Lecture 30: Introduction to Linear Transformations

Winfried Just
Department of Mathematics, Ohio University

MATH3200: Applied Linear Algebra

Why should we learn about linear transformations?

Linear transformations are important and useful:

- A lot of applications of linear algebra involve linear transformations.
- Linear algebra is much easier to understand when one looks at it through the lens of linear transformations.
- Linear transformations are not hard to understand when one thinks of them in terms of concrete examples.
- *But:* A completely formal treatment of linear transformations can easily become very dry and abstract.

Here we will develop the theory of linear transformations only as far as it directly relates to the remainder of this course and omit its more abstract aspects.

In particular, we will only study linear transformations that are functions $L : \mathbb{R}^n \to \mathbb{R}^m$ with certain properties.

Products of matrices and column vectors

Let $\mathbf{A}$ be an $m \times n$ matrix, and let $\vec{\mathbf{v}}$ be an $n \times 1$ column vector. Then $\mathbf{A}\vec{\mathbf{v}}$ is an $m \times 1$ column vector:

$$\mathbf{A}\vec{\mathbf{v}} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix} = \vec{\mathbf{w}}$$

When $\mathbf{A}$ is fixed, we can think of this operations as taking a vector $\vec{\mathbf{v}}$ in $\mathbb{R}^n$ as *input* and *transforming* it into and *output* that is a vector $\vec{\mathbf{w}} = \mathbf{A}\vec{\mathbf{v}}$ in $\mathbb{R}^m$.

In this sense, the above matrix multiplication defines a function, or map $L_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^m$ that is called a *linear transformation*.

It's output values $\vec{\mathbf{w}} = L_{\mathbf{A}}(\vec{\mathbf{v}})$ are given by $L_{\mathbf{A}}(\vec{\mathbf{v}}) = \mathbf{A}\vec{\mathbf{v}}$.

When $\mathbf{A} = [a]$, then $L_{\mathbf{A}} : \mathbb{R}^1 \to \mathbb{R}^1$ is simply the familiar linear function that sends x to $L_{\mathbf{A}}(x) = ax$.

Example 1: $L_{\mathbf{A}}$ for a 3×2 matrix

For a fixed $m \times n$ matrix $\mathbf{A}$, the product $\mathbf{A}\vec{\mathbf{v}}$ defines a function, or map $L_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^m$ that is called a *linear transformation*.

For example, let $\mathbf{A} = \begin{bmatrix} 2 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$

Question L30.1: From where to where does the transformation $L_{\mathbf{A}}$ go?

Here $L_{\mathbf{A}} : \mathbb{R}^2 \to \mathbb{R}^3$.

The space $\mathbb{R}^2$ is the *domain* of the function $L_{\mathbf{A}}$;

The space $\mathbb{R}^3$ is the *codomain* of the function $L_{\mathbf{A}}$.

The function values in our example

Let $\mathbf{A} = \begin{bmatrix} 2 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$

Question L30.2: What is $\vec{\mathbf{w}} = L_{\mathbf{A}}\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right)$?

Here $L_{\mathbf{A}}\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right) = \mathbf{A}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

Question L30.3: What is $\vec{\mathbf{w}} = L_{\mathbf{A}}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$?

Here $L_{\mathbf{A}}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \mathbf{A}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + 2y \\ 3x + 4y \\ 5x + 6y \end{bmatrix}$

Geometric interpretation of $L_{\mathbf{A}} : \mathbb{R}^2 \to \mathbb{R}^2$

Let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ be a 2×2 matrix and let $\vec{\mathbf{v}} = \begin{bmatrix} x \\ y \end{bmatrix}$ be in $\mathbb{R}^2$.

Consider $\quad L_{\mathbf{A}}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix}$

The coordinates of the function value $L_{\mathbf{A}}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$ are

$x' = a_{11}x + a_{12}y, \qquad y' = a_{21}x + a_{22}y.$

To build up some intuition about a given transformation $L_{\mathbf{A}}$, we can mark the coordinates x, y on a sheet of material and see what kind of manipulation of the sheet would send them to the coordinates x', y' of $L_{\mathbf{A}}(\vec{\mathbf{v}})$.

Example 2: A linear transformation $L : \mathbb{R}^2 \to \mathbb{R}^2$

Let α be an angle and consider the matrix

$$\mathbf{R}_\alpha \vec{\mathbf{v}} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

We will write shorthand L_α for $L_{\mathbf{R}_\alpha}$. Then

$$L_\alpha\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \mathbf{R}_\alpha \vec{\mathbf{v}} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\alpha - y\sin\alpha \\ x\sin\alpha + y\cos\alpha \end{bmatrix}$$

The transformation $L_\alpha(\vec{\mathbf{v}}) = \mathbf{R}_\alpha \vec{\mathbf{v}}$ represents *rotation by an angle α*.

Note that we could rotate a sheet of paper without crumbling it, tearing it apart. This means that this linear transformation *preserves distances between points* and hence *preserves or areas of regions.*

Moreover, we don't need to flip the paper over to perform this transformation. We say that it *preserves orientation.*

Example 3: Another linear transformation $L : \mathbb{R}^2 \to \mathbb{R}^2$

Let $\mathbf{A} = \begin{bmatrix} 3 & 0 \\ 0 & 0.5 \end{bmatrix}$ and consider the linear transformation $L_{\mathbf{A}}$:

$$L_{\mathbf{A}}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \mathbf{A}\vec{\mathbf{v}} = \begin{bmatrix} 3 & 0 \\ 0 & 0.5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x \\ 0.5y \end{bmatrix}$$

This transformation corresponds to a threefold *stretch* in the horizontal (x-) direction and a twofold *compression* in the vertical (y-) direction.

