math108b lecture notes

Alice Bob

Table of Contents

1. January 7

1.0.1. Remark

Expectations for this class

Read, write, think, discuss math in a rigorous way

Notations and their importance

John Von Neumann: in mathematics you don't understand things, you just get used to them

Some applications: eigenvalues $\rightarrow$ google search (page rank)

Start with chapter 6, inner product spaces (go back to eignstuff later)

Review chapter 3 about product space, quotient space, dual space

1.1. Inner products and Norms (chapter 6A)

Inner products encode information about angles between vectors.

1.1.1. Remark

Notation: $(\mathbb{R}^n, \beta)$. Where β is a basis for $\mathbb{R}^n$.

Standard ordered basis $e_i = (0, ..., 0, 1, 0, ..., 0)$

$[v]_\beta$ (recall this coordinate vecotr notation)

v_1 can be unique written as $v_1 = x_1 e_1 + ... + x_n e_n$ then $[v_1]_\beta = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

If $v_2 = y_1 e_1 + ... + y_n e_n$ then the inner product $\langle v_1, v_2 \rangle$ is defined as

$\langle v_1, v_2 \rangle = x_1 y_1 + ... + x_n y_n$ or in general as a function: $\langle , \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$

1.1.2. *Proposition*

Some properties we can immediately see based on 1.1.1

1. $\langle v, v \rangle \geq 0$, with equality only if $v = 0$ (zero vector)
2. $\langle , u \rangle : \mathbb{R}^n \rightarrow \mathbb{R}$ maps $v \in V \rightarrow \langle v, u \rangle$ for some fixed $u \in V$ is a linear map from V to $\mathbb{R}$
3. $\langle u, v \rangle = \langle v, u \rangle$

Proof:

Claim: the linear map defined in ii) is a linear map
Proof: excercise

$\square$

1.1.3. Definition: Norm

The norm of a vector v is defined as $\|v\| = \sqrt{\langle v, v \rangle}$

Note that $\|v\|^2 = \langle v, v \rangle$

The norm is a map $\| \ \| : \mathbb{R}^n \to \mathbb{R}$ which maps $v \in \mathbb{R}^n \to \| \ v \ \|$

1.1.4. *Proposition*

> Norm properties
> 1. $\|v\| \geq 0$ since $\langle v, v \rangle \geq 0$
> 2. $\| \ \|$ is NOT a linear map.

Proof:
2. To be linear, we expect $\|v_1 + v_2\| = \|v_1\| + \|v_2\|$. But interpretation of norms as lengths of vectors disproves this.

$\square$

1.1.5. Definition: Inner Product

An inner product is a function with two input vectors and one output scalar defined as

$$\langle , \rangle : V \times V \to \mathbb{F}$$

Properties of inner products:

1. positivity $\langle v, v \rangle \geq 0$ (equality only when v is zero vector)
2. additivity in first slot: $\langle v_1 + v_2, u \rangle = \langle v_1, u \rangle + \langle v_2, u \rangle$
3. homogenity in first slot $\langle \lambda v, u \rangle = \lambda \langle v, u \rangle$
4. conjugate symmetry: $\langle u, v \rangle = \overline{\langle v, u \rangle}$

1.1.6. Remark

Combining properties 2 and 3 from 1.1.5 we can see that $\langle , u \rangle$ is a linear map.

The fact that these properties it is in the first slot is by convention (if defined in other slots, it will not be considered the standard inner product)

For example we can define in the second slot:

1.1.7. Remark

Recall that for any real x, $x = \overline{x}$ so that $\langle v, u \rangle = \overline{\langle v, u \rangle} = \langle u, v \rangle$

1.1.8. Remark

The difference between definitions of inner products in 1.1.1 and 1.1.5 is that the first one tells you the computation and the latter is more general. The inner product defined in 1.1.1 is often called the dot product.

1.1.9. Remark

We choose a definition of inner product (called the standard inner product) for $\mathbb{R}^n$ by the usual definition.

For example, for $\mathbb{R}^2$ has standard inner product $x_1 y_1 + x_2 y_2$.

1.1.10. Example: Weighted Inner Product

Define the weighted inner product $\langle v_1, v_2 \rangle_w$ to be $3x_1 x_2 + y_1 y_2$

excercise: verify properties 1-4 in 1.1.5

Two methods: kernel-method, and filter-method which have applications in ML and signal processing

1.1.11. Example: Not an Inner Product

The inner product in $\mathbb{R}^2$ defined by $\langle v_1, v_2 \rangle = 2x_1 x_2 + (-1)y_1 y_2$ is not an inner product since this definition would violate property 1

As a concrete example, if $v_1 = \binom{0}{1}$ then $\langle v_1, v_2, \rangle = -1 \leq 0$

1.1.12. Example: application amazon/netflix

in the lecture notes

1.1.13. Remark

The fourth property from 1.1.5 comes from consideration of complex numbers.

let $v_1 = a + bi$

in the lecture notes

1.1.14. Example: LADR example 6.3

Consider the vector space $V = P([0, 1])$ be polynmials defined on [0,1] with degree n.

Recall the bijection $\mathbb{R}^n \cong P_{n-1}([0, 1])$

Some possible definitions of inner products include

$\langle f, g \rangle = \int_0^1 f(x)g(x) \mathrm{d}x$
(we can see immediately that the positivity conditions (condition 1) is satisified)

or

$\langle f, g \rangle_e = \int_0^1 f(x)g(x)e^{-x} \mathrm{d}x$
("kernel function")

1.1.15. Definition: Induced Norm

The induced norm $\|v\| = \sqrt{\langle v, v \rangle}$

Remark: once $\langle \cdot, \cdot \rangle$ is defined, $| \cdot |$ is fixed.

