Linear Transformations

This week we will learn about:

- Linear transformations,
- The standard matrix of a linear transformation,
- Composition and powers of linear transformations, and
- Change of basis for linear transformations.

Extra reading and watching:

- Section 1.2.3 in the textbook
- Lecture videos 9, 10, 11, and 12 on YouTube
- Linear map at Wikipedia
- Transformation matrix at Wikipedia

Extra textbook problems:

$\star$ 1.2.3

$\star\star$ 1.2.6, 1.2.11, 1.2.32

$\star\star\star$ 1.2.12, 1.2.28, 1.2.30

☠ none this week

Last week, we learned that we could use bases to represent vectors in (finite-dimensional) vector spaces very concretely as tuples in $\mathbb{R}^n$ (or $\mathbb{F}^n$, where $\mathbb{F}$ is the field you're working in), thus turning almost any vector space problem into one that you learned how to solve in the previous course.

We will now introduce linear transformations between general vector spaces, and see that bases let us similarly think of any linear transformation (on finite-dimensional vector spaces) as a matrix in $\mathcal{M}_{m,n}$.

Definition 3.1 — Linear Transformations

Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over the same field $\mathbb{F}$. A **linear transformation** is a function $T : \mathcal{V} \to \mathcal{W}$ that satisfies the following two properties:

a) $T(\vec{v}+\vec{w}) = T(\vec{v}) + T(\vec{w})$ for all $\vec{v}, \vec{w} \in \mathcal{V}$, and

b) $T(c\vec{v}) = cT(\vec{v})$ for all $\vec{v} \in \mathcal{V}$ and $c \in \mathbb{F}$.

Example. *Every matrix transformation is a linear transformation. That is,*

If $A \in M_{m,n}$ then the function $T_A : \mathbb{R}^n \to \mathbb{R}^m$, defined by $T_A(\vec{v}) = A\vec{v}$, is a linear transformation.

Let's check:

a) $T_A(\vec{v}+\vec{w}) = A(\vec{v}+\vec{w}) = A\vec{v} + A\vec{w} = T_A(\vec{v}) + T_A(\vec{w})$ ✓

b) $T_A(c\vec{v}) = A(c\vec{v}) = cA\vec{v} = cT_A(\vec{v})$ ✓

Yay!

Example. *Is the function $T : \mathcal{M}_{m,n} \to \mathcal{M}_{n,m}$ that sends a matrix to its transpose a linear transformation?*

Again, just check the two properties:

a) $(A+B)^T = A^T + B^T$ ✓

b) $(cA)^T = cA^T$ ✓

} by previous course

Example. *Is the function* det : $\mathcal{M}_n \to \mathbb{R}$ *that sends a matrix to its determinant a linear transformation?*

Again, just check the two properties:

a) $\det(A+B) \stackrel{?}{=} \det(A) + \det(B)$

No! If $A = B = I$ then $\det(I+I) = \det(2I) = 2^n$,

but $\det(I) + \det(I) = 1 + 1 = 2.$ ← not the same!

Example. *Is the differentiation map* $D : \mathcal{D} \to \mathcal{F}$*, which sends a differentiable function to its derivative, a linear transformation?*

vector space of differentiable functions

Again, just check the two properties:

a) $D(f+g) = (f+g)' = f' + g' = D(f) + D(g)$ ✓

b) $D(cf) = (cf)' = cf' = cD(f)$ ✓

(Both properties are learned in first-year calculus)

$\therefore$ Yes, D is a linear transformation.

Before proceeding to prove things about linear transformations, we make some notes:

- We can sometimes consider the same linear transformation as acting on different vector spaces. For example, we can similarly consider D as a linear transformation from $\mathcal{P}^3$ to $\mathcal{P}^2$.
- For all linear transformations $T : \mathcal{V} \to \mathcal{W}$, it is true that $T(\mathbf{0}) = \mathbf{0}$.

Proof: $T(\vec{0}) = T(0\vec{v}) = 0T(\vec{v}) = \vec{0}.$ (scalar)

- The **zero transformation** $O : \mathcal{V} \to \mathcal{W}$ is the one defined by $O(\vec{v}) = \vec{0}$ for all $\vec{v} \in \mathcal{V}$.
- The **identity transformation** $I : \mathcal{V} \to \mathcal{V}$ is the one defined by $I(\vec{v}) = \vec{v}$ for all $\vec{v} \in \mathcal{V}$.

The Standard Matrix

We now do for linear transformations what we did for vectors last week: we give them "coordinates" so that we can explicitly write them down using numbers in the ground field.

