Chapter 6: Inner Product Spaces

Linear Algebra Done Right (4th Edition), by Sheldon Axler

Last updated: December 20, 2024

Contents

6A: Inner Products and Norms

Definition 1 (dot product)**.** *For $x, y \in \mathbb{R}^n$, the **dot product** of x and y, denoted by $x \cdot y$, is defined by*

$$x \cdot y = x_1 y_1 + \cdots + x_n y_n,$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$.

Definition 2 (inner product)**.** *An **inner product** on V is a function that takes each ordered pair (u, v) of elements of V to a number $\langle u, v\rangle \in \mathbb{F}$ and has the following properties:*

(a) ***positivity**: $\langle v, v\rangle \geq 0$ for all $v \in V$.*

(b) ***definiteness**: $\langle v, v\rangle = 0$ if and only if $v = 0$.*

(c) ***additivity in first slot**: $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$.*

(d) ***homogeneity in first slot**: $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$ for all $\lambda \in \mathbb{F}$ and all $u, v \in V$.*

(e) ***conjugate symmetry**: $\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$.*

Definition 3 (inner product space)**.** *An **inner product space** is a vector space V along with an inner product on V.*

Corollary 4 (basic properties of an inner product)**.** *(a) For each fixed $v \in V$, the function that takes $u \in V$ to $\langle u, v\rangle$ is a linear map from V to $\mathbb{F}$.*

(b) $\langle 0, v\rangle = 0$ for every $v \in V$.

(c) $\langle v, 0\rangle = 0$ for every $v \in V$.

(d) $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$ for all $u, v, w \in V$.

(e) $\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle$ for all $\lambda \in \mathbb{F}$ and $u, v \in V$.

Definition 5 (norm, $\|v\|$)**.** *For $v \in V$, the **norm** of v, denoted by $\|v\|$, is defined by*

$$\|v\| = \sqrt{\langle v, v\rangle}$$

Corollary 6 (basic properties of the norm)**.** *Suppose $v \in V$.*

(a) $\|v\| = 0$ if and only if $v = 0$.

(b) $\|\lambda v\| = |\lambda|\|v\|$ for all $\lambda \in \mathbb{F}$.

Remark 7. *Working with norms squared is usually easier than working directly with norms.*

Definition 8 (orthogonal)**.** *Two vectors $u, v \in V$ are called **orthogonal** if $\langle u, v\rangle = 0$.*

Corollary 9 (orthogonality and 0). *(a) 0 is orthogonal to every vector in V.*

(b) 0 is the only vector in V that is orthogonal to itself.

Theorem 10 (Pythagorean Theorem). *Suppose $u, v \in V$. If u and v are orthogonal, then*

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2$$

Lemma 11 (orthogonal decomposition). *Suppose $u, v \in V$, with $v \neq 0$. Set $c = \frac{\langle u,v\rangle}{\|v\|^2}$ and $w = u - \frac{\langle u,v,\rangle}{\|v\|^2} v$. Then*

$$u = cv + w \quad \text{and} \quad \langle w, v\rangle = 0.$$

Theorem 12 (Cauchy-Schwarz inequality). *Suppose $u, v \in V$. Then*

$$|\langle u, v\rangle| \leq \|u\|\|v\|$$

This inequality is an equality if and only if one of u, v is a scalar multiple of the other.

Theorem 13 (triangle inequality). *Suppose $u, v \in V$. Then*

$$\|u+v\| \leq \|u\| + \|v\|.$$

This inequality is an equality if and only if one of u, v is a nonnegative real multiple of the other.

Theorem 14 (parallelogram equality). *Suppose $u, v \in V$. Then*

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

Problem 1

Prove or give a counter example: If $v_1, \ldots, v_m \in V$, then

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\langle v_j, v_k\rangle \geq 0.$$

Proof. By linearity of inner products,

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\langle v_j, v_k\rangle = \left\langle \sum_{j=1}^{m} v_j, \sum_{k=1}^{m} v_k \right\rangle \geq 0$$

since the two terms equal other and the conclusion follows by positivity of inner products. □

Problem 2

Suppose $S \in \mathcal{L}(V)$. Define $\langle \cdot, \cdot \rangle_1$ by

$$\langle u, v\rangle_1 = \langle Su, Sv\rangle$$

for all $u, v \in V$. Show that $\langle \cdot, \cdot \rangle_1$ is an inner product on V if and only if S is injective.

Proof. $\langle \cdot, \cdot \rangle_1$ is inner product $\iff \langle v, v\rangle_1 = \langle Sv, Sv\rangle = 0$ if and only if $v = 0$ $\iff S$ is injective. (Other properties are omitted for checking) □

Problem 3

(a) Show that the function taking an ordered pair $((x_1, x_2), (y_1, y_2))$ of elements of $\mathbb{R}^2$ to $|x_1 y_1| + |x_2 y_2|$ is not an inner product on $\mathbb{R}^2$.

(b) Show that the function taking an ordered pair $((x_1, x_2, x_3), (y_1, y_2, y_3))$ of elements of $\mathbb{R}^3$ to $x_1 y_1 + x_3 y_3$ is not an inner product on $\mathbb{R}^3$.

Proof. (a) Consider $x = (2, -2)$ and $y = (-2, 2)$ and $z = (1, 1)$. Then $\langle x, z\rangle = \langle y, z\rangle = 4$, but $\langle x + y, z\rangle = 0$.

(b) We have $\langle (0, 1, 0), (0, 1, 0)\rangle = 0$ but the element is nonzero. □

Problem 4

Suppose $T \in \mathcal{L}(V)$ is such that $\|Tv\| \leq \|v\|$ for every $v \in V$. Prove that $T - \sqrt{2}I$ is injective.

Proof. Suppose for contradiction that $T - \sqrt{2}I$ is not injective and therefore $\sqrt{2}$ is an eigenvalue of T, so $Tv = \sqrt{2}v$ for some nonzero v. Taking the norm yields that

$$\|Tv\| = \sqrt{2}\|v\|$$

which violates the assumption that $\|Tv\| \leq \|v\|$. $\square$

Problem 5

Suppose V is a real inner product space.

(a) Show that $\langle u+v, u-v\rangle = \|u\|^2 - \|v\|^2$ for every $u, v \in V$.

(b) Show that if $u, v \in V$ have the same norm, then $u+v$ is orthogonal to $u-v$.

(c) Use (b) to show that the diagonals of a rhombus are perpendicular to each other.

Proof. (a) We have that

$$\begin{aligned}\langle u+v, u-v\rangle &= \langle u, u\rangle - \langle v, v\rangle - \langle u, v\rangle + \langle v, u\rangle \\ &= \|u\|^2 - \|v\|^2\end{aligned}$$

(b) We know $\|u\| = \|v\|$, then

$$\langle u+v, u-v\rangle = \|u\|^2 - \|v\|^2 = 0$$

which shows that they are orthogonal.

(c) omitted. $\square$

Problem 6

Suppose $u, v \in V$. Prove that $\langle u, v\rangle = 0 \iff \|u\| \leq \|u + av\|$ for all $a \in \mathbb{F}$.

