Math 223: Linear Algebra
Lecture Notes

Lior Silberman

These are rough notes for the course in winter 2021. Problem sets solutions were posted on an internal website.

Contents

Introduction

Lior Silberman, lior@Math.UBC.CA, https://www.math.ubc.ca/~lior
Office: Math Annex 1112
Phone: 604-827-3031

0.1. Administrivia

Syllabus posted online. Key points:

- Problem sets will be posted on the course website. Solutions will be posted on a secure system.
  - The grader may only mark selected problems. Solutions will be complete.
- Absolutely essential to
  - ASK QUESTIONS IN CLASS
  - Read ahead according to the posted schedule. Lectures after the first will assume that you had done your reading.
  - Do homework.
- Office hours, Piazza.
- Course website has notes, problem sets, announcements, reading assignments etc.

0.2. Course plan (subject to revision)

Four aspects:

- Calculation ("matrix algebra")
- Language ("linear algebra in the wild")
- Linear Algebra
- Metamathematics

Topics

- Vector spaces
- Linear maps
- Linear Equations
- Determinants
- Eigenvectors and diagonalization
- Inner product spaces

0.3. Change-of-language

- Signal processing example: plug two guitars into amp. Ideally, output is sum of inputs, and rescaling inputs rescales output.
- Go from statements about functions to statements about sets of functions – see worksheet.

CHAPTER 1

Vector spaces and Linear maps

1.1. Vector spaces

DEFINITION 1. A (real) vector space is a triple $V = (V, +_V, \cdot_V)$ where:

(0) V is a set; $+_V : V \times V \to V$ and $\cdot_V : \mathbb{R} \times V \to V$ are operations.
(1) (addition of vectors) For every two elements ("vectors") $\underline{u}, \underline{v} \in V$, there is a vector $\underline{u} +_V \underline{v} \in V$ and:
   (a) ("Associativity") For all $\underline{u}, \underline{v}, \underline{w} \in V$ we have $(\underline{u} +_V \underline{v}) +_V \underline{w} = \underline{u} +_V (\underline{v} +_V \underline{w})$.
   (b) ("Commutativity") For all $\underline{u}, \underline{v} \in V$ we have $\underline{u} +_V \underline{v} = \underline{v} +_V \underline{u}$.
   (c) ("Zero") There is a vector $\underline{0}_V \in V$ such that for all $\underline{u} \in V$, $\underline{u} +_V \underline{0}_V = \underline{u}$.
   (d) ("Negatives") For all $\underline{u} \in V$ there is a vector $\underline{u}' \in V$ such that $\underline{u} +_V \underline{u}' = \underline{u}' +_V \underline{u} = \underline{0}_V$.
(2) (scalar multiplication) For every ("scalar")$a \in \mathbb{R}$ and every vector $\underline{u} \in V$ there is a vector $a \cdot_V \underline{u} \in V$ and:
   (a) ("Associativity") For all $a, b \in \mathbb{R}$ and $\underline{u} \in V$ we have $a \cdot_V (b \cdot_V \underline{u}) = (a \cdot_V b) \cdot_V \underline{u}$.
   (b) ("One") For all $\underline{u} \in V$ we have $1 \cdot_V \underline{u} = \underline{u}$.
(3) (distributive laws)
   (a) For all $a \in \mathbb{R}$, $\underline{u}, \underline{v} \in V$ we have $a \cdot_V (\underline{u} +_V \underline{v}) = (a \cdot_V \underline{u}) +_V (a \cdot_V \underline{v})$.
   (b) For all $a, b \in \mathbb{R}$,$\underline{u} \in V$ we have $(a+b) \cdot_V \underline{u} = (a \cdot_V \underline{u}) +_V (b \cdot_V \underline{u})$.

NOTATION 2. From now on we drop the subscript $_V$ and the dot from products.

EXAMPLE 3. $\{\underline{0}\}$, $\mathbb{R}$, $\mathbb{R}^n$.

PROBLEM 4. Decide whether any of the following is a vector space. If not, identify an axiom that fails.

(1) $V = \mathbb{R}^n$, usual addition, $a\underline{v} = \underline{0}$ for all $a, \underline{v}$.
(2) $V = \mathbb{R}^n$, usual scalar multiplication,
(3) $V = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 + 2x_2 = 0 \right\}$, addition and multiplication as in $\mathbb{R}^2$
(4) $V = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 + 2x_2 = 1 \right\}$, addition and multiplication as in $\mathbb{R}^2$.

LEMMA 5. *For any nonzero $a \in \mathbb{R}^\times$ and any $\underline{b}, \underline{c} \in V$ the equation $a\underline{x} + \underline{b} = \underline{c}$ has the unique solution* $\underline{x} = a^{-1}\left(\underline{c} + \underline{b}'\right)$.

COROLLARY 6 (Zero).

(1) *If $\underline{u} + \underline{u} = \underline{u}$ then $\underline{u} = \underline{0}$.*
(2) *There is a unique zero vector.*
(3) *For all $a \in \mathbb{R}$, $a \cdot \underline{0} = \underline{0}$.*
(4) *For all $\underline{u} \in V$, $0 \cdot \underline{u} = \underline{0}$.*

COROLLARY 7 (Elementary properties).

(1) *Every vector has a unique negative, to be denoted $-\underline{u}$. We will use the shorthand* $\underline{u} - \underline{v} \stackrel{def}{=} \underline{u} + (-\underline{v})$.

(2) $(-1)\underline{u} = -\underline{u}$

EXAMPLE 8. $\mathbb{R}^X$ (in the book: $\mathcal{F}(X,\mathbb{R})$), hence $\mathbb{R}^n$, $M_{n\times m}(\mathbb{R}) \stackrel{\text{def}}{=} \mathbb{R}^{[n]\times[m]}$, $\mathbb{R}[x]$ (in the book $P(\mathbb{R})$), $\mathbb{R}[x]^{\leq n}$ (in the book $P_n(\mathbb{R})$).

1.2. Subspaces, examples

DEFINITION 9. Let $(V,+,\cdot)$ be a vector space. A subset $U \subset V$ is called a *subspace* of V if it is a vector space under the operations $+,\cdot$.

Note that every subspace must be non-empty, because it must contain a zero vector.

LEMMA 10. *Let $U \subset V$ be a subspace. Then $\underline{0}_V \in U$.*

PROOF. Let $\underline{u} \in U$ (exists since U is non-empty). Then $\underline{0}_V = 0 \cdot_V \underline{u} \in U$ by closure. □

LEMMA 11. *To check if U is a subspace of V it is necessary and sufficient to check that $\underline{0} \in U$ and that one of the following conditions:*

(1) *For every $a \in \mathbb{R}$, $\underline{u},\underline{v} \in U$ we have $a\cdot\underline{u}, \underline{u}+\underline{v} \in U$.*
(2) *For every $a,b \in \mathbb{R}$, $\underline{uv} \in U$ we have $a\underline{u}+b\underline{v} \in U$.*

PROOF. Problem Set 1 □

EXAMPLE 12 (Subspaces). (0) Every vector space V has the "trivial" subspaces $\{\underline{0}_V\}$ and V itself (check!).

(1) $\{\underline{v} \in \mathbb{R}^n \mid v_1 = 0\}$ is a subspace, but $\{\underline{v} \in \mathbb{R}^n \mid v_1 = 1\}$ is not.
(2) $\{\underline{v} \in \mathbb{R}^n \mid \sum_{i=1}^n v_i = 0\}$ is a subspace, but $\{\underline{v} \in \mathbb{R}^n \mid \sum_{i=1}^n v_i = n\}$ is not.
(3) $\left\{(x,y) \in \mathbb{R}^2 \mid y = e^x\right\}$ is not a subspace.
(4) $\{f\colon [-1,1] \to \mathbb{R} \mid f \text{ is continuous at } 0\} \subset \mathbb{R}^{[-1,1]}$.
(5) $C([a,b]) \subset \mathbb{R}^{[a,b]}$.
(6) More generally, if $I \subset \mathbb{R}$ is an interval then $C^k(I) \subset \mathbb{R}^I$ is a subspace.
(7) $\ell^\infty(X) \subset \mathbb{R}^X$ is a subspace. (PS 1)

LEMMA 13. *Let $\{U_i\}_{i\in I}$ be a family of subspaces of a space V. Then $\bigcap_i U_i$ is a subspace as well.*

PROOF. Let $W = \bigcap_i U_i$. By Lemma 11, $\underline{0} \in U_i$ for all i so $\underline{0} \in W$. Also, let $\underline{u},\underline{v} \in W$, let $a,b \in \mathbb{R}$, and consider $a\underline{u}+b\underline{v}$. For all i, $\underline{u},\underline{v} \in U_i$ since $W \subset U_i$. By the Lemma again it follows that $a\underline{u}+b\underline{v} \in U_i$. Since this is true for every i, $a\underline{u}+b\underline{v} \in W$ and we are done. □

LEMMA 14. *Let $U \subset V$ be a subspace, $n \geq 0$, $\{a_i\}_{i=1}^n \subset \mathbb{R}$, $\{\underline{u}_i\}_{i=1}^n \subset U$. Then $\sum_{i=1}^n a_i\underline{u}_i \in U$.*

PROOF. PS1. □

DEFINITION 15. A sum $\sum_{i=1}^n a_i\underline{u}_i$ is called a *linear combination.*
Let $S \subset V$ be a subset. The *span* $\mathrm{Span}(S)$ is the set of linear combinations of elements of S.

REMARK 16. Note that $\underline{0} \in \mathrm{Span}(S)$ for all S, as the value of the *empty sum.*

THEOREM 17. $\mathrm{Span}(S)$ *is a subspace of V. In fact,*

$$\mathrm{Span}(S) = \bigcap \{U \mid S \subset U \subset V \textit{ and } U \textit{ is a subspace}\}\,.$$

PROOF. By the Remark, $\mathrm{Span}(S)$ contains zero. Closure under rescaling is automatic, under addition by concatenation of sequences. That $\mathrm{Span}(S) \subset \bigcap\{U \mid S \subset U \subset V \text{ and } U \text{ is a subspace}\}$ is Lemma 14. For the reverse note that $\mathrm{Span}(S)$ is a subspace containing S. □

1.3. Linear dependence and independence

1.3.1. Linear dependence and independence. Fix a vector space V and a set $S \subset V$.

DEFINITION 18. Say that $\underline{v} \in V$ *depends linearly on* S if there are $\{\underline{v}_i\}_{i=1}^n \subset S$ and scalars $\{a_i\}_{i=1}^n$ such that $\sum_{i=1}^n a_i \underline{v}_i = \underline{v}$. Otherwise say that $\underline{v}$ is *linearly independent* of S.

EXAMPLE 19 (Linear dependence). (0) The zero vector depends on every set S (via the empty combination)
(1) No non-zero vector depends on $\{\underline{0}\}$.
(2) $\underline{v}$ depends on S iff $\underline{v} \in \mathrm{Span}(S)$.
(3) (Useful to prove independence) If can find a subspace W such that $S \subset W$ and $\underline{v} \notin W$ then $\underline{v}$ is independent of S.

DEFINITION 20. The set S is *linearly dependent* if some $\underline{v} \in S$ depends on $S \setminus \{\underline{v}\}$; linearly independent otherwise.

LEMMA 21. *S is linearly independent iff whenever $\{\underline{v}_i\}_{i=1}^n \subset S$ are distinct and $\{a_i\}_{i=1}^n \subset \mathbb{R}$ are such that $\sum_{i=1}^n a_i \underline{v}_i = \underline{0}$ we have $a_i = 0$ for all i.*

PROOF. Solve linear dependence for a vector with a non-zero coefficient. □

1.3.2. Bases.

LEMMA 22. *S linearly independent and $v \notin \mathrm{Span}(S)$ implies $S \cup \{v\}$ independent.*

PROOF. Suppose $\sum_{i=1}^n a_i \underline{v}_i + a\underline{v} = \underline{0}$ where $\{\underline{v}_i\}_{i=1}^n \subset S$ are distinct. If $a \neq 0$ we'd have $\underline{v} = \sum_{i=1}^n (-a^{-1} a_i) \underline{v}_i \in \mathrm{Span}(S)$, a contradiction. Thus $a = 0$. Then $\sum_{i=1}^n a_i \underline{v}_i = \underline{0}$ so all the other $a_i = 0$ by independence of S. □

COROLLARY 23. *S maximal linearly independent then spanning.*

PROOF. Contrapositive of Lemma: if not spanning, then there is a vector independent of S. □

DEFINITION 24. A spanning independent set is called a *basis*.

ALGORITHM 25. *Find bases by adding vectors.*

LEMMA 26. *S spanning and minimal then independent.*

PROOF. If there is a dependence then can remove a vector without affecting span. □

ALGORITHM 27. *Find bases by subtraction.*

COROLLARY 28. *Every finitely generated vector space has a basis.*

AXIOM 29. *Every vector space has a basis.*

1.3.3. Dimension. Standard basis of $\mathbb{R}^n$; bases for space of polynomials. Bases for space of solutions of system of equations.