Recall that the norm is a function map. In function notation: $\| \cdot \| : V \to \mathbb{R}$

1.1.16. Definition: Properties of Norm

Recall from 1.1.4 that

1. $\| \ v \ \| \geq 0$ with equality only if $v = 0$
2. $\| \cdot \|$ is NOT linear

And now note that (let $\|$ denote the absolute value)

3. $\| \ \lambda u \ \| = |\lambda| \ \|u\|$ but $\| \ \lambda u \ \| \neq \lambda \ \|u\|$

1.1.17. *Theorem:* Triangle Inequality

$\|u + v\| \leq \|u\| + \|v\|$ (for $\mathbb{R}$)

Proof:

$$\| \ u + v \ \| = \sqrt{\langle u + v, u + v \rangle} \Rightarrow$$
$$\| \ u + v \ \|^2 = \langle u + v, u + v \rangle = \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle = \|u\|^2 + \|v\|^2 + 2\langle u, v \rangle$$

Note that $(\|u\| + \|v\|)^2 = \|u\|^2 + 2 \ \|u\| \ \|v\| + \|v\|^2$.

Therefore we want to show that $2\langle u, v \rangle \leq 2 \ \|u\| \ \|v\|$

Considering only $u, v \in \mathbb{R}^2$ then we know that $\langle v, u \rangle = \|u\| \cdot \|v\| \cos \theta \leq \|u\| \ \|v\|$

We will revisit this later to prove in general for any vector space.

$\square$

1.1.18. Definition: More general definition of norm

Norm of V is a function: $V \to \mathbb{F}$ if it satisifes

1. $\|v\| \geq 0$ with equality only if $v = 0$
2. $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$ for every $v_1, v_2 \in V$
3. $\|\lambda v\| = |\lambda| \ \|v\|$ for every $\lambda \in \mathbb{F}, v \in V$

1.1.19. Definition: Different norms

Let $v = (x, y) \in \mathbb{R}^2$

$\|v\|_1 = |x| + |y|$

$\|v\|_2 = \sqrt{x^2 + y^2}$

$\|v\|_{\max} = \|v\|_\infty = \max \{|x|, |y|\}$

different norms perfer different "features"

$\|v\|_1$ treats each coordinate "equally"

$\|v\|_2$ treates each coordinate "equally" BUT prefers large coordinates

For example it is true that $x > y > 1$ means that $\|(x + 1, y)\|_2 \geq \|(x, y + 1)\|_2$

1.1.20. Remark

All norms are equivalent to eachother in finite dimensional v.s. but not all equivalent in infinite dimensional v.s.

We say that two norms $\|v\|_1, \|v\|_2$ are equivalent if there exists c_1, c_2 constanst it holds that

$$c_1 \ \|v\|_1 \leq \|v\|_2 \leq c_2 \ \|v\|_2$$

So,

This means that for the 1 and 2 norms in $\mathbb{R}^2$ that we need to find some universal c_1, c_2 (not depends on choice of v) such that

$c_1(|x| + |y|) \leq \sqrt{x^2 + y^2}$ and $\sqrt{x^2 + y + 2} \leq c_2(|x| + |y|)$

Proof for this case:

Let $c_1 = \frac{1}{2}, c_2 = 2$. We would like to show that

$\frac{|x| + |y|}{2} \leq \sqrt{x^2 + y^2}$ and $\sqrt{x^2 + y^2} \leq 2(|x| + |y|)$

Continued in lecture notes

1.2. Jan 16 (recorded lecture)

1.2.1. Remark

Last time

1. Define mechanism (algorithm) such as $x \cdot y = x_1 y_1 + \ldots + x_n y_n$

then prove properties: $\langle \cdot, y \rangle$ is linear, $\langle x, x \rangle \geq 0$ with equality if and only if $x = 0$, conjugate symmetry, +more properties (textbook section 6 and homework)

2. Define inner-product by (which satisfies) axioms

Examples include standard inner-product and weighted inner-product

After doing this definition, this introduces a norm associated with the defined inner product

For example for the standard inner-product $\langle \cdot, \cdot \rangle$ then $\|x\|^2 = \langle x, x \rangle$ ("standard norm")

For weighted norm $\langle \cdot, \cdot \rangle_w = x_1 w_1 y_1 + \ldots + x_n w_n y_n$ then $\|x\|_w^2 = \langle x, x \rangle_w$

We then can prove properties about the norm such as $\|x\| \geq 0$ with equality if and only if $x = 0$, $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality), and $\|\lambda x\| = |\lambda| \cdot \|x\|$

3. Define norm by norm axioms

this means we will "lost algorithm to compute norm", though we can later introduce a computation method such as the 1 norm, 2 norm, infinity norm, etc.

here, norm is not necessarily induced from inner-product

1.2.2. *Proposition*

> given a norm $\| \cdot \|$, it is not always possible to define the particular kind of inner-product, $\langle \cdot, \cdot \rangle$ such that $\|v\|^2 = \langle v, v \rangle$

1.2.3. *Theorem:* Parallelogram equality

> Given an inner product $\langle \cdot, \cdot \rangle$ if the norm $\| \cdot \|$ is induced from $\langle \cdot, \cdot \rangle$ then the norm satisfies the following:
>
> $$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$$
>
> Remark: this has strong geometric meaning (reference the picture above 6.21 in LADR-4)

Proof:

$\langle u + v, u + v \rangle + \langle u - v, u - v \rangle = \langle u, u + v \rangle + \langle v, u + v \rangle + \langle u, u - v \rangle + \langle -v, u - v \rangle = \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle + \langle u, u \rangle + \langle u, -v \rangle + \langle -v, -v \rangle + \langle -v, u \rangle = 2\ \|u\|^2 + 2\|v\|2$

□

1.2.4. *Corollary*

> Given a norm $\|\cdot\|$ if it satisfies the triangle inequality for any $u, v \in V$ then the norm is induced from some inner-product

Proof:

(not given)

Remark: (hw) $\|\cdot\|_1$ cannot be induced by any inner-product.

