MATH 2135 Linear Algebra

Chapter 6 Inner Product Spaces

Alyssa Motas

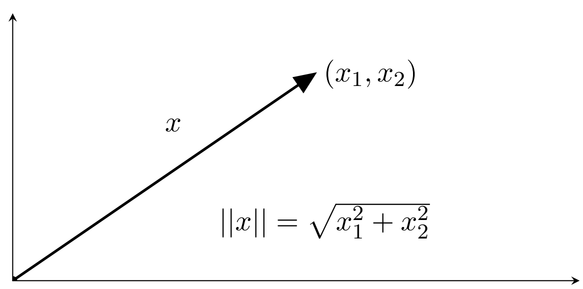
April 7, 2021

Contents

1 6.A Inner Products and Norms

To motivate the concept of inner product, think of vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ as arrows with initial point at the origin. The length of a vector x in $\mathbb{R}^2$ or $\mathbb{R}^3$ is called the *norm* of x, denoted $||x||$. Thus for $x = (x_1, x_2) \in \mathbb{R}^2$, we have $||x|| = \sqrt{x_1^2 + x_2^2}$. The generalization to $\mathbb{R}^n$ is: we defined the norm of $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ by

$$||x|| = \sqrt{x_1^2 + \cdots + x_n^2}.$$



The norm is not linear on $\mathbb{R}^n$.

1.1 Definition of dot product

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

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

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. Note that the dot product of two vectors in $\mathbb{R}^n$ is a number, not a vector.

An inner product is a generalization of the dot product. Recall that if $\lambda = a + bi$, where $a, b \in \mathbb{R}$, then

- the absolute value of λ, denoted $|\lambda|$, is defined by $|\lambda| = \sqrt{a^2 + b^2}$;
- the complex conjugate of λ, denoted $\overline{\lambda}$, is defined by $\overline{\lambda} = a - bi$;
- $|\lambda|^2 = \lambda\overline{\lambda}$.

For $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$, we define the norm of z by

$$||z|| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}.$$

The absolute values are needed because we want $||z||$ to be nonegative number. Note that

$$||z||^2 = z_1\overline{z_1} + \cdots + z_n\overline{z_n}.$$

We want to think of $||z||^2$ as the inner product of z with itself. The equation above suggests that the inner product of $w = (w_1, \ldots, w_n) \in \mathbb{C}^n$ with z should equal

$$w_1\overline{z_1} + \cdots + w_n\overline{z_n}.$$

If the roles of w and z were interchanged, the expression above would be its complex conjugate. We should expect that the inner product of w with z equals the complex conjugate of the inner product of z with w.

1.2 Definition of inner product

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

positivity

$$\langle v, v\rangle \geq 0 \text{ for all } v \in V;$$

definiteness

$$\langle v, v\rangle = 0 \text{ if and only if } v = 0;$$

additivity in first slot

$$\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle \text{ for all } u, v, w \in V;$$

homogeneity in first slot

$$\langle \lambda u, v\rangle = \lambda\langle u, v\rangle \text{ for all } \lambda \in \mathbf{F} \text{ and all } u, v \in V;$$

conjugate symmetry

$$\langle u, v\rangle = \overline{\langle v, v\rangle} \text{ for all } u, v \in V.$$

Every real number equals its complex conjugate. If we are dealing with a real vector space, then the last condition can be $\langle u, v\rangle = \langle v, u\rangle$ for all $v, w \in V$.

1.2.1 Examples

(a) The ***Euclidean inner product*** on $\mathbf{F}^n$ is defined by

$$\langle (w_1, \ldots, w_n), (z_1, \ldots, z_n) \rangle = w_1 \overline{z_1} + \cdots + w_n \overline{z_n}.$$

(b) If $c_1, \ldots, c_n$ are positive numbers, then an inner product can be defined on $\mathbf{F}^n$ by

$$\langle (w_1, \ldots, w_n), (z_1, \ldots, z_n) \rangle = c_1 w_1 \overline{z_1} + \cdots + c_n w_n \overline{z_n}.$$

(c) An inner product can be defined on the vector space of continuous real-valued functions on the interval $[-1, 1]$ by

$$\langle f, g \rangle = \int_{-1}^{1} f(x)\overline{g(x)} dx.$$

This is an inner product since for example: additivity in the left slot is defined as

$$\begin{aligned} \langle f + h, g \rangle &= \int_{-1}^{1} (f(x) + h(x))\overline{g(x)} dx \\ &= \int_{-1}^{1} f(x)\overline{g(x)} + \int_{-1}^{1} h(x)\overline{g(x)} dx \\ &= \langle f, g \rangle + \langle h, g \rangle. \end{aligned}$$

