

Math 247: Applied Linear Algebra

Week 4 Course Notes

Main topics: direct sums, linear maps, kernel, image, injective, the fundamental theorem/dimension formula.

1 Direct Sums

Definition 1.1. Say we have a vector space V over k and $U \leq V, W \leq V$. We then define the following:

$$U + W := \{U + W \mid u \in U, w \in W\}$$

We have already proven that a regular sum is a subspace. However, a sum of subspaces is called **direct**, in symbols, $U \oplus W$, if the following holds: $\forall v \in U + W$, the representation $v = u + w$ where $u \in U, w \in W$ is **uniquely determined**.

This concept can be extended to multiple sums: a sum $U_1 + \dots + U_n$ is called direct if $\forall v \in U_1 + \dots + U_n$, the representation

$$v = u_1 + \dots + u_n, u_j \in U_j \quad \forall i \leq j \leq n$$

is unique.

Theorem 1. A sum $U + W$ of subspaces is direct iff $U \cap W = \{0\}$.

Proof. Will prove both directions.

[\Rightarrow] Assume: the sum $U + W$ is direct. So, we will pick an arbitrary vector intersection and prove that it must be zero.

Let $u \in U \cap W$. Then, to prove uniqueness, we will express the zero vector in more than one way:

$$0 = 0 + 0 = v + (-v)$$

We have that the zero on the LHS is in the sum $U + W$. This can be broken down into $0 + 0$, which the first zero contained in U and the second zero contained in W . Because they are subspaces, we know that they are closed under multiplication; hence, we can also say that this is equal to $v + (-v)$. v is in the intersection, and especially in U . $-v$ is especially in W since it's closed under scalar multiplication.

So, we now have two representations of zero. However, we assumed that each representation is unique. This forces v to be zero and we are done.

[\Leftarrow] Let $U \cap W = \{0\}$. Consider an arbitrary vector in the sum. We need to show that this representation is unique.

Let $v \in U + W$ be arbitrary. Then, by the definition of the sum, we obtain:

$$v = u_1 + w_1 \text{ and } v = u_2 + w_2$$

We need to show that $u_1 = u_2$ and $w_1 = w_2$. The trick is to subtract the equations to obtain zero:

$$\begin{aligned} 0 &= (u_1 - u_2) + (w_1 - w_2) \\ &\Rightarrow u_1 - u_2 = w_2 - w_1 \end{aligned}$$

The LHS is contained inside U and the RHS is contained inside W . Due to equality, both sides are contained in the intersection. But, we can assume that the intersection is trivial. Hence:

$$u_1 - u_2 = 0 = w_2 - w_1 \Rightarrow u_1 = u_2 \text{ and } w_1 = w_2$$

□

2 Linear Maps

Motivation: in algebra, once we define a set of objects, the next task is to describe mappings between these objects. Here, we consider **linear maps**. Generally speaking, maps between objects preserve their algebraic structures.

Definition 2.1. Let V and W be VS over k (must be over the SAME k). A function $L : V \rightarrow W$ is called **linear** if it behaves well with respect to the following operations:

1. **Respecting addition:** $L(v_1 + v_2) = L(v_1) + L(v_2)$
2. **Respecting scalar multiplication:** $L(kv_1) = k(L(v_1))$

Theorem 2. *This theorem is useful when proving things are not linear maps. Let $L : V \rightarrow W$ be linear. Then, the zero vector is always preserved:*

$$L(0) = 0$$

This theorem is so big that we demonstrate two proofs (each follow from one of the properties from the definition).

Proof. Exploit **property 1**:

$$\begin{aligned} L(0) &= L(0 + 0) = L(0) + L(0) \\ &= 0 + L(0) = L(0) + L(0) \\ &\Rightarrow 0 = L(0) \text{ (cancellation law)} \end{aligned}$$

□

Proof. Exploit **property 2**:

$$L(0) = L(0 \cdot 0) = 0 \cdot L(0) = 0$$

□

CAUTION

in calculus, the so-called “linear maps” $f : x \mapsto ax + b$ is only linear if $b = 0$. This is similar to the fact that lines in general are not subspaces.

