Module 7: Linear Algebra I

Operational math bootcamp

Statistical Sciences
UNIVERSITY OF TORONTO

Emma Kroell

University of Toronto

July 21, 2022

Outline

Last time:

- Vector space
- Subspace

Today:

- Linear independence and bases
- Linear maps, null space, range, inverses
- Matrices as linear maps

Linear combinations

$\mathbb{F}$-vector space

Definition

A linear combination of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of vectors in V is a vector of the form

$$\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n = \sum_{k=1}^{n} \alpha_k \mathbf{v}_k$$

where $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$.

Span

Definition

The set of all linear combinations of a list of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{\cancel{m}\,n}$ in V is called the **span** of $\mathbf{v}_1, \ldots, \mathbf{v}_{\cancel{m}\,n}$ denoted $\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. In other words,

$$\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} = \{\alpha_1 \mathbf{v}_1 + \ldots + \alpha_m \mathbf{v}_n : \alpha_1, \ldots, \alpha_n \in \mathbb{F}\}$$

The span of the empty list is defined to be $\{\mathbf{0}\}$.

Basis

Definition

A system of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is called a basis (for the vector space V) if any vector $\mathbf{v} \in V$ admits a unique representation as a linear combination

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n = \sum_{k=1}^{n} \alpha_k \mathbf{v}_k .$$

Example

$= \mathbb{R}^n$

- For $\mathbb{F}^n$, $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $e_n = (0, \ldots, 0, 1)$ is a basis
- The monomials $1, x, x^2, \ldots, x^n$ form a basis for $\mathbb{P}_n$.

$\mathbb{R}^2 : (0,1), (1,0)$

Linear independence

Definition

A system of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in V is called *linearly independent* if

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}$$

implies $\alpha_i = 0$ for all $i = 1, \ldots, n$.

Otherwise, we call the system *linearly dependent*.

Linear combinations $\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n$ such that $\alpha_k = 0$ for every k are called trivial.

Spanning set

Definition

A system of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in V is called *spanning* if any vector in V can be written as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$. In other words,

$$V = \mathrm{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}.$$

Such a system is also often called generating or complete. The next proposition relates spanning and linearly independent to a basis.

Proposition

A system of vectors $\mathbf{v}_1, \ldots \mathbf{v}_n \in V$ is a basis if and only if it is linearly independent and spanning.

Proof.

($\Rightarrow$) Let $v_1, \ldots, v_n$ be a basis for V.
By definition, any $v \in V$ has a unique representation as a linear combination of $v_1, \ldots, v_n$. $\therefore$ $v_1, \ldots, v_n$ is spanning.
Since the representation is unique for each $v \in V$, $\vec{0} = 0v_1 + \ldots + 0v_n$, this must be the only way to $= \vec{0}$. Thus $v_1, \ldots, v_n$ are lin. ind.

Proof continued

($\Leftarrow$) Suppose $v_1, \dots, v_n$ is lin. ind & span V. Let $v \in V$.
Since $v_1, \dots, v_n$ spanning, $\exists \alpha_i \in \mathbb{F}$ s.t. $v = \sum_{i=1}^{n} \alpha_i v_n$.
To show that this is unique, suppose $\exists \beta_i \in \mathbb{F}$
s.t. $\sum_{i=1}^{n} \beta_i v_i = v$.

Then $\vec{0} = v - v = \sum_{i=1}^{n} (\alpha_i - \beta_i) v_i$

By linear ind, $\alpha_i = \beta_i \ \forall i = 1, \dots, n$. $\therefore$ it is unique

$\therefore v_1, \dots, v_n$ is a basis

Proposition

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ be spanning. Then $\mathbf{v}_1, \ldots, \mathbf{v}_n$ contains a basis.

Sketch of proof.

If $v_1, \ldots, v_n$ are linearly ind, the we're done. Otherwise, we find one that can be written as a combination of the others & remove it. Keep going until we have a basis.

Definition

An $\mathbb{F}$-vector space V is called *finite dimensional* if there exists a finite list of vectors that span it, i.e. there exist $n \in \mathbb{N}$ and $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ such that $V = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Otherwise, we call V *infinite dimensional*.

