Matrix Calculations: Linear maps, bases, and matrices

A. Kissinger

Institute for Computing and Information Sciences Radboud University Nijmegen

Version: autumn 2018



Outline

Linear maps

Basis of a vector space

From linear maps to matrices



From last time

- Vector spaces V, W,... are special kinds of sets whose elements are called *vectors*.
- Vectors can be added together, or multiplied by a real number, For *v*, *w* ∈ *V*, *a* ∈ ℝ:

$$\mathbf{v} + \mathbf{w} \in V$$
 $a \cdot \mathbf{v} \in V$

• The simplest examples are:

$$\mathbb{R}^n := \{(a_1, \ldots, a_n) \mid a_i \in \mathbb{R}\}$$

Maps between vector spaces

We can send vectors $\mathbf{v} \in V$ in one vector space to other vectors $\mathbf{w} \in W$ in another (or possibly the same) vector space?

V, W are vector spaces, so they are sets with extra stuff (namely: +, ·, 0).

A common theme in mathematics: study functions $f : V \rightarrow W$ which preserve the extra stuff.



Functions

- A function f is an operation that sends elements of one set X to another set Y.
 - in that case we write $f: X \to Y$ or sometimes $X \stackrel{f}{\to} Y$
 - this f sends $x \in X$ to $f(x) \in Y$
 - X is called the domain and Y the codomain of the function f
- Example. f(n) = 1/(n+1) can be seen as function N → Q, that is from the natural numbers N to the rational numbers Q
- On each set X there is the identity function id: X → X that does nothing: id(x) = x.
- Also one can compose 2 functions $X \xrightarrow{f} Y \xrightarrow{g} Z$ to a function:

$$g \circ f \colon X \longrightarrow Z$$
 given by $(g \circ f)(x) = g(f(x))$

Linear maps

A linear map is a function that preserves the extra stuff in a vector space:

Definition

Let V, W be two vector spaces, and $f: V \to W$ a map between them; f is called linear if it preserves both:

• addition: for all
$$\mathbf{v}, \mathbf{v}' \in V$$
,

$$f(\underbrace{\mathbf{v}+\mathbf{v}'}_{\text{in }V}) = \underbrace{f(\mathbf{v})+f(\mathbf{v}')}_{\text{in }W}$$

• scalar multiplication: for each $\mathbf{v} \in V$ and $\mathbf{a} \in \mathbb{R}$,

$$f(\underbrace{a \cdot v}_{\text{in } V}) = \underbrace{a \cdot f(v)}_{\text{in } W}$$

Linear maps

Basis of a vector space From linear maps to matrices

Radboud University Nijmegen



Linear maps preserve zero and minus

Theorem

Each linear map $f: V \rightarrow W$ preserves:

- zero: f(0) = 0.
- minus: $f(-\mathbf{v}) = -f(\mathbf{v})$

Proof:

$$f(\mathbf{0}) = f(0 \cdot \mathbf{0}) \qquad f(-\mathbf{v}) = f((-1) \cdot \mathbf{v}) \\ = 0 \cdot f(\mathbf{0}) \qquad = (-1) \cdot f(\mathbf{v}) \\ = \mathbf{0} \qquad = -f(\mathbf{v})$$

Radboud University Nijmegen

Linear map examples I

 \mathbb{R} is a vector space. Let's consider maps $f : \mathbb{R} \to \mathbb{R}$. Most of them are *not linear*, like, for instance:

- f(x) = 1 + x, since $f(0) = 1 \neq 0$
- $f(x) = x^2$, since $f(-1) = 1 = f(1) \neq -f(1)$.

So: linear maps $\mathbb{R} \to \mathbb{R}$ can only be very simple.

Theorem

Each linear map $f : \mathbb{R} \to \mathbb{R}$ is of the form $f(x) = c \cdot x$, for some $c \in \mathbb{R}$.

Proof:

$$f(x) = f(x \cdot 1) = x \cdot f(1) = f(1) \cdot x = c \cdot x$$
, for $c = f(1)$.