PROPOSITION 30 (Steinitz replacement lemma). *Let $S \subset V$ be a generating set, and let $T \subset V$ be linearly independent. Suppose that $T \not\subset S$ and let $\underline{u} \in T \setminus S$. Then there is$\underline{v} \in S \setminus T$ so that $S \setminus \{\underline{v}\} \cup \{\underline{u}\}$ is also a generating set.*

PROOF. Let $\underline{u} \in T \setminus S$. Then $\underline{u} \in \mathrm{Span}(S)$ and therefore there are $\{\underline{v}_i\}_{i=1}^n \subset S$ and $\{a_i\}_{i=1}^n \subset \mathbb{R}$ such that $\underline{u} = \sum_{i=1}^n a_i \underline{v}_i$. Suppose that for every i, $a_i = 0$ or $\underline{v}_i \in T$. Then, omitting the zero contributions, we'd have that $\underline{u}$ depends on $T \setminus \{\underline{u}\}$ ($\underline{v}_i \in S$ so they aren't equal to vu), contradicting the independence of T). It follows that there is j for which $a_j \neq 0$ and $\underline{v}_j \notin T$. We then have

$$\underline{v}_j = \sum_{\substack{i=1 \\ i \neq j}}^n (-a_j^{-1} a_i) \underline{v}_i + a_j^{-1} \underline{u} .$$

It follows that $\underline{v}_j \in \mathrm{Span}\left(S \setminus \left\{\underline{v}_j\right\} \cup \{\underline{u}\}\right)$, and hence that $V \supset \mathrm{Span}\left(S \setminus \left\{\underline{v}_j\right\} \cup \{\underline{u}\}\right) \supset \mathrm{Span}(S) = V$ so the span is as claimed. □

THEOREM 31. *Let $S \subset V$ be a finite generating set, and let $T \subset V$ be linearly independent. Then $\#T \leq \#S$.*

PROOF. We repeatedly replace vectors of T into S until $T \subset S$.

Formally, let $A \subset S$ be minimal such that there is $A' \subset T$ such that $A \cup A'$ is a generating set of size at most $\#S$ (such A exist – take $A = S$). Then A is disjoint from T (otherwise we could reduce A and increase A' by moving the vectors over). If $A' = T$, $\#T \leq \#(A \cup A') \leq \#S$, so suppose that $A' \neq T$. Then there is $\underline{u} \in T \setminus A'$ so that $\underline{u} \notin A \cup A'$ and by thte Proposition there is $\underline{v} \in (A \cup A') \setminus T = A$ so that $(A \setminus \{\underline{v}\}) \cup (A' \cup \{\underline{u}\})$ is also generating, of size at most $\#S$. This contradicts the minimality of A. □

COROLLARY 32. *Let V be finitely generated. Then any two bases of B have the same size.*

DEFINITION 33. Let V be finitely generated. Them $\dim V$ is the size of any basis of V (these exist by Corollary 28).

1.4. Geometric picture

The Euclidean plane.

(1) vectors, addition and the parallelogram law.
(2) points and lines: subspaces and affine subspaces

Euclidean 3-space. points, lines, planes.

Rotations in the plane. Given two vectors $\underline{a} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \underline{b} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ in the plane, their *Euclidean distance* is defined by

$$\mathrm{dist}(\underline{a}, \underline{b}) = \left((x_2 - x_1)^2 + (y_2 - y_1)^2\right)^2 .$$

Note that the distance only depends on the *displacement* $\underline{b} - \underline{a}$.

The *Euclidean plane* is $\mathbb{E}^2$ the plane $\mathbb{R}^2$ equipped with this distance function.

DEFINITION 34. A *Euclidean isometry* is a distance-preserving function $f\colon \mathbb{E}^2 \to \mathbb{E}^2$. The set of all such functions is called the *isometry group* or the *Euclidean group*.

This certainly includes the *translations*

$$T_{\underline{a}}(\underline{v}) = \underline{v} + \underline{a}.$$

Up to translation we can assume that an isometry preserves the origin, so let R be a Euclidean isometry such that $R\underline{0} = \underline{0}$.

- Goal: classify all such maps.

Let R be a Euclidean isometry such that $R\underline{0} = \underline{0}$.

We first note that $\mathrm{dist}(\underline{e}_1, \underline{0}) = \mathrm{dist}(\underline{e}_2, \underline{0}) = 1$ (they are "unit vectors"). Thus

$$\mathrm{dist}(R\underline{e}_1, \underline{0}) = \mathrm{dist}(R\underline{e}_1, R\underline{0}) = \mathrm{dist}(\underline{e}_1, \underline{0}) = 1 .$$

Thus $R\underline{e}_1 = \begin{pmatrix} x \\ y \end{pmatrix}$ where $x^2 + y^2 = 1$, and there is a unique angle θ so that $R\underline{e}_1 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$.

For the same reason we have $R\underline{e}_2 = \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}$ for some angle ϕ. What is the distance between two unit vectors at these angles?

$$\begin{aligned}\operatorname{dist}\left(\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix},\begin{pmatrix}\cos\phi\\ \sin\phi\end{pmatrix}\right) &= \sqrt{(\cos\phi-\cos\theta)^2+(\sin\phi-\sin\theta)^2}\\ &= \sqrt{\cos^2\phi+\cos^2\theta-2\cos\phi\cos\theta+\sin^2\phi+\sin^2\theta-2\sin\phi\sin\theta}\\ &= \sqrt{2}\sqrt{1-\cos(\theta-\phi)}\end{aligned}$$

(this is called the "law of cosines"). But we must have

$$\operatorname{dist}(R\underline{e}_1,R\underline{e}_2)=\operatorname{dist}(\underline{e}_1,\underline{e}_2)=\sqrt{2}.$$

Since the distance between two unit vectors depends only on the angle between them, it follows that a distance -preserving function must preserve the angles. So in our case $\cos(\phi-\theta)=0$ and hence $\phi-\theta=\pm\frac{\pi}{2}$. We consider the case $\phi=\theta+\frac{\pi}{2}$ first, where

$$R\underline{e}_2=\begin{pmatrix}\cos(\theta+\frac{\pi}{2})\\ \sin(\theta+\frac{\pi}{2})\end{pmatrix}=\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}.$$

Now the formula above shows that the distance between two unit vectors only dpeends on the angle between them, so our map R must *preserve angles*. Thus if we take the vector $\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix}$ (which has angle α to $\underline{e}_1$ and $\frac{\pi}{2}-\alpha$ to $\underline{e}_2$) it must map to the unit vector at angle $\alpha+\theta$. We conclude that

$$\begin{aligned}R\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix} &= \begin{pmatrix}\cos(\alpha+\theta)\\ \sin(\alpha+\theta)\end{pmatrix}\\ &= \begin{pmatrix}\cos\theta\cos\alpha-\sin\theta\sin\alpha\\ \sin\theta\cos\alpha+\cos\theta\sin\alpha\end{pmatrix}.\end{aligned}$$

Now any vector can be recaled to be a unit vector, and R must respect this scaling (since that's the distance to the origin). So for any vector x,y let $r=\sqrt{x^2+y^2}=\operatorname{dist}\left(\begin{pmatrix}x\\ y\end{pmatrix},\underline{0}\right)$ and choose α so that $\frac{y}{x}=\tan\alpha$. Then

$$\begin{aligned}R\begin{pmatrix}x\\ y\end{pmatrix}=rR\begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix} &= r\begin{pmatrix}\cos\theta\cos\alpha-\sin\theta\sin\alpha\\ \sin\theta\cos\alpha+\cos\theta\sin\alpha\end{pmatrix}\\ &= \begin{pmatrix}\cos\theta\cdot r\cos\alpha-\sin\theta\cdot r\sin\alpha\\ \sin\theta\cdot r\cos\alpha+\cos\theta\cdot r\sin\alpha\end{pmatrix}\\ &= \begin{pmatrix}\cos\theta\cdot x-\sin\theta\cdot y\\ \sin\theta\cdot x+\cos\theta\cdot y\end{pmatrix}\end{aligned}$$

(check for $(x,y)=(1,0)$ and $(x,y)=(0,1)$!).

OBSERVATION 35. *Surprise: R is a* linear *function! the coordinates of $R\underline{v}$ are* linear *in the coordinates of $\underline{v}$. The investigation of such maps will be the next topic.*

CHAPTER 2

Linear Transformations

2.1. Linear Transformations

2.1.1. Definition; basic properties. The key definition for this course:

DEFINITION 36. Let U,V be vector spaces. A function $T\colon U \to V$ is a *linear transformation* (or *linear map* or *homomorphism of vector spaces*) if for all $\underline{u},\underline{v} \in U$ and scalars a,b we have

$$T\left(a\underline{u}+b\underline{v}\right) = aT\underline{u}+bT\underline{v}\,.$$

REMARK 37. Note the notation for applying the function: no parentheses around the argument.

EXAMPLE 38 (Liner maps). (0) The zero map $f(\underline{u}) = \underline{0}_V$ for all $\underline{u}$ is linear.

(1) Recaling: for $a \in \mathbb{R}$let $Z_a\colon U \to U$ be given by $Z_a\underline{u} = a\underline{u}$ (linearity follows from axioms of vector space).
(2) Identity map: $\mathrm{Id}\,\underline{u} = \underline{u}$ ($\mathrm{Id} = Z_1$).
(3) Calculus: $\frac{d}{dx}$, $f \mapsto (x \mapsto \int_a^x f(t)\,\mathrm{d}t)$, say on $C^\infty(a,b)$.
(4) The *shift* on $\mathbb{R}^{\mathbb{N}}$ and $\mathbb{R}^{\mathbb{Z}}$.
(5) Linear functionals.
    (a) Evaluation of functions: $\delta_x(f) \stackrel{\text{def}}{=} f(x)$ as a map $\delta_x\colon \mathbb{R}^X \to \mathbb{R}$.
    (b) Limits of sequences: $\lim\colon c \to \mathbb{R}$.

LEMMA 39. *Let T be a linear map. Then*

(1) $T\underline{0} = \underline{0}$.
(2) $T(-\underline{u}) = -T\underline{u}$.

PROOF. Either multiplication by scalars ($0,-1$ respectively) or use that $T\underline{0}+T\underline{0} = T(\underline{0}+\underline{0}) = T\underline{0}$ and that $T\underline{u}+T(-\underline{u}) = T\underline{0} = \underline{0}$. ☐

LEMMA 40. $T\left(\sum_{i=1}^n a_i\underline{v}_i\right) = \sum_{i=1}^n a_iT\underline{v}_i$.

PROOF. Induction on n. ☐

COROLLARY 41. *For any $S \subset U$, $T(\mathrm{Span}(S)) = \mathrm{Span}\left(T(S)\right)$.*

2.1.2. Range and kernel; rank-nullity.

PROPOSITION 42. *Let $T\colon U \to V$ be linear.*

(1) *Let $W \subset U$ be a subspace. Then the* image $T(W) = \{T\underline{w} \mid \underline{w} \in U\}$ *is a subspace of V.*
(2) *Let $X \subset V$ be a subspace. Then the* inverse image $T^{-1}X = \{\underline{u} \in U \mid T\underline{u} \in X\}$ *is a subspace of U.*

DEFINITION 43. $\mathrm{Image}(T) = \mathrm{Im}(T) \stackrel{\text{def}}{=} T(V)$ is called the *image* of T, $\mathrm{Ker}(T) = \{\underline{u} \in U \mid T\underline{u} = \underline{0}\} = T^{-1}\left(\{\underline{0}\}\right)$ is called the *kernel* of T.

COROLLARY 44. *The Kernel of T is a subspace of U, the Image of T is a subspace of V.*

THEOREM 45 (Rank-nullity). *Let U be finite-dimensional and let $T\colon U \to V$ be linear. Then* $\dim \mathrm{Ker}\, T + \dim \mathrm{Im}\, T = \dim U$.

EXAMPLE 46 (PS3). The case of a non-zero linear funtional, where $\dim \operatorname{Im} T = 1$.

PROOF. Let $n = \dim U$, $s = \dim \operatorname{Ker} T$, $r = \dim \operatorname{Im} T$. Let $\{\underline{u}_i\}_{i=1}^s \subset \operatorname{Ker} T$ and $\{\underline{v}_j\}_{j=1}^r \subset \operatorname{Im} T$ be bases for their respective spaces. For each $\underline{v}_j$ choose $\underline{w}_j \in U$ such that $T\underline{w}_j = \underline{v}_j$. We claim that $\{\underline{u}_i\}_{i=1}^s \cup \{\underline{w}_j\}_{j=1}^r$ is a basis for U, so that $\dim U = s + r$ as claimed.