(d) An inner product can be defined on $\mathcal{P}(\mathbb{R})$ by

$$\langle p, q \rangle = \int_{0}^{\infty} p(x)q(x)e^{-x} dx.$$

(e) The dot product on $\mathbb{R}^n$

$$\langle v, w \rangle = v \cdot w = x_1 y_1 + \cdots + x_n y_n$$

and

$$\langle v, v \rangle = v \cdot v = x_1^2 + \cdots + x_n^2 \geq 0.$$

1.3 Definition of inner product space

An ***inner product space*** is a vector space V along with an inner product on V. For the rest of this chapter, V denotes an inner product space over $\mathbf{F}$.

1.4 Basic properties of an inner product

(a) For each fixed $u \in V$, the function that takes v to $\langle v, u\rangle$ is a linear map from V to $\mathbf{F}$.

Proof.

- $f(v+v') = \langle v+v', u\rangle = \langle v, u\rangle + \langle v', u\rangle = f(v) + f(v')$
- $f(\lambda v) = \cdots = \lambda f(v)$.

$\square$

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

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

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

Proof. This is additivity in the second slot.

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

$\square$

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

Proof. This is homogeneity in the second slot.

$$\begin{aligned}\langle u, \lambda v\rangle &= \overline{\langle \lambda v, u\rangle}\\ &= \overline{\lambda\langle v, u\rangle}\\ &= \overline{\lambda}\,\overline{\langle v, u\rangle}\\ &= \overline{\lambda}\langle u, v\rangle.\end{aligned}$$

$\square$

1.5 Definition of norm, $||v||$

For $v \in V$, the ***norm*** of v, denoted $||v||$, is defined by

$$||v|| = \sqrt{\langle v, v\rangle} \geq 0.$$

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

1.6 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 $\lambda \in \mathbf{F}$.

Proof. (a) The desired result holds because $\langle v, v\rangle = 0$ if and only if $v = 0$.

(b) Suppose $\lambda \in \mathbf{F}$. then

$$\begin{aligned} ||\lambda v||^2 &= \langle \lambda v, \lambda v\rangle \\ &= \lambda\langle v, \lambda v\rangle \\ &= \lambda\overline{\lambda}\langle v, v\rangle \\ &= |\lambda|^2||v||2. \end{aligned}$$

Taking square roots now gives the desired equality.

□

1.7 Definition of orthogonal

Two vectors $u, v \in V$ are called ***orthogonal*** if $\langle u, v\rangle = 0$. We write $u \perp v$ to mean "u is orthogonal to v."

1.8 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.

Proof. If $v \in V$ and $\langle v, v\rangle = 0$, then $v = 0$ (by definition of inner product). □

(c) $u \perp v \Leftrightarrow v \perp u$

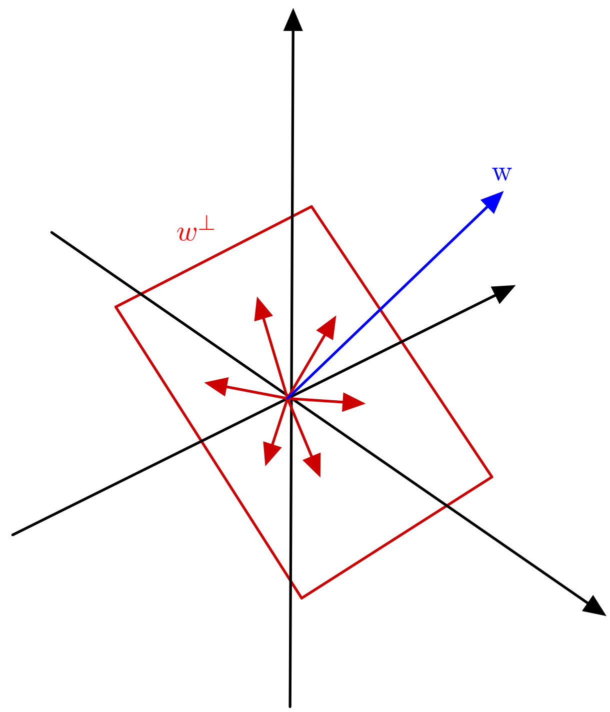
(d) $u \perp w$ and $v \perp w \Rightarrow (u + v) \perp w$.

