College of Engineering & Applied Sciences

CSPB 2820

Linear Algebra With Computer Science Applications

Study Guide 2 - Linear Functions

Taylor Larrechea

2024

Study Guide 2

Linear Functions

Study Guide Instructions

- 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...

Method Selection

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.

Problem 1

Problem Statement

Select one page or section of Chapter Two of VMLS to annotate. Include a screenshot of your annotation here.

For this problem I will be using **Method 2**.

For this weeks annotation, I have chosen to annotate **VMLS Section 2.2 - Taylor Approximation**. These annotations can be found on the following pages.

w_1	w_2	w_3	Measured sag	Predicted sag
1	0	0	0.12	—
0	1	0	0.31	—
0	0	1	0.26	—
0.5	1.1	0.3	0.481	0.479
1.5	0.8	1.2	0.736	0.740

Table 2.1 Loadings on a bridge (first three columns), the associated measured sag at a certain point (fourth column), and the predicted sag using the linear model constructed from the first three experiments (fifth column).

the steel used to construct it. This is always done during the design of a bridge. The vector c can also be *measured* once the bridge is built, using the formula (2.3). We apply the load $w = e_1$, which means that we place a one ton load at the first load position on the bridge, with no load at the other positions. We can then measure the sag, which is c_1. We repeat this experiment, moving the one ton load to positions $2, 3, \ldots, n$, which gives us the coefficients $c_2, \ldots, c_n$. At this point we have the vector c, so we can now *predict* what the sag will be with any other loading. To check our measurements (and linearity of the sag function) we might measure the sag under other more complicated loadings, and in each case compare our prediction (*i.e.*, $c^T w$) with the actual measured sag.

Table 2.1 shows what the results of these experiments might look like, with each row representing an experiment (*i.e.*, placing the loads and measuring the sag). In the last two rows we compare the measured sag and the predicted sag, using the linear function with coefficients found in the first three experiments.

2.2 Taylor approximation

In many applications, scalar-valued functions of n variables, or relations between n variables and a scalar one, can be *approximated* as linear or affine functions. In these cases we sometimes refer to the linear or affine function relating the variables and the scalar variable as a *model*, to remind us that the relation is only an approximation, and not exact.

Differential calculus gives us an organized way to find an approximate affine model. Suppose that $f : \mathbf{R}^n \to \mathbf{R}$ is differentiable, which means that its partial derivatives exist (see §C.1). Let z be an n-vector. The (first-order) *Taylor approximation* of f near (or at) the point z is the function $\hat{f}(x)$ of x defined as

$$\hat{f}(x) = f(z) + \frac{\partial f}{\partial x_1}(z)(x_1 - z_1) + \cdots + \frac{\partial f}{\partial x_n}(z)(x_n - z_n),$$

where $\frac{\partial f}{\partial x_i}(z)$ denotes the partial derivative of f with respect to its ith argument, evaluated at the n-vector z. The hat appearing over f on the left-hand side is

First order approximation, full approximation: $\sum_{n=0}^{\infty} \frac{f^{(n)}(x-a)^n}{n!}$

Estimation will worsen as x_i dives away from z_i

a common notational hint that it is an approximation of the function f. (The approximation is named after the mathematician Brook Taylor.)

The first-order Taylor approximation $\hat{f}(x)$ is a very good approximation of $f(x)$ when all x_i are near the associated z_i. Sometimes $\hat{f}$ is written with a second vector argument, as $\hat{f}(x; z)$, to show the point z at which the approximation is developed. The first term in the Taylor approximation is a constant; the other terms can be interpreted as the contributions to the (approximate) change in the function value (from $f(z)$) due to the changes in the components of x (from z).

Taylor Approx.

Creates estimate of fn around point

Equivalent

Evidently $\hat{f}$ is an affine function of x. (It is sometimes called the *linear approximation* of f near z, even though it is in general affine, and not linear.) It can be written compactly using inner product notation as

$$\hat{f}(x) = f(z) + \nabla f(z)^T (x - z), \tag{2.5}$$

First order approximation evaluated at z

where $\nabla f(z)$ is an n-vector, the *gradient* of f (at the point z),

$$\nabla f(z) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(z) \\ \vdots \\ \frac{\partial f}{\partial x_n}(z) \end{bmatrix}. \tag{2.6}$$

Take partial derivative, evaluate at z

For n values of z

The first term in the Taylor approximation (2.5) is the constant $f(z)$, the value of the function when $x = z$. The second term is the inner product of the gradient of f at z and the *deviation* or *perturbation* of x from z, *i.e.*, $x - z$.

