

Week 3b

Wednesday, September 30, 2020 11:38 PM

https://olincollege.sharepoint.com/:p:/r/sites/QEA1aGroup2/_layouts/15/Doc.aspx?sourcedoc=%7BF25F52B2-368A-4FF1-84C8-B64EA9BBADD6%7D&file=week3b.pptx&action=edit&mobileredirect=true&DefaultItemOpen=1

Chapter 8

Week 3b: Linear Independence, Span, Basis

Schedule

8.1	Linear Independence and Span [30 mins]	74
8.2	Orthogonality [30 mins]	77
8.3	Decomposition [30 mins]	77

In the last class, you were introduced to the idea of linear independence, span and basis. Today, we will dig deeper into these ideas. We will start with some review of the main ideas from the previous class, followed by an introduction of some new ideas.

8.1 Linear Independence and Span [30 mins]

Exercise 8.1

1. In words, describe what it means for a set of 3×1 vectors to be linearly independent.
2. What is the maximum number of vectors in a set of 3×1 vectors that are linearly independent?

Exercise 8.2

Next, let's do a problem similar to what you saw in last class in MATLAB. Consider the following matrix:

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & 4 \\ 1 & 1 & 3 \end{bmatrix} \quad (8.1)$$

The third column of this matrix equals the second column plus twice the first column. Hence these three vectors lie on some plane (not the xy -plane as in the previous part).

1. Open up MATLAB and using the `quiver3` command together with `hold on`, please plot the vectors corresponding to the three columns of \mathbf{B} . Note that typing » `quiver3(0,0,0,`

1. This
indepe
cannot
linear
others
2. 3 ve

- `1, 1, 1)`; in MATLAB, plots an arrow from the origin to the point $(1,1,1)$, i.e. it plots the vector corresponding to the first column of **B**. Typing `» hold on` in MATLAB results in subsequent calls to `quiver` appearing on the same axes, without erasing previous arrows.
- Using the "rotate 3D" function on the MATLAB figure window, rotate the figure around so that it appears as if all three arrows overlap. This should indicate that the vectors lie on a plane.
 - Using `det` compute the determinant of matrix **B**. Does this make sense?

The det
matrix i
three v

The fundamental property here is that the columns of the **A** and **B** matrices are not *linearly independent*. We shall next define the idea of linearly independent vectors more formally, for an arbitrary number of dimensions (recall that we went through this for 3 dimensional vectors in the last class).

- A finite set $S = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ of vectors in \mathbf{R}^n is said to be *linearly dependent* if there exist scalars c_1, c_2, \dots, c_m which are not all zero, such that

$$c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + \dots + c_m \mathbf{x}_m = \mathbf{0}.$$

Note that \mathbf{R}^n here refers to the set of all n -dimensional vectors that are made up of real numbers. (For example, \mathbf{R}^1 is the real line and \mathbf{R}^2 is the plane.) For any value of n , \mathbf{R}^n is an example of a *vector space* - we will meet different examples of vector spaces in the future. We can also express this equation using a matrix **A**, whose columns are $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$.

$$[\mathbf{x}_1 \quad \mathbf{x}_2 \quad \dots \quad \mathbf{x}_m] \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = \mathbf{0}. \quad (8.2)$$

If a non-zero solution exists to $\mathbf{A}\mathbf{c} = \mathbf{0}$ then the set of vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ is linearly dependent. In the case of a square matrix ($n = m$), the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ are linearly dependent if and only if $\det(\mathbf{A}) = 0$. Otherwise, the only way to satisfy the equation above is if $c_1 = c_2 = \dots = c_m = 0$. Figure 8.1 illustrates two examples of three vectors that are in 3D space, but are linearly dependent, since in each case, all three vectors are on a plane.

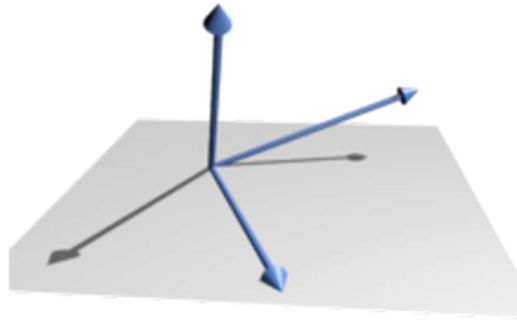


Figure 8.1: Linearly dependent vectors in \mathbf{R}^3 . (from Wikimedia Commons).

- The set of vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ is *linearly independent* if it is not linearly dependent. In other words, the set of vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ is linearly independent if

$$c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + \dots + c_m \mathbf{x}_m = \mathbf{0} \quad (8.3)$$

only when $c_1 = c_2 = \dots = c_m = 0$. In other words, if the only solution to $\mathbf{A}\mathbf{c} = \mathbf{0}$ is $\mathbf{c} = \mathbf{0}$, the set of vectors made up of the columns of **A** is linearly independent. For a square matrix this means the set is linearly independent if and only if $\det(\mathbf{A}) \neq 0$.

