Fall 2019 Math 201-003 at Drexel: blackboard notes of 2019-12-04

Darij Grinberg

December 7, 2019

1. Vector spaces

1.1. Finite-dimensional vector spaces and dimension (continued)

As before, $\mathbb{K}$ denotes a field.

Recall the following from last lecture:

Definition 1.1.1. Let V be a vector space. We say that V is **finite-dimensional** (for short, **fin-dim**) if there exists a finite list of vectors $(v_1, v_2, \ldots, v_k)$ that spans V.

Theorem 1.1.2 (Steinitz's theorem)**.** Let V be a vector space. Let $n \in \mathbb{N}$. Let $v_1, v_2, \ldots, v_n$ be n vectors in V. Then, any $n+1$ combinations of $v_1, v_2, \ldots, v_n$ are dependent.

We proved this theorem last time. Now, let us draw some consequences from it:

Definition 1.1.3. Let V be a fin-dim vector space. Then, the **dimension** of V is defined to be the smallest length of a list of vectors $(v_1, v_2, \ldots, v_k)$ that spans V.

We denote the dimension of V by $\dim V$.

(The **length** of a list is simply its number of entries. For example, the list $(3, 2, 4, 4, 4)$ of numbers has length 5.)

Next, let us define the notion of a basis of a vector space. This is just a straightforward generalization of how we defined a basis of $\mathbb{R}^n$ (in the class notes from 2019-10-09):

Definition 1.1.4. Let V be a vector space. A **basis** of V means a list $(v_1, v_2, \ldots, v_m)$ of vectors in V that is both independent and spans V.

Let us also define "spanning lists" and "independent lists" (just shorthands for lists that span V or are linearly independent, respectively):

Definition 1.1.5. Let V be a vector space. A **spanning list** of V means a list $(v_1, v_2, \ldots, v_m)$ of vectors in V that spans V. We shall also use "spanning" as an adjective (i.e., we say that a list $(v_1, v_2, \ldots, v_m)$ of vectors in V is **spanning** if it is a spanning list).

Definition 1.1.6. Let V be a vector space. An **independent list** of V means a list $(v_1, v_2, \ldots, v_m)$ of vectors in V that is independent.

Thus, a basis of a vector space V is the same as an independent list that is a spanning list at the same time.

Theorem 1.1.7. Let V be a fin-dim vector space with $\dim V = n$.

(a) Any spanning list of V can be shrunk to a basis. ("Shrinking" a list of vectors means removing some vectors from it. The word "some" includes the options "none" and "all". For example, shrinking (v_1, v_2, v_3, v_4) can lead to (v_1, v_3) or to (v_2, v_3, v_4) or to the empty list $()$ or to the full list (v_1, v_2, v_3, v_4).)

(b) Any independent list of V can be extended to a basis. ("Extending" a list of vectors means inserting some new vectors at the end of it. Again, "some" includes the option "none".)

(c) Any basis of V has length n.

(d) There exists a basis of V.

(e) Any spanning list of V has length $\geq n$.

(f) Any spanning list of V that has length n must be a basis of V.

(g) Any independent list of V has length $\leq n$.

(h) Any independent list of V that has length n must be a basis of V.

Proof. **(e)** Recall that $\dim V = n$. Hence, the definition of dimension shows that n is the smallest length of a spanning list of V. So any spanning list of V has length $\geq n$.

(a) Let us first prove the following fact:

> *Fact 1:* If a spanning list of V is dependent, then there is a vector in it that we can remove and still obtain a spanning list of V.

[*Proof of Fact 1:* Let $(v_1, v_2, \ldots, v_k)$ be a spanning list of V that is dependent. Thus, Proposition 1.1.1 in the class notes from 2019-11-25 shows that one of $v_1, v_2, \ldots, v_k$ is a combination of the others. Let v_i be this one vector. Thus,

$$v_i = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k$$

for some scalars $\alpha_j \in \mathbb{K}$. Now, I claim that $(v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k)$ is a spanning list of V.

