UNIT 4

Inner Product Spaces

Unit 4 · (Sections 3.1 – 3.2) and 6.1 – 6.3

MATH 244 — Linear Algebra
Elementary Linear Algebra, Anton & Rorres, 11th edition
King Saud University · Department of Mathematics

What you should be able to do

Sections 3.1 – 3.2

- Add, subtract and scale vectors, as arrows in $\mathbb{R}^2$ and by components in $\mathbb{R}^n$; compute norms, unit vectors, distances, dot products and angles, and use their properties.

Section 6.1

- State the four inner product axioms and check them for a given rule, such as a weighted Euclidean inner product.
- Compute inner products, norms and distances in $\mathbb{R}^n$, M_{nn}, P_n and $C[a, b]$, and simplify expressions such as $\langle \mathbf{u} - 2\mathbf{v},\ 3\mathbf{u} + 4\mathbf{v} \rangle$.

Section 6.2

- Use the Cauchy–Schwarz and triangle inequalities, find the angle between two vectors, and decide whether two vectors are orthogonal for a given inner product.

Section 6.3

- Recognize orthogonal and orthonormal sets and bases, find coordinates relative to them with inner products, and apply the Gram–Schmidt process.

SECTIONS 3.1 – 3.2 (KEY RESULTS)

Vectors in $\mathbb{R}^n$: Length, Distance and Dot Product

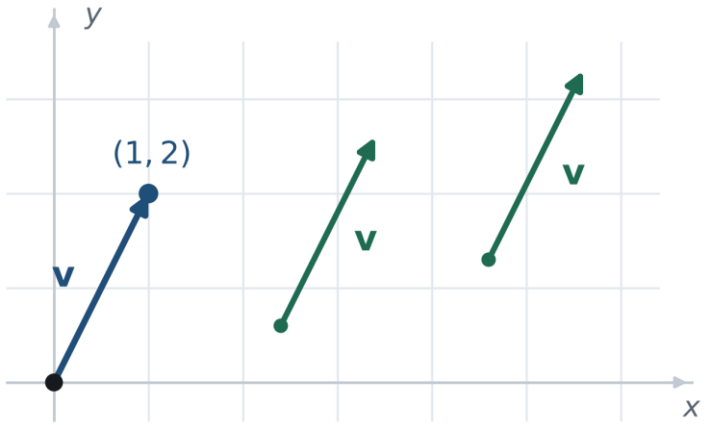
GOAL See vectors as arrows in $\mathbb{R}^2$ and as n-tuples in $\mathbb{R}^n$, and measure them: length, distance, dot product and angle. The dot product is the model for the inner products of Section 6.1.

1. Vectors as arrows: addition, subtraction and scalar multiples
2. Vectors in $\mathbb{R}^n$ and the rules for their operations
3. Norm (length) and unit vectors
4. Distance in $\mathbb{R}^n$
5. The dot product and its properties
6. Cauchy–Schwarz, angles and the triangle inequality

Vectors as arrows

In 2-space and 3-space a vector is drawn as an ARROW: its tail is the INITIAL POINT and its tip is the TERMINAL POINT. Only its length and its direction matter.

- Arrows with the same length and direction are EQUIVALENT: they represent the SAME vector, wherever they are drawn.



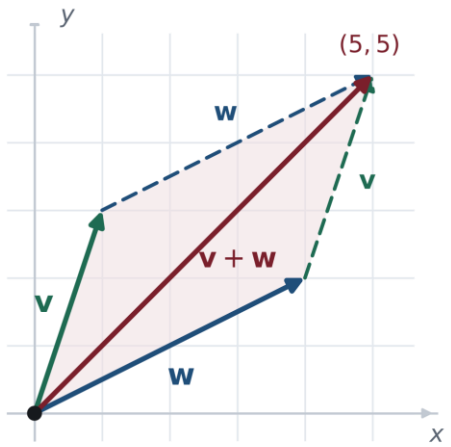
- The ZERO VECTOR $\mathbf{0}$ has length 0: its initial and terminal points coincide.
- Unless we say otherwise, a vector starts at the origin. Its terminal point then has the coordinates of the vector, so (v_1, v_2) can be read as a POINT or as a VECTOR.

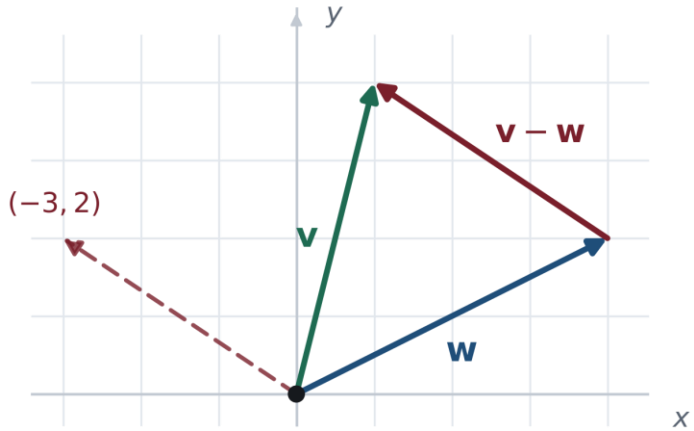
The vector from P_1 to P_2

Terminal point minus initial point: $\overrightarrow{P_1P_2} = (x_2 - x_1,\ y_2 - y_1,\ z_2 - z_1)$.

Example 1: $P_1(2, -1, 4)$, $P_2(7, 5, -8)$: $\overrightarrow{P_1P_2} = (7 - 2,\ 5 - (-1),\ -8 - 4) = (5, 6, -12)$.

Adding and subtracting vectors geometrically





Addition

PARALLELOGRAM RULE: draw $\mathbf{v}$ and $\mathbf{w}$ from the same point; $\mathbf{v}+\mathbf{w}$ is the diagonal of the parallelogram they form. TRIANGLE RULE: put the tail of $\mathbf{w}$ at the tip of $\mathbf{v}$; then $\mathbf{v}+\mathbf{w}$ runs from the tail of $\mathbf{v}$ to the tip of $\mathbf{w}$. Both give $\mathbf{v}+\mathbf{w}=\mathbf{w}+\mathbf{v}$.

$$(1,3)+(4,2)=(5,5)$$

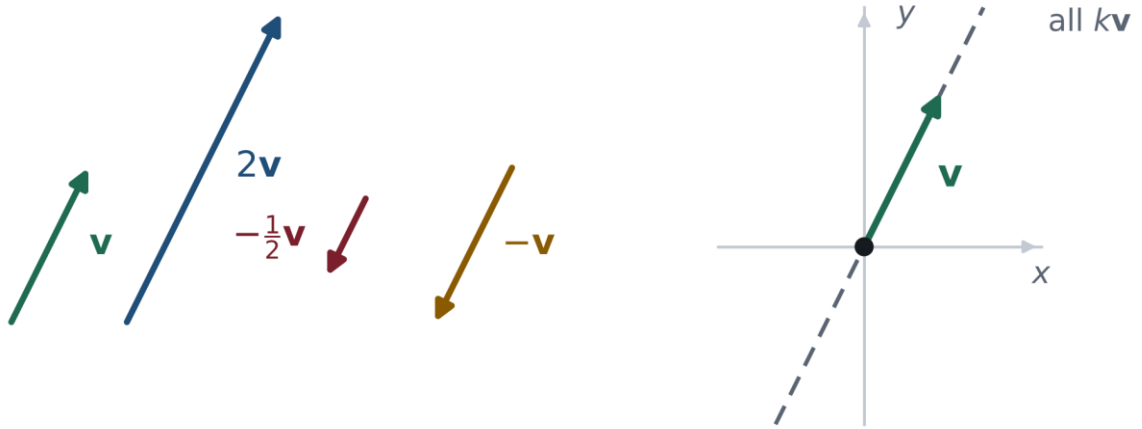
Subtraction

$\mathbf{v}-\mathbf{w}=\mathbf{v}+(-\mathbf{w})$, where $-\mathbf{w}$ has the length of $\mathbf{w}$ and the opposite direction. With $\mathbf{v}$ and $\mathbf{w}$ drawn from the same point, $\mathbf{v}-\mathbf{w}$ is the arrow from the tip of $\mathbf{w}$ to the tip of $\mathbf{v}$.

$$(1,4)-(4,2)=(-3,2)$$

Scalar multiples of a vector

SCALAR MULTIPLE For a scalar k, the vector $k\mathbf{v}$ has length $|k|$ times the length of $\mathbf{v}$. It has the direction of $\mathbf{v}$ if $k > 0$ and the opposite direction if $k < 0$; $0\mathbf{v} = \mathbf{0}$. In components, $k(v_1, v_2) = (kv_1, kv_2)$.



