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 $V, W, \ldots$ are special kinds of sets whose elements are called vectors.

• Vectors can be added together, or multiplied by a real number, For $v, w \in V$, $a \in \mathbb{R}$:

\[
\begin{align*}
\quad v + w & \in V \\
\quad a \cdot v & \in V
\end{align*}
\]

• 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: $+, \cdot, 0$).

A common theme in mathematics: study functions $f : V \to W$ 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 \( X \xrightarrow{f} Y \)
  - 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. \( f(n) = \frac{1}{n+1} \) can be seen as function \( \mathbb{N} \to \mathbb{Q} \), that is from the **natural** numbers \( \mathbb{N} \) to the **rational** numbers \( \mathbb{Q} \)

- On each set \( X \) there is the **identity** function \( \text{id} : X \to X \) that does nothing: \( \text{id}(x) = x \).

- Also one can compose 2 functions \( X \xrightarrow{f} Y \xrightarrow{g} Z \) to a function:

\[
g \circ f : X \to Z \quad \text{given by} \quad (g \circ f)(x) = g(f(x))
\]
A linear map is a function that preserves the extra stuff in a vector space:

**Definition**

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

- **addition**: for all $v, v' \in V$,
  \[
  f(\underbrace{v + v'}_{\text{in } V}) = \underbrace{f(v)}_{\text{in } W} + \underbrace{f(v')}_{\text{in } W}
  \]

- **scalar multiplication**: for each $v \in V$ and $a \in \mathbb{R}$,
  \[
  f(\underbrace{a \cdot v}_{\text{in } V}) = \underbrace{a}_{\text{in } \mathbb{R}} \cdot \underbrace{f(v)}_{\text{in } W}
  \]
Linear maps preserve zero and minus

Theorem

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

- **zero:** $f(0) = 0$.
- **minus:** $f(-v) = -f(v)$

**Proof:**

\[
\begin{align*}
    f(0) &= f(0 \cdot 0) \\
    &= 0 \cdot f(0) \\
    &= 0 \\

    f(-v) &= f((-1) \cdot v) \\
    &= (-1) \cdot f(v) \\
    &= -f(v)
\end{align*}
\]
$\mathbb{R}$ 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:**

$$f(x) = f(x \cdot 1) = x \cdot f(1) = f(1) \cdot x = c \cdot x, \quad \text{for } c = f(1).$$
Linear maps $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ start to get more interesting:

\[
\begin{align*}
    s\left( \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right) &= \begin{pmatrix} av_1 \\ v_2 \end{pmatrix} \\
    t\left( \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right) &= \begin{pmatrix} v_1 \\ bv_2 \end{pmatrix}
\end{align*}
\]

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

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

\[
\begin{align*}
    f(v + w) &= f(v) + f(w) \\
    f(a \cdot v) &= a \cdot f(v)
\end{align*}
\]
Consider the map \( f : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \) given by

\[
f\left( \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right) = \begin{pmatrix} v_1 \cos(\varphi) - v_2 \sin(\varphi) \\ v_1 \sin(\varphi) + v_2 \cos(\varphi) \end{pmatrix}
\]

This map describes rotation in the plane, with angle \( \varphi \):

We can also check linearity equations.
These extend naturally to 3D, i.e. linear maps $\mathbb{R}^3 \rightarrow \mathbb{R}^3$:

$$
\begin{align*}
  s_x \left( \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \right) &= \begin{pmatrix} a v_1 \\ v_2 \\ v_3 \end{pmatrix}, \\
  s_y \left( \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \right) &= \begin{pmatrix} v_1 \\ b v_2 \\ v_3 \end{pmatrix}, \\
  s_z \left( \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \right) &= \begin{pmatrix} v_1 \\ v_2 \\ c v_3 \end{pmatrix}
\end{align*}
$$

**Q:** How do we do rotation?

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

$$
\begin{align*}
  r_z \left( \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \right) &= \begin{pmatrix} v_1 \cos(\varphi) - v_2 \sin(\varphi) \\ v_1 \sin(\varphi) + v_2 \cos(\varphi) \\ v_3 \end{pmatrix}
\end{align*}
$$
And it works!

