MATH-UA 140 - Linear Algebra

Lecture 27: Linear transformations and their matrix representation

In this course, we have taken matrices as objects of interest, which we have studied in great detail, but only occasionally (least squares, projections) explained how they relate to more general objects in mathematics and the sciences.
The purpose of this lecture is to show that matrices are a natural way to represent a very general family of transformations, called linear transformations.

I) Linear transformations

1) Definition

In the context of this class, a linear transformation is a rule that takes any vector $\vec{u}$ in $\mathbb{R}^n$ and maps it to a vector $\vec{w}$ in $\mathbb{R}^m$, satisfying two linearity conditions:

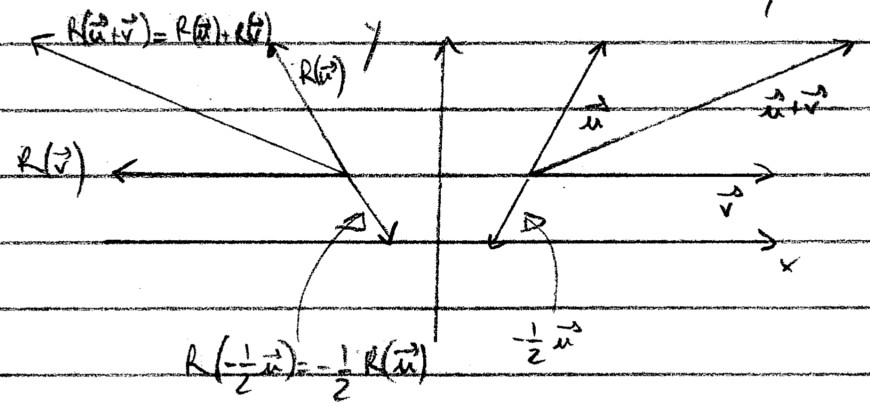
- For all vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$, $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$
- For all vectors $\vec{u}$ in $\mathbb{R}^n$ and all scalars c (in $\mathbb{R}$), $T(c\vec{u}) = c\,T(\vec{u})$

Note that the definition implies that $T(\vec{0}) = \vec{0}$. Indeed, take $c=0$ above, and any $\vec{u}$: $T(0\vec{u}) = 0\,T(\vec{u})$

$$\Leftrightarrow T(\vec{0}) = \vec{0}$$

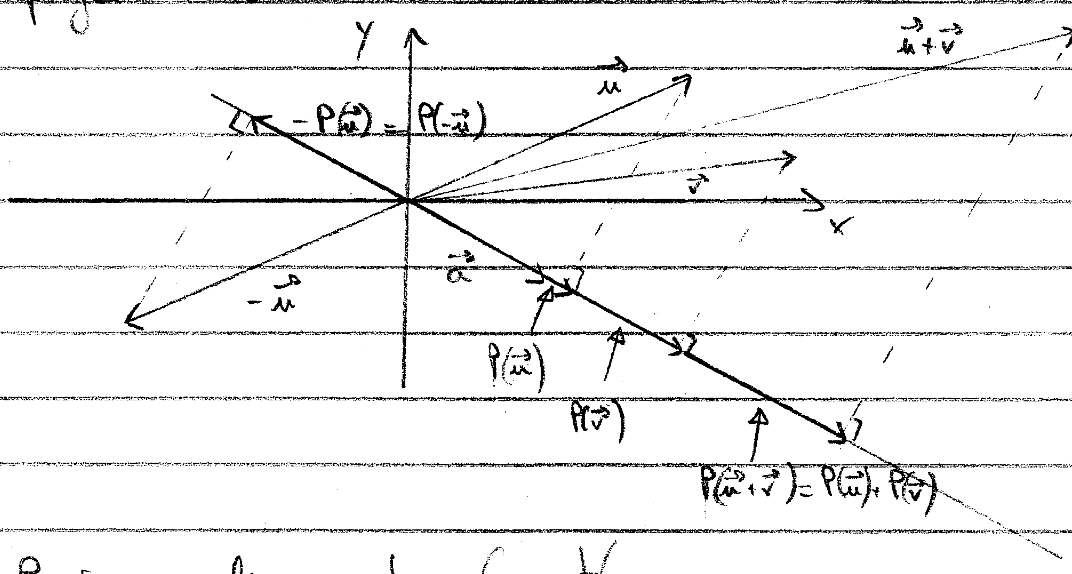
2) Examples

* Consider $\mathbb{R}^n = \mathbb{R}^2$ and $\mathbb{R}^m = \mathbb{R}^2$ in the definition, and the reflection R across the y-axis:



The reflection R is a linear transformation.

