Linear Algebra: The Foundation of Image Processing and Artificial Intelligence

Zafran A. Arif

2023

Linear algebra is a branch of mathematics that deals with vector spaces and linear mappings between those spaces. It is a fundamental area of mathematics that has numerous applications in various fields, including physics, engineering, computer science, economics, and more. Here are some key concepts and topics that are covered in this book:

1. **Vectors**
   Vectors are mathematical objects that have both magnitude and direction. They can be represented as lists of numbers or points in n-dimensional space.

2. **Vector Spaces**
   A vector space is a set of vectors that is closed under vector addition and scalar multiplication. Vector spaces have properties like closure, associativity, commutativity, and distributivity.

3. **Subspaces**
   A subspace of a vector space is a subset that is itself a vector space, containing the zero vector and closed under vector addition and scalar multiplication.

4. **Linear Combinations**
   A linear combination is a mathematical operation involving multiplying vectors by scalars and then adding them together to create a new vector.

5. **Spans**
   The span of a set of vectors is the set of all possible vectors that can be formed by taking linear combinations of the original vectors from the set.

6. **Linear Independence**
   A set of vectors is linearly independent if no vector in the set can be written as a linear combination of the others. Linearly independent sets of vectors are essential for forming bases.

7. **Bases and Dimension**
   A basis for a vector space is a set of linearly independent vectors that spans the entire space. The dimension of a vector space is the number of vectors in its basis.

8. **Linear Transformations**
   A linear transformation is a function that takes vectors from one vector space and maps them to vectors in another vector space while preserving vector addition and scalar multiplication.

9. **Matrices**
   Matrices are rectangular arrays of numbers that can be used to represent linear transformations and solve systems of linear equations.

10. **Eigenvalues and Eigenvectors**
    Eigenvalues and eigenvectors are associated with square matrices and represent properties of linear transformations, such as stretching or rotating.

11. **Inner Products and Norms**
    Inner products and norms are ways to measure the ”size” and ”angle” between vectors in a vector space. They are essential for defining concepts like orthogonality and distance.

12. **Determinants**
    The determinant of a square matrix is a scalar value that represents a linear transformation's scaling factor. It plays a crucial role in solving systems of linear equations and computing areas/volumes.

13. **Orthogonality**
    Orthogonal vectors are perpendicular to each other, and orthogonal bases simplify various calculations.

Application of Linear Algebra
Linear algebra is a foundational subject in mathematics, and its concepts and techniques are used extensively in many areas of science and engineering to model, analyze, and solve a wide range of real-world problems. It provides the mathematical framework for understanding and manipulating multi-dimensional data and transformations.

1 Vectors

In mathematics, particularly in the field of linear algebra, a vector is a fundamental mathematical object that represents both magnitude and direction. Vectors can be thought of as arrows in space, with a starting point (the ”tail”) and an ending point (the ”head”). They are used to represent quantities that have both a magnitude and a direction, such as force, velocity, displacement, and more.

A vector is typically represented as an ordered list of numbers, often arranged in a column, or as a point in n-dimensional space. In two-dimensional space, a vector might be written as:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

where v_1 and v_2 are real numbers, and $\mathbf{v}$ represents a vector in two-dimensional Euclidean space. In three-dimensional space, it might be written as:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

Vectors can be added together and multiplied by scalars, leading to various vector operations. Some of the common operations on vectors include vector addition, scalar multiplication, dot product, cross product (in three dimensions), and vector subtraction.

Vectors play a central role in linear algebra, where they are used to define vector spaces, study linear transformations, and solve systems of linear equations. They have applications in numerous areas of mathematics, physics, engineering, computer science, and various other fields.

2 Vector Spaces

2.1 Definition

A **vector space** V over a field $\mathbb{F}$, a set on which two operations (addition and scalar multiplication) are defined so that for $x, y \in V$, there's a unique element $x + y \in V$ The following properties must be satisfied:

1. **Addition is Commutative**
   Let f and g be two even functions, then $f + g = g + f$
   $$(f + g)(t) = f(t) + g(t) = g(t) + f(t) = (g + f)(t)$$
2. **Addition is Associative**
   Let f, g, and h be even functions, then $(f + g) + h = f + (g + h)$
   $$((f + g) + h)(t) = (f(t) + g(t)) + h(t) = f(t) + (g(t) + h(t)) = (f + (g + h))(t)$$
3. **Existence of Additive Identity**
   Let the additive identity for a function be $f(t) = 0$, thus for any even function f,
   $$f(t) = (f + 0)(t) = f(t) + 0(t)$$
4. **Existence of Additive Inverse**
   For any function f, let there be an inverse function $-f$ where the sum of them is 0,
   $$(f + (-f))(t) = f(t) + (-f(t)) = 0$$
5. **Existence of Multiplicative Identity**
   For every function f, let there be a scalar $a \in \mathbb{R}$ such that $af = f$. Let $a = 1$, then
   $$a \cdot f = 1 \cdot f = f$$
6. **Scalar Multiplication is Associative**
   Let f be an even function, there exist scalars a and b where $(ab)f = a(bf)$
   $$(ab)f(t) = (a \cdot b) \cdot f(t) = a \cdot (b \cdot f(t)) = a(bf(t))$$
7. **Distributive Law holds for Vectors**
   Let f and g be two even functions, there exist a scalar a where $a(f + g) = af + ag$
   $$a \cdot (f + g)(t) = af(t) + ag(t)$$
8. **Distributive Law holds for Scalars**
   Let f be an even function, there exist scalars a and b where $(a + b)f = af + bf$
   $$(a + b) \cdot f(t) = af(t) + bf(t)$$

2.2 Lemma

2.2.1

Let V be a vector space over a field $\mathbb{F}$. The additive identity of V is unique.

Proof

Step 1: Figure out what an additive identity is.
Step 2: Come up with a strategy for showing something is unique.

Suppose the additive identity is not unique. Then there are at least two distinct element x and y, who are additive identities of V. This means that for every $v \in V$,

$x + v = v$ and $v = v + y$ by VS3

First set $v = y$ to see that $x + y = y$, then set $v = x$ to see that $x + y = x$. Thus $x = x + y = y$, contradicting that they are distinct. Hence the additive identity is unique.

2.2.2

Let V be a vector space over a field $\mathbb{F}$. $\forall x \in V$, there is a *unique* element $y \in V$ s.t. $x + y = y + x = 0$. Note: this establishes that the additive inverse is unique.

Proof

Fix an $x \in V$. Suppose that the additive inverse is not unique. There are at least two distinct element a and b that serve as the additive inverse of x.

Notice $x + a = 0$ and $x + b = 0$ by VS4.

So $x + a = x + b$ by the first lemma.

Then adding a to both expressions: $a + (x + a) = a + (x + b)$

So $(a + x) + a = (a + x) + b$ by VS2 and $0 + a = 0 + b$ by VS4.

Thus $a = b$ by VS3.

2.2.3

Let V be a vector space over a field $\mathbb{F}$. Let $x, y, z \in V$. If $x + z = y + z$, then $x = y$.

Proof

Let $x, y, z \in V$ such that $x + z = y + z$, and let a be the additive inverse of z. Then adding a to the right side of each expressions we get:

$(x + z) + a = (y + z) + a$ and $x + (z + a) = y + (z + a)$ by VS2.

Thus $x = y$

2.2.4

Let V be a vector space over a field $\mathbb{F}$. Then
1. For $0 \in \mathbb{F}$ and $\forall x \in V$, $0x = 0$.
2. $\forall a \in \mathbb{F}$ and $\forall x \in V$, $(-a)x = a(-x) = -(ax)$.
3. $\forall a \in \mathbb{F}$ and $0 \in V, a0 = 0$.

Proof

Think: If I multiply the additive identity of the field by a vector, I get the additive identity of the vector space.

1. Notice $0x = (0 + 0)x = 0x + 0x$ by VS8 and properties of the additive inverse. By VS4, $0x$ has an additive inverse $-(0x)$.
Thus $0x + (-(0x)) = (0x + 0x) + (-(0x))$.
So $0x + (-(0x)) = 0x + (0x + (-(0x))$ by VS2.
$0 = 0x + 0$ by VS4, and finally $0 = 0x$ by VS3.