Proof. $\Rightarrow$ Given $\langle u, v\rangle = 0$, then

$$\begin{aligned}\|u+av\|^2 &= \langle u+av, u+av\rangle \\ &= \|u\|^2 + \overline{a}\langle u, v\rangle + |a|^2\langle v, v\rangle \\ &= \|u\|^2 + |a|^2\langle v, v\rangle \\ &\geq \|u\|^2\end{aligned}$$

$\Leftarrow$ If $v = 0$, then it's trivial. Consider $v \neq 0$. Let $a = \frac{\langle u,v\rangle}{\|v\|^2}$. Then we have that

$$\begin{aligned}
\left\| u - \frac{\langle u,v\rangle}{\|v\|^2} v \right\|^2 &= \left\langle u - \frac{\langle u,v\rangle}{\|v\|^2} v, u - \frac{\langle u,v\rangle}{\|v\|^2} v \right\rangle \\
&= \|u\|^2 - \frac{\overline{\langle u,v\rangle}}{\|v\|^2}\langle u,v\rangle - \frac{\langle u,v\rangle}{\|v\|^2}\langle v,u\rangle + \left|\frac{\langle u,v\rangle}{\|v\|^2}\right|^2 \|v\|^2 \\
&= \|u\|^2 - 2\frac{|\langle u,v\rangle|^2}{\|v\|^2} + \frac{|\langle u,v\rangle|^2}{\|v\|^2} \\
&= \|u\|^2 - \frac{|\langle u,v\rangle|^2}{\|v\|^2} \geq \|u\|^2
\end{aligned}$$

This implies that

$$\frac{|\langle u,v\rangle|^2}{\|v\|^2} = 0$$

Since $v \neq 0$, $\langle u,v\rangle = 0$. $\square$

Problem 7

Suppose $u, v \in V$. Prove that $\|au + bv\| = \|bu + av\|$ for all $a, b \in \mathbb{R}$ if and only if $\|u\| = \|v\|$.

Proof. Notice that

$$\begin{aligned}
\|au + bv\|^2 &= \langle au + bv, au + bv\rangle \\
&= |a|^2\|u\|^2 + a\bar{b}\langle u,v\rangle + b\overline{a}\langle v,u\rangle + |b|^2\|v\|^2 \\
&= |a|^2\|u\|^2 + |b|^2\|v\|^2 + ab(\langle u,v\rangle + \langle v,u\rangle)
\end{aligned}$$

At the same time we have

$$\|bu + av\|^2 = |b|^2\|u\|^2 + |a|^2\|v\|^2 + ab(\langle u,v\rangle + \langle v,u\rangle)$$

Then this means $\|au + bv\| = \|bu + av\|$ for all $a, b \in \mathbb{R}$ iff $|a|^2\|u\|^2 + |b|^2\|v\|^2 = |b|^2\|u\|^2 + |a|^2\|v\|^2$ for all $a, b \in \mathbb{R}$ iff $\|u\| = \|v\|$. $\square$

Problem 8

Suppose $a, b, c, x, y \in \mathbb{R}$ and $a^2 + b^2 + c^2 + x^2 + y^2 \leq 1$. Prove that $a + b + c + 4x + 9y \leq 10$.

Proof. Let

$$u = (a, b, c, x, y) \quad v = (1, 1, 1, 4, 9)$$

and consider the standard real euclidean inner product. Then we can apply the Cauchy-Schwarz:

$$|\langle u, v\rangle|^2 = \left(\sum_{i=1}^{5} u_i v_i\right)^2 \leq \left(\sum_{i=1}^{5} u_i\right)^2 \left(\sum_{i=1}^{5} v_i\right)^2 = \|u\|^2 \|v\|^2$$

Expanding this gives that

$$(a + b + c + 4x + 9y)^2 \leq (a^2 + b^2 + c^2 + x^2 + y^2)(1 + 1 + 1 + 16 + 81) \leq 100$$

Therefore, we have that

$$a + b + c + 4x + 9y \leq 10$$

□

Problem 9

Suppose $u, v \in V$ and $\|u\| = \|v\| = 1$ and $\langle u, v\rangle = 1$. Prove that $u = v$.

Proof. Suppose for contradiction that $u \neq v$, then $u - v \neq 0$. Then

$$\|u - v\|^2 = \langle u - v, u - v\rangle = \|u\|^2 + \|v\|^2 - \langle u, v\rangle - \langle v, u\rangle = 2 - 2 = 0$$

forming a contradiction. Therefore, $u = v$. □

Problem 10

Suppose $u, v \in V$ and $\|u\| \leq 1$ and $\|v\| \leq 1$. Prove that

$$\sqrt{1 - \|u\|^2}\sqrt{1 - \|v\|^2} \leq 1 - |\langle u, v\rangle|.$$

Proof. Notice that by the Cauchy-Schwarz $|\langle u, v\rangle| \leq \|u\|\|v\| = 1$. Hence, we have that

$$1 - \|u\|\|v\| \leq 1 - |\langle u, v\rangle|$$

So now it suffices to show that

$$(1 - \|u\|^2)(1 - \|v\|^2) \leq (1 - \|u\|^2\|v\|^2)$$

This is not hard to see, as r.h.s - l.h.s = $(\|u\| - \|v\|)^2 \geq 0$. □

Problem 12

Suppose a, b, c, d are positive numbers.

(a) Prove that $(a+b+c+d)(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}) \geq 16$.

(b) For which positive numbers a, b, c, d is the inequality above an equality?

Proof. (a) Let $u = (\sqrt{a}, \sqrt{b}, \sqrt{c}, \sqrt{d})$ and $v = (\frac{1}{\sqrt{a}}, \frac{1}{\sqrt{b}}, \frac{1}{\sqrt{c}}, \frac{1}{\sqrt{d}})$. Then applying the Cauchy-Schwarz yields that

$$\langle u, v\rangle^2 = 16 \leq= (a+b+c+d)(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}) = \|u\|^2\|v\|^2$$

(b) By the Cauchy-Schwarz, this is an equality iff $u = cv$, i.e. $(\sqrt{a}, \sqrt{b}, \sqrt{c}, \sqrt{d}) = c(\frac{1}{\sqrt{a}}, \frac{1}{\sqrt{b}}, \frac{1}{\sqrt{c}}, \frac{1}{\sqrt{d}})$, which holds if $a = b = c = d$. □

Problem 13

Show that the square of an average is less than or equal to the average of the squares. More precisely, show that if $a_1, \ldots, a_n \in \mathbb{R}$, then the square of the average of $a_1, \ldots, a_n$ is less than or equal to the average of $a_1^2, \ldots, a_n^2$.

Proof. We try to prove

$$\left(\frac{1}{n}\sum_{i=1}^{n} a_i\right)^2 \leq \frac{1}{n}\sum_{i=1}^{n} a_i^2$$

Take $u = (a_1, \ldots, a_n)$ and $v = (\frac{1}{n}, \ldots, \frac{1}{n})$. Then applying the Cauchy-Schwarz yields that

$$\langle u, v\rangle^2 = \left(\frac{1}{n}\sum_{i=1}^{n} a - i\right)^2 \leq \left(\sum_{i=1}^{n} a_i^2\right)\frac{1}{n} = \|u\|^2\|v\|^2$$

□

Problem 15

Suppose u, v are nonzero vectors in $\mathbb{R}^2$. Prove that

$$\langle u, v\rangle = \|u\|\|v\| \cos\theta,$$

where θ is the angle between u and v.