Example

- $\mathbb{F}^n$, $M_{m\times n}$, $\mathbb{P}_n$ are examples of finite dimensional vector spaces
- The $\mathbb{F}$-vector space $\mathbb{P} = \{\sum_{i=1}^{n} \alpha_i x^i \;:\; n \in \mathbb{N}, \alpha_i \in \mathbb{F}, i = 1, \ldots, n\}$ is infinite dimensional.

Why? Suppose it is finite. Then $\exists p_1, \ldots, p_n$ polynomials that span $\mathbb{P}$. But $p_1, \ldots, p_n$ must have a maximum degree; call it N.
Then $x^{N+1} \notin \operatorname{span}\{p_1, \ldots, p_n\} \Rightarrow \Leftarrow$

Corollary

Every finite dimensional vector space has a basis.

This follows from the fact that every spanning set for a vector space contains a basis.

This can also be extended to infinite dimensional vector spaces, i.e. when we do not assume that there exists a finite spanning set. However, this relies on the Axiom of Choice and is beyond the scope of this course.

Proposition

Every linearly independent list of vectors in a finite-dimensional vector space can be extended to a basis of the vector space.

Proof.

Let $u_1, \ldots, u_n$ be linearly independent vectors in U. Add the basis of U, $v_1, \ldots, v_n$.
Then $u_1, \ldots, u_n, v_1, \ldots, v_n$ spans V.
We can reduce it by Prop. 2.31 in book, to a basis that contains the u's.

Statistical Sciences
UNIVERSITY OF TORONTO

Dimension

Proposition

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and $\mathbf{u}_1, \ldots, \mathbf{u}_m$ be a basis for V. Then $m = n$.

The proof is omitted,. It relies on the fact that the number of elements in linearly independent systems are always less than or equal to the number of elements in spanning systems.

Definition

Let V be a finite dimensional $\mathbb{F}$-vector space. The number of elements in a basis of V is called the *dimension* of V and is denoted $\dim(V)$.

By the previous definition, the notion of dimension is well-defined.

Dimension

Example

- $\dim(\mathbb{F}^n) = n$
- $\dim(\mathbb{P}_n) = n + 1$
- $\dim\{\mathbf{0}\} = 0$

Linear Maps

Definition

A map from a vector space U to a vector space V is **linear** if

$$T(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha T(\mathbf{u}) + \beta T(\mathbf{v}) \quad \text{for any } \mathbf{u}, \mathbf{v} \in V,\ \alpha, \beta \in \mathbb{F}$$

Notation: $\mathcal{L}(U, V)$ is the set of all linear maps from $\mathbb{F}$-vector space U to $\mathbb{F}$-vector space V

Statistical Sciences
UNIVERSITY OF TORONTO

Example

- Zero map $0: U \to V$, $u \in U$, then $0u = \vec{0}$

- Identity map $I: V \to V$, $Iv = v$ for $v \in V$

- Differentiation $D \in \mathcal{L}(\underbrace{\mathcal{P}(\mathbb{R})}_{\to \text{ polynomials}}, \mathcal{P}(\mathbb{R}))$

$Dp = p'$

$\frac{d}{dx}(\alpha f(x) + \beta g(x)) = \alpha f'(x) + \beta g'(x)$

Theorem

Suppose $\mathbf{u}_1, \ldots, \mathbf{u}_n$ is a basis for U and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis for V. Then there exists a unique linear map $T : U \to V$ such that $T\mathbf{u}_j = \mathbf{v}_j$ for $j = 1, \ldots, n$.

Proof in book

Theorem

Let $S, T \in \mathcal{L}(U, V)$ and $\alpha \in \mathbb{F}$. $\mathcal{L}(U, V)$ is a vector space with addition defined as the sum $S + T$ and multiplication as the product αT.

The proof follows from properties of linear maps and vector spaces. Note that the additive identity is the zero map.

$S+T(v)$

Lemma

Let $T \in \mathcal{L}(U, V)$. Then $T(\mathbf{0}) = \mathbf{0}$.