Figure 8.2: Linearly independent vectors in \mathbf{R}^3 . (from Wikimedia Commons).

- The *span* of S is the set of all linear combinations of its vectors. In other words, the span of the set S is the set of all possible vectors of the form

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_m\mathbf{x}_m$$

The *span* is usually denoted by $\text{span}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m)$.

- A finite set $S = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ of vectors is said to form a *basis* of a vector space V , if the vectors in S are linearly independent, and every point in V can be expressed as a linear combination of the vectors in the set S . Hence, if a set of vectors S is linearly independent those vectors form a *basis* of the set which is the span of those vectors.

Let's solidify our understanding of linear dependence, bases and span by working on a few problems by hand.

Exercise 8.3

- In words, describe the span of the vectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.
- In words, describe the span of the vectors $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ which are all in 3-dimensional Euclidean space.

1. The span of the vectors is a line.

2. The span of the vectors is a plane.

8.2 Orthogonality [30 mins]

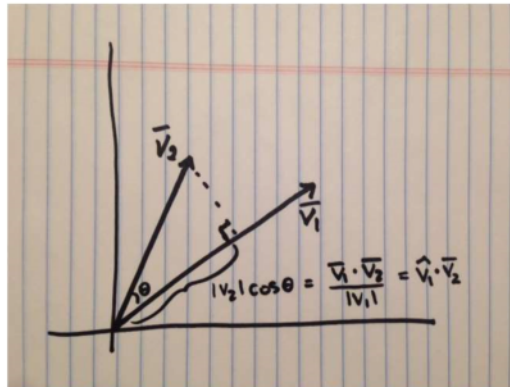


Figure 8.3: Projection

By trigonometry, if we have two vectors \mathbf{v}_1 and \mathbf{v}_2 which have an angle of θ between them, the component of \mathbf{v}_2 which lies along the direction of \mathbf{v}_1 is $|\mathbf{v}_2| \cos \theta$. Since the dot product of the two vectors can be expressed as $|\mathbf{v}_1||\mathbf{v}_2| \cos \theta$, this component (referred to as the projection) can be written as $\mathbf{v}_1 \cdot \mathbf{v}_2 / |\mathbf{v}_1|$. If the projection is zero, the vectors are *orthogonal*, and $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. If the vectors are unit length, in addition to being normal, the vectors are said to be *orthonormal*. Additionally, if a basis set is made up of orthonormal vectors, it is known as an orthonormal basis.

A square matrix with columns of unit vectors which are orthogonal to each other is known as an orthogonal matrix. An orthogonal matrix \mathbf{A} has the property that $\mathbf{A}^T = \mathbf{A}^{-1}$.

Exercise 8.4

Which of the following pairs of vectors are orthogonal or orthonormal?

1. $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$

Neither orthogonal nor orthonormal

2. $\begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$

Orthogonal

3. $\begin{bmatrix} \frac{2}{\sqrt{13}} \\ \frac{-2}{\sqrt{13}} \\ \frac{3}{\sqrt{13}} \end{bmatrix}, \begin{bmatrix} \frac{-3}{\sqrt{13}} \\ \frac{3}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} \end{bmatrix}$

Neither orthogonal nor orthonormal

8.3 Decomposition [30 mins]

Suppose we have a set (collection) of m basis vectors $\{\mathbf{v}_i\}$ which are normalized ($|\mathbf{v}_i| = 1$), mutually orthogonal ($\mathbf{v}_i^T \mathbf{v}_j = 0$ unless $i = j$) and span our space (every point can be written as some linear combination of the vectors $\{\mathbf{v}_i\}$). How do we actually find the linear combination which is equal to a given vector in our space?

Let's say we have a vector \mathbf{w} which we are interested in expressing as a linear combination of our set of orthonormal vectors $\{\mathbf{v}_i\}$. We can write this linear combination as

$$\mathbf{w} = \sum_{i=1}^m c_i \mathbf{v}_i \quad (8.4)$$

and our problem is now to find the coefficients c_i in this expression.

The obvious option is to pack the basis vectors \mathbf{v}_i into the columns of a matrix \mathbf{A} , and find solutions of

$$\mathbf{A}\mathbf{c} = \mathbf{w}$$

Since the columns of \mathbf{A} are formed from basis vectors they are linearly independent and a non-zero solution exists and can be determined by the usual methods.