Definition 2.2. Each linear map is associated to two very important subspaces. If we let $L : V \rightarrow W$ be linear. Then, we can define:

1. the **kernel**, symbolized as $\text{Ker}(L)$, is defined as

$$\text{Ker}(L) := \{v \in V \mid L(v) = 0\} \subseteq V \quad (1)$$

in other words: it's the set of all vectors that are sent to zero by the linear transformation.

2. the **image**, symbolized as $\text{Im}(L)$, is defined as:

$$\text{Im}(L) := \{L(v) \mid v \in V\} \subseteq W \quad (2)$$

Theorem 3. *Let $L : V \rightarrow W$ be linear. Then:*

1. *The $\text{Ker}(L) \subseteq V$ is a subspace of V .*
2. *The $\text{Im}(L) \subseteq W$ is a subspace of W .*

Proof. This is a pretty standard proof by now. We just check the standard subspace criterion.

Kernel:

1. We know that zero is contained in the kernel because we have just proven that $L(0) = 0$.
2. **Closure under addition:** Let $v_1, v_2 \in \text{Ker}(L)$. Then:

$$\begin{aligned} L(v_1 + v_2) &= L(v_1) + L(v_2) \\ &= 0 + 0 = 0 \\ &\Rightarrow v_1 + v_2 \in \text{Ker}(L) \end{aligned}$$

3. **Closure under scalar multiplication:**

$$\begin{aligned} L(kv_1) &= kL(v_1) \text{ (Linearity)} \\ &= k \cdot 0 \\ &= 0 \\ &\Rightarrow kv_1 \in \text{Ker}(L) \end{aligned}$$

Image:

1. We know 0 is in the image under L , since $L(0) = 0$.
2. **Closure under addition:** Let $w_1, w_2 \in \text{Im}(L)$. Then:

$$\exists v_1, v_2 \in V \text{ with } L(v_1) = w_1 \text{ and } L(v_2) = w_2$$

We need to show that $w_1 + w_2$ is contained in the image:

$$L(v_1 + v_2) = L(v_1) + L(v_2) = w_1 + w_2 \in \text{Im}(L)$$

3. **Closure under scalar multiplication:**

$$L(kv_1) = kL(v_1) = kw_1 \in \text{Im}(L)$$

□

Example 2.1.

$$\begin{aligned} L : K^n &\rightarrow K^m \\ x &\mapsto Ax \end{aligned}$$

is a linear transformation. In this case, x is an $n \times 1$ matrix. A is a $m \times n$ matrix; hence, the RHS becomes $m \times 1$. Checking the conditions:

Well-behaved with respect to addition:

$$A(x + \tilde{x}) = Ax + A\tilde{x} = L(x + \tilde{x}) = L(x) + L(\tilde{x})$$

Well-behaved with respect to SM:

$$A(kx) = k(Ax)$$

$$L(kx) = kL(x)$$

In this case, $\text{Ker}(L) = \text{nullspace}(A)$ and $\text{Im}(L) = \text{colspace}(A)$.

Example 2.2. Derivative map: CLAIM - this is a linear map.

$$\begin{aligned} D : P(\mathbb{R}) &\rightarrow P(\mathbb{R}) \\ p &\mapsto p' \end{aligned}$$

Checking the behaviour:

$$\begin{aligned} D(p + q) &= (p + q)' = kp' = D(p) + D(q) \\ D(kp) &= (kp)' = kp' = kD(p) \end{aligned}$$

Key subspaces:

$$\text{Ker}(D) = P_0(\mathbb{R}) \text{ (The set of all constant polynomials)}$$

$$\text{Im}(D) = P(\mathbb{R}) \text{ (everything)}$$

Example 2.3. We can be even more precise with our subspaces.

$$\begin{aligned} D : P_n(\mathbb{R}) &\rightarrow P_{n-1}(\mathbb{R}) \\ p &\mapsto p' \end{aligned}$$

We proved above that D is linear. The kernel remains the same, but, now the image changes. It becomes: $\text{Im}(D) = P_{n-1}(\mathbb{R})$. This makes the function now **surjective**, since every element in the co-domain is mapped to.