(e) $u \perp w$ and $\lambda \in \mathbf{F} \Rightarrow (\lambda u) \perp w$.

The last two properties imply that the set

$$w^{\perp} = \{v \mid v \perp w\}$$

is a subspace of V, called the *orthogonal complement of* V.



1.9 Pythagorean Theorem

Suppose u and v are orthogonal vectors in V. Then

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

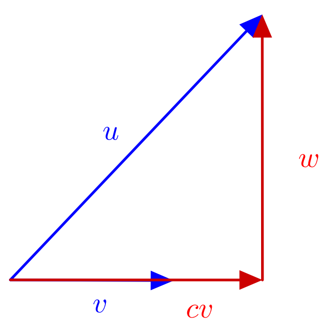
Proof. We have

$$
\begin{aligned}
||u+v||^2 &= \langle u+v, u+v\rangle \\
&= \langle u, u+v\rangle + \langle v, u+v\rangle \\
&= \langle u, u\rangle + \underbrace{\langle u, v\rangle + \langle v, u\rangle}_{0} + \langle v, v\rangle \\
&= \langle u, u\rangle + \langle v, v\rangle \\
&= ||u||^2 + ||v||^2,
\end{aligned}
$$

as desired. □

1.10 Orthogonal Decomposition (Projection)

Given $u, v \in V$, assuming $v \neq 0$. Then we can write u as a sum of two vectors, the first of which is parallel to v and the second is orthogonal to v.



Let $c = \dfrac{\langle u, v\rangle}{||v||^2} = \dfrac{\langle u, v\rangle}{\langle v, v\rangle}$ and let $w = u - cv$. Then $\langle w, v\rangle = 0$ and $u = cv + w$.

Proof. We know $u = cv + w$ holds by the definition of w. We also know that cv is parallel to v by the definition of "parallel." To prove that w is orthogonal to v, we can calculate:

$$
\begin{aligned}
\langle w, v\rangle &= \langle u - cv, v\rangle \\
&= \langle u, v\rangle - c\langle v, v\rangle \\
&= \langle u, v\rangle - \frac{\langle u, v\rangle}{\langle v, v\rangle}\langle v, v\rangle \\
&= \langle u, v\rangle - \langle u, v\rangle = 0.
\end{aligned}
$$

Therefore, $w \perp v$. $\square$

1.11 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.

Proof. Consider two cases:

Case 1. $v = 0$ and in this case, $\langle u, v\rangle = 0, ||u|| \cdot ||v|| = ||u|| \cdot 0 = 0$. So the inequality holds.

Case 2. $v \neq 0$. Consider the orthogonal decomposition
$$u = cv + w$$
where $c = \dfrac{\langle u, v\rangle}{\langle v, v\rangle}$ and $w = u - cv$. We know that $w \perp v$. By Pythagoras' Theorem,
$$\begin{aligned} ||u||^2 &= ||cv||^2 + ||w||^2 \\ &\geq ||cv||^2 \\ &= |c|^2||v||^2 \\ &= \left|\frac{\langle u, v\rangle}{||v||^2}\right|^2 ||v||^2 \\ &= \frac{|\langle u, v\rangle|^2}{||v||^4} \cdot ||v||^2 \\ &= \frac{|\langle u, v\rangle|^2}{||v||^2}. \end{aligned}$$
We just proved that
$$||u||^2 \geq \frac{|\langle u, v\rangle|^2}{||v||^2}.$$
Multiply both sides of the equation by $||v||^2$ and we get
$$||u||^2||v||^2 \geq |\langle u, v\rangle|^2.$$
Take the square root of both sides of the equation and we get
$$||u|| \cdot ||v|| \geq |\langle u, v\rangle|$$
which is the Cauchy-Schwarz inequality. $\square$

1.11.1 Examples of the Cauchy-Schawrz Inequality

(a) If $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$ then

$$|x_1 y_1 + \cdots + x_n y_n|^2 \leq (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2).$$

(b) If f, g are continuous real-valued functions on $[-1, 1]$, then

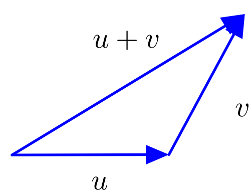
$$\left| \int_{-1}^{1} f(x)g(x)dx \right|^2 \leq \left(\int_{-1}^{1} (f(x))^2 dx \right) \left(\int_{-1}^{1} (g(x))^2 dx \right).$$

1.12 Triangle Inequality

The Triangle Inequality implies that the shortest path between two points is a line segment. 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 multiple of the other.