Proof.

$$T(\vec{0}) = T(\vec{0} + \vec{0}) = T(\vec{0}) + T(\vec{0})$$

add $-T(\vec{0})$ to both sides,

$$\Rightarrow \vec{0} = T(\vec{0})$$

$\square$

Statistical Sciences
UNIVERSITY OF TORONTO

Null space and range

Definition

Let $T : U \to V$ be a linear transformation. We define the following important subspaces:

- *Kernel or null space*: $\text{null } T = \{\mathbf{u} \in U : T\mathbf{u} = 0\}$
- *Range*: $\text{range } T = \{\mathbf{v} \in V : \exists \mathbf{u} \in U \text{ such that } \mathbf{v} = T\mathbf{u}\}$

The dimensions of these spaces are often called the following:

- *Nullity*: $\text{nullity}(T) = \dim(\text{null}(T))$
- *Rank*: $\text{rank}(T) = \dim(\text{range}(T))$

Proposition

Let $T : U \to V$. The null space of T is a subspace of U and the range of T is a subspace of V.

Proof.

Since $T(0) = 0$, 0 is in null$T$.

$u, v \in$ null$T$ then $T(u+v) = T(u) + T(v) = 0 + 0 = 0$

$\alpha \in \mathbb{F}$, $v \in$ null$T$, then $T(\alpha v) = \alpha T(v) = \alpha 0$
$= 0$

range: $T(\vec{0}) = \vec{0}$, $\vec{0} \in$ range$T$

Suppose $v_1, v_2 \in$ range$T$: Then $\exists u_1, u_2 \in U$

s.t $T(u_1) = v_1$ & $T(u_2) = v_2$.

So $T(u_1 + u_2) = T(u_1) + T(u_2) = v_1 + v_2$

$\square$

Statistical Sciences
UNIVERSITY OF TORONTO

Example

Zero map from a vector space U to a vector space V:

- The null space is U
- The range is $\{\vec{0}\}$

Differentiation map from $\mathbb{P}(\mathbb{R})$ to $\mathbb{P}(\mathbb{R})$:

- The null space is constants
- The range is $\mathbb{P}(\mathbb{R})$

Definition (Injective and surjective)

Let $T : U \to V$. T is *injective* if $T\mathbf{u} = T\mathbf{v}$ implies $\mathbf{u} = \mathbf{v}$ and T is *surjective* if $\forall \mathbf{u} \in U$, $\exists \mathbf{v} \in V$ such that $\mathbf{v} = T\mathbf{u}$, i.e. if range $T = V$.

Theorem

$T \in \mathcal{L}(U, v)$ is injective if and only if null $T = \{\mathbf{0}\}$.

Proof.

($\Rightarrow$) Suppose T is injective. We know that $\vec{0} \in \operatorname{null} T$ because $T(0) = 0$. Suppose that $\exists\, v \in \operatorname{null} T$. Then $T(v) = 0 = T(0)$, since T is injective, $v = 0$. $\therefore \operatorname{null} T = \{\vec{0}\}$.

($\Leftarrow$) Suppose $\operatorname{null} T = \{\vec{0}\}$. Let $Tu = Tv$, $u, v \in U$. We want to show $u = v$.

$$Tu = Tv \Rightarrow T(u - v) = 0 \Rightarrow u - v \in \operatorname{null} T$$

$$\Rightarrow u - v = 0 \Rightarrow u = v$$

$\square$

Theorem (Rank Nullity Theorem)

Let $T : U \to V$ be a linear transformation, where U and V are finite-dimensional vector spaces. Then

$$\text{rank } T + \text{nullity } T = \dim U.$$

$\dim \text{range} T + \dim \text{null} T = \dim U$

Proof.

Let $u_1, \ldots, u_m$ be a basis for null T. We can extend it to a basis for U. Suppose $w_1, \ldots, w_n$ is added to $u_1, \ldots, u_m$ to have a basis for U.

Then $\dim U = m + n$, $\dim \text{null} = m$.

We show that $Tw_1, \ldots, Tw_n$ is a basis for range T.

