

Always include this title page with your PDF. Include your name above.

- Submit your work in Gradescope as a PDF - you will identify where your "questions are."
- Identify the question number as you submit. Since we grade "blind" if the questions are NOT identified, the work WILL NOT BE GRADED and a 0 will be recorded. Always leave enough time to identify the questions when submitting.
- One section per page (if a page or less) - We prefer to grade the main solution in a single page, extra work can be included on the following page.
- Long instructions may be removed to fit on a single page.
- **Do not start a new question in the middle of a page.**
- Solutions to book questions are provided for reference.
- You may NOT submit given solutions - this includes minor modifications - as your own.
- Solutions that do not show individual engagement with the solutions will be marked as no credit and can be considered a violation of honor code.
- If you use the given solutions you must reference or explain how you used them, in particular...

For full credit, EACH book exercise in the Study Guides must use one or more of the following methods and FOR EACH QUESTION. Identify the number the method by number to ensure full credit.

Method 1 - Provide original examples which demonstrate the ideas of the exercise in addition to your solution.

Method 2 - Include and discuss the specific topics needed from the chapter and how they relate to the question.

Method 3 - Include original Python code, of reasonable length (as screenshot or text) to show how the topic or concept was explored.

Method 4 - Expand the given solution in a significant way, with additional steps and comments. All steps are justified. This is a good method for a proof for which you are only given a basic outline.

Method 5 - Attempt the exercise without looking at the solution and then the solution is used to check work. Words are used to describe the results.

Method 6 - Provide an analysis of the strategies used to understand the exercise, describing in detail what was challenging, who helped you or what resources were used. The process of understanding is described.

1. (20pts) Select one page or section of Chapter Three of VMLS to annotate. Include a screenshot of your annotation here.

Chapter 3

Norm and distance

cannot use pythagorean here cus not right angle
but in lecture, it was said that $c = \text{norm of } a + b \dots$ I wonder why that is?

In this chapter we focus on the norm of a vector, a measure of its magnitude, and on related concepts like distance, angle, standard deviation, and correlation.

3.1 Norm

The *Euclidean norm* of an n -vector x (named after the Greek mathematician Euclid), denoted $\|x\|$, is the squareroot of the sum of the squares of its elements,

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \quad \sqrt{a^2 + b^2} = c^2$$

The Euclidean norm can also be expressed as the squareroot of the inner product of the vector with itself, i.e., $\|x\| = \sqrt{x^T x}$.

The Euclidean norm is sometimes written with a subscript 2, as $\|x\|_2$. (The subscript 2 indicates that the entries of x are raised to the second power.) Other less widely used terms for the Euclidean norm of a vector are the *magnitude*, or *length*, of a vector. (The term *length* should be avoided, since it is also often used to refer to the dimension of the vector.) We use the same notation for the norms of vectors of different dimensions.

As simple examples, we have

$$\left\| \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix} \right\| = \sqrt{9} = 3, \quad \left\| \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\| = 1.$$

When x is a scalar, i.e., a 1-vector, the Euclidean norm is the same as the absolute value of x . Indeed, the Euclidean norm can be considered a generalization or extension of the absolute value or magnitude, that applies to vectors. The double bar notation is meant to suggest this. Like the absolute value of a number, the norm of a vector is a (numerical) measure of its magnitude. We say a vector is *small* if its norm is a small number, and we say it is *large* if its norm is a large number. (The numerical values of the norm that qualify for *small* or *large* depend on the particular application and context.)

reminds me of the pythagorean theorem!

you can think of it in terms of pythagorean because the euclidean norm calculates the length of vector in norm euclidean space so!