Indeed, $(v_1, v_2, \ldots, v_k)$ is a spanning list of V. Hence, every vector in V is a combination of $v_1, v_2, \ldots, v_k$. In other words, every vector in V can be written as

$$\begin{aligned}
&\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k \\
&= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{i-1} v_{i-1} \\
&\qquad + \lambda_i \underbrace{v_i}_{=\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k} \\
&\qquad + \lambda_{i+1} v_{i+1} + \cdots + \lambda_k v_k \\
&= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{i-1} v_{i-1} \\
&\qquad + \lambda_i \left(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k\right) \\
&\qquad + \lambda_{i+1} v_{i+1} + \cdots + \lambda_k v_k \\
&= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{i-1} v_{i-1} \\
&\qquad + \lambda_i \alpha_1 v_1 + \lambda_i \alpha_2 v_2 + \cdots + \lambda_i \alpha_{i-1} v_{i-1} + \lambda_i \alpha_{i+1} v_{i+1} + \cdots + \lambda_i \alpha_k v_k \\
&\qquad + \lambda_{i+1} v_{i+1} + \cdots + \lambda_k v_k \\
&= \left(\lambda_1 + \lambda_i \alpha_1\right) v_1 + \left(\lambda_2 + \lambda_i \alpha_2\right) v_2 + \cdots + \left(\lambda_{i-1} + \lambda_i \alpha_{i-1}\right) v_{i-1} \\
&\qquad + \left(\lambda_{i+1} + \lambda_i \alpha_{i+1}\right) v_{i+1} + \cdots + \left(\lambda_k + \lambda_i \alpha_k\right) v_k,
\end{aligned}$$

which means that it is a combination of $v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k$. Thus, $(v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k)$ is a spanning list of V. This proves Fact 1.]

Now, start with any spanning list of V. Fact 1 shows that if this list is dependent, then we can shrink it by removing some vector from it and still obtain a spanning list. By iterating this, we can keep shrinking our spanning list further and further until it becomes independent. This procedure cannot go on forever, because the list gets shorter by 1 at each step. Thus, this procedure must eventually come to an end. At the end, we obtain a spanning list of V that is independent – i.e., a basis of V.

(d) Since V is fin-dim, we know that V has a spanning list. Thus, part **(a)** yields that we can shrink this spanning list to a basis. Hence, V has a basis.

(g) Assume the contrary. Thus, there is an independent list of length $> n$. This independent list must thus have length $\geq n+1$. Hence, there is an independent list of length $n+1$ (since any sublist of an independent list is still an independent list).

On the other hand, there exists a spanning list $(v_1, v_2, \ldots, v_n)$ of V that has length n (since $n = \dim V$). Thus, each vector in V is a combination of $v_1, v_2, \ldots, v_n$. But Steinitz's theorem shows that any $n+1$ combinations of $v_1, v_2, \ldots, v_n$ are dependent. In other words, any $n+1$ vectors in V are dependent (since each vector in V is a combination of $v_1, v_2, \ldots, v_n$). This contradicts the fact that there is an independent list of length $n+1$.

This contradiction shows that our assumption was wrong. Thus, Theorem 1.1.7 **(g)** is proved.

(b) Let us first prove the following fact:

> *Fact 2:* If an independent list of V is not spanning, then there is a vector in V that we can insert into this list and still obtain an independent list of V.

[*Proof of Fact 2:* Let $(v_1, v_2, \ldots, v_k)$ be an independent list of V that is not spanning. Thus, there exists a vector $v \in V$ that is **not** a combination of $v_1, v_2, \ldots, v_k$. Hence, Proposition 1.1.2 the class notes from 2019-11-25 shows that the $k+1$ vectors $v_1, v_2, \ldots, v_k, v$ are independent. Thus, we can insert the vector v into our list $(v_1, v_2, \ldots, v_k)$ and still obtain an independent list of V. This proves Fact 2.]

Now, start with any independent list of V. Fact 2 shows that if this list is not spanning, then we can extend it by adding some vector to it and still obtain an independent list. By iterating this, we can keep expanding our independent list further and further until it becomes spanning. This procedure cannot go on forever, because the list gets longer by 1 at each step, but part **(g)** shows that it cannot become longer than n (without losing its independence). Thus, this procedure must eventually come to an end. At the end, we obtain an independent list of V that is spanning – i.e., a basis of V.

(c) A basis of V is simultaneously a spanning list and an independent list. Hence, part **(e)** shows that its length is $\geq n$, while part **(g)** shows that its length is $\leq n$. Combining these, we conclude that its length is n.