Linear map examples II

Linear maps $\mathbb{R}^2 \to \mathbb{R}^2$ start to get more interesting:

$$s\begin{pmatrix} v_1\\ v_2 \end{pmatrix} = \begin{pmatrix} av_1\\ v_2 \end{pmatrix} \qquad t\begin{pmatrix} v_1\\ v_2 \end{pmatrix} = \begin{pmatrix} v_1\\ bv_2 \end{pmatrix}$$

...these scale a vector on the X- and Y-axis.

We can show these are linear by checking the two linearity equations:

$$f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w}) \qquad \qquad f(a \cdot \mathbf{v}) = a \cdot f(\mathbf{v})$$



Linear map examples III

Consider the map $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f\left(\begin{array}{c} v_1\\ v_2 \end{array}\right) = \begin{pmatrix} v_1\cos(\varphi) - v_2\sin(\varphi)\\ v_1\sin(\varphi) + v_2\cos(\varphi) \end{pmatrix}$$

This map describes rotation in the plane, with angle φ :



We can also check linearity equations.

Linear map examples IV

These extend naturally to 3D, i.e. linear maps $\mathbb{R}^3 \to \mathbb{R}^3$:

$$sx\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} av_1 \\ v_2 \\ v_3 \end{pmatrix} \qquad sy\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} v_1 \\ bv_2 \\ v_3 \end{pmatrix} \qquad sz\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

Q: How do we do rotation?

A: Keep one coordinate fixed (axis of rotation), and 2D rotate the other two, e.g.

$$rz\begin{pmatrix} v_1\\ v_2\\ v_3 \end{pmatrix} = \begin{pmatrix} v_1 \cos(\varphi) - v_2 \sin(\varphi)\\ v_1 \sin(\varphi) + v_2 \cos(\varphi)\\ v_3 \end{pmatrix}$$

And it works!

These kinds of linear maps are the basis of all 3D graphics, animation, physics, etc.



Getting back to matrices

Q: So what is the relationship between this (cool) linear map stuff, and the (lets face it, kindof boring) stuff about matrices and linear equations from before?

A: Matrices are a convenient way to represent linear maps!

To get there, we need a new concept: basis of a vector space

Basis in space

• In \mathbb{R}^3 we can distinguish three special vectors:

$$(1,0,0)\in \mathbb{R}^3$$
 $(0,1,0)\in \mathbb{R}^3$ $(0,0,1)\in \mathbb{R}^3$

- These vectors form a basis for \mathbb{R}^3 , which means:
 - **1** These vectors span \mathbb{R}^3 , which means each vector $(x, y, z) \in \mathbb{R}^3$ can be expressed as a linear combination of these three vectors:

$$\begin{array}{ll} (x,y,z) &=& (x,0,0) + (0,y,0) + (0,0,z) \\ &=& x \cdot (1,0,0) + y \cdot (0,1,0) + z \cdot (0,0,1) \end{array}$$

- 2 Moreover, this set is as small as possible: no vectors are can be removed and still span \mathbb{R}^3 .
- Note: condition (2) is equivalent to saying these vectors are linearly independent

Basis

Definition

Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ form a basis for a vector space V if these $\mathbf{v}_1, \ldots, \mathbf{v}_n$

- are linearly independent, and
- span V in the sense that each w ∈ V can be written as linear combination of v₁,..., v_n, namely as:

 $\boldsymbol{w} = a_1 \boldsymbol{v}_1 + \cdots + a_n \boldsymbol{v}_n$ for some $a_1, \ldots, a_n \in \mathbb{R}$

- These scalars a_i are uniquely determined by $\boldsymbol{w} \in V$ (see below)
- A space V may have several bases, but the number of elements of a basis for V is always the same; it is called the dimension of V, usually written as dim(V) ∈ N.