(1) To see that they span U, let $\underline{u} \in U$. Then $T\underline{u} \in \operatorname{Im} T$ so there are b_j for which $T\underline{u} = \sum_{j=1}^r b_j \underline{v}_j = \sum_{j=1}^r b_j T\underline{w}_j$. Then
$$T\left(\underline{u} - \sum_{j=1}^r b_j \underline{w}_j\right) = \underline{0},$$
that is $\underline{u} - \sum_{j=1}^r b_j \underline{w}_j \in \operatorname{Ker} T$. It follows that there are a_i for which $\underline{u} - \sum_{i=1}^r b_j \underline{w}_j = \sum_{i=1}^s a_i \underline{u}_i$ and then
$$\underline{u} = \sum_{i=1}^s a_i \underline{u}_i + \sum_{i=1}^r b_j \underline{w}_j .$$

(2) To see that the vectors are independent, let $\{a_i\}_{i=1}^s$ and $\{b_j\}_{j=1}^r$ be such that $\sum_{i=1}^s a_i \underline{u}_i + \sum_{i=1}^r b_j \underline{w}_j = \underline{0}$. Applying T to both sides, we have
$$\begin{aligned} \underline{0} &= \sum_{i=1}^s a_i T\underline{u}_i + \sum_{i=1}^r b_j T\underline{w}_j \\ &= \sum_{i=1}^r b_j \underline{v}_j . \end{aligned}$$
The independence of $\{\underline{v}_j\}_{j=1}^r$ now shows that $b_j = 0$ for all j. Accordingly we have $\sum_{i=1}^s a_i \underline{u}_i = \underline{0}$ and since the $\underline{u}_i$ are also independent we see that the $a_i = 0$ as well.

□

DEFINITION 47. $\dim \operatorname{Ker} T$ is called the *nullity* of T, sometimes denoted $n(T)$. $\dim \operatorname{Im}(T)$ is called the *rank* of T, will be denoted $r(T)$.

COROLLARY 48. $r(T) \leq \dim U$. *Suppose* $\dim V < \dim U$. *Then* $\dim \operatorname{Ker} T > 0$.

EXAMPLE 49. Every system of homogenous linear equations with more unknowns than equations has a non-trivial solution.

7

2.2. Matrices

DEFINITION 50. $\operatorname{Hom}(U,V)$ is the space of linear maps from U to V.

LEMMA 51. $\operatorname{Hom}(U,V)$ *is a vector space under pointwise addition and scalar multiplication.*

LEMMA 52. *Let* $T \in \operatorname{Hom}(\mathbb{R}^m, \mathbb{R}^n)$. *Then there are numbers* $a_{i,j}$ *for* $1 \leq i \leq n$, $1 \leq j \leq m$ *such that* $(T\underline{x})_i = \sum_{j=1}^m a_{ij} x_j$, *that is such that*
$$T\underline{x} = \begin{pmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m \\ \vdots \\ a_{i,1}x_1 + \cdots + a_{i,m}x_m \\ \vdots \\ a_{n,1}x_1 + \cdots + a_{n,m}x_m \end{pmatrix} .$$

PROOF. For each j define $a_{i,j}$ by $T\underline{e}_j = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{i,j} \\ \vdots \\ a_{n,j} \end{pmatrix}$. Use linearity. □

DEFINITION 53. $M_{n,m}(\mathbb{R}) = \mathbb{R}^{[n]\times[m]}$. Given $A \in M_{n,m}(\mathbb{R})$ write A_{ij} or a_{ij} for the entries.

LEMMA 54. *Given $A \in M_{n,m}(\mathbb{R})$, the map* $L_A\underline{x} \stackrel{def}{=} \begin{pmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m \\ \vdots \\ a_{i,1}x_1 + \cdots + a_{i,m}x_m \\ \vdots \\ a_{n,1}x_1 + \cdots + a_{n,m}x_m \end{pmatrix}$ *is linear.*

PROPOSITION 55. *The map $A \mapsto L_A$ is an isomorphism $M_{n,m}(\mathbb{R}) \to \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$.*

- Philosophy: describe linear maps and calculate using matrices.

Now let U,V be vector spaces, and ordered bases $\{\underline{u}_j\}_{j=1}^m$ of U, $\{\underline{v}_i\}_{i=1}^n$ of V (here $m = \dim U$, $n = \dim V$). Given $T \in \mathrm{Hom}(U,V)$ define a matrix $A \in M_{n,m}(\mathbb{R})$ by setting $T\underline{u}_j = \sum_{i=1}^n a_{ij}\underline{v}_i$.

DEFINITION 56. Call this *the matrix of T with respect to the ordered bases* $\{\underline{u}_j\}_{j=1}^m$, $\{\underline{v}_i\}_{i=1}^n$.

LEMMA 57. *This is a well-defined linear map* $\mathrm{Hom}(U,V) \to M_{n,m}(\mathbb{R})$.

To see that this is an isomorphism we construct the inverse map: given A we define T by $T\left(\sum_j x_j\underline{u}_j\right) = \sum_{i=1}^n \left(\sum_{j=1}^m a_{ij}x_j\right)\underline{v}_i$ (compare with above!)

REMARK 58. Existence of a_{ij} uses that $\underline{v}_i$ are spanning, well-defined uses that they are independent. Where did we use info about $\underline{u}_j$?

PROPOSITION 59. *Let U be a vector space with basis B, V another vector space and $f\colon B \to V$. Then there is a unique linear map $T\colon U \to V$ extending f.*

PROOF. Suppose T is such a map. For each $\underline{u} \in U$ we can write $\underline{u} = \sum_{i=1}^n x_i\underline{u}_i$ for some $\underline{u}_i \in B$ and $x_i \in \mathbb{R}$ since B is spanning. Then

$$T\underline{u} = T\left(\sum_{i=1}^n x_i\underline{u}_i\right) = \sum_{i=1}^n x_i T\underline{u}_i = \sum_{i=1}^n x_i f(\underline{u}_i)$$

so T, if it exists, is unique. Conversely, define T by the relation above. This is OK since every $\underline{u}$ has a *unique* representation in the basis. Linearity easy to check. □

COROLLARY 60. *We have an isomorphism* $\mathrm{Hom}(U,V) \simeq M_{n,m}(\mathbb{R})$.

COROLLARY 61. *Since $M_{n,m}(\mathbb{R}) \simeq R^{nm}$,$\dim \mathrm{Hom}(U,V) = \dim M_{n,m}(\mathbb{R}) = nm = \dim U \cdot \dim V$.*

2.3. Composing linear maps, multiplying matrices, space of endomorphisms

- Heisenberg discovers formula for matrix multiplication.
- Challenge: show associativity
- Go back: where did this come from?
- Compose linear maps

2.4. Linear equations

2.4.1. What is a linear equation?

DEFINITION 62. A *linear equation* is an equation $T\underline{x} = \underline{b}$ where $T \in \mathrm{Hom}(U,V)$ and $\underline{b} \in V$. If $\underline{b} = \underline{0}$ we call the equation *homogenous.*

EXAMPLE 63 (Linear equations).

(1) $\begin{cases} 2x + y &= 1 \\ x + 2y &= 3 \end{cases}$

(2) $\frac{df}{dx} = e^{-x^2}$

(3) $-\frac{1}{2}\psi''(x) + \frac{1}{2}x^2\psi(x) = E\psi(x)$

(4) $F_{n+1} = F_n + F_{n-1}$

REMARK 64 (Linearity).

- The set of solutions to a homogenous equation is the kernel of a linear map, so any linear combination of solution is again a solution. In phsyics this called the *principle of superposition.*
- Recognizing that an equation is linear is important.
- For diff eq the choice of function space in which to define the equation is technical.

LEMMA 65. *The equation has solutions iff* $\underline{b} \in \mathrm{Im}\,T$*. If* $\underline{v}_0$ *is any solution then the set of solutions is* $\underline{v}_0 + \mathrm{Ker}\,T$.

DEFINITION 66. If $E \in \mathrm{Hom}(U,U)$ is invertible we call the equations $T\underline{x} = \underline{b}$ and $ET\underline{x} = E\underline{b}$ *equivalent.*

LEMMA 67. *Equivalent equations have same solutions, this is an equivalence relation. Equivalence preserves image, hence rank=column rank.*

2.4.2. Gaussian Elimination. Now concentrate on the case $T = L_A \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n)$.

NOTATION 68. Augmented matrix

LEMMA 69. *Solution to diagonal equations*

Better:

LEMMA 70. *Solution to row echelon form.*

DEFINITION 71. An *elementary row operation* is rescaling a row, or adding a multiple of one row to another row.

LEMMA 72. *These are equivalences. They preserve row space, hence row rank as well.*

PROOF. Achieved by multiplication by $\mathrm{diag}(1,\ldots,d_i,\ldots,1)$ $(d_i \neq 0)$ and by $I_n + cE^{ij}$, which are invertible. □

ALGORITHM 73. *Find first column with a non-zero entry, exchange rows to make it in first row. Subtract multiples to make zeroes below. Find next column ...*

COROLLARY 74. *Every equation* $A\underline{x} = \underline{b}$ *is equivalent to an equation* $(MA)\underline{x} = M\underline{b}$ *where* M *is a product of elementary matrices and* MA *is in row eschelon form (or row-reduced form).*

DEFINITION 75. Pivot.

In row-reduced form, a variable without pivot is called *free*. General solution obtained by arbitrarily valuing the free variables (gives new proof of dimension formula).

2.4.3. Inverting matrices using Gaussian Elimination. Start with the pair of matrices $A_0 = A, B_0 = I$. At each step multiply both sides by an elementary matrix E_n to get $A_{n+1} = E_{n+1}A_n$, $B_{n+1} = E_{n+1}B_n$.

OBSERVATION 76. *The product $B_n^{-1}A_n$ is an* invariant *of the algorithm: suppose that $B_{n+1}^{-1}A_{n+1} = B_n^{-1}E_{n+1}^{-1}E_{n+1}A_n = B_n^{-1}A_n$.*

Now $B_0^{-1}A_0 = A$, so if $A_n = I$ we'll get $B_n^{-1} = A$, that is $B_n =$ also.

CHAPTER 3

Determinants

3.1. The determinant of a matrix

Notation: for a square matrix $A \in M_n(\mathbb{R})$ write a_{ij} for the entries, A_{ij} for the *minor*, the matrix $A_{ij} \in M_{n-1}(\mathbb{R})$ obtained by deleting the ith row and jth column.

DEFINITION 77. If $A \in M_1(\mathbb{R})$ set $\det A = a_{11}$. If $A \in M_n(\mathbb{R})$ for $n \geq 1$ set $\det A = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}$.

EXAMPLE 78. $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$.

EXAMPLE 79. $\det(\mathrm{diag}(a_1, \ldots, a_n)) = a_1 \det(\mathrm{diag}(a_2, \ldots, a_n))$ so by induction $\det(\mathrm{diag}(a_1, \ldots, a_n)) = \prod_{i=1}^{n} a_i$.

DEFINITION 80. A matrix A is called *lower-triangular* if its entries above the main diagonal are zero: $a_{ij} = 0$ if $j > i$.

LEMMA 81. *Let A be lower-triangular. Then* $\det A = \prod_{i=1}^{n} a_{ii}$.

PROOF. Every $A \in M_1(\mathbb{R})$ is lower triangular and the formula holds by definition. Suppose this holds for $n \times n$ matrices and let $A \in M_{n+1}(\mathbb{R})$ be upper-triangular. Then every entry except a_{11} in the first row is zero, so $\det A = a_{11} \det A_{11}$. Now A_{11} is a lower-triangular matrix as well, and its diagonal consists of $a_{22}, \ldots, a_{n+1,n+1}$. Thus $\det A = a_{11} \prod_{i=2}^{n+1} a_{ii} = \prod_{i=1}^{n+1} a_{ii}$. □

3.2. Determinants of linear maps

3.2.1. Area of paralellograms in the plane.

3.2.2. Area forms in the plane.

Let V be a two-dimensional vector space.

DEFINITION 82. A function $A\colon V^2 \to \mathbb{R}$ is an *area form* if

(1) ("bilinearity") A is linear in each argument separately.
(2) ("alternativity") For all $\underline{u} \in V$, $A(\underline{u}, \underline{u}) = 0$.

LEMMA 83. *The space of area forms is a subspace of* $\mathbb{R}^{V^2}$.

REMARK 84. Warning about multilinearity: this means $A(\underline{a}+\underline{b}, \underline{c}+\underline{d}) = A(\underline{a}, \underline{c}) + A(\underline{a}, \underline{d}) + A(\underline{b}, \underline{c}) + A(\underline{b}, \underline{d}) \neq A(\underline{a}, \underline{c}) + A(\underline{b}, \underline{d})$. Note that multiplication is multilnear.