Proof. We have

$$\begin{aligned}
||u+v||^2 &= \langle u+v, u+v \rangle \\
&= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle \\
&= \langle u, u \rangle + \langle u, v \rangle + \overline{\langle u, v \rangle} + \langle v, v \rangle \\
&\leq \langle u, u \rangle + 2|\langle u, v \rangle| + \langle v, v \rangle \\
&\leq \langle u, u \rangle + 2||u||||v|| + \langle v, v \rangle \quad \text{(Cauchy-Schwarz)} \\
&= ||u||^2 + 2||u||||v|| + ||v||^2 \\
&= (||u|| + ||v||)^2
\end{aligned}$$

Taking the square roots:

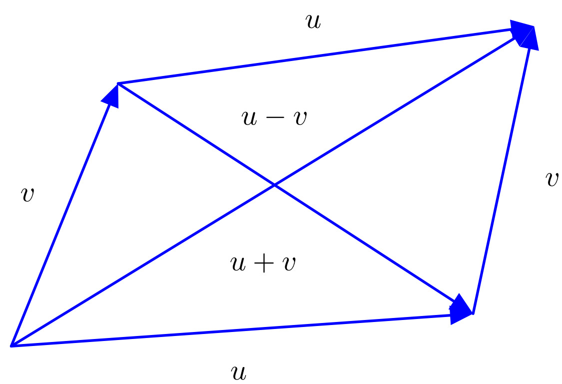
$$||u + v|| \leq ||u|| + ||v||,$$

thus we get the triangle inequality. □

1.13 Parallelogram Equality

In every parallelogram, the sum of the squares of the lengths of the diagonals equals the sum of the squares of the lengths of the four sides.
Suppose $u, v \in V$. Then

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



Proof. We have

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

as desired. □

2 6.B Orthonormal Bases

2.1 Definition of 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.
- In other words, a list $e_1, \dots, e_m$ of vectors in V is orthonormal if

$$\langle e_j, e_k \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

2.1.1 Examples

- The standard basis in $\mathbf{F}^n$ is an orthonormal list.
- In $\mathbb{R}^2$, we have

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

  v_1, v_2, v_3 is orthogonal but *not* orthonormal.
- In $\mathbb{R}^3$, we have

$$w_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, w_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, w_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

  w_1, w_2, w_3 is orthonormal.

Note: If $v_1, \dots, v_n$ is orthogonal and all are non-zero, then

$$\frac{v_1}{||v_1||}, \frac{v_2}{||v_2||}, \dots, \frac{v_n}{||v_n||}$$

is orthonormal.

2.2 The norm of an orthonormal linear combination

If $e_1, \dots, e_m$ 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 $a_1, \dots, a_m \in \mathbf{F}$.

Proof.

$$\begin{aligned}
||v||^2 &= \langle v, v\rangle \\
&= \langle a_1e_1 + \cdots + a_ne_n, a_1e_1 + \cdots + a_ne_n\rangle \\
&= a_1\overline{a_1}\langle e_1, e_1\rangle + a_1\overline{a_2}\langle e_1, e_2\rangle + \cdots + a_1\overline{a_n}\langle e_1, e_n\rangle \\
&\quad + a_2\overline{a_1}\langle e_2, e_1\rangle + a_2\overline{a_3}\langle e_2, e_3\rangle + \cdots + a_2\overline{a_n}\langle e_2, e_n\rangle \\
&\quad + \ldots \\
&\quad + a_n\overline{a_1}\langle e_n, e_1\rangle + a_n\overline{a_2}\langle e_n, e_2\rangle + \cdots + a_n\overline{a_n}\langle e_n, e_n\rangle \\
&= a_1\overline{a_1} + a_2\overline{a_2} + \cdots + a_n\overline{a_n} \\
&= |a_1|^2 + |a_2|^2 + \cdots + |a_n|^2.
\end{aligned}$$

□

2.3 An orthonormal list is linearly independent

Every orthonormal list of vectors is linearly independent.

