VMLS - Study Guide Chapter 2 Name: Alex Cullen

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 r 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. (10 pts) Select one page or section of Chapter One of VMLS to annotate. Include a screenshot of your annotation here.

Scalar-Valued Function means the output of the function is a real number

airplane, as a function of the 3-vector x, where x_1 is the angle of attack of the airplane (*i.e.*, the angle between the airplane body and its direction of motion), x_2 is its air speed, and x_3 is the air density.

The inner product function. Suppose a is an n-vector. We can define a scalar-valued function f of n-vectors, given by

$$f(x) = a^T x = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n \qquad (2.1)$$

Inner Product:
a = n-vector
x = same size n-vector

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$
a x

$a^T x = ax + by + cz$

for any n-vector x. This function gives the inner product of its n-vector argument x with some (fixed) n-vector a. We can also think of f as forming a weighted sum of the elements of x; the elements of a give the weights used in forming the weighted sum.

Superposition and linearity. The inner product function f defined in (2.1) satisfies the property

$$\begin{aligned} f(\alpha x + \beta y) &= a^T(\alpha x + \beta y) && \text{Substitute } \alpha x + \beta y \text{ for } x \\ &= a^T(\alpha x) + a^T(\beta y) && \text{Distribute} \\ &= \alpha(a^T x) + \beta(a^T y) && \text{Associative prop} \\ &= \alpha f(x) + \beta f(y) && \text{Substitute } f(x) \text{ for } a^T x \end{aligned}$$

for all n-vectors x, y, and all scalars α, β. This property is called *superposition*. A function that satisfies the superposition property is called *linear*. We have just shown that the inner product with a fixed vector is a linear function.

α, β = scalars
x, y = n-vectors

The superposition equality

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) \qquad (2.2)$$

looks deceptively simple; it is easy to read it as just a re-arrangement of the parentheses and the order of a few terms. But in fact it says a lot. On the left-hand side, the term $\alpha x + \beta y$ involves *scalar-vector* multiplication and *vector addition*. On the right-hand side, $\alpha f(x) + \beta f(y)$ involves ordinary *scalar multiplication* and *scalar addition*.

f(x) = scalar
αx = scalar-vector product
x+y = vector add
α·f(x) = scalar · scalar
f(x)+f(y) = scalar + scalar

If a function f is linear, superposition extends to linear combinations of any number of vectors, and not just linear combinations of two vectors: We have

$$f(\alpha_1 x_1 + \cdots + \alpha_k x_k) = \alpha_1 f(x_1) + \cdots + \alpha_k f(x_k),$$

for any n vectors $x_1, \ldots, x_k$, and any scalars $\alpha_1, \ldots, \alpha_k$. (This more general k-term form of superposition reduces to the two-term form given above when $k = 2$.) To see this, we note that

$$\begin{aligned} f(\alpha_1 x_1 + \cdots + \alpha_k x_k) &= \alpha_1 f(x_1) + f(\alpha_2 x_2 + \cdots + \alpha_k x_k) \\ &= \alpha_1 f(x_1) + \alpha_2 f(x_2) + f(\alpha_3 x_3 + \cdots + \alpha_k x_k) \\ &\;\;\vdots \\ &= \alpha_1 f(x_1) + \cdots + \alpha_k f(x_k). \end{aligned}$$

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$
α x_1 x_k

$f(x) = a^T x$

$f(\alpha_1 x_1 + \alpha_2 x_2) = a^T(\alpha_1 x_1 + \alpha_2 x_2)$
$= a^T \alpha_1 x_1 + a^T \alpha_2 x_2$
$= \alpha_1 f(x_1) + \alpha_2 f(x_2)$

2. (10 pts) Solve the Chapter 2 Random exercise from the video and Piazza in your own words here.

Q2: Random Exercise (problem 2.7)

Approach - Method 5

I am going to attempt to solve this problem on my own and then review the solution from the instructor and reflect on where I was right and wrong.

