Elementary Linear Algebra

Anton & Rorres, 9th Edition

Lecture Set – 06

Chapter 6:

Inner Product Spaces

Chapter Content

- Inner Products
- Angle and Orthogonality in Inner Product Spaces
- Orthonormal Bases; Gram-Schmidt Process; QR-Decomposition
- Best Approximation; Least Squares
- Orthogonal Matrices; Change of Basis

6-1 Inner Product Space

- An inner product on a real vector space V
  - a function that associates a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ with each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V in such a way that the following axioms are satisfied for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and all scalars k.
  - $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
  - $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
  - $\langle k\mathbf{u}, \mathbf{v} \rangle = k \langle \mathbf{u}, \mathbf{v} \rangle$
  - $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$

  A real vector space with an inner product is called a real inner product space.

6-1 Example

- Euclidean Inner Product on R^n
  - If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_\mathrm{n})$ are vectors in R^n, then the formula

$$\langle \mathbf{v}, \mathbf{u} \rangle = \mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$$

  defines $\langle \mathbf{v}, \mathbf{u} \rangle$ to be the Euclidean product on R^n.

- The four inner product axioms hold by Theorem 4.1.2.

6-1 Preview of Inner Product

- Theorem 4.1.2
  - If $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are vectors in R^n and k is any scalar, then
    - $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
    - $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
    - $(k\,\mathbf{u}) \cdot \mathbf{v} = k\,(\mathbf{u} \cdot \mathbf{v})$
    - $\mathbf{v} \cdot \mathbf{v} \geq \mathbf{0}$; Further, $\mathbf{v} \cdot \mathbf{v} = \mathbf{0}$ if and only if $\mathbf{v} = 0$

- Example
  - $(3\mathbf{u} + 2\mathbf{v}) \cdot (4\mathbf{u} + \mathbf{v})$
    $= (3\mathbf{u}) \cdot (4\mathbf{u} + \mathbf{v}) + (2\mathbf{v}) \cdot (4\mathbf{u} + \mathbf{v})$
    $= (3\mathbf{u}) \cdot (4\mathbf{u}) + (3\mathbf{u}) \cdot \mathbf{v} + (2\mathbf{v}) \cdot (4\mathbf{u}) + (2\mathbf{v}) \cdot \mathbf{v}$
    $=12(\mathbf{u} \cdot \mathbf{u}) + 11(\mathbf{u} \cdot \mathbf{v}) + 2(\mathbf{v} \cdot \mathbf{v})$

6-1 Weighted Euclidean Inner Product

- If $w_1, w_2, \ldots, w_n$ are positive real numbers
  - We call weights
- If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are vectors in R^n, then the formula

  $$\langle \mathbf{v}, \mathbf{u} \rangle = \mathbf{u} \cdot \mathbf{v} = w_1 u_1 v_1 + w_2 u_2 v_2 + \ldots + w_n u_n v_n$$

  is called the weighted Euclidean inner product with weights $w_1, w_2, \ldots, w_n$.

6-1 Example 2

- Let $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ be vectors in R^2. Verify that the weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 2u_2v_2$ satisfies the four product axioms.

- Solution:

  <1> $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.

  <2> If $\mathbf{w} = (w_1, w_2)$, then $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = (3u_1w_1 + 2u_2w_2) + (3v_1w_1 + 2v_2w_2) = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$

  <3> $\langle k\mathbf{u}, \mathbf{v} \rangle = 3(ku_1)v_1 + 2(ku_2)v_2 = k(3u_1v_1 + 2u_2v_2) = k \langle \mathbf{u}, \mathbf{v} \rangle$

  <4> $\langle \mathbf{v}, \mathbf{v} \rangle = 3v_1v_1 + 2v_2v_2 = 3v_1^2 + 2v_2^2$ .Obviously , $\langle \mathbf{v}, \mathbf{v} \rangle = 3v_1^2 + 2v_2^2 \geq 0$. Furthermore, $\langle \mathbf{v}, \mathbf{v} \rangle = 3v_1^2 + 2v_2^2 = 0$ if and only if $v_1 = v_2 = 0$, That is , if and only if $\mathbf{v} = (v_1, v_2) = 0$.

6-1 Norm & Length

- If V is an inner product space, then the norm (or length) of a vector $\mathbf{u}$ in V is denoted by $\|\mathbf{u}\|$ and is defined by

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2}$$

- The distance between two points (vectors) $\mathbf{u}$ and $\mathbf{v}$ is denoted by $d(\mathbf{u},\mathbf{v})$ and is defined by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

6-1 Example 3

- Norm and Distance in R^n
  - If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are vectors in R^n with the Euclidean inner product, then

$$\|\mathrm{u}\| = <\mathbf{u},\ \mathbf{u}>^{1/2} = (\mathbf{u},\ \mathbf{u})^{1/2} = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$$

$$d(\mathbf{u},\ \mathbf{v}) = \|\mathbf{u}-\mathbf{v}\| = <\mathbf{u}-\mathbf{v},\ \mathbf{u}-\mathbf{v}>^{1/2} = \left[(\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v})\right]^{1/2}$$
$$= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \ldots + (u_n - v_n)^2}$$

6-1 Example 4
(Weighted Euclidean Inner Product)

- The norm and distance depend on the inner product used.
  - For example, for the vectors $\mathbf{u} = (1,0)$ and $\mathbf{v} = (0,1)$ in R^2 with the Euclidean inner product, we have

$$\|\mathbf{u}\| = 1 \quad \text{and} \quad d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u}-\mathbf{v}\| = \|(1,-1)\| = \sqrt{1^2 - (-1)^2} = \sqrt{2}$$

  - However, if we change to the weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 2u_2v_2$, then we obtain

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2} = (3\cdot 1\cdot 1 + 2\cdot 0\cdot 0)^{1/2} = \sqrt{3}$$
$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u}-\mathbf{v}\| = \langle (1,-1), (1,-1) \rangle^{1/2} = [3\cdot 1\cdot 1 + 2\cdot(-1)\cdot(-1)]^{1/2} = \sqrt{5}$$

6-1 Unit Circles and Spheres in IPS

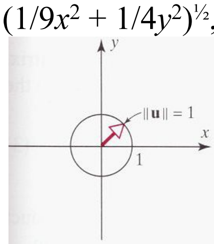
- If V is an inner product space, then the set of points in V that satisfy

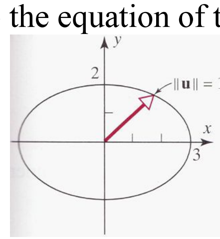
$$\|\mathbf{u}\| = 1$$

is called the unite sphere or sometimes the unit circle in V. In R^2 and R^3 these are the points that lie 1 unit away form the origin.