2. ***FINISH THIS***

3 Subspaces and Subsets

3.1 Definition: Subspace

A **subspace** of a vector space over a field $\mathbb{F}$, is a nonempty subset of the vector space that is also a vector space over $\mathbb{F}$

Notice: VS1, 2, 5, 6, 7, 8 hold for any nonempty subset of V, but $0 \in V$ might not be in the subset, nor the additive inverse, ***FINISH THIS***

3.2 Definition: Subset

A **subset** S of a vector space V generates (or spans) V provided Span(S) $= V$. We might also phrase this as **the set S spans V.**

3.3 Lemma

3.3.1

Let V be a vector space over a field $\mathbb{F}$. Let W be a subset of V. Then W is a subspace of V iff the following holds:
1. $0 \in W$
2. $\forall x, y \in W, x + y \in W$
3. $\forall x \in W, \forall a \in \mathbb{F}, ax \in W$

Proof

$\Rightarrow$ Assume W is a subspace. Because a subspace is itself a vector space, by VS3, there's an additive identity, and W must be closed under addition and multiplication.

$\Leftarrow$ From above, we need to show that $\forall x \in W$, there's an additive inverse also in W. Let -1 be the additive inverse of $1 \in \mathbb{F}$. Using the third property, we know that $(-1)x \in W$. Notice $(-1)x + x = (-1)x + (1)x = (-1 + 1)x = 0x = 0$, so $(-1)x$ is the additive inverse of x. We used VS5, 8, and the lemma. Thus VS4 is true for W. As observed above, to establish that is itself a vector space, we need closure under addition and scalar multiplication, as well as VS4, since VS1, 2, 5, 6, 7, 8 hold for any subset of V.

Standard template to establish if W is a subspace:
A. $0 \in W$ because ...
B. Let $x, y \in W$, then $x + y \in W$ because ...
C. Let $x \in W$ and let $a \in \mathbb{F}$, then $ax \in W$ because ...

3.3.2

Any intersection of subspaces of a vector space V is a subspace of V.

Proof
Let V_i be a collection of subspaces where $i \in J$ for some index set J. That is $W = \bigcap_{i \in J} V_i$

A. $0 \in W$ because $\underline{\text{0 must be in every subspace.}}$
B. Let $x, y \in W$, then $x + y \in W$ because x, y are in every V_i, and the V_i are closed under addition, so $x + y \in V_i$. Hence it's in W.
C. Let $x \in W$ and let $a \in \mathbb{F}$, then $ax \in W$ because x is in every V_i, and the V_i are closed under scalar multiplication, so $ax \in V_i$. Hence it's in W.

3.4 Examples of Subspace

3.4.1

Is $W = \{f \in \mathbb{P}_3(\mathbb{R} : f(1) = 0\}$ a subspace of $\mathbb{P}_3(\mathbb{R})$?
$\Rightarrow$
A. $0 \in W$ because $\underline{\text{because the zero polynomial evaluated at the point 1 is indeed 0.}}$
B. Let $f, g \in W$, then $f + g \in W$ $\underline{\text{because } (f+g)(1) = f(1) + g(1) = 0 + 0 = 0}$
C. Let $x \in W$ and let $a \in \mathbb{F}$, then $\underline{af \in W \text{ because } af(1) = a \times f(1) = a \times 0 = 0}$

$\therefore W$ is a subspace.

3.4.2

Is $\mathbb{P}_2$ a subspace of $\mathbb{P}_3$?
$\Rightarrow$Verbally, we agree that it is a subspace of $\mathbb{P}_3$

3.4.3

Is W = the zero polynomial a subspace of $\mathbb{P}$?
$\Rightarrow$ W is indeed a subspace of $\mathbb{P}$.

3.4.4

Is $\mathbb{R}^2$ a subspace of $\mathbb{R}^3$?
$\Rightarrow$
$\binom{2}{3}$ is in $\mathbb{R}^2$. $(2:3:4)$ is in $\mathbb{R}^3$. They are nothing alike. So it is not correct. A plane through the origin in $\mathbb{R}^3$ is two dimensional subspace of $\mathbb{R}^3$. It is not $\mathbb{R}^2$.

3.4.5

Fix $q \in \mathbb{R}$. Let $W = \{(a : b : q)|a, b \in \mathbb{R}$. Is W a subspace of $\mathbb{R}^3$?
$\Rightarrow$
If $q \neq 0$, then W is not a subspace since $(0 : 0 : q) \neq (0 : 0 : 0)$
If $q = 0$, then W is a subspace since $(0 : 0 : q) = (0 : 0 : 0)$

3.4.6

$\{[0, r]|r \in \mathbb{R}\}$ is an uncountable set of sets.

3.5 Example of Subset

3.6

Let $S = \{1 + x, x + x^2, 1 + x + x^2\}$. Does S span $\mathbb{P}_1$? span $\mathbb{P}_2$? span $\mathbb{P}_3$?

$\Rightarrow$
No. Because x^2 is not in $\mathbb{P}_1$
Yes. With a little work you can show every vector in $\mathbb{P}_2$ can be written as a linear combo of things in S, and $S \subseteq P2$
No. There is no way to get x^3 by taking linear combination of vectors in S.

4 Linear Combinations and Span

4.1 Definition: Linear Combination

Let V be a vector space over a field F. Let $S \subseteq V$ be nonempty. A vector $v \in V$ is called a **linear combination** of the vectors of S if there exist $u_1, u_2, ..., u_m \in S$ and $a_1, a_2, ..., a_m \in \mathbb{F}$ s.t. $v = a_1u_1 + a_2u_2 + ... + a_mu_m$. Note: m is a finite positive integer.

Given a set S of actual vectors in $\mathbb{R}^n$, you should be able to tell if a given vector is a linear combination of the vectors in S. (Hint: put them is a matrix and row reduce)

4.2 Definition: Span

Let V be a vector space over a field $\mathbb{F}$. Let $S \subseteq V$ be nonempty. The **Span** of S, denoted Span(S), is the set of all finite linear combinations of vectors in S. If $S = \{\}$, by convention, we say span$(S) = \{0\}$.

4.3 Theorem

Let V be a vector space over a field $\mathbb{F}$. If $S \subseteq V$, then span(S) is a subspace of V that contains S. Moreover, any subspace of V that contains S also contains span(S).

Proof

Assume $S \subseteq V$. Show that span(S) is a subspace by satisfying these properties:
A. $0 \in span(S)$ because if we take any $x \in S$, then $0x = 0 \in span(S)$
B. Let $x, y \in span(S)$, then $x + y \in span(S)$ because X, y are finite linear combinations of vectors from S, and adding together these two creates another finite linear combinations of elements in S.
C. Let $x \in span(S)$ and let $a \in \mathbb{F}$, then $ax \in span(S)$ because ax can be created by writing x as a linear combinations of elements in S and then multiplying each scalar by a to get another linear combination of vectors in S.

Next we observe that $S \subseteq span(S)$ since if $x \in S$, then $1x \in span(S)$.

Finally, let W be a subspace of V s.t. $S \subseteq W$. Let $u_1, u_2, ..., u_m \in S$ and $a_1, a_2, ..., a_m \in \mathbb{F}$. We know that $u_1, u_2, ..., u_m \in W$ since $S \subseteq W$, and $a_1u_1, a_2u_2, ..., a_mu_m \in W$ since W is closed under scalar multiplication.

Next, we observe $(...((a_1u_1 + a_2u_2) + a_3u_3)... + a_mu_m)$ is in W because each time we are adding two vectors in W to get another vector in W, which we then add to the next vector in W until the last vector has been combined into the sum. Hence, $v = a_1u_1 + a_2u_2 + ... + a_mu_m \in W$.