General forumale for affine functions. Verify that formula 2.4 holds for any affine function $f : R^n -> R$

Formula (2.4) page 33:

$f(x) = f(0) + x_1(f(e_1 - f(0)) + ... + x_n(f(e_n - f(0))$

My attempt:

Becuase we know that f(x) is affine $f(x) = ax + b$ for some n-vector a and scalar b. ax can also be described as the inner product of n-vectors a and x or $f(x) = a^T x + b$

The inner product of a and x can be expressed as $a_1 x_i + ... + a_k x_k$

e_i is a unit vector.

when x is equal to the unit vector e_i, then:

$f(e_i) = a^T * e_i + b$

the inner product of a and e_i is a_i

$f(e_i) = a_i + b$

subtract b from both sides to derive the value of a_{i} at a point

$f(e_i) - b = a_i$

when x is 0, f(x) = b

$f(0) = a^T * (0) + b = b$

substiting a_{i} and b in the original equation of f(x):

$f(x) = a^T x + b$

$f(x) = b + \sum_{i=1}^{n} a_i * x_i$

$b = f(0), a_i = f(e_i) - b$

$f(x) = f(0) + \sum_{i=1}^{n}((f(e_i) - b) * x_i$

$b = f(0), a_i = f(e_i) - b$

$f(x) = f(0) + \sum_{i=1}^{n}((f(e_i) - b) * x_i$

$f(x) = f(0) + \sum_{i=1}^{n}((f(e_i) - f(0)) * x_i$

Hence, we can see that $f(0) + \sum_{i=1}^{n}((f(e_i) - f(0)) * x_i = a^T x + b$

3. (20 pts) Explain the solution to 2.1 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

Q3: Problem 2.1

Approach - Method 5

Determine whether each of the following scalar-valued functions of n-vectors is linear. Give inner product representation (n-vector a for which $f(x) = a^T x$))

6]:
```
#Part A
```

The spread of values of the vector defined as $f(x) = max_k x_k - min_k x_k$

if f(x) is linear then $f(max_k x_k - min_k x_k) = a^T x$ when $x = max_k x_k - min_k x_k$, we will prove homogenity and additivity through inner product representation.

$a^T x = a^T (max_k x_k - min_k x_k)$

$= a^T max_k x_k - a^T min_k x_k$

$= a^T x_k max_k - a^T x_k mink$

$= f(x_k) max_k - f(x_k) min_k$

My original answer: f(x) is linear

Actual soluiont: f(x) is not linear when x = (1, 0), y = (0, 1), α = β = 1/2. Then αx + βy = (1/2, 1/2), f(x) = f(y) = 1, and

$$f(\alpha x + \beta y) = 06 = \alpha f(x) + \beta f(y) = 1.$$

Reflection after reviewing the solution:

Going into this problem, I was not properly understanding the implication of the max and min functions and how they affect the output of f(x). I thought the original equation was scaling vector x by the maximum value of vectorx, which is not what the equation was stating. I was not familiar with this notation previously. I can see now that the max(*ax* = *by*) is not equal to a*(max(x)) + b*(max(y)).

In [1]:
```
#Part B
```

I will attempt to approve homogeneity and additivity for f(x): $f(x) = x_n - x_1 = f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$

$= (\alpha x_n + \beta y_n) - (\alpha x_1 + \beta y_1)$ <-- substitute $\alpha x + \beta y$ for x

$= (\alpha x_n - \alpha x_1) + (\beta y_n - \beta y_1)$ <-- rearrage through associativity and commutatitivity of addition

$= \alpha(x - n - x_1) + \beta(y_n - y_1)$ <-- distributive property

$\alpha f(x) + \beta f(y)$ <-- closure f(x) substituted for x_n - x_1, f(y) substituted for y_n - y_1

My original answer: $\alpha f(x) + \beta f(y) = f(\alpha x + \beta y)$**, hence f(x) is linear**

Give an inner product representation for which $f(x) = a^T x$ for all x.

because f(x) = x_n - x_1, a must have a value for a_n and a_1 and sparse everywhere else.

a = [-1,0 ... 0,1] for any size of n.

Reflection after viewing answer:

My solution was correct!