Example 2.4. We can define $D : P_n(\mathbb{R}) \rightarrow P_n(\mathbb{R})$. This is also linear. Compared to the previous example, there is more freedom, especially when it comes to the target. The image then becomes: $\text{Im}(D) = P_{n-1}(\mathbb{R})$. Here, the image becomes the entire space itself; hence, D is no longer surjective.

Definition 2.3. Let I be an interval $I \subseteq \mathbb{R}$. Define $C^o(I) :=$ set of all continuous functions on I . Then, we obtain:

$$C^1(I) := \text{the set of all continuously differentiable functions}$$

$$\vdots$$

$$C^n(I) := \text{the set of all n-times continuously differentiable functions}$$

$$C^\infty(I) := \text{the set of all functions on } I \text{ with derivatives of all orders.}$$

These are all vector spaces over \mathbb{R} . Additionally, we obtain the following subspace chain:

$$C^\infty(I) \leq \dots \leq C^n(I) \leq \dots \leq C^2(I) \leq C(I) \leq C^0(I)$$

In light of this definition, we can re-interpret D in another way:

$$D : C^n(I) \rightarrow C^{n-1}(I) \text{ or}$$

$$D : C^\infty(I) \rightarrow C^\infty(I)$$

Maths fun fact: this is very important in analysis.

Example 2.5. Integration: there are two types of integration maps. The general structure:

$$V = P(\mathbb{R}); a, b \in \mathbb{R}$$

Type I: Definite Integral

$$L : P(\mathbb{R}) \rightarrow \mathbb{R}$$

$$p \mapsto \int_a^b p dx$$

This takes a **function or polynomial** and produces a **real number**. The proof is easy (just use the properties from Calc II).

Type II: Indefinite integral the input here is a polynomial x^n and the output is a polynomial x^{n+1} . It's a map between two function spaces. Symbolically:

$$L : P(\mathbb{R}) \rightarrow P(\mathbb{R})$$

$$p \mapsto \int_a^x p(t) dt$$

The proof that this is linear is really easy.

Type III: indefinite with degree restriction:

$$L : P_n(\mathbb{R}) \rightarrow P_{n+1}(\mathbb{R})$$

$$p \mapsto \int_a^x p(t) dt$$

The kernel is the trivial kernel ($\{0\}$). The image is NOT $P(\mathbb{R})$; instead, it is the set of all polynomials except for those with a root at $x = a$.

Example 2.6. Consider the VS of all sequences ($V=S$), $s = (s_1, s_2, s_3, \dots) : S = \{(x_1, x_2, x_3, \dots) \mid x_i \in \mathbb{R} \forall i\}$. There are two useful maps that we can define:

- **Left-shift:** occur when we shift all sequences to the left.

$$(x_1, x_2, x_3, \dots) \mapsto (x_2, x_3, x_4, \dots)$$

- **Right-shift:** occur when we shift all sequences to the right.

$$(x_1, x_2, x_3, \dots) \mapsto (0, x_1, x_2, x_3, \dots)$$

We need to keep the 0 in the first position to maintain linearity. This is useful for eigenvalues.

Theorem 4. Let V, W , be VS over K . Define: $L : V \rightarrow W$ (linear). Then, L is said to be **injective** $\iff \text{Ker}(L) = \{0\}$. This allows us to quickly check if a linear map is injective. In short: **a linear map is injective if it is injective at 0 no non-trivial component of the intersection.**

Recall that an **injective function** is one-to-one. $f(x) = f(y) \Rightarrow x = y$.

Proof. Two directions.

[\Rightarrow] Let L be injective. We need to prove that the kernel is trivial. Let $v \in \text{Ker}(L)$. Then, $L(v) = 0 = L(0)$. This means that we have two elements in v with the same image. However, by **injectivity**, they must be the same. Hence, $v = 0 \Rightarrow \text{Ker}(L) = \{0\}$.