Proof. By law of cosines, we have that

$$\|u-v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\|\cos\theta$$

This means that

$$\begin{aligned} 2\|u\|\|v\|\cos\theta &= \|u\|^2 + \|v\|^2 - \|u-v\|^2 \\ &= \|u\|^2 + \|v\|^2 + 2\langle u, v\rangle - \|u\|^2 - \|v\|^2 \\ &= 2\langle u, v\rangle \end{aligned}$$

□

Problem 17

Prove that

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leq \left(\sum_{k=1}^{n} k a_k^2\right)\left(\sum_{k=1}^{n} \frac{b_k^2}{k}\right)$$

Proof. Consider $u = (a_1, \sqrt{2}a_2, \ldots, \sqrt{n}a_n)$ and $v = \left(b_1, \frac{b_2}{\sqrt{2}}, \ldots, \frac{b_n}{\sqrt{n}}\right)$. Applying the Cauchy-Schwarz solves the problem. □

Problem 19

Suppose $v_1, \ldots, v_n$ is a basis of V and $T \in \mathcal{L}(V)$. Prove that if λ is an eigenvalue of T, then

$$|\lambda|^2 \leq \sum_{j=1}^{n}\sum_{k=1}^{n} |\mathcal{M}(T)_{j,k}|^2,$$

where $\mathcal{M}(T)_{j,k}$ denotes the entry in row j, column k of the matrix of T wrt. the basis $v_1, \ldots, v_n$.

Proof.

$$|\lambda|^2\|v\|^2 = ||\mathcal{M}(T)v||^2 \leq ||\mathcal{M}(T)||_F^2\|v\|^2$$

for nonzero eigenvector v. Then expanding the Frobenius norm of $\mathcal{M}(T)$ gets the desired inequality. □

Problem 20

Prove the **reverse triangular inequality**: if $u, v \in V$, then $|\|u\| - \|v\|| \leq \|u - v\|$.

Proof.

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

Taking off the square yields the expected solution. □

Problem 21

Suppose $u, v \in V$ such that

$$\|u\| = 3, \quad \|u+v\| = 4, \quad \|u-v\| = 6.$$

What number does $\|v\|$ equal?

Proof. We know that

$$\begin{aligned}\|v\| &\geq \|u+v\| - \|u\| = 1 \\ \|v\|^2 &= (\|u+v\|^2 + \|u-v\|^2)/2 - \|u\|^2 = (16+36)/2 - 9 = 17\end{aligned}$$

So $\|v\| = \sqrt{17}$. □

Problem 22

Show that if $u, v \in V$, then

$$\|u+v\|\|u-v\| \leq \|u\|^2 + \|v\|^2.$$

Proof. Let $a = \|u+v\|$, $b = \|u-v\|$, then we know that

$$a^2 + b^2 = 2(\|u\|^2 + \|v\|^2)$$

We have that

$$(a-b)^2 \geq 0$$

Expanding it gives that

$$(a-b)^2 = (a^2+b^2) - 2ab \geq 0$$

equivalently,

$$\|u\|^2 + \|v\|^2 \geq \|u+v\|\|u-v\|$$

□

Problem 23

Suppose $v_1, \dots, v_m \in V$ are such that $\|v_k\| \le 1$ for each $k = 1, \dots, m$. Show that there exists $a_1, \dots, a_m \in \{1, -1\}$ such that

$$\|a_1 v_1 + \cdots + a_m v_m\| \le \sqrt{m}.$$

Proof. We consider a probabilistic approach: Let $a_1, \dots, a_m$ be the iid Rademacher variables such with $a_i = 1$ w.p. $1/2$ and $a_i = 0$ w.p. $1/2$. Then we can define a random vector

$$X = \sum_{i=1}^{m} a_i v_i$$

and we can compute the expected value

$$\mathbb{E}\left[\|X\|^2\right] = \mathbb{E}\left[\left\|\sum_{i=1}^{m} a_i v_i\right\|^2\right] = \mathbb{E}\left[\left(\sum_{i=1}^{m} a_i v_i \sum_{j=1}^{m} a_j v_j\right)\right] = \sum_{i=1}^{m}\sum_{j=1}^{m}(v_i \cdot v_j)\mathbb{E}[a_i a_j]$$

Note that here $\mathbb{E}[a_i a_j] = \delta_{ij}$ and that

$$\mathbb{E}\left[\|X\|^2\right] = \sum_{k=1}^{m} \mathbb{E}[a_k^2](v_k \cdot v_k) = \sum_{k=1}^{m} \|v_k\|^2 \le m$$

which gives that

$$\mathbb{E}\left[\|X\|\right] \le \sqrt{m}$$

and shows the existence proof. $\square$

Problem 25

Suppose $p > 0$. Prove that there is an inner product on $\mathbb{R}^2$ such that the associated norm is given by

$$\|(x, y)\| = (|x|^p + |y|^p)^{1/p}$$

for all $(x, y) \in \mathbb{R}^2$ if and only if $p = 2$.

Proof. $\Leftarrow$ Given $p = 2$, the natural euclidean dot product induces a well-defined norm, e.g. $\|(x, y)\| = (x^2 + y^2)^{1/2}$.

$\Rightarrow$ Note that the parallelogram equalities need to hold. Thus pick $u = (1, 0), v = (0, 1)$, and then

$$\|u + v\|^2 + \|u - v\|^2 = 2 \cdot 4^{1/p}$$

and

$$2(\|u\|^2 + \|v\|^2) = 4$$

They only equal each other when

$$2 \cdot 4^{1/p} = 4$$

which holds only if $p = 2$. □

Problem 26

Suppose V is a real inner product space. Prove that

$$\langle u, v \rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}$$

for all $u, v \in V$.

Proof.

$$\begin{aligned}\|u+v\|^2 - \|u-v\|^2 &= \langle u+v, u+v \rangle - \langle u-v, u-v \rangle \\ &= (\|u\|^2 + 2\langle u, v \rangle + \|v\|^2) - (\|u\|^2 - 2\langle u, v \rangle + \|v\|^2) \\ &= 4\langle u, v \rangle\end{aligned}$$

□

Problem 29

Suppose $V_1, \ldots, V_m$ are inner product spaces. Show that the equation

$$\langle (u_1, \ldots, u_m), (v_1, \ldots, v_m) \rangle = \langle u_1, v_1 \rangle + \cdots + \langle u_m, v_m \rangle$$

defines an inner product on $V_1 \times \cdots \times V_m$.

