ECS509U - Probability & Matrices

Tassos Tombros

Week 10

Week 10: Learning Objectives

- At the end of week 10 you should be able to:
  - calculate the dot product between vectors
  - compute the cosine of the angle between vectors
  - define vector spaces and subspaces (briefly)
  - examine and understand linear combinations of a set of vectors
  - determine the linear dependence and independence between vectors
  - find spanning sets of vectors

Dot (or inner) product

- The **dot** (or **inner**) **product** between two vectors **u** and **v** is denoted as **u·v** and it is **a real number**
  - it is defined only when the vectors are of the same "dimensions", i.e. have the same number of components
- Let us see the definition for 2 and 3 dimensional vectors:
- Let **u** and **v** be vectors in 3-space: **u**=(u_1, u_2, u_3), **v**=(v_1, v_2, v_3)
  - then the dot product $\boxed{\mathbf{u}\cdot\mathbf{v}=u_1v_1+u_2v_2+u_3v_3}$
- Likewise, if **u** and **v** are vectors in 2-space:**u**=(u_1, u_2), **v**=(v_1, v_2)
  - then the dot product $\boxed{\mathbf{u}\cdot\mathbf{v}=u_1v_1+u_2v_2}$
- The definition can generalise for vectors in *n*-space

Examples

- Calculate the dot products below:
  - **a**=(9,-2) and **b**=(4,18)

    a·**b** = (9)(4) + (-2)(18) = 0
  - **u**=(3,-6,6) and **v**=(4,2,0)

    u·**v** = (3)(4) + (-6)(2) + (6)(0) = 0
  - **b** = (0, -1, -2) and **c** = (0, 3, -1)

    b·**c** = (0)(0) + (-1)(3) + (-2)(-1) = -3 + 2 = -1
  - **w**=(2,3,-4) and **a**=(0,-1)

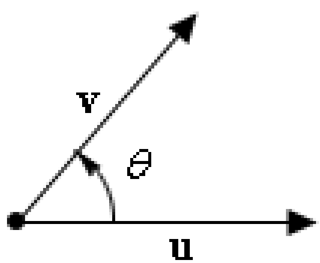
    w·**a** can not be defined because the vectors have a different number of components
- **Note that the dot product is similar to calculating the matrix product of one row vector (e.g. 1xn) and one column vector (nx1)**

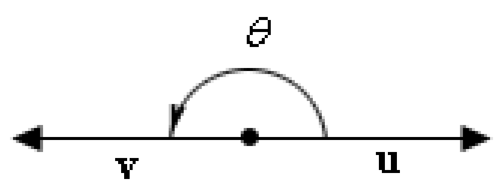
Properties of the dot product

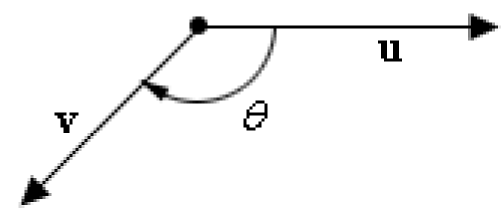
- For any vectors **u**, **v**, **w**, and any scalar *c*, the following hold:
  - $(\mathbf{u}+\mathbf{v})\cdot\mathbf{w} = \mathbf{u}\cdot\mathbf{w} + \mathbf{v}\cdot\mathbf{w}$
  - $(c\mathbf{u})\cdot\mathbf{v} = c(\mathbf{u}\cdot\mathbf{v})$
  - $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$
  - $\mathbf{u}\cdot\mathbf{u}\geq 0$ and $\mathbf{u}\cdot\mathbf{u}=0$ if and only if $\mathbf{u}=\mathbf{0}$

The angle between two vectors

- Suppose that **u**, **v** are vectors in 2-space or 3-space, placed so that their initial point is the same
- Then the angle between **u** and **v** is the **angle θ** such that **$0 \leq \theta \leq \pi$ ($0^o \leq \theta \leq 180^o$)**
- There are always two angles formed by the two vectors - the one we will always choose is the one that satisfies **$0 \leq \theta \leq \pi$**







The angle between two vectors

- We can get a simple formula for the **cosine of the angle** between two vectors as follows, assuming that **u** and **v** are non-zero vectors

$$\cos q = \frac{u \times v}{\|u\|\|v\|}$$

- Notice that the angle between the two vectors has a physical interpretation
  - It shows the **degree of similarity between the two vectors** (i.e. how close they are) - more on this later