We can express the first-order Taylor approximation as a linear function plus a constant,

$$\hat{f}(x) = \nabla f(z)^T x + (f(z) - \nabla f(z)^T z),$$

First order affine Taylor Approximation

but the form (2.5) is perhaps easier to interpret.

The first-order Taylor approximation gives us an organized way to construct an affine approximation of a function $f : \mathbf{R}^n \to \mathbf{R}$, near a given point z, when there is a formula or equation that describes f, and it is differentiable. A simple example, for $n = 1$, is shown in figure 2.3. Over the full x-axis scale shown, the Taylor approximation $\hat{f}$ does not give a good approximation of the function f. But for x near z, the Taylor approximation is very good.

$\frac{\partial f}{\partial x}(e^{ax}) = ae^{x}$

Example. Consider the function $f : \mathbf{R}^2 \to \mathbf{R}$ given by $f(x) = x_1 + \exp(x_2 - x_1)$, which is not linear or affine. To find the Taylor approximation $\hat{f}$ near the point $z = (1, 2)$, we take partial derivatives to obtain

Original function

$\frac{\partial f}{\partial x_1} = 1 + (0-1)e^{(x_2 - x_1)} = 1 - e^{(x_2 - x_1)}$

$\frac{\partial f}{\partial x_1}\Big|_z = 1 - e^{(z_2 - z_1)}$

$\frac{\partial f}{\partial x_2} = 0 + (1-0)e^{(x_2 - x_1)} = e^{(x_2 - x_1)}$

$\frac{\partial f}{\partial x_2}\Big|_z = e^{(z_2 - z_1)}$

$$\nabla f(z) = \begin{bmatrix} 1 - \exp(z_2 - z_1) \\ \exp(z_2 - z_1) \end{bmatrix},$$

Vector for estimation

which evaluates to $(-1.7183, 2.7183)$ at $z = (1, 2)$. The Taylor approximation at $z = (1, 2)$ is then

Approximation values centered around $z = (1,2)$

$$\begin{aligned} \hat{f}(x) &= 3.7183 + (-1.7183, 2.7183)^T (x - (1, 2)) \\ &= 3.7183 - 1.7183(x_1 - 1) + 2.7183(x_2 - 2). \end{aligned}$$

Vector form

Scalar form after inner product

Table 2.2 shows $f(x)$ and $\hat{f}(x)$, and the approximation error $|\hat{f}(x) - f(x)|$, for some values of x relatively near z. We can see that $\hat{f}$ is indeed a very good approximation of f, especially when x is near z.

Function will not be a good estimation when x is not near z

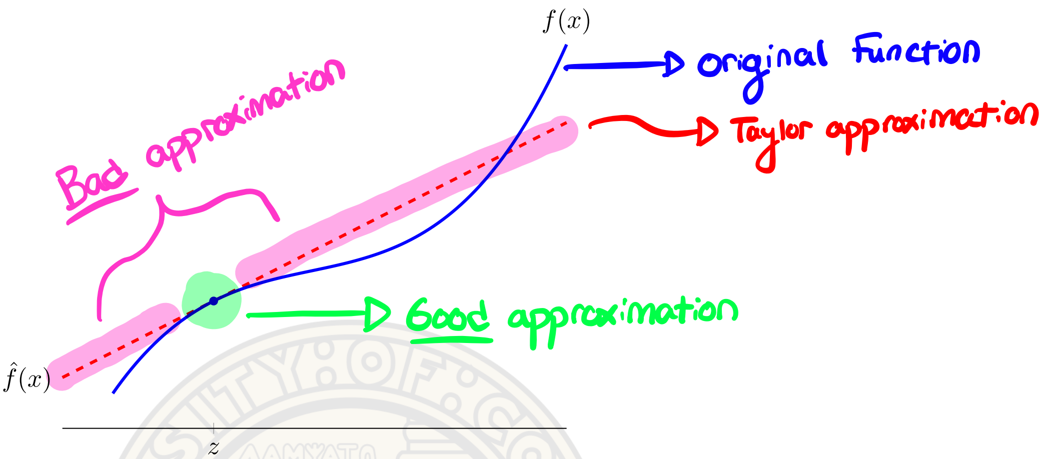

Figure 2.3 A function f of one variable, and the first-order Taylor approximation $\hat{f}(x) = f(z) + f'(z)(x - z)$ at z.