Proof. We check this by definition. Let $u, v, w \in V_1 \times \cdots \times V_m$.

positivity: $\langle v, v \rangle = \langle v_1, v_1 \rangle + \cdots + \langle v_m, v_m \rangle \geq 0$ as each of them ≥ 0.

definiteness: Suppose that $\langle v, v \rangle = \langle v_1, v_1 \rangle + \cdots + \langle v_m, v_m \rangle = 0$. Then as each of the individual element ≥ 0, the only solution is $v = 0$. Conversely, if $v = 0$, then $\langle v, v \rangle = 0$.

additivity in first slot: $\langle u+v, w \rangle = \langle u_1 + v_1, w_1 \rangle + \cdots + \langle u_m + v_m, w_m \rangle = (\langle u_1, w_1 \rangle + \cdots + \langle u_m, w_m \rangle) + (\langle v_1, w_1 \rangle + \cdots + \langle v_m, w_m \rangle) = \langle u, w \rangle + \langle v, w \rangle$

homogeneity in first slot: follows similarly as above.

conjugate symmetry: follows similarly as above. □

Problem 31

Suppose $u, v, w \in V$. Prove that

$$\left\| w - \frac{1}{2}(u+v) \right\|^2 = \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4}.$$

Proof. Let $w - u = a$ and $w - v = b$, then we have

$$\begin{aligned}
\text{l.h.s} &= \left\|a/2 + b/2\right\|^2 \\
&= 2(\left\|a/2\right\|^2 + \left\|b/2\right\|^2) - \left\|a/2 - b/2\right\|^2 \\
&= \frac{\|a\|^2 + \|b\|^2}{2} - \frac{a-b}{4} = \text{r.h.s}
\end{aligned}$$

Substituting a and b gets the desired result. □

Problem 32

SUppose that E is a subset of V with the property that $u, v \in E$ implies $\frac{1}{2}(u+v) \in E$. Let $w \in V$. Show that there is at most one point in E that is cloest to w. In other words, show that there is at most one $u \in E$ such that

$$\|w - u\| \leq \|w - x\|$$

for all $x \in E$.

Proof. Suppose for contradiction that there is another $v \in E, v \neq u$ such that

$$\|w - v\| \leq \|w - x\|$$

for all $x \in E$. Then we have that

$$\left\|w - \frac{1}{2}(u+v)\right\| = \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4}$$

by problem 31. Notice that

$$\frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4} \leq \|w - x\| - \frac{\|u-v\|^2}{4} \leq \|w - x\|$$

for all $x \in E$, reaching a contradiction $(u = v)$. □

6B: Orthonormal Bases

Definition 15 (orthonormal)**.** *A list of vectors is called **orthonormal** if each vector in the list has norm 1 and is orthogonal to all the other vectors in the list.*

Corollary 16. *Suppose $e_1, \dots, e_m$ is an orthonormal list of vectors in V. Then*

$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

for all $a_1, \dots, a_m \in \mathbb{F}$.

Corollary 17. *Every orthonormal list of vectors is linearly independent.*

Theorem 18 (Bessel's inequality)**.** *Suppose $e_1, \dots, e_m$ is an orthonormal list of vectors in V. If $v \in V$ then*

$$|\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_m \rangle|^2 \leq \|v\|^2$$

Definition 19 (orthonormal basis)**.** *An **orthonormal basis** of V is an orthogonal list of vectors in V that is also a basis of V.*

Corollary 20. *Suppose V is finite-dimensional. Then every orthonormal list of vectors in V of length $\dim V$ is an orthonormal basis of V.*

Remark 21. *Usually we write $v = \sum_{i=1}^{n} a_i v_i$, but with orthonormal basis we can just take $a_k = \langle v, e_k \rangle$.*

Lemma 22 (writing a vector as a linear combination of an orthonormal basis)**.** *Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $u, v \in V$. Then*

(a) $v = \langle v, e_1 \rangle e_1 + \langle v, e_n \rangle e_n$*,*

(b) $\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2$*,*

(c) $\langle u, v \rangle = \langle u, e_1 \rangle \overline{\langle v, e_1 \rangle} + \cdots + \langle u, e_n \rangle \overline{\langle v, e_n \rangle}$*.*

Theorem 23 (Gram-Schmidt procedure)**.** *Suppose $v_1, \dots, v_m$ is a linearly independent list of vectors in V. Let $f_1 = v_1$. For $k = 2, \dots, m$, define f_k inductively by*

$$f_k = v_k - \frac{\langle v_k, f_1 \rangle}{\|f_1\|^2} f_1 - \cdots - \frac{\langle v_k, f_{k-1} \rangle}{\|f_{k-1}\|^2} f_{k-1}.$$

For each $k = 1, \dots, m$, let $e_k = \frac{f_k}{\|f_k\|}$. Then $e_1, \dots, e_m$ is an orthonormal list of vectors in V such that

$$span(v_1, \dots, v_k) = span(e_1, \dots, e_k)$$

for each $k = 1, \dots, m$.

Corollary 24. *Every finite-dimensional inner product space has an orthornormal basis.*

Corollary 25. *Suppose V is finite-dimensional. Then every orthonormal list of vectors in V can be extended to an orthonormal basis of V.*

Lemma 26 (upper-triangular matrix with respect to some orthonormal basis)**.** *Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some orthonormal basis of V if and only if the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$.*

Theorem 27 (Schur's theorem)**.** *Every operator on a finite-dimensional complex inner product space has an upper-triangular matrix with respect to some orthonormal basis.*

Theorem 28 (Riesz representation theorem)**.** *Suppose V is finite-dimensional and φ is a linear functional on V. Then there is a unique vector $v \in V$ such that*

$$\varphi(u) = \langle u, v \rangle$$

for every $u \in V$.

Problem 1

Suppose $e_1, \dots, e_m$ is a list of vectors in V such that

$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

for all $a_1, \dots, a_m \in \mathbb{F}$. Show that $e_1, \dots, e_m$ is an orthonormal list.

Proof. First, note that if we choose $a_j = 1$ and $a_k = 0, k \neq j$, we have $\langle e_j, e_j \rangle = \|e_j\|^2 = 1$.

Further, for all $a_1, \dots, a_m \in \mathbb{F}$

$$\begin{aligned}\sum_{i=1}^{m} |a_i|^2 &= \left\| \sum_{i=1}^{m} a_i e_i \right\|^2 = \left\langle \sum_{i=1}^{m} a_i e_i, \sum_{i=1}^{m} a_i e_i \right\rangle = \sum_{i=1}^{m} \sum_{j=1}^{m} a_i \overline{a_j} \langle e_i, e_j \rangle \\ &= \sum_{i=1}^{m} |a_i|^2 \langle e_i, e_i \rangle + \sum_{i \neq j}^{m} a_i \overline{a_j} \langle e_i, e_j \rangle = \sum_{i=1}^{m} |a_i|^2 + \sum_{i \neq j}^{m} a_i \overline{a_j} \langle e_i, e_j \rangle.\end{aligned}$$

That is, for all $a_1, \dots, a_m \in \mathbb{F}$

$$\sum_{i \neq j}^{m} a_i \overline{a_j} \langle e_i, e_j \rangle = 0.$$

Choosing $a_i = 1, a_j = 1$ and $a_k = 0$ for $i \neq j, k$, we get $\langle e_j, e_k \rangle = 0$. Thus, we have $\langle e_i, e_j \rangle = \delta_{ij}$ which means that $e_1, \dots, e_m$ is orthonormal. □

Problem 3

Suppose $e_1, \dots, e_m$ is an orthonormal list in V and $v \in V$. Prove that

$$\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_m \rangle|^2 \iff v \in \operatorname{span}(e_1, \dots, e_m)$$

Proof. $\Rightarrow$ We can decompose v into two parts, one is $v_{proj} = \sum_{i=1}^{m} \langle v, e_i \rangle e_i$, which is the orthogonal projection of v onto the subspace spanned by $e_1, \dots, e_m$. We claim that $v - v_{proj}$ is orthogonal to v_{proj}. This can be seen as

$$\begin{aligned}\langle v_{proj}, v - v_{proj} \rangle &= \left\langle \sum_{i=1}^{m} \langle v, e_i \rangle e_i, v - \sum_{j=1}^{m} \langle v, e_j \rangle e_j \right\rangle \\ &= \sum_{i=1}^{m} |\langle v, e_i \rangle|^2 - \sum_{i=1}^{m} |\langle v, e_i \rangle|^2 = 0\end{aligned}$$

Then by Pythagorean theorem we have

$$\|v\|^2 = \|v_{proj}\|^2 + \|v - v_{proj}\|^2$$

where $\|v\|^2 = \left\|v_{proj}^2\right\|$ and thus $v = v_{proj}$. Equivalently, $v \in \text{span}(e_1, \dots, e_m)$.