[\Leftarrow] This is what really sets linear maps apart from other maps. Assume that $\text{Ker}(L) = \{0\}$. Let $v, \tilde{v} \in V$ s.t. $L(v) = L(\tilde{v})$. We need to prove that the two v 's are identical.

$$L(v) = L(\tilde{v}) \Rightarrow L(v) - L(\tilde{v}) = 0 \Rightarrow L(v - \tilde{v}) = 0 \Rightarrow v - \tilde{v} \in \text{ker}(L)$$

By assumption, however, we know that the kernel is trivial. Hence:

$$v - \tilde{v} = 0 \Rightarrow v = \tilde{v}$$

Hence, L is injective and we are done. \square

Theorem 5. This theorem will tell us how we can go about constructing linear maps. Let V be a finite-dimensional VS over k , W and vs over K . Let $\{v_1, \dots, v_n\}$ be a basis for V .

Let $\{w_1, \dots, w_n\}$ be arbitrary vectors in W . Then, there exists a **uniquely determined** linear map $L : V \rightarrow W$ such that:

$$L(v_1) = w_1, \dots, L(v_n) = w_n$$

The crux: the linear map is uniquely determined by the basis.

Proof. This is a big proof that will contain proofs inside it. The overall aim of the proof: to prove **existence and uniqueness**.

First, we will **define L**: v_1, \dots, v_n is a basis. So, we can pick a vector $v \in V$. The fact that it's a basis implies that $\exists k_1, \dots, k_n \in K$ s.t.:

$$v = k_1 v_1 + \dots + k_n v_n$$

We define $L(v)$ to be:

$$L(v) := k_1 w_1 + \dots + k_n w_n$$

Remark: this only makes sense since the coefficients k_1 to k_n are **UNIQUELY** determined; else, we would not have a function.

Linearity of L: To prove that L is linear, then we need to go through the properties: Let $v, \tilde{v} \in V$. We express this in terms of a basis:

$$\begin{aligned} v &= k_1 v_1 + \dots + k_n v_n \\ \tilde{v} &= l_1 v_1 + \dots + l_n v_n \end{aligned}$$

The proof that these are well-behaved with respect to scalar multiplication and addition are straightforward. So, we have proven linear. Now, we need to prove uniqueness.

Proving uniqueness:

Pick an arbitrary linear map.

Let $\tilde{L} : V \rightarrow W$ be linear with the property that $\tilde{L}(v_1) = w_1$ and $\tilde{L}(v_n) = w_n$.

To prove uniqueness, we need to show that $L = \tilde{L}$.

$$\begin{aligned} \Rightarrow \tilde{L}(k_1 v_1) &= k_1 w_1, \dots, \tilde{L}(k_n v_n) = k_n w_n \Rightarrow \tilde{L}(k_1 v_1 + \dots + k_n v_n) = \tilde{L}(k_1 v_1) + \dots + \tilde{L}(k_n v_n) \\ &\Rightarrow k_1 w_1 + \dots + k_n w_n \end{aligned}$$

Recall the definition of L . They are identical. This means that they are identical: $\tilde{L}(v) = L(v) \forall v \in V$.

$\therefore \tilde{L} = L \Rightarrow L$ is uniquely determined.

□

The most important theorem of the section:

Theorem 6 (The Dimension Formula/The Fundamental Theorem). *Let V be a finite-dimensional VS over k , W VS over K , and $L : V \rightarrow W$ be linear. Last class, we learned that with any linear map, we obtain the kernel and image. This formula relates the sizes of those subspaces to the size of V .*

$$\dim(K) + \dim(\text{Im}) = \dim(V) \quad (3)$$

This theorem implies that a big kernel means that we have a highly non-injective mapping. A large image means that the mapping is close to injective. This formula describes the balance between the objects; the “proof is one of the pretty proofs.”

Proof. We will determine the bases by looking at dimension. Define the following:

Let $\{u_1, \dots, u_k\}$ be a basis for $\text{Ker}(L)$

Let $\{w_1, \dots, w_n\}$ be a basis for $\text{Im}(L)$

We need to construct a basis for V . This is the tough part; the w 's live in the wrong space. So, we somehow need to transport the w 's back into v . However, we do know that the image is the set of all elements in W with a pre-image in V .