4.4 Example

Let $\mathbb{P}$ be the vector space of all polynomials with coefficients from $\mathbb{R}$. Let $S = \{1 + 2x, 2x + x^3, 1 + 3x^2\}$. Let $f = 2 + 4x + 3x^2 + x^3$ and $g = 3 + 5x + 3x^2 + x^3$.

a) Determine if $f \in span(S)$

b) Determine if $g \in span(S)$

Let $a, b, c \in \mathbb{R}$. We want to test if $f = 2 + 4x + 3x^2 + x^3 = a(1 + 2x) + b(2x + x^3) + c(1 + 3x^2)$
$2 = a + 0b + c$
$4 = 2a + 2b + 0c$
$3 = 0a + 0b + 3c$
$1 = 0a + b + 0c$

In the matrix form: $\begin{pmatrix} 1 & 0 & 1 & 2 \\ 2 & 2 & 0 & 4 \\ 0 & 0 & 3 & 3 \\ 0 & 1 & 0 & 1 \end{pmatrix} \xrightarrow{RowReduce} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

$a = 1; b = 2; c = 1$; Thus $f = 2 + 4x + 3x^2 + x^3 = 1(1 + 2x) + 2(2x + x^3) + 1(1 + 3x^2)$

Now, let's find $g = 3 + 5x + 3x^2 + x^3 = a(1 + 2x) + b(2x + x^3) + c(1 + 3x^2)$

In the matrix form: $\begin{pmatrix} 1 & 0 & 1 & 3 \\ 2 & 2 & 0 & 5 \\ 0 & 0 & 3 & 3 \\ 0 & 1 & 0 & 1 \end{pmatrix} \xrightarrow{RowReduce} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

The last row of the reduced echelon form of our matrix indicates that there is no solution for this system of equations. Thus, $g \notin \text{Span}(S)$.

4.5 Subset

4.5.1 Definition

A **subset** S of a vector space V generates (or spans) V provided Span(S) $= V$. We might also phrase this as **the set S spans V.**

5 Linear Dependence and Linear Independence

5.1 Definition: Linearly Dependent

Let S be a subset of a vector space V over a field $\mathbb{F}$. The set S is **linearly dependent** (LD) provided there are vectors $x_1, x_2, ..., x_n \in S$ and <u>nonzero</u> coefficients $c_1, c_2, ..., c_n \in \mathbb{F}$, such that $c_1x_1 + c_2x_2 + ... + c_nx_n = 0$.

5.2 Definition: Linearly Independent

Let S be a subset of a vector space V over a field $\mathbb{F}$. The set S is **linearly independent** (LI) if it is not linearly dependent. One of the ways we show a set is LI is to show that any time we write

$$c_1x_1 + c_2x_2 + ... + c_nx_n = 0$$

for $x_1, ..., x_n \in S$ and $c_1, .., c_n \in \mathbb{F}$, then $c_1, .., c_n$ must all be zero.

5.3 Examples

5.3.1

A permutation matrix is any matrix obtained from the identity by interchanging rows. Is the set of $n \times n$ permutation matrices a LI subset of $M_{n\times n}(\mathbb{R}$ for $n = 2$? What about $n = 3$?

$\Rightarrow n = 2; a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

Equating the 1,1 entries $a(1) + b(0) = 0$, which implies $a = 0$.

Equating the 1,2 entries $a(0) + b(1) = 0$, which implies $b = 0$.

Thus $\{\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix}\}$ is linearly independent.

$$\Rightarrow n=3;\{\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix},\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix},\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix},\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix},\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}\}$$

$$a\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}+b\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}+c\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}+d\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}+e\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}+f\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}=\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}$$

$$-c\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}+c\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}+c\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}+c\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}-c\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}-c\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}=\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}$$

If we choose c to be any nonzero number, we get a linear combination that adds to zero. Therefore the set is linearly dependent.

5.3.2 GQE 2014 Exam Question

A permutation matrix is any matrix obtained from the identity by interchanging rows. Show that the set of $n\times n$ permutation matrices is a LD subset of $M_{n\times n}(\mathbb{R})$, for $n\geq 4$

$\Rightarrow$ Using what we have done already,
Take $c\neq 0$, then

$$-c\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&I\end{pmatrix}+c\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&I\end{pmatrix}+c\begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&I\end{pmatrix}+c\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&1&0\\0&0&0&I\end{pmatrix}-c\begin{pmatrix}0&0&1&0\\1&0&0&0\\0&1&0&0\\0&0&0&I\end{pmatrix}-c\begin{pmatrix}0&1&0&0\\0&0&1&0\\1&0&0&0\\0&0&0&I\end{pmatrix}=\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

Three important problem solving ideas are used here:

1. Guess and Check
2. Start small and generalize
3. What works for 2×2 may not be true in general

5.3.3

Is $S=\{1,x,x^2,...\}$ a linearly independent subset of $\mathbb{P}$

$\Rightarrow$ Yes. Take any finite sum $a_1x^{m_1}+a_2x^{m_2}+...+a_nx^{m_n}=0$ with the m_i unique.
Recall that two polynomials are equal exactly when every coefficient of every power of x agrees. Thus $a_1=a_2=...=a_n=0$.

5.3.4

Let $S = \{1, i\}$ where $i^2 = -1$ and $W_\mathbb{R} = \text{Span}(S)$ over $\mathbb{R}$ and $W_\mathbb{C} = \text{Span}(S)$ over $\mathbb{C}$
$W_\mathbb{R} = \{a + bi | a, b \in \mathbb{R}$ and $W_\mathbb{C} = \{a + bi | a, b \in \mathbb{C}$

1. Are $W_\mathbb{R}$ and $W_\mathbb{C}$ equal as sets? YES!
2. Are $W_\mathbb{R}$ and $W_\mathbb{C}$ equal as vector spaces? Although V is the same for both, the field $\mathbb{F}$ is not the same. They are NOT the same vector space.
3. Is S LI in $W_\mathbb{R}$? YES! $a \neq -bi$ if a and b are both real.
4. Is S LI in $W_\mathbb{C}$? NO! $-i(1) + 1(i) = 0$, a nonzero linear combination that adds to zero.

5.4 Theorem

Let V be a vector space over a field $\mathbb{F}$. Let $S_1 \subseteq S_2 \subseteq V$. If S_1 is a linearly dependent subset, then S_2 is a linearly dependent subset.

Proof

6 Bases and Dimension

6.1 Definition: Bases

Let V be a vector space over a field $\mathbb{F}$. A subset β of V is a **basis** for V provided β is linearly independent and spans V. We say β spans V if $\text{Span}(\beta) = V$.

6.2 Theorem

Let V be a vector space over a field $\mathbb{F}$. Suppose S is a subset of V with n vectors such that S spans V. Let L be a linearly independent subset of V containing m vectors. Then $m \leq n$ and there is a subset R of S containing $n - m$ vectors such that $L \cup R$ spans V.

Proof
We will induct on m =number of vectors in L. Inductive hypothesis: For any linearly independent set with m elements contained in a vector space V for which there is a finite spanning set n, we see that $m \leq n$ and there is a subset R of S containing $m - n$ elements such that $L \cup R$ is a spanning set for V.

Base case: $m = 0$
Let $L = \{\}$. The result follows trivially by setting $R = S$.
For a little more certainty, let's also establish the case $m = 1$. Assume $L = \{x_1\}$ and $S = \{u_1, ..., u_n\}$
Since S spans V and L is a subset of V ($L \subseteq V$) with one element, we know $m = 1 \leq n$.
Moreover, since S spans V, $x_1 = a_1u_1 + ... + a_nu_n$ where at leas one a_j is nonzero (since $x_1 \neq 0$ as it is part of a LI set). WLOG (without loss of generality) a_1 is nonzero, hence $u_1 = \frac{x_1-(a_2u_2+...+a_nu_n)}{a_1}$. Set $R = \{u_2, ..., u_n\}$, a set with $n - 1$ elements and notice that $L \cup R$ spans V since u_1 is a linear combo of elements in this set and it also contains $\{u_2, ..., u_n\}$

Inductive step:
Assume IH is true for m. Let L be linearly independent set with $m + 1$ elements in a vector space V with a finite spanning set S. $L = \{x_1, ..., x_m, x_{m+1}\}$. Let $K = \{x_1, ..., x_m\}$. The result is true for K. Thus, $m \leq n$ and there exist $Q \subseteq S$ so that $K \cup Q$ spans V.