Proof. Let $e_1, \ldots, e_n$ be orthonormal. Take any $a_1, \ldots, a_n$ and assume

$$a_1e_1 + \cdots + a_ne_n = 0.$$

We need to show that $a_i = 0$ for all i. Consider any index i and we have

$$\begin{aligned}
0 = \langle 0, e_i\rangle &= \langle a_1e_1 + \cdots + a_ne_n, e_i\rangle \\
&= a_1\langle e_1, e_i\rangle + a_2\langle e_2, e_i\rangle + \cdots + a_i\langle e_i, e_i\rangle + \cdots + a_n\langle e_n, e_i\rangle \\
&= 0a_1 + 0a_2 + \cdots + 1a_1 + \cdots + 0a_n \\
&= a_i.
\end{aligned}$$

Since i was arbitrary, we have $a_1 = 0, \ldots, a_n = 0$ as desired. □

2.4 Definition of orthonormal basis

An ***orthonormal basis*** of V is an orthonormal list of vectors in V that is also a basis of V. For example, the standard basis is an orthonormal basis of $\mathbf{F}^n$.

2.5 An orthonormal list of the right length is an orthonormal basis

Every orthonormal list of vectors in V with length dim V is an orthonormal basis of V.

Proof. Since an orthonormal list is linearly independent, any such list must be linearly independent as well; because it has the right length, it is a basis. □

2.5.1 Example

Show that

$$\frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\1\\-1\\1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}-1\\1\\-1\\1\end{bmatrix}$$

is an orthonormal basis of $\mathbf{F}^4$.

Solution: We have

$$\left|\left|\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)\right|\right| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = 1.$$

Similarly, the other three vectors also have norm 1. The inner product of any two distinct vectors in the list above equals 0. Thus, the list above is orthonormal. Because we have an orthonormal list of length four in the four-dimensional vector space $\mathbf{F}^4$, this list is an orthonormal basis. □

2.6 Writing a vector as linear combination of orthonormal basis

Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V and $v \in V$. Then

$$v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n$$

and

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

Proof. Because $e_1, \ldots, e_n$ is a basis of V, there exist coordinates $a_1, \ldots, a_m$ such that

$$v = a_1 e_1 + \cdots + a_n e_n.$$

Then

$$\begin{aligned}\langle v, e_i\rangle &= \langle a_1, e_1 + \ldots a_n e_n, e_i\rangle \\ &= a_1\langle e_1, e_i\rangle + \cdots + a_n\langle e_n, e_i\rangle \\ &= a_i.\end{aligned}$$

Plugging $a_i = \langle v, e_i \rangle$ into v, we get

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

The second claim follows from the fact that the norm of an orthonormal linear combination. $\square$

2.7 Fourier coefficients

Let $e_1, \ldots, e_n$ be an orthonormal basis, and let $v \in V$. In the formula,

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n$$

the coefficients $\langle v, e_1 \rangle, \ldots, \langle v, e_n \rangle$ are the **Fourier coefficients** of v.

Note: Almost the same trick works if $e_1, \ldots, e_n$ is an orthogonal basis (instead of an orthonormal one). In that case, the formula changes slightly as follows: When $v = a_1 e_1 + \cdots + a_n e_n$ then we get

$$\begin{aligned}
\langle v, e_i \rangle &= \langle a_1 e_1 + \cdots + a_n e_n, e_i \rangle \\
&= a_1 \langle e_1, e_i \rangle + \cdots + a_i \langle e_i, e_i \rangle + \cdots + a_n \langle e_n, e_i \rangle \\
&= a_i \langle e_i, e_i \rangle.
\end{aligned}$$

Then the formula for the Fourier coeffiecients becomes

$$a_i = \frac{\langle v, e_i \rangle}{\langle e_i, e_i \rangle}$$

and so we have

$$v = \frac{\langle v, e_1 \rangle}{\langle e_1, e_1 \rangle} e_1 + \cdots + \frac{\langle v, e_n \rangle}{\langle e_n, e_n \rangle} e_n$$

which are the Fourier coefficients for orthogonal basis.

2.8 Existence of orthonormal basis

Every finite-dimensional inner product space has an orthonormal basis. (And therefore, it also has an orthonormal basis as well, by normalizing the basis vector.)

Proof. Since V is finite-dimensional, it has some basis, say $v_1, \ldots, v_n$. However, this basis may not be orthogonal. We will give a procedure for "turning" $v_1, \ldots, v_n$ into an orthogonal basis $w_1, \ldots, w_n$. (The Gram-Schmidt Procedure) $\square$

2.9 Gram-Schmidt Procedure

Note: This version is different than the one in 6.31. I compute an orthogonal basis and Axler computes and orthonormal basis. Axler's computation uses a lot of square roots, and mine does not!

