Linear Algebra
Week 3

G-02

8 X 2026

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. We are going to see that matrix transformations and linear transformations are equivalent. At this point we can 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 2.19 (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 two linearity axioms hold for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$ and all $\lambda \in \mathbb{R}$.*

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

(ii) $T(\lambda \mathbf{x}_1) = \lambda T(\mathbf{x}_1)$

As an immediate observation we can say:
Observation 2.20. *Every matrix transformation is a linear transformation.*

Because we defined a matrix transofrmation as $T_A(\mathbf{x}) = A\mathbf{x}$ and we know that $A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2$ and $A(\lambda\mathbf{x}_1) = \lambda A(\mathbf{x}_1)$, we know that matrix transformations satisfy the linearity axioms.

Lemma 2.21 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ / $T : \mathbb{R}^n \to \mathbb{R}$ be a linear transformation / linear functional. Then

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

This is very important. So repeat once more:

$\mathbf{0}$ is always mapped to $\mathbf{0}$ 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 2.22. 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

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

1.0.1 Examples of Linear Transformations

In this subsection we will list some examples and non-examples of linear transformations. As an exercise, you can prove that they are linear transformations. The recipe is always the same, nothing fancy: Prove that the two linearity axioms hold. We are going to provide an example proof and counterexamples.

Linear Transformations

Scalar product Let $\mathbf{v} \in \mathbb{R}^m$ be a vector. Then $T : \mathbb{R}^m \to \mathbb{R}$

$$T(\mathbf{x}) = \sum_{i=1}^{m} v_i x_i = \mathbf{v}^\top \mathbf{x}$$

is a linear functional.

Identity The identity function $\mathrm{id} : \mathbb{R}^n \to \mathbb{R}^n$

$$\mathrm{id}(\mathbf{x}) = \mathbf{x}$$

is a linear transformation.

Constant Zero Let $\mathbf{0} \in \mathbb{R}^n$ be the zero vector. Then $T : \mathbb{R}^m \to \mathbb{R}^n$:

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

is a linear transformation.

Differentiation(informal) Let $\mathbb{R}[x]$ be the vector space of polynomials in one variable as defined in *Def 4.4* of the lecture notes. Further let $P \subseteq \mathbb{R}[x]$ define subspace of polynomials of degree at most 3. The differentiation function $T_d : P \to P$:

$$T_d(\mathbf{p}) = \frac{d}{dx}\mathbf{p}$$

that gives the derivative of the polynomial is a linear transformation.

Proof. Let $\mathbf{p} = a_3x^3 + a_2x^2 + a_1x + a_0$ and $\mathbf{q} = b_3x^3 + b_2x^2 + b_1x + b_0$ be two polynomials of degree at most 3. Then we have:

$$\begin{aligned}
T_d(\mathbf{p}+\mathbf{q}) &= T_d((a_3+b_3)x^3 + (a_2+b_2)x^2 + (a_1+b_1)x + (a_0+b_0)) \\
&= \frac{d}{dx}((a_3+b_3)x^3 + (a_2+b_2)x^2 + (a_1+b_1)x + (a_0+b_0)) \\
&= 3\cdot(a_3+b_3)x^2 + 2\cdot(a_2+b_2)x^1 + (a_1+b_1) + 0 \\
&= 3\cdot a_3x^2 + 3\cdot b_3x^2 + 2\cdot a_2x^1 + 2\cdot b_2x^1 + a_1 + b_1 + 0 + 0 \\
&= [3\cdot a_3x^2 + 2\cdot a_2x^1 + a_1 + 0] + [3\cdot b_3x^2 + 2\cdot b_2x^1 + b_1 + 0] \\
&= \frac{d}{dx}(a_3x^3 + a_2x^2 + a_1x + a_0) + \frac{d}{dx}(b_3x^3 + b_2x^2 + b_1x + b_0) \\
&= \frac{d}{dx}\mathbf{p} + \frac{d}{dx}\mathbf{q} \\
&= T_d(\mathbf{p}) + T_d(\mathbf{q})
\end{aligned}$$

and for $\lambda \in \mathbb{R}$ we have

$$\begin{aligned}
T_d(\lambda\mathbf{p}) &= T_d(\lambda\,(a_3x^3 + a_2x^2 + a_1x + a_0)) \\
&= T_d(\lambda a_3x^3 + \lambda a_2x^2 + \lambda a_1x + \lambda a_0) \\
&= \frac{d}{dx}(\lambda a_3x^3 + \lambda a_2x^2 + \lambda a_1x + \lambda a_0) \\
&= 3\cdot\lambda a_3x^2 + 2\cdot\lambda a_2x^1 + 1\cdot\lambda a_1 + \lambda\cdot 0 \\
&= \lambda\,(3a_3x^2 + 2a_2x^1 + 1a_1 + 0) \\
&= \lambda\,\frac{d}{dx}(a_3x^3 + a_2x^2 + a_1x + a_0) \\
&= \lambda\,T_d(\mathbf{p})
\end{aligned}$$