Proof continued

Let $u \in U$. Then $\exists\, \alpha_i, \beta_j$ $i=1,\dots,m$, $j=1,\dots,n$, such that $u = \alpha_1 u_1 + \cdots + \alpha_m u_m + \beta_1 w_1 + \cdots + \beta_n w_n$

Apply T: $Tu = \alpha_1 Tu_1 + \cdots + \alpha_m Tu_m + \beta_1 Tw_1 + \cdots + \beta_n Tw_n$

$= \beta_1 Tw_1 + \cdots + \beta_n Tw_n$

$\therefore Tw_1, \dots, Tw_n$ span range T

Let $c_1, \dots, c_n \in \mathbb{F}$. Let $0 = c_1 Tw_1 + \cdots + c_n Tw_n$

$= T(c_1 w_1 + \cdots + c_n w_n)$

$c_1 w_1 + \cdots + c_n w_n \in \text{null}\, T$

Proof continued

Since $u_1, \dots, u_m$ is a basis for null T

$\Rightarrow \exists\, d_1, \dots, d_m \in \mathbb{F}$ s.t.

$$c_1 w_1 + \dots + c_n w_n = d_1 u_1 + \dots + d_m u_m$$

Since $u_1, \dots, u_m, w_1, \dots, w_n$ is a basis for V.

$$\therefore c_1 = \dots = c_n = d_1 = \dots = d_m = 0$$

Since c_i's $= 0$, then $Tw_1, \dots, Tw_n$ are lin. ind.

$\therefore Tw_1, \dots, Tw_n$ is a basis for range T $\therefore$ dim range $T = n$.

Definition (Product of linear maps)

Let $S \in \mathcal{L}(U, V)$ and $T \in \mathcal{L}(V, W)$. We define the product $ST \in \mathcal{L}(U, W)$ for $\mathbf{u} \in U$ as $ST(\mathbf{u}) = S(T(\mathbf{u}))$.

Definition

A linear map $T : U \to V$ is *invertible* if there exists a linear map $S : V \to U$ such that ST is the identity map on U and TS is the identity map on V. Such a map S is called the *inverse* of T.

If T is invertible, we denote the inverse by T^{-1}. This is justified by the fact that the inverse is unique:

$TT^{-1} = I \quad , \quad T^{-1}T = I$

Proposition

Any invertible linear map has a unique inverse.

Proof.

Let $T: U \to V$. Suppose S_1 & S_2 are both inverses for T.

Then $S_1 = S_1 I = \underbrace{S_1 T}_{S_1 \text{ is inv}} S_2 = I S_2 = S_2$.

$$S_1 T S_2 = S_1 \underbrace{T S_2}_{S_2 \text{ is inv}}$$

$\square$

Theorem

A linear map is invertible if and only if it is injective and surjective.

See proof in the book.

Definition

An invertible linear map is called an *isomorphism*. If there exists an isomorphism from one vector space to another, we say that the vector spaces are *isomorphic*.

Theorem

Two finite-dimensional vector spaces over $\mathbb{F}$ are isomorphic if and only if they have the same dimension.

Proof.

($\Rightarrow$) Suppose $\exists T : U \to V$ which is invertible.

By Thm, T is both injective & surjective.

$$\Rightarrow \text{null}\, T = \{\vec{0}\}, \quad \text{range}\, T = V$$

$$\dim \text{range}\, T + \dim \text{null}\, T = \dim U$$

$$\dim V + \underbrace{\dim \{\vec{0}\}}_{=0} = \dim U$$

$$\Rightarrow \dim V = \dim U$$

$\square$

Proof continued

($\Leftarrow$) Let U, V have the same dimension. Let $u_1, \dots, u_n$ be a basis for U and $v_1, \dots, v_n$ be a basis for V.

Define the map $T: U \to V$

$$T(c_1u_1 + \dots + c_nu_n) = c_1v_1 + \dots + c_nv_n$$

Since $v_1, \dots, v_n$ span V, map T is surjective and since they are linearly ind, $\text{null}\, T = \{\vec{0}\}$.

