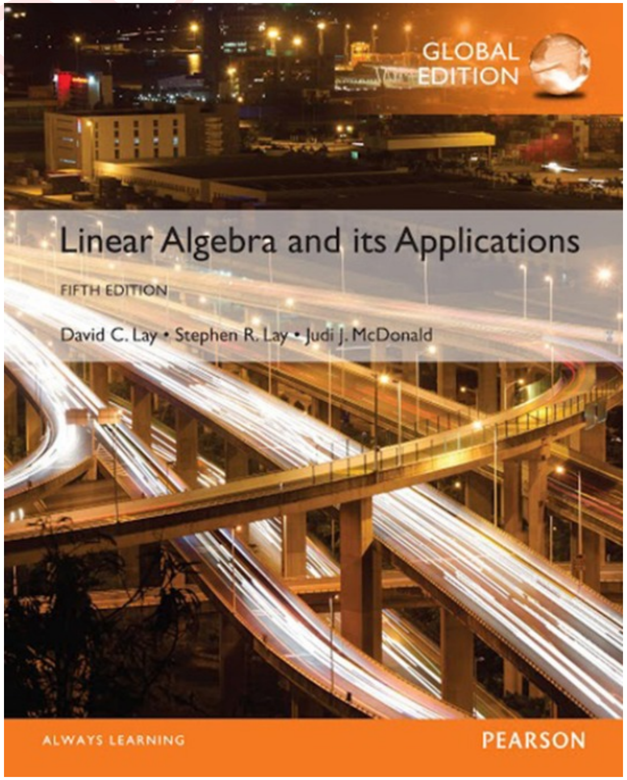
6 Orthogonality and Least Squares

6.2

ORTHOGONAL SETS



ORTHOGONAL SETS

- 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, that is, if $u_i \cdot u_j = 0$ whenever $i \neq j$.

- **Theorem 4:** If $S = \{\mathrm{u}_1,\ldots,\mathrm{u}_p\}$ 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.

ORTHOGONAL SETS

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

$$0 = 0 \cdot u_1 = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p) \cdot u_1$$

$$= (c_1\mathbf{u}_1) \cdot u_1 + (c_2\mathbf{u}_2) \cdot u_1 + \cdots + (c_p\mathbf{u}_p) \cdot u_1$$

$$= c_1(\mathbf{u}_1 \cdot u_1) + c_2(\mathbf{u}_2 \cdot u_1) + \cdots + c_p(\mathbf{u}_p \cdot u_1)$$

$$= c_1(\mathbf{u}_1 \cdot u_1)$$

because $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2,\ldots,\mathbf{u}_p$.

- Since $\mathbf{u}_1$ is nonzero, $u_1 \cdot u_1$ is not zero and so $c_1 = 0$
- Similarly, $c_2,\ldots,c_p$ must be zero.

ORTHOGONAL SETS

- Thus S is linearly independent.
- **Definition:** An **orthogonal basis** for a subspace W of $\mathbb{R}^n$ is a basis for W that is also an orthogonal set.

- **Theorem 5:** 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

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

are given by

$$c_j = \frac{y \cdot u_j}{u_j \cdot u_j} \qquad (j = 1,\ldots,p)$$

ORTHOGONAL SETS

- **Proof:** The orthogonality of $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ shows that

$$y \cdot u_1 = (c_1\mathrm{u}_1 + c_2\mathrm{u}_2 + \cdots + c_p\mathrm{u}_p) \cdot u_1 = c_1\, u_1 \cdot u_1$$

- Since $u_1 \cdot u_1$ is not zero, the equation above can be solved for c_1.

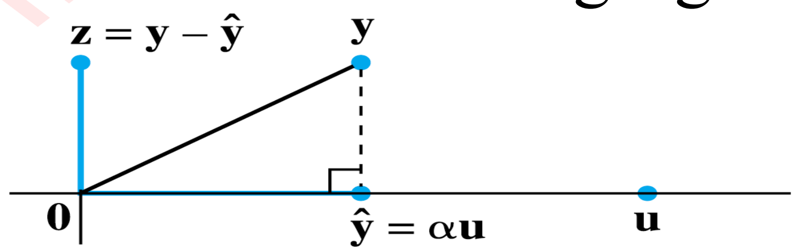
- To find c_j for $j = 2,\ldots, p$, compute $y \cdot u_j$ and solve for c_j.