Since T_d satisfies both linearity axioms, T_d is a linear transformation. □ (*Note that this is an informal example since we do not define the differentiation operation formally in the lecture notes.*)

Non-linear Transformations

Nonzero Constant Let $\mathbf{c} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ be a nonzero vector. Then $T : \mathbb{R}^m \to \mathbb{R}^n$:

$$T(\mathbf{x}) = \mathbf{c}$$

is not a linear transformation.

Euclidean Norm The transformation $T : \mathbb{R}^m \to \mathbb{R}$

$$T(\mathbf{x}) = \sqrt{\sum_{i=1}^{n} x_i^2} = \|\mathbf{x}\|$$

is not a linear transformation. *Counterexample.*

$$T\left(\begin{pmatrix}1\\1\end{pmatrix} + \begin{pmatrix}-1\\-1\end{pmatrix}\right) = T(\mathbf{0}) = 0 \neq \sqrt{2} + \sqrt{2} = T\left(\begin{pmatrix}1\\1\end{pmatrix}\right) + T\left(\begin{pmatrix}-1\\-1\end{pmatrix}\right)$$

1.1 The Matrix of a linear Transformation

We have seen that every matrix A defines a linear transformation T_A above (Observation 2.20). This theorem states the other direction: Every linear transformation $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is of the form $T = T_A$ for a unique $m \times n$ matrix A.

Theorem 2.23. Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation. There is exactly one $m \times n$ matrix A such that $T = T_A$ (meaning that $T(\mathbf{x}) = T_A(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$). This matrix is:

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

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 $i \in [n]$. And multiplying a matrix with the i-th unit vector means to choose the i-th column of that matrix. 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.

□

We knew that all matrix transformations are linear. Now we have proven that every linear transformation is also a matrix transformation. This means that the two types of transformations are *equivalent*! This fundamental result applies to many fields. One example is algebra. Remember how we showed above that differentiation is a linear transformation for polynomials of degree at most 3? For a polynomial of degree at most 3, represented as the vector of its coefficients $\mathbf{p} = (a_0\, a_1\, a_2\, a_3)^\top$ the matrix corresponding to differentiation

is:

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Try it for a polynomial! Write its coefficients as a vector and multiply it with this matrix.

2 Matrix Multiplication

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

Definition 2.31 (Composition of linear transformations). Let $T : \mathbb{R}^n \to \mathbb{R}^k$ and $S : \mathbb{R}^k \to \mathbb{R}^m$ be two linear transformations. The composition $S \circ T : \mathbb{R}^n \to \mathbb{R}^m$ is defined by $(S \circ T)(\mathbf{x}) = S(T(\mathbf{x}))$

Take a close look at the input and output domains. The transformation T gives outputs in the correct domain that S accepts as input. We can further establish that composition of two linear transformations is a linear transformation:

Lemma 2.32. Let $T : \mathbb{R}^n \mapsto \mathbb{R}^k$ and $S : \mathbb{R}^k \mapsto \mathbb{R}^m$ be two linear transformations. Then $S \circ T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is another linear transformation.

You can show the linearity axioms for $S \circ T$ using $(S \circ T)(\mathbf{x}) = S(T(\mathbf{x}))$ and the fact that T and S are linear transformations. Since we showed that matrix transformations are equivalent to linear transformations, we can say that the composition of two matrix transformations $T_A : \mathbb{R}^k \to \mathbb{R}^m$, $T_B : \mathbb{R}^n \to \mathbb{R}^k$ is another matrix transformation: $T_A \circ T_B : \mathbb{R}^n \to \mathbb{R}^m$. The matrix corresponding to the composition is the multiplication of matrices A and B.

Lemma 2.33 (Matrix of the composition). Let A be an $m \times k$ matrix and

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

a $k \times n$ matrix. The $m \times n$ matrix

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

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 2.34 (Matrix multiplication in column notation). Let A be an $m \times k$ matrix and

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

a $k \times n$ matrix. The $m \times n$ matrix

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

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: $T_A \circ T_B = T_{AB}$ (*See Corollary 2.35 for a more formal version.*)

How to multiply matrices?