Notice: We need $m + 1 \geq n$ and to put m_{x+1} back into the mix amending $K \cup Q \Rightarrow L \cup R$.
Since $K \cup Q$ is a spanning set, $x_{m+1} = a_1x_1 + ... + a_mx_m + b_1q_1 + ... + b_kq_k$ where $q1, ..., q_k \in Q \subseteq S$.
Since L is LI, we know that x_{m+1} is NOT a linear combo of the other x_j's so at least one of the b_l's is nonzero.
First this tells us that $Q \neq \{\}$, and hence $m < n$. WLOG $b_1 \neq 0$, so $q_1 = \frac{x_{m+1}-(a_1x_1+...+a_mx_m+b_1q_1+...+b_kq_k)}{b_1}$
Set $R = Q - \{q_1\}$. Then $L \cup R$ is a spanning set for V.

6.3 Corollary

Let V be a vector space with a finite basis over a field $\mathbb{F}$. Then every basis for V has the same number of elements.

Proof
Suppose β and γ are two bases for V.
Since β is LI and γ is a spanning set, β can't have more elements than γ.
Since γ is LI and β is a spanning set, γ can't have more elements than β.

6.4 Definition: Dimension

Let V be a vector space over a field $\mathbb{F}$. The vector space V is said to be **finite dimensional** if it has a basis consisting of a finite number of vectors. The unique number of vectors in a basis is called the **dimension** of V and is denoted by $\dim(V)$. A vector space that is not finite dimensional is called infinite-dimensional.

6.5 Theorem (two out of three ain't bad)

Let V be an n dimensional vector space over a field $\mathbb{F}$. Let S be a subset of V. Any two of the following properties implies the third:

(i) S spans V

(ii) S is linearly independent

(iii) S has n elements

Proof:

1. (i) and (ii) $\Rightarrow$ (iii):
   A LI spanning set is a basis, and all bases have the same number elements denoted as the dimension, so there must be n elements in the set.
2. (i) and (iii) $\Rightarrow$ (ii):
   If S is a spanning set with n elements, we know that S contains a basis. However, if we need to remove one or more vectors, our set will be too small to basis, thus it's already linearly independent.
3. (ii) and (iii) $\Rightarrow$ (i):
   If S is a linearly independent set with n elements, we know we can extend S to a basis, possibly by adding elements. If need to add elements to get a basis, well have too many elements, thus S must already have been a basis and hence a spanning set.

6.6 Theorem

Let V be a finite dimensional vector space over a field $\mathbb{F}$. If W is a subspace of V, then $\dim(W) \leq \dim(V)$. If they're equal, then $W = V$.

Proof:
Let β be a basis for W. Then β is a linearly independent subset of V, so it can be extended to a basis doe

V. Hence, $\dim(W) \leq \dim(V)$. If $\dim(W) = \dim(V)$, then β is a LI set with the right number of elements to be a basis for V, and hence $\text{span}\beta = W$ and $\text{span}\beta = V$, so $W = V$.

6.7 Example

Consider $W = \{(a; b; a\}|a, b \in \mathbb{R}\}$, a subspace of $\mathbb{R}^3$. A basis for W is $\{(1; 0; 1), (0; 1; 0)\}$. The standard basis for $\mathbb{R}^3$ is $\{(1; 0; 1), (0; 1; 0), (0; 0; 1)\}$

Can we remove a vector from the basis for $\mathbb{R}^3$ to get a basis for W?

Solution:
No. Any two will leave us with a row that us always zero when we take the span of these two vectors.

6.8 Corollary:

Let V be a finite dimensional vector space over a field $\mathbb{F}$. If β is a basis for a subspace W of V, then β can be extended to a basis for V.

7 Maximal Linearly Independent Sets

7.1 The Axiom of Choice

Given a (possibly infinite) collection of sets, you can choose one element from each set, and make a new set. It tends to be very useful in Abstract Algebra, but is somewhat controversial for analytical techniques because you can use it to create sets with no measure and duplicate a sphere (the Banack-Tarski Paradox).

7.2 Definition: Poset

A partially ordered set (poset) is a set $\mathcal{F}$ together with a relation $\subseteq$ that is reflexive, antisymmetric, and transitive.

7.2.1 Example

Let $S = \{\{a, b\}, \{a, b, c\}, \{a, b, c, d\}, \{a, b, d, e\}\}$. Let the relaation on S be the usual subset relation. So, for example $\{a, b\} \subseteq \{a, b, c\}$ and $\{a, b\} \subseteq \{a, b, d\}$, but $\{a, b, c\}$ and $\{a, b, d\}$ are not comparable.

7.3 Definition: Maximal

Let $\mathcal{F}$ be a poset. An element $M \in \mathcal{F}$ is said to be **maximal** in $\mathcal{F}$ if $M \subseteq N \in \mathcal{F}$ implies $M = N$.

7.3.1 Example

Working with the example above, $\{a, b\}$ is not maximal since $\{a, b\} \subseteq \{a, b, c\}$, however $\{a, b, c, d\}$ and $\{a, b, d, e\}$ are both maximal in S.

$T = S \cup \{\{c, e\}\}$. Is $\{c, e\}$ maximal in T? Yes, it's not contained in any of the sets listed in S.

7.4 Definition: Chain

A collection of elements (subset of) in $\mathcal{F}$ is called a **chain** C, if for each $A, B \in C$, either $A \subseteq B$ or $B \subseteq A$.

7.4.1 Example

Type...

7.5 Upper Bound

Let C be a chain in a poset $\mathcal{F}$. An element $U \in \mathcal{F}$ is an **upper bound** for C provided $\forall A \in C, A \subseteq U$.

7.5.1 Example

Type...

7.6 Zorn's Lemma (Also attributed to Hausdorff)

Let $\mathcal{F}$ be a partially order set. If every chain in $\mathcal{F}$ has an upper bound, then $\mathcal{F}$ contains a maximal element. In the linear algebra context, $\mathcal{F}$ will be sets of LI subsets of a vector space V, ad the relation $\subseteq$ will be actually be the subset relation. Zorn's lemma is equivalent to the Axiom of Choice.

7.7 Theorem

7.7.1

Let V be a vector space over a field $\mathbb{F}$. If β is a maximal LI subset of V, then β is a basis for V.

Proof:
Assume β is a maximal linearly independent subset of V. Let $v \in V$. If v can't be written as a finite linear combination of elements in β, then we could take $\beta \cup \{v\}$ and get a LI set that properly contains β and hence show that β is not a maximal linearly independent set, contradicting our assumptions.

7.7.2

Let V be a vector space over a field $\mathbb{F}$. If S is a linearly independent subset of V, then S is contained in a maximal linear independent subset of V, hence S is contained in a basis for V.

Proof:
Let S be linearly independent subset in V. Let $\mathcal{F}$ be the set of all sets of LI vectors that contain S.

Let C be a chain of LI subsets, where each subset contains S. Let U be the union of the sets in the chain C. So U is an upper bound with respect to subset inclusion, however we need to show it is linearly independent.

Suppose U is LD. Then there are vectors $v_1, ..., v_n \in U$ and nonzero scalars $a_1, ..., a_n$ so that there is a linear combination $a_1v_1 + ... + a_nv_n = 0$.
Since the chain is ordered by inclusion and the vectors are in the union, there must be a set in the chain that contains all the vectors. But this contradicts that the chain consisted of LI sets.

Thus, U itself must be LI and hence every chain of LI vectors has an upper bound. By Zorn's Lemma, there is a maximal LI subset β. By the previous theorem, β is a basis for V.

8 Linear Transformations, Null Spaces, and Ranges

8.1 Definition: Linear Transformation

Let V and W be vector spaces over a field $\mathbb{F}$. A function $T : V \to W$ is a **linear transformation** provided $\forall x, y \in V, \forall a \in \mathbb{F}$,

(a) $T(x + y) = T(x) + T(y)$

(b) $T(ax) = aT(x)$

8.2 Examples

8.2.1

When is the line: $T(x) = mx + b$, where $m, b \in \mathbb{F}$, a linear transformation?
$\Rightarrow$

$$T(x_1 + x_2) = m(x_1 + x_2) + b$$
$$T(x_1) + T(x_2) = (m(x_1) + b) + (m(x_2) + b)$$

These two will be equal when $b = 0$.
If $b = 0$, then $T(cx) = m(cx) = c(mx) = cT(x)$. Thus $T(x) = mx + b$ is a linear transformation if and only if $b = 0$.