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

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

• 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:

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

  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
**Definition**

Vectors \( v_1, \ldots, v_n \in V \) form a **basis** for a vector space \( V \) if these \( v_1, \ldots, v_n \)

- are **linearly independent**, and
- span \( V \) in the sense that each \( w \in V \) can be written as linear combination of \( v_1, \ldots, v_n \), namely as:

\[
    w = a_1 v_1 + \cdots + a_n v_n \quad \text{for some} \quad a_1, \ldots, a_n \in \mathbb{R}
\]

- These scalars \( a_i \) are uniquely determined by \( w \in V \) (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 \( \dim(V) \in \mathbb{N} \).
The standard basis for $\mathbb{R}^n$

- For the space $\mathbb{R}^n = \{ (x_1, \ldots, x_n) | x_i \in \mathbb{R} \}$ there is a standard choice of basis vectors:
  
  $e_1 := (1, 0, 0 \ldots, 0), \ e_2 := (0, 1, 0, \ldots, 0), \ \cdots, \ e_n := (0, \ldots, 0, 1)$

- $e_i$ has a 1 in the $i$-th position, and 0 everywhere else.
- We can easily check that these vectors are independent and span $\mathbb{R}^n$.
- This enables us to state precisely that $\mathbb{R}^n$ is $n$-dimensional.
• 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:
    \[
    \begin{cases}
    a + b = 0 \\
    a - b = 0
    \end{cases}
    \text{ and thus } \begin{cases}
    a = 0 \\
    b = 0
    \end{cases}
    \]

• 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 $x \in V$ can be represented in two ways:
  
  $$x = a_1 v_1 + \cdots + a_n v_n \quad \text{and also} \quad x = b_1 v_1 + \cdots + b_n v_n$$

Then: $a_1 = b_1$ and \ldots and $a_n = b_n$.

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

$$0 = x - x = (a_1 v_1 + \cdots + a_n v_n) - (b_1 v_1 + \cdots + b_n v_n)
= (a_1 - b_1) v_1 + \cdots + (a_n - b_n) v_n.$$

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

- Fixing a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ therefore gives us a *unique* way to represent a vector $v \in V$ as a list of numbers called *coordinates*:

  $$v = a_1v_1 + \cdots + a_nv_n$$

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

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

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

- ...but if $\mathcal{B}$ is not the standard basis, this can be different

- ...and if $V \neq \mathbb{R}^n$, 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:

\[ \mathbf{v} = 5 \cdot \mathbf{up} + 3 \cdot \mathbf{right} - 2 \cdot \mathbf{forward} \]

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 \((1, 0, 0), (0, 1, 0), (0, 0, 1) \in \mathbb{R}^3\), namely:
  \[
f((1, 0, 0)) = (1, 0) \quad f((0, 1, 0)) = (-1, 1) \quad f((0, 0, 1)) = (0, 1).
  \]
- Indeed, using linearity:
  \[
f((x_1, x_2, x_3)) \\
= f\left((x_1, 0, 0) + (0, x_2, 0) + (0, 0, x_3)\right) \\
= f\left(x_1 \cdot (1, 0, 0) + x_2 \cdot (0, 1, 0) + x_3 \cdot (0, 0, 1)\right) \\
= f\left(x_1 \cdot (1, 0, 0)\right) + f\left(x_2 \cdot (0, 1, 0)\right) + f\left(x_3 \cdot (0, 0, 1)\right) \\
= x_1 \cdot f((1, 0, 0)) + x_2 \cdot f((0, 1, 0)) + x_3 \cdot f((0, 0, 1)) \\
= x_1 \cdot (1, 0) + x_2 \cdot (-1, 1) + x_3 \cdot (0, 1) \\
= (x_1 - x_2, x_2 + x_3)
  \]
“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:
  - $f((1, 0, 0)) = (1, 0)$
  - $f((0, 1, 0)) = (-1, 1)$
  - $f((0, 0, 1)) = (0, 1)$

- We can organise this data into a $2 \times 3$ matrix:
  $$
  \begin{pmatrix}
  1 & -1 & 0 \\
  0 & 1 & 1 \\
  \end{pmatrix}
  $$

  The vector $f(v_i)$, for basis vector $v_i$, appears as the $i$-th 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 \( \mathbf{v} = (x_1, \ldots, x_n) \), \( \mathbf{w} = (y_1, \ldots, y_n) \) \( \in \mathbb{R}^n \) define their inner product (or dot product) as the real number:

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

Definition

If \( \mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \) 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 \( \mathbf{A} \) with the (input) vector \( \mathbf{v} \).
Matrix-vector multiplication, explicitly

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

$$A \cdot b = c$$

is a column vector of length $m$.