The standard basis for \mathbb{R}^n

For the space ℝⁿ = {(x₁,...,x_n) | x_i ∈ ℝ} there is a standard choice of basis vectors:

$$m{e}_1 := (1, 0, 0 \dots, 0), \ m{e}_2 := (0, 1, 0, \dots, 0), \ \cdots, \ m{e}_n := (0, \dots, 0, 1)$$

- **e**_i has a 1 in the *i*-th position, and 0 everywhere else.
- We can easily check that these vectors are independent and span ℝⁿ.
- This enables us to state precisely that \mathbb{R}^n is *n*-dimensional.

An alternative basis for \mathbb{R}^2

- The standard basis for \mathbb{R}^2 is (1,0), (0,1).
- But many other choices are possible, eg. (1,1), (1,-1)
 - independence: if $a \cdot (1,1) + b \cdot (1,-1) = (0,0)$, then:

$\int a + b = 0$	and thus	∫ a = 0
$\int a - b = 0$		$\int b = 0$

spanning: each point (x, y) can written in terms of (1, 1), (1, -1), namely:

$$(x,y) = \frac{x+y}{2}(1,1) + \frac{x-y}{2}(1,-1)$$

Uniqueness of representations

Theorem

- Suppose V is a vector space, with basis v_1, \ldots, v_n
- assume $\mathbf{x} \in V$ can be represented in two ways:

 $\mathbf{x} = a_1 \mathbf{v}_1 + \dots + a_n \mathbf{v}_n$ and also $x = b_1 \mathbf{v}_1 + \dots + b_n \mathbf{v}_n$ Then: $a_1 = b_1$ and \dots and $a_n = b_n$.

Proof: This follows from independence of v_1, \ldots, v_n since:

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = (a_1 \mathbf{v}_1 + \dots + a_n \mathbf{v}_n) - (b_1 \mathbf{v}_1 + \dots + b_n \mathbf{v}_n) \\ = (a_1 - b_1) \mathbf{v}_1 + \dots + (a_n - b_n) \mathbf{v}_n.$$

Hence $a_i - b_i = 0$, by independence, and thus $a_i = b_i$.



Representing vectors

Fixing a basis B = {v₁,..., v_n} therefore gives us a *unique* way to represent a vector v ∈ V as a list of numbers called *coordinates*:

$$\mathbf{v} = a_1 \mathbf{v}_1 + \dots + a_n \mathbf{v}_n$$

New notation: $\mathbf{v} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}_{\mathcal{B}}$

• If $V = \mathbb{R}^n$, and \mathcal{B} is the standard basis, this is just the vector itself:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}_{\mathcal{B}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

- ...but if ${\mathcal B}$ is not the standard basis, this can be different
- …and if V ≠ ℝⁿ, a list of numbers is meaningless without fixing a basis.



What does it mean?

"The introduction of numbers as coordinates is an act of violence." – Hermann Weyl



What does it mean?

- Space is (probably) real
- ...but coordinates (and hence bases) only exist in our head
- Choosing a basis amounts to fixing some directions in space we decide to call "up", "right", "forward", etc.
- Then a linear combination like:

$$v = 5 \cdot up + 3 \cdot right - 2 \cdot forward$$

describes a point in space, mathematically.

• ...and it makes working with *linear maps* a *lot* easier

\$

Linear maps and bases, example I

- Take the linear map $f((x_1, x_2, x_3)) = (x_1 x_2, x_2 + x_3)$
- Claim: this map is entirely determined by what it does on the basis vectors (1,0,0), (0,1,0), (0,0,1) ∈ ℝ³, namely:

f((1,0,0)) = (1,0) f((0,1,0)) = (-1,1) f((0,0,1)) = (0,1).

Indeed, using linearity:

$$f((x_1, x_2, x_3)) = f((x_1, 0, 0) + (0, x_2, 0) + (0, 0, x_3))$$

= $f(x_1 \cdot (1, 0, 0) + x_2 \cdot (0, 1, 0) + x_3 \cdot (0, 0, 1))$
= $f(x_1 \cdot (1, 0, 0)) + f(x_2 \cdot (0, 1, 0)) + f(x_3 \cdot (0, 0, 1))$
= $x_1 \cdot f((1, 0, 0)) + x_2 \cdot f((0, 1, 0)) + x_3 \cdot f((0, 0, 1))$
= $x_1 \cdot (1, 0) + x_2 \cdot (-1, 1) + x_3 \cdot (0, 1)$
= $(x_1 - x_2, x_2 + x_3)$

Linear maps and bases, geometrically

"If we know how to transform any set of axes for a space, we know how to transform everything."