LEMMA 85. *Let V be a vector space of any dimension. A bilinear form $A\colon V^2 \to \mathbb{R}$ is alternating iff its* antisymmetric*, in that $A(\underline{u}_1, \underline{u}_2) = -A(\underline{u}_2, \underline{u}_1)$ for all $\underline{u}_1, \underline{u}_2 \in V$.*

PROOF. Suppose A is alternating and consider $A(\underline{u}_1 + \underline{u}_2, \underline{u}_1, \underline{u}_1 + \underline{u}_2)$. We have

$$\begin{aligned} 0 &= A(\underline{u}_1 + \underline{u}_2, \underline{u}_1, \underline{u}_1 + \underline{u}_2) \\ &= A(\underline{u}_1, \underline{u}_1) + A(\underline{u}_1, \underline{u}_2) + A(\underline{u}_2, \underline{u}_1) + A(\underline{u}_2, \underline{u}_2) \\ &= A(\underline{u}_1, \underline{u}_2) + A(\underline{u}_2, \underline{u}_1). \end{aligned}$$

Conversely, suppose A is antisymmetric. Then exchanging the two arguments gives $A(\underline{u}, \underline{u}) = -A(\underline{u}, \underline{u})$ and hence $2A(\underline{u}, \underline{u}) = 0$ and $A(\underline{u}, \underline{u}) = 0$. □

REMARK 86. The argument depended on $1+1=2$ being non-zero. There are situations where this isn't the case, which is why alternativity is the stronger hypothesis.

Now let $\{\underline{v}_1,\underline{v}_2\}$ be an ordered basis for V. Then given $(\underline{u}_1,\underline{u}_2)\in V^2$ there are a_{ij} such that $\underline{u}_1 = a_{11}\underline{v}_1+a_{12}\underline{v}_2$ and $\underline{u}_2=a_{21}\underline{v}_1+a_{22}\underline{v}_2$ so that for any area form A, by the distributive remark above

$$\begin{aligned} A(\underline{u}_1,\underline{u}_2) &= a_{11}a_{21}A(\underline{v}_1,\underline{v}_1)+a_{11}a_{22}A(\underline{v}_1,\underline{v}_2)+a_{12}a_{21}A(\underline{v}_2,\underline{v}_1)+a_{12}a_{22}A(\underline{v}_2,\underline{v}_2) \\ &= a_{11}a_{22}A(\underline{v}_1,\underline{v}_2)-a_{12}a_{21}A(\underline{v}_1,\underline{v}_2) \\ &= [a_{11}a_{22}-a_{12}a_{21}]A(\underline{v}_1,\underline{v}_2)\,. \end{aligned}$$

It follows that A is determined by the single number $A(\underline{v}_1,\underline{v}_2)$. Since evaluation is a linear map on functions, we have shown that the following linear map is injective:

$$\begin{aligned} \{\text{Area forms}\} &\to \mathbb{R} \\ A &\mapsto A(\underline{v}_1,\underline{v}_2)\,. \end{aligned}$$

COROLLARY 87. *The space of area forms is at most* 1*-dimensional.*

LEMMA 88. *The map* $\left(\begin{pmatrix} a \\ c\end{pmatrix},\begin{pmatrix} b \\ d\end{pmatrix}\right)\mapsto ad-bc$ *is a non-zero area form on* $\mathbb{R}^2$.

COROLLARY 89. *The space of area form on a 2-dimensional space is exactly 1-dimensional.*

Now let $T\in \mathrm{End}(V)$. Then for any area form A, $(\underline{u}_1,\underline{u}_2)\mapsto A(T\underline{u}_1,T\underline{u}_2)$ is also an area form (this is checked in PS8), and this map on area forms is linear. Since that space is 1d, there is c such that $A(T\underline{u}_1,T\underline{u}_2)=cA(\underline{u}_1,\underline{u}_2)$ for all area forms A and vectors $\underline{u}_1,\underline{u}_2$.

DEFINITION 90. The *determinant* $\det T$ is that number.

EXERCISE 91. Let T be represented by the matrix A. Then $\det T=\det A$.

3.2.3. Volume forms.

Fix an n-dimensional vector space V.

DEFINITION 92. Let $f\colon V^n\to\mathbb{R}$ be a function.

(1) Call f *multi-linear* if it's linear in every argument separately.
(2) Call f alternating if exchanging any two arguments reverses the sign.

Call alternating multi-linear functions *volume forms.*

LEMMA 93. *Let* $f\colon V^k\to V^k$ *be multilinear. Then* f *is alternating iff* f *has the property that whenever some argument vanishes [or two are equal]* f *vanishes.*

PROOF. Given $i\neq j$ $g(\underline{u}_i,\underline{u}_j)\stackrel{\text{def}}{=} f(\underline{a}_1,\ldots,\underline{a}_{i-1},\underline{u}_i,\underline{a}_{i+1},\ldots,\underline{a}_{j-1},\underline{u}_j,\underline{a}_{j+1},\ldots,\underline{a}_n)$ is a bilinear function. Now apply Lemma 85. □

Now fix an ordered basis $\{\underline{v}_i\}_{i=1}^n\subset V$. Then given $(\underline{u}_1,\ldots,\underline{u}_n)\in V^n$ let a_{ij} be such that $\underline{u}_i=\sum_{j=1}^n a_{ij}\underline{v}_j$. Then for any volume form f we have

$$\begin{aligned} f(\underline{u}_1,\ldots,\underline{u}_n) &= f\left(\sum_{j_1=1}^n a_{1,j_1}\underline{v}_{j_1},\cdots,\sum_{j_n=1}^n a_{n,j_n}\underline{v}_{j_n}\right) \\ &= \sum_{(j_1,\ldots,j_n)\in\{1,\cdots,n\}^n}\left(\prod_{\ell=1}^n a_{\ell,j_\ell}\right) f\left(\underline{v}_{j_1},\cdots,\underline{v}_{j_n}\right) \\ &= \sum_{\sigma\text{ a rearrangement}}\left(\prod_{\ell=1}^n a_{\ell,\sigma(\ell)}\right) f\left(\underline{v}_{\sigma(1)},\cdots,\underline{v}_{\sigma(n)}\right) \\ &= \sum_{\sigma}(-1)^{\sigma}\left(\prod_{\ell=1}^n a_{\ell,\sigma(\ell)}\right) f(\underline{v}_1,\ldots,\underline{v}_n) \end{aligned}$$

so again f is determined by $f(\underline{v}_1,\ldots,\underline{v}_n)$.

COROLLARY 94. *The space of volume forms is a most* 1*-dimensional.*

THEOREM 95. *The map* $f \mapsto f(\underline{v}_1,\cdots,\underline{v}_n)$ *is an isomorphism of vector spaces.*

PROOF. We show that the map $A \mapsto \det A$ of Definition 77 is a non-zero volume form on $\mathbb{R}^n$ (thought of as a function of the columns), by induction on n.

The case $n = 1$ is easy, and $n = 2$ was done above. Now try $n+1$. We first show that the function is multilinear: let $A \in M_{n+1}(\mathbb{R})$ and suppose that $a_{i,k} = \beta b_i + \gamma c_i$ for some particular k. Let B be the matrix where every column is the same as A except the kth column is (b_i) and similarly define C. Then for $j \neq k$, the minors A_{1j}, B_{1j}, C_{1j} have all columns the same except the one coming from the kth column – that column in A_{1j} is the *combination* of the respective columns in B_{1j}, C_{1j}. By induction $\det(A_{1j}) = \beta \det B_{1j} + \gamma \det C_{1j}$. For $j = k$ we see that A, B, C all have the same minor. It follows that

$$\begin{aligned}
\det A &= \sum_{j=1}^{n+1} (-1)^{1+j} a_{1j} \det A_{1j} \\
&= \sum_{j\neq k} (-1)^{1+j} a_{1j} \left(\beta \det B_{1j} + \gamma \det C_{1j}\right) + (-1)^{1+k}(\beta b_1 + \gamma c_1) \det A_{1k} \\
&= \beta \left[\sum_{j\neq k} (-1)^{1+j} a_{1j} \det B_{1j} + (-1)^{1+k} b_1 \det B_{1k}\right] + \gamma \left[\sum_{j\neq k} (-1)^{1+j} a_{1j} \det C_{1j} + (-1)^{1+k} c_1 \det C_{1k}\right] \\
&= \beta \det B + \gamma \det C .
\end{aligned}$$

Now suppose that the kth and ℓth columns of A are equal. Then the same is true in every minor A_{1j} unless $j = k$ or $j = \ell$. It follows that

$$\det A = (-1)^{1+k} a_{1k} \det A_{1k} + (-1)^{1+\ell} a_{1\ell} \det A_{1\ell} .$$

Now $a_{1k} = a_{1\ell}$, and $A_{1\ell}$ is obtained from A_{1k} by repeatedly taking the ℓth column and exchanging it with its left neighbour $\ell - k - 1$ times. It follows that

$$\det A = a_{1k}(-1)^{1+k} \left(\det A_{1k} + (-1)^{\ell-k}(-1)^{\ell-k-1} \det A_{1k}\right) = 0 .$$

Finally it's easy to check by induction that $\det I_n = 1$. □

REMARK 96. The textbook shows by a very similar induction that this is also a volume form when considered as a function of the *rows*.

COROLLARY 97. *Let* $T\colon V \to V$*. Then* $f \mapsto f(T\cdot,\cdots,T\cdot)$ *is a linear map on a 1d space, hence of the form* $f \mapsto cf$*.*

DEFINITION 98. Call this constant the *determinant* of T.

PROPOSITION 99. *Let* A *be a matrix of* T *wrt the basis* $\{\underline{v}_i\}_{i=1}^n \subset V$*. Then* $\det A = \det T$*. Here* $\det A$ *is given by Definition 77 while* $\det T$ *is given by Definition 98.*

PROOF. $f \in (V^n)^*$ be the volume form defined as follows: given $\{\underline{u}_i\}_{i=1}^n$ let a_{ij} be such that $\underline{u}_j = \sum_{i=1}^n a_{ij} \underline{v}_i$ and let $f(\underline{u}_1,\ldots,\underline{u}_n) = \det\left((a_{ij})_{i,j=1}^n\right)$ (check that this is a volume form!), Then $f(\underline{v}_1,\ldots,\underline{v}_n) = \det(I_n) = 1$. We thus have

$$\det T = (\det T) f(\underline{v}_1,\ldots,\underline{v}_n) = f(T\underline{v}_1,\ldots,T\underline{v}_n) = \det A$$

Since the matrix (a_{ij}) such that $T\underline{v}_j = \sum_i a_{ij} \underline{v}_i$ is exactly the matrix A of T in the given basis. □

3.3. Properties of determinants

Fix an n-dimensional space V.

PROPOSITION 100. *For all* $T,S \in \mathrm{End}(V)$

(1) $\det \mathrm{Id}_V = 1$.
(2) $\det(TS) = \det T \cdot \det S$.
(3) $\det T \neq 0$ *iff* T *is invertible*

PROOF. (1) Any volume form is unchanged by composition with the identity transformation. For (2) let f be a volume form, and let f_T be the form $f_T(\underline{u}_1,\ldots,\underline{u}_n) = f(T\underline{u}_1,\ldots,T\underline{u}_n)$. Then

$$(\det(TD))f(\underline{u}_1,\ldots,\underline{u}_n) = f(TS\underline{u}_1,\ldots,TS\underline{u}_n) = f_T(S\underline{u}_1,\ldots,S\underline{u}_n) = (\det S)\,f_T(\underline{u}_1,\ldots,\underline{u}_n) = (\det S)(\det T)f(\underline{u}_1,\ldots,\underline{u}_n$$

Finally, if $TS = \mathrm{Id}$ then $(\det T)(\det S) = 1$ no neither is zero. If T is not invertible let $\underline{v}_1 \in \mathrm{Ker}\,T$ and extend this to a basis. Let f any volume form. Then $(\det T)f(\underline{v}_1,\cdots,\underline{v}_n) = f(T\underline{v}_1,\cdots,T\underline{v}_n) = f(\underline{0},\cdots) = 0$. Since a non-zero volume form is non-zero on a basis we see that $\det T = 0$. □

THEOREM 101. *Let* A,B *be two matrices related by a sequence of row or column combinations. Then* $\det A = \det B$.

PROOF. If $B = E_r E_{r-1}\cdots E_1 A E_1' \cdot E_s'$ then $\det A = \det B$ since $\det(I_n + cE^{ij}) = 1$ where $i \neq j$. □

PROPOSITION 102. $\det A = \det A^t$.

PROOF. Exercise. □

COROLLARY 103. *Let* A *be upper-triangular, then* $\det A = \prod_{i=1}^n a_{ii}$.

COROLLARY 104 (Minor expansion). $\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij})$.