$\square$

1.2.5. Remark

Recall: given an inner product we showed that the induced norm satisfies the triangle inequality, but only in $\mathbb{R}^2$ (today we prove a more general version)

Recall: a norm defined by axioms satisfies the triangle inequality automatically

1.2.6. Definition: Orthogonal

$u, v \in V$ are orthogonal to each other if $\langle u, v \rangle = 0$

1.2.7. *Theorem:* Pythagorean Theorem

> Let u, v be orthogonal then the following is true:
> $$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

Proof:

$\langle u + v, u + v \rangle = \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle = \|u\|^2 + \|v\|^2$

Which is very nice as it only requires definition of inner product, norm, and orthogonal

$\square$

1.2.8. *Theorem:* Orthogonal decomposition

> Given $u, v \in V, v \neq 0$ then $u = cv + w$ where cv is parallel to v and w is normal to v i.e. $\langle w, v \rangle = 0$
> In fact we can explicitly compute c and w by $c = \frac{\langle u,v \rangle}{\|v\|^2}$ and $w = u - \frac{\langle u,v \rangle}{\|v\|^2} v$

Proof:

Definte $w = u - cv$ then

$$0 = \langle w, v \rangle = \langle u - cv, v \rangle = \langle u, v \rangle - c\ \|v\|^2 \Rightarrow c = \frac{\langle u, v \rangle}{\|v\|^2}$$

and the computation for w follows (plug in c)

□

1.2.9. *Theorem:* Cauchy - Schwarz Inequality

> For $u, v \in V$ then $|\langle u, v \rangle| \leq \|u\|\ \|v\|$.
> Note that $\| \cdot \|$ must be induced from an inner-product

Proof:

Note that

$$u = \frac{\langle u, v \rangle}{\|v\|^2} v + w$$

then by Pythagorean theorem since v, w ,

$$\|u\|^2 = \| \frac{\langle u, v \rangle}{\|v\|^2} v\|^2 + \|w\|^2 = \frac{|u + v|^2}{\|v\|^2} + \|w\|^2 \geq \frac{|u + v|^2}{\|v\|^2}$$

Then rearrange the far left and far right, take square root, done. Additionally note that

$$\| \frac{\langle u, v \rangle}{\|v\|^2} v\ \|^2 = \langle \frac{\langle u, v \rangle}{\|v\|^2} v, \frac{\langle u, v \rangle}{\|v\|^2} v \rangle = \left(\frac{\langle u, v \rangle}{\|v\|^2} \right)^2\ \|v\|^2 = \frac{|u + v|^2}{\|v\|^2}$$

□

1.3. January 21 "From Inner-Product to Orthonormal"

1.3.1. Definition: Orthonormal Vectors

Set of vectors, denoted $\left\{ e_j \right\}_{i=1}^{n}$

1. Must be orthogonal $\langle e_i, e_j \rangle = 0$ for $i \neq j$
2. Normalized: $\langle e_i, e_i \rangle = \|e_i\|^2 = 1 = \|e_i\|$

1.3.2. Example: Orthonormal Sets

standard basis in $\mathbb{R}^3$ is orthonormal
the set $\left\{\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right)\right\}$ is orthonormal

1.3.3. *Corollary*

If a list of vectors are orthonormal, then that list of vectors are linearly independent

Proof:

Let $\{e_i\}_{i=1}^n$ be orthonormal.

Proof is in the notes, and the book

$\square$

1.3.4. Definition: Orthonormal basis

A list of vectors is orthonormal basis if they are orthonormal, and they are a basis for V
Recall that a basis is linearly independent and spans V
Also recall that for finite dimensional vector spaces that if the number of orthonormal vectors is equal to $\dim(V)$ then those vectors are a bssis (linear independence lemma)

1.3.5. Remark

Let V be a vector space with dimension n. Let $\{e_i\}_{i=1}^n$ be an orthonormal basis and $\{v_i\}_{i=1}^n$ be a basis (for V).

The benefits of orthnormal basis: easy to verify that a list is orthonormal basis

Another benefit is writing any arbitrary vector in V as a linear combination of the orthonormal basis (LADR 6.29). This is also called "projecting onto the list of vectors" i.e. if you have a vector $v = a_1 e_1 + \ldots + a_n e_n = b_1 v_1 + \ldots + b_n v_n$ then $a_i = \langle v, e_i \rangle = \langle a_i e_i + \ldots + a_n e_n, e_i \rangle = a_i \langle e_i, e_i \rangle$. In contrast, to compute $b_1, \ldots, b_n$ we need to solve the big system of linear equations. Computationally, the running time is $O(n)$ and $O(n^3)$ if use gaussian elimination and $O(n), O(n^2)$ space respectively

1.3.6. *Proposition*

If $v = a_1 e_1 + \ldots + a_n e_n$ then $\|v\|_2^2 = |a_1|^2 + \ldots\ |a_n|^2$ (LADR 6.30b)

Proof: By the pythagorean theorem (prove this yourself)

$\square$

1.3.7. *Theorem*

Any linearly independent list can be turn into an orthnormal list, where both lists have the same span

Proof: By the Grahm-Schmidt procedure (LADR 6.32)

$\square$

1.3.8. Remark

Applicationof orthonormal bases: discretization of wave function (such as the recording of a voice).
Lengdre polynmoials

1.4. January 23

1.4.1. *Theorem*

> Every finite-dimensional inner product space has an orthonormal basis.
> Suppose V is finite-dimensional. Then every orthonormal list of vectors in V can be extended to an orthonormal basis of V.

Proof: See LADR 6.35 and 6.36

□

1.4.2. Definition: Linear Functional

A linear functional on V is a linear map from V to $\mathbb{F}$. Denote the space of linear functions as $L(V, \mathbb{F})$

1.4.3. Remark

1. Range is a “number”
2. $L(V, W) \to L(V, \mathbb{F})$ if we set $W = \mathbb{F}$
3. $\| \cdot \| : V \to \mathbb{R}$ is in general not linear
4. $\langle v, u \rangle$: fix u choose $v \in V$, then this inner product is a linear functional (recall that it will be linear in the first (as well as second in the real case) slots)

1.4.4. Example

Define the map $\phi : P_2([-1, 1]) \to \mathbb{R}$ where $P(x) = \int_{-1}^{1} p(x) \cos(\pi x) \mathrm{d}x$

Excercise: show that ϕ is linear.

