College of Engineering & Applied Sciences

CSPB 2820

Linear Algebra With Computer Science Applications

Study Guide 5 - Linear Independence

Taylor Larrechea

2023

Sections

Linear Independence 2
Problem 1. 3
Problem 2. 8
Problem 3. 11
Problem 4. 13
Problem 5. 15
Problem 6. 17
Problem 7. 19

Study Guide 5

Linear Independence

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

Pick a section of Chapter 5 to annotate.

For annotations this week, I have chosen to annotate section 5.1 of VMLS. The annotations for this problem can be seen on the following page.

Chapter 5

Linear independence

In this chapter we explore the concept of linear independence, which will play an important role in the sequel.

5.1 Linear dependence

A collection or list of n-vectors $a_1, \ldots, a_k$ (with $k \geq 1$) is called *linearly dependent* if

$$\beta_1 a_1 + \cdots + \beta_k a_k = 0$$

At least one β that is non zero

holds for some $\beta_1, \ldots, \beta_k$ that are not all zero. In other words, we can form the zero vector as a linear combination of the vectors, with coefficients that are not all zero. Linear dependence of a list of vectors does not depend on the ordering of the vectors in the list.

When a collection of vectors is linearly dependent, at least one of the vectors can be expressed as a linear combination of the other vectors: If $\beta_i \neq 0$ in the equation above (and by definition, this must be true for at least one i), we can move the term $\beta_i a_i$ to the other side of the equation and divide by β_i to get

$$a_i = (-\beta_1/\beta_i)a_1 + \cdots + (-\beta_{i-1}/\beta_i)a_{i-1} + (-\beta_{i+1}/\beta_i)a_{i+1} + \cdots + (-\beta_k/\beta_i)a_k.$$

Can construct vector based off of others

The converse is also true: If any vector in a collection of vectors is a linear combination of the other vectors, then the collection of vectors is linearly dependent.

Following standard mathematical language usage, we will say "The vectors $a_1, \ldots, a_k$ are linearly dependent" to mean "The list of vectors $a_1, \ldots, a_k$ is linearly dependent". But it must be remembered that linear dependence is an attribute of a *collection* of vectors, and not individual vectors.

Linearly independent vectors. A collection of n-vectors $a_1, \ldots, a_k$ (with $k \geq 1$) is called *linearly independent* if it is not linearly dependent, which means that

$$\beta_1 a_1 + \cdots + \beta_k a_k = 0 \tag{5.1}$$

Linearly dependent → at least one β is zero

Linearly independent → all β's are zero

} For 5.1 to be true

only holds for $\beta_1 = \cdots = \beta_k = 0$. In other words, the only linear combination of the vectors that equals the zero vector is the linear combination with all coefficients zero.

As with linear dependence, we will say "The vectors $a_1, \ldots, a_k$ are linearly independent" to mean "The list of vectors $a_1, \ldots, a_k$ is linearly independent". But, like linear dependence, linear independence is an attribute of a collection of vectors, and not individual vectors.

Set of vectors

It is generally not easy to determine by casual inspection whether or not a list of vectors is linearly dependent or linearly independent. But we will soon see an algorithm that does this.

Examples.

- A list consisting of a single vector is linearly dependent only if the vector is zero. It is linearly independent only if the vector is nonzero.
- Any list of vectors containing the zero vector is linearly dependent.
- A list of two vectors is linearly dependent if and only if one of the vectors is a multiple of the other one. More generally, a list of vectors is linearly dependent if any one of the vectors is a multiple of another one.
- The vectors

β's not all zero

$$a_1 = \begin{bmatrix} 0.2 \\ -7.0 \\ 8.6 \end{bmatrix}, \quad a_2 = \begin{bmatrix} -0.1 \\ 2.0 \\ -1.0 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 0.0 \\ -1.0 \\ 2.2 \end{bmatrix}$$

are linearly dependent, since $a_1 + 2a_2 - 3a_3 = 0$. We can express any of these vectors as a linear combination of the other two. For example, we have $a_2 = (-1/2)a_1 + (3/2)a_3$.