x	$f(x)$	$\hat{f}(x)$	$\lvert\hat{f}(x) - f(x)\rvert$
$(1.00, 2.00)$	3.7183	3.7183	0.0000
$(0.96, 1.98)$	3.7332	3.7326	0.0005
$(1.10, 2.11)$	3.8456	3.8455	0.0001
$(0.85, 2.05)$	4.1701	4.1119	0.0582
$(1.25, 2.41)$	4.4399	4.4032	0.0367

Table 2.2 Some values of x (first column), the function value $f(x)$ (second column), the Taylor approximation $\hat{f}(x)$ (third column), and the error (fourth column).

Problem 1 Summary

Procedure

- Annotate examples from the chapter that cover the topic of Taylor Series

Key Concepts

- This section covers the topic of what a Taylor series expansion is how it works in the context of linear algebra

Variations

- Since this is a problem is a problem on annotation, the only thing that can change is annotating a different section

Problem 2

Problem Statement

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

Original Question

General formula for affine functions. Verify that formula (2.4) holds for any affine function $f : \mathbf{R}^n \to \mathbf{R}$. You can use the fact that $f(x) = a^T x + b$ for some n-vector a and scalar b.

Solution

For this problem, I will be using **Method 4**.

The formula that is referenced in the problem statement (2.4) is

$$f(x) = f(0) + x_1(f(e_1) - f(0)) + \cdots + x_n(f(e_n) - f(0)) = f(0) + \sum_{i=1}^{n} x_i(f(e_i) - f(0)). \tag{1}$$

We know that by definition

$$f(x) = a^T x + b, \tag{2}$$

this implies that $f(0)$ is going to result in an inner product of a with the zero vector. This means we can say that $f(0)$ is then

$$f(0) = a^T\mathbf{0} + b = \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} + b = 0(a_1) + \cdots + 0(a_n) + b = b. \tag{3}$$

From equation (3) we see that $f(0)$ is just the scalar value b. We now aim to express the other terms in equation (1). Starting with the general form $x_n(f(e_n) - f(0))$ we can first define $f(e_n)$ to be

$$f(e_n) = a^T e_n + b = \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} + b = a_1(0) + \cdots + a_n(1) + b = a_n(1) + b = a_n + b. \tag{4}$$

In equation (4), e_n is the n^{th} basis vector (a vector filled with zeros for elements with the exception of the n^{th} element is a 1). Combining equation (4) with the result from (3) we can justifiably say that $f(e_n) - f(0)$ is thus

$$f(e_n) - f(0) = a_n + b - b = a_n. \tag{5}$$

This in turn implies that $x_n(f(e_n) - f(0))$ is

$$x_n(f(e_n) - f(0)) = x_n a_n. \tag{6}$$

We can further say from the result of equation (6) that

$$f(x) = f(0) + \sum_{i=1}^{n} x_i(f(e_i) - f(0)) = b + \sum_{i=1}^{n} x_i a_i = b + a^T x = a^T x + b. \tag{7}$$

We used the commutative property to re-arrange our final result to match the definition of $f(x)$ as well as the commutative property of inner products to say $\sum_{i=1}^{n} x_i a_i = a^T x$. Consequently (2.4) holds for any affine function.

Problem 2 Summary

Procedure

- Take the formula (2.4) from the problem statement and show that the zero vector evaluates to the scalar b
- Take the same formula (2.4) from the problem statement and show that it evaluates to $a_n + b$ for the unit vector (e_n)
- Show what the term $f(e_n) - f(0)$ evaluates to
- Show what the term $x_n(f(e_n) - f(0)) = x_n a_n$ evaluates to
- Piece together the following terms to show what the formula evaluates to

Key Concepts

- Affine functions are of the form $f(x) = a^T x + b$
- Formula (2.4) has the form $f(x) = f(0) + \sum_{i=1}^{n} x_i(f(e_i) - f(0)) = a^T x + b$ for any affine function

Variations

- We could be asked to show that an affine function would hold for a different formula
  - We would then have to go through the same procedure to show that it holds

Problem 3

Problem Statement

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).

