Linear Algebra
Week 3

G-17

9 X 2025

1 Linear Transformations

Linear Transformations (and linear functionals) are one of the fundamental concepts of linear algebra. Under the hood matrices are tools to represent linear transformations. At this point we should consider matrices as functions.
We already have the intuition on linear transformations since they are equivalent to the function a matrix represents. We will prove this fact but before that here is a formal definition from the lecture notes:

Definition 2.21 (Linear transformation, linear functional) *A function* $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ *is called a lienar transformation and a function* $\mathbb{R}^n \mapsto \mathbb{R}$ *is called a linear functional, if the following linearity axiom hold for all* $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$ *and all* $\lambda_1, \lambda_2 \in \mathbb{R}$.

$$T(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2) = \lambda_1 T(\mathbf{x}_1) + \lambda_2 T(\mathbf{x}_2)$$

As an immediate observation we can say:
Observation 2.2. *Every matrix transformation us a linear transformation.*
Because we defined a matrix transofrmation as $T_A(\mathbf{x}) = A\mathbf{x}$ and we know that $A(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2) = \lambda_1 A\mathbf{x}_1 + \lambda_2 A\mathbf{x}_2$

In the equation $T(\lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2) = \lambda_1 T(\mathbf{x}_1)+\lambda_2 T(\mathbf{x}_2)$ we have two properties. One says that we can first add the inputs and then feed them into the function or first feed the inputs separately to two function instances and then add the outputs. The result will be the same. Second property says we can first multiply the input with a scalar and then feed the product into the function or first give the input to the function and then multiply the output with a scalar. The two results will still be the same. The following lemma takes those two properties apart and offers an alternative definition for linear transformations.

Lemma 2.23. A function $T : \mathbb{R}^n \to \mathbb{R}^m$ / $T : \mathbb{R}^n \to \mathbb{R}$ is a linear transformation / linear functional if and only if the following two linearity axioms hold for all $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ and all $\lambda \in \mathbb{R}$:

(i) $T(\mathbf{x}+\mathbf{x}') = T(\mathbf{x}) + T(\mathbf{x}')$, and

(ii) $T(\lambda\mathbf{x}) = \lambda T(\mathbf{x})$.

Proof: We can show that the two definitions from ***Definition 2.21*** and ***Lemma 2.23.*** are equivalent.

$Def 2.21 \Rightarrow Lemma 2.23$: Choose $\mathbf{x}_1 = \mathbf{x}$ and $\mathbf{x}_2 = \mathbf{x}'$ and $\lambda_1 = \lambda_2 = 1$. Then **Definition 2.21** tells us that

$$T(\mathbf{x}+\mathbf{x}') = T(\lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2) \overset{Def 2.21}{=} \lambda_1 T(\mathbf{x}_1)+\lambda_2 T(\mathbf{x}_2) = T(\mathbf{x}) + T(\mathbf{x}')$$

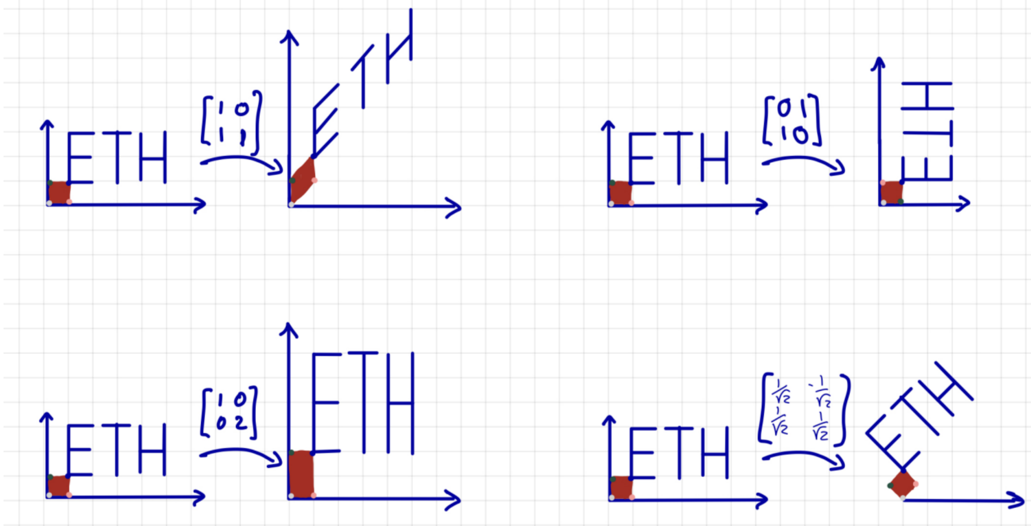
This shows *(i)* from the lemma and similarly you can show *(ii)* by choosing $\mathbf{x}_1 = \mathbf{x}, \mathbf{x}_2 = \mathbf{0}, \lambda_1 = \lambda, \lambda_2 = 0$.

