ITCS 531: Linear Algebra - Linear maps and matrices

Rob Egrot

What is a linear map?

Definition 1
Let V and W be vector spaces over the same field $\mathbb{F}$. A function $T : V \to W$ is a *linear map* if the following linearity conditions are satisfied:

1. $T(u + v) = T(u) + T(v)$ for all $u, v \in V$.
2. $T(\lambda v) = \lambda T(v)$ for all $v \in V$ and for all $\lambda \in \mathbb{F}$.

- Think of an equation $y = ax$ defining a straight line through the origin in the Euclidean plane.
- Think of two real numbers x_1 and x_2. Then, at the point $x_1 + x_2$, the value of y is given by $a(x_1 + x_2) = ax_1 + ax_2$.
- Similarly, if b is another real number then the value of y at the point bx_1 is given by $a(bx_1)$, which is equal to $ba(x_1)$.
- Linear maps 'behave like' straight lines.

Examples of linear maps

Definition 2 ($\mathcal{L}(V, W)$)

If V and W are vector spaces over the same field, then we denote the set of all linear maps from V to W by $\mathcal{L}(V, W)$.

Example 3

1. A straight line $y = ax$ can be thought of as a linear map from $\mathbb{R}$ (or any other field $\mathbb{F}$) to itself.
2. For any vectors spaces V and W over the same field, define the **zero map** $0 : V \to W$ by $0(v) = 0$ for all $v \in V$.
3. For any vector space V define the **identity map** $I : V \to V$ by $I(v) = v$ for all $v \in V$.
4. Remember $\mathbb{R}[x]$ is the vector space of polynomials over $\mathbb{R}$ with the variable x. The map $D : \mathbb{R}(x) \to \mathbb{R}(x)$ defined by taking first derivatives is a linear map.

Linear maps and zero

Lemma 4

If $T : V \to W$ is a linear map, then $T(0) = 0$.

Proof.

$0 = 0 + 0$, so, by linearity,

$$T(0) = T(0 + 0) = T(0) + T(0),$$

and so by subtracting $T(0)$ from both sides we see $T(0) = 0$ as required. $\square$

Defining linear maps

- Given arbitrary vector spaces V and W (over the same field), how can we define a linear map T between them?
- Do we have to specify the value $T(v)$ for every vector $v \in V$?
- Fortunately the answer to this is no, at least, it is no so long as we know a basis for V.

Proposition 5

Let V be a finite dimensional vector space over $\mathbb{F}$, and let $(v_1, \ldots, v_n)$ be a basis for V. Let W be another vector space over $\mathbb{F}$. Then for any $w_1, \ldots, w_n \in W$, there is a unique linear map $T : V \to W$ such that $T(v_i) = w_i$ for all $i \in \{1, \ldots, n\}$.

Defining linear maps - proof

- By the definition of a basis, given an element $v \in V$ we have $v = a_1 v_1 + \ldots + a_n v_n$ for some $a_1, \ldots, a_n \in \mathbb{F}$.
- Then the requirement that T is linear tells us what value T must take at on v.
- I.e.

$$T(v) = a_1 T(v_1) + \ldots + a_n T(v_n) = a_1 w_1 + \ldots + a_n w_n.$$

- It's straightforward to show that T defined in this way is a linear map (we just have to check the two conditions from definition 1).

More on defining linear maps

- Proposition 5 tells us that a linear map from a finite dimensional vector space V is completely determined by what it does to the basis vectors of V.
- Also, a linear map can take any values on the basis vectors of V.
- I.e. if $\{v_1, \ldots, v_n\}$ provides a basis for V, then the set of linear maps $\mathcal{L}(V, W)$ is in bijection with the set of functions from $\{v_1, \ldots, v_n\}$ to W.

Composition of linear maps

Definition 6

If $S \in \mathcal{L}(U, V)$, and $T \in \mathcal{L}(V, W)$, then it's easy to check that the composition $TS \in \mathcal{L}(U, W)$, where TS is defined by $TS(u) = T(S(u))$ for all $u \in U$.