- The vectors

All β's must be zero

$$a_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \quad a_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

are linearly independent. To see this, suppose $\beta_1 a_1 + \beta_2 a_2 + \beta_3 a_3 = 0$. This means that

Can't construct vectors from others

$$\beta_1 - \beta_3 = 0, \quad -\beta_2 + \beta_3 = 0, \quad \beta_2 + \beta_3 = 0.$$

Adding the last two equations we find that $2\beta_3 = -0$, so $\beta_3 = 0$. Using this, the first equation is then $\beta_1 = 0$, and the second equation is $\beta_2 = 0$.

- The standard unit n-vectors $e_1, \ldots, e_n$ are linearly independent. To see this, suppose that (5.1) holds. We have

$$0 = \beta_1 e_1 + \cdots + \beta_n e_n = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix},$$

so we conclude that $\beta_1 = \cdots = \beta_n = 0$.

The set of linearly independent unit vectors is linearly independent

Linear combinations of linearly independent vectors. Suppose a vector x is a linear combination of $a_1, \ldots, a_k$,

$$x = \beta_1 a_1 + \cdots + \beta_k a_k.$$

When the vectors $a_1, \ldots, a_k$ are linearly independent, the coefficients that form x are *unique*: If we also have

$$x = \gamma_1 a_1 + \cdots + \gamma_k a_k,$$

then $\beta_i = \gamma_i$ for $i = 1, \ldots, k$. This tells us that, in principle at least, we can find the coefficients that form a vector x as a linear combination of linearly independent vectors.

To see this, we subtract the two equations above to get

$$0 = (\beta_1 - \gamma_1)a_1 + \cdots + (\beta_k - \gamma_k)a_k.$$

Since $a_1, \ldots, a_k$ are linearly independent, we conclude that $\beta_i - \gamma_i$ are all zero.

The converse is also true: If each linear combination of a list of vectors can only be expressed as a linear combination with one set of coefficients, then the list of vectors is linearly independent. This gives a nice interpretation of linear independence: A list of vectors is linearly independent if and only if for any linear combination of them, we can infer or deduce the associated coefficients. (We will see later how to do this.) must be able to determine co-efficients

Supersets and subsets. If a collection of vectors is linearly dependent, then any superset of it is linearly dependent. In other words: If we add vectors to a linearly dependent collection of vectors, the new collection is also linearly dependent. Any nonempty subset of a linearly independent collection of vectors is linearly independent. In other words: Removing vectors from a collection of vectors preserves linear independence.

5.2 Basis

Independence-dimension inequality. If the n-vectors $a_1, \ldots, a_k$ are linearly independent, then $k \leq n$. In words:

A linearly independent collection of n-vectors can have at most n elements.

Put another way:

Any collection of $n + 1$ or more n-vectors is linearly dependent.

As a very simple example, we can conclude that any three 2-vectors must be linearly dependent. This is illustrated in figure 5.1.

We will prove this fundamental fact below; but first, we describe the concept of basis, which relies on the independence-dimension inequality.

Problem 1 Summary

Procedure

- Annotate a chapter from the textbook by adding notes and insights

Key Concepts

- This chapter encapsulates what it means for a collection of vectors to be linearly independent or dependent
- For a collection of vectors to be linearly independent it must follow that

$$\beta_1 a_1 + \cdots + \beta_k a_k = 0$$

  where the only way the above is true is if **every** β is 0
- For a collection of vectors to be linearly dependent it must follow that

$$\beta_1 a_1 + \cdots + \beta_k a_k = 0$$

  where the above is true is if there is at least **one** β that is non zero

Variations

- We could be asked to annotate a different chapter or section of the textbook

Problem 2

Problem Statement

Solve and explain the solution to 5.1 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

Original Question:

Linear independence of stacked vectors. Consider the stacked vectors

$$c_1 = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix}, \ldots, c_k = \begin{bmatrix} a_k \\ b_k \end{bmatrix},$$

where $a_1, \ldots, a_k$ are n-vectors and $b_1, \ldots, b_k$ are m-vectors.