(f) Let $(v_1, v_2, \ldots, v_n)$ be a spanning list of V that has length n. Then, part **(a)** shows that we can shrink this list to a basis $\mathbf{b}$ of V. But this latter basis $\mathbf{b}$ must, too, have length n (by part **(c)**). But the only way to get a length-n list by shrinking a length-n list is to leave it unchanged (i.e., to remove no entries)[1]. Thus, we must have obtained our basis $\mathbf{b}$ from our spanning list $(v_1, v_2, \ldots, v_n)$ by leaving it unchanged. Hence, $\mathbf{b} = (v_1, v_2, \ldots, v_n)$. Thus, the original list $(v_1, v_2, \ldots, v_n)$ is a basis of V (since $\mathbf{b}$ is a basis of V).

(h) This is proved similarly to part **(f)**, but using part **(b)** instead of **(a)**. □

Example 1.1.8. Let $n \in \mathbb{N}$. The vector space $\mathbb{K}^n$ (of column vectors of size n) has dimension n.

To see this, let us construct a basis of $\mathbb{K}^n$:

For each $i \in \{1, 2, \ldots, n\}$, let e_i be the vector $(0, 0, \ldots, 0, 1, 0, 0, \ldots, 0)^T$ (with the 1 being placed in position i). Then, $(e_1, e_2, \ldots, e_n)$ is a basis of $\mathbb{K}^n$, known as the **standard basis** of $\mathbb{K}^n$. Thus, Theorem 1.1.7 **(c)** shows that the dimension of $\mathbb{K}^n$ is n (since this basis has length n).

Example 1.1.9. Let $n, m \in \mathbb{N}$. The vector space $\mathbb{K}^{n \times m}$ (of $n \times m$-matrices) has dimension nm.

Indeed, we can construct a basis as follows:

[1] since any removal of entries would make it shorter

For each $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$, we let $E_{i,j}$ be the $n \times m$-matrix whose (i, j)-th entry is 1 and whose all other entries are 0. For example, if $n = 4$ and $m = 5$, then

$$E_{2,3} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

This $E_{i,j}$ is called the **(i, j)-th matrix unit**. There is a total of nm matrix units in $\mathbb{K}^{n \times m}$. Listing up all these nm matrix units, we obtain a basis of $\mathbb{K}^{n \times m}$. Since this basis has length nm, we conclude that the dimension of $\mathbb{K}^{n \times m}$ is nm.

Let us show this basis in full in the case $n = 2$ and $m = 2$: In this case, we have

$$E_{1,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad E_{1,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$
$$E_{2,1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad E_{2,2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus, the list $(E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$ is a basis of the 4-dimensional vector space $\mathbb{K}^{2 \times 2}$.

Example 1.1.10. Fix $k \in \mathbb{N}$. Let P_k be the $\mathbb{R}$-vector space of polynomials of degree $\leq k$ with real coefficients. Thus,

$$P_k = \left\{ a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k \mid a_0, a_1, \ldots, a_k \in \mathbb{R} \right\}.$$

We have $\dim P_k = k + 1$. Indeed, the list $\left(1, x, x^2, \ldots, x^k\right)$ of length $k + 1$ is a basis of P_k.

The next two examples show that one and the same set can underlie two different vector spaces of two different dimensions:

Example 1.1.11. Recall that $\mathbb{C}$ is an $\mathbb{R}$-vector space. Its dimension is 2, since the list $(1, i)$ is a basis of this vector space.

Example 1.1.12. Recall that $\mathbb{C}$ is a $\mathbb{C}$-vector space. Its dimension is 1, since the list (1) is a basis of this vector space. More generally, any field $\mathbb{K}$ is a $\mathbb{K}$-vector space of dimension 1, with basis (1).

Proposition 1.1.13. **(a)** Any subspace of a fin-dim vector space is fin-dim.

(b) If U is a subspace of a fin-dim vector space V, then U is fin-dim and satisfies $\dim U \leq \dim V$.

(c) If U is a subspace of a fin-dim vector space V that satisfies $\dim U = \dim V$, then $U = V$.

Proof of Proposition 1.1.13 (sketched). **(b)** Let U be a subspace of a fin-dim vector space V.