Examples

- Consider the following pairs of vectors in 3- and 2-space. What is the angle between the two vectors?

  a) **u**=(3,-1,6), **v**=(4,2,0)   b) **a**=(9,-2), **b**=(4,18)

(a) We know that : $\cos q = \frac{u \times v}{\|u\|\|v\|} = \frac{(3)(4)+(-1)(2)+(6)(0)}{\sqrt{3^2+(-1)^2+6^2}\sqrt{4^2+2^2+0^2}} =$

$$\frac{10}{\sqrt{46}\sqrt{20}} = 0.3296902$$

From this we can calculate that the angle $q = 70.75^{o}$ (note that we need a calculator for this and this is an optional step unless specifically asked by an exercise)

(b) Similarly, we find that $\cos q = \frac{a \times b}{\|a\|\|b\|} = \frac{(9)(4)+(-2)(18)}{\sqrt{85}\sqrt{340}} = \frac{0}{\sqrt{85}\sqrt{340}} = 0$

and from this we know (without calculator) that $q = 90^{o}$

Some important properties

- We (should) know that **-1 ≤ cosθ ≤ 1**
- This holds in the vector case because someone (someone called Schwarz) has proven that:

$$\boxed{-1 \le \frac{u \cdot v}{\|u\|\|v\|} \le 1 \quad \textbf{Schwarz inequality}}$$

- **Orthogonal vectors:** Two non-zero vectors **u**, **v** are called orthogonal if and only if **their dot product is equal to zero: u·v** = 0
  - note that if **u·v** = 0 then cosθ=0 which means that the angle θ between them will be 90º

A different example

- Consider a very very very small language vocabulary with only two words: *big* and *fish*

  Consider two very very very small documents from that language:

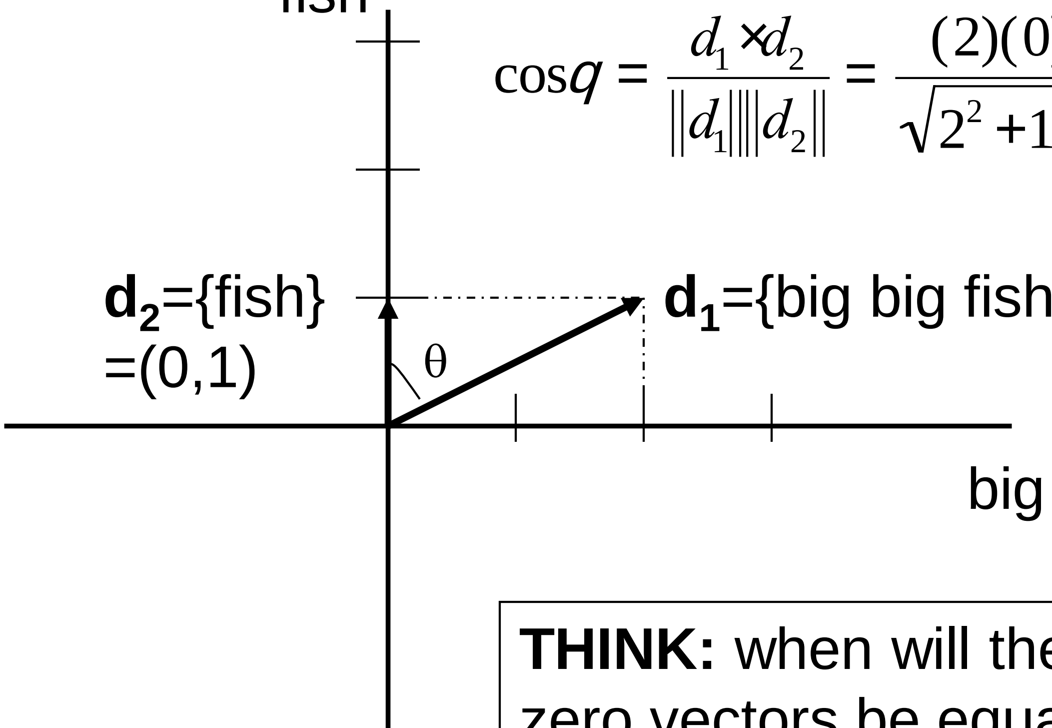
  Document $\mathbf{d_1}$contains: {big big fish} and document $\mathbf{d_2}$ contains: {fish}

  a) Represent the two documents graphically in a two-dimensional space, where the axis are the words of the vocabulary

  b) Find how similar $\mathbf{d_1}$ is to $\mathbf{d_2}$

Example continued

$$\cos\theta = \frac{d_1 \times d_2}{\|d_1\|\|d_2\|} = \frac{(2)(0)+(1)(1)}{\sqrt{2^2+1^2}\sqrt{0^2+1^2}} = \frac{1}{\sqrt{5}\sqrt{1}} = \frac{1}{\sqrt{5}} = 0.4472$$



d_2={fish}
=(0,1)

d_1={big big fish}=(2,1)