$\Leftarrow$ This means that $v = \sum_{i=1}^m a_i e_i$. However, we know that $a_i = \langle v, e_i \rangle$, so $\|v\|^2 = \sum_{j=1}^m |\langle v, e_j \rangle|^2$ by repeatedly applying the Pythagorean theorem. $\square$

Problem 4

Suppose n is a positive integer. Prove that

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots \frac{\sin nx}{\sqrt{\pi}}$$

is an orthonormal list of vectors in $C[-\pi, \pi]$, the vector space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} fg.$$

Proof. First, we show each of the element has norm 1.

$$\left\|\frac{1}{\sqrt{2\pi}}\right\| = \sqrt{\int_{-\pi}^{\pi} \frac{1}{2\pi} dx} = 1$$

$$\left\|\frac{\cos nx}{\sqrt{\pi}}\right\| = \sqrt{\int_{-\pi}^{\pi} \frac{\cos^2 nx}{\pi} dx} = \sqrt{\frac{1}{\pi}[\frac{x}{2} + \frac{\sin(2nx)}{4n}]_{-\pi}^{\pi}} = 1$$

$$\left\|\frac{\sin nx}{\sqrt{\pi}}\right\| = \sqrt{\int_{-\pi}^{\pi} \frac{\sin^2 nx}{\pi} dx} = \sqrt{\frac{1}{\pi}[\frac{x}{2} - \frac{\cos(2nx)}{4n}]_{-\pi}^{\pi}} = 1$$

Next, we show that each element is orthogonal to each other, there are many different cases, we begin examine here:

$$\langle \frac{1}{\sqrt{2\pi}}, \frac{\cos nx}{\sqrt{\pi}} \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \cos nx dx = \frac{1}{\sqrt{2\pi}} [\frac{\sin nx}{n}]_{-\pi}^{\pi} = 0$$

$$\langle \frac{1}{\sqrt{2\pi}}, \frac{\sin nx}{\sqrt{\pi}} \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \sin nx dx = \frac{1}{\sqrt{2\pi}} [-\frac{\cos nx}{n}]_{-\pi}^{\pi} = 0$$

Similarly, one can derive between every different pairs of element, their inner product is 0 for different index. The derivation is omitted. $\square$

Problem 6

Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V.

(a) Prove that if $v_1, \ldots, v_n$ are vectors in V such that

$$\|e_k - v_k\| < \frac{1}{\sqrt{n}}$$

for each k, then $v_1, \ldots, v_n$ is a basis of V.

(b) Show that there exist $v_1, \ldots, v_n \in V$ such that

$$\|e_k - v_k\| \leq \frac{1}{\sqrt{n}}$$

for each k, but $v_1, \ldots, v_n$ is not linearly independent.

Proof. (a) Suppose for contradiction that $v_1, \ldots, v_n$ is not a basis and thus linearly dependent. Then there exists scalars $a_1, \ldots, a_n \in \mathbb{F}$ not all zero such that $\sum_{i=1}^n a_i v_i = 0$. Then we have that

$$\sum_{i=1}^{n} a_i(v_i - e_i) + \sum_{i=1}^{n} a_i e_i = 0$$

which means that

$$\left\| \sum_{i=1}^{n} a_i(v_i - e_i) \right\| = \left\| \sum_{i=1}^{n} a_i e_i \right\|$$

Note that

$$\left\| \sum_{i=1}^{n} a_i(v_i - e_i) \right\| \leq \sum_{i=1}^{n} \|a_i(v_i - e_i)\| = \sum_{i=1}^{n} |a_i| \|v_i - e_i\| < \sum_{i=1}^{n} \frac{|a_i|}{\sqrt{n}} \leq \left(\sum_{i=1}^{n} |a_i|^2 \right)^{1/2}$$

where the last inequality is shown by the Cauchy-Schwarz. This reaches a contradiction, as

$$\left\| \sum_{i=1}^{n} a_i e_i \right\| = \left(\sum_{i=1}^{n} |a_i|^2 \right)^{1/2} < \left(\sum_{i=1}^{n} |a_i|^2 \right)^{1/2}$$

(b) Suppose $v_1 = e_1 + \frac{1}{\sqrt{n}} e_2$ and $v_j = e_j$ for $2 \leq j \leq n$. Hence we have

$$\|e_1 - v_1\| = \left\| \frac{1}{\sqrt{n}} e_2 \right\| = \frac{1}{\sqrt{n}}$$

where other conditions hold trivially. However, we can clearly tell that $v_1, \ldots, v_n$ is not linearly independent. $\square$

Problem 9

Suppose $e_1, \ldots, e_m$ is the result of applying the Gram-Schmidt procedure to a linearly independent list $v_1, \ldots, v_m$ in V. Prove that $\langle v_k, e_k \rangle > 0$ for each $k = 1, \ldots, m$.

Proof. In the Gram-Schmidt process, we decompose v_k into v_{proj} and f_k where v_{proj} is the Orthogonal projection of v_k onto the $\operatorname{span}(v_1, \ldots, v_{k-1}) = \operatorname{span}(e_1, \ldots, e_{k-1})$. To show $\langle v_k, e_k \rangle > 0$, it's equivalent to show $\langle v_k, f_k \rangle > 0$, which naturally holds as

$$\langle v_k, f_k \rangle = \langle f_k + v_{proj}, f_k \rangle = \langle f_k, f_k \rangle > 0$$

$\square$

Problem 11

Find a polynomial $q \in \mathcal{P}_2(\mathbb{R})$ such that $p(\frac{1}{2}) = \int_0^1 pq$ for every $p \in \mathcal{P}_2(\mathbb{R})$.

Proof. Define $\varphi \in \mathcal{L}(\mathcal{P}_2(\mathbb{R}))$ to be $\varphi(p) = p(\frac{1}{2})$ and consider the inner product $\langle p, q \rangle = \int_0^1 pq$. Following the Riesz representation theorem, we can derive that

$$q = \overline{\varphi(e_1)}e_1 + \overline{\varphi(e_2)}e_2 + \overline{\varphi(e_3)}e_3$$

where we can consider the orthonormal basis $\sqrt{\frac{1}{2}}, \sqrt{\frac{3}{2}}x, \sqrt{\frac{45}{8}}(x^2 - \frac{1}{3})$. Then

$$\begin{aligned}
q(x) &= \sqrt{\frac{1}{2}}\sqrt{\frac{1}{2}} + \sqrt{\frac{3}{2}}\frac{1}{2}\sqrt{\frac{3}{2}}x + \sqrt{\frac{45}{8}}\left(\frac{1}{4} - \frac{1}{3}\right)\sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right) \\
&= \frac{1}{2} + \frac{3}{4}x + \frac{5}{32} - \frac{15}{32}x^2 \\
&= -\frac{15}{32}x^2 + \frac{3}{4}x + \frac{21}{32}
\end{aligned}$$

$\square$

Problem 13

Show that a list $v_1, \ldots, v_m$ of vectors in V is linearly dependent if and only if the Gram-Schmidt formula produces $f_k = 0$ for some $k \in \{1, \ldots, m\}$.