Original Question

Linear or *not*? Determine whether each of the following scalar-valued functions of n-vectors is linear. If it is a linear function, give its inner product representation, i.e., an n-vector a for which $f(x) = a^T x$ for all x. If it is not linear, give specific x, y, α, and β for which superposition fails, i.e.,

$$f(\alpha x + \beta y) \neq \alpha f(x) + \beta f(y).$$

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

(b) The difference of the last element and the first, $f(x) = x_n - x_1$.

(c) The median of an n-vector, where we will assume $n = 2k+1$ is odd. The median of the vector x is defined as the $(k+1)^{\text{st}}$ largest number among the entries of x. For example, the median of $(-7.1, 3.2, -1.5)$ is (-1.5).

(d) The average of the entries with odd indices, minus the average of the entries with even indices. You can assume that $n = 2k$ is even.

(e) Vector extrapolation, defined as $x_n + (x_n - x_{n-1})$, for $n \geq 2$. (This is a simple prediction of what x_{n+1} would be, based on a straight line drawn through x_n and x_{n-1}.)

Solution - Part (a)

For this problem, I will be using **Method 4**.

Function for this part: $f(x) = \max_k x_k - \min_k x_k$.

VMLS Solution:

Not linear. Choose $x = (1,0), y = (0,1), \alpha = \beta = 1/2$. Then $\alpha x + \beta y = (1/2, 1/2)$, $f(x) = f(y) = 1$, and

$$f(\alpha x + \beta y) = 0 \neq \alpha f(x) + \beta f(y) = 1. \tag{1}$$

Explanation:

The fundamental reason as to why this function is not linear is because there exists a given set of vectors that will cause superposition (the addition of two vectors into one) to fail. In this example, the authors chose the vectors $x = (1,0)$ and $y = (0,1)$ along with $\alpha = \beta = 1/2$ to demonstrate this point. When we calculate the minimum value being subtracted from the maximum value of a given vector, it can be shown that

$$f(\alpha x + \beta y) = f\left(\frac{1}{2} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{2} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = f\left(\begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{2} \end{bmatrix}\right) = f\left(\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}\right) = \max - \min = 1/2 - 1/2 = 0. \tag{2}$$

Where, if we now try to test the law of superposition to prove that this specific example is linear we have

$$f(x) = f\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \max - \min = 1 - 0 = 1 \ , \ f(y) = f\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \max - \min = 1 - 0 = 1 \tag{3}$$

where we then calculate

$$\alpha f(x) + \beta f(y) = \frac{1}{2}(1) + \frac{1}{2}(1) = 1 \ \rightarrow \ f(\alpha x + \beta y) \neq \alpha f(x) + \beta f(y) \ \rightarrow \ 1 \neq 0 \tag{4}$$

and thus superposition fails and this is not linear.

Solution - Part (b)

For this problem, I will be using **Method 4**.

Function for this part: $f(x) = x_n - x_1$.

VMLS Solution:

Linear. The function can be written as an inner product $f(x) = a^T x$ for

$$a = (-1, 0, \dots, 0, 1). \tag{5}$$

Explanation:

The simplest explanation for why this is indeed linear is that when calculating the difference between the first and last element of a vector, there is no scalar value (α, β) that when applied to the vectors that will cause the difference between the scaled vector and original vector to be different. Namely

$$f(\alpha x + \beta y) = \alpha(x_n) - \alpha(x_1) + \beta(y_n) - \beta(y_1) = \alpha(x_n - x_1) + \beta(y_n - y_1) = \alpha f(x) + \beta f(y). \tag{6}$$

Equation (6) satisfies superposition and thus this function is indeed linear. This vector would then be represented as

$$a = \begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{7}$$

where the middle elements must be zero so that we end up with $f(x) = x_n - x_1$ for our function.

Solution - Part (c)

For this problem, I will be using **Method 4**.

Function for this part: Median of vector.