Linear maps and bases, example I (cntd)

•
$$f((x_1, x_2, x_3)) = (x_1 - x_2, x_2 + x_3)$$
 is totally determined by:

- $f((1,0,0)) = (1,0) \qquad f((0,1,0)) = (-1,1) \qquad f((0,0,1)) = (0,1)$
- We can organise this data into a 2×3 matrix:

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

The vector $f(\mathbf{v}_i)$, for basis vector \mathbf{v}_i , appears as the *i*-the column.

• Applying *f* can be done by a new kind of multiplication:

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \cdot x_1 + -1 \cdot x_2 + 0 \cdot x_3 \\ 0 \cdot x_1 + 1 \cdot x_2 + 1 \cdot x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_2 + x_3 \end{pmatrix}$$

Matrix-vector multiplication: Definition

Definition

For vectors $\mathbf{v} = (x_1, \dots, x_n)$, $\mathbf{w} = (y_1, \dots, y_n) \in \mathbb{R}^n$ define their inner product (or dot product) as the real number:

$$\langle \mathbf{v}, \mathbf{w} \rangle = x_1 y_1 + \dots + x_n y_n = \sum_{i=1}^n x_i y_i$$

Definition

If
$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$
 and $\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$, then $\mathbf{w} := \mathbf{A} \cdot \mathbf{v}$

is the vector whose *i*-th element is the dot product of the *i*-th row of matrix \boldsymbol{A} with the (input) vector \boldsymbol{v} .

Matrix-vector multiplication, explicitly

For **A** an $m \times n$ matrix, **B** a column vector of length *n*:

 $\mathbf{A} \cdot \mathbf{b} = \mathbf{c}$

is a column vector of length m.

$$\begin{pmatrix} \vdots & \vdots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \vdots & \vdots \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} \vdots \\ a_{i1}b_1 + \cdots + a_{in}b_n \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ c_i \\ \vdots \\ \vdots \end{pmatrix}$$

$$c_i = \sum_{k=1}^n a_{ik} b_k$$

:

Radboud University Nijmegen 🐺

Another example, to learn the mechanics

$$\begin{pmatrix} 9 & 3 & 2 & 9 & 7 \\ 8 & 5 & 6 & 6 & 3 \\ 4 & 5 & 8 & 9 & 3 \\ 3 & 4 & 3 & 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 9 \\ 5 \\ 2 \\ 5 \\ 7 \end{pmatrix}$$

$$= \begin{pmatrix} 9 \cdot 9 + 3 \cdot 5 + 2 \cdot 2 + 9 \cdot 5 + 7 \cdot 7 \\ 8 \cdot 9 + 5 \cdot 5 + 6 \cdot 2 + 6 \cdot 5 + 3 \cdot 7 \\ 4 \cdot 9 + 5 \cdot 5 + 8 \cdot 2 + 9 \cdot 5 + 3 \cdot 7 \\ 3 \cdot 9 + 4 \cdot 5 + 3 \cdot 2 + 3 \cdot 5 + 4 \cdot 7 \end{pmatrix}$$

$$= \begin{pmatrix} 81 + 15 + 4 + 45 + 49 \\ 72 + 25 + 12 + 30 + 21 \\ 36 + 25 + 16 + 45 + 21 \\ 27 + 20 + 6 + 15 + 28 \end{pmatrix} = \begin{pmatrix} 194 \\ 160 \\ 143 \\ 96 \end{pmatrix}$$

Representing linear maps

Theorem

For every linear map $f : \mathbb{R}^n \to \mathbb{R}^m$, there exists an $m \times n$ matrix **A** where:

$$f(\mathbf{v}) = \mathbf{A} \cdot \mathbf{v}$$

(where " \cdot " is the matrix multiplication of **A** and a vector **v**)