Proof. At each step k, the formula aims at decomposes $v_k = v_{proj} + f_k$, where v_{proj} is the orthogonal projection of v_k onto $\operatorname{span}(e_1, \ldots, e_{k-1})$. $f_k = 0$ equivalently means that $v_k = v_{proj}$, which means that $v \in \operatorname{span}(e_1, \ldots, e_{k-1}) = \operatorname{span}(v_1, \ldots, v_{k-1})$ and therefore renders the list to be linearly dependent. $\square$

Problem 14

Suppose V is a real inner product space and $v_1, \ldots, v_m$ is a linearly independent list of vectors in V. Prove that there exist exactly 2^m orthonormal lists $e_1, \ldots, e_m$ of vectors in V such that

$$\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(e_1, \ldots, e_k)$$

for all $k \in \{1, \ldots, m\}$.

Proof. We prove this statement through induction on m. For base case, consider $\operatorname{span}(v_1)$ for nonzero $v_1 \in V$, there are only two nonzero vectors in $\operatorname{span}(v_1)$: $\pm\frac{v_1}{\|v_1\|}$. So there are exactly $2^1 = 1$ orthonormal list of vectors.

For induction, assume that for $v_1, \ldots, v_{k-1}$ linearly independent list of vectors in V, there exist exactly 2^{k-1} orthonormal lists $e_1, \ldots, e_{k-1}$ of vectors in V such that

$$\operatorname{span}(v_1, \ldots, v_{k-1}) = \operatorname{span}(e_1, \ldots, e_{k-1})$$

For k, by the Gram-Schmidt, we have the e_k such that

$$\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(e_1, \ldots, e_k)$$

Suppose such choice of e_k is not unique and there's other e'_k also satisfies

$$\operatorname{span}(e_1, \ldots, e'_k) = \operatorname{span}(e_1, \ldots, e_k)$$

which indicates that $e'_k = \sum_{i=1}^{k} \langle e'_k, e_i \rangle e_i = \langle e'_k, e_k \rangle e_k$ and that

$$1 = \|e'_k\| = |\langle e'_k, e_k \rangle|$$

so $e'_k = \pm e_k$ and this gives $2 * 2^{m-1} = 2^m$ choices of orthonormal lists of vectors. $\square$

Problem 15

Suppose $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are inner products on V such that $\langle u, v \rangle_1 = 0$ if and only if $\langle u, v \rangle_2 = 0$. Prove that there is a positive number c such that $\langle u, v \rangle_1 = c \langle u, v \rangle_2$ for every $u, v \in V$.

Proof. It suffices to prove that $c = \frac{\langle u, v \rangle_1}{\langle u, v \rangle_2}$ for every $u, v \in V$ is a constant number.

First, pick nonzero $u \in V$. Then we know that $\langle u, u \rangle_1 > 0, \langle u, u \rangle_2 > 0$. Pick $v \in V$ s.t. $\langle u, v \rangle_1 \neq 0, \langle u, v \rangle_2 \neq 0$ (i.e. they are not orthogonal). So we have that

$$\left\langle u - \frac{\langle u, v \rangle_2}{\langle v, v \rangle_2} v, v \right\rangle_1 = 0 = \left\langle u - \frac{\langle u, v \rangle_2}{\langle v, v \rangle_2} v, v \right\rangle_2$$

by the orthogonal decomposition of u. This gives that

$$\frac{\langle u, v\rangle_1}{\langle u, v\rangle_2} = \frac{\langle v, v\rangle_1}{\langle v, v\rangle_2}$$

Similarly, we have

$$\left\langle u, v - \frac{\langle u, v\rangle_2}{\langle v, v\rangle_2} u \right\rangle_1 = 0 = \left\langle u, v - \frac{\langle u, v\rangle_2}{\langle v, v\rangle_2} u \right\rangle_2$$

and that

$$\frac{\langle u, v\rangle_1}{\langle u, v\rangle_2} = \frac{\langle u, u\rangle_1}{\langle u, u\rangle_2} = \frac{\langle v, v\rangle_1}{\langle v, v\rangle_2} = c$$

which yields the desired solution. □

Problem 16

Suppose V is finite-dimensional. Suppose $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are inner products on V with corresponding norms $\|\cdot\|_1$ and $\|\cdot\|_2$. Prove that there exists a positive number c such that $\|v\|_1 \leq c\|v\|_2$ for every $v \in V$.

Proof. Let $e_1, \ldots, e_n$ be an orthonormal basis of V wrt. $\langle \cdot, \cdot \rangle_1$ and $f_1, \ldots, f_n$ be an orthonormal basis of V wrt. $\langle \cdot, \cdot \rangle_2$. Pick nonzero $v \in V$. Then there exists $\varphi \in V'$ such that

$$\|v\|_1^2 = \sum_{i=1}^{n} |\langle v, e_i \rangle_1|^2 \leq |\varphi(v)|^2$$

We can proceed with that

$$\begin{aligned}
\|v\|_1^2 &\leq |\varphi(v)|^2 \\
&= |\langle v, \overline{\varphi(f_1)} f_1 + \cdots + \overline{\varphi(f_n)} f_n \rangle_2|^2 \\
&\leq \left\| \sum_{i=1}^{n} \overline{\varphi(f_i)} f_i \right\|_2^2 \|v\|_2^2
\end{aligned}$$

which completes the proof. □

Problem 17

Suppose $\mathbb{F} = \mathbb{C}$ and V is finite-dimensional. Prove that if T is an operator on V such that 1 is the only eigenvalue of T and $\|Tv\| \leq \|v\|$ for all $v \in V$, then T is the identity operator.

Proof. By Schur's theorem, there exists an orthonormal basis $e_1, \ldots, e_n$ such that the matrix of T is upper-triangular. Then 1 is the only component on the diagonal entries. Hence,

$$\|Te_k\| \leq \|e_k\| = 1$$

Note that $Te_k = \sum_{i=1}^{k-1} a_i e_i + e_k$ since we know the upper-triangular matrix has diagonal term to be 1. This gives that

$$\left\| \sum_{i=1}^{k-1} a_i e_i + e_k \right\| = \|e_k\| + \sum_{i=1}^{k-1} |a_i| \|e_i\| \leq \|e_k\|$$

so for each e_k, the off-diagonal entries a_i are all 0 and thus the matrix of T is the identity matrix, and T is the identity operator. □

Problem 18

Suppose $u_1, \ldots, u_m$ is a linearly independent list in V. Show that there exists $v \in V$ such that $\langle u_k, v \rangle = 1$ for all $k \in \{1, \ldots, m\}$.