(a) some multiples of $\mathbf{v}$ (b) all multiples: a line

$$\mathbf{v} = (1,2): \qquad 2\mathbf{v} = (2,4), \qquad -1/2\,\mathbf{v} = (-1/2\,, -1), \qquad (-1)\mathbf{v} = -\mathbf{v} = (-1, -2)$$

$\mathbf{v}$ and $k\mathbf{v}$ are PARALLEL (collinear). The multiples of a nonzero $\mathbf{v}$ fill the line through the origin in the direction of $\mathbf{v}$.

Vectors in $\mathbb{R}^n$

DEFINITION $\mathbb{R}^n$ is the set of all ordered n-tuples $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ of real numbers. Two vectors are EQUAL when all their corresponding components are equal. For $\mathbf{v}, \mathbf{w}$ in $\mathbb{R}^n$ and a scalar k:

1) $\mathbf{v} + \mathbf{w} = (v_1 + w_1,\ v_2 + w_2,\ \ldots,\ v_n + w_n)$
2) $k\mathbf{v} = (kv_1,\ kv_2,\ \ldots,\ kv_n)$ and $-\mathbf{v} = (-v_1,\ -v_2,\ \ldots,\ -v_n)$
3) $\mathbf{v} - \mathbf{w} = \mathbf{v} + (-\mathbf{w}) = (v_1 - w_1,\ v_2 - w_2,\ \ldots,\ v_n - w_n)$

Example 3: $\mathbf{v} = (1, -3, 2)$, $\mathbf{w} = (4, 2, 1)$

$$\mathbf{v} + \mathbf{w} = (5, -1, 3), \qquad 2\mathbf{v} = (2, -6, 4), \qquad -\mathbf{w} = (-4, -2, -1), \qquad \mathbf{v} - \mathbf{w} = (-3, -5, 1)$$

For $n > 3$ we cannot draw the arrows, but the algebra is the same. A vector may also be written as a row $[\,v_1 \ \ v_2 \ \ \cdots \ \ v_n\,]$ or as a column, as in the old slides.

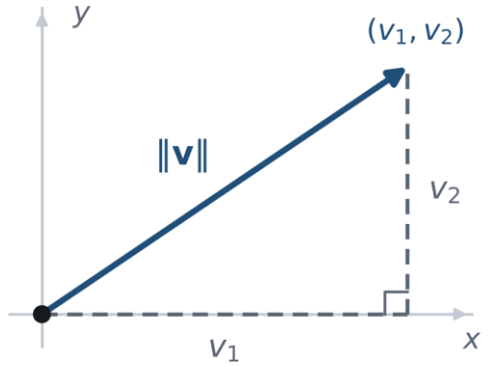
The norm (length) of a vector

DEFINITION 1 The NORM (length, magnitude) of $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$ is the number $\| \mathbf{v} \| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$.

Examples

$$\| (-3,2,1) \| = \sqrt{9 + 4 + 1} = \sqrt{14}$$

$$\| (0,-2,1,4,-2) \| = \sqrt{0 + 4 + 1 + 16 + 4} = \sqrt{25} = 5$$



In $\mathbb{R}^2$ the formula is the Theorem of Pythagoras.

THEOREM 3.2.1 For $\mathbf{v}$ in $\mathbb{R}^n$ and a scalar k:

(a) $\| \mathbf{v} \| \geq 0$

(b) $\| \mathbf{v} \| = 0$ if and only if $\mathbf{v} = \mathbf{0}$

(c) $\| k\mathbf{v} \| = |k| \ \| \mathbf{v} \|$

Proof of (c): $\| k\mathbf{v} \| = \sqrt{k^2 v_1^2 + \cdots + k^2 v_n^2} = \sqrt{k^2}\sqrt{v_1^2 + \cdots + v_n^2} = |k| \ \| \mathbf{v} \|$. ■

Unit vectors and normalizing

A vector of norm 1 is a UNIT VECTOR. If $\mathbf{v} \neq \mathbf{0}$, then $\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}$ is the unit vector in the direction of $\mathbf{v}$: by Theorem 3.2.1(c), $\| \mathbf{u} \| = \| \mathbf{v} \| / \| \mathbf{v} \| = 1$. Finding $\mathbf{u}$ is called NORMALIZING $\mathbf{v}$.

Example: normalize $\mathbf{v} = (3, -1, 2)$

$$\| \mathbf{v} \| = \sqrt{9 + 1 + 4} = \sqrt{14}, \qquad \mathbf{u} = \frac{1}{\sqrt{14}}(3, -1, 2) = \left(\frac{3}{\sqrt{14}}, -\frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}} \right)$$

$$\| \mathbf{u} \| = \sqrt{\frac{9}{14} + \frac{1}{14} + \frac{4}{14}} = 1$$

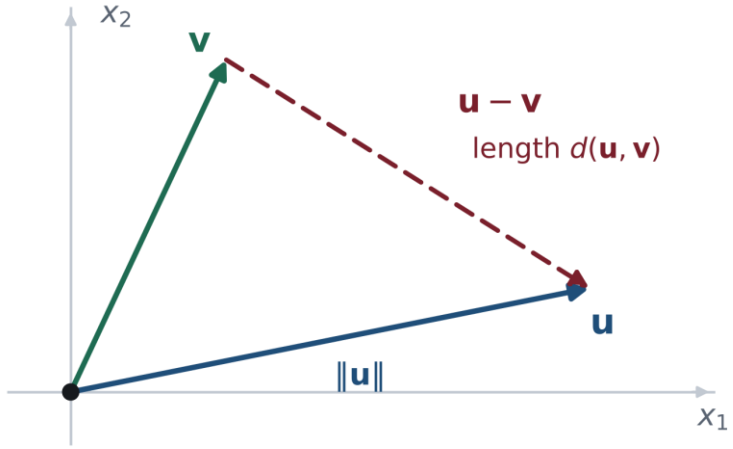
The STANDARD UNIT VECTORS of $\mathbb{R}^n$ are $\mathbf{e}_1 = (1, 0, \ldots, 0)$, …, $\mathbf{e}_n = (0, \ldots, 0, 1)$, and $\mathbf{v} = v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n$; for example $(7, 3, -4, 5) = 7\mathbf{e}_1 + 3\mathbf{e}_2 - 4\mathbf{e}_3 + 5\mathbf{e}_4$.

Distance in $\mathbb{R}^n$

DEFINITION 2 The DISTANCE between $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ in $\mathbb{R}^n$ is

$$d(\mathbf{u}, \mathbf{v}) = \| \mathbf{u} - \mathbf{v} \| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}.$$

In $\mathbb{R}^2$ and $\mathbb{R}^3$ this is the usual distance between the points $\mathbf{u}$ and $\mathbf{v}$: the length of the arrow $\mathbf{u} - \mathbf{v}$ from the tip of $\mathbf{v}$ to the tip of $\mathbf{u}$.



Example 4: $\mathbf{u} = (1,3,-2,7)$, $\mathbf{v} = (0,7,2,2)$

$$d(\mathbf{u}, \mathbf{v}) = \sqrt{(1-0)^2 + (3-7)^2 + (-2-2)^2 + (7-2)^2} = \sqrt{58}$$

By Theorem 3.2.1: $d(\mathbf{u}, \mathbf{v}) \geq 0$, $d(\mathbf{u}, \mathbf{v}) = 0$ if and only if $\mathbf{u} = \mathbf{v}$, and $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$.

The dot product in $\mathbb{R}^2$ and $\mathbb{R}^3$

DEFINITION 3 If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ and θ is the angle between them, $0 \le \theta \le \pi$, their DOT PRODUCT is $\mathbf{u} \cdot \mathbf{v} = \| \mathbf{u} \| \| \mathbf{v} \| \cos\theta$. If $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$, then $\mathbf{u} \cdot \mathbf{v} = 0$.

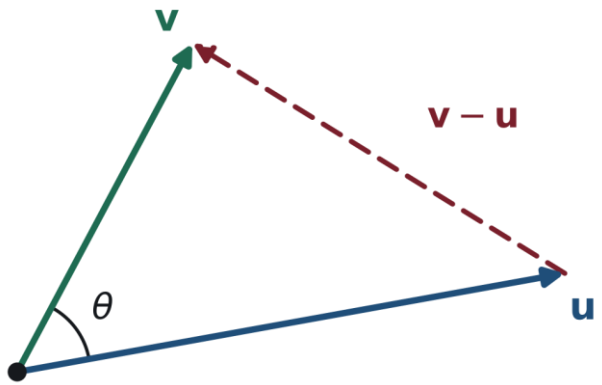
From the angle to the components

The law of cosines in the triangle with sides $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{v} - \mathbf{u}$ gives

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

so $\mathbf{u} \cdot \mathbf{v} = \frac{1}{2}\left(\| \mathbf{u} \|^2 + \| \mathbf{v} \|^2 - \| \mathbf{v} - \mathbf{u} \|^2 \right)$. Writing the norms with components and simplifying:

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 \text{ in } \mathbb{R}^2, \qquad \mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3 \text{ in } \mathbb{R}^3$$



The dot product in $\mathbb{R}^n$

DEFINITION 4 For $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$ in $\mathbb{R}^n$, the DOT PRODUCT (Euclidean inner product) is the NUMBER $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$: multiply corresponding components and add.