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

$$c_i = \sum_{k=1}^{n} a_{ik} b_k$$
Another example, to learn the mechanics

\[
\begin{pmatrix}
9 & 3 & 2 & 9 & 7 \\
8 & 5 & 6 & 6 & 3 \\
4 & 5 & 8 & 9 & 3 \\
3 & 4 & 3 & 3 & 4 \\
\end{pmatrix}
\begin{pmatrix}
9 \\
5 \\
2 \\
5 \\
7 \\
\end{pmatrix}
= \\
\begin{pmatrix}
9 \cdot 9 + 3 \cdot 5 + 2 \cdot 2 + 9 \cdot 5 + 7 \cdot 7 \\
8 \cdot 9 + 5 \cdot 5 + 6 \cdot 2 + 6 \cdot 5 + 3 \cdot 7 \\
4 \cdot 9 + 5 \cdot 5 + 8 \cdot 2 + 9 \cdot 5 + 3 \cdot 7 \\
3 \cdot 9 + 4 \cdot 5 + 3 \cdot 2 + 3 \cdot 5 + 4 \cdot 7 \\
\end{pmatrix}
= \\
\begin{pmatrix}
81 + 15 + 4 + 45 + 49 \\
72 + 25 + 12 + 30 + 21 \\
36 + 25 + 16 + 45 + 21 \\
27 + 20 + 6 + 15 + 28 \\
\end{pmatrix}
= \\
\begin{pmatrix}
194 \\
160 \\
143 \\
96 \\
\end{pmatrix}
\]
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:

\[
f(v) = A \cdot v
\]

(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:

\[
A \cdot e_i = \begin{pmatrix}
a_{1i}0 + \ldots + a_{1i}1 + \ldots + a_{1n}0 \\
\vdots \\
a_{mi}0 + \ldots + a_{mi}1 + \ldots + a_{mn}0
\end{pmatrix} = \begin{pmatrix}
a_{1i} \\
\vdots \\
a_{mi}
\end{pmatrix} = f(e_i)
\]

Since it is enough to check basis vectors and \( f(e_i) = A \cdot e_i \), we are done.
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 \( \mathbb{R}^3 \), so we need to evaluate at 3 basis vectors \( e_1, e_2, e_3 \):

\[ f\left( \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad f\left( \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad f\left( \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \]

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

\[ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \rightarrow \quad \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \]
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:

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

- If we stick this new matrix ‘inside’ $f$, with matrix multiplication, then *viola*:

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

- 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 $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$

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

$f$ maps 3d space to the the 2d plane.

The matrix of $f$ is the following $2 \times 3$ 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 \( \varphi \) 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

\[
\begin{pmatrix}
\cos(\varphi) & -\sin(\varphi) \\
\sin(\varphi) & \cos(\varphi)
\end{pmatrix}.
\]
Example: systems of equations

\[ a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \]
\[ \vdots \]
\[ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \]

⇒

\[ A \cdot x = b \]

\[ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \]

\[ a_{11}x_1 + \cdots + a_{1n}x_n = 0 \]
\[ \vdots \]
\[ a_{m1}x_1 + \cdots + a_{mn}x_n = 0 \]

⇒

\[ A \cdot x = 0 \]

\[ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \]
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 \( f : \mathbb{R}^n \rightarrow \mathbb{R}^m \) can be represented by a matrix, and every matrix represents a linear map:

\[
f(v) = A \cdot v
\]

- The \( i \)-th column of \( A \) is \( f(e_i) \), written in terms of the standard basis \( e'_1, \ldots, e'_m \) of \( \mathbb{R}^m \).
- (Next time, we’ll see the matrix of \( f \) depends on the choice of basis: for different bases, a different matrix is obtained)