* Again, take $\mathbb{R}^n = \mathbb{R}^2$ and $\mathbb{R}^m = \mathbb{R}^2$, and consider the projection P onto the vector $\vec{a}$:



P is a linear transformation

* (Counterexample) The transformation which to any vector $\vec{u}$ adds the constant, nonzero vector $\vec{a}$, $T: \vec{u} \longmapsto \vec{u}+\vec{a}$, is <u>not</u> a linear transformation, because $T(\vec{0}) = \vec{a} \neq \vec{0}$

* (Counterexample) The transformation which takes any vector $\vec{u}$ and maps it to its norm, $T: \vec{u} \longmapsto \|\vec{u}\|$, is <u>not</u> a linear transformation, because in general $T(\vec{u}+\vec{v}) = \|\vec{u}+\vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\| = T(\vec{u}) + T(\vec{v})$

3) <u>Important properties of linearity</u>

* Consider two vectors $\vec{u}$ and $\vec{v}$, and the line going through the tips of each vector: $(1-t)\vec{u} + t\vec{v}$, $t \in \mathbb{R}$

Applying a linear transformation to the whole line, we find,

$$T((1-t)\vec{u} + t\vec{v}) = T((1-t)\vec{u}) + T(t\vec{v})$$
$$= (1-t)\, T(\vec{u}) + t\, T(\vec{v}), \quad t \in \mathbb{R}$$

This is the line going through the tips of $T(\vec{u})$ and $T(\vec{v})$:

<u>A linear transformation transforms a line into a line.</u>

* Consider the middle point between the tips of the vectors $\vec{u}$ and $\vec{v}$, given by the tip of the vector $\vec{w} = \frac{1}{2}\vec{u} + \frac{1}{2}\vec{v}$

$$T(\vec{w}) = T\left(\frac{1}{2}\vec{u} + \frac{1}{2}\vec{v}\right) = \frac{1}{2}T(\vec{u}) + \frac{1}{2}T(\vec{v})$$

$T(\vec{w})$ is the middle point between $T(\vec{u})$ and $T(\vec{v})$.

<u>A linear transformation transforms equispaced points into equispaced</u>

points. (Note that the spacing between the equispaced points may be larger or smaller than the original spacing.)

* Consider the linear combination of n vectors,

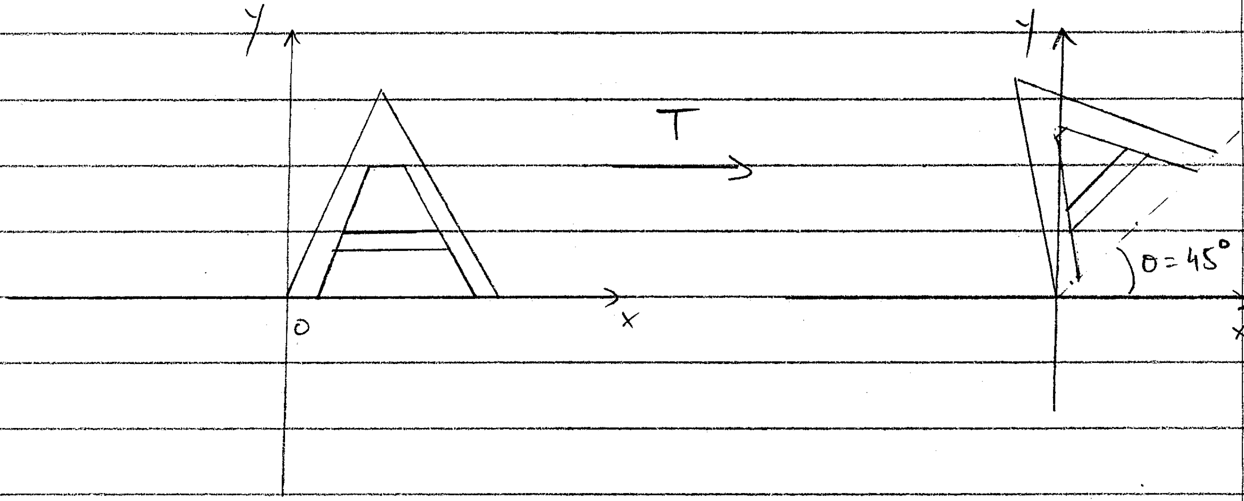
$$\vec{v} = c_1 \vec{u}_1 + c_2 \vec{u}_2 + \dots + c_n \vec{u}_n$$

If T is a linear transformation,

$$T(\vec{v}) = T(c_1 \vec{u}_1 + c_2 \vec{u}_2 + \dots + c_n \vec{u}_n) = T(c_1 \vec{u}_1) + T(c_2 \vec{u}_2) + \dots + T(c_n \vec{u}_n)$$
$$= c_1 T(\vec{u}_1) + c_2 T(\vec{u}_2) + \dots + c_n T(\vec{u}_n)$$

If one wants to understand the action of a linear transformation, it is often best to visualize it's action on an entire object rather than one vector at a time.