Examples

$$(1,2,0,-3) \cdot (3,-2,4,2) = (1)(3) + (2)(-2) + (0)(4) + (-3)(2) = -7$$

$$(-1,3,5,7) \cdot (-3,-4,1,0) = (-1)(-3) + (3)(-4) + (5)(1) + (7)(0) = -4$$

With $\mathbf{u} = \mathbf{v}$: $\mathbf{v} \cdot \mathbf{v} = v_1^2 + v_2^2 + \cdots + v_n^2 = \| \mathbf{v} \|^2$, so $\| \mathbf{v} \| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$.

Written as columns, $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \mathbf{v}^T \mathbf{u}$ (a 1×1 matrix). $\mathbb{R}^n$ with these operations, the norm and the dot product is EUCLIDEAN n-SPACE.

Properties of the dot product

THEOREM 3.2.2 For $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathbb{R}^n$ and a scalar k:

(a) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ (symmetry)

(b) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ (distributive)

(c) $k(\mathbf{u} \cdot \mathbf{v}) = (k\mathbf{u}) \cdot \mathbf{v}$ (homogeneity)

(d) $\mathbf{v} \cdot \mathbf{v} \geq 0$, and $\mathbf{v} \cdot \mathbf{v} = 0$ if and only if $\mathbf{v} = \mathbf{0}$ (positivity)

THEOREM 3.2.3 Also:

(a) $\mathbf{0} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{0} = 0$

(b) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$

(c) $\mathbf{u} \cdot (\mathbf{v} - \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{w}$

(d) $(\mathbf{u} - \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} - \mathbf{v} \cdot \mathbf{w}$

(e) $k(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (k\mathbf{v})$

Proof of 3.2.2(c): $k(\mathbf{u} \cdot \mathbf{v}) = k(u_1v_1 + \cdots + u_nv_n) = (ku_1)v_1 + \cdots + (ku_n)v_n = (k\mathbf{u}) \cdot \mathbf{v}$. For (d): $\mathbf{v} \cdot \mathbf{v} = \| \mathbf{v} \|^2$, then Theorem 3.2.1. The other parts are proved in the same way.

Theorem 3.2.2 is the model: its four properties become the AXIOMS of an inner product in Section 6.1.

EXAMPLE

Calculating with dot products

Let $\mathbf{u} = (2, -2)$, $\mathbf{v} = (5,8)$ and $\mathbf{w} = (-4,3)$. Then

(a) $\mathbf{u} \cdot \mathbf{v} = (2)(5) + (-2)(8) = -6$

(b) $(\mathbf{u} \cdot \mathbf{v})\mathbf{w} = -6\mathbf{w} = -6(-4,3) = (24, -18)$

(c) $\mathbf{u} \cdot (2\mathbf{v}) = 2(\mathbf{u} \cdot \mathbf{v}) = 2(-6) = -12$

(d) $\| \mathbf{w} \|^2 = \mathbf{w} \cdot \mathbf{w} = (-4)(-4) + (3)(3) = 25$

(e) $\mathbf{u} \cdot (\mathbf{v} - 2\mathbf{w}) = \mathbf{u} \cdot (5 - (-8),\ 8 - 6) = \mathbf{u} \cdot (13,2) = (2)(13) + (-2)(2) = 26 - 4 = 22$

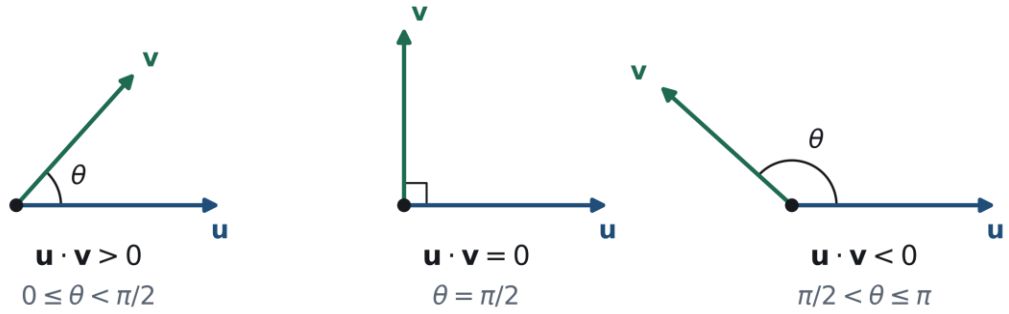
In (b) the answer is a VECTOR: the number $\mathbf{u} \cdot \mathbf{v}$ times $\mathbf{w}$. An expression such as $\mathbf{u} \cdot (\mathbf{v} \cdot \mathbf{w})$ makes no sense: $\mathbf{v} \cdot \mathbf{w}$ is a number, and a dot product needs two vectors.

The Cauchy–Schwarz inequality and angles in $\mathbb{R}^n$

THEOREM 3.2.4 (Cauchy–Schwarz) For $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$: $|\mathbf{u} \cdot \mathbf{v}| \leq \| \mathbf{u} \| \| \mathbf{v} \|$. It is proved in Section 6.2, for every inner product.

So $-1 \leq \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \leq 1$ for nonzero $\mathbf{u}, \mathbf{v}$, and the ANGLE θ between them is defined by

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\| \mathbf{u} \| \| \mathbf{v} \|}, \qquad 0 \leq \theta \leq \pi$$



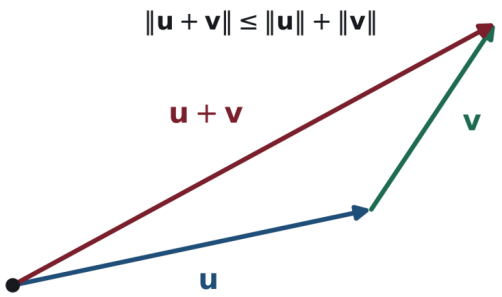
Example: $\mathbf{u} = (-4,0,2,-2)$, $\mathbf{v} = (2,0,-1,1)$: $\cos\theta = \frac{-12}{\sqrt{24}\sqrt{6}} = -1$, so $\theta = \pi$ (indeed $\mathbf{u} = -2\mathbf{v}$). The angle with the zero vector is not defined.

The triangle inequality and the parallelogram equation

THEOREM 3.2.5 For $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathbb{R}^n$:

(a) $\| \mathbf{u} + \mathbf{v} \| \leq \| \mathbf{u} \| + \| \mathbf{v} \|$ (triangle inequality)

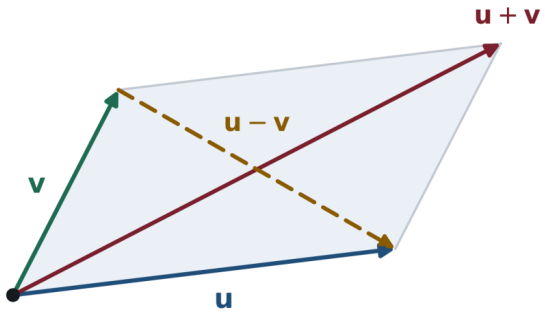
(b) $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$



Key step of (a), by Cauchy–Schwarz: $\| \mathbf{u} + \mathbf{v} \|^2 = \| \mathbf{u} \|^2 + 2\,\mathbf{u} \cdot \mathbf{v} + \| \mathbf{v} \|^2 \leq \| \mathbf{u} \|^2 + 2 \| \mathbf{u} \| \| \mathbf{v} \| + \| \mathbf{v} \|^2 = (\| \mathbf{u} \| + \| \mathbf{v} \|)^2$.

THEOREM 3.2.6 (parallelogram equation) For $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$:

$$\| \mathbf{u} + \mathbf{v} \|^2 + \| \mathbf{u} - \mathbf{v} \|^2 = 2\left(\| \mathbf{u} \|^2 + \| \mathbf{v} \|^2 \right)$$



Key step: add $\| \mathbf{u} \pm \mathbf{v} \|^2 = \| \mathbf{u} \|^2 \pm 2\,\mathbf{u} \cdot \mathbf{v} + \| \mathbf{v} \|^2$.

SECTION 6.1

Inner Products

GOAL Extend the dot product of $\mathbb{R}^n$ to any vector space. Four properties of the dot product become the axioms of an inner product, and every inner product brings its own length and distance.