8.2.2

What about the quadratic: $T(x) = ax^2 + bx + c$?
$\Rightarrow$ T is a linear transformation if and only if both $a = 0$ and $c = 0$.

8.2.3

Is the differential operator $D : \mathbb{P} \to \mathbb{P}$ a linear transformation?
$\Rightarrow$ Yes.

8.2.4

How about integration on polynomials: i.e. $T(f) = \int f(x)dx = F(x) + C$, where $F' = f$ and C is an arbitrary constant?
$\Rightarrow$
The indefinite integral is not unique - it depends on the choice of C, thus this T is not even a function.
The identity transformation is $T(x) = x$, and it is a linear transformation from V to V.
The zero transformation $T_0(x) = 0$ is a linear transformation from any vector space V into any vector space W.

8.3 Definition: Null Spaces

Let V and W be vector spaces over a field $\mathbb{F}$ and $T : V \to W$ is a linear transformation. The **null space (or kernel)** of T, denoted $N(T)$, is the set of all $x \in V$ such that $T(x) = 0$. (the set of all vectors/things that maps to zero)

$$N(T) = \{x \in V | T(x) = 0\}$$

8.4 Definition: Ranges

Let V and W be vector spaces over a field $\mathbb{F}$ and $T : V \to W$ is a linear transformation. The **range** of T, denoted $R(T)$, is the set of all $y \in W$ such that $\exists x \in V$, so that $T(x) = y$.

$$R(T) = \{T(x) | x \in V\}$$

8.5 Theorem

Let V and W be vector spaces over a field $\mathbb{F}$ and $T : V \to W$ is a linear transformation. The null space and range of T are subspace of V and W respectively.

Proof:
In each case we need to show

(i) $0 \in$ set

(ii) $x, y \in$ set implies $x + y \in$ set

(iii) $x \in$ set, $a \in \mathbb{F}$ implies $ax \in$ set

Details left to reader.

8.6 Definition: Nullity and Rank

Let V and W be vector spaces over a field $\mathbb{F}$ and $T : V \to W$ is a linear transformation. The **nullity** of T is the dimension of the null space of T and the **rank** of T is the dimension of the range of T.

8.7 Theorem

Let V and W be vector spaces over a field $\mathbb{F}$ and $T : V \to W$ is a linear transformation. If β is a basis for V, then $T(\beta)$ spans the range of T.

Proof:

Let $w \in R(T)$. Then there is a $v \in V$ so that $w = T(v)$. Since β is a basis for V, there exist vectors $u_1, ..., u_n \in \beta$ and scalars $a_1, ..., a_n \in \mathbb{F}$ so that $v = a_1u_1 + ... + a_nu_n$. Notice

$$w = T(v) = T(a_1u_1 + ... + a_nu_n) = T(a_1u_1) + ... + T(a_nu_n) = a_1T(u_1) + ... + a_nT(u_n)$$

so w is a linear combination of elements from $T(\beta)$, establishing the result.

8.8 Example

Let $T : \mathbb{R}^3 \to \mathbb{R}^4$ be defined by $T \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a \\ 0 \\ c \\ 0 \end{bmatrix}$

$\beta = \{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \}$, $\beta = \{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \}$

Notice that $T(\beta)$ is linearly dependent.

Aside: $N(T) = \{ \begin{pmatrix} 0 \\ b \\ 0 \end{pmatrix} | b \in \mathbb{R}\}$ and $R(T) = \{ \begin{pmatrix} a \\ 0 \\ c \\ 0 \end{pmatrix} | a, c \in \mathbb{R}\}$

8.9 Theorem

Let V and W be vector spaces over a field $\mathbb{F}$ and $T : V \to W$ is a linear transformation. If V is finite dimensional, then

$$rank(T) + nullity(T) = dim(V)$$

Proof:
Let α be a basis for $N(T)$. So the number of elements in α is equal to the nullity of T. Let $\alpha = \{v_1, v_2, ..., v_n\}$. Extend α to a basis β for V. Write $\beta = \{v_1, v_2, ..., v_n, v_{n+1}, ..., v_p\}$. Here $n = nullity(T)$ and $p = dim(V)$. What happens if we apply T to the elements of β? $T(v_1) = 0, ..., T(v_n) = 0$ since we are applying T to vectors in the null space of T. Set $\gamma = \{T(v_{n+1}), ..., T(v_p)\}$. We will show that gamma is a basis for $R(T)$. Since span of γ = span T(β), we know γ spans $R(T)$.

Suppose γ is linearly dependent, then there exist $a_{p+1}, ..., a_n$ not all zero so that

$$a_{p+1}T(v_{p+1}) + ... + a_nT(v_n) = 0$$

Since T is linear,

$$T(a_{p+1}v_{p+1} + ... + a_nv_n) = 0$$

But then $a_{p+1}v_{p+1} + ... + a_nv_n \in N(T)$. Since α is a basis for $N(T)$, we have

$$a_{p+1}v_{p+1} + ... + a_nv_n = b_1v_1 + ... + b_pv_p$$

But then $b_1v_1 + ... + b_pv_p - (a_{p+1}v_{p+1} + ... + a_nv_n) = 0$, but this is a linear combination of the elements in the basis β, hence all the coefficients are zero, contradicting our earlier assumption that γ was linearly dependent. Thus, γ is a linearly independent spanning set for $R(T)$ and hence the $rank(T) = n-p = dim(V)-nullity(T)$ completing the proof.

8.10 Theorem

Let V and W be vector spaces over a field $\mathbb{F}$ and $T : V \to W$ is a linear transformation. The transformation T is one-to-one if and only if $N(T) = \{0\}$.

Proof:
Instead we could try to prove T is not one-to-one if and only if $N(T) \neq \{0\}$.
T is not one-to-one if and only if there exist $u, v \in V$ so that $u \neq v$ yet $T(u) = T(v)$.
If and only if there exist $u, v \in V$ so that $u \neq v$ yet $T(u) - T(v) = 0$.
If and only if there exist $u, v \in V$ so that $u \neq v$ yet $T(u - v) = 0$.
If and only if there exist $u, v \in V$ so that $u \neq v$ and $(u - v) \in N(T)$.

If and only if there exist $u, v \in V$ so that $u \neq v$, and thus $w = u - v \neq 0$ but $T(w) = 0$
If and only if there exist $w \in V, w \neq 0$ such that $T(w) = 0$
If and only if $N(T) \neq \{0\}$

8.11 Theorem

Let V and W be finite dimensional vector spaces over a field $\mathbb{F}$ and $T : V \to W$ a linear transformation. If $dim(V) = dim(W)$, then the following are equivalent

1. T is one-to-one
2. T is onto
3. $rank(T) = dim(V) = dim(W)$

Proof:
Throughout the proof we assume $dim(V) = dim(W)$. The result is NOT true otherwise. We will employ the proof strategy of showing (1) $\Rightarrow$ (3) $\Rightarrow$ (2) $\Rightarrow$ (1).

(1) $\Rightarrow$ (3) : Assume T is one-to-one. Then $N(T) = \{0\}$ and hence nullity$(T) = 0$ and we know $rank(T) + nullity(T) = dim(V) = dim(W)$, and thus $rank(T) = dim(V) = dim(W)$.

(3) $\Rightarrow$ (2) : Assume $rank(T) = dim(V) = dim(W)$. Since $R(T) \subseteq W$ and $rank(T) = dim(W)$, we can conclude that $R(T) = W$ establishing that T is onto.

(2) $\Rightarrow$ (1) : Assume T is onto. Then $rank(T) = dim(W) = dim(V)$, and hence since $nullity(T) = dim(V) - rank(T) = 0$, we can conclude that $N(T) = \{0\}$ and hence T is one-to-one.

9 The Matrix Representation of a Linear Transformation

9.1 Definition: Ordered Basis

Let V be a vector space over a field $\mathbb{F}$. A basis with a fixed ordering on the elements is called an **ordered basis**. We abuse the standard definition of a set from here on in.