However, our basis vectors form an orthogonal set (collection) which permits a more direct calculation. Consider a particular vector \mathbf{v}_k in our basis set, and let's take the dot product between \mathbf{v}_k and our vector \mathbf{w} :

$$\mathbf{v}_k^T \mathbf{w} = \mathbf{v}_k^T \sum_{i=1}^m c_i \mathbf{v}_i \quad (8.5)$$

Distributing the dot product into the summation we have:

$$\mathbf{v}_k^T \mathbf{w} = \sum_{i=1}^m c_i \mathbf{v}_k^T \mathbf{v}_i \quad (8.6)$$

But from orthogonality we know that the dot product of any two different vectors in our orthonormal set is zero, so all terms in the sum where $k \neq i$ are zero. This leads to the following simplification

$$\mathbf{v}_k^T \mathbf{w} = c_k \mathbf{v}_k^T \mathbf{v}_k \quad (8.7)$$

In addition, since our set of vectors is normalized, we know that $\mathbf{v}_k^T \mathbf{v}_k = 1$, leaving us with

$$\mathbf{v}_k^T \mathbf{w} = c_k \quad (8.8)$$

This gives us a very nice, simple way of decomposing a vector into a linear combination of the vectors within our basis set. The dot product of each basis vector with our target vector will result in the coefficient of that term in the linear decomposition.

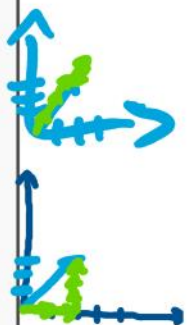
Exercise 8.5

1. There are many (in general, an infinite number) of bases for a given set V . Hence, we can describe elements in the set V as linear combinations of vectors from different bases. Consider the following two basis sets which form bases for 2-dimensional space.

$$\bullet \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ and}$$

$$\bullet \mathbf{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Express the vector $\mathbf{w} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ as a linear combination of the first basis set (i.e., a sum of scaled versions of each vector in the basis set). Repeat for the second. Please make two different drawings of $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$, one expressed as a sum of scaled vectors in the first basis set and another for the vectors from the second basis set. Please label the lengths of each vector in the set.



; 1 2 2];

2. Suppose that you wish to write the vector $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$ as a linear combination of the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} \text{ and } \mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}.$$

Please write a matrix equation to find the coefficients of the linear combination, and solve for the coefficients using MATLAB if possible.

3. Representing vectors using different bases is a very powerful technique that we will keep coming back to in this class (in both semesters). Vectors described in different bases can give us insight that may not be so obvious when viewed in the original basis. Representing vectors in different bases can also be used for dimensionality reduction, which is an important technique that is used to speed up computations and compress data in a number of different fields. Here we will consider a problem of lossy data compression using a change of basis. Lossy compression refers to methods of representing data more efficiently, but with a loss of accuracy. Examples of lossy data compression include jpg images, and mp3 audio files. If care is taken in lossy compression, the effects of the data loss can be kept at acceptable levels (this is of course subjective and dependent on the application). We will start with a toy example and then move to more complicated ones in subsequent homework problems. Consider a set of four 2-dimensional data variables stored in the following vectors:

$$\mathbf{d}_1 = \begin{bmatrix} 2.2 \\ 1.2 \end{bmatrix}, \mathbf{d}_2 = \begin{bmatrix} 1 \\ 0.6 \end{bmatrix}, \mathbf{d}_3 = \begin{bmatrix} 1.5 \\ 0.7 \end{bmatrix}, \mathbf{d}_4 = \begin{bmatrix} 1.7 \\ 0.8 \end{bmatrix} \quad (8.9)$$

- (a) In MATLAB, plot the data using points (without lines connecting them) by typing `plot([2.2 1 1.5 1.7], [1.2 0.6 0.7 0.8], 'o')`; You will find that these points lie close to the line through the origin with slope $1/2$.
- (b) Define a unit vector that points in the direction $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and call it \mathbf{u}_1 . Find another unit vector that is orthogonal to \mathbf{u}_1 and call it \mathbf{u}_2 . These vectors form a basis in 2 dimensional space.
- (c) Rather than storing the original data, we are now going to express the original data in terms of the new basis that we have defined. To do that, write $\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3$ and \mathbf{d}_4 , as a linear combination of \mathbf{u}_1 and \mathbf{u}_2 . You can use MATLAB here to find the coefficients.
- (d) In this toy example, we are going to "compress" our data by only keeping the coefficients corresponding to \mathbf{u}_1 . i.e. we will discard the coefficient corresponding to \mathbf{u}_2 . Suppose that we wish to recover approximations to $\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \mathbf{d}_4$, from the four coefficients. These approximations, which you should denote by $\tilde{\mathbf{d}}_1, \dots, \tilde{\mathbf{d}}_4$, are all scaled versions of \mathbf{u}_1 . In your axes from part a, please plot the points corresponding to $\tilde{\mathbf{d}}_1, \dots, \tilde{\mathbf{d}}_4$. Do you think they make good approximations?
- (e) We can describe how well our compressed data represents our original data. One way to do this is to calculate the difference between our original and compressed data, and call this error vector $\mathbf{f}_i = \mathbf{d}_i - \tilde{\mathbf{d}}_i$. Now, compute the size of this error using $\text{norm}(\mathbf{f}_i)$ for $i = 1, 2, 3, 4$. Then, summarize the error by finding the root-mean-square (RMS) error between your approximations and the true data points. The RMS function squares the errors, takes the mean, and then takes the square root. This quantity is a single number that can be used to measure how well or poorly your compressed data represents your original data. You may find MATLAB's `norm` and `rms` functions helpful here.