(a) Suppose $a_1, \ldots, a_k$ are linearly independent. (We make no assumptions about the vectors $b_1, \ldots, b_k$) Can we conclude that the stacked vectors $c_1, \ldots, c_k$ are linearly independent?

(b) Now suppose that $a_1, \ldots, a_k$ are linearly dependent. (Again, with no assumptions about $b_1, \ldots, b_k$.) Can we conclude that the stacked vectors $c_1, \ldots, c_k$ are linearly dependent?

Solution - Part (a)

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

VMLS Solution:

(a) Yes, the stacked vectors are always independent. Suppose that $\beta_1 c_1 + \cdots + \beta_k c_k = 0$. Then we have $\beta_1 a_1 + \cdots + \beta_k a_k = 0$ and $\beta_1 b_1 + \cdots + \beta_k b_k = 0$. Since $a_1, \ldots, a_k$ are linearly independent, we must have $\beta_1 = \cdots = \beta_k = 0$. So $c_1, \ldots, c_k$ are linearly independent.

Explanation:

We are told in the problem statement that the vector a is linearly independent. Mathematically this means

$$\sum_{i=1}^{k} \beta_i a_i = 0 \tag{1}$$

where the above is only valid if all values of β are zero. This is the definition of linear independence. Now for the vector b we have

$$\sum_{i=1}^{k} \beta_i b_i = 0 \tag{2}$$

but the question is if all the values of β in equation (2) are non-zero or zero. Applying a similar method to vector c as we did to both vectors a and b we have

$$\sum_{i=1}^{k} \beta_i c_i = 0 \tag{3}$$

where if we expand equation (3) we then have

$$\sum_{i=1}^{k} \beta_i c_i = \sum_{i=1}^{k} \beta_i \begin{bmatrix} a_i \\ b_i \end{bmatrix} = 0. \tag{4}$$

Because the same β is being multiplied to vector b as it is to vector a, and we know that all the β's must be zero for a to be linearly independent, this implies that the vector c must also be ***linearly independent*** as well.

Solution - Part (b)

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

VMLS Solution:

(b) No, we can't conclude that they are dependent. We can find examples with a_i linearly dependent, and c_i linearly independent, and also examples with a_i linearly dependent, and c_i linearly dependent. For the first example, take $a_1 = a_2 = 1$ (they are 1-vectors), and $b_1 = 1, b_2 = 0$. For the second example, take the same a_i and $b_1 = b_2 = 1$.

Explanation:

We are told in the problem statement that the vector a is linearly dependent. Mathematically this means

$$\sum_{i=1}^{k} \beta_i a_i = 0 \tag{5}$$

where for equation (5) to be true, there must be at least one value of β that is non-zero for a to be classified as linearly dependent. Because of this fact, we can't implicitly say whether or not vector c is linearly independent or linearly dependent. There may still be a combination of values for b to be considered linearly independent or dependent, and thus we can't conclude anything about vector c.

Essentially, b could be constructed in manner so that it is either linearly independent or linearly dependent and thus we can't make any direct assumptions about vector c. We could also have a length of b that would violate the linear independence inequality as well.

Problem 2 Summary

Procedure

(a) Part (a)

- Show that if one collection of vectors is linearly independent we can conclude what the values of β must be for both concatenated vectors
- Show that with this value of β that both concatenated vectors are linearly independent and the entire vector is indeed linearly independent

(b) Part (b)

- Show that if one collection of vectors is linearly dependent, this means that we can't deduce what the full vector is in terms of linearly independent and dependent because the other vector is unknown

Key Concepts

- If one vector is shown to be linearly independent in a stacked vector this then means that each vector is linearly independent and the entire vector is linearly independent
- If one vector is shown to be linearly dependent, we can't conclude what the stacked vector is linearly independent or dependent

Variations

- We could be given a different set of stacked vectors where we are asked what the relationship between them is
  - We then would need to reason what the coefficients of these vectors are to make a determination about the stacked vector

Problem 3

Problem Statement

Solve and Explain the solution to 5.2 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

Original Question:

A surprising discovery. An intern at a quantitative hedge fund examines the daily returns of around 400 stocks over one year (which has 250 trading days). She tells her supervisor that she has discovered that the returns of one of the stocks, Google (GOOG), can be expressed as a linear combination of the others, which include many stocks that are unrelated to Google (say, in a different type of business or sector).