6-1 Example 5 (Unit Circles in R^2)

- Sketch the unit circle in an xy-coordinate system in R^2 using the Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + u_2v_2$
- Sketch the unit circle in an xy-coordinate system in R^2 using the Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 1/9u_1v_1 + 1/4u_2v_2$
- Solution
  - If $\mathbf{u} = (x,y)$, then $\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2} = (x^2 + y^2)^{1/2}$, so the equation of the unit circle is $x^2 + y^2 = 1$.
  - If $\mathbf{u} = (x,y)$, then $\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2} = (1/9x^2 + 1/4y^2)^{1/2}$, so the equation of the unit circle is $x^2/9 + y^2/4 = 1$.





6-1 Inner Product Generated by Matrices

- Let $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ be vectors in R^n (expressed as $n\times 1$ matrices), and let A be an invertible $n\times n$ matrix.
- If $\mathbf{u} \cdot \mathbf{v}$ is the Euclidean inner product on R^n, then the formula

$$\langle \mathbf{u}, \mathbf{v}\rangle = A\mathbf{u} \cdot A\mathbf{v}$$

  defines an inner product; it is called the inner product on R^n generated by A.
- The Euclidean inner product $\mathbf{u} \cdot \mathbf{v}$ can be written as the matrix product $\mathbf{v}^T\mathbf{u}$, the above formula can be written in the alternative form $\langle \mathbf{u}, \mathbf{v}\rangle = (A\mathbf{v})^T A\mathbf{u}$, or equivalently,

$$\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{v}^T A^T A\mathbf{u}$$

6-1 Example 6

(Inner Product Generated by the Identity Matrix)

- The inner product on R^n generated by the $n\times n$ identity matrix is the Euclidean inner product: Let $A = I$, we have $\langle \mathbf{u}, \mathbf{v}\rangle = I\mathbf{u} \cdot I\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$
- The weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v}\rangle = 3u_1v_1 + 2u_2v_2$ is the inner product on R^2 generated by

$$A = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \end{bmatrix}$$

  since

$$\langle u, v\rangle = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 3u_1v_1 + 2u_2v_2$$

- In general, the weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v}\rangle = w_1u_1v_1 + w_2u_2v_2 + \ldots + w_nu_nv_n$ is the inner product on R^n generated by

$$A = \text{diag}(\sqrt{w_1}, \sqrt{w_2}, \cdots, \sqrt{w_n})$$

6-1 Example 7 (An Inner Product on M_{22})

- If $U = \begin{bmatrix} u_1 & u_2 \\ u_3 & u_4 \end{bmatrix}$ and $V = \begin{bmatrix} v_1 & v_2 \\ v_3 & v_4 \end{bmatrix}$
  are any two 2×2 matrices, then
  $$\langle U, V\rangle = \mathrm{tr}(U^TV) = \mathrm{tr}(V^TU) = u_1v_1 + u_2v_2 + u_3v_3 + u_4v_4$$
  defines an inner product on M_{22}
- For example, if $U = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $V = \begin{bmatrix} -1 & 0 \\ 3 & 2 \end{bmatrix}$
  then $\langle U, V\rangle = 16$
- The norm of a matrix U relative to this inner product is
  $$\|U\| = <U,\ U>^{1/2} = \sqrt{u_1^2 + u_2^2 + u_3^2 + u_4^2}$$
  and the unit sphere in this space consists of all 2×2 matrices U whose entries satisfy the equation $\|U\| = 1$, which on squaring yields $u_1^2 + u_2^2 + u_3^2 + u_4^2 = 1$

6-1 Example 8 (An Inner Product on P_2)

- If $\mathbf{p} = a_0 + a_1x + a_2x^2$ and $\mathbf{q} = b_0 + b_1x + b_2x^2$ are any two vectors in P_2,
  - An inner product on P_2:
  $$\langle \mathbf{p}, \mathbf{q}\rangle = a_0b_0 + a_1b_1 + a_2b_2$$
- The norm of the polynomial $\mathbf{p}$ relative to this inner product is
  $$\|\mathbf{p}\| = <\mathbf{p},\ \mathbf{p}>^{1/2} = \sqrt{a_0^2 + a_1^2 + a_2^2}$$
- The unit sphere in this space consists of all polynomials $\mathbf{p}$ in P_2 whose coefficients satisfy the equation $\|\mathbf{p}\| = 1$, which on squaring yields
  $$a_0^2 + a_1^2 + a_2^2 = 1$$

Theorem 6.1.1 (Properties of Inner Products)

- If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are vectors in a real inner product space, and k is any scalar, then:
  - $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$
  - $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
  - $\langle \mathbf{u}, k\mathbf{v} \rangle = k \langle \mathbf{u}, \mathbf{v} \rangle$
  - $\langle \mathbf{u} - \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle - \langle \mathbf{v}, \mathbf{w} \rangle$
  - $\langle \mathbf{u}, \mathbf{v} - \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{u}, \mathbf{w} \rangle$

6-1 Example 11

- $\langle \mathbf{u} - 2\mathbf{v}, 3\mathbf{u} + 4\mathbf{v} \rangle$
  $= \langle \mathbf{u}, 3\mathbf{u} + 4\mathbf{v} \rangle - \langle 2\mathbf{v}, 3\mathbf{u}+4\mathbf{v} \rangle$
  $= \langle \mathbf{u}, 3\mathbf{u} \rangle + \langle \mathbf{u}, 4\mathbf{v} \rangle - \langle 2\mathbf{v}, 3\mathbf{u} \rangle - \langle 2\mathbf{v}, 4\mathbf{v} \rangle$
  $= 3 \langle \mathbf{u}, \mathbf{u} \rangle + 4 \langle \mathbf{u}, \mathbf{v} \rangle - 6 \langle \mathbf{v}, \mathbf{u} \rangle - 8 \langle \mathbf{v}, \mathbf{v} \rangle$
  $= 3 \| \mathbf{u} \|^2 + 4 \langle \mathbf{u}, \mathbf{v} \rangle - 6 \langle \mathbf{u}, \mathbf{v} \rangle - 8 \| \mathbf{v} \|^2$
  $= 3 \| \mathbf{u} \|^2 - 2 \langle \mathbf{u}, \mathbf{v} \rangle - 8 \| \mathbf{v} \|^2$