Let $n = \dim V$. Any independent list of U is clearly an independent list of V, and thus has length $\leq n$ (by Theorem 1.1.7 **(g)**). Now, we claim:

> *Claim 1:* Any independent list of U can be extended to a basis of U.

[*Proof of Claim 1:* If we knew that U is fin-dim, then this would follow from Theorem 1.1.7 **(b)** (applied to U instead of V). But we don't know yet that U is fin-dim. Nevertheless, we can apply the same argument that we used to prove Theorem 1.1.7 **(b)** to U instead of V. The only difference is that we need a new reason why the procedure (of gradually extending our list by inserting new vectors into it) cannot go on forever. But we have such a reason: We know that any independent list of U has length $\leq n$; thus, we cannot insert more than n vectors into our list without breaking its independence. Thus, Claim 1 is proved.]

But the empty list $()$ is clearly an independent list of U. Thus, Claim 1 shows that this empty list $()$ can be extended to a basis of U. Hence, there exists a basis of U. Thus, U is fin-dim (since a basis of U is clearly a list of vectors that spans U).

To prove that $\dim U \leq \dim V$, we start with a basis of U and expand it to a basis of V (Theorem 1.1.7 **(b)** tells us that we can do this, since any basis of U is an independent list of V). Since a list can only grow in length when it is being expanded, we thus conclude that the length of our basis of U is $\leq$ to the length of our basis of V. In other words, $\dim U \leq \dim V$ (since Theorem 1.1.7 **(c)** shows that the dimension of a fin-dim vector space is the length of any basis).

(a) This is just the first half of part **(b)**.

(c) Left to the reader. □

Example 1.1.14. Here are the subspaces of $\mathbb{R}^3$ again:

- The one-element subspace $\left\{ \overrightarrow{0} \right\}$. This has dimension 0.
- The lines through the origin. They have dimension 1.
- The planes through the origin. They have dimension 2.
- The whole space $\mathbb{R}^3$. It has dimension 3.

Remark 1.1.15. The notions of "basis" and of "dimension" can be extended even to vector spaces that are not fin-dim. But this would require modifying them (e.g., the dimension would not longer be a nonnegative integer, but rather a cardinal number). If we did that, we would be able to claim one of the most celebrated results of linear algebra: Every vector space has a basis (in the extended sense of the word).

1.2. The fundamental subspaces of a matrix

Any matrix gives rise to four interesting subspaces:

Definition 1.2.1. Let $n, m \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times m}$ be an $n \times m$-matrix.

(a) We define the **column space** of A to be

$$\operatorname{Col} A = (\text{the span of the columns of } A) = \left\{ Av \mid v \in \mathbb{K}^{m \times 1} \right\}.$$

(b) We define the **row space** of A to be

$$\operatorname{Row} A = (\text{the span of the rows of } A) = \left\{ wA \mid w \in \mathbb{K}^{1 \times n} \right\}.$$

(c) We define the **kernel** (or **nullspace**) of A to be

$$\operatorname{Ker} A = \left\{ v \in \mathbb{K}^{m \times 1} \mid Av = 0_{n \times 1} \right\}.$$

(d) We define the **left kernel** (or **left nullspace**) of A to be

$$\operatorname{LKer} A = \left\{ w \in \mathbb{K}^{1 \times n} \mid wA = 0_{1 \times m} \right\}.$$

It is easy to see:

Proposition 1.2.2. The four sets we just defined are subspaces of the respective vector spaces:

- $\operatorname{Col} A$ is a subspace of $\mathbb{K}^{n \times 1}$.
- $\operatorname{Row} A$ is a subspace of $\mathbb{K}^{1 \times m}$.
- $\operatorname{Ker} A$ is a subspace of $\mathbb{K}^{m \times 1}$.
- $\operatorname{LKer} A$ is a subspace of $\mathbb{K}^{1 \times n}$.

Moreover, the equality signs in parts **(a)** and **(b)** of Definition 1.2.1 hold.

These four sets are called the **four fundamental subspaces** of the matrix A.

Note that $\operatorname{Ker} A$ is the set of all solutions to the system of equations $Av = 0$. Gaussian elimination yields a basis of $\operatorname{Ker} A$.

1.3. The rank of a matrix

Definition 1.3.1. The **rank** of a matrix $A \in \mathbb{K}^{n \times m}$ is defined to be $\dim(\operatorname{Col} A)$. It is denoted by $\operatorname{rank} A$.