Theorem 3.1 — Standard Matrix of a Linear Transformation

Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces with bases B and D, respectively, where $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $\mathcal{W}$ is m-dimensional. A function $T : \mathcal{V} \to \mathcal{W}$ is a linear transformation if and only if there exists a matrix $[T]_{D\leftarrow B} \in \mathcal{M}_{m,n}$ for which

$$[T(\vec{v})]_D = [T]_{D\leftarrow B}[\vec{v}]_B \quad \text{for all } \vec{v} \in \mathcal{V}.$$

Furthermore, the unique matrix $[T]_{D\leftarrow B}$ with this property is called the **standard matrix** of T with respect to the bases B and D, and it is

$$[T]_{D\leftarrow B} = \left[\,[T(\vec{v}_1)]_D \mid [T(\vec{v}_2)]_D \mid \cdots \mid [T(\vec{v}_n)]_D\,\right].$$

Before proving this theorem, we make some notes:

- The matrix $[T]_{D\leftarrow B}$ tells us how to convert coordinate vectors of $\mathbf{v} \in \mathcal{V}$ to coordinate vectors of $T(\mathbf{v}) \in \mathcal{W}$.
- Using this theorem, we can think of every linear transformation $T : \mathcal{V} \to \mathcal{W}$ as a matrix.
- The standard matrix looks different depending on the bases B and D,

Just like coordinate vectors.

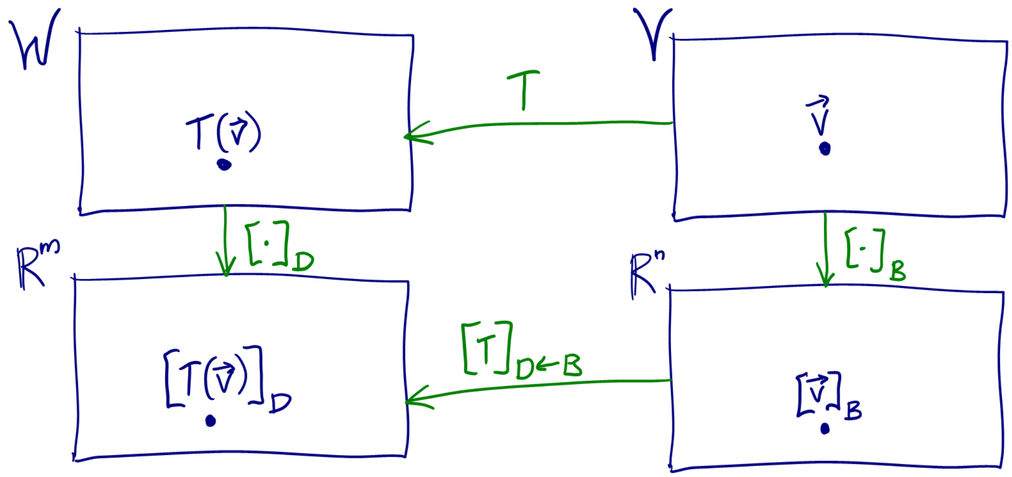
Proof of Theorem 3.1. We just do block matrix multiplication:

Write $\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n$, so $[\vec{v}]_B = (c_1, c_2, \ldots, c_n)$.

Then

$$\begin{aligned}
[T]_{D\leftarrow B}[\vec{v}]_B &= \left[\,[T(\vec{v}_1)]_D \mid [T(\vec{v}_2)]_D \mid \cdots \mid [T(\vec{v}_n)]_D\,\right]\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \\
&= c_1[T(\vec{v}_1)]_D + c_2[T(\vec{v}_2)]_D + \cdots + c_n[T(\vec{v}_n)]_D \\
&= [c_1T(\vec{v}_1) + c_2T(\vec{v}_2) + \cdots + c_nT(\vec{v}_n)]_D \\
&= [T(c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n)]_D = [T(\vec{v})]_D. \quad \blacksquare
\end{aligned}$$

Standard matrices can perhaps be made a bit simpler to understand if we draw a schematic of how they work:



Example. *Find the standard matrix of the transpose map on $\mathcal{M}_2$ with respect to the standard basis $\{E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2}\}$.* $= B = D$

$$[T]_B = \Big[[T(E_{1,1})]_D \;\Big|\; [T(E_{1,2})]_D \;\Big|\; [T(E_{2,1})]_D \;\Big|\; [T(E_{2,2})]_D \Big]$$

abbreviated notation, since $B = D$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Side calculation:
$T(E_{1,1}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $T(E_{1,2}) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, etc, so
$[T(E_{1,1})]_D = (1,0,0,0)$, $[T(E_{1,2})]_D = (0,0,1,0)$, etc.