We compute:

$$
\begin{aligned}
w_1 &= v_1 \\
w_2 &= v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1 \\
w_3 &= v_3 - \frac{\langle v_3, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle v_3, w_2\rangle}{\langle w_2, w_2\rangle} w_2 \\
&\vdots \\
w_n &= v_n - \frac{\langle v_n, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \cdots - \frac{\langle v_n, w_{n-1}\rangle}{\langle w_{n-1}, w_{n-1}\rangle} w_{n-1}
\end{aligned}
$$

Claim: $w_1, \ldots, w_n$ is an orthogonal basis of V.

Proof. To prove the claim above, we must show that

1. $w_1, \ldots, w_n$ are orthogonal.
2. $\operatorname{span}(w_1, \ldots, w_n) = \operatorname{span}(v_1, \ldots, v_n)$

For (1), we can prove by induction on n.

- Base case: $n = 1$. Then w_1 is an orthogonal list of vectors (because there is only one vector).
- Induction step: By the induction hypotheses, $w_1, \ldots, w_{n-1}$ are orthogonal. We need to show that $w_1, \ldots, w_n$ are orthogonal. So specifically, we only need to prove that w_n is orthogonal to each $w_1, \ldots, w_{n-1}$. Let

$i \in \{1, \dots, n-1\}$. We compute:

$$\begin{aligned}
\langle w_n, w_i \rangle &= \left\langle v_n - \frac{\langle v_n, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \dots - \frac{\langle v_n, w_{n-1} \rangle}{\langle w_{n-1}, w_{n-1} \rangle} w_{n-1}, w_i \right\rangle \\
&= \langle v_n, w_i \rangle - \frac{\langle v_n, w_i \rangle}{\langle w_1, w_1 \rangle} \langle w_1, w_i \rangle - \dots - \frac{\langle v_n, w_i \rangle}{\langle w_i, w_i \rangle} \langle w_i, w_i \rangle - \dots \\
&\quad - \frac{\langle v_n w_{n-1} \rangle}{\langle w_{n-1}, w_{n-1} \rangle} \langle w_{n-1}, w_i \rangle \\
&= \langle v_n, w_i \rangle - \frac{\langle v_n, w_i \rangle}{\langle w_i, w_i \rangle} \langle w_i, w_i \rangle \\
&= \langle v_n, w_i \rangle - \langle v_n, w_i \rangle = 0.
\end{aligned}$$

For (2), we can prove by induction on n.

- Base case: When $n = 1$, $w_1 = v_1$, so true.
- Induction step: Assume $\operatorname{span}(w_1, \dots, w_{n-1}) = \operatorname{span}(v_1, \dots, v_{n-1})$. We need to show that $\operatorname{span}(w_1, \dots, w_n) = \operatorname{span}(v_1, \dots, v_n)$. We only need to show that $w_n \in \operatorname{span}(v_1, \dots, v_n)$ and $v_n \in \operatorname{span}(w_1, \dots, w_n)$.

  But from
  $$w_n = v_n - \frac{\langle v_n, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \dots - \frac{\langle v_n, w_{n-1} \rangle}{\langle w_{n-1}, w_{n-1}} w_{n-1}$$
  we get
  $$w_n \in \operatorname{span}(w_1, \dots, w_{n-1}, vn) = \operatorname{span}(v_1, \dots, v_{n-1}, v_n).$$
  Also, from
  $$v_n = w_n + \frac{\langle v_n, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \dots + \frac{\langle v_n, w_{n-1} \rangle}{\langle w_{n-1}, w_{n-1} \rangle} w_{n-1}$$
  we have $v_n \in \operatorname{span}(w_1, \dots, w_n)$.

Thus, we proved (1) and (2). $\square$

2.10 Examples

(a) Let $V = \mathbb{R}^3$ with the dot product $\langle v, w \rangle = v \cdot w$. Let

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Use the Gram-Schmidt method to get an orthogonal basis (and then an orthonormal basis) of $\mathbb{R}^3$.