VMLS Solution:

Not linear. Choose $x = (-1, 0, 3), y = (2, -1, 0), \alpha = \beta = 1$. Then $\alpha x + \beta y = (1, -1, 2), f(x) = f(y) = 0$, and

$$f(\alpha x + \beta y) = 1 \neq \alpha f(x) + \beta f(y) = 0. \tag{8}$$

Explanation:

This function fails superposition and is not linear. In this example, we are grabbing the median value (which is the second largest element in a vector by the problems definition) of a given vector. It can be shown for this specific example that there exists an α and β such that superposition will fail and thus the function will be non linear. In our case, with the values that the other chose for the vectors we see that

$$f(x) = f\left(\begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}\right) = 0 \ , \ f(y) = f\left(\begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}\right) = 0 \ \rightarrow \ \alpha f(x) + \beta f(y) = 1(0) + 1(0) = 0. \tag{9}$$

Consequently, it can also be shown that with $\alpha = \beta = 1$ that

$$f(\alpha x + \beta y) = f\left(1 \cdot \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix} + 1 \cdot \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}\right) = f\left(\begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}\right) = 1. \tag{10}$$

By the law of superposition we have just shown $f(\alpha x + \beta y) \neq \alpha f(x) + \beta f(y)$ and thus this function is not linear.

Solution - Part (d)

For this problem, I will be using **Method 4**.

Function for this part: Average of odd indices minus average of even indices.

VMLS Solution:

Linear. If the size of the vector is even ($n = 2m$), there are m even and m odd indices, and $f(x) = a^T x$ with

$$a = (1/m, -1/m, 1/m, -1/m, \dots, 1/m, -1/m). \tag{11}$$

Explanation:

This function passes superposition and is thus linear. If we write out the function in mathematical terms, we have

$$f(x) = \frac{1}{k}\left(\sum_{i=i,3,\dots,2k-1} x_i - \sum_{i=2,4,\dots,2k} x_i \right) \tag{12}$$

and no matter the value of α and β we can show

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) \tag{13}$$

and the function will scale linearly. This vector is scaled with a factor of the value k and is thus represented as

$$a = \begin{bmatrix} 1/k \\ -1/k \\ \vdots \\ 1/k \\ -1/k \end{bmatrix} \tag{14}$$

where when a scalar value is applied to this vector, the averages of the odd and even indices will scale accordingly.

Solution - Part (e)

For this problem, I will be using **Method 4**.

Function for this part: Vector extrapolation ($x_n + (x_n - x_{n-1})$) for $n \geq 2$.

VMLS Solution:

Linear. We have $f(x) = a^T x$ for

$$a = (0, 0, \dots, 0, -1, 2). \tag{15}$$

Explanation:

This function is linear because it satisfies the law of superposition. For any given value of α or β the scaled vector will match the combination of scaled vectors independently. For instance, if we choose $x = (0, 1, 2)$ and $y = (-1, 1, 3)$ with $\alpha = \beta = 1$

$$f(\alpha x + \beta y) = f\left(1 \cdot \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + 1 \cdot \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix} \right) = f\left(\begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix} \right) = 8. \tag{16}$$

Consequently, for this specific example we can show

$$\alpha f(x) + \beta f(y) = 1 \cdot f\left(\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right) + 1 \cdot f\left(\begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix} \right) = 1(3) + 1(5) = 8 \tag{17}$$

and thus

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y). \tag{18}$$

Although equations (16) through (18) are for a specific example, any arbitrary set of vectors (x, y) with any arbitrary scalar values (α, β) will satisfy the law of superposition and thus the function is linear. For the inner product representation of this function, we constitute that a must be

$$a = \begin{bmatrix} 0 \\ \vdots \\ -1 \\ 2 \end{bmatrix}. \tag{19}$$

We require the trailing elements in a to be zero so that when the inner product of a is taken with x we get the resulting function $f(x) = x_n + (x_n - x_{n-1}) = 2x_n - x_1$. The scalar values (-1,2) are chosen for the last elements of a such that when the inner product is carried out with an arbitrary vector x, we get the previous expression.