So, we can say that
similarity(d_1, d_2)=0.4472

THINK: when will the similarity between two non-zero vectors be equal to zero? Equal to one?

zero: when there is nothing in common between the two vectors

one: when the two vectors are identical

The cosine as a similarity measure

- We can use the cosine of the angle between two vectors to **measure their similarity**
  - **a similarity of 1** will mean that the two vectors are identical
  - **a similarity of 0** will mean that the two vectors have nothing in common (when we only use positive values for components like in the previous example)
- **Important properties** of similarity between objects (similarity(u,v) = $\cos\theta$) :
  - **similarity(u,v) = similarity(v,u)**
  - **similarity(u,u)=1** (a vector is always identical to itself)

The cosine as a similarity measure

- The cosine as a similarity measure is called the **cosine coefficient**
- The cosine coefficient is a very popular **measure of similarity** between objects in multi-dimensional spaces
  - Think of customers of an e-commerce system and their ratings of e.g. films (dimensions: number of films)
  - Think of a set of web pages and the terms they contain (dimensions: number of terms)
  - Many more applications in bioinformatics, image processing and analysis, etc. etc.

The heart of Linear Algebra

- This part of the module goes in the heart of Linear Algebra, and will eventually draw the link between matrices and vectors
- We will talk about:
  - Vector spaces and subspaces (**only briefly**)
  - **Linear combinations of vectors**
  - **Linear dependence and independence** between vectors
  - **Spanning sets**, **basis** and **dimension** of vector spaces and subspaces, and **changing the basis** of spaces
- We will start today and conclude this material in Week 11

Vector spaces

- What are **Vector Spaces**?
  - We have already seen some in our examples: $\mathbf{R}^2$ and $\mathbf{R}^3$
  - In fact, the most important vector spaces are denoted by $\mathbf{R}^1$, $\mathbf{R}^2$, $\mathbf{R}^3$, …. $\mathbf{R}^n$ - there is one for every positive integer
  - For example the space $\mathbf{R}^2$ consists of all vectors on the *x*-*y* plane where the components of the vectors are the *x*-*y* coordinates
- So vector spaces are spaces where we have a set of vectors, and where we can:
  - **add** any two vectors together
  - **multiply** vectors **by scalars** (real numbers)

  and where the result will still be a vector in the same space
- So, **vector spaces are closed under addition and scalar multiplication**

Vector spaces

- We know how to do these operations, we learnt this in Week 9
- Formally, a real vector space is a set of vectors as defined previously, **together with rules for vector addition and scalar multiplication** (we saw these rules in week 9)
- Normally our vectors will belong in one of the vectors $\mathbf{R}^n$
  - we will start seeing them as **column vectors**
- However, the formal definition of vector spaces allows us to see other 'things' as vectors
  - for example, the space of any $n \times m$ matrices

Vector subspaces

- A **vector subspace** of a vector space, is a nonempty subset of the vector space that satisfies two requirements:
  - **If we add any two vectors** in the subspace, their sum is still in the subspace
  - **If we multiply any vector in the subspace by a scalar**, the result is still in the subspace
- In other words, the subspace is **a subset of the vector space** that is **also"closed" under addition and scalar multiplication**

Linear combinations of vectors

- We say that a vector **w** in some vector space *V* is **a linear combination** of the vectors $\mathbf{v_1}$, $\mathbf{v_2}$, …, $\mathbf{v_n}$, all from the same vector space *V*, if there are scalars c_1, c_2, .., c_n so that **w** can be written as:

  $\mathbf{w} = c_1\mathbf{v_1}+c_2\mathbf{v_2}+\ldots+c_n\mathbf{v_n}$

  Note that in linear combinations the operations involved are **vector addition** and **scalar multiplication**
- You can think of this as 'mixing' vectors $\mathbf{v_1}$, $\mathbf{v_2}$, …, $\mathbf{v_n}$ in the 'right proportions' in order to get vector **w**
  - the 'right proportions' are given by the values of the scalars c_1, c_2, .., c_n

Example

- Is **w**=(-12,20) a linear combination of $\mathbf{v_1}$=(-1,2) and $\mathbf{v_2}$=(4,-6)?