We have

$$w_1 = v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

For w_2, let's compute:

$$\begin{aligned} \langle v_2, w_1 \rangle &= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 1 \\ \langle w_1, w_1 \rangle &= \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 2. \end{aligned}$$

Then we have

$$\begin{aligned} w_2 &= v_2 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = v_2 - \frac{1}{2} w_1 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1. \end{bmatrix} \end{aligned}$$

For w_3, let's compute:

$$\begin{aligned} \langle v_3, w_1 \rangle &= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 2 \\ \langle w_1, w_1 \rangle &= 2 \\ \langle v_3, w_2 \rangle &= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix} = 1 \\ \langle w_3, w_2 \rangle &= \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix} = \frac{3}{2}. \end{aligned}$$

Then we have

$$\begin{aligned} w_3 &= v_3 - \frac{2}{2}w_1 - \frac{1}{\frac{3}{2}}w_2 \\ &= v_3 - w_1 - \frac{2}{3}w_2 \\ &= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{3}\begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ \frac{1}{3}. \end{bmatrix} \end{aligned}$$

So the orthogonal basis are:

$$w_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, w_2 = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix}, w_3 = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ \frac{1}{3} \end{bmatrix}.$$

To get an orthonormal basis, we just have to normalize the vectors:

$$\begin{aligned} e_1 &= \frac{w_1}{||w_1||} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \\ e_2 &= \frac{w_2}{||w_2||} = \sqrt{\frac{2}{3}}\begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix} = \frac{1}{\sqrt{6}}\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \\ e_3 &= \frac{w_3}{||w_3||} = \sqrt{3}\begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ \frac{1}{3} \end{bmatrix} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}. \end{aligned}$$

□

(b) Let $V = \mathbb{R}^3$, with the inner product

$$\langle v, w \rangle = x_1y_1 + 2x_2y_2 + 3x_3y_3.$$

Let

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Apply the Gram-Schmidt Procedure to get an orthogonal basis.

We have

$$w_1 = v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

For w_2, let's compute:

$$\begin{aligned} \langle v_2, w_1 \rangle &= 0 \cdot 1 + 2 \cdot 1 \cdot 1 + 3 \cdot 1 \cdot 0 = 2 \\ \langle w_1, w_1 \rangle &= 1 \cdot 2 \cdot 1 \cdot 1 + 3 \cdot 0 \cdot 0 = 3 \end{aligned}$$

Then we have

$$\begin{aligned} w_2 &= v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = v_2 - \frac{2}{3} w_1 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{3} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} \\ \frac{1}{3} \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}. \end{aligned}$$

For w_3, let's compute:

$$\begin{aligned} \langle v_3, w_1 \rangle &= 1 \cdot +2 \cdot 1 \cdot 1 + 3 \cdot 1 \cdot 0 = 3 \\ \langle v_3, w_2 \rangle &= 1 \left(-\frac{2}{3} \right) + 2 \cdot 1 \cdot \frac{1}{3} + 3 \cdot 1 \cdot 1 = 3 \\ \langle w_2, w_2 \rangle &= \left(-\frac{2}{3} \right)^2 + 2 \left(\frac{1}{3} \right)^2 + 3 \cdot 1^2 = \frac{11}{3}. \end{aligned}$$

Then we have

$$\begin{aligned} w_3 &= v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1} w_1 - \frac{\langle v_3, w_2}{\langle w_2, w_2 \rangle} w_2 \\ &= v_3 - \frac{3}{3} w_1 - \frac{3}{\frac{11}{3}} w_2 \\ &= v_3 - w_1 - \frac{9}{11} w_2 \\ &= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{9}{11} \cdot \frac{1}{3} \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} \frac{6}{11} \\ -\frac{3}{11} \\ \frac{2}{11} \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 6 \\ -3 \\ 2 \end{bmatrix}. \end{aligned}$$

□

2.11 Legendre Polynomials

Let $V = \mathcal{P}(\mathbb{R})$ be the space of polynomials in one variable with the inner product

$$\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)dx.$$

Apply the Gram-Schmidt Procedure to find an orthogonal basis of $\operatorname{span}(1, x, x^2)$.