Is the supervisor right? Did the intern make a mistake? Give a very brief explanation.

Solution

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

VMLS Solution:

The supervisor is wrong, and the intern is most likely correct. She is examining 400 250-vectors, each representing the daily returns of a particular stock over one year. By the independence-dimension inequality, any set of $n + 1$ or more n vectors is linearly dependent, so we know that this set of vectors is linearly dependent. That is, there exist some vectors that can be expressed as a linear combination of the others. It is quite likely that the returns of any particular stock, such as GOOG, can be expressed as a linear combination of the returns other stocks.

Even if Google's return can be expressed as a linear combination of the others, by the independence-dimension inequality, this fact is not as useful as it might seems. For example, Google's future returns would not be given by the same linear combination of other asset returns. So although the intern is right, and the supervisor is wrong, the observation cannot be monetized.

Explanation:

From VMLS, the independence-dimension equality is:

Independence-dimension inequality. If the n-vectors $a_1, \ldots, a_k$ are linearly independent, then $k \leq n$. In words:

A linearly independent collection of n-vectors can have at most n elements.

Put another way:

Any collection of n + 1 or more n-vectors is linearly dependent.

From the problem statement, the intern is dealing with 400 (this is the value of n) 250-vectors (this is the value of k) that represent the returns of stocks for 400 different stocks. From the independence-inequality we know that this collection of vectors is ***linearly dependent*** because $k = 250 \leq n = 400$.

Translated to English, this means that the stock returns of one stock, for example GOOG in the problem statement, can be expressed as linear combinations of others. So, by this logic the intern is correct and the supervisor is wrong.

Problem 3 Summary

Procedure

- Use the independence-dimension inequality to reason if the intern is correct

Key Concepts

- Using the independence-dimension inequality we can reason that this collection of vectors is linearly dependent
- The independence-dimension inequality states that

A linearly independent collection of n-vectors can have at most n elements.

Any collection of n + 1 or more n-vectors is linearly dependent.

Variations

- We could be given a different set of vectors to determine if they are linearly independent or dependent
  - We would then reason with the independence-dimension inequality to determine if these vectors are linearly independent or dependent

Problem 4

Problem Statement

Solve and Explain the solution to 5.3 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

Original Question:

Replicating a cash flow with single-period loans. We continue the example described on page 93. Let $\mathbf{c}$ be any $\mathbf{n}$-vector representing a cash flow over $\mathbf{n}$ periods. Find the coefficients $\alpha_1, \dots, \alpha_n$ of $\mathbf{c}$ in its expansion in the basis $\mathbf{e}_1, \mathbf{l}_1, \dots, \mathbf{l}_{n-1}$, i.e.,

$$\mathbf{c} = \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{l}_1 + \dots + \alpha_n \mathbf{l}_{n-1}.$$

Verify that α_1 is the net present value (NPV) of the cash flow $\mathbf{c}$, defined on page 22, with interest rate r. Hint. Use the same type of argument that was used to show that $\mathbf{e}_1, \mathbf{l}_1, \dots, \mathbf{l}_{n-1}$ are linearly independent. Your method will find α_n first, then α_{n-1}, and so on.

Solution

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

VMLS Solution:

As noted on page 93,

$$\alpha_1 e_1 + \alpha_2 l_1 + \cdots + \alpha_n l_{n-1} = \begin{bmatrix} \alpha_1 + \alpha_2 \\ \alpha_3 - (1+r)\alpha_2 \\ \vdots \\ \alpha_n - (1+r)\alpha_{n-1} \\ -(1+r)\alpha_n \end{bmatrix}.$$