The following theorem is far from obvious (see, e.g., [Strickland, §22] for proofs of some of its parts[2]):

Theorem 1.3.2. Let A be an $n \times m$-matrix.
(a) We have $\operatorname{rank} A = \dim(\operatorname{Col} A) = \dim(\operatorname{Row} A)$.
(b) We have $\operatorname{rank} A = \operatorname{rank}(A^T)$.
(c) We have $\dim(\operatorname{Ker} A) = m - \operatorname{rank} A$ and $\dim(\operatorname{LKer} A) = n - \operatorname{rank} A$.
(d) We have $\operatorname{rank} A = (\text{\# of pivots in the RREF of } A)$.

For the next theorem, we need the notion of **column operations**:

Definition 1.3.3. We define **column operations** in the same way as we defined row operations, just replacing each word "row" by "column". Thus, the column operations are the following operations on a matrix:

- **ECO1:** Exchange two columns.
- **ECO2:** Scale a column by a nonzero constant.
- **ECO3:** Add a multiple of a column to another column.

Now, we state the theorem (whose proof can be found in [Strickland, Proposition 22.11][3]):

Theorem 1.3.4. Let A be an $n \times m$-matrix. By a sequence of row **and** column operations, we can transform A into what is called the **rank normal form**: an $n \times m$-matrix whose "diagonal" entries (i.e., the entries in cells $(1,1), (2,2), \ldots, (k,k)$, where $k = \min\{n, m\}$) are

$$\underbrace{1, 1, \ldots, 1}_{\operatorname{rank} A \text{ many 1's}}, 0, 0, \ldots, 0 \qquad \text{(in this order)}$$

and whose all other entries are 0.

Example 1.3.5. Here is an example for Theorem 1.3.4:

Let A be the 3×4-matrix $\begin{pmatrix} 1 & 5 & 2 & 8 \\ -1 & -5 & 1 & 1 \\ 1 & 5 & 1 & 5 \end{pmatrix}$ over the field $\mathbb{Q}$. Its rank is $\operatorname{rank} A = 2$. Thus, Theorem 1.3.4 claims that by a sequence of row and column

[2] Note that [Strickland, §22] writes $\operatorname{img}(A)$ for what we call $\operatorname{Col} A$, and writes $\ker(A)$ for what we call $\operatorname{Ker} A$.

[3] What we call "rank normal form" is being just called "normal form" in [Strickland, §22].

operations, we can transform A into the rank normal form – i.e., into the matrix $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ (since $\operatorname{rank} A = 2$).

Let us see how to actually do this. First, we know how to transform A into RREF using row operations:

$$A = \begin{pmatrix} 1 & 5 & 2 & 8 \\ -1 & -5 & 1 & 1 \\ 1 & 5 & 1 & 5 \end{pmatrix} \xrightarrow{\substack{\text{a sequence of} \\ \text{row operations}}} \begin{pmatrix} \boxed{1} & 5 & 0 & 2 \\ 0 & 0 & \boxed{1} & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(As usual, the pivots are boxed.) Now, let us use column operations (specifically, ECO3) to clear out the nonzero entries to the right of the pivots:

$$\begin{pmatrix} \boxed{1} & 5 & 0 & 2 \\ 0 & 0 & \boxed{1} & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{from column 2}]{\text{subtract } 5 \cdot \text{column 1}} \begin{pmatrix} \boxed{1} & 0 & 0 & 2 \\ 0 & 0 & \boxed{1} & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
$$\xrightarrow[\text{from column 4}]{\text{subtract } 2 \cdot \text{column 1}} \begin{pmatrix} \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{from column 4}]{\text{subtract } 3 \cdot \text{column 3}} \begin{pmatrix} \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We have obtained a matrix whose only nonzero entries are its pivots. Finally, we use column operations (specifically, ECO1) in order to put the pivots into the first two columns:

$$\begin{pmatrix} \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{\text{swap columns 2 and 3}} \begin{pmatrix} \boxed{1} & 0 & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

This is the rank normal form of A.

(As you might have noticed, column operation ECO2 is not needed in this procedure.)

We notice another important property of the rank:

Proposition 1.3.6. The rank of a matrix does not change when we apply row operations or column operations to the matrix.

