LECTURE 18

Abstract Vector Spaces and Concrete Examples

Our discussion of linear algebra so far has been devoted to discussing the relations between systems of linear equations, matrices, and vectors. As such, topics like the notions of subspaces, bases, linear independence, etc. tend to come off as a bunch of formal nonsense. To see the real utility of these ideas, it's actually helpful to first abstract things further; this will allow us to bring more interesting examples into the picture.

DEFINITION 18.1. *A vector space over $\mathbb{R}$ is a set V for which the following operations are defined*

- ***scalar multiplication**: for every $r \in \mathbb{R}$ and $\mathbf{v} \in V$ we have a map $(r, \mathbf{v}) \to r \cdot \mathbf{v} \in V$.*
- ***vector addition:** for every pair of vectors $\mathbf{u}, \mathbf{v} \in V$ we have a map $(\mathbf{u}, \mathbf{v}) \to \mathbf{u} + \mathbf{v} \in V$*

  and satisfy the following properties

  (1) $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$ *(associativity of vector addition)*
  (2) $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$ *(commutativity of vector addition)*
  (3) *There exists an element $\mathbf{0} \in V$ such that $\mathbf{v}+\mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$. (additive identity.)*
  (4) *For each $\mathbf{v} \in V$ there exists an element $-\mathbf{v} \in V$ such that $\mathbf{v}+(-\mathbf{v}) = \mathbf{0}$. (additive inverses)*
  (5) $r(\mathbf{u}+\mathbf{v}) = r\mathbf{u}+r\mathbf{v}$ *(scalar multiplication is distributive with respect to vector addition).*
  (6) $(r+s)\mathbf{v} = r\mathbf{v}+s\mathbf{v}$. *(scalar multiplication is distributive with respect to addition of scalars)*
  (7) $r(s\mathbf{v}) = (rs)\mathbf{v}$ *(scalar multiplication preserves associativity of multiplication in $\mathbb{R}$.)*
  (8) $(1)\mathbf{v} = \mathbf{v}$ *(preservation of scale).*

DEFINITION 18.2. *A **subspace** W of a vector space V is a subset of V that is closed under scalar multiplication and vector addition. Equivalently, a subspace of a vector space V is a subset W of V that is also a vector space (using the scalar multiplication and vector addition it inherits from V).*

EXAMPLE 18.3. $V = \{m \times n \text{ matrices}\}$

EXAMPLE 18.4. $V = \{\text{polynomials of degree } n\}$

This is an important example, not just for its applications, but also because it illustrates how this notion of an abstract vector can be used to connect a disparent part of mathematics with linear algebra. We'll work this example out in detail.

Let P be the set of polynomials in one variable x of degree $\leq n$. We define **scalar multiplication** in P by

$$\text{if } p = a_n x^n + \cdots a_1 x + a_0 \in P \quad \text{and} \quad \lambda \in \mathbb{R} \quad , \quad \text{then} \quad \lambda p = \lambda a_n x^n + \cdots + \lambda a_1 x + \lambda a_0 \in P$$

We define **vector addition** in P by

$$\begin{aligned} \text{if} \quad p &= a_n x^n + \cdots a_1 x + a_0 \quad \text{and} \quad p' = a_n' x^n + \cdots a_1' x + a_0' \quad , \\ \text{then} \quad p + p' &= (a_n + a_n') x^n + \cdots + (a_1 + a_1') x + a_0 + a_0' \in P \end{aligned}$$

It is easily verified that P endowed with this scalar multiplication and this vector addition satisfies all eight of the axioms above. So P is a vector space.

The set P_n consisting of polynomials of degree $\leq n$ is a subspace of P.