Note that $\phi(p) \overset{\text{def}}{=}$ (continued in lecture notes)

1.4.5. *Theorem:* Riesz Representation Theorem

> Given a finite-dimensional inner product space V. Let ϕ be a linear functional on V. Then there exist a unique element $u \in V$ such that $\phi(v) = \langle v, u \rangle$ for any $v \in V$
> We say that u is the representatin of ϕ in V.

Proof:

Find a u that satisfies the theorem, and prove that u is unique.

Assume $\{e_i\}_{i=1}^{n}$ is an orthonormal basis of V.

Let $v = \langle v, e_1 \rangle e_1 + ... + \langle v, e_n \rangle e_n$

Then $\phi(v) = \phi(\langle v, e_1 \rangle e_1 + ... + \langle v, e_n \rangle e_n) = \langle v, e_1 \rangle \phi(e_1) + ... + \langle v, e_n \rangle \phi(e_n) = \langle v, \overline{\phi(e_1)} e_1 \rangle + ... + \langle v, \overline{\phi(e_n)} e_n \rangle = \langle v, u \rangle$

If we define $u = \overline{\phi(e_1)} e_1 + ... + \overline{\phi(e_n)} e_n \in V$

Excercise: prove that u is unique.

Let u_1, u_2 and $u_1 \neq u_2$

Consider $u_1 - u_2 = 0$ and find a contradiction (hint: in the book)

$\square$

1.4.6. Remark

u_ϕ is the representation of ϕ, u_φ is representation of linear functional φ, etc.

Note the following bijection: $L(V, \mathbb{F}) \cong V$. In other words $\phi \in L(V, \mathbb{F}) \leftrightarrow u_\phi$ is a one to one map. It follows that for f.d.v.s that $\dim(L(V, \mathbb{F})) = \dim(V)$

For a fixed ϕ, its represenation u_ϕ is uniform regardless of input v

1.4.7. Example: possibilities of Riesz Rep Theorem

consider $f(x) : \mathbb{R}^n \to \mathbb{R} \leftrightarrow u = (x_1, x_2, ..., x_n)$

Another example: consider the previous linear functioal $\phi(p)$:

$\phi(p) = \int_{-1}^{1} p(x) \cos(\pi x) \mathrm{d}x \leftrightarrow q \in P_2([-1, 1])$

Then recall that $u_\phi = \phi(e_1)e_1 + \phi(e_2)e_2 + \phi(e_3)e_3 \in P_2([-1, 1])$ where $\{e_1, e_2 e_3\}$ form an ON basis.

Using grahm schmidt we convert $\{1, x, x^2\} \to \left\{\sqrt{\frac{1}{2}}, \sqrt{\frac{3}{2}}x, \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right)\right\}$

And compute that $u_\phi = -\frac{45}{2\pi^2}\left(x^2 - \frac{1}{3}\right)$

Then $\phi(p) = \langle p, \cos(\pi x) \rangle = \langle p(x), -\frac{45}{2\pi^2}\left(x^2 - \frac{1}{3}\right) \rangle$

In fact, we have $\cos(\pi x) = u_\phi + w(x)$ such that $w(x)$ is orthogonal to the orthonormal basis $\{e_1, e_2, e_3\}$

i.e. that $\int_{-1}^{1} e_i(x) w(x) = 0$ for $e_1 = 1, e_2 = x, e_3 = x^2$

Recall that the definition of inner product over the space of $P_2([-1, 1])$ is $\langle p_1(x), p_2(x) \rangle = \int_{-1}^{1} p_1(x) p_2(x) \mathrm{d}x$

Note that $\langle \cos(\pi x), e_1 \rangle = \langle u_\phi + w(x), e_1 \rangle = \langle u_\phi, e_1 \rangle + \langle w(x), e_1 \rangle = \langle u_\phi, e_1 \rangle$ and since $w(x)$ is orthogonal we can simply remove $w(x)$ from $\cos(\pi x)$ and still obtain the same vector in $P_2([-1, 1])$ (expand to all e_i)

1.4.8. Remark

Remark: the Riesz Rep theorem can be generalized to infinite-dimensional space

1.5. January 28

1.5.1. Remark

Midterm 1

LADR 6A, 6B (no upper triangular matrix parts e.g. schur's theorem)

inner products, norm, orthogonality, orthonormal bases, gram-schmidt

Some reminders:

Some inner products are defined explicitly (such as the dot product)

Recall different norms: $\|u\|_1, \|u\|_2, \|u\|_{\max}$

Examples of norm squared include $\|f\|^2 = \int_{-1}^{1} f^2 \mathrm{d}x$ or $\|f\|^2 = \int_{-\infty}^{\infty} f^2 e^{-\frac{x}{2}} \mathrm{d}x$ and etc.

Recall the 4 axioms of inner product (5 in textbook) and 3 axioms of norms

Cauchy-Schwarz, parallelogram, triangle, pythagorean, etc.

Review linearity, injectivity, surjectivity

compute gram schmidt by hand (!)

compute Riesz represenation by hand (!)

other review: some integration (calculation), basis, dimension, linear independence

true false, proof, computation

1.5.2. Remark

Riesz represenation: questions for future lectures (not on final):

Recall that our construction for the "representation" of a linear functional depends on the choice of orthonormal basis

For the same functional, its representation varies depending on the choice of V

We now turn to section (3E) product/quotient space and (6C) (orthogonal complements)

1.5.3. Note

The inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ has a domain $V \times V$, where $V \times V$ is a product space

1.5.4. *Proposition*

The product space $V_1 \times \ldots \times V_n$ is a vector space.
Where we define $V_1 \times \ldots \times V_n = \{(v_1, v_2, \ldots, v_n) : v_i \in V_i\}$.

And for this set, vector addition and scalar multiplication is coordinate wise / distributes to each coordinate as usual.