1.4. Linear maps

Definition 1.4.1. Let V and W be two vector spaces (over a field $\mathbb{K}$). Let $f : V \to W$ be a map. Then, we say that f is **linear** if and only if it satisfies the following three axioms:

- We have $f(v+w) = f(v) + f(w)$ for all $v, w \in V$.
- We have $f(\lambda v) = \lambda \cdot f(v)$ for all $\lambda \in \mathbb{K}$ and $v \in V$.
- We have $f\left(\overrightarrow{0}\right) = \overrightarrow{0}$. (More precisely: Applying f to the zero vector of V yields the zero vector of W.)

Linear maps are also called **homomorphisms of vector spaces**. (Hefferon uses this latter terminology in [Heffer16].)

Example 1.4.2. Let $m, n \in \mathbb{N}$. Let $A \in \mathbb{K}^{n \times m}$. Then, we define the map

$$L_A : \mathbb{K}^m \to \mathbb{K}^n, \qquad v \mapsto Av.$$

This map L_A is linear, because:

- For all $v, w \in \mathbb{K}^m$, we have
$$L_A(v+w) = A(v+w) = Av + Aw = L_A(v) + L_A(w).$$
- For all $\lambda \in \mathbb{K}$ and $v \in \mathbb{K}^m$, we have
$$L_A(\lambda v) = A(\lambda v) = \lambda \cdot Av = \lambda \cdot L_A(v).$$
- We have
$$L_A(0_{m\times 1}) = A \cdot 0_{m \times 1} = 0_{n \times 1}.$$

Moreover, it can be showed that **any** linear map from $\mathbb{K}^m$ to $\mathbb{K}^n$ has the form L_A for an $n \times m$-matrix A.

Thus, the linear maps from $\mathbb{K}^m$ to $\mathbb{K}^n$ are in 1-to-1 correspondence with the $n \times m$-matrices over $\mathbb{K}$. Linear maps in general (i.e., not just between $\mathbb{K}^m$ and $\mathbb{K}^n$) can thus be regarded as "generalized matrices".

A straightforward computation proves the following proposition:

Proposition 1.4.3. Let A be an $n \times m$-matrix. Let B be an $m \times p$-matrix. Then,

$$L_{AB} = L_A \circ L_B.$$

This proposition motivates the definition of the product of two matrices: it was tailored to correspond to the composition of the corresponding linear maps.

When a matrix A is invertible, left-invertible or right-invertible, what does this mean for the corresponding map L_A? The following proposition (whose proof we, again, omit) answers this question:

Proposition 1.4.4. Let A be an $n \times m$-matrix. Then:

(a) The matrix A is invertible if and only if the map L_A is invertible (i.e., bijective). Moreover, if B is the inverse of A, then L_B is the inverse of L_A.

(b) The matrix A is left-invertible if and only if the map L_A is injective.

(c) The matrix A is right-invertible if and only if the map L_A is surjective.

Here are some more examples of linear maps.

Example 1.4.5. Let V be any vector space. Then, the identity map $\operatorname{id} : V \to V$ is always linear. Also linear are the maps

$$V \to V, \qquad v \mapsto 0$$

and

$$V \to V, \qquad v \mapsto -v.$$

More generally, for any $\lambda \in \mathbb{R}$, the map

$$V \to V, \qquad v \mapsto \lambda v$$

is linear.

Example 1.4.6. Let P be the $\mathbb{R}$-vector space of all polynomials in one variable x with real coefficients.

(a) The map $P \to P$, $f \mapsto f + 1$ is **not** linear. (Indeed, for example, it fails to send 0 to 0.)

(b) The map $P \to P$, $f \mapsto f^2$ is **not** linear. (Indeed, for example, it fails to satisfy the first axiom, since two polynomials f and g do not usually satisfy $(f+g)^2 = f^2 + g^2$.)

(c) The map $P \to P$, $f \mapsto f(0)$ is linear. (Because, for example, $(f+g)(0) = f(0) + g(0)$ for any two polynomials f and g.)

(d) The map $P \to P$, $f \mapsto f' = \frac{d}{dx} f$ is linear. (Because, for example, $(f+g)' = f' + g'$ for any two polynomials f and g.)