Chapter Content

- Inner Products
- Angle and Orthogonality in Inner Product Spaces
- Orthonormal Bases; Gram-Schmidt Process; QR-Decomposition
- Best Approximation; Least Squares
- Orthogonal Matrices; Change of Basis

Theorems 6.2.1 & 6.2.2

- Theorem 6.2.1 (Cauchy-Schwarz Inequality)
  - If **u** and **v** are vectors in a real inner product space, then

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \le \|\mathbf{u}\|\,\|\mathbf{v}\|$$

- Theorem 6.2.2 (Properties of Length)
  - If **u** and **v** are vectors in an inner product space V, and if k is any scalar, then :
    - $\|\mathbf{u}\| \ge 0$
    - $\|\mathbf{u}\| = 0$ if and only if $\mathbf{u} = \mathbf{0}$
    - $\|k\mathbf{u}\| = |k|\,\|\mathbf{u}\|$
    - $\|\mathbf{u} + \mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|$ (Triangle inequality)

Theorem 6.2.3 (Properties of Distance)

- If **u**, **v**, and **w** are vectors in an inner product space V, and if k is any scalar, then:
  - $d(\mathbf{u}, \mathbf{v}) \geq 0$
  - $d(\mathbf{u}, \mathbf{v}) = 0$ if and only if $\mathbf{u} = \mathbf{v}$
  - $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$
  - $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$ (Triangle inequality)

6-2 Angle Between Vectors

- The Cauchy-Schwarz inequality for R^n (Theorem 4.1.3) follows as a special case of Theorem 6.2.1 by taking $\langle \mathbf{u}, \mathbf{v} \rangle$ to be the Euclidean inner product $\mathbf{u} \cdot \mathbf{v}$.

- The angle between vectors in general inner product spaces can be defined as

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|} \quad \text{and} \quad 0 \leq \theta \leq \pi$$

- Example 2
  - Let R^4 have the Euclidean inner product. Find the cosine of the angle θ between the vectors $\mathbf{u} = (4, 3, 1, -2)$ and $\mathbf{v} = (-2, 1, 2, 3)$.

6-2 Orthogonality

- Two vectors **u** and **v** in an inner product space are called orthogonal if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

- Example 3
  - If M_{22} has the inner project defined previously, then the matrices

$$U = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \text{ and } V = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$$

are orthogonal, since $\langle U, V \rangle = 1(0) + 0(2) + 1(0) + 1(0) = 0$.

6-2 Example 4 (Orthogonal Vectors in P_2)

- Let P_2 have the inner product $<\mathbf{p}, \mathbf{q}> = \int_{-1}^{1} p(x)q(x)dx$ and let $\mathbf{p} = x$ and $\mathbf{q} = x^2$.

- Then

$$\|\mathbf{p}\| = <\mathbf{p}, \mathbf{p}>^{1/2} = \left[\int_{-1}^{1} xxdx\right]^{1/2} = \left[\int_{-1}^{1} x^2dx\right]^{1/2} = \sqrt{\frac{2}{3}}$$

$$\|\mathbf{q}\| = <\mathbf{q}, \mathbf{q}>^{1/2} = \left[\int_{-1}^{1} x^2x^2dx\right]^{1/2} = \left[\int_{-1}^{1} x^4dx\right]^{1/2} = \sqrt{\frac{2}{5}}$$

$$<\mathbf{p}, \mathbf{q}> = \int_{-1}^{1} xx^2dx = \int_{-1}^{1} x^3dx = 0$$

because $\langle \mathbf{p}, \mathbf{q} \rangle = 0$, the vectors $\mathbf{p} = x$ and $\mathbf{q} = x^2$ are orthogonal relative to the given inner product.

Theorem 6.2.4 (Generalized Theorem of Pythagoras)

- If **u** and **v** are orthogonal vectors in an inner product space, then

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

6-2 Example 5

- Since $\mathbf{p} = x$ and $\mathbf{q} = x^2$ are orthogonal relative to the inner product $<\mathbf{p}, \mathbf{q}> = \int_{-1}^{1} p(x)q(x)dx$ on P_2.
- It follows from the Theorem of Pythagoras that

$$\| \mathbf{p} + \mathbf{q} \|^2 = \| \mathbf{p} \|^2 + \| \mathbf{q} \|^2$$

- Thus, from the previous example:

$$\|\mathbf{p}+\mathbf{q}\|^2 = (\sqrt{\frac{2}{3}})^2 + (\sqrt{\frac{2}{5}})^2 = \frac{2}{3} + \frac{2}{5} = \frac{16}{15}$$

- We can check this result by direct integration:

$$\|\mathbf{p}+\mathbf{q}\|^2 = <\mathbf{p}+\mathbf{q}, \mathbf{p}+\mathbf{q}> = \int_{-1}^{1} (x+x^2)(x+x^2)dx$$

$$= \int_{-1}^{1} x^2 dx + 2\int_{-1}^{1} x^3 dx + \int_{-1}^{1} x^4 dx = \frac{2}{3} + 0 + \frac{2}{5} = \frac{16}{15}$$

6-2 Orthogonality

- Let W be a subspace of an inner product space V.
  - A vector **u** in V is said to be orthogonal to W if it is orthogonal to every vector in W, and
  - the set of all vectors in V that are orthogonal to W is called the orthogonal complement of W.

Theorem 6.2.5

(Properties of Orthogonal Complements)

- If W is a subspace of a finite-dimensional inner product space V, then:
  - $W^{\perp}$ is a subspace of V.
  - The only vector common to W and $W^{\perp}$ is $\mathbf{0}$; that is , $W \cap W^{\perp} = \mathbf{0}$.
  - The orthogonal complement of $W^{\perp}$ is W; that is , $(W^{\perp})^{\perp} = W$.

Theorem 6.2.6

- If A is an $m \times n$ matrix, then:
  - The nullspace of A and the row space of A are orthogonal complements in R^n with respect to the Euclidean inner product.
  - The nullspace of A^T and the column space of A are orthogonal complements in R^m with respect to the Euclidean inner product.