PROOF. By Prop 102, □

EXAMPLE 105. $\begin{vmatrix}1&2&3\\3&1&0\\3&3&-2\end{vmatrix} \overset{R_3-R_2}{=} \begin{vmatrix}1&2&3\\3&1&0\\0&2&-2\end{vmatrix} = 2\begin{vmatrix}1&2&3\\3&1&0\\0&1&-1\end{vmatrix} \overset{R_1+3R_3}{=} 2\begin{vmatrix}1&5&0\\3&1&0\\0&1&-1\end{vmatrix} = -2\begin{vmatrix}1&5\\3&1\end{vmatrix} = -2(1-15) = 28$.

$\begin{vmatrix}1&2&3\\3&1&0\\3&3&-2\end{vmatrix} \overset{R_1-2R_2}{=} \begin{vmatrix}-5&0&3\\3&1&0\\3&3&-2\end{vmatrix} \overset{R_3-3R_2}{=} \begin{vmatrix}-5&0&3\\3&1&0\\-6&0&-2\end{vmatrix} = -2\begin{vmatrix}-5&0&3\\3&1&0\\3&0&1\end{vmatrix} \overset{R_1-3R_3}{=} -2\begin{vmatrix}-14&0&0\\3&1&0\\3&0&1\end{vmatrix} = -2(-14\cdot 1\cdot 1) = 28$.

CHAPTER 4

Eigenvalues, eigenvectors, and diagonalization

4.1. Similarity and change of basis

Let V be a vector space with ordered basis $\mathcal{B} = \{\underline{v}_i\}_{i=1}^n$. To a linear map $T \in \mathrm{End}(V)$ we associated a *matrix* A consisting of the coefficients of the $T\underline{v}_j$ in the basis:

$$T\underline{v}_j = \sum_{i=1}^n a_{ij}\underline{v}_i \,.$$

QUESTION 106. *What happens if we instead use a different basis?*

So let $\mathcal{C} = \{\underline{u}_k\}_{k=1}^n \subset V$ be another basis. We can expand each $\underline{u}_k$ in the original basis, obtaining the *change of basis matrix* S, whose entries are defined by

$$\underline{u}_\ell = \sum_{j=1}^n s_{j\ell}\underline{v}_j \,.$$

Note that the columbs of S are exactly the expansions of the elements of $\mathcal{C}$ in the basis $\mathbf{B}$, and that S is the matrix with respect to $\mathcal{B}$ of the linear map $R \in \mathrm{End}(V)$ defined by $R\underline{v}_\ell = \underline{u}_\ell$.

Applying T to both sides we get:

$$\begin{aligned} T\underline{u}_\ell &= T\left(\sum_{j=1}^n s_{j\ell}\underline{v}_j\right) \\ &= \sum_{j=1}^n s_{j\ell}T\underline{v}_j \\ &= \sum_{j=1}^n s_{j\ell}\sum_{i=1}^n a_{ij}\underline{v}_i \\ &= \sum_{i=1}^n\left(\sum_{j=1}^n a_{ij}s_{j\ell}\right)\underline{v}_i \,. \end{aligned}$$

Note that the prentheses are exactly give the $i\ell$th entry of the matrix AS. Next, expand $\underline{v}_i$ in the basis $\mathcal{C}$. Suppose $\underline{v}_i = \sum_{k=1}^n t_{ki}\underline{u}_\ell$ (so the t_{ki} are the entries of the reverse change-of-basis matrix). We then have

$$\begin{aligned} \underline{v}_i = \sum_{k=1}^n t_{ki}\underline{u}_\ell &= \sum_{k=1}^n t_{ki}\sum_{j=1}^n s_{jk}\underline{v}_j \\ &= \sum_{j=1}^n\left(\sum_{k=1}^n s_{jk}t_{ki}\right)\underline{v}_j \,. \end{aligned}$$

But vectors have a unique representation in the basis, and we get

$$\sum_{k=1}^n s_{jk}t_{ki} = \begin{cases} 1 & j=i \\ 0 & j\neq i \end{cases} = \delta_{ij}$$

is the ijth entry of the identity matrix. In other words, the t_{ki} are the entries of the inverse matrix S^{-1} and the relation is

$$\underline{v}_i = \sum_{\ell=1}^{n} (s^{-1})_{ki}\underline{u}_k$$

and hence

$$\begin{aligned} T\underline{u}_\ell &= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}s_{j\ell}\right)\underline{v}_i \\ &= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}s_{j\ell}\right)\sum_{k=1}^{n}(s^{-1})_{ki}\underline{u}_k \\ &= \sum_{k=1}^{n}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}(s^{-1})_{ki}a_{ij}s_{j\ell}\right)\underline{u}_k \\ &= \sum_{k=1}^{n}\left(S^{-1}AS\right)_{k\ell}\underline{u}_k\,. \end{aligned}$$

But on the other hand by definition the coefficients here define the matrix of T in the basis $\mathcal{C}$, and we have proved

PROPOSITION 107. *Let $A, B \in M_n(\mathbb{R})$ be the matrices of $T \in \mathrm{End}(V)$ wrt the bases $\{\underline{v}_i\}_{i=1}^n, \{\underline{u}_k\}_{k=1}^n \subset V$ respectively. Let $S \in M_n(\mathbb{R})$ be the change-of-basis matrix*, defined by $\underline{u}_\ell = \sum_{j=1}^n s_{j\ell}\underline{v}_j$. Then

$$B = S^{-1}AS\,.$$

DEFINITION 108. Two matrices $A, B \in M_n(\mathbb{R})$ are *similar* if there is a matrix $S \in M_n(\mathbb{R})$ such that $B = S^{-1}AS$. Two linear maps $S, T \in \mathrm{End}(V)$ are *similar* if there is an invertible map $R \in \mathrm{End}(V)$ such that $S = R^{-1}TR$.

OBSERVATION 109. *Two matrices are similar iff they represent the same linear map in different bases.*

EXERCISE 110. (PS6) Similarity is an equivalence relation.

4.2. Motivation

Fix a vector space V.

DEFINITION 111. Let $T \in \mathrm{End}(V)$. Suppose we have a scalar λ and a non-zero $\underline{v} \in V$ such that $T\underline{v} = \lambda\underline{v}$. We then say that λ is an *eigenvalue* of T, and that $\underline{v}$ is an *eigenvector* corresponding to the eigenvalue λ.

REMARK 112. The equation is *non-linear*! [but linear in $\underline{v}$ for λ fixed]

Why care?

4.2.1. Diagonlization. Suppose we have a basis consisting of eigenvectors. Then the matrix is diagonal, hence simple (for example we can easily find the maqtrix of T^2 in that basis).

4.2.2. Solve differential equations. 1,$\cos(2\pi kx)$, $\sin(2\pi kx)$ are a basis for functions on the circle on which $\frac{d^2}{dx^2}$ acts by scalars. This is a good basis in which to study differential equations.

Note that eigenvalues are given by non-positive reals.

4.2.3. Solve difference equations. Let $L\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ be the left-shift operator. Then for any $r \in \mathbb{R}$, $(r^n)_{n\geq 0}$ is an eigenvector with eigenvalue r.

Now let F_n be the Fibbonaci sequence satisfying $F_{n+2} = F_{n+1} + F_n$. From our earlier work on difference equations we know to write this as $\left(L^2 - L - 1\right)\underline{F} = \underline{0}$ and that the space of solutions is two-dimensional. Now let $r_{1,2} = \frac{1\pm\sqrt{5}}{2}$ be the two roots of $r^2 - r - 1 = 0$. It follows that $\left\{ \left(r_1^n\right)_{n\geq 0}, \left(r_2^n\right)_{n\geq 0} \right\}$ both belong to $\operatorname{Ker}\left(L^2 - L - 1\right)$. They are not propoprtional, hence a basis. We have proven:

THEOREM 113. *Let F_n be the Fibonnacci sequence with $F_0 = 0$, $F_1 = 1$. Then $F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right)$.*

PROOF. There are A, B such that $Ar_1^n + Br_2^n = F_n$ for all n. Specifically for $n = 0, 1$ we see:

$$\begin{aligned} A + B &= 0 \\ Ar_1 + Br_2 &= 1 \end{aligned}$$

and this has the solution $A = -B = \frac{1}{\sqrt{5}}$. □

COROLLARY 114. *$F_n^{1/n} \to \frac{1+\sqrt{5}}{2}$ and $\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n$ is exponentially close to being an integer.*

4.2.4. Quantum Mechanics. Observables are linear operator

4.2.5. PCA = FA.

4.3. The characteristic polynomial, trace and determinant

4.3.1. Work by hand.

EXAMPLE 115. Let $A = \begin{pmatrix} 4 & 3 \\ 1 & 2 \end{pmatrix}$. Then we need to solve $\begin{pmatrix} 4 & 3 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$, that is $\begin{cases} (4-\lambda)x + 3y &= 0 \\ x + (2-\lambda)y &= 0 \end{cases}$.

Suppose (x, y, λ) is a solution. Then $x = (\lambda - 2)y$ so

$$(4-\lambda)(\lambda-2)y + 3y = 0$$

that is

$$\left(\lambda^2 - 6\lambda + 5\right)y = 0\,.$$

Thus either $y = 0$ at which point $x = 0$ and λ can be arbitrary, or $y \neq 0$ at which point $\lambda \in \{1, 5\}$ and $x = (\lambda - 2)y$ for arbitrary y (check that these are solutions!)

CONCLUSION 116. The eigenvalues of A are $1, 5$ and the corresponding eigenspaces are $\operatorname{Span}\left\{\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\}$, $\operatorname{Span}\left\{\begin{pmatrix} 3 \\ 1 \end{pmatrix}\right\}$.

4.3.2. The char poly. Recap:

- The eigenvalue problem is non-linear – we found λ as roots of a polynomial.
- Given λ, the problem is purely linear.

Fix V, $T \in \operatorname{End}(V)$. Then λ is an eigenvalue iff there is non-zero $\underline{v}$ such that $T\underline{v} = \lambda\underline{v}$ that is iff $\operatorname{Ker}(\lambda \operatorname{Id}_V - T) \neq \{\underline{0}\}$. If V is finite-dimensional this is equivalent to $\lambda - T$ being non-invertible and hence to $\det(\lambda \operatorname{Id}_V - T) = 0$.

DEFINITION 117. Let V be finite dimensional. The *characteristic polyonomial* of $T \in \operatorname{End}(V)$ is $p(x) = p_T(x) = \det(x \operatorname{Id}_V - T)$.

REMARK 118. Best to think of the matrix of $x \operatorname{Id}_V - T$ as a matrix in $M_n(\mathbb{R}[x])$, showing that the determinant is indeed a polynomial.

We have proved:

THEOREM 119. *λ is an eigenvalue of T iff λ is a root of $p_T(x)$.*

REMARK 120. In practice this is a terrible way of finding eigenvalues – can't find roots of polynomials.

EXERCISE 121. The characteristic polynomial is always monic of degree is $\dim V$. Any such polynomial is the char poly of a linear map.

EXAMPLE 122. The characteristic polynomial of $\begin{pmatrix} 4 & 3 \\ 1 & 2 \end{pmatrix}$ is $x^2 - 6x + 5$.

PROOF. Can do direct calculation, but also note that must be monic and divisible by $(x-5)(x-1)$ since those are eigenvalues. □

REMARK 123. The polynomial doesn't have to have real roots (consider $\begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$), but the F.T.Algebra says that always factor over complex numbers (exercize: the 2x2 case).

The following aside will help us identify "obvious" roots:

THEOREM 124 (Rational roots). *Let $p(x) \in \mathbb{Z}[x]$ satisfy $p(x) = \sum_{i=0}^{d} a_i x^i$ with $a_0 a_d \neq 0$. Suppose that $p(\frac{r}{s}) = 0$ with $r, s \in \mathbb{Z}$ relatively prime, $s \neq 0$. Then $r | a_0$, $s | a_d$.*

PROOF. Clear denominators to get $\sum_{i=0}^{d} a_i s^{d-i} r^i = 0$. Now $a_0 s^d = -r\left(\sum_{i=1}^{d} a_i s^{d-i} r^{i-1}\right)$ and $a_d r^d = -s\left(\sum_{i=0}^{d-1} a_i s^{(d-1)-i} r^i\right)$. □

4.4. Properties and diagonalization

Fix V, $T \in \mathrm{End}(V)$.

DEFINITION 125. $V_\lambda = \mathrm{Ker}(V - \lambda\,\mathrm{Id}_V)$; $\mathrm{Spec}(\lambda) = \{\lambda \in \mathbb{C} \mid V_\lambda \neq \{\underline{0}\}\}$.

LEMMA 126. *Let $\{\underline{v}_i\}_{i=1}^{r}$ be eigenvectors of T with distinct eigenvalues λ_i. Then the $\{\underline{v}_i\}$ are independent.*

PROOF. Consider a minimal dependence and apply T. □

COROLLARY 127. *The sum $\sum_\lambda V_\lambda$ is direct. In particular, at most n distinct eigenvalues (also follows from char poly).*

DEFINITION 128. Let λ be a scalar. The *algebraic* multiplicity of λ as an eigenvalue is the maximum r such that $(x-\lambda)^r$ divides $p_T(\lambda)$. The *geometric* multiplicity is $\dim V_\lambda$.