Example. *Find the standard matrix of the differentiation map $D : \mathcal{P}^3 \to \mathcal{P}^3$ with respect to the standard basis $\{1, x, x^2, x^3\} \subset \mathcal{P}^3$.* $= B$

$D(1) = 0$, $D(x) = 1$, $D(x^2) = 2x$, $D(x^3) = 3x^2$,
$[D(1)]_B = (0,0,0,0)$, $[D(x)]_B = (1,0,0,0)$, $[D(x^2)]_B = (0,2,0,0)$, $[D(x^3)]_B = (0,0,3,0)$.

$$\therefore\ [D]_B = \Big[[D(1)]_B \;\Big|\; [D(x)]_B \;\Big|\; [D(x^2)]_B \;\Big|\; [D(x^3)]_B \Big] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Check: If $p(x) = a + bx + cx^2 + dx^3$ then $D(p(x)) = b + 2cx + 3dx^2$.
$[p]_B = (a,b,c,d)$ $\qquad [D(p)]_B = (b, 2c, 3d, 0)$.

Composition and Powers of Linear Transformations

It is often useful to consider the effect of applying two or more linear transformations to a vector, one after another. Rather than thinking of these linear transformations as separate objects that are applied in sequence, we can combine their effect into a single new function that is called their **composition**:

$$(S \circ T)(\vec{v}) = S(T(\vec{v})) \quad \text{for all } \vec{v} \in \mathcal{V}.$$

The following theorem tells us that we can find the standard matrix of the composition of two linear transformations simply via matrix multiplication (as long as the bases "match up").

Theorem 3.2 — Composition of Linear Transformations

Suppose $\mathcal{V}$, $\mathcal{W}$, and $\mathcal{X}$ are finite-dimensional vector spaces with bases B, C, and D, respectively. If $T : \mathcal{V} \to \mathcal{W}$ and $S : \mathcal{W} \to \mathcal{X}$ are linear transformations then $S \circ T : \mathcal{V} \to \mathcal{X}$ is a linear transformation, and its standard matrix is

$$[S \circ T]_{D \leftarrow B} = [S]_{D \leftarrow C}[T]_{C \leftarrow B}.$$

Proof. We just need to show that $[(S \circ T)(\mathbf{v})]_D = [S]_{D\leftarrow C}[T]_{C\leftarrow B}[\mathbf{v}]_B$ for all $\mathbf{v} \in \mathcal{V}$. To this end,

we just use Theorem 3.1 twice:

$$[S]_{D\leftarrow C}[T]_{C\leftarrow B}[\vec{v}]_B \underset{\text{by Theorem 3.1}}{=} [S]_{D\leftarrow C}[T(\vec{v})]_C \underset{\text{by Theorem 3.1}}{=} [S(T(\vec{v}))]_D = [(S\circ T)(\vec{v})]_D.$$

Since $[S\circ T]_{D\leftarrow B}[\vec{v}]_B = [(S\circ T)(\vec{v})]_D$ for all $\vec{v} \in \mathcal{V}$ too (again, by Theorem 3.1), we see that $[S]_{D\leftarrow C}[T]_{C\leftarrow B} = [S\circ T]_{D\leftarrow B}$. ■

In the special case when the linear transformations that we are composing are equal to each other, we get **powers** of those transformations:

$T^k = T \circ T \circ \cdots \circ T.$ (k times)

In this special case, the previous theorem tells us that we can find the standard matrix of a power of a linear transformation by computing the corresponding power of the standard matrix of the original linear transformation.

Example. *Use standard matrices to compute the fourth derivative of $x^2e^x + 2xe^x$.*

Let $B = \{e^x, xe^x, x^2e^x\}$, $V = \text{span}(B)$, and
$D: V \to V$ be the differentiation map.

$D(e^x) = e^x$, $D(xe^x) = e^x + xe^x$, $D(x^2e^x) = 2xe^x + x^2e^x$.

$[D(e^x)]_B = (1,0,0)$, $[D(xe^x)]_B = (1,1,0)$, $[D(x^2e^x)]_B = (0,2,1)$.

$$\therefore \quad [D]_B = \Big[[D(e^x)]_B \,\Big|\, [D(xe^x)]_B \,\Big|\, [D(x^2e^x)]_B \Big] = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$[D^2]_B = [D]_B^2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}, \quad [D^4]_B = [D^2]_B^2 = \begin{bmatrix} 1 & 4 & 12 \\ 0 & 1 & 8 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$\therefore [D^4(x^2e^x + 2xe^x)]_B = [D^4]_B [x^2e^x + 2xe^x]_B = \begin{bmatrix} 1 & 4 & 12 \\ 0 & 1 & 8 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 20 \\ 10 \\ 1 \end{bmatrix}.$$

$$\therefore \quad D^4(x^2e^x + 2xe^x) = 20e^x + 10xe^x + x^2e^x.$$

Later on in this course, we will learn how to come up with a formula for powers of arbitrary matrices, which will let us (for example) find a formula for the n-th derivative of $x^2e^x + 2xe^x$.