6-2 Example 6 (Basis for an Orthogonal Complement)

- Let W be the subspace of R^5 spanned by the vectors $\mathbf{w}_1=(2, 2, -1, 0, 1)$, $\mathbf{w}_2=(-1, -1, 2, -3, 1)$, $\mathbf{w}_3=(1, 1, -2, 0, -1)$, $\mathbf{w}_4=(0, 0, 1, 1, 1)$. Find a basis for the orthogonal complement of W.
- Solution
  - The space W spanned by $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$, and $\mathbf{w}_4$ is the same as the row space of the matrix

$$A = \begin{bmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 1 & 1 & -2 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Theorem 6.2.7 (Equivalent Statements)

- If A is an $m\times n$ matrix, and if $T_A : R^n \to R^n$ is multiplication by A, then the following are equivalent:
  - A is invertible.
  - $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
  - The reduced row-echelon form of A is I_n.
  - A is expressible as a product of elementary matrices.
  - $A\mathbf{x} = \mathbf{b}$ is consistent for every $n\times 1$ matrix $\mathbf{b}$.
  - $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n\times 1$ matrix $\mathbf{b}$.
  - $\det(A)\neq 0$.
  - The range of T_A is R^n.
  - T_A is one-to-one.
  - The column vectors of A are linearly independent.
  - The row vectors of A are linearly independent.
  - The column vectors of A span R^n.
  - The row vectors of A span R^n.
  - The column vectors of A form a basis for R^n.
  - The row vectors of A form a basis for R^n.
  - A has rank n.
  - A has nullity 0.
  - The orthogonal complement of the nullspace of A is R^n.
  - The orthogonal complement of the row of A is $\{\mathbf{0}\}$.

Chapter Content

- Inner Products
- Angle and Orthogonality in Inner Product Spaces
- Orthonormal Bases; Gram-Schmidt Process; QR-Decomposition
- Best Approximation; Least Squares
- Change of Basis
- Orthogonal Matrices

6-3 Orthonormal Basis

- A set of vectors in an inner product space is called an orthogonal set if all pairs of distinct vectors in the set are orthogonal.
- An orthogonal set in which each vector has norm 1 is called orthonormal.

- Example 1
  - Let $\mathbf{u}_1 = (0, 1, 0)$, $\mathbf{u}_2 = (1, 0, 1)$, $\mathbf{u}_3 = (1, 0, -1)$ and assume that R^3 has the Euclidean inner product.
  - It follows that the set of vectors $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is orthogonal since

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle \mathbf{u}_1, \mathbf{u}_3 \rangle = \langle \mathbf{u}_2, \mathbf{u}_3 \rangle = 0.$$

6-3 Example 2

- Let $\mathbf{u}_1 = (0, 1, 0)$, $\mathbf{u}_2 = (1, 0, 1)$, $\mathbf{u}_3 = (1, 0, -1)$
  - The Euclidean norms of the vectors are

$$\|\mathbf{u}_1\| = 1, \quad \|\mathbf{u}_2\| = \sqrt{2}, \quad \|\mathbf{u}_3\| = \sqrt{2}$$

  - Normalizing $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ yields

$$\mathbf{v}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = (0,1,0), \quad \mathbf{v}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = (\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}}), \quad \mathbf{v}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} = (\frac{1}{\sqrt{2}},0,-\frac{1}{\sqrt{2}})$$

  - The set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is orthonormal since

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathbf{v}_3 \rangle = \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0 \quad \textit{and} \quad \|\mathbf{v}_1\| = \|\mathbf{v}_2\| = \|\mathbf{v}_3\| = 1$$

6-3 Orthonormal Basis

- Theorem 6.3.1*
  - If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthonormal basis for an inner product space V, and $\mathbf{u}$ is any vector in V, then
$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \cdots + \langle \mathbf{u}, \mathbf{v}_n \rangle \mathbf{v}_n$$
- Remark
  - The scalars $\langle \mathbf{u}, \mathbf{v}_1 \rangle, \langle \mathbf{u}, \mathbf{v}_2 \rangle, \ldots, \langle \mathbf{u}, \mathbf{v}_n \rangle$ are the coordinates of the vector $\mathbf{u}$ relative to the orthonormal basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and
$$(\mathbf{u})_S = (\langle \mathbf{u}, \mathbf{v}_1 \rangle, \langle \mathbf{u}, \mathbf{v}_2 \rangle, \ldots, \langle \mathbf{u}, \mathbf{v}_n \rangle)$$
    is the coordinate vector of $\mathbf{u}$ relative to this basis

6-3 Example 3

- Let $\mathbf{v}_1 = (0, 1, 0)$, $\mathbf{v}_2 = (-4/5, 0, 3/5)$, $\mathbf{v}_3 = (3/5, 0, 4/5)$.
  It is easy to check that $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal basis for R^3 with the Euclidean inner product.
  Express the vector $\mathbf{u} = (1, 1, 1)$ as a linear combination of the vectors in S, and find the coordinate vector $(\mathbf{u})_s$.

- Solution:
  - $\langle \mathbf{u}, \mathbf{v}_1 \rangle = 1$, $\langle \mathbf{u}, \mathbf{v}_2 \rangle = -1/5$, $\langle \mathbf{u}, \mathbf{v}_3 \rangle = 7/5$
  - Therefore, by Theorem 6.3.1 we have $\mathbf{u} = \mathbf{v}_1 - 1/5\, \mathbf{v}_2 + 7/5\, \mathbf{v}_3$
  - That is, $(1, 1, 1) = (0, 1, 0) - 1/5\,(-4/5, 0, 3/5) + 7/5\,(3/5, 0, 4/5)$
  - The coordinate vector of $\mathbf{u}$ relative to S is
$$(\mathbf{u})_s = (\langle \mathbf{u}, \mathbf{v}_1 \rangle, \langle \mathbf{u}, \mathbf{v}_2 \rangle, \langle \mathbf{u}, \mathbf{v}_3 \rangle) = (1, -1/5, 7/5)$$

Theorem 6.3.2

- If S is an orthonormal basis for an n-dimensional inner product space, and if $(\mathbf{u})_s = (u_1, u_2, \ldots, u_n)$ and $(\mathbf{v})_s = (v_1, v_2, \ldots, v_n)$ then:

$$\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

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

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

- Example 4 (Calculating Norms Using Orthonormal Bases)
  - u=(1, 1, 1)

6-3 Coordinates Relative to Orthogonal Bases

- If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal basis for a vector space V, then normalizing each of these vectors yields the orthonormal basis

$$S' = \left\{ \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}, \ldots, \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|} \right\}$$

- Thus, if $\mathbf{u}$ is any vector in V, it follows from theorem 6.3.1 that

$$\mathbf{u} = \left\langle \mathbf{u}, \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \right\rangle \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} + \left\langle \mathbf{u}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} \right\rangle \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} + \cdots + \left\langle \mathbf{u}, \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|} \right\rangle \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|}$$

  or

$$\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2} \mathbf{v}_2 + \cdots + \frac{\langle \mathbf{u}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|^2} \mathbf{v}_n$$

- The above equation expresses $\mathbf{u}$ as a linear combination of the vectors in the orthogonal basis S.

