Introduction to Linear Transformations

A. Havens

Department of Mathematics
University of Massachusetts, Amherst

February 9, 2018

Outline

The Map $\mathbf{x} \mapsto \mathrm{A}\mathbf{x}$

Let A be an $m \times n$ matrix. We can view the assignment, to any $\mathbf{x} \in \mathbb{R}^n$, of the matrix-vector product

$$\mathbf{x} \mapsto \mathrm{A}\mathbf{x}$$

as a function which takes vectors $\mathbf{x} \in \mathbb{R}^n$ to vectors in $\mathbb{R}^m$.

Thus there is a *transformation*

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

such that any $T(\mathbf{x}) = \mathrm{A}\mathbf{x} \in \mathbb{R}^m$ is the linear combination of the columns of A using the entries of $\mathbf{x}$ as weights.

Functions, Transformations, Maps

Our perspective in this section is that we can regard this assignment $\mathbf{x} \mapsto T(\mathbf{x}) = \mathrm{A}\mathbf{x}$ as a *function*, sending $\mathbb{R}^n$ to some subset of $\mathbb{R}^m$.

We wish to study properties of such maps, their *images*, and various geometric effects of such transformations.

The terms "transformation", "map" and "function" are used interchangeably. Let us give a precise statement of what is meant by this language.

What is Meant by "Transformation"?

Definition

A *function*, also called a *map* or *transformation* of a set X with values in a set Y is an assignment of an element of Y to each element $x \in X$.

Often, one chooses a letter such as f or T to label the function.

The set containing the "inputs," X, is called the *domain* of the function, and the set Y containing possible outputs is called the *codomain*.

To completely specify a function, one must give the domain, codomain, and an assignment rule.

Specifying a Transformation

Definition

This is often notated as

$$\begin{aligned} f : X &\to Y \\ x &\mapsto y = f(x) \end{aligned}$$

where $f(x)$ would be an explicit rule giving an element of Y.

Note that for a given $x \in X$, a function assigns only one element $y \in Y$.

Our Primary Example

Example

The primary example of our concern is the case where $X = \mathbb{R}^n$ for some $n \geq 1$, $Y = \mathbb{R}^m$ for some $m \geq 1$, and the function is a transformation defined by mapping a vector $\mathbf{x}$ to its matrix vector product with some $m \times n$ matrix A.

In the notation of our definition, we specify this transformation of $\mathbb{R}^n$ by writing

$$\begin{aligned} T : \mathbb{R}^n &\to \mathbb{R}^m \\ \mathbf{x} &\mapsto \mathbf{y} = \mathrm{A}\mathbf{x} \end{aligned}$$

Brief Anatomy of a Matrix Map

Example

Observe that the domain is $\mathbb{R}^n$, and the codomain is $\mathbb{R}^m$.

The assignment rule $\mathbf{y} = T(\mathbf{x})$ in this case is given by $\mathbf{y} = A\mathbf{x}$. Sometimes, one simply describes the rule by writing $T(\mathbf{x}) = A\mathbf{x}$.

Given an explicit matrix, we obtain an explicit example of a transformation. We will see a number of explicit examples shortly, and we'll examine their effects geometrically in each case.

Domains and Images

A useful perspective is that a transformation specifies how to *transform* or *map* a set, the domain, to some subset of the codomain.

One then studies transformations in part by understanding what happens to a given point in the domain. For transformations of spaces, this often entails studying geometry (and sometimes topology).

Geometry/Topology

When the domain and codomain are the same space (e.g. $\mathbb{R}^n$), the geometry is sometimes captured by considering how points are "pushed" or moved according to the map, and studying distances, angles, coordinates, and other related measures of "where things go" in relation to each other.

When the codomain is different from the domain, it can be helpful to imagine that points are moved from the domain space to the codomain space, with points possibly being collapsed or glued together. For "discontinuous" maps, the domain may be cut or pulled apart and then stitched or scattered onto some points or regions of the codomain.

Geometry/Topology

Topology concerns intrinsic properties of spaces and maps necessary to discuss continuity. Geometry concerns the intrinsic and extrinsic, metric properties used to study position, size, and spatial relations.

The functions we study in this course (linear maps) have certain rigidity properties forcing them to be continuous; they act on the domains without breaking them apart or scattering relatively nearby pieces to relatively faraway ends of the codomain.