If it is a linear combination, then it means that it can be expressed as: w = $c_1\mathbf{v_1}+c_2\mathbf{v_2}$, so

$(-12,20)=c_1(-1,2) + c_2(4,-6)$

We then arrive at a 2x2 system of equations:

$-c_1 + 4c_2 = -12$

$2c_1 - 6c_2 = 20$

If the system is consistent (i.e. it has at least one solution) then **w** can be expressed as a linear combination of $\mathbf{v_1}$ and $\mathbf{v_2}$

You can verify that the system has a unique solution: c_1=4, c_2=-2, so **w** is a linear combination of $\mathbf{v_1}$ and $\mathbf{v_2}$

Another view of the same example

- We can express the linear combination from the previous example also in terms of column vectors:

$$\begin{bmatrix} -1 \\ 2 \end{bmatrix} c_1 + \begin{bmatrix} 4 \\ -6 \end{bmatrix} c_2 = \begin{bmatrix} -12 \\ 20 \end{bmatrix} \xrightarrow{\text{Equivalent to:}} \boxed{\begin{aligned} -c_1 + 4c_2 &= -12 \\ 2c_1 - 6c_2 &= 20 \end{aligned}}$$

- This is another way to express systems of linear equations, of **any** size
- We will use this notation quite a lot from now on

Spanning sets

- Now assume that we have a set of vectors $S=\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ in a vector space, and we **take all of their possible linear combinations**
- The set of all the linear combinations W is called **the spanning set of *S***: $W=span(S)$, or $W=span(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n})$

  We can also say that vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ **span** W

Example

- Describe the span of the following set of vectors:

  $\mathbf{v_1}$=(1,0,0) and $\mathbf{v_2}$=(0,1,0)

The span of this set of vectors, is the set of all their linear combinations, and we can write down a general linear combination for these two vectors as:

$a\mathbf{v_1}+b\mathbf{v_2} = (a,0,0)+(0,b,0) = (a,b,0)$

So it looks like the span of $\mathbf{v_1}$ and $\mathbf{v_2}$ will be the set of vectors in $\mathbf{R}^3$ of the form $(a,b,0)$ for **any** values of *a*, *b* (for any real numbers a, b)

Spanning sets example no. 2

- Consider the following three vectors: $\mathbf{v_1}$=(1,2,-1), $\mathbf{v_2}$=(3,-1,1), and $\mathbf{v_3}$=(-3,8,-5). Determine if these three vectors will span $\mathbf{R^3}$.

If they span $\mathbf{R^3}$, then if we pick a random vector **u** in $\mathbf{R^3}$, **u**=(a,b,c) it can be expressed as a linear combination of $\mathbf{v_1}$, $\mathbf{v_2}$, $\mathbf{v_3}$. So essentially we need to determine whether we can find scalars c_1, c_2, c_3 such that: $\mathbf{u} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + c_3\mathbf{v_3} \Rightarrow (a,b,c) = c_1(1,2,-1) + c_2(3,-1,1) + c_3(-3,8,-5) \Rightarrow (a,b,c) = (c_1, 2c_1,-c_1) + (3c_2,-c_2, c_2) + (-3 c_3, 8c_3, -5c_3) \Rightarrow (a,b,c)=(c_1+3c_2-3c_3, 2c_1-c_2+8c_3, -c_1+c_2-5c_3)$. From here we get a 3x3 system of equations, where the unknowns are c_1, c_2, c_3:

$$\left.\begin{aligned} c_1 + 3c_2 - 3c_3 &= a \\ 2c_1 - c_2 + 8c_3 &= b \\ -c_1 + c_2 - 5c_3 &= c \end{aligned}\right\} \begin{bmatrix} 1 & 3 & -3 \\ 2 & -1 & 8 \\ -1 & 1 & -5 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

If we calculate the determinant of the coefficient matrix in this system, we will find that it is zero, so the system will not have a unique solution.

Spanning sets example no. 2 cntd.