Remark: These are not orthogonal to begin with. For example:

$$\begin{aligned} \langle 1, x^2 \rangle &= \int_{-1}^{1} 1 \cdot x^2 dx \\ &= \frac{x^3}{3}\Big|_{-1}^{1} = \frac{2}{3}. \end{aligned}$$

Let $v_1 = 1, v_2 = x, v_3 = x^2$. Then we have

$$u_1 = v_1 = 1.$$

For u_2, let us compute:

$$\begin{aligned} \langle v_2, u_1 \rangle &= \int_{-1}^{1} x \cdot 1 dx = \int_{-1}^{1} x dx \\ &= \frac{x^2}{2}\Big|_{-1}^{1} = \frac{1}{2} - \frac{1}{2} = 0 \\ \langle u_1, u_1 \rangle &= \langle 1, 1 \rangle = \int_{-1}^{1} 1 \cdot 1 dx = 2. \end{aligned}$$

Then we have

$$u_2 = v_2 - \frac{\langle v_2, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 = v_2 = x.$$

For u_3, let us compute:

$$\begin{aligned} \langle v_3, u_1 \rangle &= \langle x^2, 1 \rangle = \frac{2}{3} \\ \langle v_3, u_2 \rangle &= \langle x^2, x \rangle = \int_{-1}^{1} x^2 \cdot x dx = \int_{-1}^{1} x^3 dx = 0 \end{aligned}$$

Then we have

$$\begin{aligned} u_3 &= v_3 - \frac{\langle v_3, u_1\rangle}{\langle u_1, u_1\rangle} u_1 - \frac{\langle v_3, u_2\rangle}{\langle u_2, u_2\rangle} u_2 \\ &= v_3 - \frac{\frac{2}{3}}{2} u_1 = v_3 - \frac{1}{3} u_1 \\ &= x^2 - \frac{1}{3}. \end{aligned}$$

The orthogonal basis are

$$u_1 = 1, u_2 = x, u_3 = x^2 - \frac{1}{3}.$$

□

Similarly, apply the Gram-Schmidt Procedure to the polynomials

$$1, x, x^2, x^3, x^4, x^5, x^6, \ldots, ...$$

Answer:

$1, x, x^2 - \frac{1}{3}, x^3 - \frac{3}{5}x, x^4 - \frac{6}{7}x^2 + \frac{3}{35}, x^5 - \frac{10}{9}x^3 + \frac{15}{8}x, x^6 - \frac{15}{11}x^4 + \frac{5}{11}x^2 - \frac{5}{231}, \ldots$

These are called the *Legendre polynomials*.

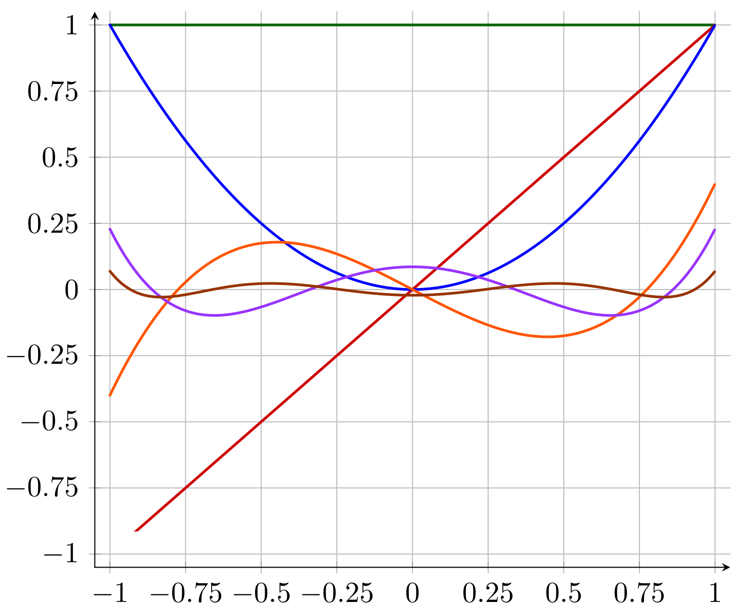


Figure 1: First six Legendre polynomials.

2.12 Orthonormal list extends to 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. Suppose $e_1, \ldots, e_m$ is an orthonormal list of vectors in V. Then $e_1, \ldots, e_m$ is linearly independent. Hence this list can be extended to a basis $e_1, \ldots, e_m, v_1, \ldots, v_n$ of V. Now apply the Gram-Schmidt Procedure to $e_1, \ldots, e_m, v_1, \ldots, v_n$, producing an orthonormal list

$$e_1, \ldots, e_m, f_1, \ldots, f_n;$$

here the formula given by the Gram-Schmidt Procedure leaves the first m vectors unchanged because they are already orthonormal. The list above is an orthonormal basis of V. $\square$

3 6.C Orthogonal Complements and Minimization Problems