Proof:

We need to check the vector space axioms

□

1.5.5. *Theorem*

$\dim(V_1 \times \ldots \times V_n) = \dim(V_1) + \ldots + \dim(V_n)$

Proof:

In LADR 3.92

Remark: observe that $\mathbb{R}^4 = \mathbb{R} \times \mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}^2$ (all of these product space are equal). i.e. $(a, b, c, d) = (a, b, c) \times (d) = (a, b) \times (c, d)$

We can consturct the basis

$$(e_1, 0, \ldots, 0), (e_2, 0, \ldots, 0), \ldots, \left(e_{n_1}, 0, \ldots, 0\right)$$
$$(0, f_1, \ldots, 0), (0, f_2, 0 \ldots, 0), \ldots, \left(0, f_{n_2}, 0, \ldots, 0\right)$$

Where e_i is a basis for V_1, f_i is a basis for V_2 etc. and their cooresponding dimensions are $n_1, n_2, \ldots$

Alternatively, we can consider some isomorphisms

Consider $V_1 \cong \mathbb{R}^{n_1}, \ldots, V_n \cong \mathbb{R}^{n_n}$ by LADR 3.70

Then we can try to see if $V_1 \times \ldots \times V_n$ is isomphic to $\mathbb{R}^{n_1} \times \ldots \times \mathbb{R}^{n_n} = \mathbb{R}^{n_1 + \ldots + n_n}$

□

1.5.6. Definition: Orthogonal complement

Let U, V be vecor spaces such that $U \subseteq V$ Denote the orthogonal complement of U to be $U^\perp$ defined by

$$U^\perp = \{v \in V : \langle v, u \rangle = 0 (\forall u \in U)\}$$

Properties (LADR 6.48)

$U^\perp$ is a subspace (note that $0 \in U^\perp$)

$\{0\}^\perp = V$

$V^\perp = \{0\}$

$U \cap U^\perp \subseteq \{0\}$. In particular if U is a subspace then $U + U^\perp = U \oplus U^\perp$ is a direct sum. Recall that sum of U, V is a direct sum iff $U \cap V = \{0\}$

If U, W are subsets of V with $U \subseteq W$ then $W^\perp \subseteq U^\perp$

1.5.7. Example: Orthogonal complement (has potential issues)

Consider the z axis (represented by U)

Let $U = \{k(0,0,1) : k \in \mathbb{R}\}$ then the orthogonal complement is $U^\perp = \left\{v \in \mathbb{R}^3 : v_1 \cdot 0 + v_2 \cdot 0 + v_3 \cdot k = 0\right\} = \left\{ \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix} : v_1, v_2 \in \mathbb{R} \right\}$ (i.e. the complement is the x-y plane)

1.6. Feb 4

1.6.1. Note

Midterm 1

Median 14.4/18

Mean 13.82/18

SD 3.7/18

reflection: present in lecture earn 2.4/18 extra, OH earn 1.6/18 extra

midterm 1 bonus: present between feb 4 and feb 13 in lectures, feb 4 and feb 20 in OH

if not selected for lectures, can present in office hours

Continuation of LADR 6.C + 3.E

1.6.2. *Proposition*

> If U, W are subsets of V with $U \subseteq W$ then $U^{\perp} \supseteq W^{\perp}$

Proof:

WTS: if $v \in W^{\perp}$ show that $v \in U^{\perp}$

According the definition of orthogonal complement, the equivalent statements are that for any $w \in W$, $\langle v, w \rangle = 0$ and for any $u \in U$, $\langle v, u \rangle = 0$

Note that for any $u \in U$, since $U \subseteq W$, $u \in W$. Given $v \in W^{\perp}$, $\langle v, w \rangle = 0$ for any $w \in W$, so $\langle v, u \rangle = 0$ since $u \in W \forall u$

$\square$

1.6.3. *Theorem:* Decomposite Vector Space

> Given a subspace (finite dim) U of vector space V we can decompose V as follows:
>
> $$V = U \oplus U^{\perp}$$
>
> (We call U the "main part" and $U^{\perp}$ the "residue" part)

Proof:

Let an orthonormal basis of U be $\{e_1, e_2, ..., e_m\}$ (guaranteed to exist by gram-schmidt)

First we should prove that $V = U \oplus U^{\perp}$ by showing that each $v \in V$ can be written as a sum of an element from U and and an element from $U^{\perp}$

Define $v = \langle v, e_1 \rangle e_1 + ... + \langle v, e_m \rangle e_m + v - \langle v, e_1 \rangle e_1 - ... - \langle v, e_m \rangle e_m$

Consider the inner product $\langle w, e_k \rangle = \langle v, e_k \rangle - \langle v, e_k \rangle = 0$

Thus we know that w is orthogonal to every vector in $\text{span}\{e_1, ..., e_m\}$ which shows that $w \in U^{\perp}$

Second, we can not that $U \cap U^{\perp} = \{0\}$ and therefore the sum is a direct sum

$\square$

1.6.4. *Corollary*

> if V is f.d.v.s. and U is a subspace of V then
>
> $$\dim(U) + \dim(U^{\perp}) = \dim(V)$$

Proof:

Follows from the direct sum properties (LADR 3.94)

$\square$

1.6.5. Example: Orthogonal Complement

Let $U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : x, y \in \mathbb{R} \right\}$ (note that this is a subset, not a subspace)

Let the orthogonal complement be $U^\perp = \left\{ v = \begin{pmatrix} a \\ b \\ c \end{pmatrix} : \langle v, u \rangle = ax + by + 2c = 0 \forall x, y \in \mathbb{R} \right\}$

Since $a = b = c = 0$ by necessity then $U^\perp = \{(0, 0, 0)\}$

Note that we cannot compute $\dim(\mathbb{R}^3) = \dim(U) + \dim(U^\perp)$ since U is only a subset (not a subspace)

"Affine space?" (TODO)

1.6.6. Example: Orthogonal Complement

The orthogonal complement of the xy plane is the z axis

Call them $W, W^\perp$ respectively then $W = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\}$ and $W^\perp = \left\{ \begin{pmatrix} 0 \\ 0 \\ k \end{pmatrix} : k \in \mathbb{R} \right\}$