For linear transformations any topological considerations, though simple, will be largely ignored in our study. The geometry on the other hand is illuminating, and so we will make use of geometric notions as needed.

A Mapmaker's Function

A useful visual, motivating the terminology "map", is to think about making a cartographical map of a piece of the earth.

The domain is the set of points $\mathscr{R}$ on the earth (the "region") that are in the area over which the cartographer is constructing a map.

The codomain is the paper or vellum $\mathscr{V}$ on which the cartographer draws the map.

A Mapmaker's Function

Let m be the rule specifying how a point of $\mathscr{V}$ corresponds to a point of $\mathscr{R}$. For a given point $p \in \mathscr{R}$ on the globe, the corresponding point $m(p) \in \mathscr{V}$ is called *the image of p*.

Thus, the cartographer is tasked with constructing a function $m : \mathscr{R} \to \mathscr{V}$ that captures the curved region $\mathscr{R}$ of earth as a flat "image" on $\mathscr{V}$. (A deep result, Gauss's Theorema Egregium, states that the curvature of earth forces any such image to be distorted).

It's Not Even a Metaphor

The set of all points of $\mathscr{V}$ which correspond to points $p \in \mathscr{R}$ is called variously *the image of* $\mathscr{R}$ *under* m, *the image of the map* $m : \mathscr{R} \to \mathscr{V}$, or the *the range of the function* m. This "total image" or range is frequently denoted $m(\mathscr{R})$.

A Remark on Notation and Terminology

The letters themselves in the above example are of course irrelevant and may be swapped out for any other collection denoting sets and a function rule.

I remark here that the textbook we are using reserves the term "image" only for the image of a point, and uses "range" for the set of all images. It is not uncommon to hear mathematicians use the term image more broadly as I have in the previous slide.

A Remark on Notation and Terminology

There is something evocative about considering "the image of a map" or "the image of a transformation" while envisioning how a transformation rule bends and stretches a piece of the globe onto a flat page, or how a digital photo is rotated in image manipulation software, or how a portion of space is molded and contorted into a piece of another space.

For this reason, I prefer the term image to range (range suggests one-dimensional images to me, as a collection of real numbers arising from the functions studied in single variable calculus).

For a function $f : X \to Y$, I'll write $y = f(x)$ for the image of a point x, and $f(X) = \mathrm{Image}(f)$ for the image of the whole domain X by the function f. What is meant will always be clear from context.

Linear Endomorphisms of $\mathbb{R}^2$

The first examples we'll consider are maps of the plane constructed via matrices. A fancy term for such maps is *linear endomorphisms of* $\mathbb{R}^2$, but we can also just call them *linear transformations of the plane.*

Since we want maps $T : \mathbb{R}^2 \to \mathbb{R}^2$, we need to send a 2-vector to a 2-vector. What size matrix should we use?

Since an $m \times n$ matrix A sends an n-vector to an m vector, we need both m and n equal to two.

Linear Endomorphisms of $\mathbb{R}^2$

A general linear endomorphism of $\mathbb{R}^2$ can thus be described by a map $\mathbf{x} \mapsto T(\mathbf{x}) = \mathrm{A}\mathbf{x}$ for some 2×2 matrix A.
We can write

$$T(\mathbf{x}) = \mathrm{A}\mathbf{x} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}.$$

2 Easy Examples

Example

Consider the matrices

$$\mathrm{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \mathrm{R} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

What are their respective actions on the plane?

2 Easy Examples

Example

The map $\mathbf{x} \mapsto \mathrm{I}_2\mathbf{x}$ is the "identity map" taking the vector $\mathbf{x}$ to itself:

$$\mathrm{I}_2\mathbf{x} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{x}\,.$$

What of the other map, $\mathbf{x} \mapsto \mathrm{R}\mathbf{x}$?

2 Easy Examples

Example

The map $\mathbf{x} \mapsto \mathrm{R}\mathbf{x}$ rotates the vector $\mathbf{x}$ by an angle of $\pi/2$ radians counterclockwise:

$$\mathrm{R}\mathbf{x} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix}.$$

We can easily confirm that $\mathbf{x} \perp \mathrm{R}\mathbf{x}$: $\mathbf{x} \cdot \mathrm{R}\mathbf{x} = x(-y) + (y)(x) = 0$.

Rotations of $\mathbb{R}^2$

Example

The latter matrix is a special case of general rotation matrices.