Problem 3 Summary

Procedure

(a) Part (a)
- Pick two vectors that can disprove the law of superposition
- Go through the machinery of the law of superposition for the given function of the problem

(b) Part (b)
- Go through the machinery of the law of superposition for the given function of the problem
- Show that the law of superposition holds for this specific function

(c) Part (c)
- Pick two vectors that can disprove the law of superposition
- Go through the machinery of the law of superposition for the given function of the problem

(d) Part (d)
- Go through the machinery of the law of superposition for the given function of the problem
- Show that the law of superposition holds for this specific function

(e) Part (e)
- Go through the machinery of the law of superposition for the given function of the problem
- Show that the law of superposition holds for this specific function

Key Concepts

- For a function to be considered linear, it must be of the form

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

for all values of α and β

Variations

- We could be given a different function to prove it is linear or not
  - Go through the same machinery of reasoning if the law of superposition holds for a given function

Problem 4

Problem Statement

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).

Original Question

Linear function? The function ϕ: $\mathbf{R}^3 \to \mathbf{R}$ satisfies

$$\phi(1,1,0) = -1, \quad \phi(-1,1,1) = 1, \quad \phi(1,-1,-1) = 1.$$

Choose one of the following, and justify your chose: ϕ must be linear; ϕ could be linear; ϕ cannot be linear.

Solution

For this problem, I will be using **Method 4**.

VMLS Solution:

The correct answer is: ϕ *cannot be linear.* To see this, we note that the third point $(1,-1,-1)$ is the negative of the second point $(-1,1,1)$. If ϕ were linear, the two values of ϕ would need to be negatives of each other. But they are 1 and 1, not negatives of each other.

Explanation:

For the mapping of a function from one space to another, it must satisfy *homogeneity* and *additivity*. We first check additivity to show

$$\phi(x+y) = \phi\left(\begin{bmatrix}1\\1\\0\end{bmatrix} + \begin{bmatrix}1\\-1\\-1\end{bmatrix}\right) = \phi\left(\begin{bmatrix}2\\0\\-1\end{bmatrix}\right) \;,\; \phi(x+y) = \phi\left(\begin{bmatrix}1\\1\\0\end{bmatrix} + \begin{bmatrix}-1\\1\\1\end{bmatrix}\right) = \phi\left(\begin{bmatrix}0\\2\\1\end{bmatrix}\right) \tag{1}$$

that these functions ϕ fail the additivity property. This is because the resulting functions that are produced after using the additivity property do not exist in this space that we are provided. Therefore the additivity property cannot be confirmed. (Due to the associativity property of vector sums I do not need to show the other combinations, the result will be the same.)

We can then check the homogeneity property with $(\alpha = -1)$ of these functions and show

$$\phi(\alpha x) = \phi\left(-1 \cdot \begin{bmatrix}1\\1\\0\end{bmatrix}\right) = \phi\left(\begin{bmatrix}-1\\-1\\0\end{bmatrix}\right) =? \neq \alpha\phi(x) = (-1)\phi(1,1,0) = 1 \tag{2}$$

$$\phi(\alpha x) = \phi\left(-1 \cdot \begin{bmatrix}-1\\1\\1\end{bmatrix}\right) = \phi\left(\begin{bmatrix}1\\-1\\-1\end{bmatrix}\right) = 1 \neq \alpha\phi(x) = (-1)\phi(-1,1,1) = -1 \tag{3}$$

$$\phi(\alpha x) = \phi\left(-1 \cdot \begin{bmatrix}1\\-1\\-1\end{bmatrix}\right) = \phi\left(\begin{bmatrix}-1\\1\\1\end{bmatrix}\right) = 1 \neq \alpha\phi(x) = (-1)\phi(1,-1,-1) = -1 \tag{4}$$

that these functions do not satisfy homogeneity and cannot be linear functions.

Problem 4 Summary

Procedure

- For the functions that are presented to us, test them with the homogeneity and additivity principle
- Test two of the functions with the additivity principle and show that the resulting function does not exist in this space
- Choose a value for α and show that it doesn't hold for the homogeneity principle

Key Concepts

- For a function to be considered linear, they must follow the following principles

$$f(\alpha x) = \alpha f(x) \qquad \text{(Homogeneity)}$$
$$f(x+y) = f(x) + f(y) \qquad \text{(Additivity)}$$

- These functions that are presented to us are not considered linear because they do not uphold the homogeneity and additivity principle
- The additivity principle fails because the resulting function does not have a value in this space

Variations

- We could be given different functions to prove if they are linear or not
  - In this case we would use the homogeneity and additivity principle to test if the functions are linear or not

Problem 5

Problem Statement

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).