Then $\dim(\mathbb{R}^3) + \dim(W) + \dim(W^\perp) = 2 + 1 = 3$

1.6.7. *Theorem*

> (LADR 6.52) If U is a finite-dim subspace of V then $U = (U^\perp)^\perp$

Proof:

1. First, let $u \in U$ then for all $w \in U^\perp$ we have $\langle u, w \rangle = 0$. Since each u is orthogonal to any element in $U^\perp$, it follows that $u \in (U^\perp)^\perp$ and therefore $U \subseteq (U^\perp)^\perp$
2. Second, let $v \in (U^\perp)^\perp$ then $v = u + w$ where $u \in U, w \in U^\perp$ (since $V = U \oplus U^\perp$)

First observe that $u \in U \Rightarrow u \in (U^\perp)^\perp$ and therefore $v - u \in (U^\perp)^\perp$

Next observe that $v - u = w \in U^\perp$

Hence $v - u \in U^\perp \cap (U^\perp)^\perp \Rightarrow v - u = 0 \Rightarrow v = u \Rightarrow v \in U$. Therefore $(U^\perp)^\perp \subseteq U$

Since we have proven both inclusions the claim follows.

$\square$

1.6.8. Definition: Orthogonal projection of V onto a subspace U

Define the orthogonal projection of V onto a subspace U by $P_U : V \to V$ which maps $v \in V \to u \in U$ where $v = u + w$ such that $u \in U, w \in U^\perp$

Note that $P_U(u) = u$ (u is already in U). In particular, $U = \{ku : u \in V, k \in \mathbb{F}\}$

An explicit formula is $P_U(v) = \frac{\langle v,u \rangle}{\langle u,u \rangle} u$ (note the connection to the orthogonal decomposition)

1.7. Feb 6

1.7.1. Note

Last Time: orthogonal projection

Given a subspace $U \subseteq V$, $v = u + w$ (vectors), $V = U \oplus U^{\perp}$ (for vector spaces)

where $u \in U, w \in U^{\perp}$

We have the map $P_U : V \to V$ which maps $v \to u = P_U(v)$

1.7.2. Note

Orthogonal Projection is one type of projection. In general, we can have more general projections onto some subspace.

It's just that in orthogonal projection, the projection must be done such that it is orthogonal to the subspace being projected onto

1.7.3. Definition: Projection

Consider $P : V \to V$, then P is a projection if $P(P(v)) = P(v)$ i.e. $P^2 = P$ ("two consecutive operations equivalent to one operation")

1.7.4. Definition: Orthogonal Projection

P is an orthogonal projection if

$$\begin{cases} P^2 = P \\ (\langle P(v), u\rangle = \langle v, P(u)\rangle)(\text{this is not true in the case of general projection}) \end{cases}$$

A geometric interpretation is $\langle v, P(u)\rangle = \|v\| \cdot \cos\theta \; \|P(u)\| = \|P(v)\| \cdot \|P(u)\| = \langle P(v), u\rangle$

1.7.5. Remark

LADR 6.57 lists some properties of orthogonal projection (some properties also hold for general projections)

One strategy to prove these definitions might be to let $v = P_U(v) + w$ with $P_U(v) \perp w$ then this implies that $\langle P_U(v), v - P_U(v)\rangle = \langle P_U(v), w\rangle = 0$ (see the definition of orthogonal projection)

1.7.6. ***Proposition***

$\|P_U(v)\| \leq \|v\|$

Proof:

Note the pythagorean theorem: $\|P_U(v)\|^2 + \|v - P_U(v)\|^2 = \|v\|^2$

□

1.7.7. ***Proposition***

Let $U = \texttt{span}\{e_1, e_2, ..., e_n\}$ be an orthonormal basis. Then $P_U(v) = \langle v, e_1\rangle e_1 + ... + \langle v, e_n\rangle e_n$

Proof:

Not given in lecture (but should be obvious considering the representation of elements in vector space with respect to orthonormal bases)

□

1.7.8. ***Theorem***

Let U be a subspace of V, for any $v \in V$ then

$$\|v - P_U(v)\| = \min_{u \in U} \| \ v - u\| \leq \|v - u\| \text{ for any } u \in U$$

1.7.9. Remark

$\text{dist}(v, u) = \min_{u \in U} \|v - u\|$ and $P_U(v) = \text{argmin}_{u \in U} \ \|v - u\|$

Proof:

$$\|v - P_U(v)\|^2 \leq \|v - P_U(v)\|^2 + \|P_U(v) - u \ \|^2 = \|v - P_U(v) + P_U(v) - u\|^2 = \|v - u\|^2$$

(LADR 6.61)

□

1.7.10. Notation

P_U : orthogonal projection onto a subspace U

P_C : orthogonal projection onto a closed convex subset C (which is a generalization of P_U)

Such P_C is fundamental in convex analysis/optimization

1.7.11. Definition: P_C

define P_C to be

$$P_C = \arg\min_{y \in C} \ \|y - x\| = \arg\min_{y \in C} \ \|y - x\|^2$$

We can also write $P_C(x) = \arg\min_{y \in C} \frac{1}{2} \ \|y - x\|^2$ for cleaner derivatives

1.7.12. Problem 1

> Example Problem: Given $x = (x_1, x_2, ..., x_n)$
>
> (i) find $\min \|y - x\|^2$ s.t. $y \in \mathbb{R}^n_{\geq 0}$
>
> (ii) find $\min \|y - x\|^2$ s.t. $y \in$ unit ball i.e. $\|y\| \leq 1$

Solution

We can reformulate the problem: $C = \{y \in \mathbb{R}^n : y \text{ subject to constraint}\}$

Then the constrained optimization problem becomes $\operatorname{dist}(x, c)$ = optimal obj value.