1. The dot product as the model
2. The inner product axioms
3. Weighted Euclidean inner products
4. Norm, distance and unit circles
5. Inner products on M_{nn}, P_n and $C[a, b]$
6. Algebraic properties of inner products

The dot product is the model

Section 3.2 measured everything with the dot product: $\| \mathbf{v} \| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$, $d(\mathbf{u}, \mathbf{v}) = \| \mathbf{u} - \mathbf{v} \|$, $\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \, \|\mathbf{v}\|}$. Its four basic properties become axioms:

	Dot product on $\mathbb{R}^n$ (Theorems 3.2.2, 3.2.3)	**Inner product on V (Definition 1)**
Symmetry	$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$	$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
Additivity	$(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$	$\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
Homogeneity	$(k\mathbf{u}) \cdot \mathbf{v} = k(\mathbf{u} \cdot \mathbf{v})$	$\langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle$
Positivity	$\mathbf{v} \cdot \mathbf{v} \geq 0$, and $= 0$ only for $\mathbf{v} = \mathbf{0}$	$\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, and $= 0$ only for $\mathbf{v} = \mathbf{0}$

> Section 6.1 keeps these four properties as AXIOMS. Any rule $\langle \mathbf{u}, \mathbf{v} \rangle$ on a vector space V that has them gives a norm, a distance and an angle in the same way, and the results of Section 3.2 carry over (Section 6.2).

DEFINITION 1

Inner products and inner product spaces

DEFINITION 1 An INNER PRODUCT on a real vector space V is a rule that assigns a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ to each pair of vectors $\mathbf{u}, \mathbf{v}$ in V so that, for all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in V and all scalars k:

1) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ (symmetry)

2) $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ (additivity)

3) $\langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle$ (homogeneity)

4) $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$ (positivity)

A real vector space with an inner product is a REAL INNER PRODUCT SPACE.

Example: the Euclidean inner product

$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \cdots + u_n v_n$ satisfies the four axioms (Theorems 3.2.2 and 3.2.3). It is the EUCLIDEAN (standard) inner product, and $\mathbb{R}^n$ with it is EUCLIDEAN n-SPACE.

To show that a rule is NOT an inner product, one counterexample to one axiom is enough.

EXAMPLE 1

A weighted Euclidean inner product on $\mathbb{R}^2$

Show that $\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 2u_2v_2$ is an inner product on $\mathbb{R}^2$. Let $\mathbf{w} = (w_1, w_2)$ be a third vector.

Axiom 1

Swapping $\mathbf{u}$ and $\mathbf{v}$ gives $3v_1u_1 + 2v_2u_2$, the same number.

Axiom 2

$$\begin{aligned} \langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle &= 3(u_1 + v_1)w_1 + 2(u_2 + v_2)w_2 \\ &= (3u_1w_1 + 2u_2w_2) + (3v_1w_1 + 2v_2w_2) \\ &= \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle \end{aligned}$$

Axiom 3

$$\langle k\mathbf{u}, \mathbf{v} \rangle = 3(ku_1)v_1 + 2(ku_2)v_2 = k\langle \mathbf{u}, \mathbf{v} \rangle$$

Axiom 4

$\langle \mathbf{v}, \mathbf{v} \rangle = 3v_1^2 + 2v_2^2 \geq 0$, and it is 0 only when $v_1 = v_2 = 0$, that is, when $\mathbf{v} = \mathbf{0}$.

The weights 3 and 2 are what matter: both are POSITIVE. Here w_1, w_2 are the components of $\mathbf{w}$, not weights.

When does a weighted sum give an inner product?

WEIGHTED EUCLIDEAN INNER PRODUCT If $w_1, w_2, \ldots, w_n$ are POSITIVE real numbers (the weights), then $\langle \mathbf{u}, \mathbf{v} \rangle = w_1 u_1 v_1 + w_2 u_2 v_2 + \cdots + w_n u_n v_n$ is an inner product on $\mathbb{R}^n$. All weights equal to 1 give the Euclidean inner product.

A rule that is NOT an inner product

On $\mathbb{R}^3$, $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 - 2u_2 v_2 + u_3 v_3$ has a negative weight. Axioms 1–3 hold, but for $\mathbf{v} = (1,2,1)$:

$$\langle \mathbf{v}, \mathbf{v} \rangle = (1)(1) - 2(2)(2) + (1)(1) = -6 < 0$$

so the positivity axiom fails, and the rule is not an inner product.

A zero weight fails too: for $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1$ on $\mathbb{R}^2$, the vector $\mathbf{v} = (0,1) \neq \mathbf{0}$ has $\langle \mathbf{v}, \mathbf{v} \rangle = 0$.

Norm and distance in an inner product space

DEFINITION 2 In a real inner product space V, the NORM of $\mathbf{v}$ is $\| \mathbf{v} \| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$, and the DISTANCE between $\mathbf{u}$ and $\mathbf{v}$ is $d(\mathbf{u}, \mathbf{v}) = \| \mathbf{u} - \mathbf{v} \| = \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle}$. A vector of norm 1 is a UNIT VECTOR.

THEOREM 6.1.1 For $\mathbf{u}, \mathbf{v}$ in V and a scalar k:

(a) $\| \mathbf{v} \| \geq 0$, with equality if and only if $\mathbf{v} = \mathbf{0}$;

(b) $\| k\mathbf{v} \| = |k| \ \| \mathbf{v} \|$;

(c) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$;

(d) $d(\mathbf{u}, \mathbf{v}) \geq 0$, with equality if and only if $\mathbf{u} = \mathbf{v}$.

Proof of (b)

$\| k\mathbf{v} \|^2 = \langle k\mathbf{v}, k\mathbf{v} \rangle = k\langle \mathbf{v}, k\mathbf{v} \rangle = k\langle k\mathbf{v}, \mathbf{v} \rangle = k^2 \langle \mathbf{v}, \mathbf{v} \rangle = k^2 \| \mathbf{v} \|^2$ (homogeneity, symmetry, homogeneity), and $\sqrt{k^2} = |k|$. ■

EXAMPLE 2

Norm and distance depend on the inner product

Take $\mathbf{u} = (1,0)$ and $\mathbf{v} = (0,1)$ in $\mathbb{R}^2$, so $\mathbf{u} - \mathbf{v} = (1,-1)$.

	Euclidean: $u_1v_1 + u_2v_2$	**Weighted:** $3u_1v_1 + 2u_2v_2$
$\| \mathbf{u} \|$	$\sqrt{1^2 + 0^2} = 1$	$\sqrt{3(1)(1) + 2(0)(0)} = \sqrt{3}$
$d(\mathbf{u}, \mathbf{v})$	$\sqrt{1^2 + (-1)^2} = \sqrt{2}$	$\sqrt{3(1)(1) + 2(-1)(-1)} = \sqrt{5}$

The same two vectors have different norms and a different distance. When you compute a norm, a distance or an angle, always say which inner product you are using.

Unit circles

DEFINITION 3 The UNIT CIRCLE (or unit sphere) of an inner product space V is the set $\{\mathbf{u} \in V : \| \mathbf{u} \| = 1\}$.



(a) Euclidean: $x^2 + y^2 = 1$ (b) $\frac{1}{9}u_1v_1 + \frac{1}{4}u_2v_2$: $\frac{x^2}{9} + \frac{y^2}{4} = 1$

(a) Euclidean: $\| (x, y) \| = \sqrt{x^2 + y^2} = 1$ gives the circle $x^2 + y^2 = 1$.

(b) $\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{9}u_1v_1 + \frac{1}{4}u_2v_2$: $\| (x, y) \| = \sqrt{\frac{x^2}{9} + \frac{y^2}{4}} = 1$ gives the ellipse $\frac{x^2}{9} + \frac{y^2}{4} = 1$.

Every point of the ellipse is at distance 1 from the origin for this inner product: the weights change what "length" means.

Inner products on $\mathbb{R}^n$ generated by a matrix

If A is an INVERTIBLE $n \times n$ matrix and $\mathbf{u}, \mathbf{v}$ are written as columns, then $\langle \mathbf{u}, \mathbf{v} \rangle = A\mathbf{u} \cdot A\mathbf{v} = \mathbf{v}^T A^T A \mathbf{u}$ is an inner product on $\mathbb{R}^n$, the inner product GENERATED BY A.

Examples 4 and 5

$A = I$ gives the Euclidean inner product. A diagonal matrix with entries $\sqrt{w_1}, \ldots, \sqrt{w_n}$ gives a weighted one:

$$A = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \end{bmatrix}: \quad A\mathbf{u} \cdot A\mathbf{v} = 3u_1v_1 + 2u_2v_2$$

A matrix that is not diagonal

For $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$, $\mathbf{u} = (1,2)$, $\mathbf{v} = (3,-1)$:

$$A\mathbf{u} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad A\mathbf{v} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \quad \langle \mathbf{u}, \mathbf{v} \rangle = 6 - 8 = -2$$

Why A must be invertible: $\langle \mathbf{v}, \mathbf{v} \rangle = \| A\mathbf{v} \|^2$ is 0 only when $A\mathbf{v} = \mathbf{0}$, and that forces $\mathbf{v} = \mathbf{0}$ exactly when A is invertible.