Consider for example the linear transformation corresponding to the rotation in $\mathbb{R}^2$, around the origin by an angle $\theta = 45°$. It transformes the Eiffel tower as follows:



II) Using matrices to characterize linear transformations

Geometric descriptions of linear transformations can be very powerful, as in our Eiffel tower case: it is easier to visualize how a 45° rotation acts on an object than to look at the matrix representation of the linear transformation and see its action on the Eiffel tower.

This is however not always the case. It will often be easier to understand a linear transformation by looking at its matrix representation. This matrix representation is also a powerful tool to compute things we want to know regarding transformations

1) Preamble

Consider an $m \times n$ matrix A. The transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by: $T: \vec{u} \in \mathbb{R}^n \longmapsto \vec{w} = A\vec{u}$ in $\mathbb{R}^m$ is a linear transformation.

Indeed, by the rules of matrix vector product,

- For two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$, $A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v}$
- If c is a scalar, $A(c\vec{u}) = cA\vec{u}$

We will now see that any linear transformation has a matrix representation. This simply requires the introduction of a basis and coordinates.

2) Representing linear transformations with matrices

Imagine we want to fully characterize a linear transformation acting on vectors in $\mathbb{R}^n$. Then all we need is to look at how the linear transformation transforms the n vectors of any basis of $\mathbb{R}^n$.

Indeed, consider a basis $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$ of $\mathbb{R}^n$. Any vector $\vec{v}$ in $\mathbb{R}^n$ can be uniquely written as

$$\vec{v} = c_1 \vec{u}_1 + c_2 \vec{u}_2 + \dots + c_n \vec{u}_n$$

where the c_i are the coordinates of $\vec{v}$ in the basis $\vec{u}_1, \dots, \vec{u}_n$. By linearity, we can write the following for the linear transformation T:

$$T(\vec{v}) = c_1 T(\vec{u}_1) + c_2 T(\vec{u}_2) + \dots + c_n T(\vec{u}_n)$$

If we know $T(\vec{u}_1), T(\vec{u}_2), \dots, T(\vec{u}_n)$, we know $T(\vec{v})$ for any $\vec{v}$.

Now, in order to have a matrix representation of T, we also need a basis for $\mathbb{R}^m$, the space T maps to: $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_m$

We can then write
$$T(\vec{u}_1) = a_{11}\vec{w}_1 + a_{21}\vec{w}_2 + \dots + a_{m1}\vec{w}_m$$
$$T(\vec{u}_2) = a_{12}\vec{w}_1 + a_{22}\vec{w}_2 + \dots + a_{m2}\vec{w}_m$$
$$\vdots$$
$$T(\vec{u}_n) = a_{1n}\vec{w}_1 + a_{2n}\vec{w}_2 + \dots + a_{mn}\vec{w}_m$$

If we write $T(\vec{v}) = A\vec{v}$, then

$$\begin{aligned} A\vec{v} = \; & a_{11}c_1\vec{w_1} + a_{21}c_1\vec{w_2} + \cdots + a_{m1}c_1\vec{w_m} \\ & + a_{12}c_2\vec{w_1} + a_{22}c_2\vec{w_2} + \cdots + a_{m2}c_2\vec{w_m} \\ & \vdots \\ & + a_{1n}c_n\vec{w_1} + a_{2n}c_n\vec{w_2} + \cdots + a_{mn}c_n\vec{w_m} \end{aligned}$$

So taking $\vec{w_1}, \ldots, \vec{w_m}$ as a basis for $\mathbb{R}^m$ and $\vec{u_1}, \ldots, \vec{u_n}$ as a basis for $\mathbb{R}^n$, we have

$$A\vec{v} = A\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n \\ \vdots \\ a_{m1}c_1 + a_{m2}c_2 + \cdots + a_{mn}c_n \end{bmatrix}$$

We see that A must then be: $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$

So the rule for finding the matrix representation A of T given a basis $\vec{u_1}, \ldots, \vec{u_n}$ of $\mathbb{R}^n$ and $\vec{w_1}, \ldots, \vec{w_m}$ of $\mathbb{R}^m$ is as follows:

The j^{th} column of A is found by applying T to the j^{th} basis vector $\vec{u_j}$ of $\mathbb{R}^n$.

$T(\vec{u_j}) = a_{1j}\vec{w_1} + \cdots + a_{mj}\vec{w_m}$

Note that if we express T in different bases, A will change, even if T remains the same.