PROPOSITION 129. *Algebraic$\geq$Geometric*

PROOF. Let $\{\underline{v}_i\}_{i=1}^{r}$ be a basis for V_λ. Complete this into a basis $\{\underline{v}_i\}_{i=1}^{n}$ for V. Let A be the matrix for T in this basis. Then $p_A(x) = (x-\lambda)^r p_B(x)$ for the lower-right square B of A (repeatedly expand by columns).

Let $\{\underline{v}_i\}_{i=1}^{r}$ span the λ-eigenspace; complete to a basis. Let A be matrix by this basis, and expand $\det(xI_n - A)$ by first r columns to see that $(x-\lambda)^r$ divides $p_A(x)$. □

EXAMPLE 130. $\begin{pmatrix} & 1 & \\ & & 1 \\ & -2 & -1 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}$.

DEFINITION 131. Call T diagonable (or diagonalizable) if V has a basis consisting of eigenvectors.

This is equivalent to saying that in some basis the matrix of T is diagonal.

LEMMA 132. *Let T be a linear map, $\{\underline{v}_i\}_{i=1}^n$, $\{\underline{u}_j\}_{j=1}^n$ two bases of V. Let A, A' be the matrices of T wrt the two bases. Let B be the matrix whose jth column is the decomposition of $\underline{u}_j$ in $\underline{v}_i$. Then S is invertible and $A' = B^{-1}AB$.*

PROOF. B is the matrix of the map S such that $S\underline{v}_i = \underline{u}_i$ wrt basis $\underline{v}_i$ .Maps basis to basis so invertible. B^{-1} is matrix of $S^{-1}\underline{u}_i = \underline{v}_i$ in basis $\underline{v}_i$, so columns are decomposition of $\underline{v}_j$ in terms of $\underline{u}_i$. Now
$T\underline{u}_j = T\left(\sum_i b_{ij}\underline{v}_i\right) = \sum_i b_{ij} T\underline{v}_i = \sum_i b_{ij} \sum_k a_{ki}\underline{v}_k = \sum_{i,k,l} a_{ki} b_{ij} \sum_l (b^{-1})_{lk}\underline{u}_k = \sum_k \left(B^{-1}AB\right)_{kj}\underline{u}_k$. □

THEOREM 133. *Let $A \in M_n(\mathbb{R})$ and let $\{\underline{v}_i\}_{i=1}^n$ be linearly independent such that $A\underline{v}_i = \lambda_i\underline{v}_i$. Let S be the matrix with columns $\underline{v}_i$. Then S is invertible and $S^{-1}AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Equivalently, $A = SDS^{-1}$ where $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.*

EXAMPLE 134. $\begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix}$ has eigenvectors $\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

LEMMA 135. *Let A, B be similar matrices. Then they are the matrices of the same linear map in different bases.*

4.5. Diversion: Graph eigenvalues and PageRank

CHAPTER 5

Inner product spaces

In this chapter the field of scalars is either $\mathbb{R}$ or $\mathbb{C}$.

5.1. Inner product spaces

5.1.1. Motivation and basic examples. In Euclidean space $\mathbb{E}^n$ we have a notion of distance between points. Equivalently we have a notion of distance in $\mathbb{R}^n$.

5.1.2. Definition. For real vector spaces

DEFINITION 136. Let V be a vector space over the real field. An *inner product* on V is a map $\langle\cdot,\cdot\rangle : V\times V\to\mathbb{R}$ such that:

(1) (Bilinearity) The map is linear in the second coordinate: $\langle \underline{u}, \alpha\underline{v}+\beta\underline{w}\rangle = \alpha\langle\underline{u},\underline{v}\rangle + \beta\langle\underline{u},\underline{w}\rangle$.
(2) (Symmetry) $\langle\underline{u},\underline{v}\rangle = \langle\underline{v},\underline{u}\rangle$.
(3) (Positivity) $\langle\underline{u},\underline{u}\rangle \geq 0$ with equality iff $\underline{u}=\underline{0}$.

An *inner product space* is a pair $(V,\langle\cdot,\cdot\rangle)$ where V is a real vector space and $\langle\cdot,\cdot\rangle$ is an inner product on V.

For complex scalars positivity requires a more complicated definition:

DEFINITION 137. Let V be a vector space over the complex field. A *hermitian product* on V is a map $\langle\cdot,\cdot\rangle : V\times V\to\mathbb{C}$ such that:

(1) (Conjugate linearity) The map is linear in the second coordinate: $\langle \underline{u}, \alpha\underline{v}+\beta\underline{w}\rangle = \alpha\langle\underline{u},\underline{v}\rangle + \beta\langle\underline{u},\underline{w}\rangle$.
(2) (Conjugate symmetry) $\langle\underline{u},\underline{v}\rangle = \overline{\langle\underline{v},\underline{u}\rangle}$.
(3) (Positivity) $\langle\underline{u},\underline{u}\rangle \geq 0$ with equality iff $\underline{u}=\underline{0}$.

A *hermitian space* is a pair $(V,\langle\cdot,\cdot\rangle)$ where V is a complex vector space and $\langle\cdot,\cdot\rangle$ is a hermitian product on V.

When $\underline{u}=\underline{v}$, axiom (2) reads $\langle\underline{u},\underline{u}\rangle = \overline{\langle\underline{u},\underline{u}\rangle}$, ensuring that $\langle\underline{u},\underline{u}\rangle\in\mathbb{R}$ so that axiom 3 makes sense.

REMARK 138. We often abuse terminology and use "inner product space" in both contexts.

EXAMPLE 139. The *standard inner product* on $\mathbb{R}^n$ is the one from above. The analogues *standard inner product* on $\mathbb{C}^n$ is

$$\langle\underline{z},\underline{w}\rangle = \sum_{i=1}^{n} \overline{z_i}\cdot w_i\,.$$

EXAMPLE 140. Let $C(a,b)$ be the space of real- or complex-valued functions on the interval $[a,b]$. Setting

$$\langle f,g\rangle = \int_a^b \overline{f(x)}g(x)\,\mathrm{d}x$$

defines an inner (respectively) hermitian product on $C(a,b)$.

PROOF. Linearity is immediate from properties of the Riemann integral and conjugate symmetry is clear. If $f = g$ we have

$$\langle f, f\rangle = \int_a^b |f(x)|^2 \, \mathrm{d}x .$$

The is non-negative since the integrand is non-negative. Also, since f is continuous if f is non-zero then $|f|$ is positive on a subinterval, and the integral is strictly positive. □

EXAMPLE 141. (PS12) On $M_n(\mathbb{C})$ set $\langle A, B\rangle = \mathrm{Tr}\left(A^\dagger B\right)$ where $(A^\dagger)_{ij} = \overline{A_{ij}}$.

LEMMA 142. *Let V be an inner product space. Then the inner product is linear in the first variable (conjugate-linear if the scalars are complex).*

PROOF. We have

$$\begin{aligned}
\langle \alpha \underline{u} + \beta \underline{v}, \underline{w}\rangle &= \overline{\langle \underline{w}, \alpha \underline{u} + \beta \underline{v}\rangle} \\
&= \overline{\alpha \langle \underline{w}, \underline{u}\rangle + \beta \langle \underline{w}, \underline{v}\rangle} \\
&= \overline{\alpha}\overline{\langle \underline{w}, \underline{u}\rangle} + \overline{\beta}\overline{\langle \underline{w}, \underline{v}\rangle} \\
&= \overline{\alpha} \langle \underline{u}, \underline{w}\rangle + \overline{\beta} \langle \underline{v}, \underline{w}\rangle .
\end{aligned}$$

□

LEMMA 143 (Restriction). *Let V be an inner product space and let W be a subspace. Then $(W, \langle \cdot , \cdot \rangle \restriction_{W\times W})$ is an inner product space. If V is complex then $(V, \Re \langle \cdot , \cdot \rangle)$ is a real inner product space when we treat V as a real vector space.*

PROOF. All the axioms are universal. □

5.2. The Cauchy–Schwartz inequality

Fix an inner product space V.

DEFINITION 144 (The norm). The *norm* of $\underline{u} \in V$ is the non-negative real number $\|\underline{u}\| = \sqrt{\langle \underline{u}, \underline{u}\rangle}$ (recall that this $\langle \underline{u}, \underline{u}\rangle$ is always a non-negative real number).

By the axioms for an inner product space we have $\|\underline{u}\| = 0$ iff $\underline{u} = \underline{0}$. We also observe that the norm is 1-*homogenous*:

$$\begin{aligned}
\|\alpha \underline{u}\| &= \sqrt{\langle \alpha \underline{u}, \alpha \underline{u}\rangle} = \sqrt{\bar{\alpha}\alpha \langle \underline{u}, \underline{u}\rangle} \\
&= \sqrt{\bar{\alpha}\alpha}\sqrt{\langle \underline{u}, \underline{u}\rangle} \\
&= |\alpha| \, \|\underline{u}\| .
\end{aligned}$$

LEMMA 145 (Cauchy–Schwartz). *Let $\underline{u}, \underline{v} \in V$. Then*

$$|\langle \underline{u}, \underline{v}\rangle| \leq \|\underline{u}\| \, \|\underline{v}\| ,$$

with equality if and only if $\underline{u}, \underline{v}$ are multiples of each other.

PROOF. If $\langle \underline{u}, \underline{v}\rangle = 0$ there is nothing to prove. Otherwise there is a number α of modulus 1 such that $\alpha \langle \underline{u}, \underline{v}\rangle = |\langle \underline{u}, \underline{v}\rangle|$. Consider then the real-valued function

$$f(t) = \|t\underline{u} + \alpha \underline{v}\|^2 = \langle t\underline{u} + \alpha \underline{v}, t\underline{u} + \alpha \underline{v}\rangle .$$

Using the bilinearity we have

$$f(t) = t^2 \langle \underline{u}, \underline{u}\rangle + t \langle \underline{u}, \alpha \underline{v}\rangle + t \langle \alpha \underline{v}, \underline{u}\rangle + |\alpha|^2 \langle \underline{v}, \underline{v}\rangle .$$

Observe now that $\langle \underline{u}, \alpha \underline{v} \rangle = \alpha \langle \underline{u}, \underline{v} \rangle = |\langle \underline{u}, \underline{v} \rangle|$ is real, hence equal to its own complex conjugate, and we get that $\langle \alpha \underline{v}, \underline{u} \rangle = \langle \underline{u}, \alpha \underline{v} \rangle$. We also have $\langle \underline{u}, \underline{u} \rangle = \|\underline{u}\|^2$ and $\langle \underline{v}, \underline{v} \rangle = \|\underline{v}\|^2$ and $|\alpha| = 1$ so in the end

$$f(t) = t^2 \|\underline{u}\|^2 + 2t |\langle \underline{u}, \underline{v} \rangle| + \|\underline{v}\|^2 .$$

Note that this is a quadratic function with positive real coefficients: if $\langle \underline{u}, \underline{u} \rangle$ or $\langle \underline{v}, \underline{v} \rangle$ vanishes than one of $\underline{u}, \underline{v}$ would vanish and then their inner product would vanish as well. Completing the square, we have

$$f(t) = \left(t \|\underline{u}\| + \frac{|\langle \underline{u}, \underline{v} \rangle|}{\|\underline{u}\|} \right)^2 + \|\underline{v}\|^2 - \frac{|\langle \underline{u}, \underline{v} \rangle|^2}{\|\underline{u}\|^2} .$$

Now as the norm of a vector we have $f(t) \geq 0$ for all t, so we must have

$$\|\underline{v}\|^2 - \frac{|\langle \underline{u}, \underline{v} \rangle|^2}{\|\underline{u}\|^2} \geq 0$$

Which can be rearranged to form the desired inequality. In addition equality holds iff there is t such that $f(t) = 0$, that is iff there is t such that $t\underline{u} + \alpha \underline{v} = 0$ or (assuming $\underline{u}, \underline{v} \neq \underline{0}$) that $\underline{u} = -t^{-1} \alpha \underline{v}$. □

PROPOSITION 146 (Minkowsky's inequality; "triangle inequality"). *We have* $\|\underline{u} + \underline{v}\| \leq \|\underline{u}\| + \|\underline{v}\|$ *(and also* $\|\alpha \underline{u}\| = |\alpha| \, \underline{u}$ *and* $\|\underline{u}\| = 0 \iff \underline{u} = \underline{0}$*).*

PROOF. We apply CS:

$$\begin{aligned} \|\underline{u} + \underline{v}\|^2 &= \langle \underline{u} + \underline{v}, \underline{u} + \underline{v} \rangle \\ &= \langle \underline{u}, \underline{u} \rangle + \langle \underline{u}, \underline{v} \rangle + \langle \underline{v}, \underline{u} \rangle + \langle \underline{v}, \underline{v} \rangle \\ &\leq \|\underline{u}\|^2 + \|\underline{u}\| \, \|\underline{u}\| + \|\underline{v}\| \, \|\underline{u}\| + \|\underline{v}\|^2 \\ &= (\|\underline{u}\| + \|\underline{v}\|)^2 . \end{aligned}$$

□

COROLLARY 147. *The function* $d(\underline{u}, \underline{v}) = \|\underline{v} - \underline{u}\|$ *is a* metric *on* V*:* $d \colon V \times V \to \mathbb{R}_{\geq 0}$ *satisfies* $d(\underline{u}, \underline{v}) = d(\underline{v}, \underline{u})$*,* $d(\underline{u}, \underline{v}) = 0 \iff \underline{u} = \underline{v}$ *and* $d(\underline{u}, \underline{w}) \leq d(\underline{u}, \underline{v}) + d(\underline{v}, \underline{w})$*.*

In a real inner product space we observe that for any non-zero vectors $\underline{u}, \underline{v}$ we have

$$-1 \leq \frac{\langle \underline{u}, \underline{v} \rangle}{\|\underline{u}\| \, \|\underline{v}\|} \leq 1$$

and in particular there is a unique angle $\theta \in [0, \pi]$ such that

$$\langle \underline{u}, \underline{v} \rangle = \|\underline{u}\| \, \|\underline{v}\| \cos \theta .$$

DEFINITION 148. We call θ the *angle* between $\underline{u}, \underline{v}$.