2 2

This toy example illustrates that we can sometime be more efficient (albeit at the cost of some accuracy) in representing (or computing) data when it is expressed in certain bases.

Solution 8.1

1. A set of vectors is linearly independent if it is not possible to scale and sum these vectors to result in the all-zeros vector, except if all the scale factors are zero.
2. You can have at most 3 vectors in a set of linearly independent 3×1 vectors

Solution 8.2

1. Type the following into MATLAB:
 » `quiver3(0,0,0,1,1,1)`
 » `hold on`
 » `quiver3(0,0,0,1,2,1)`
 » `quiver3(0,0,0,3,4,3)`
- 2.
3. The determinant of \mathbf{B} is zero. Recall that a matrix is not invertible if and only if the determinant is zero. This matrix is not invertible since it collapses all vectors to a plane.

Solution 8.3

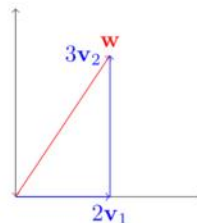
1. The span of these two vectors is all over \mathbf{R}^2 , i.e., a plane.
2. The span of these three vectors is the xy -plane in \mathbf{R}^3 .

Solution 8.4

1. The dot product of these two vectors is non-zero, so they are not orthogonal.
2. The dot product of these two vectors is zero, so they are orthogonal.
3. The dot product of these two vectors is zero, so they are orthogonal. Furthermore, each vector is unit length, so they are orthonormal.

Solution 8.5

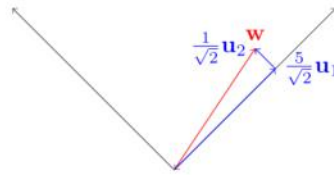
1. It's clear that $2\mathbf{v}_1 + 3\mathbf{v}_2 = \mathbf{w}$. We visualize this as



To write \mathbf{w} as a linear combination of the basis vectors \mathbf{u}_1 and \mathbf{u}_2 requires a bit more work. We can set up the matrix equation

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

and solve to learn that $\frac{5}{\sqrt{2}}\mathbf{u}_1 + \frac{1}{\sqrt{2}}\mathbf{u}_2 = \mathbf{w}$. We can visualize this as



2. First, we create a matrix in MATLAB whose columns are the vectors \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 ,
`» V=[1 3 1; 1 1 2; 1 2 2]`
 and the vector \mathbf{w} ,
`» w=[1; 2; 4]`.
 Let \mathbf{c} be the vector of coefficients. We have the equation $\mathbf{V}\mathbf{c} = \mathbf{w}$, so to solve for \mathbf{c} we compute
 $\mathbf{c} = \mathbf{V}^{-1}\mathbf{w}$. In MATLAB, we use `» inv(V)*w`. This tells us that $\mathbf{w} = -10\mathbf{v}_1 + 2\mathbf{v}_2 + 5\mathbf{v}_3$.
3. (a)
 (b) We define `» u1=[2; 1]` and `» u2=[-1; 2]`. There are other choices for \mathbf{u}_2 , but they are all constant multiples of this choice, e.g., `» u2=[-2; 4]`.
 (c) Create a 2×2 matrix with \mathbf{u}_1 and \mathbf{u}_2 as the columns,
`» U=[2 -1; 1 2]`
 and a 2×4 matrix the vectors \mathbf{d}_i as the columns
`» D=[2.2 1 1.5 1.7; 1.2 0.6 0.7 0.8]`.
 Then compute
`» inv(U)*D`
 to get the matrix of coefficients. This tells us that

$$\mathbf{d}_1 = 1.12\mathbf{u}_1 + 0.04\mathbf{u}_2, \mathbf{d}_2 = 0.52\mathbf{u}_1 + 0.04\mathbf{u}_2,$$

$$\mathbf{d}_3 = 0.74\mathbf{u}_1 - 0.02\mathbf{u}_2, \text{ and } \mathbf{d}_4 = 0.84\mathbf{u}_1 - 0.02\mathbf{u}_2.$$