We equate this to $c = (c_1, c_2, \dots, c_{n-1}, c_n)$, and determine $\alpha_n, \alpha_{n-1}, \dots, \alpha_1$ from the equality:

$$\alpha_n = -\frac{c_n}{1+r},$$

and therefore

$$\alpha_{n-1} = -\frac{c_{n-1} - \alpha_n}{1+r} = -\frac{c_{n-1}}{1+r} - \frac{c_n}{(1+r)^2},$$

and

$$\alpha_{n-2} = -\frac{c_{n-2} - \alpha_{n-1}}{1+r} = -\frac{c_{n-2}}{1+r} - \frac{c_{n-1}}{(1+r)^2} - \frac{c_n}{(1+r)^2},$$

et cetera. Continuing the pattern we find

$$\alpha_2 = -\frac{c_2}{1+r} - \frac{c_3}{(1+r)^2} - \cdots - \frac{c_n}{(1+r)^{n-1}}$$

and finally

$$\alpha_1 = c_1 - \alpha_2 = c_1 + \frac{c_2}{1+r} + \cdots + \frac{c_n}{(1+r)^{n-1}}.$$

This is exactly the net present value (NPV) of the cash flow, with interest rate r.

Problem 4 Summary

Procedure

- Follow the logic of the VMLS solution

Key Concepts

- This proof is used to show that α_1 is the NPV of the cash flow

Variations

- We could be given a different expression to prove
  - We would then need to use proper algebra and whatnot to reach the new conclusion

Problem 5

Problem Statement

Solve and Explain the solution to 5.4 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

Original Question:

Norm of linear combination of orthonormal vectors. Suppose $a_1, \ldots, a_k$ are orthonormal n-vectors, and $x = \beta_1 a_1 + \cdots + \beta_k a_k$, where $\beta_1, \ldots, \beta_k$ are scalars. Express $\|x\|$ in terms of $\beta = (\beta_1, \ldots, \beta_k)$.

Solution

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

VMLS Solution:

$$
\begin{aligned}
\|x\|^2 &= x^T x \\
&= (\beta_1 a_1 + \cdots + \beta_k a_k)^T (\beta_1 a_1 + \cdots + \beta_k a_k) \\
&= \beta_1^2 + \cdots + \beta_k^2 \\
&= \|\beta\|^2.
\end{aligned}
$$

(Going from the third to the fourth expression we use $a_i^T a_i = 1$, $a_i^T a_j = 0$ for $j \neq i$.) So we have the simple formula $\|x\| = \|\beta\|$.

Problem 5 Summary

Procedure

- Use the definition of inner product to reach the conclusion

Key Concepts

- The inner product of orthonormal vectors will evaluate to just the norm of the coefficients

Variations

- We could be given a different type of vectors to show what the relationship of them is
  - We would then have to use the proper definitions and whatnot to show this final relationship

Problem 6

Problem Statement

Solve and Explain the solution to 5.5 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

Original Question

Orthogonalizing vectors. Suppose that a and b are any n-vectors. Show that we can always find a scalar γ so that $(a - \gamma b) \perp b$, and that γ is unique if $b \neq 0$. (Give a formula for the scalar γ.) In other words, we can always subtract a multiple of a vector from another one, so that the result is orthogonal to the original vector. The orthogonalization step in the Gram–Schmidt algorithm is an application of this.

Solution

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

Two vectors are orthogonal if their inner product is zero. So we want to find γ such that $(a - \gamma b)^T b = a^T b - \gamma b^T b = 0$. If $b = 0$, then we can choose any γ, and we have $(a - \gamma b) \perp b$, since all vectors are orthogonal to 0. If $b \neq 0$, then $b^T b = \|b\|^2 \neq 0$, and we can take $\gamma = a^T b / b^T b$.

Explanation:

We know that two vectors (in this case lets call them α and β) are said to be orthogonal if their inner product is zero. Namely,

$$\alpha^T \beta = 0. \tag{1}$$

So, from the problem statement this means that

$$(a - \gamma b)^T b = 0 \tag{2}$$

must be true. Expanding equation (2) we then have

$$\begin{aligned}
(a - \gamma b)^T b &= 0 && \text{(Premise)} && (3)\\
(a^T - \gamma b^T) b &= 0 && \text{(Distribution of transpose)} && (4)\\
a^T b - \gamma b^T b &= 0 && \text{(Distribution of inner product)} && (5)\\
a^T b &= \gamma b^T b && \text{(Simplification by addition)} && (6)\\
\frac{a^T b}{b^T b} &= \gamma && \text{(Simplification by division)} && (7)\\
\gamma &= \frac{a^T b}{b^T b} && \text{(Commutative property)} && (8)
\end{aligned}$$