EXAMPLE 18.5. $V = \{\text{functions on the real line}\}$

EXAMPLE 18.6. $V = \{\text{functions from a set } S \text{ to the set of real numbers}\}$

EXAMPLE 18.7. $V = \{\text{formal power series in one variable}\}$

EXAMPLE 18.8. Let P be the vector space of polynomials in one variable x. Determine whether or not the given set of vectors is linearly dependent or independent.

$$\left\{1, 4x+3, 3x-4, x^2+2, x-x^2\right\} \tag{1}$$

Recall that a set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ of vectors in $\mathbb{R}^n$ is linearly independent if and only if the only solution of

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0} \tag{2}$$

is $c_1 = c_2 = \cdots = c_k = 0$. The notion of linear independence in P is exactly the same (as soon as we identify polynomials with abstract vectors). Thus the polynomials

$$\begin{aligned} p_1 &= 1 \\ p_2 &= 4x+3 \\ p_3 &= 3x-4 \\ p_4 &= x^2+2 \\ p_5 &= x-x^2 \end{aligned}$$

will be linearly independent if and only if

$$c_1p_1 + c_2p_2 + c_3p_3 + c_4p_4 + c_5p_5 = 0 \quad \Rightarrow \quad 0 = c_1 = c_2 = c_3 = c_4 = c_5$$

Now the left hand side of the above equation is

$$c_1+c_2(4x+3)+c_3(3x-4)+c_4(x^2+2)+c_5(x-x^2) = (c_4+c_5)\,x^2+(4c_2+3c_3+c_5)\,x+(c_1+3c_2-4c_3+2c_4)$$

so if this is to vanish we must have

$$\begin{aligned} c_1 + 3c_2 - 4c_3 + 2c_4 &= 0 \\ 4c_2 + 3c_3 + c_5 &= 0 \\ c_4 + c_5 &= 0 \end{aligned}$$

Looking at the corresponding augmented matrix

$$\left[\begin{array}{ccccc|c} 1 & 3 & -4 & 2 & 0 & 0 \\ 0 & 4 & 3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right]$$

and noting that it is already in row-echelon form with no rows corresponding to inconsistent equations, we can conclude that the corresponding linear system is consistent. But a consistent linear system with more variables than equations will always have an infinite number of solutions. Hence, there will be non-trivial solutions of (2), and so the polynomials (1) are not linearly independent.

1. Vector Space Isomorphisms

In short, a vector space over $\mathbb{R}$ is a set V upon which notions of scalar multiplication and vector addition are defined, and such that these operations behave exactly like scalar multiplication and vector addition in $\mathbb{R}^n$. The point of this definition is that it brings to our theoretical and calculational apparatus a much wider variety of underlying sets. To facilitate this, however, we need two more concepts. The first is the generalization of linear transformations $T : R^n \to \mathbb{R}^m$ to abstract vector spaces

DEFINITION 18.9. *Let V and W be two vector spaces. A mapping $T : V \to W$ is called a (**vector space**) **homomorphism** if it preserves the operations of scalar multiplication and vector addition:*

$$\begin{aligned} T(\lambda \cdot_V \mathbf{v}) &= \lambda \cdot_W T(\mathbf{v}) \\ T(\mathbf{v}_1 +_V \mathbf{v}_2) &= T(\mathbf{v}_1) +_W T(\mathbf{v}_2) \end{aligned}$$

Here, $\cdot_V$ denotes the scalar multiplication in V, $+_V$ denotes vector addition in V, etc. If $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ then a vector space homomorphism from V to W is exactly the same thing as a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Recall that a mapping $T : V \to W$ is called **one-to-one** if an inverse mapping $T^{-1} : W \to V$ exists (such that both $T^{-1} \circ T : V \to V$ and $T \circ T^{-1} : W \to W$ are the identify mappings).

DEFINITION 18.10. *A vector space homomorphism $T : V \to W$ is called **one-to-one** if*

$$f(x) = f(y) \quad \Rightarrow \quad x = y$$

LEMMA 18.11. *A vector space homomorphism $T : V \to W$ is one-to-one if and only if* $\ker(T) = \{\mathbf{0}\}$.

Proof.