Theorem 6.3.3

- If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space
  - then S is linearly independent.

- Remark
  - By working with orthonormal bases, the computation of general norms and inner products can be reduced to the computation of Euclidean norms and inner products of the coordinate vectors.

Theorem 6.3.4 (Projection Theorem)

- If W is a finite-dimensional subspace of an product space V,
  - then every vector $\mathbf{u}$ in V can be expressed in exactly one way as

$$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$$

where $\mathbf{w}_1$ is in W and $\mathbf{w}_2$ is in $W^{\perp}$.

Theorem 6.3.5

- Let W be a finite-dimensional subspace of an inner product space V.
  - If $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is an orthonormal basis for W, and $\mathbf{u}$ is any vector in V, then

$$\text{proj}_W\mathbf{u} = \langle \mathbf{u},\mathbf{v}_1\rangle\, \mathbf{v}_1 + \langle \mathbf{u},\mathbf{v}_2\rangle\, \mathbf{v}_2 + \ldots + \langle \mathbf{u},\mathbf{v}_r\rangle\, \mathbf{v}_r$$

  - If $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is an orthogonal basis for W, and $\mathbf{u}$ is any vector in V, then

$$\text{proj}_W\,\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\langle \mathbf{u}, \mathbf{v}_2\rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{u}, \mathbf{v}_r\rangle}{\|\mathbf{v}_r\|^2}\mathbf{v}_r \Leftarrow$$

Need Normalization

6-3 Example 6

- Let R^3 have the Euclidean inner product, and let W be the subspace spanned by the orthonormal vectors $\mathbf{v}_1 = (0, 1, 0)$ and $\mathbf{v}_2 = (-4/5, 0, 3/5)$.
- From the above theorem, the orthogonal projection of $\mathbf{u} = (1, 1, 1)$ on W is $\text{proj}_W\mathbf{u} = <\mathbf{u}, \mathbf{v}_1>\mathbf{v}_1 + <\mathbf{u}, \mathbf{v}_2>\mathbf{v}_2$

$$=(1)(0,\ 1,\ 0)+(-\frac{1}{5})(-\frac{4}{5},\ 0,\ \frac{3}{5})=(\frac{4}{25},\ 1,\ -\frac{3}{25})$$

- The component of $\mathbf{u}$ orthogonal to W is

$$\text{proj}_{W^\perp}\mathbf{u} = \mathbf{u} - \text{proj}_W\mathbf{u} = (1,\ 1,\ 1) - (\frac{4}{25},\ 1,\ -\frac{3}{25}) = (\frac{21}{25},\ 0,\ \frac{28}{25})$$

- Observe that $\text{proj}_{W^\perp}\mathbf{u}$ is orthogonal to both $\mathbf{v}_1$ and $\mathbf{v}_2$.

6-3 Finding Orthogonal/Orthonormal Bases

- Theorem 6.3.6
  - Every nonzero finite-dimensional inner product space has an orthonormal basis.

- Remark
  - The step-by-step construction for converting an arbitrary basis into an orthogonal basis is called the Gram-Schmidt process.

6-3 Example 7(Gram-Schmidt Process)

- Consider the vector space R^3 with the Euclidean inner product. Apply the Gram-Schmidt process to transform the basis vectors

$$\mathbf{u}_1 = (1,\ 1,\ 1),\ \mathbf{u}_2 = (0,\ 1,\ 1),\ \mathbf{u}_3 = (0,\ 0,\ 1)$$

  into an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$; then normalize the orthogonal basis vectors to obtain an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$.

- Solution:
  - Step 1: Let $\mathbf{v}_1 = \mathbf{u}_1$. That is, $\mathbf{v}_1 = \mathbf{u}_1 = (1,\ 1,\ 1)$
  - Step 2: Let $\mathbf{v}_2 = \mathbf{u}_2 - \text{proj}_{W_1}\mathbf{u}_2$. That is,

$$\mathbf{v}_2 = \mathbf{u}_2 - \text{proj}_{W_1}\mathbf{u}_2 = \mathbf{u}_2 - \frac{<\mathbf{u}_2,\ \mathbf{v}_1>}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$$

$$= (0,\ 1,\ 1) - \frac{2}{3}(1,\ 1,\ 1) = (-\frac{2}{3},\ \frac{1}{3},\ \frac{1}{3})$$

6-3 Example 7(Gram-Schmidt Process)

We have two vectors in W_2 now!

- Step 3: Let $\mathbf{v}_3 = \mathbf{u}_3 - \text{proj}_{W_2}\mathbf{u}_3$. That is,

$$\mathbf{v}_3 = \mathbf{u}_3 - \text{proj}_{W_2}\mathbf{u}_3 = \mathbf{u}_3 - \frac{<\mathbf{u}_3,\ \mathbf{v}_1>}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{<\mathbf{u}_3,\ \mathbf{v}_2>}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$$

$$= (0,\ 1,\ 1) - \frac{1}{3}(1,\ 1,\ 1) - \frac{1/3}{2/3}(-\frac{2}{3},\ \frac{1}{3},\ \frac{1}{3}) = (0,\ -\frac{1}{2},\ \frac{1}{2})$$

- Thus, $\mathbf{v}_1 = (1, 1, 1)$, $\mathbf{v}_2 = (-2/3, 1/3, 1/3)$, $\mathbf{v}_3 = (0, -1/2, 1/2)$ form an orthogonal basis for R^3. The norms of these vectors are

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

so an orthonormal basis for R^3 is

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = (\frac{1}{\sqrt{3}},\ \frac{1}{\sqrt{3}},\ \frac{1}{\sqrt{3}}),\ \ \mathbf{q}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = (-\frac{2}{\sqrt{6}},\ \frac{1}{6},\ \frac{1}{\sqrt{6}}),$$

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

Theorem 6.3.7 (*QR*-Decomposition)

- If A is an $m \times n$ matrix with linearly independent column vectors, then A can be factored as

$$A = QR$$

where Q is an $m \times n$ matrix with orthonormal column vectors, and R is an $n \times n$ invertible upper triangular matrix.