Proof. Define $\varphi \in V'$ s.t. $\varphi(u_k) = 1$ for all k. By Riesz representation theorem, there is a unique $v \in V$ s.t.

$$\varphi(u_k) = \langle u_k, v \rangle = 1$$

□

6C: Orthogonal Complements and Minimization Problems

Definition 29 (orthogonal complement, $U^\perp$)**.** *If U is a subset of V, then the* ***orthogonal complement*** *of U, denoted by $U^\perp$, is the set of all vectors in V that are orthogonal to every vector in U:*

$$U^\perp = \{v \in V : \langle u, v\rangle = 0 \text{ for every } u \in U\}.$$

Corollary 30. *Properties of orthogonal complement:*
(a) If U is a subset of V, then $U^\perp$ is a subspace of V.
(b) $\{0\}^\perp = V$.
(c) $V^\perp = \{0\}$.
(d) If U is a subset of V, then $U \cap U^\perp \subseteq \{0\}$.
(e) If G and H are subsets of V and $G \subseteq H$, then $H^\perp \subseteq G^\perp$.

Corollary 31. *Suppose U is a finite-dimensional subspace of V. Then*

$$V = U \oplus U^\perp$$

and thus $\dim U^\perp = \dim V - \dim U$. In addition,

$$U = (U^\perp)^\perp$$

Corollary 32. *Suppose U is a finite-dimensional subspace of V. Then*

$$U^\perp = \{0\} \Longleftrightarrow U = V.$$

Definition 33 (orthogonal projection, P_U)**.** *Suppose U is a finite-dimensional subspace of V. The* ***orthogonal projection*** *of V onto U is the operator $P_U \in \mathcal{L}(V)$ defined as follows: for each $v \in V$, write $v = u + w$, where $u \in U$ and $w \in U^\perp$. Then let $P_U v = u$.*

Remark 34. *Suppose $u \in V$ with $u \neq 0$ and $U = span(u)$. If $v \in V$, then*

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

Then this implies that

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

Corollary 35 (properties of orthogonal projection P_U)**.** *Suppose U is a finite-dimensional subspace of V. Then*

(a) $P_U \in \mathcal{L}(V)$;

(b) $P_U u = u$ *for every* $u \in U$;

(c) $P_U w = 0$ *for every* $w \in U^\perp$;

(d) *range* $P_U = U$;

(e) *null* $P_U = U^\perp$;

(f) $v - P_U v \in U^\perp$ *for every* $v \in V$;

(g) $P_U^2 = P_U$;

(h) $\|P_U v\| \leq \|v\|$ *for every* $v \in V$;

(i) *if* $e_1, \ldots, e_m$ *is an orthonormal basis of* U *and* $v \in V$, *then*

$$P_U v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m$$

Theorem 36 (Riesz representation theorem, revisited)**.** *Suppose* V *is finite-dimensional. For each* $v \in V$, *define* $\varphi_v \in V'$ *by*

$$\varphi_v(u) = \langle u, v\rangle$$

for each $u \in V$. *Then* $v \mapsto \varphi_v$ *is a one-to-one function from* V *to* V'.

Theorem 37 (minimizing distance to a subspace)**.** *Suppose* U *is a finite-dimensional subspace of* V, $v \in V$, *and* $u \in U$. *Then*

$$\|v - P_U v\| \leq \|v - u\|.$$

Furthermore, the inequality above is an equality if and only if $u = P_U v$.

Lemma 38 (restriction of a linear map to obtain a one-to-one and onto map)**.** *Suppose* V *is finite-dimensional and* $T \in \mathcal{L}(V, W)$. *Then* $T|_{(null\ T)^\perp}$ *is an injective map of* $(null\ T)^\perp$ *onto range* T.

Definition 39 (pseudoinverse, $T^\dagger$)**.** *Suppose that* V *is finite-dimensional and* $T \in \mathcal{L}(V, W)$. *The* ***pseudoinverse*** $T^\dagger \in \mathcal{L}(W, V)$ *of* T *is the linear map from* W *to* V *defined by*

$$T^\dagger w = (T|_{(null\ T)^\perp})^{-1} P_{range\ T} w$$

for each $w \in W$.

Corollary 40 (algebraic properties of the pseudoinverse)**.** *Suppose* V *is finite-dimensional and* $T \in \mathcal{L}(V, W)$.

(a) *If* T *is invertible, then* $T^\dagger = T^{-1}$.

(b) $TT^\dagger = P_{range\ T} =$ *the orthogonal projection of* W *onto range* T.

(c) $T^\dagger T = P_{(null\ T)^\perp} =$ *the orthogonal projection of* V *onto* $(null\ T)^\perp$.

Theorem 41 (pseudoinverse provides best approximate solution or best solution)**.** *Suppose* V *is finite-dimensional,* $T \in \mathcal{L}(V, W)$, *and* $w \in W$.

(a) If $v \in V$*, then*

$$\left\| T(T^{\dagger}w) - w \right\| \leq \|Tv - w\|$$

with equality if and only if $v \in T^{\dagger}w + \text{null}\, T$.

(b) If $v \in T^{\dagger}w + \text{null}\, T$*, then*

$$\left\| T^{\dagger}w \right\| \leq \|v\|,$$

with equality if and only if $v = T^{\dagger}w$.

Problem 1

Suppose $v_1, \ldots, v_m \in V$. Prove that

$$\{v_1, \ldots, v_m\}^{\perp} = (\operatorname{span}(v_1, \ldots, v_m))^{\perp}.$$

Proof. Denote $A = \{v_1, \ldots, v_m\}^{\perp}$ and $B = \operatorname{span}(v_1, \ldots, v_m)^{\perp}$

$\Rightarrow$ Let $v \in A$, then $\langle v, v_i \rangle = 0$ for all i. So we have

$$\left\langle v, \sum_{i=1}^{m} a_i v_i \right\rangle = \sum_{i=1}^{m} \overline{a_i} \langle v, v_i \rangle = 0$$

which means that $v \in B$.

$\Leftarrow$ Conversely, let $v \in B$, then naturally by definition $v \in A$. $\square$

Problem 4

Suppose $e_1, \ldots, e_n$ is a list of vectors in V with $\|e_k\| = 1$ for each $k = 1, \ldots, n$ and

$$\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2$$

for all $v \in V$. Prove that $e_1, \ldots, e_n$ is an orthonormal basis of V.

Proof. It now suffices to prove that $\langle e_i, e_j \rangle = \delta_{ij}$. To see this, take $v = e_i$, then we have that

$$\|v\|^2 = \|e_i\|^2 = 1 = \sum_{j \neq i}^{n} |\langle e_i, e_j \rangle|^2 + 1$$

This gives that $\langle e_i, e_j \rangle = 0$ for all $i \neq j$, completing the proof. $\square$

Problem 5

Suppose that V is finite-dimensional and U is a subspace of V. Show that $P_{U^{\perp}} = I - P_U$, where I is the identity operator on V.

Proof. Take $v \in V$, then we know $v = u + w$ for $u \in U, w \in U^{\perp}$. We have that

$$P_U v = u \quad P_{U^{\perp}} v = w = v - u = (I - P_U) v$$

$\square$

Problem 6

Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Show that

$$T = T P_{(\operatorname{null} T)^{\perp}} = P_{\operatorname{range} T} T.$$

Proof. Take arbitrary $v \in V$, then $v = u + w$ for $u \in \text{null } T$ and $w \in (\text{null } T)^\perp$. We have

$$Tv = T(u + w) = Tw = TP_{(\text{null } T)^\perp} v$$

Furthermore, since $Tv \in \text{range } T$, $P_{\text{range } T}$ acts as an identity operator for Tv, thus we have the second equality. $\square$

Problem 7

Suppose that X and Y are finite-dimensional subspaces of V. Prove that $P_X P_Y = 0$ if and only if $\langle x, y \rangle = 0$ for all $x \in X$ and all $y \in Y$.