9.2 Definition: Coordinate Vector of v relative to β

Let V be a vector space over a field $\mathbb{F}$. Let $\beta = \{u_1, u_2, ..., u_n\}$ be an ordered basis for V. $\forall v \in V, \exists$ unique $a_1, a_2, ..., a_n \in \mathbb{F}$ such that $v = \sum_{j=1}^{n} a_j u_j$. We define the **coordinate vector of v relative to β** by

$$[v]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

9.3 Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and let $\beta = \{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\}$

$A = \{1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 3 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 4 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\}$

$[A]_\beta = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$

9.4 Definition: Matrix Representation of T relative to β and γ

Let V and W be a finite dimensional vector spaces over a field $\mathbb{F}$ with ordered bases $\beta = \{v_1, v_2, ..., v_m\}$ and $\gamma = \{w_1, w_2, ..., w_n\}$ respectively. Let $T : V \to W$ be a linear transformation. The **matrix representation of T relative to β and γ is the matrix** $A = [a_{ij}]$, where

$$T(v_j) = \sum_{i=1}^{n} a_{ij} w_i$$

as denoted by $[T]_\beta^\gamma$.
i.e. The j-th column of $[T]_\beta^\gamma = A$ is formed from the coefficients of $T(v_j)$ expressed in the γ basis.

9.5 Example

Let $D : \mathbb{P}_3 \to \mathbb{P}_3$. Let $\beta = \{1, x, x^2, x^3\}$ and $\gamma = \{1, x, x^2\}$ be the derivative transformation.

$D(1) = 0 = 0(1) + 0x + 0x^2$
$D(x) = 1 = 1(1) + 0x + 0x^2$
$D(x^2) = 2x = 0(01) + 2x + 0x^2$
$D(x^3) = 3x^2 = 0(1) + 0(1) + 0x + 3x^2$

$$[D]_\beta^\gamma = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

9.6 Discussion

Let V and W be vector spaces over a field $\mathbb{F}$. We add and scale linear transformations in the standard way. With these two operations, the set of linear transformation from V to W also form a vector space over $\mathbb{F}$, denoted by $\mathcal{L}(V, W)$ or just $\mathcal{L}(V)$ when $W = V$. Moreover, if $T : V \to W$ and $U : V \to W$ are linear transformations, then for bases β and γ,

(i) $[T + U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$

(ii) $[aT]_\beta^\gamma = a[T]_\beta^\gamma$

9.7 Theorem

Let $T : U \to V$ and $S : V \to W$ be linear transformation on finite dimensional vector spaces, U, V, W, over the same field $\mathbb{F}$. Let α, β, γ be bases of each vector space respectively.

Set $B = [T]_\alpha^\beta$ and $A = [S]_\beta^\gamma$. Then $[ST]_\alpha^\gamma = AB$

9.8 HW Questions Help & Midterm info

Q2 When you are trying to show two sets are equal, say $R = S$, sometimes it works to show $R \subseteq S$ and $S \subseteq R$.
Q3
Q4 Example of vector spaces: $\mathbb{R}^n, \mathbb{R}, \mathbb{C}, \{0\}, Matrices, \mathbb{P}_n$

Cheat sheet 8.5 $\times$ 11 one sided formula sheet will be allowed. There will be choice of problem on the exam, choose 4 from 6. Last week of Oct.

10 Invertibility and Isomorphism

Recall
$I_v : V \to V$ is a linear transformation

Find $[I_V]_\alpha^\beta$ when $\alpha = \{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \}$ and $\beta = \{ \begin{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \} \end{bmatrix} \}$

$\Rightarrow$ Solution:

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 \\ 2 & 0 & 1 & 1 & -1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{RowReduce} \begin{bmatrix} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{bmatrix}$$

$$[I_V]_\alpha^\beta = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

Note: I_V beta2, cool. Iv beta alpha, ok.

10.1 Definition: Inverse

Let V and W be vector spaces over a field $\mathbb{F}$. Let $T : V \to W$ be a linear transformation. A transformation $S : W \to V$ is called the **inverse** of T if $ST = I_V$ and $TS = I_W$. If T has an inverse, it is said to be **invertible** and the inverse is denoted by T^{-1}.

Invertible and nonsingular are synonyms.
Not invertible and singular are synonyms.
Something is invertible or it's not.

10.2 Theorem

Let V and W be vector spaces over a field $\mathbb{F}$. Let $T : V \to W$ be an invertible linear transformation.

1. $T^{-1} : W \to V$ is a linear transformation.
2. $(T^{-1})^{-1} = T$
3. If V is finite dimensional, then $dim(W) = dim(V)$
4. If $S : V \to W$ is an invertible linear transformation, then $(ST)^{-1} = T^{-1}S^{-1}$
5. T is one-to-one and onto

Proof:

10.3 Midterm

Tuesday, Oct 24 and Q&A on Monday, Oct 23 during class
Sections: 1.1-1.6, 2.1-2.5, and three ways to multiply matrices
No test on 1.7 (maximal sets and Zorn's lemma, October's 3rd week). One side of one 8.5x11 sheet of notes
Day off: November 17

10.4 Definition: Isomorphism

Let V and W be vector spaces over a field $\mathbb{F}$. We say V is **isomorphic** to W if there exists an invertible linear transformation $T : V \to W$. We refer to T as an **isomorphism** (of vector spaces).

10.5 Theorem

Let V and W be vector spaces over a field $\mathbb{F}$ with ordered bases $\beta = \{v_1.v_2, ..., v_n\}$. and $\gamma = \{w_1, w_2, ..., w_n\}$ respectively. Let $T : V \to W$ be a linear transformation. The following hold:

1. T is invertible if and only if $[T]_\beta^\gamma$ is invertible.
2. $[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$

Proof:
Let T be invertible. Let A be the matrix representation of T and let B be the matrix representation of T^{-1}. From Matrix multiplication in the previous chapter, we know that

$$[T^{-1}]_\gamma^\beta [T]_\beta^\gamma = [T^{-1}T]_\beta^\beta = [I_V]_\beta^\beta = I$$

So $[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$.

Conversely, suppose $[T]_\beta^\gamma$ is invertible. Define $T^{-1}(y) = ([T]_\beta^\gamma)^{-1}y$ for any $y \in W$. Then $([T]_\beta^\gamma)^{-1}y[T]_\beta^\gamma = I$. Let $[v]_\beta$ be any vector in V written in the β basis. Then $([T]_\beta^\gamma)^{-1}[T(v)]_\gamma = [T^{-1}]_\gamma^\beta [T]_\beta^\gamma [v]_\beta = I[v]_\beta = [v]_\beta$. So T must be invertible.

10.6 Example

Let $T : \mathbb{P}_3 \to M_{2\times 2}$ be defined by $T(f) = \begin{pmatrix} f(0) & f(1) \\ f(2) & f(3) \end{pmatrix}$. Is T invertible? If so, find the inverse.

Solution:
Choose ordered bases $\beta = \{1, x, x^2, x^3\}$ and $\gamma = \{E^{ij} | 1 \le (i, j) \le 2;$ lexicographical ordering, i and j are integers$\}$.

$$T(1) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}; T(x) = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} = 0E^{11} + 1E^{12} + 2E^{21} + 3E^{22}; T(x^2) = \begin{pmatrix} 0 & 1 \\ 4 & 9 \end{pmatrix}; T(x^3) = \begin{pmatrix} 0 & 1 \\ 8 & 27 \end{pmatrix}$$

$$[T]_{\beta}^{\gamma} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{pmatrix} \xrightarrow{RowReduce} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

We can now conclude that T is invertible and $\mathbb{P}_3$ over $\mathbb{R}$ is isomorphic to $M_{2\times 2}$ over $\mathbb{R}$. T is an isomorphism.

11 Change of Coordinate Matrix

11.1 Question

A linear transformation $T : V \to W$ has different matrix representations depending on the bases we choose. What is the relationship between these matrices?