unclear

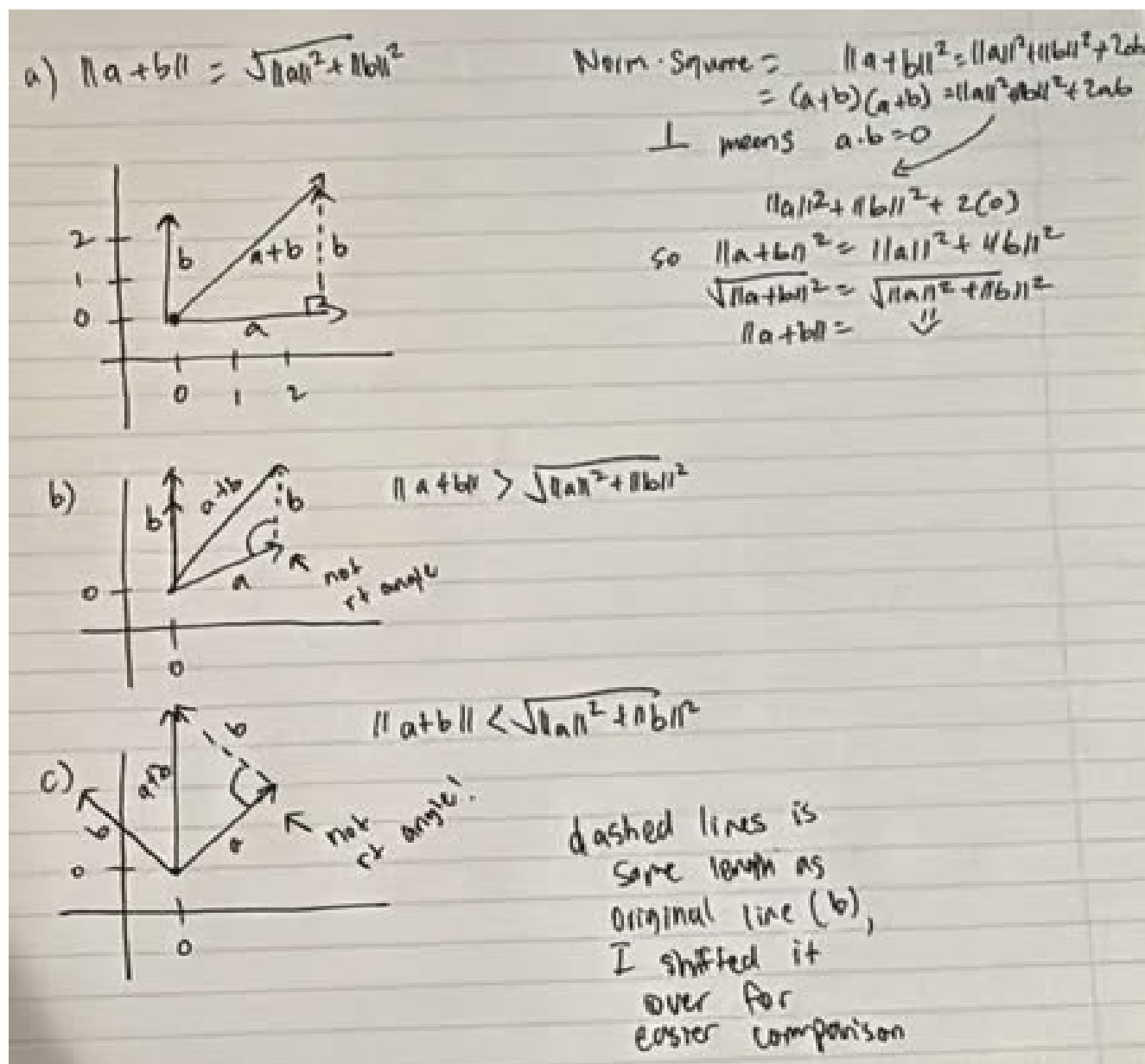
$$\sqrt{x^2} = |x|$$

2. (20 pts) Solve the Chapter 3 Random exercise from the video and Piazza in your own words here.

3.19 Norm of sum. Use the formulas (3.1) and (3.6) to show the following:

- (a) $a \perp b$ if and only if $\|a+b\| = \sqrt{\|a\|^2 + \|b\|^2}$.
 (b) Nonzero vectors a and b make an acute angle if and only if $\|a+b\| > \sqrt{\|a\|^2 + \|b\|^2}$.
 (c) Nonzero vectors a and b make an obtuse angle if and only if $\|a+b\| < \sqrt{\|a\|^2 + \|b\|^2}$.

Draw a picture illustrating each case in 2-D.



I drew the dashed line to show where "b" is supposed to match up with the endpoint for the vectors "a+b" added together. This is to better visualize that "b" meets with "a+b", but "b" might not actually be at that point. "b" can start at the origin (0,0) as well. I shifted the origin up so that it's easier to see problem a).

3. (20 pts) Explain the solution to 3.1 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

3.1 Distance between Boolean vectors. Suppose that x and y are Boolean n -vectors, which means that each of their entries is either 0 or 1. What is their distance $\|x - y\|$?

There are 4 possibilities. When $x = 0, 1$ and when $y = 0, 1$.

Hamming distance between two binary vectors is the # of positions where their corresponding elements are different. x and y are boolean vectors, which means the # of differing positions is based off of the mismatch between the two.