Change of Basis for Linear Transformations

Recall that last week we learned how to convert a coordinate vector from one basis B to another basis C. We now learn how to do the same thing for linear transformations: we will see how to convert a standard matrix with respect to bases B and D to a standard matrix with respect to bases C and E.

Fortunately, we already did most of the hard work last week when we introduced change-of-basis matrices, so we can just "stitch things together" to make them work in this setting.

Theorem 3.3 — Change of Basis for Linear Transformations

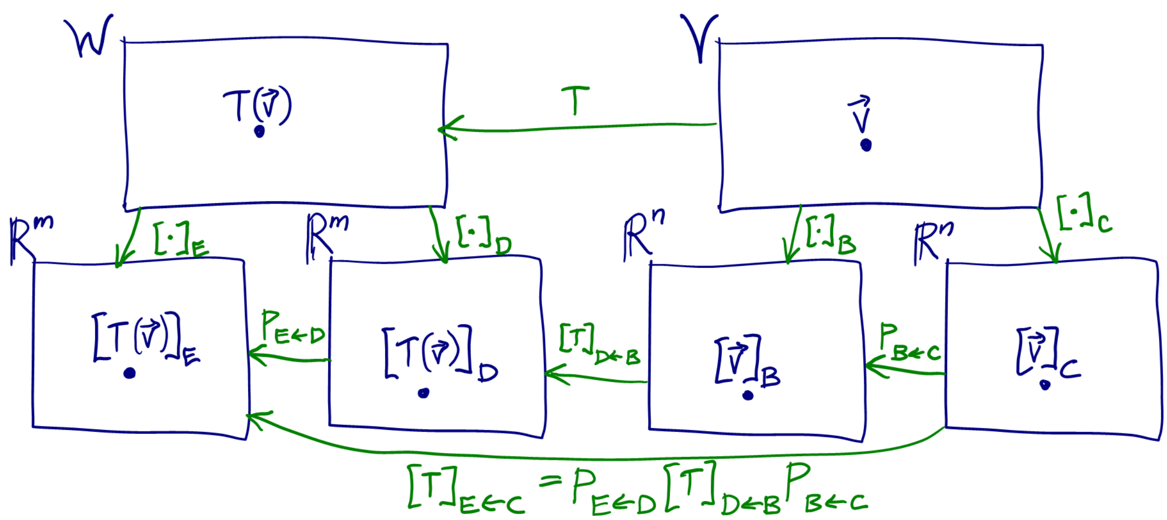
Let $T : \mathcal{V} \to \mathcal{W}$ be a linear transformation between finite-dimensional vector spaces $\mathcal{V}$ and $\mathcal{W}$, and let B and C be bases of $\mathcal{V}$, while D and E are bases of $\mathcal{W}$. Then

$$[T]_{E\leftarrow C} = P_{E\leftarrow D}[T]_{D\leftarrow B}P_{B\leftarrow C}.$$

change-of-basis matrices

The above theorem is made easier to remember by noting that adjacent subscripts always match (e.g., the two Ds are next to each other) and the outer subscripts on the left- and right-hand sides are the same (E's on the far left and C's on the far right).

We can also make sense of the theorem via a diagram:



Proof of Theorem 3.3. Let's think about what happens if we multiply $P_{E\leftarrow D}[T]_{D\leftarrow B}P_{B\leftarrow C}$ on the right by a coordinate vector $[\mathbf{v}]_C$:

$$\begin{aligned} P_{E\leftarrow D}[T]_{D\leftarrow B}P_{B\leftarrow C}[\vec{v}]_C &= P_{E\leftarrow D}[T]_{D\leftarrow B}[\vec{v}]_B \\ &= P_{E\leftarrow D}[T(\vec{v})]_D \\ &= [T(\vec{v})]_E \end{aligned}$$

However, $[T]_{E\leftarrow C}[\vec{v}]_C = [T(\vec{v})]_E$ for all $\vec{v} \in \mathcal{V}$ too, so $[T]_{E\leftarrow C} = P_{E\leftarrow D}[T]_{D\leftarrow B}P_{B\leftarrow C}$. ■

Example. *Compute the standard matrix of the* $\overbrace{\textit{transpose map}}^{T}$ *on* $\mathcal{M}_2(\mathbb{C})$ *with respect to the basis*

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right\}.$$ "Pauli basis"

We saw earlier this week that if

$E = \{E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2}\}$ then $[T]_E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

We want $[T]_B = P_{B \leftarrow E} [T]_E P_{E \leftarrow B}$.

Well, $P_{E \leftarrow B} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -i & 0 \\ 0 & 1 & i & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}$, so

$$P_{B \leftarrow E} = P_{E \leftarrow B}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}.$$

$$\therefore [T]_B = P_{B \leftarrow E} [T]_E P_{E \leftarrow B} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -i & 0 \\ 0 & 1 & i & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$