In [2]:
```
#Part C
```

We will attempt to prove homogeneity and additivity for f(x) where:

$f(x) = x_{k+1} = f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$

$= \alpha x_{k+1} + \beta y_{k+1}$ <-- substitute \alpha x + \beta y for x_{k+1}

$= \alpha f(x) + \beta f(y)$ <- closure f(x) and f(y) for x_{k+1} and y_{k+1}

My original answer: $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$ **when f(x) is the median value of vector x, hence f(x) is linear**

Actual solution:

Solution: Not linear. Choose x = (–1, 0, 2), y = (2, –1, 0), α = β = 1. Then αx + βy = (1, –1, 2), f(x) = f(y) = 0, and f(αx + βy) = 1 6= αf(x) + βf(y) = 0.

My reflection:

I assumed that the vectors were sorted in ascending order in my initial derivation, which is not a valid assumption, therefore violating additivity. f(x + y) != f(x) + f(y) when x and y are not already pre-sorted in ascending or descending order.

In [3]:
```
#Part D
```

I will attempt to prove that f(x) is linear by verifying f(x) demonstrates additivity and homogeneity:

$f(x) = avg(x_{i+1}) - avg(x_i) = f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$

I'm going to use x_{i+1} and x_i as shorthand to represent the average of odd values and average of even values over a vector.

$= average(\alpha x_{i+1} + \beta y_{i+1}) - average(x_i + \beta y_i)$ <--- substitute alpha x and beta y for x in f(x)

$= average(\alpha x_{i+1}) + average(\beta y_{i+1}) - average(\alpha x_i)$
$- average(\beta y_i)$
<- distribute average function

$= \alpha * average(x_{i+1}) - \alpha * average(x_i) + \beta * average(y_{i+1}) - \beta$
$* average(y_i)$
<-- commutatitivity and associativity, rearranging terms

$= \alpha * (average(x_{i+1}) - average(x_i)) + \beta$
$* (average(y_{i+1}) - average(y_i))$
<--- distribute out alpha and beta

$= \alpha f(x) + \beta f(y)$ <-- closure f(x) for average x_{i+1} - x_{i}

My original solution:

hence, both additivity and homogeneity hold for f(x), f(x) is linear.

the inner product representation of f(x) = sum of each element of x divided by size of x, thus

f(x) = a^Tx = sum(a_i * x_i) = sum(x)/n

a_i = 1/n

a = (1/n,1/n...1/n)

Reflection after viewing the solution:

I was correct that the function is linear, but I was not correct on the inner product representation of a. Because we are subtracting even indices averages from odd indices averages, then the sign of each odd indexed element should be positive, and the sign of each even indexed element should be negative.

odd indices averages, then the sign of each odd indexed element should be positive, and the sign of each even indexed element should be negative.

In [4]:
```
#Part E
```

I will attempt to prove that f(x) is linear by verifying f(x) demonstrates additivity and homogeneity:

$f(x) = x_n + (x_n - x_n - 1) = f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$

$= \alpha x_n + \beta y_n + (\alpha x_n + \beta y_n - \alpha x_{n-1} + \beta y_{n-1})$ <-- substitute \alpha x + \beta y for x in f(x)

$= \alpha x_n + \alpha x_n - \alpha x_{n-1} + \beta y_n + \beta y_n - \beta y_{n-1}$ <-- rearrange terms through commutatitivity and associativity of addition

$= \alpha(x_n + x_n - x_{n-1}) + \beta(y_n + y_n + y_{n-1})$ <-- distribute alpha and beta out of like terms

my original answer $= \alpha f(x) + \beta f(y)$**, hence the function f(x) is linear**

the inner product representation of a = (0,0,.. -1, 2). This is because we are adding two of the last element and subtracting one of the second to last element

Reflection after viewing solution:

I was correct!

Q4: Problem 2.4

4. (20 pts) Explain the solution to 2.4 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

Q4: Problem 2.4

Approach - Method 5:

I will attempt this myself and then review the solution

Given: $f(1,1,0) = -1, f(-1,1,1) = 1, f(1,-1,-1) = 1$

choose whether the function is linear, could be linear, or cannot be linear.

Hypothesis: if f(x) is linear then $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$ by homogeneity and additivity

f(x + y + z) must equal f(x) + f(y) + f(z) if the function is linear, where x, y, and z represent the three given vectors given above:

$$f((1,1,0) + (-1,1,1) + (1,-1,-1)) = f(1,1,0) \\ + f(-1,1,1) + f(1,-1,-1)$$

However, we know that x + y + z = (1, 1, 0) = f(1,1,0) = -1, and that f(x) + f(y) + f(z) = -1 + 1 +1 = 1.