where equation (8) is the result of several operations involving inner products. These operations are valid by the distributivity property of transposes and since inner products of vectors are scalar values. It should be noted that the result in equation (8) is only valid if b is not the zero vector. This of course means that our value for γ is then

$$\gamma = \frac{a^T b}{b^T b}. \tag{9}$$

Problem 6 Summary

Procedure

- Use the definition of orthogonal to show what the inner product must evaluate to
- Use the distributive property of transposes and inner products with algebra to solve for γ

Key Concepts

- We can find a scalar value in this inner product relationship with the use of properties and algebra

Variations

- We could be asked to find this value for a different condition of vectors
  - We would then use the properties of whatever is handed to us to find this value

Problem 7

Problem Statement

Solve and Explain the solution to 5.6 here in your own words. (Since you are given a solution, you will be graded on your ability to explain).

Original Question:

Gram–Schmidt algorithm. Consider the list of n n-vectors

$$a_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \; a_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \; \ldots, \; a_n = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

(The vector a_i has its first i entries equal to one, and the remaining entries zero.) Describe what happens when you run the Gram–Schmidt algorithm on this list of vectors, i.e., say what $q_1, \ldots, q_n$ are. Is $a_1, \ldots, a_n$ a basis?

Solution

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

VMLS Solution:

$\tilde{q}_1 = a_1$ has norm one, so we have $q_1 = a_1 = e_1$. $\tilde{q}_2 = a_2 - (q_1^T a_2)q_2 = e_2$ also has norm one, so $q_2 = e_2$. This pattern continues, and we see that $q_i = e_i$.

Yes, they are a basis. We know this since the Gram–Schmidt algorithm terminates after processing all n vectors. We can also see this directly. Suppose that $\beta_1 a_1 + \cdots + \beta_n a_n = 0$. The last entry of this vector is β_n, so we know that $\beta_n = 0$. This implies that $\beta_1 a_1 + \cdots + \beta_{n-1} a_{n-1} = 0$. The $(n-1)$ entry of this vector is β_{n-1}, so we know that $\beta_{n-1} = 0$. This continues, and we conclude that all β_i are zero. So the vectors are linearly independent, and therefore a basis.

Explanation:

The first step of the Gram–Schmidt algorithm is orthogonalization. Namely,

$$\tilde{q}_i = a_i - \sum_{j=1}^{i-1} (q_j^T a_i) q_j. \tag{1}$$

From a quick inspection of a_1 we can see that a_1 is the unit vector e_1. If we then move on to a_2 we see

$$\begin{aligned}
\tilde{q}_2 &= a_2 - \sum_{j=1}^{i-1} (q_j^T a_i) q_j \qquad (2) \\
&= a_2 - \left((q_1^T a_2) q_1 \right) \qquad (3) \\
&= \begin{bmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \left(\begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - (1) \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = e_2. \qquad (4)
\end{aligned}$$

Just as q_1 was the unit vector e_1, we see that q_2 is also a unit vector e_2. One can infer that this pattern will repeat for all vectors a_i and thus each vector a_i is a unit vector. Namely,

$$\tilde{q}_i = e_i. \tag{5}$$

Now, since each $\tilde{q}_i$ is a unit vector, the only way that

$$\sum_{i=i}^{n} \beta_i \tilde{q}_i = 0 \tag{6}$$

is a true statement is if all values of β are zero. This implies that the vectors $\tilde{q}_i$ are ***linearly independent***. This further implies that a_i form a basis.

Problem 7 Summary

Procedure

- Use the Gram–Schmidt algorithm to find what these vectors will evaluate to in the given collection

Key Concepts

- For this given set of vectors we can show that after the first step of Gram–Schmidt the vectors are unit vectors

Variations

- We could be given a different set of vectors and be asked to use the Gram–Schmidt algorithm on them
  - We would then use the Gram–Schmidt algorithm on them and conclude what they are after