6-3 Example 8
(QR-Decomposition of a 3×3 Matrix)

- Find the QR-decomposition of
$$A=\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}$$
- Solution:
  - The column vectors A are
$$\mathbf{u}_1=\begin{bmatrix}1\\1\\1\end{bmatrix},\ \mathbf{u}_2=\begin{bmatrix}0\\1\\1\end{bmatrix},\ \mathbf{u}_3=\begin{bmatrix}0\\-1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}$$
  - Applying the Gram-Schmidt process with subsequent normalization to these column vectors yields the orthonormal vectors
$$\mathbf{q}_1=\begin{bmatrix}1/\sqrt{3}\\1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix},\ \mathbf{q}_2=\begin{bmatrix}-2/\sqrt{6}\\1/\sqrt{6}\\1/\sqrt{6}\end{bmatrix},\ \mathbf{q}_3=\begin{bmatrix}0\\-1/\sqrt{2}\\1/\sqrt{2}\end{bmatrix}\quad\Rightarrow\ Q$$

6-3 Example 8

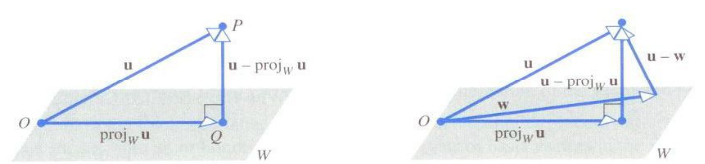
- The matrix R is
$$R=\begin{bmatrix}\langle\mathbf{u}_1,\mathbf{q}_1\rangle&\langle\mathbf{u}_2,\mathbf{q}_1\rangle&\langle\mathbf{u}_3,\mathbf{q}_1\rangle\\0&\langle\mathbf{u}_2,\mathbf{q}_2\rangle&\langle\mathbf{u}_3,\mathbf{q}_2\rangle\\0&0&\langle\mathbf{u}_3,\mathbf{q}_3\rangle\end{bmatrix}=\begin{bmatrix}3/\sqrt{3}&2/\sqrt{3}&1/\sqrt{3}\\0&2/\sqrt{6}&1/\sqrt{6}\\0&0&1/\sqrt{2}\end{bmatrix}$$
- Thus, the QR-decomposition of A is
$$\underbrace{\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}}_{A}=\underbrace{\begin{bmatrix}1/\sqrt{3}&-2/\sqrt{6}&0\\1/\sqrt{3}&1/\sqrt{6}&-1/\sqrt{2}\\1/\sqrt{3}&1/\sqrt{6}&1/\sqrt{3}\end{bmatrix}}_{Q}\underbrace{\begin{bmatrix}3/\sqrt{3}&2/\sqrt{3}&1/\sqrt{3}\\0&2/\sqrt{6}&1/\sqrt{6}\\0&0&1/\sqrt{2}\end{bmatrix}}_{R}$$

Chapter Content

- Inner Products
- Angle and Orthogonality in Inner Product Spaces
- Orthonormal Bases; Gram-Schmidt Process; QR-Decomposition
- Best Approximation; Least Squares
- Change of Basis
- Orthogonal Matrices

6-4 Orthogonal Projections Viewed as Approximations

- If P is a point in 3-space and W is a plane through the origin, then the point Q in W closest to P is obtained by dropping a perpendicular from P to W.



(a) Q is the point in W closest to P. (b) $\|\mathbf{u} - \mathbf{w}\|$ is minimized by $\mathbf{w} = \text{proj}_W\, \mathbf{u}$.

Figure 6.4.1

- If we let $\mathbf{u} = OP$, the distance between P and W is given by $\|\ \mathbf{u} - \text{proj}_W\mathbf{u}\ \|$.
- In other words, among all vectors $\mathbf{w}$ in W the vector $\mathbf{w} = \text{proj}_W\mathbf{u}$ minimize the distance $\|\ \mathbf{v} - \mathbf{w}\ \|$.

6-4 Best Approximation

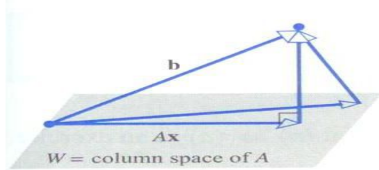
- Remark
  - Suppose **u** is a vector that we would like to approximate by a vector in W.
  - Any approximation **w** will result in an "error vector" $\mathbf{u} - \mathbf{w}$ which, unless **u** is in W, cannot be made equal to **0**.
  - However, by choosing $\mathbf{w} = \text{proj}_W\mathbf{u}$ we can make the length of the error vector
    $\|\ \mathbf{u} - \mathbf{w}\ \| = \|\ \mathbf{u} - \text{proj}_W\mathbf{u}\ \|$ as small as possible.
  - Thus, we can describe $\text{proj}_W\mathbf{u}$ as the "*best approximation*" to **u** by the vectors in W.

Theorem 6.4.1 (Best Approximation Theorem)

- If W is a finite-dimensional subspace of an inner product space V, and if **u** is a vector in V,
  - then $\text{proj}_W\mathbf{u}$ is the best approximation to **u** form W in the sense that

$$\|\ \mathbf{u} - \text{proj}_W\mathbf{u}\ \| < \|\ \mathbf{u} - \mathbf{w}\ \|$$

for every vector **w** in W that is different from $\text{proj}_W\mathbf{u}$.



6-4 Least Square Problem

- Given a linear system $A\mathbf{x} = \mathbf{b}$ of m equations in n unknowns
  - find a vector $\mathbf{x}$, if possible, that minimize $\| A\mathbf{x} - \mathbf{b} \|$ with respect to the Euclidean inner product on R^m.
  - Such a vector is called a least squares solution of $A\mathbf{x} = \mathbf{b}$.