Proof. $\Rightarrow$ take arbitrary $y \in Y$, then $P_X(y) = 0$, which means that $y = 0 + (y - 0)$ where $0 \in X$ and thus all $y \in Y$ are orthogonal to $x \in X$, completing this direction.

$\Leftarrow$ Take $v \in V$, then $v = y + y'$ for $y \in Y, y' \in Y^\perp$ and we further have $y = 0 + y$ for $0 \in X$ and $y \in X^\perp$. We now have that

$$P_X P_Y(v) = P_X(y) = 0$$

$\square$

Problem 9

Suppose V is finite-dimensional. Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and every vector in null P is orthogonal to every vector in range P. Prove that there exists a subspace U of V such that $P = P_U$.

Proof. We can simply take $U = \text{range } P$. Note that $V = \text{null } P \oplus \text{range } P$ as

$$v = Pv + (v - Pv)$$

where $P(v - Pv) = 0$ so $v - Pv \in \text{null } P$.

Then take $v = v_1 + v_2$ where $v_1 \in \text{null } P, v_2 \in \text{range } P$, then we have

$$Pv = P(v_1 + v_2) = Pv_2 = P_U v$$

$\square$

Problem 11

Suppose $T \in \mathcal{L}(U)$ and U is a finite-dimensional subspace of V. Prove that

$$U \text{ is invariant under } T \iff P_U T P_U = T P_U$$

Proof. U is invariant under $T \iff Tu \in U$ for all $u \in U \iff$ for $v = u + u^\perp \in V$, $TP_U(v) = Tu = P_U(Tu) = P_U T P_U(v)$ $\square$

Problem 13

Suppose $\mathbb{F} = \mathbb{R}$ and V is finite-dimensional. For each $v \in V$, let φ_v denote the linear functional on V defined by

$$\varphi_v(u) = \langle u, v \rangle$$

for all $u \in V$.

(a) Show that $v \mapsto \varphi_v$ is an injective linear map from V to V'.

(b) Use (a) and a dimension-counting argument to show that $v \mapsto \varphi_v$ is an isomorphism from V onto V'.

Proof. (a) denote this map $v \mapsto \varphi_v$ to be T. Then take $v \in \operatorname{null} T$, we have $T(v) = \varphi_v = 0$. By definition, since this holds for all $u \in V$, $v = 0$ and thus T is injective. To show it's also linear, $T(\lambda v_1 + v_2)(u) = \varphi_{\lambda v_1 + v_2}(u) = \langle u, \lambda v_1 + v_2 \rangle = \lambda \langle u, v_1 \rangle + \langle u, v_2 \rangle = \lambda \varphi_{v_1} + \varphi_{v_2} = \lambda T(v_1) + T(v_2)$.

(b) We know that $\dim V = \dim V'$ and combining this with (a) yields the solution. $\square$

Problem 15

In $\mathbb{R}^4$, let

$$U = \operatorname{span}((1,1,0,0),(1,1,1,2))$$

Find $u \in U$ such that $\left\| u - (1,2,3,4) \right\|$ is as small as possible.

Proof. We first find the orthonormal basis of U and apply the formula, i.e. $P_U(v) = \sum_{i=1}^{n} \langle v, e_i \rangle e_i$. Using the Gram-Schmidt, we can find that

$$\left\{ \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0 \right), \left(0, 0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right) \right\}$$

Then we can get that

$$\begin{aligned} u &= \left\langle (1,2,3,4), \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0 \right) \right\rangle \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0 \right) + \\ &\quad \left\langle (1,2,3,4), \left(0, 0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right) \right\rangle \left(0, 0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right) \\ &= \left(\frac{3}{2}, \frac{3}{2}, \frac{11}{5}, \frac{22}{5} \right) \end{aligned}$$

$\square$

Problem 19

Suppose V is finite-dimensional and $P \in \mathcal{L}(V)$ is an orthogonal projection of V onto some subspace of V. Prove that $P^\dagger = P$.

Proof. Suppose the subspace is U. Take $u \in U$, then we know that $u \in \text{range } P$, and thus

$$P^\dagger u = (P|_{(\text{null } P)^\perp})^{-1} P_{\text{range } P} u = (P|_{(\text{null } P)^\perp})^{-1} u = u = Pu$$

Take $u \in U^\perp$, then we have $Pu = 0$ and also $P^\dagger u = 0$ by definition. Thus these two operators equal each other. $\square$

Problem 20

Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Show that

$$\text{null } T^\dagger = (\text{range } T)^\perp \text{ and range } T^\dagger = (\text{null } T)^\perp.$$

Proof. We know that $T^\dagger = (T|_{(\text{null } T)^\perp})^{-1} P_{\text{range } T}$ and the first part $T|_{(\text{null } T)^\perp}$ we've shown it's bijective with the restriction in book's lemma. So for $v \in \text{null } T^\dagger$, $P_{\text{range } T} v = 0$ and thus $v \in (\text{range } T)^\perp$. Conversely, it holds by definition.

For the other equality, take $v \in \text{range } T^\dagger$, then there exists $u \in \text{range } T$ s.t. $T|_{(\text{null } T)^\perp} v = u$, so $v \in (\text{null } T)^\perp$. Conversely, take $v \in (\text{null } T)^\perp$, then there exists $u \in \text{range } T$ s.t. $Tv = u$, and we have $T^\dagger u = v$ so $v \in \text{range } T^\dagger$. $\square$

Problem 22

Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that

$$TT^\dagger T = T \text{ and } T^\dagger T T^\dagger = T^\dagger.$$

Proof. For the first equality, take $v \in \text{null } T$, then $TT^\dagger Tv = Tv = 0$. Take $v \in (\text{null } T)^\perp$, then $TT^\dagger(Tv) = T(v)$ by definition.

For the second equality, take $w \in (\text{range } T)^\perp$, then $T^\dagger T T^\dagger w = 0 = T^\dagger w$. Take nonzero $w \in \text{range } T$, then there exists $v \in (\text{null } T)^\perp$ such that $Tv = w$, hence $T^\dagger T T^\dagger w = T^\dagger T v = v = T^\dagger w$. $\square$

Problem 23

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that

$$(T^\dagger)^\dagger = T.$$

Proof. Denote $S = T^\dagger$, we have that

$$(T^\dagger)^\dagger = S^\dagger = (S|_{(\text{null } S)^\perp})^{-1} P_{\text{range } S} = (S|_{\text{range } T})^{-1} P_{(\text{null } T)^\perp}$$

$\square$

where we use the conclusion from problem 20. Note that if $v \in \text{null } T$, then naturally $(T^\dagger)^\dagger v = Tv = 0$. If $v \in (\text{null } T)^\perp$, then first note that

$$(S|_{\text{range } T})^{-1} P_{(\text{null } T)^\perp} = (S|_{\text{range } T})^{-1} v$$

Expanding the definition gives that

$$(S|_{\text{range } T})^{-1} v = (((T|_{(\text{null } T)^\perp})^{-1} P_{\text{range } T})|_{\text{range } T})^{-1} v = T|_{(\text{null } T)^\perp} v = Tv$$

Therefore, we complete the proof.