11.2 Theorem

Let V be a finite dimensional vector space over a field $\mathbb{F}$ with ordered bases $\beta = \{v_1, v_2, ..., v_n\}$ and $\gamma = \{w_1, w_2, ..., w_n\}$. Let $Q = [I_V]_\beta^\gamma$. Then

1. Q is invertible
2. for any $v \in V, [v]_\gamma = Q[v]_\beta$

NOTE: $I_V : V \to V$, so the bases have the same number of elements.
Proof:
From the homework, Q is invertible.
Observe that $Q[v]_\beta$

11.3 Example

give some examples

We have an ugly $I_V y$ matrix representations and the pretty I matrix representation for the identity transformation depending on the bases we use. Can we do this for other linear transformations?
Note that $I_v(v) = v$ for every $v \in V$. If we take the same basis for both the domain and the codomain,

$$I_V(v_i) = v_i = 0v_1 + 0v_2 + ... + 1v_i + ... + 0v_n$$

For a more general linear transformation, do we have particular vectors, say u,v,w... so that $T(u) = u, T(v) = v$, but maybe $T(w) \neq w$?

11.4 Definition

Let $T : V \to V$ be a linear operator (a linear transformation from a vector space back into itself) on a vector space V. We say $u \neq 0$ is an eigenvector of T with eigenvalue λ provided

$$T(u) = \lambda u$$

So if we can find a basis of eigenvectors, call them $u_1, u_2, ..., u_n$ for the vector space V (of dimension n), with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$, then if $\beta = \{u_1, u_2, ..., u_n\}$, then

$$[T]_\beta^\beta = \begin{pmatrix} \lambda_1 & 0 & ... & ... & 0 \\ 0 & \lambda_2 & ... & & \vdots \\ \vdots & 0 & ... & ... & \vdots \\ 0 & \vdots & ... & \lambda_{n-1} & 0 \\ 0 & 0 & ... & 0 & \lambda_n \end{pmatrix}, \quad \text{almost as pretty as } I$$

Since $T(u_1) = \lambda_1 u_1 = \lambda_1 u_1 + 0u_2 + ... + 0u_n$
$T(u_2) = \lambda_2 u_2 = 0u_1 + \lambda_2 u_2 + ... + 0u_n$

Can an eigenvalue have more than one eigenvector? Can an eigenvector have more than one eigenvalue? Can we always get a basis for V of eigenvectors of T? Can an eigenvalue have more than one eigenvector?

Suppose $T(x) = \lambda x$, what is $T(3x)$? $T(3x) = 3T(x) = 3\lambda x = \lambda(3x)$. There are infinitely many different eigenvectors for a given eigenvalue since, in particular, the nonzero scalar multiples of eigenvectors are eigenvectors.

Suppose $T(x) = \lambda x$ and $T(y) = \lambda y$, then $T(x+y) = T(x) + T(y) = \lambda x + \lambda y = \lambda(x+y)$. So adding to eigenvectors with the same eigenvalue, gives us another eigenvector (as long as the result is not zero).

Are they a subspace? When we scale and add eigenvectors, we get eigenvectors.

11.5 Definition and Observation

Let λ be an eigenvalue of a linear operator/transformation T Then $E_\lambda = \{x \in V | T(x) = \lambda x\}$ is a subspace. This subspace consists of all the eigenvectors associated with λ, as well as the zero vector. We refer to it as the eigenspace of T with eigenvalue λ.

Can an eigenvector have more than one eigenvalue?
Suppose x has both λ and μ as eigenvalues for a given operator T. $\lambda x = T(x) = \mu x$, but then $(\lambda - \mu)x = 0$. So either $\lambda - \mu = 0$ or $x = 0$. Since $T(0) = \mu 0$ for every scalar μ, it doesn't give us much information. Every scalar would be an eigenvalue for the 0 vector. Moreover, nobody wants zero to show up in a basis because a set with zero vector is always linearly dependent and hence can't be a basis.

Does every operator T on a vector space V have a basis of eigenvectors?
$T(x) = A(x)$, where $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Then $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear operator, but 1 is the only eigenvalue. To talk about this example we need to know how to find eigenvectors. If $Ax = \lambda x$, then $(A - \lambda I)x = 0$.

12 Eigenvalues and Eigenvectors

Notation: "Let V be a space over a field $\mathbb{F}$ and $T : V \to V$ a linear transformation." will be summarized with the statement "Let T be a linear operator on a vector space V"

12.1 Definition: Diagonalizable

A linear operator T on a finite dimensional vector space V is said to be **diagonalizable** if there is an ordered basis β for V such that $[T]_\beta$ is diagonal.
NOTE: $[T]_\beta = [T]_\beta^\beta$

12.2 Definition: Eigenvector & Eigenvalue

Let T be a linear operator on a vector space V. A nonzero vector $v \in V$ is said to be an **eigenvector** of T if there is a scalar λ such that $T(v) = \lambda v$. The scalar λ is referred to as the **eigenvalue** corresponding to the eigenvector v.

12.2.1 Example

Given a random vectors
$T(a + bx + cx^2 + dx^3) = (5a + b - c - d) + (a + 5b - c - d)x + (-a - b + 6c)x^2 + (-a - b + 6d)x^3$
Notice $T(1 + x + x^2 + x^3) = 4 + 4x + 4x^2 + 4^3 = 4(1 + x + x^2 + x^3)$
So for this transformation, $1 + x + x^2 + x^3$ is an eigenvector with eigenvalue 4.

12.3 Theorem

Let T be a linear operator on a finite dimensional vector space V. Then T is diagonalizable if and only if there's a basis for V consisting of eigenvectors of T. Moreover, when $\beta = \{v_1, v_2, ..., v_n\}$ is a basis for V consisting of eigenvectors of T, then $[T]_\beta = D$ where d_{jj} is the eigenvalue corresponding to the eigenvector v_j for $1 \leq j \leq n$.

Proof:
$\Rightarrow$
Assume T is diagonalizable. Then there's an ordered basis β for V such that $[T]_\beta = D$, where D is a diagonal matrix. Let $\beta = \{v_1, v_2, ..., v_n\}$, for any $1 \leq j \leq n, T(v_j) = \sum_{i=1}^n d_{ij} v_i = d_{jj} v_j$. Thus each vector is β is an eigenvector of T. Moreover, the eigenvalue associated v_j is d_{jj}. Hence β is a basis for V consisting of eigenvectors.
$\Leftarrow$
Let $\beta = \{v_1, v_2, ..., v_n\}$ be a basis for V consisting of eigenvectors of T. Since each vector v_j is an eigenvector, we know there exist scalars $\lambda_1, \lambda_2, ..., \lambda_n$ so that $T(v_j) = \lambda_j v_j$.

$$T(v_j) = 0v_1 + ... + 0v_{j-1} + \lambda_j v_j + 0v_{j+1} + ... + 0v_n$$

Notice that this implies that all the off-diagonal entries in the j-th column will be zero. Hence $[T]_\beta$ is a diagonal matrix.

12.4 Recall Theorem

Let $A \in M_{n\times n}(\mathbb{F})$. Then $Ax = 0$ has nontrivial solutions if and only if $\det(A) = 0$.

12.5 Theorem

Let $A \in M_{n\times n}(\mathbb{F})$. Then λ is an eigenvalue of A if and only if $\det(A - \lambda I) = 0$.
Proof:
$\det(A - \lambda I) = 0$ if and only if there is a nonzero vector x such that $(A - \lambda I)x = 0$
$\Leftrightarrow$
there is a nonzero vector x such that $Ax = \lambda I x$
$\Leftrightarrow$
there is a nonzero vector x so that $Ax = \lambda x$
$\Leftrightarrow$
λ is an eigenvalue of A.

12.6 Definition: Characteristic Polynomial

Let $A \in M_{n\times n}(\mathbb{F})$. The polynomial $f(\lambda) = det(A - \lambda I)$ is called the **characteristic polynomial** of A.

12.6.1 Remark

Let $A \in M_{n\times n}(\mathbb{F})$. The roots of the characteristic polynomial of A are precisely the eigenvalues of A.

So how do we move from matrices to linear operator? How would you find the characteristic polynomial of a linear operator?
We could find a matrix representation of the linear operator with respect to a basis.