$x = 0, 1$
 $y = 0, 1$

max distance $1 \rightarrow -1 = 2$

$\|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$

$x_i = 1$	$y_i = 0 \Rightarrow x_i - y_i = 1$
$= 0$	$= 1 \Rightarrow -1$
$= 1$	$= 1 \Rightarrow 0$
$= 0$	$= 0 \Rightarrow 0$

$(x_i - y_i)^2 = \begin{cases} 1 & \text{if } x_i \neq y_i \leftarrow \text{diff position} \\ 0 & \text{if } x_i = y_i \leftarrow \text{same position} \end{cases}$

sum has to be

$\sum_{i=1}^n (x_i - y_i)^2$ where $x_i \neq y_i$

$\|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$

$\|x - y\| = \sqrt{(\# \text{ of differing positions})}$

also known as Hamming distance

4. (ungraded 0 pts) Explain the solution to 3.2 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

5. (ungraded 0 pts) Explain the solution to 3.4 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

6. (20pts) Explain the solution to 3.10 here in your own words.

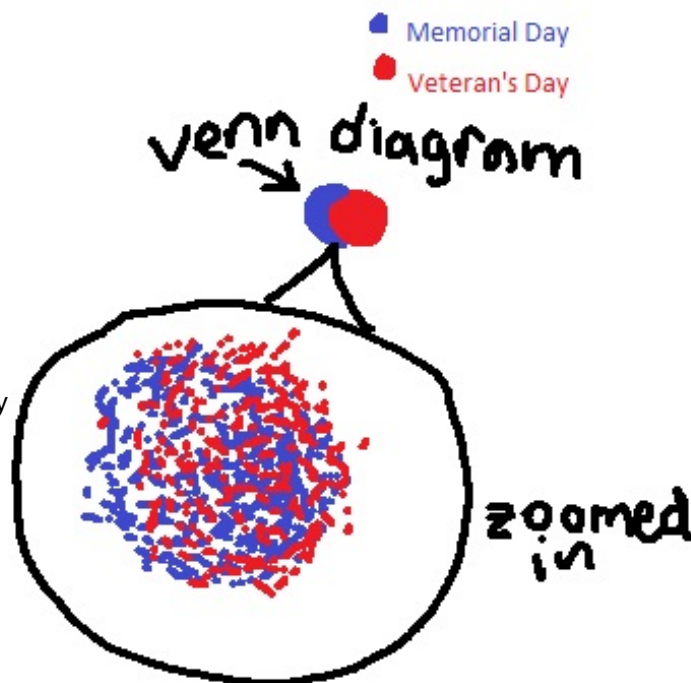
3.10 Nearest neighbor document. Consider the 5 Wikipedia pages in table 3.1 on page 51. What is the nearest neighbor of (the word count histogram vector of) 'Veterans Day' among the others? Does the answer make sense?

	Veterans Day	Memorial Day	Academy Awards	Golden Globe Awards	Super Bowl
Veterans Day	0	0.095	0.130	0.153	0.170
Memorial Day	0.095	0	0.122	0.147	0.164
Academy A.	0.130	0.122	0	0.108	0.164
Golden Globe A.	0.153	0.147	0.108	0	0.181
Super Bowl	0.170	0.164	0.164	0.181	0

Table 3.1 Pairwise word count histogram distances between five Wikipedia articles.

