vector.

Thus, if vector $\mathbf{v}$ lies in a subspace with basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m\}$ and

$$\mathbf{b}_i.\mathbf{b}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

then $\mathbf{v} = \sum_{i=1}^{m} a_i \mathbf{b}_i$ if $a_i = \mathbf{v}.\mathbf{b}_i$. This type of basis is known as an **orthonormal basis**, i.e. the basis vectors are **ortho**gonal and **normal**ised to have length one.

We might then ask how do we find an orthonormal basis for any given subspace? This can be done using an algorithm called **Gram-Schmidt orthogonalisation**. We shan't look at this but if you are interested, then look it up.

Example 5

The two 4-D unit vectors $\mathbf{b}_1 = 0.5(1, -1, -1, 1)$ and $\mathbf{b}_2 = 0.5(-1, -1, 1, 1)$ are linearly independent and orthogonal. The vector $\mathbf{v} = (2.5, 0.5, -2.5, -0.5)$ is in the subspace spanned by the two vectors and its coordinates w.r.t the basis are

$$a_1 = \mathbf{v}.\mathbf{b}_1 = 0.5((1 \times 2.5) + (-1 \times 0.5) + (-1 \times -2.5) + (1 \times -0.5)) = 0.5 \times 4 = 2$$

$$a_2 = \mathbf{v}.\mathbf{b}_2 = 0.5((-1 \times 2.5) + (-1 \times 0.5) + (1 \times -2.5) + (1 \times -0.5)) = 0.5 \times -6 = -3$$

which is correct since

$$(2.5, 0.5, -2.5, -0.5) = 2 \times 0.5(1, -1, -1, 1) - 3 \times 0.5(-1, -1, 1, 1)$$

Projections onto Subspaces

To finish off, let's return to our discussion about the dot product. Recall that we said that it could be seen as the projection of one vector onto the direction of the other. In other words, it sort of tells how much there is of one vector in the direction of the other. It turns out that it is very useful to ask a similar question about vectors and subspaces - how much of a vector is within a given subspace?

We can answer this by considering the projection of a vector onto a subspace. What does this mean? Well in the same way as with the dot product, it means the vector within the subspace which is closest to the vector being projected, i.e. the distance between them is the smallest possible across all vectors in the subspace.

Let $\mathbf{v}' = \sum_{i=1}^{m} c_i \mathbf{b}_i$ be the projection of a vector $\mathbf{v}$ onto the subspace spanned by an orthonormal basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m\}$, then since it is the closest vector to $\mathbf{v}$, the difference vector $\mathbf{v} - \mathbf{v}'$ and $\mathbf{v}'$ must be orthogonal. We can see this by considering the geometry as in the case of the dot product. We can also show (and you'll do this as well in the workshop) that the coordinates c_i of $\mathbf{v}'$ are given by $c_i = \mathbf{v}.\mathbf{b}_i$, i.e. the projection of $\mathbf{v}$ onto each of the basis vectors.

Example 6

Consider the subspace spanned by the orthonormal basis $\mathbf{b}_1$ and $\mathbf{b}_2$ in *Example 5*. We want to find the projection of the vector $\mathbf{v} = (-0.5, -1, 0, 1)$ onto the subspace. Computing the projection of $\mathbf{v}$ onto each basis vector gives

$$\mathbf{v}.\mathbf{b}_1 = 0.5(-0.5 + 1 + 0 + 1) = 0.75 \qquad \mathbf{v}.\mathbf{b}_2 = 0.5(0.5 + 1 + 0 + 1) = 1.25$$

Hence the projection is given by

$$\mathbf{v}' = 0.75 \times 0.5(1, -1, -1, 1) + 1.25 \times 0.5(-1, -1, 1, 1) = (-0.25, -1, 0.25, 1)$$

which obviously lies in the subspace. We can now check whether the difference vector is orgononal to subspace, i.e. $\mathbf{d} = \mathbf{v} - \mathbf{v}' = (-0.5, -1, 0, 1) - (-0.25, -1, 0.25, 1) = (-0.25, 0, -0.25, 0)$. Forming the dot product with each basis vector gives

$$\mathbf{b}_1.\mathbf{d} = 0.5(-0.25 + 0 + 0.25 + 0) = 0 \qquad \mathbf{b}_2.\mathbf{d} = 0.5(0.25 + 0 - 0.25 + 0) = 0$$

which shows that $\mathbf{d}$ is orthogonal to both basis vectors and hence to all vectors in the subspace (you should convince yourself of that as well). In fact this example is interesting since it illustrates the basic mechanism used in the Gram-Schmidt algorithm - it iteratively projects vectors onto subspaces, building up an orthogonal basis as it goes.