Recall: $[T]_\beta = Q^{-1}[T]_\gamma Q$ so

$$\begin{aligned} det(\lambda I - [T]_\beta) &= det(\lambda Q^{-1} I Q - Q^{-1}[T]_\gamma Q \\ &= det(Q^{-1}(\lambda I - [T]_\gamma)Q) \\ &= det(Q^{-1} det(\lambda I - [T]_\gamma) det(Q) \\ &= det(\lambda I - [T]_\gamma) \end{aligned}$$

The coefficients of the eigenvectors will look different because they are coefficients with respect to different bases. When you translate the back into the original vector space, they will be the same.

12.7 Theorem

Let T be a linear operator of a finite dimensional vector space V and let λ be an eigenvalue of T. hen v is an eigenvector of T corresponding to λ if and only if v is a nonzero vector in $N(\lambda I_V - T)$.

12.8 Definition: Eigenspace

Let T be a linear operator of a finite dimensional vector space V and let λ be an eigenvalue of T. Then $N(\lambda I_V - T)$ is referred to as the **eigenspace** of T corresponding to the eigenvalue λ.

12.9 Example

$T(a + bx + cx^2 + dx^3)$
$= (5a + b - c - d) + (a + 5b - c - d)x + (-a - b + 6c)x^2 + (-a - +6d)x^3$

Is T diagonalizable? If so, find a basis β so that $[T]_\beta$ is a diagonal matrix. How many different diagonal matrix representations can T have?
Solution:
Step 1: Turn T into its matrix representation. We have a standard basis $\alpha = \{1, x, x^2, x^3\}$
$T(1) = 5 + x - x^2 - x^3$, Set $a = 1, b = c = d = 0$
$T(x) = 1 + 5x - x^2 - x^3$, Set $b = 1, a = c = d = 0$
$T(x^2) = -1 - x + 6x^2 + 0x^3$, Set $c = 1, a = b = d = 0$
$T(x^3) = -1 - x + 0x^2 + 6x^3$, Set $d = 1, a = b = c = 0$

$$[T]_\alpha - \lambda I = \begin{pmatrix} 5-\lambda & 1 & -1 & -1 \\ 1 & 5-\lambda & -1 & -1 \\ -1 & -1 & 6-\lambda & 0 \\ -1 & -1 & 0 & 6-\lambda \end{pmatrix}$$

We don't need to row reduce the matrix because it will change the eigenvalue. Now, find the determinant of this matrix. Thus we have

$$\begin{aligned} det([T]_\alpha - \lambda I) &= det(\begin{pmatrix} 5-\lambda & 1 & -1 & -1 \\ 1 & 5-\lambda & -1 & -1 \\ -1 & -1 & 6-\lambda & 0 \\ -1 & -1 & 0 & 6-\lambda \end{pmatrix}) \\ &= 768 - 608\lambda + 176\lambda^2 - 22\lambda^3 + \lambda^4 \\ &= (-8 + \lambda)(-6 + \lambda)(-4 + \lambda)^2 \end{aligned}$$

From here, we can see that the eigenvalues are $\{8, 6, 4, 4\}$.

$$\begin{pmatrix} 5-8 & 1 & -1 & -1 \\ 1 & 5-8 & -1 & -1 \\ -1 & -1 & 6-8 & 0 \\ -1 & -1 & 0 & 6-8 \end{pmatrix} \xrightarrow{RowReduce} \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & rhs \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -x_4 \\ -x_4 \\ x_4 \\ x_4 \end{pmatrix} = x_4 \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$$

Any nonzero choice of x_4 creates an eigenvector. For example: $\begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$

Remember, this gives us the coefficients of the basis we chose to start with: $\{1, x, x^2, x^3\}$ so in the original vector space $-1 - x + x^2 + x^3$ should be an eigenvector with eigenvalue 8.

$T(a + bx + cx^2 + dx^3)$
$= (5a + b - c - d) + (a + 5b - c - d)x + (-a - b + 6c)x^2 + (-a - +6d)x^3$

Set $a = b = -1$, $c = d = 1$, then
$T(-1 - x + x^2 + x^3) = -8 - 8x + 8x^2 + 8x^3 = 8(-1 - x + x^2 + x^3)$

Now, let's do the eigenvalue $= 6$

$$\begin{pmatrix} 5-6 & 1 & -1 & -1 \\ 1 & 5-6 & -1 & -1 \\ -1 & -1 & 6-6 & 0 \\ -1 & -1 & 0 & 6-6 \end{pmatrix} \xrightarrow{RowReduce} \begin{array}{c} \begin{matrix} x_1 & x_2 & x_3 & x_4 & rhs \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

; Eigenvector $x_4 \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}$ which represents $-x^2 + x^3$

Now, let's do the eigenvalue $= 4$

$$\begin{pmatrix} 5-4 & 1 & -1 & -1 \\ 1 & 5-4 & -1 & -1 \\ -1 & -1 & 6-4 & 0 \\ -1 & -1 & 0 & 6-4 \end{pmatrix} \xrightarrow{RowReduce} \begin{array}{c} \begin{matrix} x_1 & x_2 & x_3 & x_4 & rhs \end{matrix} \\ \begin{pmatrix} 1 & 1 & 0 & -2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}; \begin{pmatrix} -x_2 + 2x_4 \\ x_2 \\ x_4 \\ x_4 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 2 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

which represents Eigenvector for $\lambda = 4 : \{-1 - x, 2 + x^2 + x^3\}$

Set $\beta = \{-1 - x + x^2 + x^3, -x^2 + x^3, -1 + x, 2 + x^2 + x^3\}$

$[T]_\beta = \begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$, if we reorder the basis, we may get a different arrangement of the eigenvalues on the diagonal.

13 Diagonalizability

13.1 Theorem

Let T be a linear operator on a vector space V. Let $\lambda_1, ..., \lambda_k$ be distinct eigenvalues of T with corresponding eigenvectors $v_1, ..., v_k$. Then $\{v_1, ..., vk\}$ is linearly independent.

Proof:
Suppose $\{v_1, ..., vk\}$ is linearly dependent. Let $S_j = \{v_1, ..., vj\}$. Certainly S_1 is linearly independent and S_k is assumed to be linearly dependent (looking for a contradiction). Choose the smallest m so that S_m is linearly dependent. Notice $1 < m \leq k$. Choose coefficients $a_1, ..., a_m$ not all zero so that $a_1v_1 + ... + a_mv_m = 0$. Since S_{m-1} is linearly independent, we know $a_m \neq 0$.

Notice $T(a_1v_1 + ... + a_mv_m) = T(0) = 0$, so $a_1T(v_1) + ... + a_mT(v_m) = 0$, and thus $a_1\lambda_1v_1 + ... + a_m\lambda_mv_m = 0$.

Subtract $\lambda_m(a_1v_1 + ... + a_mv_m) = 0$ from both sides of this equation, we get

$$a_1(\lambda_1 - \lambda_m)v_1 + ... + a_m(\lambda_m - \lambda_m)v_m = 0$$

$$a_1(\lambda_1 - \lambda_m)v_1 + ... + a_{m-1}(\lambda_{m-1} - \lambda_m)v_{m-1} = 0$$

Because the λ_j are distinct, $\lambda_j - \lambda_m \neq 0$ for all $a \leq j \leq m$. Since $v_m \neq 0$, at least one value of a_j with $1 \leq j \leq m - 1$ is nonzero, so such an $a_j(\lambda_j - \lambda_m)$ is nonzero, contradicting that S_{m-1} was a linearly independent set. Hence $\{v_1, .., v_k\}$ can't be linearly dependent.

13.2 Corollary

If T has $\dim(V)$ distinct eigenvalues, then T is diagonalizable.
REMARK: T having $dim(V) = n$ distinct eigenvalues is sufficient, but not necessary for T to be diagonalizable.

13.3 Examples

....

13.4 Definition: Splits

A polynomial f **splits** over $\mathbb{F}$ if there are scalars $c, a_1, a_2, ..., a_n$ such that $f(t) = c(t - a_1)(t - a_2)...(t - a_n)$. The a_j are the roots (zeros) of the polynomial.

13.5 Example

$f(t) = t^2 + 1 = (t - i)(t + i)$, so $f(t)$ splits over $\mathbb{C}$, but does not split over $\mathbb{R}$.