Note that this transformation $L_{\mathbf{A}}$ *changes distances between points* and cannot be performed by manipulating a sheet of paper without crumbling it or tearing it apart.

However, this transformation $L_{\mathbf{A}}$ can be performed by manipulating a sheet of elastic material that lies flat on a surface. It *preserves orientation.*

Example 4: One more linear transformation $L : \mathbb{R}^2 \to \mathbb{R}^2$

Let $\mathbf{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and consider the linear transformation $L_{\mathbf{B}}$:

$$L_{\mathbf{B}}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \mathbf{B}\vec{\mathbf{v}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$$

This transformation corresponds to flipping over a sheet of material along the diagonal $x = y$ so that the roles of the x- and y-coordinates will be switched.

This transformation $L_{\mathbf{B}}$ be performed by manipulating a sheet of paper without crumbling it or tearing it apart. It *preserves* distances between points and areas.

However, this transformation $L_{\mathbf{B}}$ cannot be performed by manipulating a sheet of paper that lies flat on a surface. We would have to lift it and flip it over, so that we would end up looking at the other side. We say that this transformation *changes orientation.*

Products of matrices and column vectors: Properties

Let $\mathbf{A}$ be a matrix of order $m \times n$ and let $\vec{\mathbf{v}}$ be an $n \times 1$ column vector.

$$\mathbf{A}\vec{\mathbf{v}} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix}$$

Then $\mathbf{A}\vec{\mathbf{v}}$ is an $m \times 1$ column vector.

This defines a function $L_{\mathbf{A}} : \mathbb{R}^n \to \mathbb{R}^m$ by $L(\vec{\mathbf{v}}) = \mathbf{A}\vec{\mathbf{v}}$.

By general properties of matrix multiplication:

(i) $L_{\mathbf{A}}(\lambda\vec{\mathbf{v}}) = \mathbf{A}(\lambda\vec{\mathbf{v}}) = \lambda\mathbf{A}\vec{\mathbf{v}} = \lambda L_{\mathbf{A}}(\vec{\mathbf{v}})$

(ii) $L_{\mathbf{A}}(\vec{\mathbf{v}} + \vec{\mathbf{w}}) = \mathbf{A}(\vec{\mathbf{v}} + \vec{\mathbf{w}}) = \mathbf{A}\vec{\mathbf{v}} + \mathbf{A}\vec{\mathbf{w}} = L_{\mathbf{A}}(\vec{\mathbf{v}}) + L_{\mathbf{A}}(\vec{\mathbf{w}})$.

These two properties assure us that the function $L_{\mathbf{A}}$ is a **linear transformation.**

Linear transformations: Our definition

Definition (For the purpose of this course)

Let n, m be positive integers and let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a function. Then L is called a *linear transformation* if it satisfies both of the following conditions for all vectors $\vec{\mathbf{v}}, \vec{\mathbf{w}}$ in $\mathbb{R}^n$ and all scalars λ in $\mathbb{R}$:

(i) $L(\lambda\vec{\mathbf{v}}) = \lambda L(\vec{\mathbf{v}})$

(ii) $L(\vec{\mathbf{v}} + \vec{\mathbf{w}}) = L(\vec{\mathbf{v}}) + L(\vec{\mathbf{w}})$.

These two simple properties have many important consequences. They apply to a much broader class of functions $L : V \to W$, where V, W are *abstract vector spaces* and the scalars are allowed to be also complex numbers or members of any *algebraic field.*

This more general theory is rather abstract and gives linear transformations a reputation for being a difficult concept.

Our definition will do here, but you should know that it is just a special case and everything works just the same way if, for example, the scalars are complex numbers.

Take-home message

Linear transformations are important tools in linear algebra.

The general theory is a bit abstract, here we confine ourselves to the case of functions $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$, where n, m are given positive integers. Such an L is called a *linear transformation* if it satisfies both of the following conditions for all vectors $\vec{\mathbf{v}}, \vec{\mathbf{w}}$ in $\mathbb{R}^n$ and all scalars λ in $\mathbb{R}$:

(i) $L(\lambda\vec{\mathbf{v}}) = \lambda L(\vec{\mathbf{v}})$

(ii) $L(\vec{\mathbf{v}} + \vec{\mathbf{w}}) = L(\vec{\mathbf{v}}) + L(\vec{\mathbf{w}})$.

When $\mathbf{A}$ is an $m \times n$ matrix, then the function $L_{\mathbf{A}} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ that is defined by $L(\vec{\mathbf{v}}) = \mathbf{A}\vec{\mathbf{v}}$ is a linear transformation.

Rotations are examples of linear transformations from $\mathbb{R}^2$ into $\mathbb{R}^2$ that *preserve distances and areas.*

Rotations also *preserve orientation*, while some other linear transformations from $\mathbb{R}^2$ into $\mathbb{R}^2$ *change orientation* and cannot be performed on a sheet of material that lies flat on a surface.