EXAMPLE 6

The standard inner product on M_{nn}

EXAMPLE 6 For matrices U, V in M_{nn}, $\langle U, V\rangle = \operatorname{tr}(U^T V)$ is an inner product, the STANDARD inner product on M_{nn}. It multiplies corresponding entries and adds them; for 2×2 matrices:

$$\left\langle \begin{bmatrix} u_1 & u_2 \\ u_3 & u_4 \end{bmatrix}, \begin{bmatrix} v_1 & v_2 \\ v_3 & v_4 \end{bmatrix} \right\rangle = u_1 v_1 + u_2 v_2 + u_3 v_3 + u_4 v_4$$

For $U = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $V = \begin{bmatrix} -1 & 0 \\ 3 & 2 \end{bmatrix}$:

$$\langle U, V\rangle = 1(-1) + 2(0) + 3(3) + 4(2) = 16$$

$$\| U \| = \sqrt{1 + 4 + 9 + 16} = \sqrt{30}, \qquad \| V \| = \sqrt{1 + 0 + 9 + 4} = \sqrt{14}$$

The old slides write $\operatorname{tr}(AB^T)$; it gives the same number, the sum of the products of corresponding entries.

EXAMPLE 7

The standard inner product on P_n

EXAMPLE 7 For $\mathbf{p} = a_0 + a_1x + \cdots + a_nx^n$ and $\mathbf{q} = b_0 + b_1x + \cdots + b_nx^n$ in P_n, the STANDARD inner product multiplies corresponding coefficients: $\langle \mathbf{p}, \mathbf{q} \rangle = a_0b_0 + a_1b_1 + \cdots + a_nb_n$, so $\| \mathbf{p} \| = \sqrt{a_0^2 + a_1^2 + \cdots + a_n^2}$.

Example in P_2

$\mathbf{p} = 1 - 2x^2$ and $\mathbf{q} = 4 - 2x + x^2$ have the coefficient lists $(1,0,-2)$ and $(4,-2,1)$.

$$\langle \mathbf{p}, \mathbf{q} \rangle = (1)(4) + (0)(-2) + (-2)(1) = 2, \qquad \| \mathbf{q} \| = \sqrt{4^2 + (-2)^2 + 1^2} = \sqrt{21}$$

$$\mathbf{p} - \mathbf{q} = -3 + 2x - 3x^2, \qquad d(\mathbf{p}, \mathbf{q}) = \sqrt{(-3)^2 + 2^2 + (-3)^2} = \sqrt{22}$$

The evaluation inner product on P_n

EXAMPLE 8 Choose $n+1$ distinct numbers $x_0, x_1, \ldots, x_n$ (the sample points). Then $\langle \mathbf{p}, \mathbf{q} \rangle = p(x_0)q(x_0) + p(x_1)q(x_1) + \cdots + p(x_n)q(x_n)$ is an inner product on P_n, the EVALUATION inner product.

Positivity: if $\langle \mathbf{p}, \mathbf{p} \rangle = p(x_0)^2 + \cdots + p(x_n)^2 = 0$, then $\mathbf{p}$ has the $n+1$ roots $x_0, \ldots, x_n$. A nonzero polynomial of degree at most n has at most n roots, so $\mathbf{p} = \mathbf{0}$.

Example 9: sample points $-2,\ 0,\ 2$ in P_2, $\mathbf{p} = x^2$, $\mathbf{q} = 1 + x$

$$\langle \mathbf{p}, \mathbf{q} \rangle = p(-2)q(-2) + p(0)q(0) + p(2)q(2) = (4)(-1) + (0)(1) + (4)(3) = 8$$

$$\| \mathbf{p} \| = \sqrt{p(-2)^2 + p(0)^2 + p(2)^2} = \sqrt{16 + 0 + 16} = \sqrt{32} = 4\sqrt{2}$$

Algebraic properties of inner products

THEOREM 6.1.2 For $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in a real inner product space and a scalar k:

(a) $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$

(b) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$

(c) $\langle \mathbf{u}, \mathbf{v} - \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{u}, \mathbf{w} \rangle$

(d) $\langle \mathbf{u} - \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle - \langle \mathbf{v}, \mathbf{w} \rangle$

(e) $k\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, k\mathbf{v} \rangle$

Proof of (b)

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle &= \langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle && \text{symmetry} \\ &= \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle && \text{additivity} \\ &= \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle && \text{symmetry} \end{aligned}$$

Proof of (a)

$\langle \mathbf{0}, \mathbf{v} \rangle = \langle 0\mathbf{v}, \mathbf{v} \rangle = 0\langle \mathbf{v}, \mathbf{v} \rangle = 0$ (homogeneity with $k = 0$); then symmetry gives $\langle \mathbf{v}, \mathbf{0} \rangle = 0$.

EXAMPLE 12

Calculating with inner products

Expand $\langle \mathbf{u} - 2\mathbf{v},\ 3\mathbf{u} + 4\mathbf{v} \rangle$ using only the axioms and Theorem 6.1.2:

$$\begin{aligned} \langle \mathbf{u} - 2\mathbf{v},\ 3\mathbf{u} + 4\mathbf{v} \rangle &= \langle \mathbf{u}, 3\mathbf{u} + 4\mathbf{v} \rangle - \langle 2\mathbf{v}, 3\mathbf{u} + 4\mathbf{v} \rangle \\ &= \langle \mathbf{u}, 3\mathbf{u} \rangle + \langle \mathbf{u}, 4\mathbf{v} \rangle - \langle 2\mathbf{v}, 3\mathbf{u} \rangle - \langle 2\mathbf{v}, 4\mathbf{v} \rangle \\ &= 3\langle \mathbf{u}, \mathbf{u} \rangle + 4\langle \mathbf{u}, \mathbf{v} \rangle - 6\langle \mathbf{v}, \mathbf{u} \rangle - 8\langle \mathbf{v}, \mathbf{v} \rangle \\ &= 3 \parallel \mathbf{u} \parallel^2 -2\langle \mathbf{u}, \mathbf{v} \rangle - 8 \parallel \mathbf{v} \parallel^2 \end{aligned}$$

For instance, if $\parallel \mathbf{u} \parallel = 2$, $\parallel \mathbf{v} \parallel = 1$ and $\langle \mathbf{u}, \mathbf{v} \rangle = -1$, the value is $3(4) - 2(-1) - 8(1) = 6$.

It works like multiplying out $(a - 2b)(3a + 4b)$, with $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ to combine the middle terms.

SECTION 6.2

Angle and Orthogonality in Inner Product Spaces

GOAL Prove the Cauchy–Schwarz inequality for every inner product, use it to define angles, and study orthogonality, which now depends on the inner product in use.

1. The Cauchy–Schwarz inequality
2. The angle between two vectors
3. The triangle inequality
4. Orthogonal vectors
5. The generalized Theorem of Pythagoras
6. Vectors orthogonal to a set

THEOREM 6.2.1

The Cauchy–Schwarz inequality

THEOREM 6.2.1 (Cauchy–Schwarz) If $\mathbf{u}$ and $\mathbf{v}$ are vectors in a real inner product space, then $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \| \mathbf{u} \| \, \| \mathbf{v} \|$.

Proof

If $\mathbf{u} = \mathbf{0}$, both sides are 0. Let $\mathbf{u} \neq \mathbf{0}$. For every real number t, the positivity axiom gives

$$0 \leq \langle t\mathbf{u} + \mathbf{v},\ t\mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle t^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle t + \langle \mathbf{v}, \mathbf{v} \rangle$$

A quadratic $at^2 + bt + c$ with $a > 0$ that is never negative has no two distinct real roots, so $b^2 - 4ac \leq 0$:

$$4\langle \mathbf{u}, \mathbf{v} \rangle^2 - 4\langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle \leq 0, \quad \text{that is,} \quad \langle \mathbf{u}, \mathbf{v} \rangle^2 \leq \| \mathbf{u} \|^2 \, \| \mathbf{v} \|^2$$

Taking square roots gives $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \| \mathbf{u} \| \, \| \mathbf{v} \|$. ■

Equality holds exactly when $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent (Exercise 48).

The angle between two vectors

For nonzero $\mathbf{u}$ and $\mathbf{v}$, Cauchy–Schwarz gives $-1 \le \frac{\langle \mathbf{u},\mathbf{v} \rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|} \le 1$, so there is exactly one angle θ with

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\| \mathbf{u} \| \, \| \mathbf{v} \|}, \qquad 0 \le \theta \le \pi$$

This θ is the ANGLE between $\mathbf{u}$ and $\mathbf{v}$ (for the given inner product).