(e) The map $P \to P$, $f \mapsto x \cdot f$ is linear.

1.5. Representing linear maps by matrices

We have seen above that the linear maps from $\mathbb{K}^m$ to $\mathbb{K}^n$ correspond to $n \times m$-matrices. But even more generally, linear maps between fin-dim vector spaces correspond to matrices, although we need to choose bases of both domain and codomain in order to set up the correspondence (and the correspondence we obtain will depend on the basis we chose). Here is how this works:

Definition 1.5.1. Let V and W be two fin-dim vector spaces.

Let $\mathbf{v} = (v_1, v_2, \ldots, v_m)$ be a basis of V.

Let $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ be a basis of W.

Let $f : V \to W$ be a linear map.

The **matrix repesenting f with respect to the bases $\mathbf{v}$ and $\mathbf{w}$** is the $n \times m$-matrix $M_{\mathbf{v},\mathbf{w},f}$ defined as follows: For each $j \in \{1, 2, \ldots, m\}$, we write the vector $f(v_j)$ as a linear combination of the vectors $w_1, w_2, \ldots, w_n$ as follows:

$$f(v_j) = a_{1,j}w_1 + a_{2,j}w_2 + \cdots + a_{n,j}w_n.$$

Then, $M_{\mathbf{v},\mathbf{w},f}$ is defined to be the $n \times m$-matrix $(a_{i,j})_{1 \le i \le n,\ 1 \le j \le m}$. (Note that the coefficients $a_{i,j}$ exist and are uniquely determined, since $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ is a basis of W.)

Note that Hefferon, in [Heffer16], denotes this matrix $M_{\mathbf{v},\mathbf{w},f}$ by $\operatorname{Rep}_{\mathbf{v},\mathbf{w}}(f)$.

The point of this construction is: Unlike the map f, the matrix $M_{\mathbf{v},\mathbf{w},f}$ is a finite object that we can calculate with. It uniquely determines f (as long as $\mathbf{v}$ and $\mathbf{w}$ are known).

This way of constructing a matrix $M_{\mathbf{v},\mathbf{w},f}$ from a linear map f is inverse to the above construction of a linear map L_A from a matrix A, if the bases are chosen appropriately.

Proposition 1.5.2. Let $m, n \in \mathbb{N}$. Let $\mathbf{e} = (e_1, e_2, \ldots, e_m)$ be the standard basis of $\mathbb{K}^m$. Let $\mathbf{f}$ be the standard basis of $\mathbb{K}^n$. Let $h : \mathbb{K}^m \to \mathbb{K}^n$ be a linear map, and let $A \in \mathbb{K}^{n \times m}$ be an $n \times m$-matrix. Then, we have the following equivalence:

$$(M_{\mathbf{e},\mathbf{f},h} = A) \iff (h = L_A).$$

Also, composition of linear maps corresponds to multiplication of the matrices representing them, provided that we choose the "right" bases for it:

Theorem 1.5.3. Let U, V and W be three fin-dim vector spaces. Let $f : V \to W$ and $g : U \to V$ be two linear maps. Let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ be bases of U, V and W, respectively. Then,

$$M_{\mathbf{v},\mathbf{w},f} \circ M_{\mathbf{u},\mathbf{v},g} = M_{\mathbf{u},\mathbf{w},f \circ g}.$$

Here we have to stop, since the quarter is over.

Further reading:

- Strickland's [Strickland, §19–§22] for a hands-on study of the subspaces of $\mathbb{R}^n$ (including algorithms for finding bases).

- Strickland's [Strickland, §23] for an introduction to the notions of orthogonal and symmetric matrices. (This relies on the notion of **dot product**, which he defines at the beginning of [Strickland, §3].)
- Hefferon's [Heffer16] (especially Chapters Two, Three and Five) for a thorough treatment of vector spaces. (In particular, [Heffer16, §Five.II] explains diagonalization as finding a basis in which a given linear map has a particularly simple representation.)

References

[Heffer16] Jim Hefferon, *Linear Algebra*, 3rd edition 2017.
`http://joshua.smcvt.edu/linearalgebra/`

[Strickland] Neil Strickland, *MAS201 Linear Mathematics for Applications*, lecture notes, 28 September 2013.
`http://neil-strickland.staff.shef.ac.uk/courses/MAS201/`