Theorem 6.4.2

- For *any* linear system $A\mathbf{x} = \mathbf{b}$, the associated normal system

$$A^TA\mathbf{x} = A^T\mathbf{b}$$

is *consistent*, and all solutions of the normal system are least squares solutions of $A\mathbf{x} = \mathbf{b}$.

Moreover, if W is the column space of A, and $\mathbf{x}$ is any least squares solution of $A\mathbf{x} = \mathbf{b}$, then the orthogonal projection of $\mathbf{b}$ on W is

$$\text{proj}_W\mathbf{b} = A\mathbf{x}$$

(or you can treat it as $A\mathbf{x} - \text{proj}_W\mathbf{b} = \mathbf{0}$)

Theorem 6.4.3

- If A is an $m \times n$ matrix, then the following are equivalent.
  - A has linearly independent column vectors.
  - $A^T A$ is invertible.

Theorem 6.4.4

- If A is an $m \times n$ matrix with linearly independent column vectors,
  - then for every $m \times 1$ matrix $\mathbf{b}$, the linear system $A\mathbf{x} = \mathbf{b}$ has a unique least squares solution.
  - This solution is given by

$$\mathbf{x} = (A^T A)^{-1} A^T \mathbf{b}$$

  - Moreover, if W is the column space of A, then the orthogonal projection of $\mathbf{b}$ on W is

$$\text{proj}_W \mathbf{b} = A\mathbf{x} = A(A^T A)^{-1} A^T \mathbf{b}$$

6-4 Example 1 (Least Squares Solution)

- Find the least squares solution of the linear system $A\mathbf{x} = \mathbf{b}$ given by

$$\begin{aligned} x_1 - \ \ x_2 &= 4 \\ 3x_1 + 2x_2 &= 1 \\ -2x_1 + 4x_2 &= 3 \end{aligned}$$

  and find the orthogonal projection of **b** on the column space of A.
- Solution:

$$A = \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

  - Observe that A has linearly independent column vectors, so we know in advance that there is a unique least squares solution.

6-4 Example 1

- We have

$$A^T A = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

  so the normal system $A^T A\mathbf{x} = A^T\mathbf{b}$ in this case is $\begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$

- Solving this system yields the least squares solution

$$x_1 = 17/95, \ x_2 = 143/285$$

- The orthogonal projection of **b** on the column space of A is

$$A\mathbf{x} = \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 17/95 \\ 143/285 \end{bmatrix} = \begin{bmatrix} -92/285 \\ 439/285 \\ 94/57 \end{bmatrix}$$

6-4 Example 2 (Orthogonal Projection on a Subspace)

- Find the orthogonal projection of the vector $\mathbf{u} = (-3,-3,8,9)$ on the subspace of R^4 spanned by the vectors

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

- Solution:
  - The subspace spanned by $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$, is the column space of

$$A = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix}$$

  - If $\mathbf{u}$ is expressed as a column vector, we can find the orthogonal projection of $\mathbf{u}$ on W by finding a least squares solution of the system $A\mathbf{x} = \mathbf{u}$.
  - $\text{proj}_W\mathbf{u} = A\mathbf{x}$ from the least square solution.

6-4 Example 2

- From Theorem 6.4.4, the least squares solution is given by

$$\mathbf{x} = (A^TA)^{-1}A^T\mathbf{u}$$

- That is,

$$\mathbf{x} = \left(\begin{bmatrix} 3 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ -1 & 0 & 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & 1 & -1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 1 & -1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 3 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ -1 & 0 & 2 & -1 \end{bmatrix} \begin{bmatrix} -3 \\ -3 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$$

- Thus, $\text{proj}_W\mathbf{u} = A\mathbf{x} = [-2\ 3\ 4\ 0]^T$
- Second method: using Gram-Schmidt process

6-4 Orthogonal Projection

- If W is a subspace of R^m,
  - then the transformation $P: R^m \rightarrow W$ that maps each vector $\mathbf{x}$ in R^m into its orthogonal projection $\text{proj}_W\mathbf{x}$ in W is called orthogonal projection of R^m on W.

- The orthogonal projections are linear operators
  - The standard matrix for the orthogonal projection of R^m on W is $[P] = A(A^TA)^{-1}A^T$

  where A is constructed using any basis for W as its column vectors

6-4 Example 3

- Verifying formula $[P] = A(A^TA)^{-1}A^T$
- The standard matrix for the orthogonal projection of R3 on the xy-plane :

6-4 Example 4

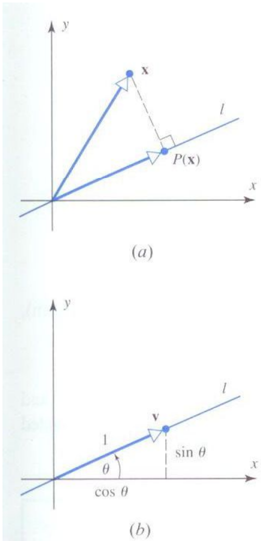


Figure 6.4.3

Theorem 6.4.5 (Equivalent Statements)

- If A is an $m \times n$ matrix, and if $T_A : R^n \rightarrow R^n$ is multiplication by A, then the following are equivalent:
  - A is invertible.
  - $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
  - The reduced row-echelon form of A is I_n.
  - A is expressible as a product of elementary matrices.
  - $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.
  - $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.
  - $\det(A) \neq 0$.
  - The range of T_A is R^n.
  - T_A is one-to-one.
  - The column vectors of A are linearly independent.
  - The row vectors of A are linearly independent.

6-4 Example 4

- Find the standard matrix for the orthogonal projection P of R^2 on the line that passes through the origin and makes an angle θ with the positive x-axis.

Theorem 6.4.5 (Equivalent Statements)

- The column vectors of A span R^n.
- The row vectors of A span R^n.
- The column vectors of A form a basis for R^n.
- The row vectors of A form a basis for R^n.
- A has rank n.
- A has nullity 0.
- The orthogonal complement of the nullspace of A is R^n.
- The orthogonal complement of the row space of A is $\{\mathbf{0}\}$.
- A^TA is invertible.