Example 1: M_{22}, standard inner product

For $U = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $V = \begin{bmatrix} -1 & 0 \\ 3 & 2 \end{bmatrix}$ of Example 6 in 6.1:

$$\cos\theta = \frac{16}{\sqrt{30}\sqrt{14}} = \frac{16}{\sqrt{420}} \approx 0.78$$

Cauchy–Schwarz in $\mathbb{R}^3$ (Euclidean)

$\mathbf{u} = (1,-1,3)$, $\mathbf{v} = (2,0,-1)$: $\mathbf{u} \cdot \mathbf{v} = 2 + 0 - 3 = -1$, and

$$|-1| = 1 \le \| \mathbf{u} \| \, \| \mathbf{v} \| = \sqrt{11}\sqrt{5} = \sqrt{55}$$

THEOREM 6.2.2

The triangle inequality

THEOREM 6.2.2 For $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in a real inner product space:

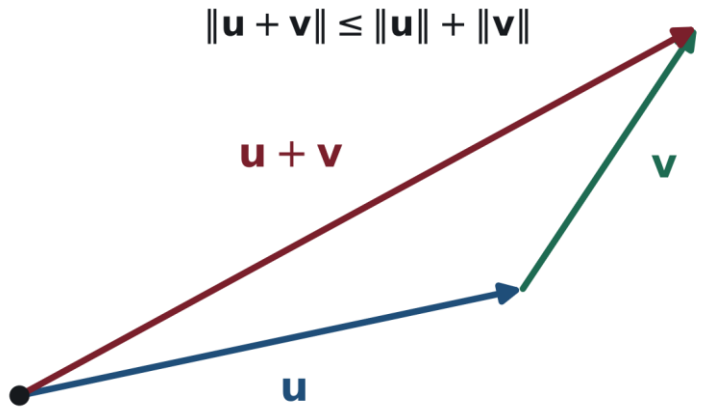
(a) $\| \mathbf{u} + \mathbf{v} \| \leq \| \mathbf{u} \| + \| \mathbf{v} \|$ (triangle inequality for vectors)

(b) $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$ (triangle inequality for distances)

Proof of (a)

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

The last step is Cauchy–Schwarz. Take square roots. ■



(b) is (a) applied to $\mathbf{u} - \mathbf{w}$ and $\mathbf{w} - \mathbf{v}$, whose sum is $\mathbf{u} - \mathbf{v}$.

Orthogonal vectors

DEFINITION 1 Two vectors $\mathbf{u}$ and $\mathbf{v}$ in an inner product space are ORTHOGONAL if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. For nonzero vectors this means $\theta = \pi/2$; the zero vector is orthogonal to every vector.

Example 2: orthogonality depends on the inner product

$\mathbf{u} = (1,1)$ and $\mathbf{v} = (1,-1)$: $\mathbf{u} \cdot \mathbf{v} = 1 - 1 = 0$, so they are orthogonal for the Euclidean inner product; but for $\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 2u_2v_2$, $\langle \mathbf{u}, \mathbf{v} \rangle = 3 - 2 = 1 \neq 0$, so they are NOT orthogonal.

Example 3: orthogonal matrices in M_{22} (standard inner product)

$$\left\langle \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \right\rangle = 1(0) + 0(2) + 1(0) + 1(0) = 0$$

The generalized Theorem of Pythagoras

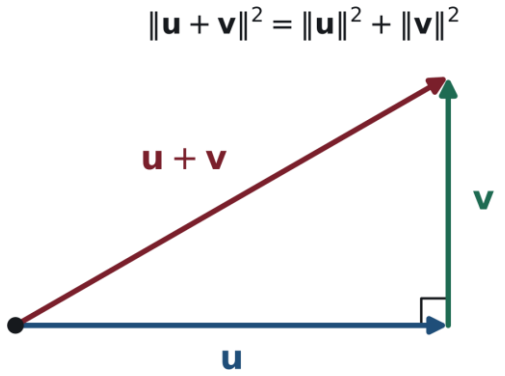
THEOREM 6.2.3 If $\mathbf{u}$ and $\mathbf{v}$ are orthogonal vectors in a real inner product space, then

$$\| \mathbf{u}+\mathbf{v} \|^2 = \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2.$$

Proof

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

since $\langle \mathbf{u},\mathbf{v} \rangle = 0$. ■



The proof shows more: $\| \mathbf{u}+\mathbf{v} \|^2 = \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2$ holds IF AND ONLY IF $\langle \mathbf{u},\mathbf{v} \rangle = 0$, as the old slides state it.

The parallelogram identity

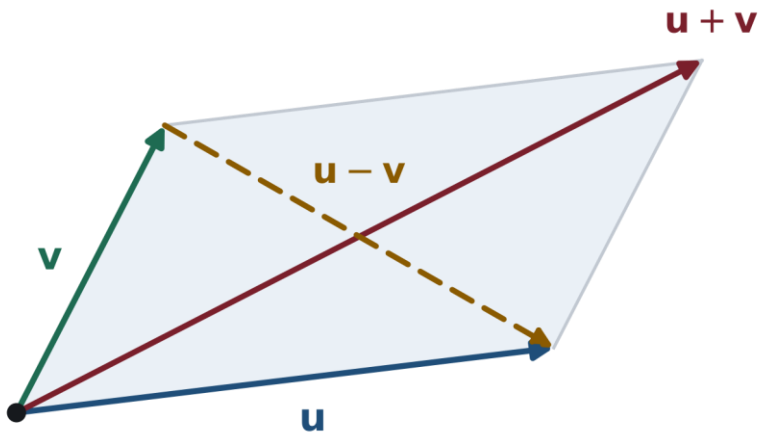
PARALLELOGRAM IDENTITY For all $\mathbf{u}, \mathbf{v}$ in a real inner product space:

$$\| \mathbf{u}+\mathbf{v} \|^2 + \| \mathbf{u}-\mathbf{v} \|^2 = 2 \| \mathbf{u} \|^2 + 2 \| \mathbf{v} \|^2.$$

Proof

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

Add the two lines: the terms $\pm 2\langle \mathbf{u}, \mathbf{v} \rangle$ cancel. ■



In a parallelogram, the squares of the two diagonals add up to the squares of the four sides.

A vector orthogonal to several vectors

> If $\mathbf{w}$ is orthogonal to each of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$, then $\mathbf{w}$ is orthogonal to every vector in $\text{span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$.

Proof: $\langle k_1\mathbf{u}_1 + \cdots + k_r\mathbf{u}_r, \mathbf{w}\rangle = k_1\langle \mathbf{u}_1, \mathbf{w}\rangle + \cdots + k_r\langle \mathbf{u}_r, \mathbf{w}\rangle = 0$ (additivity and homogeneity). ■

Example: vectors of $\mathbb{R}^3$ orthogonal to $\mathbf{u}_1 = (1,0,1)$ and $\mathbf{u}_2 = (0,1,1)$ (Euclidean)

$\mathbf{x} = (x_1, x_2, x_3)$ must satisfy $x_1 + x_3 = 0$ and $x_2 + x_3 = 0$. With $x_3 = t$ free:

$$\{\, t(-1,-1,1) : t \in \mathbb{R} \,\}$$

Each of these vectors is orthogonal to the whole plane $\text{span}\{\mathbf{u}_1, \mathbf{u}_2\}$.

In $\mathbb{R}^2$ (old slides): the vectors orthogonal to $(4,2)$ are $\{(-t/2\,,\, t) : t \in \mathbb{R}\}$, a line through the origin.

SECTION 6.3

Gram–Schmidt Process

GOAL Work with bases whose vectors are mutually orthogonal. Coordinates then come from inner products, and the Gram–Schmidt process turns any basis into an orthonormal one.

1. Orthogonal and orthonormal sets
2. Orthogonal sets are linearly independent
3. Coordinates relative to orthogonal and orthonormal bases
4. Removing a component: the key step
5. The Gram–Schmidt process
6. Gram–Schmidt in $\mathbb{R}^3$, with a weighted inner product, and in P_2

Orthogonal and orthonormal sets

DEFINITION 1 A set of two or more vectors in a real inner product space is ORTHOGONAL if every two distinct vectors of the set are orthogonal. An orthogonal set in which every vector has norm 1 is ORTHONORMAL.

Dividing each vector of an orthogonal set of NONZERO vectors by its norm (normalizing) keeps the set orthogonal and makes it orthonormal.

Examples 1 and 2: Euclidean $\mathbb{R}^3$

$$\mathbf{v}_1 = (0,1,0),\ \ \mathbf{v}_2 = (1,0,1),\ \ \mathbf{v}_3 = (1,0,-1): \qquad \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathbf{v}_3 \rangle = \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0$$

The norms are $\| \mathbf{v}_1 \| = 1$ and $\| \mathbf{v}_2 \| = \| \mathbf{v}_3 \| = \sqrt{2}$, so normalizing gives the orthonormal set

$$\mathbf{u}_1 = (0,1,0), \qquad \mathbf{u}_2 = \left(1/\sqrt{2},\ 0,\ 1/\sqrt{2}\right), \qquad \mathbf{u}_3 = \left(1/\sqrt{2},\ 0,\ -1/\sqrt{2}\right)$$

THEOREM 6.3.1

Orthogonal sets of nonzero vectors are linearly independent

THEOREM 6.3.1 If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is an orthogonal set of NONZERO vectors in an inner product space, then S is linearly independent.