The venn diagrams overlap by 90.5% ($1 - 0.095$) ----->

I'm pretty sure 0.095 is how much the document deviates from each other



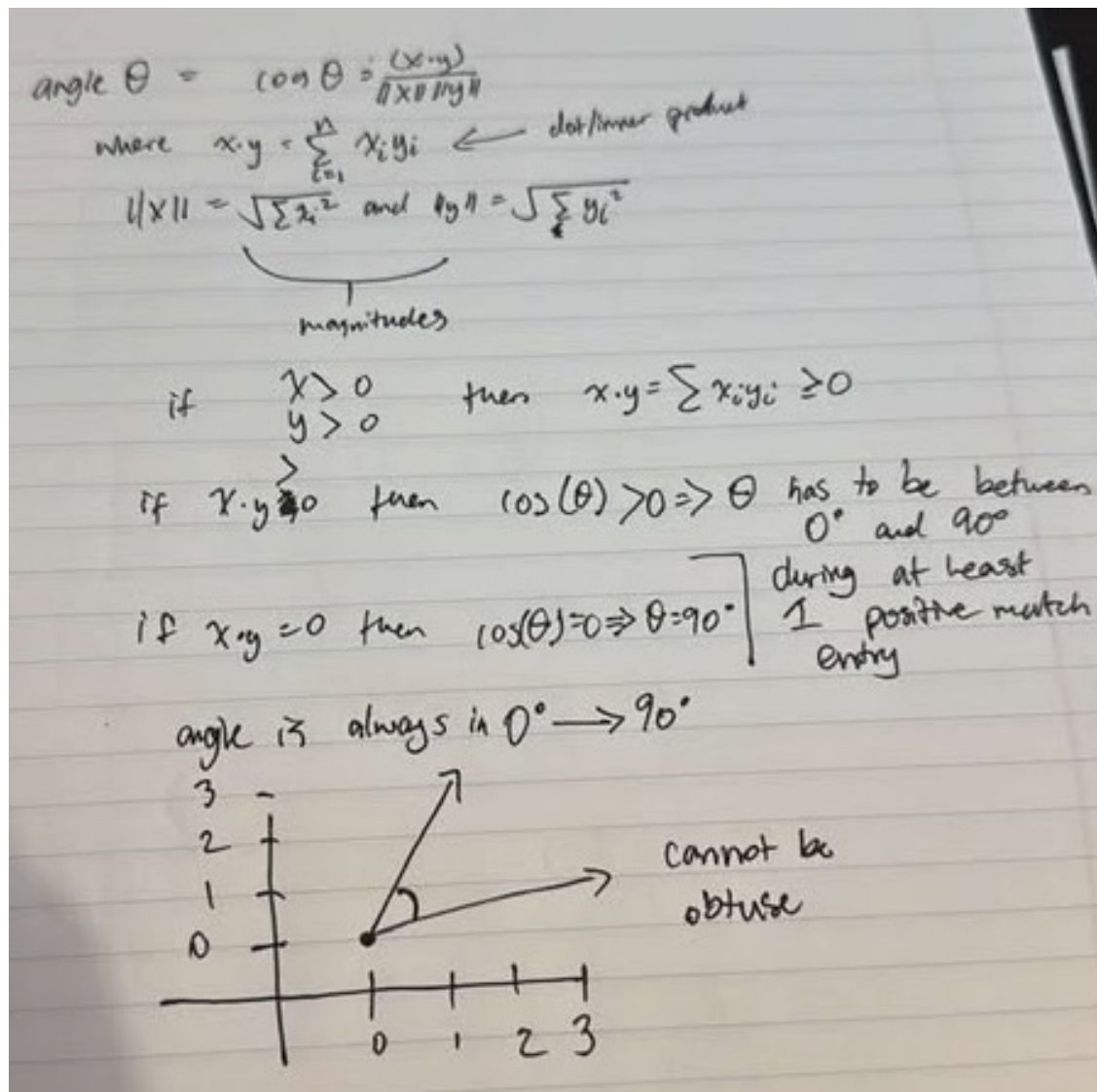
"Nearest neighbor" makes me think of "K-nearest neighbor" which makes me think of clustering algorithms and how close datapoints or nodes in a vector space are together. Usually these similarities are measured by contents or what each node represents. If each word was a node, and both Veteran's day and Memorial day used a ton of shared words, that would make sense because both documents probably are talking about honoring the military's accomplishments. I picture Veteran's day as a round cluster, and Memorial day as another round cluster, and both overlap a ton (like a venn diagram).

The word frequency of the two histograms are more similar than others because Veteran's day probably has little connection with the Super Bowl. But maybe a Super Bowl ad came up and was like "Thank you America for hosting such an amazing sport event!" and the tokens "Thank you" and "America" would be overlapping with Veteran's day (but the other tokens might not be).

7. (ungraded 0 pts) Explain the solution to 3.20 here in your own words. (Since you are given a solution, you will be graded on your ability to explain) Show and justify all steps.

8. (20 pts) Explain the solution to 3.23 here in your own words. (Since you are given a solution, you will be graded on your ability to explain)

3.23 *Angle between two nonnegative vectors.* Let x and y be two nonzero n -vectors with nonnegative entries, i.e., each $x_i \geq 0$ and each $y_i \geq 0$. Show that the angle between x and y lies between 0 and 90° . Draw a picture for the case when $n = 2$, and give a short geometric explanation. When are x and y orthogonal?



$$x \cdot y = 0 \Rightarrow x_i y_i = 0 \text{ for all } i$$

the angle theta between 2 non negative vectors is always going to be in the top right quadrant between 0 degrees and 90

9. (Ungraded Bonus - not exam question)

Explain the solution to 3.24 here in your own words.

- Create your own example for part a)
- Justify each of the steps in the proof.