Professor Jennifer Balakrishnan, *jbala@bu.edu*

What is on today

1 Inner product, length, and orthogonality

Lay–Lay–McDonald §6.1 pp. 332 – 338

Today we explore length, distance, and perpendicularity for vectors in $\mathbb{R}^n$. All three ideas are defined in terms of the *inner product* of two vectors.

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, then we can think of them as $n \times 1$ matrices. The transpose $\mathbf{u}^T$ is a $1 \times n$ matrix, and the matrix product $\mathbf{u}^T\mathbf{v}$ is a 1×1 matrix, which is a scalar. This scalar is called the *inner product* of $\mathbf{u}$ and $\mathbf{v}$ and is often written as $\mathbf{u} \cdot \mathbf{v}$ and called the *dot product*. If $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$, then the inner product of $\mathbf{u}$ and $\mathbf{v}$ is

$$[u_1 \quad \cdots \quad u_n] \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + \cdots + u_n v_n.$$

Example 1. *Let* $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$. *Compute* $\mathbf{u} \cdot \mathbf{v}$ *and* $\mathbf{v} \cdot \mathbf{u}$.

$u \cdot v = (2)(3) + (-5)(2) + (-1)(-3) = 6 - 10 + 3 = -1$

$v \cdot u = (3)(2) + (2)(-5) + (-3)(-1) = 6 - 10 + 3 = -1$

Here are properties of the inner product:

Theorem 2. *Let* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *be vectors in* $\mathbb{R}^n$, *and let* c *be a scalar. Then*

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
3. $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$
4. $\mathbf{u} \cdot \mathbf{u} \geq 0$ *and* $\mathbf{u} \cdot \mathbf{u} = 0$ *if and only if* $\mathbf{u} = \mathbf{0}$.

(if $u = \vec{0} = (0, 0, \ldots, 0)$, then $u \cdot u$

$= 0 \cdot 0 + 0 \cdot 0 + \cdots + 0 \cdot 0 = 0$.

Suppose $\underset{(u_1, u_2, \ldots, u_n)}{u} \cdot u = 0$

$u_1^2 + u_2^2 + \cdots + u_n^2 = 0 \Rightarrow u_i = 0$ for all i)

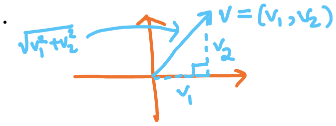
Definition 3. *Let* $\mathbf{v}$ *be a vector in* $\mathbb{R}^n$ *with entries* $v_1, \ldots, v_n$. *The length (or norm) of* $\mathbf{v}$ *is the nonnegative scalar* $||\mathbf{v}||$ *defined by*

$$||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + \cdots + v_n^2},$$

$v = (v_1, \ldots, v_n)$

and $||\mathbf{v}||^2 = \mathbf{v} \cdot \mathbf{v}$.

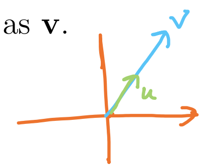
Note that if $\mathbf{v} \in \mathbb{R}^2$, then $||v||$ coincides with the standard notion of the length of the line segment from the origin to $\mathbf{v}$ by the Pythagorean Theorem.



For any scalar c, we have

$$||c\mathbf{v}|| = |c|||\mathbf{v}||.$$

A vector whose length is 1 is called a *unit vector*. If we divide a nonzero vector $\mathbf{v}$ by its length, we obtain a unit vector $\mathbf{u}$ because the length of $\mathbf{u}$ is $\left(\frac{1}{||\mathbf{v}||}\right)||\mathbf{v}||$. The process of creating $\mathbf{u}$ from $\mathbf{v}$ is called *normalizing* $\mathbf{v}$ and we say that $\mathbf{u}$ is in the same direction as $\mathbf{v}$.



Example 4. *Let* $\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix}$. *Find a unit vector* $\mathbf{u}$ *in the same direction as* $\mathbf{v}$.

$||v|| = \sqrt{v \cdot v} = \sqrt{1+4+4} = 3$

$u = \frac{v}{||v||} = \frac{1}{3} \cdot \begin{pmatrix} 1 \\ -2 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/3 \\ -2/3 \\ 2/3 \\ 0 \end{pmatrix}$

Example 5. *Let* W *be the subspace of* $\mathbb{R}^2$ *spanned by* $\mathbf{x} = \begin{bmatrix} \frac{2}{3} \\ 1 \end{bmatrix}$. *Find a unit vector* $\mathbf{z}$ *that is a basis for* W.