Fix an angle θ, measured counterclockwise from the x axis.

Then the matrix R_θ given by

$$\mathrm{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

acts on $\mathbf{x} \in \mathbb{R}^2$ by rotating it through an angle of θ counterclockwise.

Projections in $\mathbb{R}^2$

Example

Another kind of map is a *projection* onto a line. For example, the map

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \mathrm{R}\mathbf{x} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}$$

projects the vector $\mathbf{x}$ onto the x-axis, leaving only its x component.

What is the image of the projection? What subset of $\mathbb{R}^2$ is sent to $\mathbf{0}$?

Projections in $\mathbb{R}^2$

Example

A useful exercise in vector geometry is to convince yourself of the following general formula for projections. Fix a vector $\mathbf{u}$. the projection onto the line spanned by $\mathbf{u}$ is

$$\mathbf{x} \mapsto \frac{\mathbf{x}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}\,.$$

Another good exercise: represent this map as a matrix, and show that the result depends only on $\mathrm{Span}\,\{\mathbf{u}\}$, in the sense that projecting onto any other nonzero collinear vector yields the same map.

We will eventually encounter a theorem that lets us systematically compute a matrix representing a linear map.

Reflections in $\mathbb{R}^2$

Example

Try to describe a reflection through the x-axis. Can you write down a matrix which accomplishes this?

Reflections in $\mathbb{R}^2$

Example

Try to describe a reflection through the x-axis. Can you write down a matrix which accomplishes this?

The map

$$\mathrm{M}\mathbf{x} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}$$

is a reflection through the x-axis.

Reflections in $\mathbb{R}^2$

Example

Given a vector $\mathbf{v} \in \mathbb{R}^2$, can you describe a general reflection through the line $\mathrm{Span}\{\mathbf{v}\}$?

Using projections onto a line, you can build reflections (perhaps after drawing the right picture).

You should be able to write down a formula for the reflection of $\mathbf{x}$ through $\mathrm{Span}\{\mathbf{v}\}$ using projections and dot products in terms of $\mathbf{v}$, and also a matrix in terms of the components of $\mathbf{v}$.

Shear Transforms

Example

Consider the matrix

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

What are the images of the vectors

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 \\ 1 \end{bmatrix}?$$

This matrix *shears* the plane along the x-axis.

Dilations and Contractions

Example

Let $s \in \mathbb{R}$ be a positive scalar. A transformation

$$\mathbf{x} \mapsto s\mathbf{x} = \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} \mathbf{x}$$

is called a dilation if $s > 1$ and a contraction if $0 < s < 1$. Correspondingly, it either dilates (expands) or contracts (shrinks) areas of the plane, respectively.

Similarity transformations

Some linear maps of the plane are a mixture of the above examples.

Maps which decompose as a collection of reflections, rotations, and dilations are a subset of the *similarity transforms* of the plane. These preserve the Euclidean shapes and angles, but not sizes, of shapes in the plane.

However, there are nonlinear similarity transforms: translations. A translation shifts the zero vector's location, but a map $\mathbf{x} \mapsto \mathrm{A}\mathbf{x}$ will always send $\mathbf{0}$ to itself.

A transformation of the form $\mathbf{x} \mapsto \mathrm{A}\mathbf{x} + \mathbf{b}$ for a matrix A and a constant vector $\mathbf{b}$ are called *affine transformations*, and are a natural generalization of linear transformations.

Similarity transformations

Example

Consider the map

$$\mathbf{x} \mapsto \begin{bmatrix} y - x \\ x + y \end{bmatrix}.$$

What is the matrix of this map? What does this map do?

I claim that this is a similarity transformation which admits a decomposition involving a rotation, a dilation, and a reflection!

Similarity transformations

Example

Consider the map

$$\mathbf{x} \mapsto \begin{bmatrix} y - x \\ x + y \end{bmatrix}.$$

What is the matrix of this map? What does this map do?

I claim that this is a similarity transformation which admits a decomposition involving a rotation, a dilation, and a reflection!

You can check that it is the result of first rotating by $\pi/4$ counterclockwise, then dilating by $\sqrt{2}$, and then reflecting through the y-axis. (Hint: consider the images of the corners of the unit square with vertices $(0,0)$, $(1,0)$, $(1,1)$, and $(0,1)$.)