5.3. Orthogonality

5.3.1. Intro. Fix an inner product space V. We identify particular configurations of vectors which are convenient for linear algebra. Three ideas:

(1) Orthogonal and orthonormal systems
(2) The Gram–Schmidt procedure
(3) The orthogonal complement

DEFINITION 149. Two vectors $\underline{u}, \underline{v} \in V$ are *orthogonal* if $\langle \underline{u}, \underline{v} \rangle = 0$, (for non-zero vectors, if the angle between them if $\frac{\pi}{2}$). In that case we write $\underline{u} \perp \underline{v}$

5.3.2. Orthogonal and orthonormal systems.

DEFINITION 150. A set of vectors $B \subset V$ is an *orthogonal system* if the vectors are non-zero and mutually orthogonal.

LEMMA 151. *An orthogonal system is linearly independent.*

PROOF. Suppose we have a linear combination

$$\sum_{i=1}^{n} a_i \underline{v}_i = \underline{0}$$

where a_i are scalars and $\underline{v}_i \in B$ are distinct. Taking inner product with $\underline{v}_j$ we have

$$\begin{aligned} 0 = \langle \underline{v}_j, \underline{0} \rangle &= \left\langle \underline{v}_j, \sum_{i=1}^{n} a_i \underline{v}_i \right\rangle \\ &= \sum_{i=1}^{n} a_j \langle \underline{v}_j, \underline{v}_i \rangle && \text{linearity} \\ &= a_j \langle \underline{v}_j, \underline{v}_j \rangle && i \neq j \Rightarrow \underline{v}_j \perp \underline{v}_i\,. \end{aligned}$$

Since $\langle \underline{v}_j, \underline{v}_j \rangle > 0$ we have $a_j = 0$. □

Orthogonality is powerful because determining the coefficients of a vector with respect to an orthogonal system does not require solving systems of linear equations:

LEMMA 152. *Let B be an orthogonal system, and let $v \in \operatorname{Span}(B)$ have the form $\underline{v} = \sum_{i=1}^{n} a_i \underline{v}_i$ where $\underline{v}_i \in B$ are distinct. Then*

$$a_j = \frac{\langle \underline{v}_j, \underline{v} \rangle}{\langle \underline{v}_j, \underline{v}_j \rangle}\,.$$

PROOF. Same calculation as above. □

OBSERVATION 153. *The coefficient a_j depends only on $\underline{v}_j$ – not on the whole system. Note that this is* completely false *for general basis B.*

Clearly rescaling the vectors does not change orthogonality, and helps with calculation above. Accordingly we set

DEFINITION 154. An orthorgonal system B is *orthonormal* if every $\underline{v} \in B$ has norm 1. It is *complete* if the only vector orthogonal to the entire system is the zero vector.

REMARK 155. Complete orthonormal systems are often called “orthonormal bases”; this is often abbreviated o.n.b.

LEMMA 156. *Suppose* $\dim V = n < \infty$. *Then every complete orthonormal system in V is a basis.*

PROOF. Let $B = \{\underline{v}_i\}_{i=1}^{m} \subset V$ be a complete orthonormal system and let $\underline{v} \in V$. For each $\underline{v}_i \in B$ let $a_i = \langle \underline{v}_i, \underline{v} \rangle$ as above, and consider the vector

$$\underline{v} - \sum_{i=1}^{m} a_i \underline{v}_i\,.$$

We will verify that this vector is orthogonal to B, and it will follow that it is the zero vector, in other words that $\underline{v} = \sum_{i=1}^{m} a_i \underline{v}_i$.

Indeed

$$\left\langle \underline{v}_j, \underline{v} - \sum_{i=1}^{m} a_i \underline{v}_i \right\rangle = \left\langle \underline{v}_j, \underline{v} \right\rangle - \sum_{i=1}^{m} a_i \left\langle \underline{v}_j, \underline{v}_i \right\rangle$$
$$= a_j - a_j = 0\,.$$

□

EXAMPLE 157. The infinite-dimensional situation is more complicated, because the natural notions of span involve series of vectors. As an example, let $C(\mathbb{R}/\mathbb{Z})$ denote the space of continuous complex-valued functions on $\mathbb{R}$ which are $\mathbb{Z}$-periodic, that is functions such that $f(x+1) = f(x)$ for all x, equipped with the inner product $\langle f, g \rangle = \int_0^1 \overline{f(x)} g(x)\,\mathrm{d}x$. Then the functions $e_n(x) = e^{2\pi i n x}$ are an orthonormal system in $C(\mathbb{R}/\mathbb{Z})$: it's easy to check that they are orthogonal, but completeness is more involved (given $f \in C(\mathbb{R}/\mathbb{Z})$ which is non-zero and perpendicular to all these functions, one uses the Stone–Wierestrass Theorem to product an element $g = \sum_{|n| \leq N} a_n e_n$ in the span which is close to f pointwise. Since $|(f-g)(x)| \leq \varepsilon$ for all x we have $\|f-g\|^2 \leq \varepsilon^2$ (plug into the integral). On the other hand $f \perp g$ gives

$$\langle f-g, f-g \rangle = \langle f, f \rangle + \langle g, g \rangle - \langle f, g \rangle - \langle g, f \rangle$$
$$= \langle f, f \rangle + \langle g, g \rangle \geq \langle f, f \rangle$$

producing a contradiction as soon as $\varepsilon < \|f\|$.

(If you prefer the real-valued version of this, use the orthonormal system $\{1\} \cup \left\{ \sqrt{2}\sin(2\pi n x), \sqrt{2}\cos(2\pi n x) \right\}_{n=1}^{\infty}$ instead).

5.3.3. Gram–Schmidt.

5.3.4. Orthogonality and the orthogonal complement.

DEFINITION 158. A vector $\underline{u}$ is orthogonal to a subset $S \subset V$ if $\langle \underline{u}, \underline{v} \rangle = 0$ for all $\underline{v} \in S$. The *orthogonal complement* of a subset S is $S^{\perp} = \{\underline{u} \in V \mid \underline{u} \perp S\}$.

LEMMA 159. *The orthogonal complement of a subset is a subspace.*

PROOF. By definition $S^{\perp} = \bigcap \left\{ \underline{v}^{\perp} \right\}_{\underline{v} \in S}$ so it's enough to show that the orthogonal complement of each vector is a subspace. But

$$\underline{v}^{\perp} = \{\underline{u} \mid \langle \underline{u}, \underline{v} \rangle = 0\}$$
$$= \{\underline{u} \mid \langle \underline{v}, \underline{u} \rangle = 0\}$$
$$= \operatorname{Ker}\left(\langle \underline{v}, \cdot \rangle \right)\,.$$

□

PROPOSITION 160 (Orthogonal decomposition). *Let $W \subset V$ be a subspace. Then $W \cap W^{\perp} = \{\underline{0}\}$ and $V = W \oplus W^{\perp}$.*

5.4. Linear maps and : the adjoint

As usual we fix an inner product space $(V, \langle \cdot, \cdot \rangle)$, which we assume finite-dimensional, with $n = \dim V$.

So far we've taken a *geometric* point of view: focusing on distances, angles, orthogonality, etc. We now take an *algebraic* point of view: how inner products interact with the rest of linear algebra. We begin with linear functional.

REMARK 161. The results of this section apply in the infinite-dimensional case under further *analytic* assumptions such as *completeness* of the space and *compactness* of the linear operators. Interested students should look up the theory of *Hilbert spaces*.

5.4.1. The Riesz Representation Theorem.

OBSERVATION 162. *Let $\underline{u} \in V$. We then get scalar-valued function on V by setting $\varphi_{\underline{u}}(\underline{v}) = \langle \underline{u}, \underline{v} \rangle$. This is actually a linear functional (by the definition of inner product!). The inner product thus gives an linear functional for each $\underline{u}$. In fact we get all linear functionals this way.*

LEMMA 163 (Riesz Representation Theorem, finite-dimensional case). *For each $\varphi \in V^*$ there is a unique $\underline{u} \in V$ such that $\varphi = \varphi_{\underline{u}}$. Furthermore, this bijection respects addition but is anti-linear for scalar multiplication: $\varphi_{c\underline{u}} = \bar{c}\varphi_{\underline{u}}$.*

PROOF. We prove the uniqueness first, Suppose $\varphi_{\underline{u}} = \varphi_{\underline{u}'}$, that is $\langle \underline{u}, \underline{v} \rangle = \langle \underline{u}', \underline{v} \rangle$ for all $\underline{v} \in V$. This statement is equivalent to

$$0 = \langle \underline{u}, \underline{v} \rangle - \langle \underline{u}', \underline{v} \rangle = \langle \underline{u} - \underline{u}', \underline{v} \rangle$$

and now choosing $\underline{v} = \underline{u} - \underline{u}'$ shows $\|\underline{u} - \underline{u}'\| = 0$ so $\underline{u} = \underline{u}'$.

For existence, let $\varphi \in V^*$. If φ is the zero functional then $\varphi = \varphi_{\underline{0}}$ and we are done, so suppose φ is non-zero. Then its image is 1-dimensional so by the rank-nullity theorem we have $\dim \operatorname{Ker} \varphi = n - 1$. Since $V = (\operatorname{Ker} \varphi) \oplus (\operatorname{Ker} \varphi)^{\perp}$ we have $\dim (\operatorname{Ker} \varphi)^{\perp} = 1$ so we can choose a non-zero vector $\underline{t} \in (\operatorname{Ker} \varphi)^{\perp}$ and set $\underline{u} = \frac{\overline{\varphi(\underline{t})}}{\langle \underline{t}, \underline{t} \rangle}\underline{t}$. We claim that $\varphi = \varphi_{\underline{u}}$. Indeed any $\underline{v} \in V$ can be uniquely written as a sum of two vectors, one each from $\operatorname{Ker} \varphi$ and $(\operatorname{Ker} \varphi)^{\perp}$. In other words we can write $\underline{v} = \underline{w} + c\underline{t}$ for some $\underline{w} \in \operatorname{Ker} \varphi$ and scalar c (recall that $(\operatorname{Ker} \varphi)^{\perp}$ is one-dimensional hence spanned by $\underline{t}$). Then

$$\varphi(\underline{v}) = \varphi(\underline{w} + c\underline{t}) = \varphi(\underline{w}) + c\varphi(\underline{t}) = c\varphi(\underline{t})$$

since $\underline{w} \in \operatorname{Ker} \varphi$. On the other hand $\underline{t} \in (\operatorname{Ker} \varphi)^{\perp}$ and $\underline{w} \in \operatorname{Ker} \varphi$ means $\langle \underline{t}, \underline{w} \rangle = 0$ and hence

$$\begin{aligned}
\varphi_{\underline{u}}(\underline{v}) &= \langle \underline{u}, \underline{v} \rangle \\
&= \left\langle \frac{\overline{\varphi(\underline{t})}}{\langle \underline{t}, \underline{t} \rangle}\underline{t}, \underline{w} + c\underline{t} \right\rangle \\
&= \frac{\overline{\overline{\varphi(\underline{t})}}}{\langle \underline{t}, \underline{t} \rangle} \left[\langle \underline{t}, \underline{w} \rangle + c \langle \underline{t}, \underline{t} \rangle \right] \\
&= \frac{\varphi(\underline{t})}{\langle \underline{t}, \underline{t} \rangle} c \langle \underline{t}, \underline{t} \rangle = c\varphi(\underline{t})
\end{aligned}$$

and thus $\varphi(\underline{v}) = \varphi_{\underline{u}}(\underline{v})$. □

EXERCISE 164. (PS12) Use the uniqueness to prove the claims about the (anti) linearity of the map.