$||x|| = \sqrt{x \cdot x} = \sqrt{(\frac{2}{3})^2 + 1^2} = \sqrt{\frac{4}{9}+1} = \frac{\sqrt{13}}{3}$; $z = \frac{x}{||x||} = \frac{3}{\sqrt{13}} \begin{pmatrix} 2/3 \\ 1 \end{pmatrix} = \begin{pmatrix} 2/\sqrt{13} \\ 3/\sqrt{13} \end{pmatrix}$.

Recall that if a, b are real numbers, the distance on the number line between a and b is given by the absolute value $|a - b|$. This definition of distance in $\mathbb{R}$ has a direct analogue in $\mathbb{R}^n$.

Definition 6. *For* $\mathbf{u}$ *and* $\mathbf{v}$ *in* $\mathbb{R}^n$, *the distance between* $\mathbf{u}$ *and* $\mathbf{v}$, *written as dist*$(\mathbf{u}, \mathbf{v})$, *is the length of the vector* $\mathbf{u} - \mathbf{v}$. *That is, dist*$(\mathbf{u}, \mathbf{v}) = ||\mathbf{u} - \mathbf{v}||$.

Example 7. *Compute the distance between the vectors* $\mathbf{u} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$.

$||u - v|| = \left|\left| \begin{bmatrix} 4 \\ -1 \end{bmatrix} \right|\right| = \sqrt{4^2 + (-1)^2} = \sqrt{17}$.

Definition 8. *Two vectors* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ *are orthogonal if* $\mathbf{u} \cdot \mathbf{v} = \mathbf{0}$.

Note that the zero vector is orthogonal to every vector in $\mathbb{R}^n$.

Here is a useful result about orthogonality:

Theorem 9 (Pythagorean Theorem). *Two vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *are orthogonal if and only if* $||\mathbf{u}+\mathbf{v}||^2 = ||\mathbf{u}||^2 + ||\mathbf{v}||^2$.

If a vector $\mathbf{z}$ is orthogonal to every vector in a subspace W of $\mathbb{R}^n$, then $\mathbf{z}$ is said to be orthogonal to W. The set of all vectors $\mathbf{z}$ that are orthogonal to W is called the *orthogonal complement* of W and is denoted by $W^\perp$.

Let W be a subspace of $\mathbb{R}^n$. Here are two facts about orthogonal complements.

1. A vector $\mathbf{x}$ is in $W^\perp$ iff $\mathbf{x}$ is orthogonal to every vector in a set that spans W.

2. $W^\perp$ is a subspace of $\mathbb{R}^n$.

Here is another relationship between the null space and column space of a matrix.

Theorem 10. *Let A be an $m \times n$ matrix. The orthogonal complement of the column space of A is the null space of A^T:* $(\mathrm{Col}\,A)^\perp = \mathrm{Nul}\,A^T$.

2 Orthogonal sets

Lay–Lay–McDonald §6.2 pp. 340 – 342

A set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be an *orthogonal set* if each pair of distinct vectors from the set is orthogonal.

Example 11. *Show that* $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ *is an orthogonal set, where* $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} -\frac{1}{2} \\ -2 \\ \frac{7}{2} \end{bmatrix}$.

Show: $\left\{ \begin{array}{l} u_1 \cdot u_2 = 0 \\ u_1 \cdot u_3 = 0 \\ u_2 \cdot u_3 = 0 \end{array} \right.$

$u_1 \cdot u_2 = 3(-1) + 1(2) + 1(1) = -3 + 2 + 1 = 0$

$u_1 \cdot u_3 = -3/2 - 2 + 7/2 = 0$

$u_2 \cdot u_3 = 1/2 - 4 + 7/2 = 0$ ✓

Theorem 12. *If* $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ *is an orthogonal set of nonzero vectors in* $\mathbb{R}^n$*, then S is linearly independent and hence is a basis for the subspace spanned by S.*

Proof. If $\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ for some scalars $c_1, \ldots, c_p$, then

$$\begin{aligned} 0 = \mathbf{0} \cdot \mathbf{u}_1 &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p) \cdot \mathbf{u}_1 \\ &= (c_1\mathbf{u}_1) \cdot \mathbf{u}_1 + (c_2\mathbf{u}_2) \cdot \mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p) \cdot \mathbf{u}_1 \\ &= c_1(\mathbf{u}_1 \cdot \mathbf{u}_1) + c_2(\cancel{\mathbf{u}_2 \cdot \mathbf{u}_1}) + \cdots + c_p(\cancel{\mathbf{u}_p \cdot \mathbf{u}_1}) \\ &= c_1(\mathbf{u}_1 \cdot \mathbf{u}_1) \end{aligned}$$

because $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2, \ldots, \mathbf{u}_p$. Since $\mathbf{u}_1$ is nonzero, $\mathbf{u}_1 \cdot \mathbf{u}_1$ is not zero, and so $c_1 = 0$. Similarly, $c_2, \ldots, c_p$ must be zero. Thus S is linearly independent. □

Repeat the whole process w/ $u_2, u_3 \ldots$

Definition 13. *An orthogonal basis for a subspace W of $\mathbb{R}^n$ is a basis for W that is also an orthogonal set.*

The next theorem tells us why an orthogonal basis is nicer than other bases. The weights in a linear combination can be computed easily.