$\Rightarrow$ Suppose $T : V \to W$ is one-to-one and $\ker(T) \neq \{\mathbf{0}\}$. Let $\mathbf{x}$ be a non-zero element of $\ker(T)$. Then for any $\mathbf{y} \in V$ we have

$$\mathbf{y} + \mathbf{x} \neq \mathbf{y}$$

but

$$T(\mathbf{y} + \mathbf{x}) = T(\mathbf{y}) + T(\mathbf{x}) = T(\mathbf{y}) + \mathbf{0} = T(\mathbf{y})$$

which contradicts our hypothesis that T is one-to-one. We conclude that if T is one-to-one then $\ker(T) = \{\mathbf{0}\}$.

$\Leftarrow$ Suppose $\ker(T) = \{\mathbf{0}\}$ and $\mathbf{x}, \mathbf{y}$ are such that $T(\mathbf{x}) = T(\mathbf{y})$. Then

$$\mathbf{0} = T(\mathbf{x}) - T(\mathbf{y}) = T(\mathbf{x} - \mathbf{y}) \quad \Rightarrow \quad \mathbf{x} - \mathbf{y} = \mathbf{0}$$

since $\ker(T) = \{\mathbf{0}\}$. But then

$$T(\mathbf{x}) = T(\mathbf{y}) \quad \Rightarrow \quad \mathbf{x} = \mathbf{y}$$

and so T is onto.

DEFINITION 18.12. *A vector space homomorphism $T : V \to W$ is said to be onto if*

$$W = Range(f) \equiv \{w \in W \mid w = T(x) \text{ for some } x \in V\}$$

DEFINITION 18.13. *A (**vector space**) **isomorphism** is a vector space homomorphism that is one-to-one and onto. If an isomorphism $T : V \to W$ exists between two vector spaces, then we say that V and W are **isomorphic**.*

Now let me explain how this idea of a vector space isomorphism is used in practice.

LEMMA 18.14. *If $T : V \to W$ is a vector space isomorphism then* $\dim(V) = \dim(W)$.

THEOREM 18.15. *Every n-dimensional (real) vector space is isomorphic to $\mathbb{R}^n$.*

This theorem is the key to carrying out calculations in an abstract vector space. We have already a wide range of calculational techniques for handling questions about vectors in $\mathbb{R}^n$. The definition of an abstract vector space tells already that the operations of scalar multiplication and vector addition in a vector space behave like the corresponding operations in $\mathbb{R}^n$. The theorem tells us that in fact we can consistently convert a calculation in an abstract vector space V to a corresponding calculation in $\mathbb{R}^n$, and then convert the answer back to V.

EXAMPLE 18.16. Let P_2 be the vector space of polynomials in one variable of degree ≤ 2. Determine whether or not the following vectors are a basis for P_2.

$$\left\{1, x - 1, (x - 1)^2\right\} \tag{3}$$

If (3) is a basis for P_2, then every vector (i.e. polynomial) in P_2 has a unique expression as a linear combination of $p_1 = 1$, $p_2 = x - 1$, and $p_3 = (x-1)^2$. Let

$$p = a_0 + a_1 x + a_2 x^2$$

be an arbitrary element of P_2. We want to find a unique set of coefficients c_1, c_2, c_3 such that

$$p = c_1 p_1 + c_2 p_2 + c_3 p_3$$

or

$$a_0 + a_1 x + a_2 x^2 = c_1 + c_2(x-1) + c_3(x-1)^2 = c_1 - c_2 + c_3 + (c_2 - 2c_3)\,x + c_3 x^2$$

Comparing the coefficients of like powers of x on both sides we see we must have

$$\begin{aligned} c_1 - c_2 + c_3 &= a_0 \\ c_2 - 2c_3 &= a_1 \\ c_3 &= a_2 \end{aligned}$$

The corresponding augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & a_0 \\ 0 & 1 & -2 & a_1 \\ 0 & 0 & 1 & a_2 \end{array}\right]$$

This matrix is already in row-echelon form, and it is clear that the corresponding linear system is consistent and has a unique solution. Thus, the polynomials (3) form a basis for P_2.