AN ORTHOGONAL PROJECTION

- Given a nonzero vector **u**in $\mathbb{R}^n$, consider the problem of decomposing a vector **y**in $\mathbb{R}^n$into the sum of two vectors, one a multiple of **u** and the other orthogonal to **u**.
- We wish to write

(1) $$\text{y} = \hat{\text{y}} + \text{z}$$

where $\hat{\text{y}} = \alpha \text{u}$ for some scalar α and **z** is some vector orthogonal to **u**. See the following figure.



Finding α to make $\mathbf{y} - \hat{\mathbf{y}}$ orthogonal to $\mathbf{u}$.

AN ORTHOGONAL PROJECTION

- Given any scalar α, let $z = y - \alpha u$, so that (1) is satisfied.
- Then $y - \hat{y}$ is orthogonal to **u** if an only if

  $$0 = (y - \alpha u) \cdot u = y \cdot u - (\alpha u) \cdot u = y \cdot u - \alpha(u \cdot u)$$

- That is, (1) is satisfied with **z** orthogonal to **u** if and

  only if $\alpha = \dfrac{y \cdot u}{u \cdot u}$ and $\hat{y} = \dfrac{y \cdot u}{u \cdot u} u$.

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

AN ORTHOGONAL PROJECTION

- If c is any nonzero scalar and if $\mathbf{u}$ is replaced by $c\mathbf{u}$ in the definition of $\hat{\mathbf{y}}$, then the orthogonal projection of $\mathbf{y}$ onto $c\mathbf{u}$ is exactly the same as the orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$.
- Hence this projection is determined by the *subspace L* spanned by $\mathbf{u}$ (the line through $\mathbf{u}$ and $\mathbf{0}$).
- Sometimes $\hat{\mathbf{y}}$ is denoted by $\text{proj}_L \mathbf{y}$ and is called the **orthogonal projection of y onto** L.
- That is,

$$\hat{\mathbf{y}} = \text{proj}_L \mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u} \qquad (2)$$

AN ORTHOGONAL PROJECTION

- **Example 3:** Let $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. Find the orthogonal projection of **y** onto **u**. Then write **y** as the sum of two orthogonal vectors, one in Span$\{\mathbf{u}\}$ and one orthogonal to **u**.
- **Solution:** Compute

$$y \cdot u = \begin{bmatrix} 7 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40$$

$$\mathrm{u} \cdot u = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20$$

AN ORTHOGONAL PROJECTION

- The orthogonal projection of **y** onto **u** is

$$\hat{\mathrm{y}} = \frac{y \cdot u}{u \cdot u}\mathrm{u} = \frac{40}{20}\mathrm{u} = 2\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

and the component of **y** orthogonal to **u** is

$$\mathrm{y} - \hat{\mathrm{y}} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

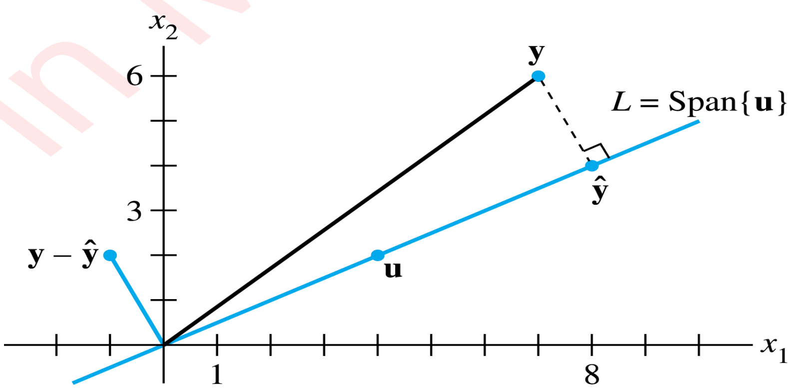
- The sum of these two vectors is **y**.

AN ORTHOGONAL PROJECTION

- That is,

$$\underset{\mathbf{y}}{\begin{bmatrix} 7 \\ 6 \end{bmatrix}} = \underset{\hat{\mathbf{y}}}{\begin{bmatrix} 8 \\ 4 \end{bmatrix}} + \underset{(\mathbf{y}-\hat{\mathbf{y}})}{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}$$

- The decomposition of **y** is illustrated in the following figure:



The orthogonal projection of **y** onto a line L through the origin.

AN ORTHOGONAL PROJECTION

- *Note:* If the calculations above are correct, then $\{\hat{\mathrm{y}}, \mathrm{y} - \hat{\mathrm{y}}\}$ will be an orthogonal set.

- As a check, compute

$$\hat{y} \cdot (y - \hat{y}) = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} = -8 + 8 = 0$$

- Since the line segment in the figure on the previous slide between **y**and $\hat{y}$ is perpendicular to L, by construction of $\hat{y}$, the point identified with $\hat{y}$is the closest point of L to **y**.