Theorem 14. *Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\mathbf{y}$ in W, the weights in the linear combination*

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

are given by

$$c_j = \frac{\mathbf{y}\cdot\mathbf{u}_j}{\mathbf{u}_j\cdot\mathbf{u}_j}, \quad j = 1, \ldots, p.$$

Example 15. *The set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ in the previous example is an orthogonal basis for $\mathbb{R}^3$. Express the vector $\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ as a linear combination of the vectors in S.*

$u_1 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, u_3 = \begin{pmatrix} -1/2 \\ -2 \\ 7/2 \end{pmatrix}, y = \begin{pmatrix} 6 \\ 1 \\ -8 \end{pmatrix}$

$c_3 = \frac{y\cdot u_3}{u_3\cdot u_3} = \frac{-3-2-28}{\frac{1}{4}+4+\frac{49}{4}} = \frac{-33}{66/4} = -2$

$c_1 = \frac{y\cdot u_1}{u_1\cdot u_1} = \frac{18+1-8}{9+1+1} = \frac{11}{11} = 1$; $c_2 = \frac{y\cdot u_2}{u_2\cdot u_2} = \frac{-6+2-8}{1+4+1} = \frac{-12}{6} = -2$

$y = u_1 - 2u_2 - 2u_3$

Given a nonzero vector $\mathbf{u}$ in $\mathbb{R}^n$, we consider the problem of decomposing a vector $\mathbf{y} \in \mathbb{R}^n$ into the sum of two vectors, one a multiple of $\mathbf{u}$ and the other orthogonal to $\mathbf{u}$. That is, we want to write

$$\mathbf{y} = \underset{\alpha\cdot u}{\hat{\mathbf{y}}} + \mathbf{z}, \qquad (1)$$

$z\cdot u = 0$

where $\hat{\mathbf{y}} = \alpha\mathbf{u}$ for some scalar α and $\mathbf{z}$ is orthogonal to $\mathbf{u}$. Equation (1) is satisfied under these constraints if and only if $\alpha = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}$ and

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}.$$

The vector $\hat{\mathbf{y}}$ is called the *orthogonal projection of* $\mathbf{y}$ *onto* $\mathbf{u}$ and the vector $\mathbf{z}$ is called the *component of* $\mathbf{y}$ *orthogonal to* $\mathbf{u}$.

Example 16. *Let $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. Find the orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$. Then write $\mathbf{y}$ as the sum of two orthogonal vectors, one in Span$\{\mathbf{u}\}$ and one orthogonal to $\mathbf{u}$.*

$\hat{y} = \frac{y\cdot u}{u\cdot u}u = \frac{7\cdot 4+6\cdot 2}{16+4}\begin{pmatrix} 4 \\ 2 \end{pmatrix} = \frac{40}{20}\begin{pmatrix} 4 \\ 2 \end{pmatrix} = \underline{\begin{pmatrix} 8 \\ 4 \end{pmatrix}} \in \text{Span}\{u\}$

$\begin{pmatrix} 4 \\ 2 \end{pmatrix}$

orthog. proj. of y onto u

$\begin{pmatrix} 7 \\ 6 \end{pmatrix} = \overset{\hat{y}=}{\begin{pmatrix} 8 \\ 4 \end{pmatrix}} + \underline{z}$ ← by construction z is orthog. to u

$\begin{pmatrix} 7 \\ 6 \end{pmatrix} - \begin{pmatrix} 8 \\ 4 \end{pmatrix} = z$

$\begin{pmatrix} -1 \\ 2 \end{pmatrix} = z$

orthog. proj of y onto u ; ⊥ to u

$\boxed{\begin{pmatrix} 7 \\ 6 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \end{pmatrix}}$

Span$\left\{\begin{pmatrix} 4 \\ 2 \end{pmatrix}\right\}$

z is this ⊥ to $\begin{pmatrix} 4 \\ 2 \end{pmatrix}$

$-4+4=0$ ✓