$Lemma 2.23 \Rightarrow Def 2.21$: If we assume **Lemma 2.23** then we can show that ***Definition 2.21*** holds via direct computation and using *(i)* and *(ii)*:

$$T(\lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2) \overset{(i)}{=} T(\lambda_1\mathbf{x}_1) + T(\lambda_2\mathbf{x}_2) \overset{(ii)}{=} \lambda_1 T(\mathbf{x}_1)+\lambda_2 T(\mathbf{x}_2).$$

To understand in what ways a matrix transforms a vector we are specifically interested in the standard unit vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. To intuitively sum up why, it is because we can express everything in a space $\mathbb{R}^n$ using the set of standard unit vectors. If we know what A does to this set of vectors, we can generalize the effect of A to all vectors in $\mathbb{R}^n$.

This is demonstrated below. If you know what happens to the unit square (little red guy), then you can generalize that effect to any other shape or *e.g.* to any set of vectors.



Example from HS23

Lemma 2.24 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ / $T : \mathbb{R}^n \to \mathbb{R}$ be a linear transformation / linear functional. Then

$$T(\mathbf{0}) = \mathbf{0} \ / \ T(\mathbf{0}) = 0.$$

This is very important. So repeat once more:

$\boxed{\mathbf{0} \text{ is always mapped to } \mathbf{0} \text{ by a linear transformation/functional: } T(\mathbf{0}) = \mathbf{0} \ T(\mathbf{0}) = 0}$

An important lemma that formalizes the axiom of linearity for more than two vectors is the following which comes with a proof by induction that I recommend reading from the lecture notes.

Lemma 2.25. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ / $T : \mathbb{R}^n \to \mathbb{R}$ be a linear transformation / linear functional, let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_\ell \in \mathbb{R}^n$ and $\lambda_1, \lambda_2, \ldots, \lambda_\ell \in \mathbb{R}$. Then

$$T\left(\sum_{j=1}^{\ell} \lambda_j \mathbf{x}_j\right) = \sum_{j=1}^{\ell} \lambda_j T(\mathbf{x}_j).$$

1.1 The Matrix of a linear Transformation

We have seen that every matrix A defines a linear transformation T_A above (Observation 2.22). This theorem states the other direction: Every linear transformation $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ (Definition 2.27) is of the form $T = T_A$ for a unique $m \times n$ matrix A.

Theorem 2.29. Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation. There exists a unique $m \times n$ matrix A such that $T = T_A$. This matrix is:

$$A = \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \ldots & T(\mathbf{e}_n) \\ | & | & & | \end{bmatrix}$$

Proof: For a given vector we want to create the same effect as a linear transformation using a matrix. An important set of vectors to observe is the set of unit vectors. Because as mentioned earlier, we can generalize the effect of a linear transformation if we know what happens to the unit cube.

We want an A s.t. $T(\mathbf{x}) = T_A(\mathbf{x}) = A\mathbf{x}$. In particular we want this to hold for the unit vectors: $T(\mathbf{e}_i) = A\mathbf{e}_i$ for $i \in [n]$. But multiplying a matrix with the i-th unit vector means to choose the i-th column of that matrix. And since we want this matrix to have the same effect as the given linear transformation, we put the result of $T(\mathbf{e}_i)$ to the i-th column of the matrix.

$$A = \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \dots & T(\mathbf{e}_n) \\ | & | & & | \end{bmatrix}$$

Now we definitely have $T(\mathbf{e}_i) = A\mathbf{e}_i$. How about other vectors that are not the standard unit vectors? We can deconstruct them as a linear combination of the unit vectors! An example from 3 dimensions is:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{\mathbf{e}_1} + x_2 \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{\mathbf{e}_2} + x_3 \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{\mathbf{e}_3} .$$

So in general we can write $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$. Using this fact and the linearity axiom for linear transformations we can write:

$$T_A(\mathbf{x}) = A\mathbf{x} = \sum_{j=1}^{n} x_j T(\mathbf{e}_j) = T\left(\sum_{j=1}^{n} x_j \mathbf{e}_j\right) = T(\mathbf{x}).$$

The second equality writes the matrix vector multiplication as the linear combination of the columns of the matrix A. Using that A is our candidate matrix from above, we can write $T(\mathbf{e}_j)$ instead of v_j as the j-th column of A. This brings us to the third equation which is just an implementation of the linearity axioms. Last but not least, the fourth equation is justified because $\sum_{j=1}^{n} x_j \mathbf{e}_j$ is equivalent to multiplying the vector $\mathbf{x}$ with the identity matrix.