Note: The monomials $\left\{1, x, x^2\right\}$ also form a basis for P_2 and the fact that every polynomial in P_2 can be uniquely expressed as either

$$p = a_0 + a_1 x + a_2 x^2$$

or

$$p = c_1 + c_2(x-1) + c_3(x-1)^2$$

can be understood as an example of Taylor's theorem (the first expression corresponds to a Taylor expansion of p about $x = 0$ and the second expression corresponds to the Taylor expansion of p about $x = 1$.)

Example 18.17. Find a basis for the span of the following polynomials.

$$\left\{x^2 - 1, x^2 + 1, 4, 2x - 3\right\}$$

Recall that a basis for the span of a set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ of vectors in $\mathbb{R}^n$ can be found by identifying the vectors with the columns of a $n \times k$ matrix $\mathbf{A}$, reducing $\mathbf{A}$ to a row-echelon form $\mathbf{A}'$, and then selecting the vectors $\mathbf{v}_i$ corresponding to columns of $\mathbf{A}'$ with pivots. We can do same thing here if we first translate the polynomials above to vectors in $\mathbb{R}^3$. Thus, we set

$$\begin{array}{rcl} x^2 - 1 & \to & \mathbf{v}_1 = (-1, 0, 1) \\ x^2 + 1 & \to & \mathbf{v}_2 = (1, 0, 1) \\ 4 & \to & \mathbf{v}_3 = (4, 0, 0) \\ 2x - 3 & \to & \mathbf{v}_4 = (-3, 2, 0) \end{array}$$

(Here we are interpreting the coefficient of x^{i-1} as the i^{th} component of the corresponding vector.) This leads us to

$$\mathbf{A} = [\mathbf{v}_1|\mathbf{v}_2|\mathbf{v}_3|\mathbf{v}_4] = \left[\begin{array}{cccc} -1 & 1 & 4 & -3 \\ 0 & 0 & 0 & 2 \\ 1 & 1 & 0 & 0 \end{array}\right]$$

This matrix row reduces to

$$\begin{bmatrix} -1 & 1 & 4 & -3 \\ 0 & 0 & 0 & 2 \\ 1 & 1 & 0 & 0 \end{bmatrix} \quad \xrightarrow{R_3 \to R_3 + R_1} \quad \begin{bmatrix} -1 & 1 & 4 & -3 \\ 0 & 0 & 0 & 2 \\ 0 & 2 & 4 & -3 \end{bmatrix} \quad \xrightarrow{R_2 \longleftrightarrow R_3} \quad \begin{bmatrix} -1 & 1 & 4 & -3 \\ 0 & 2 & 4 & -3 \\ 0 & 0 & 0 & 2 \end{bmatrix} = \mathbf{A}'$$

The pivots of $\mathbf{A}'$ occur in the first, second, and fourth columns. Hence, the vectors $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ form a basis for the span$(\mathbf{v}_1, \mathbf{v}_2\mathbf{v}_3, \mathbf{v}_4)$. And so the polynomials $\left\{x^2-1, x^2+1, 2x-3\right\}$ form a basis for the span of $\left\{x^2-1, x^2+1, 4, 2x-3\right\}$.

Example 18.18. Find the coordinates of the vector x^2 with respect to the basis $\left\{1, x-1, (x-1)^2\right\}$

We have already verified that $\left\{1, x-1, (x-1)^2\right\}$ is a basis for P_2. We now seek to explicitly identify the coefficients c_1, c_2, c_3 such that

$$x^2 = c_1 + c_2(x-1) + c_3(x-1)^2$$

Expanding the right hand side in powers of x we have

$$x^2 = c_1 - c_2 + c_3 + (c_2 - 2c_3)\,x + c_3 x^2$$

Comparing the coefficients of like powers of x we see that we must have

$$\begin{aligned} c_1 - c_2 + c_3 &= 0 \\ c_2 - 2c_3 &= 0 \\ c_3 &= 1 \end{aligned}$$

This system is easily solved by back-substitution:

$$c_3 = 1 \quad \Rightarrow \quad c_2 = 2 \quad \Rightarrow \quad c_1 = 1$$

Thus,

$$x^2 = 1 + 2(x-1) + (x-1)^2$$

Note: the right hand side is interpretable as the Taylor expansion of $f(x) = x^2$ about $x = 1$.