3) Examples

* Consider the linear transformation T corresponding to the projection onto the vector $\vec{v} = \begin{bmatrix} 1 \\ \sqrt{3} \end{bmatrix}$

- We take the standard basis $\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \vec{w}_1$, $\vec{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \vec{w}_2$

$$T(\vec{u}_1) = \frac{1}{2} \frac{\vec{v}}{\|\vec{v}\|} = \begin{bmatrix} \frac{1}{4} \\ \frac{\sqrt{3}}{4} \end{bmatrix} \quad ; \quad T(\vec{u}_2) = \frac{\sqrt{3}}{2} \frac{\vec{v}}{\|\vec{v}\|} = \begin{bmatrix} \frac{\sqrt{3}}{4} \\ \frac{3}{4} \end{bmatrix}$$

Thus, the projection matrix is $P = \begin{bmatrix} \frac{1}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{3}{4} \end{bmatrix}$

(same result as $P = \dfrac{\vec{v}\,\vec{v}^T}{\|\vec{v}\|^2}$)

- Let us now express P in a different basis

$P\vec{v} = \vec{v}$, so $\vec{v}$ is an eigenvector with eigenvalue 1

Let $\vec{t} = \begin{bmatrix} -\sqrt{3} \\ 1 \end{bmatrix}$, constructed to be orthonormal to $\vec{v}$

$P\vec{t} = \vec{0}$, so $\vec{t}$ is an eigenvector with eigenvalue 0.

$\vec{v}$ and $\vec{t}$ are linearly independent, so they are a basis of $\mathbb{R}^2$, the eigenvector basis for P. If we use that basis for both $\vec{u}_1$ and $\vec{u}_2$ and $\vec{w}_1$ and $\vec{w}_2$, the matrix representation of P is:

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

* The counterclockwise rotation of angle θ is a linear transformation. Let us see its matrix representation in the standard basis $\vec{u_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \vec{w_1}$ and $\vec{u_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \vec{w_2}$

$$T(\vec{u_1}) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \quad \text{and} \quad T(\vec{u_2}) = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

so the matrix is $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

as we have seen in Lecture 17.

III) Change of basis

We have said that coordinates and the matrix corresponding to a linear transformation change when we have a change of basis. We now formalize this.

1) Vectors

Imagine we have a vector $\vec{v}$ in a given basis: each entry of the column vector is a coordinate of $\vec{v}$ in that basis.
What would be the column vector $\vec{c}$ corresponding to the vector $\vec{v}$ written in a new basis $\vec{w_1}, \dots, \vec{w_n}$?
The entries c_i in $\vec{c}$ are such that

$$\vec{v} = c_1 \vec{w_1} + c_2 \vec{w_2} + \dots + c_n \vec{w_n}$$

In other words, $\vec{v} = W\vec{c}$, where W is the matrix whose columns are the vectors of the new basis.

Example: Consider the basis $\vec{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\vec{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and the vector $\vec{v}$ with coordinates $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ in that basis.

What is the vector representation $\vec{w}$ of $\vec{v}$ in the standard basis $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$?

$$\vec{w} = \underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 3 & 1 \end{bmatrix}}_{W} \underbrace{\begin{bmatrix} 2 \\ -1 \end{bmatrix}}_{\vec{c}} = \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix}$$

2) Matrices

Imagine now that we have a matrix A corresponding to a linear transformation T written in a basis $(\vec{u}_1, \ldots, \vec{u}_n)$ of $\mathbb{R}^n$. What is the matrix B for T written in a different basis $(\vec{w}_1, \ldots, \vec{w}_n)$?

We know how T acts on any vector written in $(\vec{u}_1, \ldots, \vec{u}_n)$:
$T(\vec{a}) = A\vec{a}$
So if we have a vector $\vec{v}$ written in the basis $(\vec{w}_1, \ldots, \vec{w}_n)$, the idea is to first write it as $\vec{a}$ in the old basis $(\vec{u}_1, \ldots, \vec{u}_n)$:

$W\vec{v} = \vec{a}$, where $W = [\vec{w}_1 \cdots \vec{w}_n]$

Then, we know that $T(\vec{a}) = A\vec{a}$
So, when the output is written in the $(\vec{u}_1, \ldots, \vec{u}_n)$, and the input $\vec{v}$ is written in the basis $(\vec{w}_1, \ldots, \vec{w}_n)$, we can write

$$T(\vec{v}) = AW\vec{v}$$

But what we really want is the output written in the basis $W = (\vec{w}_1, \ldots, \vec{w}_n)$ as well. To obtain, we need to apply W^{-1} to the vector $W^{-1}AW\vec{v}$.
We conclude that the matrix B for T written in the basis $(\vec{w}_1, \ldots, \vec{w}_n)$ is given by:

$$B = W^{-1}AW$$

B and A are similar matrices.