Since w_1, \dots, w_n are in $\text{Im}(L)$, we know that $\exists v_1, \dots, v_n \in V$ with $L(v_1) = w_1, \dots, L(v_n) = w_n$. However, the choices of v_1, \dots, v_n are generally NOT uniquely determined (there could be an infinite number vectors...). If this is the case, then we can pick an arbitrary vector v ; the choice is completely irrelevant.

We need to prove that all the u 's with the v 's are linearly independent. We also need to prove that they form a basis for V (more important).

Let $B := \{u_1, \dots, u_k, v_1, \dots, v_n\}$. Claim: B is a basis for V . To prove something is a basis for a VS, we need to prove linear independence and spanning.

Proof. **B is linearly independent:** consider the arbitrary combination of the vectors that produce zero:

$$a_1 u_1 + \dots + a_k u_k + b_1 v_1 + \dots + b_n v_n = 0 \quad (4)$$

We need to prove that all the coefficients in this equation are zero. The trick is to apply the linear map to both sides of the equation and use linearity:

$$\begin{aligned} L(a_1u_1 + \dots + a_ku_k + b_1v_1 + \dots + b_nv_n) &= L(0) = 0 \\ aL(u_1) + \dots + a_kL(u_k) + b_1L(v_1) + \dots + b_nL(v_n) &= 0 \end{aligned}$$

We know that the u 's form a basis for the kernel; hence, the transformation will produce zero. We know that the v 's form a basis for the pre-image; hence, the transformation will map them to the image vectors: the w 's. So, we then obtain:

$$b_1w_1 + \dots + b_nw_n = 0$$

The w 's form a basis for the image. So, they are especially linearly independent. So, we have that all the b 's must be zero. If we plug this into the main equation (4), we obtain:

$$a_1u_1 + \dots + a_ku_k + 0 + \dots + 0 = 0$$

We know that the vectors u_1, \dots, u_k form a basis for the kernel. So, they are especially linearly independent. Hence, all their coefficients must be zero. So, we have proven that all scalars involved are zero, which forces B to be linearly independent.

B is spanning: to do this, we take an arbitrary vector in V and prove that we can write it as a linear combination of those vectors. This is a difficult proof, but as always, there is a trick: **map a vector under L** . Such a vector would live in $im(L)$, which we already have a basis for.

Let $v \in V$ be arbitrary. Let $w := L(v)$. Then, $w \in Im(L)$. Thus, we can express that vector as a linear combination of the basis vectors of the image: $\exists b_1, \dots, b_n \in K$ s.t.:

$$w = b_1w_1 + \dots + b_nw_n$$

Let \tilde{v} denote an arbitrary vector in the image. We can replace the w 's in the previous equation with v 's:

$$\tilde{v} = b_1v_1 + \dots + b_nv_n$$

To figure out the relationship between these two vectors, we can consider the difference between v and \tilde{v} and map that under L :

$$L(v - \tilde{v}) = L(v) - L(\tilde{v})$$

Consider $L(v)$. That's precisely going to be in the image. We know from the previous formulas that $L(\tilde{v})$ is going to be equal to $b_1w_1 + \dots + b_nw_n = w$. We obtain the same image:

$$w - w = 0 \Rightarrow v - \tilde{v} \in Ker(L)$$

We have a basis for the kernel: so we can represent $v - \tilde{v}$ as a linear combination of u 's:

$$v - \tilde{v} = a_1u_1 + \dots + a_ku_k$$

Moving \tilde{v} to the LHS and adding the representation of \tilde{v} to both sides of the equation, we then obtain:

$$v = a_1u_1 + \dots + a_ku_k + b_1v_1 + \dots + b_nv_n$$

We have just expressed an arbitrary vector v as a linear combination of the u 's and v 's; hence, B is spanning. Since we have proven both parts, we know that it is a basis. \square

Back to the main proof... since B is a basis for V , we have the following:

$$\dim(B) = |B| = k + n$$

- k is the number of vectors in the basis of the kernel ($k = \dim(\ker(L))$).
- n is the number of elements in the basis of the image ($n = \dim(\text{Im}(L))$).
- So we obtain: $\dim(B) = \dim(\text{Ker}(L)) + \dim(\text{Im}(L))$

□

Example 2.7. $D : P_n(k) \rightarrow P_n(k)$ (differentiation map).