Example 18.19. Show that the mapping $D : P_2 \to P_2 : p(x) \mapsto x\frac{dp}{dx} - 2p$ is a linear transformation of P_2. Find the kernel and range of $\frac{d}{dx}$.

First let's demonstrate that the mapping D is linear. If $\lambda \in \mathbb{R}$, and $p \in P_2$ then

$$D(\lambda p) = x\frac{d}{dx}(\lambda p) - 2(\lambda p) = \lambda x\frac{dp}{dx} - 2\lambda p = \lambda\left(x\frac{dp}{dx} - 2p\right) = \lambda D(p)$$

so D preserves scalar multiplication. If $p_1, p_2 \in P_2$, then

$$D(p_1 + p_2) = x\frac{d}{dx}(p_1 + p_2) - 2(p_1 + p_2) = x\frac{dp_1}{dx} - 2p_1 + x\frac{dp_2}{dx} - 2p_2 = D(p_1) + D(p_2)$$

so D preserves vector addition. Since D is a mapping between two vector spaces (albeit, the same one) that preserves both scalar multiplication and vector addition, D is a linear transformation.

Now recall the a linear mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is interpretable as a $m \times n$ matrix $\mathbf{A}$ multiplying an n-dimensional vector from the left, and the kernel of T corresponds to the null space of $\mathbf{A}$ and the range of T corresponds to the column space of $\mathbf{A}$. To make use of these facts, we must first reinterpret the polynomials in P_2 as vectors in $\mathbb{R}^3$. We thus map

$$p(x) = a_2 x^2 + a_1 x + a_0 \quad \Rightarrow \quad \mathbf{v} = (a_0, a_1, a_2)$$

Now to construct the matrix $\mathbf{A}$ we examine the action of T on the standard basis vectors $\mathbf{e}_1 = (1,0,0)$, $\mathbf{e}_2 = (0,1,0)$, and $\mathbf{e}_3 = (0,0,1)$. This would correspond to looking at the action of D on the "standard basis vectors" $p_1 = 1$, $p_2 = x$, and $p_3 = x^2$ of P_2. We have

$$\begin{aligned}
D(p_1) &= x\frac{d}{dx}(1) - 2(1) = -2 = -2p_1 \quad = -2p_1 + (0)p_2 + (0)p_3 \\
D(p_2) &= x\frac{d}{dx}(x) - 2(x) = x - 2x = -p_2 = (0)p_1 - p_2 + (0)p_3 \\
D(p_3) &= x\frac{d}{dx}(x^2) - 2(x^2) = 2x^2 - 2x^2 = 0 = (0)p_1 + (0)p_2 + (0)p_3
\end{aligned}$$

Thus we take the matrix $\mathbf{A}$ to be

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

The solution to the corresponding homogeneous linear system

$$\mathbf{Ax} = \mathbf{0} \quad \Rightarrow \quad \begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

corresponding to this matrix is evidently

$$\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix} \in span \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \Rightarrow \quad p = span(p_3) = span\left(x^2\right)$$

So the kernel of D is the $\left\{p \in P_2 \mid p = cx^2\right\}$.

The range of D will correspond to the column space of $\mathbf{A}$, which in turn would be span of the column vectors of $\mathbf{A}$ corresponding to the columns of a row-echelon form of $\mathbf{A}$ that contain pivots. Luckily, $\mathbf{A}$ is already in row-echelon form, and the columns containing pivots are columns 1 and 2. Thus,

$$\begin{aligned}
Range(D) &\approx ColSp(\mathbf{A}) \\
&= span\left(\begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix}, \begin{matrix} 0 \\ -1 \\ 0 \end{matrix}\right) \\
&\approx span(-2, -x) \\
&= span(1, x)
\end{aligned}$$