In other words $P_C(x)$ = minimizer solution (follows from previous theorem)

Therefore we would like to study operators such as $P_C : V \to V$

Examples of possible C might be $C = \{ku : k \in \mathbb{R}\} = \operatorname{span}\{u\}$ (note that $P_C(v) = \frac{\langle v,u \rangle}{\langle u,u \rangle} u$) or $C = \{\|y\| \leq 1\}$ (note that here,

$$P_{C(x)} = \left\{ x \text{ if } \|x\| \leq 1; \frac{x}{\|x\|} \text{ if } \|x\| > 1 \right\}$$

) (either pick the point itself is in the ball, or normalize the vector to length 1)

1.7.13. *Theorem:* Orthogonal Projection characterization

> Given a closed convex subset C, for every $x \in V$: $z = P_C(x)$ if and only if $\langle y - z, x - z \rangle = \|y - z\| \cdot \|x - z\| \cos\theta \leq 0$ for any $y \in C$
>
> In other words, the angle between $y - z$ and $x - z$ is greater than 90 degrees

1.8. Feb 11

1.8.1. Note

Last time:

$P_U(v) = \arg\min_{u \in U} \|u - v\|^2 = \arg\min_{u \in U} \|u - v\| = \arg\min_{u \in U} \frac{1}{2} \|u - v\|^2$

$\operatorname{dist}(v, U) = \min_{u \in U} \|u - v\|$

Generalize: $P_C(x) = \arg\min(y \in C)\ \|x - y\|^2$

1.8.2. *Theorem:* characterization theorem for orthogonal projection

> $z = P_C(x)$ if and only if $\langle y - z, x - z \rangle \leq 0 \forall y \in C$
>
> Recall this means the angle if $\geq 90^{\circ}$ or $\frac{\pi}{2}$

Proof:

Assume that $\langle y - z, x - z \rangle \leq 0$

Note that

$\|y - x\|^2 = \|y - z + z - x\|^2 = \|y - z\|^2 + \|z - x\|^2 + 2\langle y - z, z - x \rangle = \|y - z\|^2 + \|z - x\|^2 - 2\langle y - z, x - z \rangle$

We conclude with $\|x - z\|^2 \leq \|x - y\|^2\ \forall y \in C$ i.e. $\|x - z\| \leq \|x - y\|$

Assume $z = P_C(x)$, that is $\|x - z\|^2 \leq \|x - y\|^2\ \forall y \in C$

create a $y_t = ty + (1 - t)z$ which represents a point in C. Because we need a convex combination we require that $0 \leq t \leq 1$

Note that $\|x - z\|^2 \leq \|x - y_t\|^2$ for any y and any $0 \leq t \leq 1$

Definte $F(t) = \|x - y_t\|^2$. F is continuous/differentiable in terms of t. If $t = 0, y_t = 0 \cdot y + (1 - 0)z = z$. By assumption $F(t)$ achieve its minimum at $t = 0 \Leftrightarrow \frac{dF(t)}{d} t \mid_{t=0} = 0$

Then, $F(t) = \|x - (y + (1 - y)z)\|^2 = \|x - z - t(y - z)\|^2 = \|x - z\|^2 + t^2\|y - z\|^2 - 2t\langle x - z, y - z \rangle$

$\left(D \frac{F(t)}{\mathrm{d}t}\right)_{t=0} = 2\|y - z\|^2\ t - 2\langle s - z, y - z \rangle$

$\left(\frac{dF(t)}{F(t)}\right)_{t=0} \geq 0 \Leftrightarrow -2\langle x - z, y - z \rangle \geq 0 \Leftrightarrow \langle x - z, y - z \rangle \leq 0$

We need to show the above line ($F(t)$ is increasing at $t = 0$ for any y)

$\square$

1.8.3. Definition: Convex set

any convex combination of two points is closed in the set

1.8.4. Remark

Properties of P_C

$\|P_C(x) - P_C(y)\| \leq \|x - y\|$ (consequentially nonexpansive)

$\langle P_C(x) - P_C(y), x - y \rangle \geq \|P_C(x) - P_C(y)\|^2$ (firmly nonexpansive)

$P_C : V \to V$ is a contraction mapping

We generalize projections onto vector spaces to projections onto convex sets

1.8.5. Notation

$v + U$:

for a $U \subseteq V$ we denote $v + U = \{v + u : u \in U\}$ which is a subset of V

1.8.6. Remark

If $v \notin U$ then $0 \notin v + U \Rightarrow v + U$ is not a subspace

1.8.7. Remark

Let $U = \{y = 2x\}$ and $v = (2, 0)$ in $\mathbb{R}^2$

We say that $v + U$ is parallel to U and also say that $v + U$ is an affine subset

We can also observe that if we let $w = v + u$ for some $u \in U$ that $v + U = w + U$

1.8.8. *Proposition*

> These are equivalent
> (a) $v + U = w + U$
> (b) $w - v \in U$
> (c) $(v + U) \cap (w + U) \neq \emptyset$ (LADR 3.101)

Proof: in book

□

1.8.9. Definition: quotient space

The quotient space denoted $V \setminus U = \{v + U : v \in V\}$. This is a collection of subsets $\{v_1 + U, v_2 + U, \ldots\}$

1.8.10. Remark

$V \setminus U \nsubseteq V$ is not a valid statement since $V \setminus U$ has elements that are subsets of vectors, while V has elements that are vectors

1.8.11. Remark

To add $v_2 + U$ to the set $\{v_1 + U\}$ we first need to check if $v_2 - v_1 \notin U$ (otherwise $v_1 + U = v_2 + U$)

1.9. Feb 13

1.9.1. Remark

Last Time: For U a subspace of V call the set $V \setminus U = \{v + U : v \in V\}$ the set of all affine subsets parallel to U

1.9.2. Example

Let $U = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\}$ be the x-y plane. Then the quotient space (see the picture in lecture notes) of planes parallel to U

1.9.3. *Theorem:* $V \setminus U$ is a vector space

> Define addition and scalar-multiplication as follows:
>
> Let $v_1 + U, v_2 + U \in V \setminus U$ where $v_1, v_2 \in V$ and let $\lambda \in \mathbb{F}$
>
> Addition: $v_1 + U + (v_2 + U) = (v_1 + v_2) + U$ (the addition is not a binary operation on vectors)
> Scalar Multiplication: $\lambda(v + U) = \lambda v + U$