5.4.2. The adjoint. If we interpret column vectors as $n \times 1$ matrices, then the standard inner product on $\mathbb{R}^n$ is given by $\langle \underline{u}, \underline{v} \rangle = \underline{u}^T \cdot \underline{v}$ where the dot denotes matrix multiplication. Then for any linear map A we can use the formula $(AB)^T = B^T A^T$ and the associativity of multiplication to get:

$$\begin{aligned}
\langle \underline{u}, A\underline{v} \rangle &= \underline{u}^T \cdot (A \cdot \underline{v}) = \left(\underline{u}^T \cdot A\right) \cdot \underline{v} \\
&= \left(\underline{u}^T \cdot \left(A^T\right)^T\right) \cdot \underline{v} \\
&= \left(A^T \underline{u}\right)^T \cdot \underline{v} \\
&= \left\langle A^T \underline{u}, \underline{v} \right\rangle .
\end{aligned}$$

In other words, the transpose matrix A^T is the matrix such the following formula holds:

$$\langle \underline{u}, A\underline{v} \rangle = \left\langle A^T \underline{u}, \underline{v} \right\rangle$$

for all $\underline{u}, \underline{v} \in \mathbb{R}^n$.

DEFINITION 165. Let $T \in \operatorname{End}(V)$. The *adjoint* of T is the map T^* such that

$$\langle T^* \underline{u}, \underline{v} \rangle = \langle \underline{u}, T\underline{v} \rangle$$

LEMMA 166. *The adjoint exists and is unique. We have $(cS+T)^* = c^* S^* + T^*$ where c^* is the complex conjugate.*

PROOF. Observe that for each $\underline{u}$, the map

$$\underline{v} \mapsto \langle \underline{u}, T\underline{v} \rangle$$

is a linear functional. By the Reisz representation theorem there is a unique vector $T^*\underline{u}$ such that this functional equals $\varphi_{T^*\underline{u}}$, in other words such that we have

$$\langle \underline{u}, T\underline{v} \rangle = \langle T^* \underline{u}, \underline{v} \rangle$$

for all $\underline{v}$. We need to check that T^* is a linear map. For this we use the uniqueness. Given two vectors $\underline{u}, \underline{u}'$ we get two vectors $T^*\underline{u}, T^*\underline{u}'$ such that for all $\underline{v}$ we have

$$\langle \underline{u}, T\underline{v} \rangle = \langle T^* \underline{u}, \underline{v} \rangle$$
$$\left\langle \underline{u}', T\underline{v} \right\rangle = \left\langle T^* \underline{u}', \underline{v} \right\rangle .$$

We then get for all $\underline{v}$ that:

$$\begin{aligned}
\left\langle T^*(c\underline{u}+\underline{u}'), \underline{v} \right\rangle &= \left\langle c\underline{u}+\underline{u}', T\underline{v} \right\rangle && \text{def of } T^* \\
&= c^* \langle \underline{u}, T\underline{v} \rangle + \left\langle \underline{u}', T\underline{v} \right\rangle && \text{inner pdt} \\
&= c^* \langle T^*\underline{u}, \underline{v} \rangle + \left\langle T^*\underline{u}', \underline{v} \right\rangle && \text{def of } T^* \\
&= \left\langle cT^*\underline{u} + T^*\underline{u}', \underline{v} \right\rangle && \text{inner pdt}
\end{aligned}$$

and it follows that $T^*(c\underline{u}+\underline{u}') = cT^*\underline{u} + T^*\underline{u}'$. □

EXAMPLE 167. When $V = \mathbb{R}^n$ equipped with the standard inner product the adjoint of a matrix is the transpose.

EXERCISE 168. (PS12) Show that when $V = \mathbb{C}^n$ equipped with its standard Hermitian product, the adjoint of a matrix is the *conjugate transpose* $T^\dagger$ *("T dagger", written* `T^\dagger` *in* LaTeX*), given by*

$$T^\dagger_{ij} = \overline{T_{ji}} .$$

EXERCISE 169. Let $C^\infty_c(\mathbb{R})$ be the space of functions $f \colon \mathbb{R} \to \mathbb{C}$ which are infinitely differentiable and *compactly supported* in that there is an M such that $f(x) = 0$ if $|x| > M$ (i.e. the non-zero part of the graph of f happens over a finite interval). Equip $C^\infty_c(\mathbb{R})$ with the inner product $\langle f, g \rangle = \int_{-\infty}^{+\infty} \bar{f} g dx$.

Interpret the formula for integration by parts to show that the operator $D = i\frac{d}{dx}$ acting on $C^\infty_c(\mathbb{R})$ is self-adjoint, in that $D^\dagger = D$ (note the factor of i!).

5.5. The spectral theorem

We can finally discuss the notion of "an operator respecting an inner product".

DEFINITION 170. We call a linear map *self-adjoint* if $T^\dagger = T$.

EXAMPLE 171. A matrix $A \in M_n(\mathbb{R})$ is called *symmetric* if $A^T = A$; a matrix $A \in M_n(\mathbb{C})$ is called *Hermitian* if $A^\dagger = A$.

Fix an inner product space V and a self-adjoint linear map $T \in \operatorname{End}(V)$.

5.5.1. Eigenvalues and eigenvectors of self-adjoint operators.

LEMMA 172. *Suppose $T\underline{v} = \lambda\underline{v}$ and that $\underline{v}$ is non-zero. Then λ is a real number.*

PROOF. We evaluate the inner product $\langle \underline{v}, T\underline{v}\rangle$ in two different ways. On the one hand

$$\langle \underline{v}, T\underline{v}\rangle = \langle \underline{v}, \lambda\underline{v}\rangle = \lambda \langle \underline{v}, \underline{v}\rangle$$

and on the other hand

$$\begin{aligned} \langle \underline{v}, T\underline{v}\rangle &= \left\langle T^{\dagger}\underline{v}, \underline{v}\right\rangle && \text{def of adjoint} \\ &= \langle T\underline{v}, \underline{v}\rangle && \text{self-adjointness} \\ &= \langle \lambda\underline{v}, \underline{v}\rangle \\ &= \overline{\lambda} \langle \underline{v}, \underline{v}\rangle . \end{aligned}$$

Since $\langle \underline{v}, \underline{v}\rangle \neq 0$ we must have $\lambda = \bar{\lambda}$ so $\lambda \in \mathbb{R}$. □

LEMMA 173. *Suppose $T\underline{v} = \lambda\underline{v}$ and $T\underline{w} = \mu\underline{w}$ and that $\lambda \neq \mu$. Then $\underline{v} \perp \underline{w}$.*

PROOF. We evaluate $\langle \underline{v}, T\underline{w}\rangle$ in two ways: We have:

$$\langle \underline{v}, T\underline{w}\rangle = \langle \underline{v}, \mu\underline{w}\rangle = \mu \langle \underline{v}, \underline{w}\rangle$$

and also

$$\langle \underline{v}, T\underline{w}\rangle = \langle T\underline{v}, \underline{w}\rangle = \langle \lambda\underline{v}, \underline{w}\rangle = \lambda \langle \underline{v}, \underline{w}\rangle$$

since λ is real. Subtracting the two expressions we get:

$$(\mu - \lambda) \langle \underline{v}, \underline{w}\rangle = \mu \langle \underline{v}, \underline{w}\rangle - \lambda \langle \underline{v}, \underline{w}\rangle = 0$$

and since $\mu - \lambda \neq 0$ we must have $\langle \underline{v}, \underline{w}\rangle = 0$. □

5.5.2. The Spectral Theorem. Recall that we have fixed a a finite-dimensional inner product space V and a self-adjoint map $T \in \mathrm{End}(V)$.

LEMMA 174. *T has at least one eigenvalue.*

PROOF. We have seen that every linear map has a complex eigenvalue. □

PROOF. Consider the continuous function $f(\underline{v}) = \langle \underline{v}, T\underline{v}\rangle$ on the sphere $\{\underline{v} \mid \|\underline{v}\| = 1\}$. This is a differentiable function so from calculus it has a maximum. We have $\nabla f(\underline{v}) = (T + T^{\dagger})\underline{v} = 2T\underline{v}$ and $\nabla \|\underline{v}\|^2 = 2\underline{v}$. By the theory of Lagrange multiplies there is $\lambda \in \mathbb{R}$ such that at the maximum point $\underline{v}$ we have

$$\nabla f(\underline{v}) = \lambda \nabla \left(\|\underline{v}\|^2 - 1 \right)$$

or in other words

$$2T\underline{v} = 2\lambda\underline{v}.$$

□

PROOF. Consider the continuous function $f(\underline{v}) = \langle \underline{v}, T\underline{v}\rangle$ on the sphere $\{\underline{v} \mid \|\underline{v}\| = 1\}$. It attains its maximum at some point $\underline{v}$. Now let $\underline{w} \in \underline{v}^{\perp}$ be a non-zero unit vector. Then for each real number c we have $\|\underline{v} + c\underline{w}\|^2 = 1 + |c|^2$ by Pythagoras. Thus $\frac{\underline{v}+c\underline{w}}{\sqrt{1+c^2}}$ has norm 1, and we set

$$g(c) = f\left(\frac{\underline{v} + c\underline{w}}{\sqrt{1 + |c^2|}} \right),$$

which attains a local maximum at $c = 0$. We have

$$f\left(\frac{\underline{v}+c\underline{w}}{\sqrt{1+|c^2|}}\right) = \left\langle \frac{\underline{v}+c\underline{w}}{\sqrt{1+|c^2|}}, T\frac{\underline{v}+c\underline{w}}{\sqrt{1+|c^2|}} \right\rangle$$
$$= \frac{1}{1+|c|^2}\left[\langle \underline{v}, T\underline{v}\rangle + c\langle \underline{v}, T\underline{w}\rangle + c\langle \underline{w}, T\underline{v}\rangle + c^2\langle \underline{w}, T\underline{w}\rangle\right].$$

Differentating at $c = 0$ we get

$$\frac{dg}{gc}(0) = \langle \underline{v}, T\underline{w}\rangle + \langle \underline{w}, T\underline{v}\rangle$$
$$= \langle T\underline{v}, \underline{w}\rangle + \langle \underline{w}, T\underline{v}\rangle$$
$$= 2\Re\langle \underline{w}, T\underline{v}\rangle .$$

It thus follows that $\Re\langle \underline{w}, T\underline{v}\rangle = 0$. Replaing $\underline{w}$ with $i\underline{w}$ (which still has $\langle i\underline{w}, \underline{v}\rangle = 0$ we conclude that $\Im\langle \underline{w}, T\underline{v}\rangle = 0$ so $\langle \underline{w}, T\underline{v}\rangle = 0$. Since $\underline{w} \in \underline{v}^\perp$ was arbitrary we conclude that $T\underline{v} \in \left(\underline{v}^\perp\right)^\perp = \mathrm{Span}\{\underline{v}\}$ and hence that $T\underline{v} = \lambda\underline{v}$ for some $\lambda \in \mathbb{C}$. □

LEMMA 175. *Let $W \subset V$ be a T-invariant subspace (in that $T(W) \subset W$). Then $W^\perp$ is also T-invariant.*

PROOF. Let $\underline{v} \in W^\perp$. Then for all $\underline{w} \in W$ we have $\langle \underline{w}, T\underline{v}\rangle = \langle T\underline{w}, \underline{v}\rangle = 0$ since $T\underline{w} \in W$ and $\underline{v} \in W^\perp$. It follows that $T\underline{v} \in W^\perp$ as well. □

THEOREM 176. *Let V be a finite-dimensional inner product space, and let $T \in \mathrm{End}(V)$ be self-adjoint. Then T is diagonable with real eigenvalues. In fact, there is an orthonormal basis of V consisting of eigenvectors of T.*

PROOF. Let $B \subset V$ be an orthonormal system consisting of eigenvectors of T which is as large as possible (this exists since $\#B \le \dim V = n$). An d let $W = \mathrm{Span}\,B$. Suppose B is not complete, so that $B^\perp = W^\perp$ is non-zero. By Lemma 175, $T\restriction_{W^\perp} \in \mathrm{End}(W^\perp)$. This is a self-adjoint map with respect to the restricted inner product, so by Lemma 174 it has at least one eigenvector $\underline{v}$, which we may take to have norm 1. But then $\underline{v} \in W^\perp = B^\perp$ is orthogonal to B, so $B \cup \{\underline{v}\}$ is a larger orthonormal system consisting of eigenvectors of T, a contradition. □

5.5.3. Orthogonal and unitary maps.