Null spaces

Definition 7

Given $T \in \mathcal{L}(V, W)$, the *null space* of T, denoted $\mathrm{null}\, T$, is defined by

$$\mathrm{null}\, T = \{v \in V : T(v) = 0\}.$$

Lemma 8

$\mathrm{null}\, T$ *is a subspace of* V*.*

Proof.

This is exercise 3.3. $\square$

Null spaces

Lemma 9

Let $T \in \mathcal{L}(V, W)$. Then T is injective if and only if $\mathrm{null}\, T = \{0\}$.

Proof.

- Clearly if T is injective then only 0 can be mapped to 0 by T, so we have the forward implication.
- For the converse, suppose T is not injective.
- Then there are $u, v \in V$ with $u \neq v$ and $T(u) = T(v)$.
- But then by linearity of T we have $T(u - v) = T(u) - T(v) = 0$, and so $u - v \in \mathrm{null}\, T$.

$\square$

Range

Definition 10

Given $T \in \mathcal{L}(V, W)$, the *range* of T is defined by

$$\operatorname{ran} T = \{T(v) : v \in V\}.$$

Lemma 11

If $T \in \mathcal{L}(V, W)$ then $\operatorname{ran} T$ is a subspace of W.

Proof.

We just need to check the conditions for being a subspace are satisfied:

1. Since $0 = T(0)$ we have $0 \in \operatorname{ran} T$.
2. $T(u) + T(v) = T(u + v)$, so $\operatorname{ran} T$ is closed under vector addition.
3. $\lambda T(v) = T(\lambda v)$, so $\operatorname{ran} T$ is closed under scalar multiplication.

$\square$

The Rank-Nullity theorem

Theorem 12

Let V be finite dimensional, and let $T \in \mathcal{L}(V, W)$. Then $\operatorname{ran} T$ is finite dimensional, and

$$\dim V = \dim \operatorname{ran} T + \dim \operatorname{null} T.$$

In other words, the dimension of V is equal to the rank of T plus the nullity of T.

The Rank-Nullity theorem - proof

- Let $(u_1, \ldots, u_m)$ be a basis for $\mathrm{null}\, T$.
- Extend $(u_1, \ldots, u_m)$ to a basis $(u_1, \ldots, u_m, v_1, \ldots, v_n)$ for V.
- We will show $(T(v_1), \ldots, T(v_n))$ is a basis for $\mathrm{ran}\, T$.
- First check $(T(v_1), \ldots, T(v_n))$ is linearly independent.
- Suppose $0 = a_1 T(v_1) + \ldots + a_n T(v_n)$. Then $T(a_1 v_1 + \ldots + a_n v_n) = 0$, by the linearity of T.
- But this means $a_1 v_1 + \ldots + a_n v_n \in \mathrm{null}\, T$, and so there are $b_1, \ldots, b_m$ with $a_1 v_1 + \ldots + a_n v_n = b_1 u_1 + \ldots + b_m u_m$. I.e.

$$a_1 v_1 + \ldots + a_n v_n - b_1 u_1 - \ldots - b_m u_m = 0,$$

- As $(u_1, \ldots, u_m, v_1, \ldots, v_n)$ is a basis for V (and so is linearly independent), the only way this can happen is if

$$a_1 = \ldots = a_n = b_1 = \ldots = b_m = 0.$$

- In particular we have $a_1 = \ldots = a_n = 0$, and so $(T(v_1), \ldots, T(v_n))$ is indeed linearly independent.

The Rank-Nullity theorem - proof continued

- We must show that $(T(v_1), \ldots, T(v_n))$ spans $\mathrm{ran}\, T$.
- If $w \in \mathrm{ran}\, T$ then, by definition of range, there must be $v \in V$ with $w = T(v)$.
- Since $(u_1, \ldots, u_m, v_1, \ldots, v_n)$ is a basis for V, it follows that there must be $a_1, \ldots, a_m, b_1, \ldots, b_n$ with

$$v = a_1 u_1 + \ldots + a_m u_m + b_1 v_1 + \ldots + b_n v_n.$$

- Since $u_i \in \mathrm{null}\, T$ for all $i \in \{1, \ldots, m\}$, and so $T(u_i) = 0$, we have

$$w = T(v) = b_1 T(v_1) + \ldots + b_n T(v_n).$$

- So $w \in \mathrm{span}(T(v_1), \ldots, T(v_n))$ as required.