Proof:

exercise

□

1.9.4. Remark

a vector in $V \setminus U$ is affine subset

the zero vector has multiple representations: $0 \in V \setminus U = U = 0_{\in V} + U = u_{\in U} + U$

1.9.5. Definition: quotient map

Call the map $\pi : V \to V \setminus U$ to the quotient map defined by $v \to v + U$

1.9.6. Remark

The quotient map is linear

1.9.7. Example: quotient map

An example of a quotient map is $\pi : \mathbb{Z} \to \mathbb{Z} \setminus 7$ where $U = \{7k : k \in \mathbb{Z}\}$

$0 \to 0 + U$

$1 \to 1 + U$

$\vdots$

$7 \to 7 + U = 0 + U$

$8 \to 8 + U = 1 + U$

$\vdots$

1.9.8. Remark

Here π is not injective, though it is surjective

1.9.9. Definition:

Given $T \in L(V, W)$ define $T^{\sim} : V \setminus \text{Null}(T) \to W$ by $v + \text{Null}(T) \to w$ such that $T^{\sim}(v + \text{Null}(T)) = T(v)$

1.9.10. Example: application of $T^{\sim}$

Solve inhomogenous system such as $T(v) = b$ i.e. $Av = b$

Find a particular v_0 such that $T(v_0) = b$ i.e. $Av_0 = b$

(?? TODO)

1.9.11. Remark

In general T is a map if each input corresponds to one output(no multiple outputs)

We can call T well defined or not well defined respectively.

Is it possible that $T^{\sim}(v + \text{Null}(T)) \to T(v) = T^{\sim}(v + u_0 + \text{Null }(T))$ if $T(u_0) = 0$?

1.9.12. *Proposition*

$T^{\sim}$ is well defined

Proof:

Want to show that if $v_1 + \text{Null}(T) = v_2 + \text{Null }(T)$ then $T^{\sim}(v_1 + \text{Null}(T)) = T^{\sim}(v_2 + \text{Null}(T))$

(continued in lecture notes (TODO))

$\square$

1.9.13. Definition: linear functional

A linear functional is a map $\phi : V \to \mathbb{F}$ where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$

1.9.14. Note

In this class we will work with $\mathbb{R}$, i.e. $\phi(v) \in \mathbb{R}$ and $\phi \in L(V, \mathbb{R})$

1.9.15. Definition: dual space of V

The collections of linear functionals on $\mathbb{R}$

We denote the dual space of V by $V' - L(V, \mathbb{R})$

1.9.16. *Theorem:* dimension of V'

$$\dim(V') = \dim(L(V, \mathbb{R})) = \dim(V) \cdot \dim(\mathbb{R}) = \dim(V)$$

Proof: Also in the book (LADR 3.111)

□

1.9.17. *Corollary*

V is isomorphic to V' i.e. each vector in V corresponds to a linear functional in V'

Proof: By 1.9.16 and LADR 3.70

□

1.9.18. Remark

1.9.17 motivates the riesz-representation calculation:

For any functional $\phi \in V'$ find its representation in V by $V_\phi = \phi(e_1)e_1 + \ldots + \phi(e_n)e_n$ where $\{e_1, \ldots, e_n\}$ is orthonormal basis in V

1.10. Feb 18

1.10.1. Note

Midterm 2

True False + 3 free response

Covers 3E (product spaces and quotient spaces)

Covers 6C (orthogonal complement, orthogonal projection)

Old Material that may come up: inner pdocuts, orthonormal bases, riesz representation

Recall the orthogonal projection operator P_U which maps $v \to P_U(v) + w$ where $w \perp U$

Characterization of $P_U(v)$ (1.8.2): $\langle v - P_U(v), u - P_U(v) \rangle \leq 0$

Applications of P_U such as minimization problem i.e. $P_U(v) = \arg\min_{u \in U} \|v - u\|^2$

Review various properties (eg relating to dimension) and interesting maps such as $T : V_1 \times \ldots \times v_n \to V_1 \oplus \ldots \oplus V_n$ and $\pi : V \to V \setminus U$ (are these maps injective or surjective?)

Dual spaces: $V' = L(V, \mathbb{R})$ as a special case of $L(V, W)$

Dual space connection to riesz representation theorem

prop 1.8.8

Interesting topica (Not in Exams)

1.10.2. Remark

Recall orthogonal projections: for subspace $U, v \in V$ and then we can find a $w = (v - P_U(v)) \perp U$

Now consider if $U = \mathtt{span}\{a_1, a_2, \ldots, a_k\} \subseteq \mathbb{R}^n$ and matrix $A = \begin{pmatrix} \vdots & \vdots & \ldots & \vdots \\ a_1 & a_2 & \ldots & a_k \\ \vdots & \vdots & \ldots & \vdots \end{pmatrix}$ and $P_U(v) \in U$

Then $P_U(v) = x_1 a_1 + x_2 a_2 + \ldots + x_k a_k = Ax$ where $x_1, \ldots, x_k$ are "weights".

fix a constant vector $b \in \mathbb{R}^n$ then $v - P_U(v) = b - Ax$ and note that U is the column space of A

Notice that $b - Ax \in \mathrm{Null}(A)$ is equivalent to saying $b - Ax$ is orthogonal to the column space

In other words $\forall a_i, a_i \cdot (b - Ax = 0)$ i.e. $A^T(b - Ax) = 0$

We would like to study how

$$A^T(b - Ax) = 0$$
$$A^T A x = A^T b$$
$$x = (A^T A)^{-1} A^T b$$

as long as $A^T A$ is invertible.

Reconsidering $P_U : V \to A$ where $b \in V \to (A^T A)^{-1} A^T b$ then $(A^T A)^{-1} A^T$ is the (orthogonal) projection matrix

An application of this is solving $Ax = b$ when A is not invertible. We can try $x = (A^T A)^{-1} A^T b$

And $x = \arg\min \|Ay - b\|_2^2$ is probably a minimizer of some optimization problem

Think linear regression or least squared problem from machine learning