Original Question

Questionnaire Scoring. A questionnaire in a magazine has 30 questions, broken into two sets of 15 questions. Someone taking the questionnaire answers each question with 'Rarely', 'Sometimes', or 'Often'. The answers are recorded as a 30-vector a, with $a_i = 1, 2, 3$ if question i is answered Rarely, Sometimes, or Often, respectively. The total score on a completed questionnaire is found by adding up 1 point for every question answered Sometimes and 2 points for every question answered Often on questions 1-15, and by adding 2 points and 4 points for those responses on questions 16-30. (Nothing is added to the score for Rarely responses.) Express the total score s in the form of an affine function $s = w^T a + v$ where w is a 30-vector and v is a scalar (number).

Solution

For this problem, I will be using **Method 4**.

VMLS Solution:

We have $w = (\mathbf{1}_{15}, 2\mathbf{1}_{15})$ and $v = -45$. The contribution to the score from questions 1-15 is given by $\mathbf{1}^T(a_{1:15} - \mathbf{1})$. The contribution to the scores from questions 16-30 is given by $2\mathbf{1}^T(a_{16:30} - \mathbf{1})$. So, we can write the overall score as

$$(\mathbf{1}_{15}, 2\mathbf{1}_{15})^T(a - \mathbf{1}) = (\mathbf{1}_{15}, 2\mathbf{1}_{15})^T a - 45. \tag{1}$$

Explanation:

For starters, the term $(a - \mathbf{1})$ represents a 30 vector that is used to create the desired representation of the responses for each question. Here, if the participant responds with 'Rarely' then this element will be 0 (1 - 1) (since $\mathbf{1}$ is a vector of 1's in every entry), 'Sometimes' will evaluate to 1 (2 - 1) and 'Often' will evaluate to 2 (3 - 1). The term $(\mathbf{1}_{15}, 2\mathbf{1}_{15})^T$ is another 30 vector that represents the weighting of the first 15 and last 15 questions respectively. Putting this together we get a representation of the total score from the responses with

$$s = (\mathbf{1}_{15}, 2\mathbf{1}_{15})^T(a - \mathbf{1}). \tag{2}$$

Now, to make this function affine we must add a constant scalar value to the end of our inner product. For the first 15 entries, we are subtracting a value of 1 from each element and for the last 15 we are subtracting 2 from each element. This simplifies to $-1(15) - 2(15) = -45$, this is how we obtain the value of v to be -45. We can simplify the equation in equation (1) to then just take the inner product of the weighted responses for both the first 15 and last 15 responses of the survey with the scoring values found in a. Since the first 15 questions are weighted with a factor of just 1, we have the term $(\mathbf{1}_{15})$ and since the last 15 responses are weighted with a value of 2, we have the term $(2\mathbf{1}_{15})$. Stitching this together to get an affine function we have

$$s = (\mathbf{1}_{15}, 2\mathbf{1}_{15})^T a - 45. \tag{3}$$

Equations (2) and (3) are equivalent where equation (3) represents an affine function.

Problem 5 Summary

Procedure

- Generate a term that can accurately depict the responses for the questionnaire (this is represented by subtracting the value 1 from each entry)
- Generate a vector that will accurately represent the weighting of the questions from the questionnaire (this is done by concatenating a 10 vector with a 20 vector)
- Write an expression that encapsulates each of these terms in the form of an inner product
- To create an equivalent expression to the previous one, subtract the scalar 45 from the inner product of the weighting of the responses and the answers from the questionnaire

Key Concepts

- This problem shows that an inner product can be written as an affine function with correct manipulation of the original inner product

Variations

- We could be given a different question that encapsulates a different scenario and asked to do a similar operation
  - We then would use similar methods to show that the inner product could be represented as an affine function

Problem 6

Problem Statement

(a) Re-do 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?

Solution - Part (a)

For this problem, I will be using **Method 1**.