Matrices

An $m \times n$ matrix over a field $\mathbb{F}$ is an array of elements of $\mathbb{F}$. We can express matrices explicitly, using the following form:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Here the element a_{ij} is the one in the ith row and the jth column. We sometimes use the shorthand (a_{ij}) to express a matrix of the form above.

Matrix algebra - scalar multiplication

Given an $m \times n$ matrix A over $\mathbb{F}$, and $\lambda \in \mathbb{F}$, we define the scalar product λA to be

$$\begin{bmatrix} \lambda a_{11} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \dots & \lambda a_{2n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \dots & \lambda a_{mn} \end{bmatrix}$$

Matrix algebra - matrix addition

Given $m \times n$ matrices $A = (a_{ij})$ and $B = (b_{ij})$ over the same field $\mathbb{F}$, we define the sum $A + B$ to be

$$\begin{bmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & \dots & a_{2n} + b_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

Matrix algebra - matrix multiplication

- Let $A = (a_{ij})$ be an $m \times n$ matrix over $\mathbb{F}$.
- Let $B = (b_{jk})$ be an $n \times p$ matrix over $\mathbb{F}$.
- Define the matrix product AB to be the $m \times p$ matrix (c_{ik}).
- Here for each $i \in \{1, \dots m\}$ and $k \in \{1, \dots, p\}$, the entry c_{ik} is defined by

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}.$$

- I.e. the element c_{ik} is defined using the ith row of A, and the kth column of B.

Matrix algebra - matrix multiplication of a vector

In particular, if A is an $m \times n$ matrix, and v is a $n \times 1$ matrix (so v a column vector in $\mathbb{F}^n$), then the product Av is calculated using:

$$Av = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} b_1 a_{11} + b_2 a_{12} + \ldots + b_n a_{1n} \\ b_1 a_{21} + b_2 a_{22} + \ldots + b_n a_{2n} \\ \vdots \\ b_1 a_{m1} + b_2 a_{m2} + \ldots + b_n a_{mn} \end{bmatrix}$$

Matrices define linear maps

- Matrix multiplication as defined here is quite mysterious. However, as we shall soon see, it is actually extremely natural.
- Every $m \times n$ matrix over $\mathbb{F}$ defines a linear map $\mathbb{F}^n \to \mathbb{F}^m$.
- That is, an $m \times n$ matrix over $\mathbb{F}$ takes a vector from $\mathbb{F}^n$ and transforms it into a vector from $\mathbb{F}^m$.
- Moreover, this transformation is linear (easy to check).
- The correspondence between linear maps and matrices actually goes both ways.
- I.e. every linear map between finite dimensional vector spaces can be represented by a matrix.

Linear maps define matrices 1

- Let $T \in \mathcal{L}(V, W)$, and suppose V and W are both finite dimensional.
- Let $(v_1, \ldots, v_n)$ be a basis for V, and let $(w_1, \ldots, w_m)$ be a basis for W.
- The map T is defined by what it does to $v_1, \ldots, v_n$.
- As $(w_1, \ldots, w_m)$ is a basis for W, each $T(v_j)$ can be written as a linear combination of elements of $\{w_1, \ldots, w_m\}$.
- From this we can calculate the matrix associated with T with respect to the bases $(v_1, \ldots, v_n)$ and $(w_1, \ldots, w_m)$.

Linear maps define matrices 2

- For each $j \in \{1, \ldots, n\}$, we have

$$T(v_j) = a_{1j} w_1 + \ldots + a_{mj} w_m. \tag{$\dagger$}$$

- Let $A = (a_{ij})$ be the matrix defined using the a_{ij} defined in ($\dagger$), with $i \in \{1, \ldots, m\}$.

$$\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix}$$

Linear maps define matrices 3

- Given a column vector that has 1 in it's jth place and 0 everywhere else.
- What happens when we multiply this vector with A?