Proof. Let $\{e_1, \ldots, e_n\}$ be the standard basis for \mathbb{R}^n . **A** be the matrix whose *i*-th column is $f(e_i)$. Then:

$$\boldsymbol{A} \cdot \boldsymbol{e}_{i} = \begin{pmatrix} a_{11}0 + \ldots + a_{1i}1 + \ldots + a_{1n}0 \\ \vdots \\ a_{m1}0 + \ldots + a_{mi}1 + \ldots + a_{mn}0 \end{pmatrix} = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} = f(\boldsymbol{e}_{i})$$

Since it is enough to check basis vectors and $f(\mathbf{e}_i) = \mathbf{A} \cdot \mathbf{e}_i$, we are done.

A. Kissinger

Getting a matrix from a linear map

- This proof tells us how to build the matrix
- Here's how: Take a linear map and *evaluate* it at each basis vector of the input vector space. E.g. for:

$$f((x_1, x_2, x_3)) = (x_1 - x_2, x_2 + x_3)$$

...the input vector space is R³, so we need to evaluate at 3 basis vectors e₁, e₂, e₃:

$$f\begin{pmatrix} 1\\0\\0 \end{pmatrix} = \begin{pmatrix} 1\\0 \end{pmatrix} \qquad f\begin{pmatrix} 0\\1\\0 \end{pmatrix} = \begin{pmatrix} -1\\1 \end{pmatrix} \qquad f\begin{pmatrix} 0\\0\\1 \end{pmatrix} = \begin{pmatrix} 0\\1 \end{pmatrix}$$

 This gives us 3 vectors, which become the *columns* of our new matrix:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Getting a matrix from a linear map

• So, from
$$f((x_1, x_2, x_3)) = (x_1 - x_2, x_2 + x_3)$$
, we computed

$$oldsymbol{A} = egin{pmatrix} 1 & -1 & 0 \ 0 & 1 & 1 \end{pmatrix}$$

• If we stick this new matrix 'inside' *f*, with matrix multiplication, then *viola*:

$$f(\mathbf{v}) = \mathbf{A} \cdot \mathbf{v} \longrightarrow f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_2 + x_3 \end{pmatrix}$$

• What did this accomplish?

f is a whole function. A is 6 numbers.

Examples of linear maps and matrices I

Projections are linear maps that send higher-dimensional vectors to lower ones. Consider $f : \mathbb{R}^3 \to \mathbb{R}^2$

$$f\left(\begin{pmatrix} x\\ y\\ z \end{pmatrix}\right) = \begin{pmatrix} x\\ y \end{pmatrix}.$$

f maps 3d space to the the 2d plane. The matrix of f is the following 2×3 matrix:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$



Examples of linear maps and matrices II

We have already seen: Rotation over an angle φ is a linear map



This rotation is described by $f: \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$f((x,y)) = (x\cos(\varphi) - y\sin(\varphi), x\sin(\varphi) + y\cos(\varphi))$$

The matrix that describes f is

$$egin{pmatrix} \cos(arphi) & -\sin(arphi) \ \sin(arphi) & \cos(arphi) \end{pmatrix}$$

Radboud University Nijmegen 🛞

Example: systems of equations

Matrix summary

- Take the standard bases: $\{e_1,\ldots,e_n\} \subset \mathbb{R}^n$ and $\{e_1',\ldots,e_m'\} \subset \mathbb{R}^m$
- Every linear map f: ℝⁿ → ℝ^m can be represented by a matrix, and every matrix represents a linear map:

$$f(\mathbf{v}) = \mathbf{A} \cdot \mathbf{v}$$

- The *i*-th column of *A* is *f*(*e_i*), written in terms of the standard basis *e*'₁,..., *e*'_m of ℝ^m.
- (Next time, we'll see the matrix of *f* depends on the choice of basis: for different bases, a different matrix is obtained)