Proof

Suppose $k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_n\mathbf{v}_n = \mathbf{0}$. Take the inner product of both sides with $\mathbf{v}_i$:

$$k_1\langle \mathbf{v}_1, \mathbf{v}_i\rangle + \cdots + k_i\langle \mathbf{v}_i, \mathbf{v}_i\rangle + \cdots + k_n\langle \mathbf{v}_n, \mathbf{v}_i\rangle = \langle \mathbf{0}, \mathbf{v}_i\rangle = 0$$

Every term with $j \neq i$ is 0 by orthogonality, so $k_i\langle \mathbf{v}_i, \mathbf{v}_i\rangle = 0$. Since $\mathbf{v}_i \neq \mathbf{0}$, $\langle \mathbf{v}_i, \mathbf{v}_i\rangle > 0$, hence $k_i = 0$. This holds for every i. ■

"Nonzero" matters: $\{(1,0), (0,0)\}$ is an orthogonal set, but it is linearly dependent.

Every orthonormal set is linearly independent, since its vectors have norm 1.

Orthogonal and orthonormal bases

A basis that is an orthogonal set is an ORTHOGONAL BASIS; a basis that is an orthonormal set is an ORTHONORMAL BASIS. Examples: the standard basis of $\mathbb{R}^n$ (Euclidean inner product), and $\{1, x, \ldots, x^n\}$ in P_n with the standard inner product (Example 3).

> **COROLLARY** In an inner product space of dimension n, every orthogonal set of n nonzero vectors is a basis (Theorem 6.3.1 and Theorem 4.5.4).

Example: a basis for $\mathbb{R}^4$ with no system to solve

$$\mathbf{v}_1 = (2,3,2,-2), \quad \mathbf{v}_2 = (1,0,0,1), \quad \mathbf{v}_3 = (-1,0,2,1), \quad \mathbf{v}_4 = (-1,2,-1,1)$$

All six dot products are 0: $\mathbf{v}_1 \cdot \mathbf{v}_2 = 2 - 2$, $\mathbf{v}_1 \cdot \mathbf{v}_3 = -2 + 4 - 2$, $\mathbf{v}_1 \cdot \mathbf{v}_4 = -2 + 6 - 2 - 2$, $\mathbf{v}_2 \cdot \mathbf{v}_3 = -1 + 1$, $\mathbf{v}_2 \cdot \mathbf{v}_4 = -1 + 1$, $\mathbf{v}_3 \cdot \mathbf{v}_4 = 1 - 2 + 1$.

So these are four nonzero orthogonal vectors in $\mathbb{R}^4$: an orthogonal basis.

THEOREM 6.3.2

Coordinates relative to an orthogonal or orthonormal basis

THEOREM 6.3.2 Let $\mathbf{u}$ be any vector of an inner product space V.

(a) If $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is an ORTHOGONAL basis for V, then $\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{u}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|^2}\mathbf{v}_n$.

(b) If S is an ORTHONORMAL basis, then

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{u}, \mathbf{v}_n \rangle \mathbf{v}_n, \text{ so } (\mathbf{u})_S = (\langle \mathbf{u}, \mathbf{v}_1 \rangle, \dots, \langle \mathbf{u}, \mathbf{v}_n \rangle).$$

Proof of (a)

Write $\mathbf{u} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$. Then $\langle \mathbf{u}, \mathbf{v}_i \rangle = c_1\langle \mathbf{v}_1, \mathbf{v}_i \rangle + \cdots + c_n\langle \mathbf{v}_n, \mathbf{v}_i \rangle = c_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle = c_i \| \mathbf{v}_i \|^2$, because all the other terms are 0. So $c_i = \langle \mathbf{u}, \mathbf{v}_i \rangle / \| \mathbf{v}_i \|^2$. For (b), every $\| \mathbf{v}_i \| = 1$. ■

No linear system to solve: each coordinate is one inner product.

Coordinates from inner products

Example 5 (orthonormal basis)

$\mathbf{v}_1 = (0,1,0)$, $\mathbf{v}_2 = \left(-\frac{4}{5}, 0, \frac{3}{5}\right)$, $\mathbf{v}_3 = \left(\frac{3}{5}, 0, \frac{4}{5}\right)$ form an orthonormal basis S of $\mathbb{R}^3$. For $\mathbf{u} = (1,1,1)$:

$$\langle \mathbf{u}, \mathbf{v}_1 \rangle = 1, \quad \langle \mathbf{u}, \mathbf{v}_2 \rangle = -1/5\,, \quad \langle \mathbf{u}, \mathbf{v}_3 \rangle = 7/5$$

$$\mathbf{u} = \mathbf{v}_1 - 1/5\,\mathbf{v}_2 + 7/5\,\mathbf{v}_3, \qquad (\mathbf{u})_S = (1, -1/5\,, 7/5)$$

Example (old slides)

$B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ with $\mathbf{v}_1 = \left(\frac{3}{5}, \frac{4}{5}, 0\right)$, $\mathbf{v}_2 = \left(-\frac{4}{5}, \frac{3}{5}, 0\right)$, $\mathbf{v}_3 = (0,0,1)$ is an orthonormal basis of $\mathbb{R}^3$. For $\mathbf{w} = (5, -5, 2)$:

$$\begin{aligned} \langle \mathbf{w}, \mathbf{v}_1 \rangle &= 3 - 4 = -1 \\ \langle \mathbf{w}, \mathbf{v}_2 \rangle &= -4 - 3 = -7 \\ \langle \mathbf{w}, \mathbf{v}_3 \rangle &= 2 \end{aligned} \qquad [\mathbf{w}]_B = \begin{bmatrix} -1 \\ -7 \\ 2 \end{bmatrix}$$

Check of Example 5: $(0,1,0) - \frac{1}{5}\left(-\frac{4}{5}, 0, \frac{3}{5}\right) + \frac{7}{5}\left(\frac{3}{5}, 0, \frac{4}{5}\right) = \left(\frac{4}{25} + \frac{21}{25},\ 1,\ -\frac{3}{25} + \frac{28}{25}\right) = (1,1,1).$

EXAMPLE 6

An orthogonal basis: divide by the squared norms

$\mathbf{w}_1 = (0,2,0)$, $\mathbf{w}_2 = (3,0,3)$, $\mathbf{w}_3 = (-4,0,4)$ are orthogonal (all dot products are 0) and nonzero, so they form an orthogonal basis of $\mathbb{R}^3$. Write $\mathbf{u} = (1,2,4)$ in this basis with Theorem 6.3.2(a):

$$\frac{\langle \mathbf{u}, \mathbf{w}_1 \rangle}{\| \mathbf{w}_1 \|^2} = \frac{4}{4} = 1, \qquad \frac{\langle \mathbf{u}, \mathbf{w}_2 \rangle}{\| \mathbf{w}_2 \|^2} = \frac{15}{18} = \frac{5}{6}, \qquad \frac{\langle \mathbf{u}, \mathbf{w}_3 \rangle}{\| \mathbf{w}_3 \|^2} = \frac{12}{32} = \frac{3}{8}$$

$$\mathbf{u} = \mathbf{w}_1 + 5/6\,\mathbf{w}_2 + 3/8\,\mathbf{w}_3 = (0,2,0) + (5/2\,, 0, 5/2) + (-\,3/2\,, 0, 3/2) = (1,2,4)$$

Anton first normalizes: $\mathbf{v}_1 = (0,1,0)$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}(1,0,1)$, $\mathbf{v}_3 = \frac{1}{\sqrt{2}}(-1,0,1)$, and then part (b) gives

$$\mathbf{u} = 2\mathbf{v}_1 + 5/\sqrt{2}\ \mathbf{v}_2 + 3/\sqrt{2}\ \mathbf{v}_3$$

Removing the component along a vector

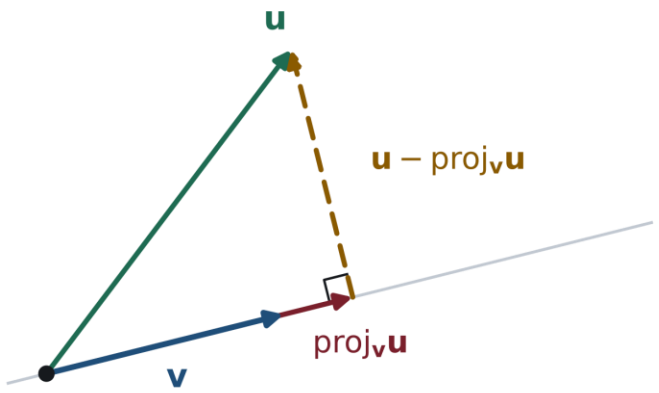
Let $\mathbf{v} \neq \mathbf{0}$. The vector $\operatorname{proj}_{\mathbf{v}}\mathbf{u} = \frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\mathbf{v}\|^2}\,\mathbf{v}$ is the ORTHOGONAL PROJECTION of $\mathbf{u}$ on $\mathbf{v}$, and what is left over is orthogonal to $\mathbf{v}$:

$$\left\langle \mathbf{u} - \frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\,\mathbf{v}\,\|^2}\mathbf{v},\ \mathbf{v} \right\rangle = \langle \mathbf{u},\mathbf{v}\rangle - \frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\,\mathbf{v}\,\|^2}\langle \mathbf{v},\mathbf{v}\rangle = 0$$

More generally, if $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are orthogonal and nonzero, then

$$\mathbf{u} - \frac{\langle \mathbf{u},\mathbf{v}_1\rangle}{\|\,\mathbf{v}_1\,\|^2}\mathbf{v}_1 - \cdots - \frac{\langle \mathbf{u},\mathbf{v}_k\rangle}{\|\,\mathbf{v}_k\,\|^2}\mathbf{v}_k$$

is orthogonal to every $\mathbf{v}_i$: the same computation, one i at a time, since $\langle \mathbf{v}_j,\mathbf{v}_i\rangle = 0$ for $j \neq i$.