This is a contradiction! and f(x + y + z) != f(x) + f(y) +f(z) $f(1,1,0)! = 1$

My original solution: ¶

Hence, this function cannot be linear.

Reflection after reviewing solution:

The book took a different approach to solving the solution, in that, in a linear function, the inverse value of an input should produce and inverse output, however, that is not the case for f(y) and f(z) given above where y = -1, 1, 1 and z = 1, -1, -1, and both outputs = 1

Q5: Problem 2.6

5. (20 pts) Explain the solution to 2.6 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

Q5: Problem 2.6

Approach - Method 5

$s = w^T a + v$

scores in a are ordinal (1, 2, 3), each value representing a contribution to the score of (0, 1, 2) and (0, 2, 4) for questions 1-15 and 16-30 respectively

for questions <= 15, score is just equal to the (ordinal value - 1), for questions greater than 15, the score is weighted by 2 * (ordinal value -1)

elements 1-15 or vector w, where w is a 30-vector, equal 1, and elements 15-30 of vector w is equal 2

scalar v is equal to 0, as the score is not offset by any constant value

My original answer to problem 2.6

$s = [1\ldots_{15}2\ldots_{15}]^T a + 0$

Reflection after reviewing the solution

I missed a step in my solution, where I forgot to continue the derivation for a and left is as a =(ordinal value - 1).

To continue that would require that the inner product of vector w with a vector a where a_i = (ordinal value - 1) = the inner product of vector w with vector a - the inner product of vector w with the inner product of vector -1 ~ (w^T-1). (w^T-1) = -45. Therefore, $s = [1_{15}\, 2_{15}]^T a - 45$

6. (20 pts)

a) Redo the proof on page 30 that an inner product satisfies superposition with 2 vectors of your own choosing (using length 3).

b) Is this still a proof?

c) Why is it useful?

Question 6

Redo proof on page 30

On page 30, the function has 3 vectors (a, x, y) , however this question only asks to specify two of the vectors. I took this to mean that we should choose values for vectors x and y.

$f(\alpha x + \beta y) = a^T(\alpha x + \beta y)$

$f(x) = a^T x$

x = [1, 2, 3], y = [-1, 0, 1] alpha = 1, beta = 1

We will prove the superposition equality of f(x) as described in the statement:

$a^T(\alpha[1,2,3] + \beta[-1,0,1]) = f(\alpha[1,2,3] + \beta[-1,0,1]) = \alpha f([1,2,3]) + \beta f([-1,0,1])$

Proof:

$a^T x = a^T(x + y)$ when (x + y) is substituted for x in f(x)

$= a_1 * 1 + a_1 * -1... = \sum a_i * x_i + \sum a_i * y_i$ is the inner product of vector a and (x + y)

The inner product of a and the sum of vectors x and y is equivalent to the sum of the inner product of a and x and the inner product of a and y

$= a^T([1,2,3]) + a^T([-1,0,1])$ = Sum of a_i * x_i = inner product of a and x, Sum of a_i * y_i = inner product of a and y

$= a_1 * 1 + a_2 * 2... + a_1 * -1 + a_2 * 0$ expressing the equivalent form of the inner products above

$= a_1 * (1 - 1) + a_2 * (2 + 0)...$ rearranging terms and pulling out like terms

$f(\alpha x + \beta y) = a^T(0,2,4) = a_1 * (0) + a_2 * (2)...$ <--- demonstrating that the inner product of a with the sum of x and y is the same as above for the sum of the inner products a with x and a with y

Hence, we have proven the superposition equality and thus the linearity of the function with x and y defined with the values above.

Is this still a proof?

It is not a proof in the general case for all values of x and y, but is a proof that when x and y are vectors with the specific values I defined above, that the inner product of any arbitrary vector a and the sum of those two vectors will be a linear function. It also can be generalized to be true for all values x and y.

Why is it useful?

It is useful because by moving from the general case to the specific, we can understand more intuitively the logic behind the proof, and better understand what is not obvious about the proof to begin with. For example, it wouldn't immediately be obvious that the product of a vector with the sum of two vectors would be equal to the sum of the products of two vectors.