$\therefore T$ is surjective & injective

$\Rightarrow T$ is an isomorphism

Linear maps and matrices

Example

Let $A \in M_{m\times n}$ be a fixed matrix. Then, we can define a linear map $T_A \colon \mathbb{F}^n \to \mathbb{F}^m$ via $T_A(\mathbf{v}) = A\mathbf{v}$, where we recall matrix vector multiplication $(A\mathbf{v})_i = \sum_{k=1}^n A_{ik} v_k$ for $i = 1, \ldots, m$.

Next we will see that we can use matrices to represent linear maps between finite dimensional vector spaces.

Definition

Let $T \in \mathcal{L}(U, V)$ where U and V are vector spaces. Let $\mathbf{u}_1, \ldots, \mathbf{u}_n$ and $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be bases for U and V respectively. The matrix of T with respect to these bases is the $m \times n$ matrix $\mathcal{M}(T)$ with entries A_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$ defined by

$$T\mathbf{u}_k = A_{1k}\mathbf{v}_1 + \cdots + A_{mk}\mathbf{v}_m$$

i.e. the kth column of A is the scalars needed to write $T\mathbf{u}_k$ as a linear combination of the basis of V:

$$T\mathbf{u}_k = \sum_{i=1}^{m} A_{ik}\mathbf{v}_i$$

Note that since a linear map $T \in \mathcal{L}(U, V)$ is uniquely determined by its image on a basis of U, we see that once we pick basis of U and V its matrix representation is uniquely determined.

Example

Let $D \in \mathcal{L}(\mathbb{P}_4(\mathbb{R}), \mathbb{P}_3(\mathbb{R}))$ be the differentiation map, $Dp = p'$. Find the matrix of D with respect to the standard bases of $\mathbb{P}_3(\mathbb{R})$ and $\mathbb{P}_4(\mathbb{R})$.

Standard basis: $1, x, x^2, x^3, (x^4)$

$T(u_1) = (1)' = 0$

$T(u_2) = (x)' = 1$

$T(u_3) = (x^2)' = 2x$

$T(u_4) = (x^3)' = 3x^2$

$T(u_5) = (x^4)' = 4x^3$

The matrix is:

$$M(D) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

- Observe that if we choose bases $\mathbf{u}_1, \ldots, \mathbf{u}_n$ and $\mathbf{v}_1, \ldots, \mathbf{v}_m$ for U, V and represent $T \in \mathcal{L}(U, V)$ as a matrix $\mathcal{M}(T)$, then the corresponding map can be obtained by just working with the coordinates of vectors in U, V with respect to the chosen basis
- If $\mathbf{u} = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i$, then the coordinates of $T(\mathbf{u})$ with respect to $\mathbf{v}_1, \ldots, \mathbf{v}_m$ can be obtained by the matrix vector multiplication $\mathcal{M}(T)\boldsymbol{\alpha}$, where $\boldsymbol{\alpha}$ is the $n \times 1$ matrix with entries α_i

Example

If we want to find the derivative of $p = x^4 + 12x^3 - 5x^2 + 7$ with respect to the standard monomial basis of $\mathbb{P}_4(\mathbb{R})$, we use $\mathcal{M}(D)$ from the previous example to obtain

$$\mathcal{M}(D)\boldsymbol{\alpha} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 7 \\ 0 \\ -5 \\ 12 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -10 \\ 36 \\ 4 \end{pmatrix}.$$

Thus, translating back into the monomial basis of $\mathbb{P}_3(\mathbb{R})$ gives $D(p) = -10x + 36x^2 + 4x^3$.

Other points

- Looking at matrices as representations of linear maps gives us an intuitive explanation for why we do matrix multiplication the way we do! In fact, we want matrix multiplication to represent composition of linear maps
- We can use matrices to solve linear systems.

Next time

- Determinants
- Eigenvalues and eigenvectors
- Inner product spaces

References

Axler S. *Linear Algebra Done Right*. 3rd ed. Undergraduate Texts in Mathematics. Springer, 2015. Available from: https://link.springer.com/book/10.1007/978-3-319-11080-6

Treil S. *Linear Algebra Done Wrong*. 2017. Available from: https://www.math.brown.edu/streil/papers/LADW/LADW.html