Endomorphisms of $\mathbb{R}^3$

Just as one can consider maps from $\mathbb{R}^2$ to itself, one can consider maps from $\mathbb{R}^3$ to itself.

We'll look at just a couple of linear endomorphisms of $\mathbb{R}^3$.

Projection to a plane in $\mathbb{R}^3$

The map

$$P(\mathbf{x}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

projects $\mathbb{R}^3$ onto the plane $z = 0$, commonly called the xy-plane. Essentially the same information is contained in the map

$$T : \mathbb{R}^3 \to \mathbb{R}^2$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix},$$

but the latter map has codomain $\mathbb{R}^2$.

Reflection Through a Line

Consider the map

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} -x \\ -y \\ z \end{bmatrix}.$$

This is a 3-dimensional reflection through the line $x = 0 = y$, i.e., a reflection through the z axis.

Can you give a matrix description for this map? Can you describe a general reflection through a line? What about a reflection through a plane?

The Many Maps I Wish We Could Examine...

Here's a few challenge problems worth exploring:

- What is the matrix of a projection onto the plane $ax + by + cz = 0$?
- Can you give a general matrix describing reflection through a plane $ax + by + cz = 0$?
- Find a matrix describing a spatial rotation by an angle θ counterclockwise around a vector $\mathbf{v}$.
- Show that any rotation (of either $\mathbb{R}^2$ or $\mathbb{R}^3$) can be written as a composition of reflections.
- What is the general form of a similarity transformation of $\mathbb{R}^2$ that fixes the origin? Can you describe a general form of a similarity transformation of $\mathbb{R}^3$?

$\mathbf{x} \mapsto A\mathbf{x}$ is Linear

What makes the above examples *linear?*

For one, we can check that they either *send lines to lines* or *crush lines to* $\mathbf{0}$.

More generally, the maps we've called linear *send linear combinations of vectors to a linear combination of the images of those vectors, with the same weights.* In a sense, this is the algebraic way to understand what it means to *preserve a linear structure*.

This happens for our examples because of the general properties of matrix-vector multiplication that we encountered before.

In particular, recall the proposition:

Proposition

Let $\mathbf{u}$ and $\mathbf{v}$ be arbitrary vectors in $\mathbb{R}^n$, let $s \in \mathbb{R}$ be any real scalar, and let A be any $m \times n$ matrix. Then the matrix vector product satisfies the following two properties:

(i.) $\mathrm{A}(\mathbf{u}+\mathbf{v}) = \mathrm{A}\mathbf{u} + \mathrm{A}\mathbf{v}$,

(ii.) $\mathrm{A}(s\mathbf{u}) = s(\mathrm{A}\mathbf{u})$.

It follows from this that $\mathrm{A}(s\mathbf{x}+t\mathbf{y}) = s(\mathrm{A}\mathbf{x}) + t(\mathrm{A}\mathbf{y})$ for any $m \times n$ matrix A, vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and scalars $s, t \in \mathbb{R}$. This generalizes to the following fact:

The image of a span of vectors will be the span of the images.

Definitions of Linear Transformations $T : \mathbb{R}^n \to \mathbb{R}^m$

Definition

A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a *linear transformation*, or a *linear map* if and only if for any vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and any scalar $s \in \mathbb{R}$ the following two properties hold:

(i.) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$,

(ii.) $T(s\mathbf{u}) = sT(\mathbf{u})$.

Remark

It follows from this definition that for a linear map $T : \mathbb{R}^n \to \mathbb{R}^m$,

$$T(c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k) = c_1 T(\mathbf{v}_1) + \ldots + c_k T(\mathbf{v}_k)\,,$$

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

Showing a Map is Linear: an Example

Example

Fix two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, and consider the map

$$T(\mathbf{x}) = (\mathbf{u} \cdot \mathbf{x})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{x}\,.$$

Use the definition of a linear transformation to show that T is linear.

The key is that we can check both properties (i.) and (ii.) by confirming that $T(a\mathbf{x} + b\mathbf{y}) = aT(\mathbf{x}) + bT(\mathbf{y})$ for any vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and any scalars $s, t \in \mathbb{R}$. That is, we test that the image of a linear combination is a linear combination of images.