□

1.2 Matrix Multiplication and Linear Transformation

We can combine linear transformations and represent the composition as matrix multiplication. We should first define what a composition of functions formally is.

Definition 2.33 (Composition of functions). Let $g : X \to Y$ and $f : Y \to Z$ be two functions where X, Y, Z are arbitrary sets. The function $h : X \to Z$,

$$h : x \mapsto f(g(x))$$

is the composition of f and g, written as $f \circ g$ (*"first apply g, then f"*).

Take a close look at the input and output domains. the function g gives outputs in the correct from the "correct" domain to f. Now in the same way we can compose matrix transformations:

Lemma 2.34. Let $T : \mathbb{R}^n \mapsto \mathbb{R}^a$ and $T_B : \mathbb{R}^a \mapsto \mathbb{R}^m$ be two matrix transformations. Then their composition $T_A \circ T_B : \mathbb{R}^n \mapsto \mathbb{R}^m$ is another matrix transformation.

How can we combine the matrices that correspond to the separate operations T_A and T_B to represent $T_C = T_A \circ T_B$? The next lemma tells us:

Lemma 2.35 (Matrix of the composition). Let A be an $a \times n$ matrix and

$$B = \begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_b \\ | & | & & | \end{bmatrix}$$

an $n \times b$ matrix. The $a \times b$ matrix

$$C = \begin{bmatrix} | & | & & | \\ A\mathbf{x}_1 & A\mathbf{x}_2 & \cdots & A\mathbf{x}_b \\ | & | & & | \end{bmatrix}$$

is the unique matrix that satisfies $T_C = T_A \circ T_B$.

You can read the short proof of this from the lecture notes to better understand why this holds. This actually gives us the first insight in matrix multiplication which we are going to define formally as the following:

Definition 2.36 (Matrix multiplication in column notation). Let A be an $a \times n$ matrix and

$$B = \begin{bmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_b \\ | & | & & | \end{bmatrix}$$

an $n \times b$ matrix. The $a \times b$ matrix

$$AB := \begin{bmatrix} | & | & & | \\ A\mathbf{x}_1 & A\mathbf{x}_2 & \cdots & A\mathbf{x}_b \\ | & | & & | \end{bmatrix}$$

is the product of A and B.

Now we can go on and define the matrix of the composition $T_A \circ T_B$ as the product of the matrix multiplication: $T_A \circ T_B = T_{AB}$
Observations about matrix multiplication:

1. You can write matrix matrix multiplication AB in different notations: in column notation as we did above, as a collection of scalar products as in Observation 2.38 or in brackets notation as in observation 2.39 from the lecture notes.

2. (***Lemma 2.40***) $(AB)^\top = B^\top A^\top$ for two matrices with corresponding dimensions.

3. (***Corollary 2.41***) Let I be the $m \times m$ identity matrix. Then $IA = A$ for all $m \times n$ matrices, and $AI = A$ for all $n \times m$ matrices.

Last but not least matrix multiplication is distributive and associative. You can verify this by keeping track of each entry in the multiplication or thinking of matrices as linear transformations:

Lemma 2.42. Let A, B, C be three matrices. Whenever the respective sums and products in the following are defined, we have

[(i)]

1. $A(B + C) = AB + AC$ and $(A + B)C = AC + BC$; (distributivity)

2. $(AB)C = A(BC)$. (associativity)

1.3 Image and Kernel

Here are the definitions for Image and Kernel of a matrix (from lecture notes). We are going to examine them closer in a couple of weeks but the keywords might already come up in external sources, assignments and sometimes in the lecture.

Definition 2.27 (Kernel and image). Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The set

$$\boldsymbol{Ker}(T) := \{ \mathbf{x} \in \mathbb{R}^n : T(\mathbf{x}) = 0 \} \subseteq \mathbb{R}^n$$

is the *kernel* of T. The set

$$\boldsymbol{Im}(T) := \{ T(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n \} \subseteq \mathbb{R}^m$$

is the *image* of T.

2 Hints

1. In Class.
2. Bonus! No hints.
3. Try to find out what the matrix transformation does to the unit vectors. One of them is already given. Find $T(\mathbf{e}_2)$.
4. This is an if and only if proof so you have to prove both directions. In one direction you prove if T is a linear transformation then $v_{n+1} = \mathbf{0}$ Think about what should hold if T is a linear transformation. In the other direction you should prove if $v_{n+1} = \mathbf{0}$ then T is a linear transformation. Remember how you prove if something is a linear transformation.
5. You can write the left hand side as a matrix. Do that and choose x, y, z such that all entries are 0. **b)** no hints. **c)** use the previous question. **d)** Don't overthink, this is direct computation. **e)** Calculate a small example e.g. 3×3. What do you notice about the $k-th$ column of T^k?
6. No hints.

mkilic

♠