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

Linear maps define matrices 4

- We can interpret the vector below as the element v_j of V.

$$\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

- And the vector below as the element of W defined by $a_{1j}w_1 + \ldots + a_{mj}w_m$. I.e. $T(v_j)$.

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

- Why? Because this is how we defined the a_{ij} values.

Linear maps define matrices 5

- So this matrix multiplication transforms v_j into $T(v_j)$.
- This is true for all $j \in \{1, \ldots, n\}$, so the matrix A corresponds to the action of T on every basis vector v_i.
- Note that

  $$A(au + bv) = aAu + bAv = aTu + bTv = T(au + bv)$$

  for all $u, v \in V$ and $a, b \in \mathbb{F}$.
- So A corresponds to the action of T on every element of V.
- A represents T with respect to the choice of bases for V, W.
- Different bases will produce a different matrix.

Linear maps define matrices 6

- This tells us why matrix multiplication is defined like it is.
- We can think of the product AB of two matrices as being the list of vectors we get from applying the transformation defined by A to each of the columns of B.
- From this perspective, the jth column of AB is the result of applying A to the jth column of B.

Composition of linear maps and matrices

A nice property of correspondence between matrices and linear maps is that it also extends to compositions of linear maps.

Proposition 13

Let $S \in \mathcal{L}(U, V)$ and let $T \in \mathcal{L}(V, W)$. Let $(u_1, \ldots, u_n)$, $(v_1, \ldots, v_m)$ and $(w_1, \ldots, w_p)$ be bases for U, V and W respectively. Suppose that B is the matrix of T with respect to $(v_1, \ldots, v_m)$ and $(w_1, \ldots, w_p)$, and that A is the matrix of S with respect to $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_m)$. Then BA is the matrix of TS with respect to $(u_1, \ldots, u_n)$ and $(w_1, \ldots, w_p)$.

Proof.

Exercise 3.4. $\square$

Thinking about linear maps

- Linear maps are linear transformations of space.
- I.e., transformations of space that keep straight lines straight.
- Think of Euclidean space $\mathbb{R}^3$.
- The vectors $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ define a *unit cube*.
- If $A = (a_{ij})$ is a 3×3 matrix, then A acting on the vectors $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ produces three new vectors.
- These vectors also define a shape in Euclidean space. This shape is the result of transforming the unit cube by the transformation defined by A.
- The linearity of A means it can stretch vectors, and change their directions, but it can't bend them.

Thinking about linear maps

Example 14

Let A be the real valued matrix

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

- Think of A as a linear transformation of the Euclidean plane.
- What does A do to $(1, 0)$?
- A takes $(1, 0)$ to $(0, 1)$, and A takes $(0, 1)$ to $(-1, 0)$.
- If $(1, 0)$ and $(0, 1)$ have their usual meanings as vectors in $\mathbb{R}^2$, then this corresponds to an anticlockwise rotation by $\frac{\pi}{2}$ radians (90°).

Determinants

- A *determinant* is a value calculated from a square matrix.
- We understood the action of a matrix by looking at its effect on a 'unit cube', e.g. in $\mathbb{R}^3$.
- The determinant of a matrix corresponds to the 'volume' of the unit cube after being transformed.
- This makes intuitive sense in 3 dimensions, and generalizes to other dimensions.
- The determinant can be positive or negative, so it gives us more information than just the volume of the transformed unit cube.
- This is known as a *signed volume*.

Determinants and inverses

Definition 15

A linear map $T \in \mathcal{L}(V, W)$ is *invertible* if there is a map $T^{-1} \in \mathcal{L}(W, V)$ such that the map $T^{-1}T$ is the identity map on V, and the map TT^{-1} is the identity map on W.

- A map $T : V \to W$ should be invertible so long as no information is 'lost' during the transformation.
- A linear map T is invertible if and only if the corresponding matrix is invertible.
- Can check whether a matrix is invertible by finding its determinant (invertible iff non-zero).
- Determinant is zero *iff* volume of transformed unit cube becomes zero *iff* linear transformation 'loses a dimension'
- I.e. *iff* not invertible.