Testing Linearity

Let $a, b \in \mathbb{R}$ be arbitrary scalars, and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ arbitrary vectors. Then

$$\begin{aligned} T(a\mathbf{x} + b\mathbf{y}) &= \big(\mathbf{u} \cdot (a\mathbf{x} + b\mathbf{y})\big)\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})(a\mathbf{x} + b\mathbf{y}) \\ &= \big(\mathbf{u} \cdot (a\mathbf{x}) + \mathbf{u} \cdot (b\mathbf{y})\big)\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})(a\mathbf{x}) - (\mathbf{u} \cdot \mathbf{v})(b\mathbf{y}) \\ &= a(\mathbf{u} \cdot \mathbf{x})\mathbf{v} + b(\mathbf{u} \cdot \mathbf{y})\mathbf{v} - a(\mathbf{u} \cdot \mathbf{v})(\mathbf{x}) - b(\mathbf{u} \cdot \mathbf{v})(\mathbf{y}) \\ &= a(\mathbf{u} \cdot \mathbf{x})\mathbf{v} - a(\mathbf{u} \cdot \mathbf{v})(\mathbf{x}) + b(\mathbf{u} \cdot \mathbf{y})\mathbf{v} - b(\mathbf{u} \cdot \mathbf{v})\mathbf{y} \\ &= a\big((\mathbf{u} \cdot \mathbf{x})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{x}\big) + b\big((\mathbf{u} \cdot \mathbf{y})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{y}\big) \\ &= aT(\mathbf{x}) + bT(\mathbf{y})\,. \end{aligned}$$

A Non-Linear Example

We can easily check that affine transformations $\mathbf{x} \mapsto \mathrm{A}\mathbf{x} + \mathbf{b}$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ are *not* linear.

For a fixed $m \times n$ matrix A and a *nonzero* vector $\mathbf{b} \in \mathbb{R}^m$, let $T(\mathbf{x}) = \mathrm{A}\mathbf{x} + \mathbf{b}$, and consider $T(s\mathbf{x})$ for any scalar $s \in \mathbb{R}$.

$$T(s\mathbf{x}) = \mathrm{A}(s\mathbf{x}) + \mathbf{b} = s\mathrm{A}\mathbf{x} + \mathbf{b} \neq s(\mathrm{A}\mathbf{x} + \mathbf{b}).$$

This proves our claim.

Representing Linear Maps

The reason we've focused so much on matrices in our examples is that *any linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be realized by some matrix-vector product formula*!

We'll encounter a precise way to compute the matrix of a linear map if we know its effect on certain special vectors, called *the standard basis vectors of $\mathbb{R}^n$*.

The standard basis vectors are just the n different n-vectors corresponding to the points a unit distance from the origin along right-angled coordinate axes. They're called a basis because they are *a linearly independent set that span the whole space* (in this case $\mathbb{R}^n$.)

Why We Must Study Matrix Algebra

A matter remains: we've not discussed how to efficiently compose linear maps.

We saw already that some maps can be constructed by applying separate linear transformations in succession (e.g. similarity transformations). We would like to know how to represent a map that results from applying several matrices in succession.

This leads naturally to the notion of *matrix products*. We will also be able to regard spaces of matrices as being analogous to spaces of vectors, with rules for scaling and addition. This perspective will enrich our study of linear transformations.

Matrices Everywhere

More generally, when we study other vector spaces (like spaces of polynomials), it will turn out that there are ways to choose *linear coordinates*, after which we can *still* represent linear maps by matrices (though the entries won't always be real numbers, if e.g. we are studying complex vector spaces, or vector spaces over finite fields).

Thus, while the definition of linear maps mirrors the properties of matrix-vector products, we can imagine most linear maps as being conveniently represented by matrices. The exceptions appear in the study of infinite dimensional vector spaces, like spaces of functions, where linear maps can be more complex, like derivative and integral operators.

Homework

- I recommend reading sections 1.7, and 1.8 for Monday 2/12, 1.9 by Wednesday 2/14 (if not by Monday), and 2.1 by Friday 2/16.
- The MyMathLab assignment on 1.7 (linear independence) is due 2/13, and 1.8 (linear transformations) is due 2/15.
- The first exam is coming up! Our section, math 235-04, meets in Hasbrouck Laboratory Addition (HASA) 124 on Tuesday night, February 27th, 7 - 9 pm.
- The first exam covers the material of sections 1.1, 1.2, 1.3, 1.4, 1.5, 1.7, 1.8, 1.9, and 2.1 in the text.