This is the step that the Gram–Schmidt process repeats.

The Gram–Schmidt process

THEOREM 6.3.5 Every nonzero finite-dimensional inner product space has an orthonormal basis.

GRAM–SCHMIDT To turn a basis $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ into an orthogonal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$:

1) $\mathbf{v}_1 = \mathbf{u}_1$

2) $\mathbf{v}_2 = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1$

3) $\mathbf{v}_3 = \mathbf{u}_3 - \frac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 - \frac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2} \mathbf{v}_2$

4) Continue: $\mathbf{v}_k$ is $\mathbf{u}_k$ minus its components along $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$.

For an orthonormal basis $\{\mathbf{q}_1, \dots, \mathbf{q}_r\}$, normalize at the end: $\mathbf{q}_k = \mathbf{v}_k / \| \mathbf{v}_k \|$.

EXAMPLE 8

Gram–Schmidt in $\mathbb{R}^3$

Euclidean $\mathbb{R}^3$, basis $\mathbf{u}_1 = (1,1,1)$, $\mathbf{u}_2 = (0,1,1)$, $\mathbf{u}_3 = (0,0,1)$.

$$\mathbf{v}_1 = \mathbf{u}_1 = (1,1,1)$$

$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\| \mathbf{v}_1 \|^2}\mathbf{v}_1 = (0,1,1) - \frac{2}{3}(1,1,1) = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

$$\mathbf{v}_3 = \mathbf{u}_3 - \frac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\| \mathbf{v}_1 \|^2}\mathbf{v}_1 - \frac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\| \mathbf{v}_2 \|^2}\mathbf{v}_2 = (0,0,1) - \frac{1}{3}(1,1,1) - \frac{1/3}{2/3}\left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right) = \left(0, -\frac{1}{2}, \frac{1}{2}\right)$$

Normalize, with $\| \mathbf{v}_1 \| = \sqrt{3}$, $\| \mathbf{v}_2 \| = \frac{\sqrt{6}}{3}$, $\| \mathbf{v}_3 \| = \frac{1}{\sqrt{2}}$:

$$\mathbf{q}_1 = 1/\sqrt{3}\,(1,1,1), \qquad \mathbf{q}_2 = 1/\sqrt{6}\,(-2,1,1), \qquad \mathbf{q}_3 = 1/\sqrt{2}\,(0,-1,1)$$

EXAMPLE

Gram–Schmidt with scaling

Euclidean $\mathbb{R}^3$, basis $\mathbf{u}_1 = (1,1,0)$, $\mathbf{u}_2 = (1,2,0)$, $\mathbf{u}_3 = (0,1,2)$.

$$\mathbf{v}_1 = (1,1,0), \qquad \mathbf{v}_2 = (1,2,0) - \frac{3}{2}(1,1,0) = \left(-\frac{1}{2}, \frac{1}{2}, 0\right)$$

A nonzero multiple of $\mathbf{v}_2$ is just as good: it is still orthogonal to $\mathbf{v}_1$. Use $2\mathbf{v}_2 = (-1,1,0)$ to avoid fractions:

$$\mathbf{v}_3 = (0,1,2) - \frac{1}{2}(1,1,0) - \frac{1}{2}(-1,1,0) = (0,0,2)$$

$$\mathbf{q}_1 = 1/\sqrt{2}\,(1,1,0), \qquad \mathbf{q}_2 = 1/\sqrt{2}\,(-1,1,0), \qquad \mathbf{q}_3 = (0,0,1)$$

EXAMPLE

Gram–Schmidt with a weighted inner product

Let $\mathbb{R}^2$ have $\langle \mathbf{u}, \mathbf{v} \rangle = 2u_1v_1 + u_2v_2$, and take the basis $\mathbf{u}_1 = (1,1)$, $\mathbf{u}_2 = (0,1)$.

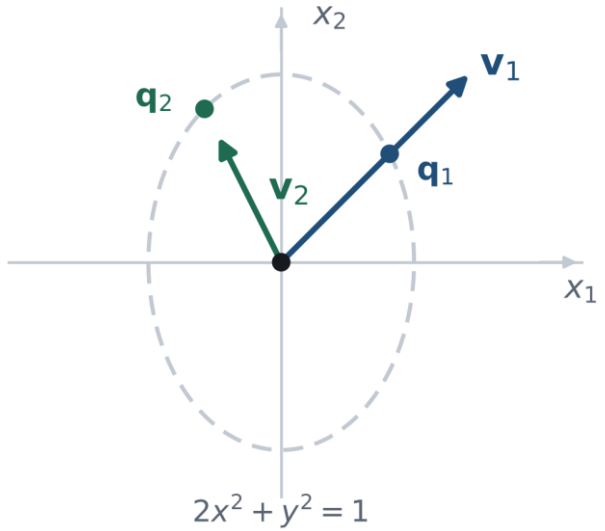
$$\mathbf{v}_1 = (1,1), \qquad \| \mathbf{v}_1 \|^2 = \langle \mathbf{v}_1, \mathbf{v}_1 \rangle = 2(1)(1) + (1)(1) = 3$$

$$\langle \mathbf{u}_2, \mathbf{v}_1 \rangle = 2(0)(1) + (1)(1) = 1$$

$$\mathbf{v}_2 = (0,1) - 1/3\,(1,1) = (-1/3\,, 2/3)$$

Check: $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 2(1)\left(-\frac{1}{3}\right) + (1)\left(\frac{2}{3}\right) = 0$, and $\| \mathbf{v}_2 \|^2 = \langle \mathbf{v}_2, \mathbf{v}_2 \rangle = 2 \cdot \frac{1}{9} + \frac{4}{9} = \frac{2}{3}$.

$$\mathbf{q}_1 = 1/\sqrt{3}\,(1,1), \qquad \mathbf{q}_2 = 1/\sqrt{6}\,(-1,2)$$



$\mathbf{v}_1 \cdot \mathbf{v}_2 = -\frac{1}{3} + \frac{2}{3} = \frac{1}{3} \neq 0$: on the page they are not perpendicular. Orthogonal always means orthogonal for the inner product in use.

SUMMARY

Unit 4 in one slide

3.1 – 3.2 Vectors in $\mathbb{R}^n$

- $\mathbf{u}\cdot\mathbf{v} = u_1v_1 + \cdots + u_nv_n$, $\| \mathbf{v} \| = \sqrt{\mathbf{v}\cdot\mathbf{v}}$, $d(\mathbf{u},\mathbf{v}) = \| \mathbf{u}-\mathbf{v} \|$; the dot product is the model for inner products.

6.1 Inner products

- An inner product satisfies symmetry, additivity, homogeneity and positivity, and gives the norm $\| \mathbf{v} \| = \sqrt{\langle \mathbf{v},\mathbf{v}\rangle}$ and the distance $d(\mathbf{u},\mathbf{v}) = \| \mathbf{u}-\mathbf{v} \|$.
- Examples: weighted Euclidean (positive weights), generated by an invertible matrix, $\mathrm{tr}(U^TV)$ on M_{nn}

6.2 Angle and orthogonality

- Cauchy–Schwarz $|\langle \mathbf{u},\mathbf{v}\rangle| \le \| \mathbf{u} \| \| \mathbf{v} \|$ gives the angle, $\cos\theta = \langle \mathbf{u},\mathbf{v}\rangle / (\| \mathbf{u} \| \| \mathbf{v} \|)$, and the triangle inequality; $\mathbf{u},\mathbf{v}$ are orthogonal exactly when $\langle \mathbf{u},\mathbf{v}\rangle = 0$, and then $\| \mathbf{u}+\mathbf{v} \|^2 = \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2$.

6.3 Orthonormal bases and Gram–Schmidt

- Orthogonal sets of nonzero vectors are linearly independent; relative to an orthogonal basis, the coordinates are $\langle \mathbf{u},\mathbf{v}_i\rangle / \| \mathbf{v}_i \|^2$.
- Gram–Schmidt: $\mathbf{v}_k$ is $\mathbf{u}_k$ minus its components along $\mathbf{v}_1,\ldots,\mathbf{v}_{k-1}$; normalize at the end.