There are many ways of thinking about the matrix multiplication. All ways of multiplying matrices yield the same result, but depending on the context one or the other might be more intuitive.

Collection of Linear Combinations *(Similar to Def 2.34)* This is the closest one to the definition. We already know that matrix vector multiplication is actually a compact representation of the linear combination of the columns of the matrix where the coefficients are the elements of the vector. When we multiply a matrix with another matrix, we do a matrix vector multiplication with each column of the matrix on the right side: Let A be an $m \times k$ matrix and

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

a $k \times n$ matrix. The $m \times n$ matrix

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

includes n different linear combinations of the columns of A.

You can find an interactive example on `linearalgebra.ch`

Collection of Scalar products *(Similar to observation 2.36)* We can extend the above way of multiplying matrices as a collection of scalar products. When you multiply the column vector $\mathbf{x}_i$ with the matrix A, you are actually taking the scalar product of $\mathbf{x}_i$ with A's rows. Hence the entry that sits at i-th row and j-th column of the product is the scalar product of the i-th row of matrix A and the j-th column of matrix B. For the matrix A

$$A = \begin{bmatrix} — & \mathbf{u}_1^\top & — \\ — & \mathbf{u}_2^\top & — \\ & \vdots & \\ — & \mathbf{u}_m^\top & — \end{bmatrix}$$

and the matrix B as above, we can write the product as:

$$AB = \begin{bmatrix} \mathbf{u}_1^\top \mathbf{x}_1 & \mathbf{u}_1^\top \mathbf{x}_2 & \cdots & \mathbf{u}_1^\top \mathbf{x}_n \\ \mathbf{u}_2^\top \mathbf{x}_1 & \mathbf{u}_2^\top \mathbf{x}_2 & \cdots & \mathbf{u}_2^\top \mathbf{x}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_m^\top \mathbf{x}_1 & \mathbf{u}_m^\top \mathbf{x}_2 & \cdots & \mathbf{u}_m^\top \mathbf{x}_n \end{bmatrix}$$

You can find an interactive example on `linearalgebra.ch`

Linear Combination of Rows You can shift the focus to the matrix B and write the matrix product as the linear combination of the rows of matrix B. Calling it a linear combination of the rows is a slight abuse of notation since we have covectors involved. However as *everything is matrix multiplication* (Section 2.3.4 of the lecture notes) this trick is legal.

$$AB = \begin{bmatrix} — & \mathbf{u}_1^\top B & — \\ — & \mathbf{u}_2^\top B & — \\ & \vdots & \\ — & \mathbf{u}_m^\top B & — \end{bmatrix}$$

Here, each row is a linear combination of the rows of B. Let B written in row notation be

$$B = \begin{bmatrix} — & \mathbf{b}_1^\top & — \\ — & \mathbf{b}_2^\top & — \\ & \vdots & \\ — & \mathbf{b}_k^\top & — \end{bmatrix}$$

Then for a vector $\mathbf{v} \in \mathbb{R}^k$, $\mathbf{v}^\top B$ is nothing else but the sum of the rows of B scaled by the coefficients as the elements of the vector $\mathbf{v}$:

$$\begin{array}{ll} v_1 \cdot & \begin{bmatrix} — & \mathbf{b}_1^\top & — \end{bmatrix} \\ v_2 \cdot & \begin{bmatrix} — & \mathbf{b}_2^\top & — \end{bmatrix} \\ \vdots & \\ v_k \cdot & \begin{bmatrix} — & \mathbf{b}_k^\top & — \end{bmatrix} \\ + & \\ \hline & \mathbf{v}^\top B \end{array}$$

Note that this is a slightly different way of viewing the covector-matrix multiplication than it is defined in Definition 2.41

You can find an interactive example on `linearalgebra.ch`

Observations about matrix multiplication:

1. You can only multiply matrices if their dimensions match. The product AB is defined if and only if A has as many columns as B has rows.
2. $(AB)^\top = B^\top A^\top$ for two matrices $A \in \mathbb{R}^{m\times k}$ and $B \in \mathbb{R}^{k\times n}$. (See formal version at ***Lemma 2.38***)
3. 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. (See formal version at ***Corollary 2.39***)

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

1. $A\,(B + C) = AB + AC$ and $(A + B)\,C = AC + BC$ (distributivity)
2. $(AB)C = A(BC)$. (associativity)

2.1 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 2.25 (Kernel and image). Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The set

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

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.

3 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. Very nice exercise, no spoilers.

mkilic