To do this proof, we first define arbitrary vectors for (x, y, a) and scalar values (α, β)

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad , \quad y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \quad , \quad a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} . \tag{1}$$

Starting with $f(\alpha x + \beta y) = a^T(\alpha x + \beta y)$ we can show

$$\begin{aligned}
f(\alpha x + \beta y) &= \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \left(\alpha \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \beta \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \right) && (2) \\
&= \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \left(\begin{bmatrix} \alpha x_1 + \beta y_1 \\ \alpha x_2 + \beta y_2 \\ \alpha x_3 + \beta y_3 \end{bmatrix} \right) && (3) \\
&= a_1(\alpha x_1 + \beta y_1) + a_2(\alpha x_2 + \beta y_2) + a_3(\alpha x_3 + \beta y_3) && (4) \\
&= a_1 \alpha x_1 + a_1 \beta y_1 + a_2 \alpha x_2 + a_2 \beta y_2 + a_3 \alpha x_3 + a_3 \beta y_3 && (5) \\
&= \alpha(a_1 x_1 + a_2 x_2 + a_3 x_3) + \beta(a_1 y_1 + a_2 y_2 + a_3 y_3) && (6) \\
&= \alpha(a^T x) + \beta(a^T y) && (7) \\
&= \alpha f(x) + \beta f(y). && (8)
\end{aligned}$$

To get to equation (8), we used the definition of an inner product, namely $f(k) = a^T k$ where k is an arbitrary vector (in our case x and y).

If we would like to show an example of this with actual values and not just placeholders for values, we would need to show equivalence between equation (5) (or equation (6)) and equation (7) (or equation (8)). For this example, we will define our vectors to be

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad , \quad y = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \quad , \quad a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \quad , \quad (\alpha, \beta) \tag{9}$$

First we use equation (5), when we do this we obtain

$$\begin{aligned}
a_1 \alpha x_1 + a_1 \beta y_1 + a_2 \alpha x_2 + a_2 \beta y_2 + a_3 \alpha x_3 + a_3 \beta y_3 &= a_1 \alpha(1) + a_1 \beta(4) + a_2 \alpha(2) + a_2 \beta(5) + a_3 \alpha(3) + a_3 \beta(6) && (10) \\
&= \alpha(a_1 + 2a_2 + 3a_3) + \beta(4a_1 + 5a_2 + 6a_3) && (11)
\end{aligned}$$

Now, for superposition to hold for our given example we require equation (11) to be equal to that of the result that will come out of equation (7). Carrying out the steps for equation (7) we find

$$\begin{aligned}
\alpha(a^T x) + \beta(a^T y) &= \alpha \cdot \left(\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right) + \beta \cdot \left(\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \right) && (12) \\
&= \alpha(a_1 + 2a_2 + 3a_3) + \beta(4a_1 + 5a_2 + 6a_3) && (13)
\end{aligned}$$

Equation (13) is equal to that of equation (11). Consequently, we have shown that for the vectors and scalars that we have defined in equation (9), that the superposition property is satisfied for our specific example.

Solution - Part (b)

For this problem, I will be using **Method 1**.

The exercise in part (a) is not a formal proof. For this example we were picking a specific set of vectors to show that the inner product of these vectors would satisfy superposition. For this to be a formal proof, we would have to generalize it for any set of vectors. What I did prior to the selecting of specific vectors was more a formal proof, but probably not with great format. Selecting a specific example like we did in the exercise only shows that the inner product of those vectors will satisfy superposition, not all vectors.

Solution - Part (c)

For this problem, I will be using **Method 1**.

This is still useful because it allows us to work out a specific example and show that the inner product of this specific example will satisfy superposition. It is useful to work out a proof like this with specific examples to show that it is valid, just so that the machinery of showing superposition gets stuck in one's head.

Problem 6 Summary

Procedure

(a) Part (a)

- Generalize the vectors x, y, a and apply them to $f(\alpha x + \beta y) = a^T(\alpha x + \beta y)$
- Carry out the algebra and show that the theorem is correct
- Show that the theorem holds for a specific set of vectors

(b) Part (b)

- Answer the prompt

(c) Part (c)

- Answer the prompt

Key Concepts

- It can be shown that the law of superposition can be written as

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

Variations

- We could be asked to show that a different theorem is true
  - We would apply the same machinery from the original problem to this new example