1. The dimension of $P_n(k)$ is $n+1$.
2. $\text{Ker}(D)$ is all constant polynomials, $P_0(k)$. This dimension is one.
3. $\text{Im}(D)$ is all polynomials of degree at most $n - 1$. Hence, this dimension is n .
4. Adding the components together, we obtain:

$$\begin{aligned} \dim(V) &= \dim(\ker(D)) + \dim(\text{Im}(L)) \\ n + 1 &= 1 + n \end{aligned}$$

This is really useful because we can use this formula to determine the dimension of the image, kernel, whole space if we only know some parts of the puzzle. This is useful for proving spanning, basis, and linear independence.

Example 2.8. Let A be a matrix, $A \in \text{mat}(n \times m, \mathbb{R})$. Consider the linear system $Ax = 0$. Here, the set of solutions is the null-space. Recall from Math 133:

- $\dim(\text{nullspace}(A)) + \dim(\text{columnspace}(A)) = m$
- $\dim(\text{colspace}(A)) = \dim(\text{rowspace}(A)) = \text{rank}(A)$.
- In Assignment 2, we:

$$Lx \mapsto Ax, L : k^m \mapsto k^n$$

- $\text{nullspace}(A) = \text{Ker}(L)$
- $\text{colspace}(A) = \text{Im}(L)$
- Since the dimension of kernel + dimension of image + $\dim V$, we obtain:
* $\dim(\text{nullspace}) + \text{rank}(A) = m$

Example 2.9. Consider $\text{tr} : \text{mat}(n \times n, k)$. Recall that the trace is the sum of all the elements on the main diagonal.

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{1n} & \dots & a_{nn} \end{bmatrix} \mapsto \sum a_{11} + \dots + a_{nn}$$

We can show that the trace is linear just by using the properties of matrix arithmetic we learned in Math 133. What's more interesting are the associated subspaces:

1. $\text{Ker}(L)$: the set of all matrices whose trace is zero. There is one constraint here: one degree of freedom. We will prove exactly what this will be using the dimension formula.
2. $\text{Im}(L)$: is the set of all traces, which is the set of all scalars in k . This is essentially \mathbb{R}^k ; so, the dimension of the image is 1.

3. $\dim(V)$ = set of all n -by- n coefficients in k would be n^2 .

4. Hence:

$$\dim(\text{Ker}(L)) = n^2 - 1$$

Example 2.10. Also related to Math 133: let $A \in \text{mat}(n \times m, \mathbb{R})$. This means n rows and m columns. We can assume $n < m$ (rows < cols). Then, the homogeneous linear system $Ax = 0$ has non-trivial solutions.

- The number of columns is equal to the number of variables.
- We have more variables than constraints, so we will obtain non-trivial solutions. Proving with row-reduction is not elegant. So, we will apply the dimension formula.

$L : x \mapsto Ax \Rightarrow Ax = 0$ the system has non-trivial solutions. Our spaces : $K^m \mapsto K^n$

The linear map has a non-trivial kernel. This implies that L has a non-trivial solution. This occurs when the dimension of the kernel is strictly greater than zero. We will prove with the dimension formula:

$$\dim(\text{Ker}(L)) + \dim(\text{Im}(L)) = \dim(V) \quad (\dim(V) = \dim(K^m))$$

$$\dim(\text{Ker}(L)) = \dim(K) - \dim(\text{Im}(L))$$

$$\dim(\text{Ker}(L)) = m - \dim(\text{colspace}(L))$$

$$\dim(\text{Ker}(L)) = m - \dim(\text{rowspace})$$

$$\dim(\text{Ker}(L)) = m - \text{rank}$$

$$\dim(\text{Ker}(L)) = m - n > 0 \therefore \text{ker}(L) \text{ is non-trivial}$$

The kernel is non-trivial implies that the system has non-trivial solutions.