So our system can have infinite solutions, or it can have no solutions. But we want to know that there will be a solution for EVERY choice of (a,b,c) on the right hand side. However with the determinant at zero, we now know that there will be at least one combination of (a,b,c) that will make our system unsolvable. Let's see how far we can get with Gauss, trying to solve the system via the augmented matrix:

$$\begin{bmatrix} 1 & 3 & -3 & a \\ 2 & -1 & 8 & b \\ -1 & 1 & -5 & c \end{bmatrix} \xrightarrow{\text{row operations ...}} \begin{bmatrix} 1 & 3 & -3 & a \\ 0 & 1 & -2 & b+2c \\ 0 & 0 & 0 & a-4b-7c \end{bmatrix}$$

Look at the last row of the matrix : it says that $0c_1+0c_2+0c_3= a-4b-7c$.
So, if there is even one set of a,b,c that makes a-4b-7c be not equal to zero, the system will have no solution (because the last row will say 0=not-zero). We can clearly find such a set of a,b,c (e.g. 1, -2, 1). So there is at least one set of a,b,c (i.e. there is at least one vector in $\mathbf{R}^3$) that makes the system not solvable (i.e. that it is not a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$). So $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ DO NOT span $\mathbf{R}^3$

Linear (in)dependence

- Suppose $S=\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$ is a non-empty set of vectors, and form the equation ($c_1, \dots, c_n$ are scalars):

  $c_1\mathbf{v_1}+c_2\mathbf{v_2}+\dots+c_n\mathbf{v_n}= \mathbf{0}$

  This equation will always have at least one solution (the trivial solution for $c_1=c_2=\dots=c_n=0$)
- **If the trivial solution is the only solution** to this equation, then the vectors in the set *S* are called **linearly independent** and the set is called a linearly independent set
- **If there is another solution**, then the vectors in the set *S* are called **linearly dependent** and the set is called a linearly dependent set

Example

- Determine if the following sets of vectors are linearly dependent or independent: **u**=(3, -1) and **v**=(-2, 2)

We take $c_1\mathbf{u}+c_2\mathbf{v}=\mathbf{0}$, which means:

$(3c_1 - 2c_2 , -c_1 + 2c_2)=\mathbf{0}$, where the right part is the zero vector, so $(3c_1 - 2c_2 , -c_1 + 2c_2) = (0,0)$

from where we get a 2x2 homogeneous system of equations:

$3c_1 - 2c_2 = 0$

$-c_1 + 2c_2 = 0$

Solving this (do it as an exercise), will give you only the unique solution c_1=0, c_2=0 (the trivial solution), and so the two vectors are linearly independent

Example

- Do the same for the vectors:
  $\mathbf{v_1}$=(2,-2,4), $\mathbf{v_2}$=(3,-5,4), $\mathbf{v_3}$=(0,1,1)

We follow the same procedure as in the previous example, and reach the following expression:

$$\begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix} c_1 + \begin{bmatrix} 3 \\ -5 \\ 4 \end{bmatrix} c_2 + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} c_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ Solving this we will get } c_1 = -3/4t, c_2 = 1/2t, c_3 = t \text{ where}$$

t is any real number

We have more than the trivial solution here, so the three vectors are linearly dependent

Notice that if we choose *t*=1 then we have: -3/4$\mathbf{v_1}$+1/2$\mathbf{v_2}$+$\mathbf{v_3}$=0 from where we get for example that: $\mathbf{v_3}$ = 3/4$\mathbf{v_1}$-1/2$\mathbf{v_2}$. The same could be done for any *t*

So, any one of the dependent vectors can be expressed as a linear combination of the others (i.e. we could also express $\mathbf{v_2}$ in terms of $\mathbf{v_1}$ and $\mathbf{v_3}$, etc.)

You could also just find the determinant of the coefficient matrix = 0.

Three important points

- A finite set of vectors that contains the zero vector will be linearly dependent
- Suppose that $S=\{\mathbf{v}_1, \mathbf{v}_2, .., \mathbf{v}_k\}$ is a set of vectors in $\mathbf{R}^n$. If $k>n$, then the vectors in S will be linearly dependent
  - for example, any 3 vectors in $\mathbf{R}^2$ will be linearly dependent, any 4 in $\mathbf{R}^3$, etc.
- If a set of vectors is linearly dependent, it means that at least one of the vectors can be expressed as a linear combination of the others (see previous example) - **extremely important !!!**

Summary of lecture

- In Week 10 we covered:
  - Dot product between vectors, angle between vectors
  - Vector spaces (very briefly)
  - Linear combinations of vectors, spans of sets of vectors
  - Linear dependence and independence
- For Friday's tutorial:
  - Come to the tutorial having attempted the week's exercises