ORTHONORMAL SETS

- A set $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ is an **orthonormal set** if it is an orthogonal set of unit vectors.

- If W is the subspace spanned by such a set, then $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ is an **orthonormal basis** for W, since the set is automatically linearly independent, by Theorem 4.

- The simplest example of an orthonormal set is the standard basis $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ for $\mathbb{R}^n$.

- Any nonempty subset of $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ is orthonormal, too.

ORTHONORMAL SETS

- **Example 2:** Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal basis of $\mathbb{R}^3$, where

$$\mathrm{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \mathrm{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \mathrm{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

- **Solution:** Compute

$$v_1 \cdot v_2 = -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0$$

$$v_1 \cdot v_3 = -3/\sqrt{726} - 4/\sqrt{726} + 7/\sqrt{726} = 0$$

ORTHONORMAL SETS

$$v_2 \cdot v_3 = 1/\sqrt{396} - 8/\sqrt{396} + 7/\sqrt{396} = 0$$

- Thus $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set.
- Also, $v_1 \cdot v_1 = 9/11 + 1/11 + 1/11 = 1$

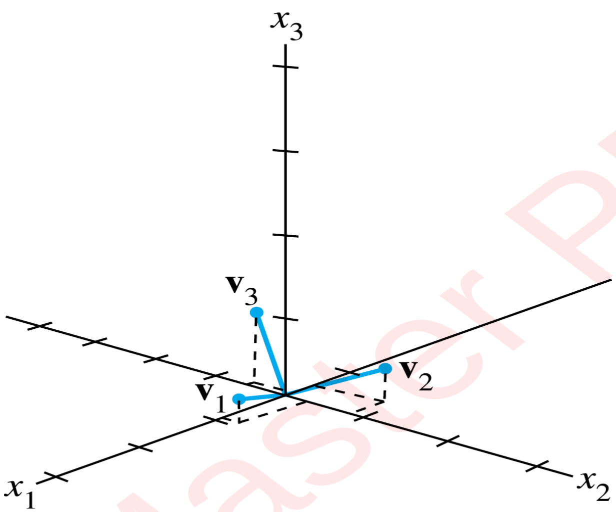
$$v_2 \cdot v_2 = 1/6 + 4/6 + 1/6 = 1$$

$$v_3 \cdot v_3 = 1/66 + 16/66 + 49/66 = 1$$

which shows that $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are unit vectors.

- Thus $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal set.
- Since the set is linearly independent, its three vectors form a basis for . See the figure on the next slide.

ORTHONORMAL SETS



- When the vectors in an orthogonal set of nonzero vectors are *normalized* to have unit length, the new vectors will still be orthogonal, and hence the new set will be an orthonormal set.

ORTHONORMAL SETS

- **Theorem 6:** An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

- **Proof:** To simplify notation, we suppose that U has only three columns, each a vector in $\mathbb{R}^m$.
- Let $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix}$ and compute

$$U^T U = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \mathbf{u}_1^T\mathbf{u}_3 \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \mathbf{u}_2^T\mathbf{u}_3 \\ \mathbf{u}_3^T\mathbf{u}_1 & \mathbf{u}_3^T\mathbf{u}_2 & \mathbf{u}_3^T\mathbf{u}_3 \end{bmatrix}$$

(4)

ORTHONORMAL SETS

- The entries in the matrix at the right are inner products, using transpose notation.
- The columns of U are orthogonal if and only if

$$u_1^T u_2 = u_2^T u_1 = 0, u_1^T u_3 = u_3^T u_1 = 0, u_2^T u_3 = u_3^T u_2 = 0 \tag{5}$$

- The columns of U all have unit length if and only if

$$u_1^T u_1 = 1,\ u_2^T u_2 = 1,\ u_3^T u_3 = 1 \tag{6}$$

- The theorem follows immediately from (4)–(6).

ORTHONORMAL SETS

- **Theorem 7:** Let U be an $m \times n$ matrix with orthonormal columns, and let **x** and **y** be in $\mathbb{R}^n$.

  Then

  $$\|U\mathrm{x}\| = \|\mathrm{x}\|$$

  $$(Ux) \cdot (Uy) = x \cdot y$$

  a. $\underbrace{(Ux) \cdot (Uy)}_{=0}$ if and only if $x \cdot y$=0

- Properties (a) and (c) say that the linear mapping $\mathrm{x} \mapsto U\mathrm{x}$ preserves lengths and orthogonality.