6-4 Coordinate Matrices

- Recall from Theorem 5.4.1 that if $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for a vector space V, then each vector $\mathbf{v}$ in V can be expressed uniquely as a linear combination of the basis vectors, say

$$\mathbf{v} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots + k_n\mathbf{v}_n$$

- The scalars $k_1, k_2, \ldots, k_n$ are the coordinates of $\mathbf{v}$ relative to S, and the vector

$$(\mathbf{v})_S = (k_1, k_2, \ldots, k_n)$$

is the coordinate vector of $\mathbf{v}$ relative to S.
- Thus, we define

$$[\mathbf{v}]_S = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}$$

to be the coordinate matrix of $\mathbf{v}$ relative to S.

Chapter Content

- Inner Products
- Angle and Orthogonality in Inner Product Spaces
- Orthonormal Bases; Gram-Schmidt Process; QR-Decomposition
- Best Approximation; Least Squares
- Change of Basis
- Orthogonal Matrices

6-5 Change of Basis

- Change of basis problem
  - change the basis for a vector space V from some old basis B to a new basis B' $[\mathbf{v}]_B$ => $[\mathbf{v}]_{B'}$?
- Solution of the change of basis problem
  - change the basis for a vector space V from some old basis $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ to some new basis $B' = \{\mathbf{u}_1', \mathbf{u}_2', \ldots, \mathbf{u}_n'\}$,

$$[\mathbf{v}]_B = P\,[\mathbf{v}]_{B'}$$

where the column of P are the coordinate matrices of the new basis vectors relative to the old basis; that is, the column vectors of P are

$$[\mathbf{u}_1']_B, [\mathbf{u}_2']_B, \ldots, [\mathbf{u}_n']_B$$

  - The matrix P is called the transition matrix form B' to B; it can be expressed in terms of its column vector as

$$P = [[\mathbf{u}_1']_B \mid [\mathbf{u}_2']_B \mid \ldots \mid [\mathbf{u}_n']_B]$$

6-5 Example 1 (Finding a Transition Matrix)

- Consider bases $B = \{\mathbf{u}_1, \mathbf{u}_2\}$ and $B' = \{\mathbf{u}_1', \mathbf{u}_2'\}$ for R^2, where

$$\mathbf{u}_1 = (1, 0),\ \mathbf{u}_2 = (0, 1);$$
$$\mathbf{u}_1' = (1, 1),\ \mathbf{u}_2' = (2, 1).$$

  Find the transition matrix from B' to B.

  Find $[\mathbf{v}]_B$ if $[\mathbf{v}]_{B'} = [-3\ 5]^T$.
- Solution:
  - find the coordinate matrices for the new basis vectors $\mathbf{u}_1'$ and $\mathbf{u}_2'$ relative to the old basis B
  - By inspection $\mathbf{u}'_1 = \mathbf{u}_1 + \mathbf{u}_2$ so that

$$[\mathbf{u}'_1]_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } [\mathbf{u}'_2]_B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

  - Thus, the transition matrix from B' to B is

$$P = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$$

6-5 Example 1
(A different view point on Example 1)

- Consider bases $B = \{\mathbf{u}_1, \mathbf{u}_2\}$ and $B' = \{\mathbf{u}_1', \mathbf{u}_2'\}$ for R^2, where

$$\mathbf{u}_1 = (1, 0),\ \mathbf{u}_2 = (0, 1);$$
$$\mathbf{u}_1' = (1, 1),\ \mathbf{u}_2' = (2, 1).$$

  Find the transition matrix from B to B'.

Theorem 6.5.1

- If P is the transition matrix from a basis B' to a basis B for a finite-dimensional vector space V, then:
  - P is invertible.
  - P^{-1} is the transition matrix from B to B'.

- Remark
  - If P is the transition matrix from a basis B' to a basis B, then for every $\mathbf{v}$ the following relationships hold:

$$[\mathbf{v}]_B = P\,[\mathbf{v}]_{B'}$$
$$[\mathbf{v}]_{B'} = P^{-1}\,[\mathbf{v}]_B$$

Chapter Content

- Inner Products
- Angle and Orthogonality in Inner Product Spaces
- Orthonormal Bases; Gram-Schmidt Process; QR-Decomposition
- Best Approximation; Least Squares
- Change of Basis
- Orthogonal Matrices

6-6 Orthogonal Matrix

- *A* square matrix A with the property

  $$A^{-1} = A^T$$

  is said to be an orthogonal matrix.

- Remark
  - A square matrix A is orthogonal if and only if $AA^T = I$ or $A^TA = I$.

6-6 Example 1 & 2

- Example 1

$$A=\begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix} \qquad A^T A=\begin{bmatrix} \frac{3}{7} & -\frac{6}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{3}{7} & \frac{6}{7} \\ \frac{6}{7} & \frac{2}{7} & -\frac{3}{7} \end{bmatrix}\begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Example 2
  - Rotation and reflection matrices is orthogonal.

$$A=\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Theorem 6.6.1

- The following are equivalent for an $n \times n$ matrix A.
  - A is orthogonal.
  - The row vectors of A form an orthonormal set in R^n with the Euclidean inner product.
  - The column vectors of A form an orthonormal set in R^n with the Euclidean inner product.

Theorem 6.6.2

- The inverse of an orthogonal matrix is orthogonal.
- A product of orthogonal matrices is orthogonal.
- If A is orthogonal, then $\det(A) = 1$ or $\det(A) = -1$.

6-6 Example 3

- The matrix

$$A = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

  is orthogonal since its row (and column) vectors form orthonormal sets in R^2.
- We have $\det(A) = 1$.
- Interchanging the rows produces an orthogonal matrix for which $\det(A) = -1$.

Theorem 6.6.3 (Orthogonal Matrices as Linear Operators)

- If A is an $n \times n$ matrix, then the following are equivalent.
  - A is orthogonal.
  - $\| A\mathbf{x} \| = \| \mathbf{x} \|$ for all $\mathbf{x}$ in R^n.
  - $A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}$ and $\mathbf{y}$ in R^n.

- Remark:
  - If $T: R^n \to R^n$ is multiplication by an orthogonal matrix A, then T is called an orthogonal operator on R^n.
  - It follows from the preceding theorem that the orthogonal operator on R^n are precisely those operators that leave the length of all vectors unchanged.

Theorem 6.6.4

- If P is the transition matrix from one orthonormal basis to another orthonormal basis for an inner product space,
  - then P is an orthogonal matrix; that is,

$$P^{-1} = P^T$$

6-6 Example 4
(Application to Rotation of Axes in 2-Space)

- A change from basis B={u_1, u_2} to B'={u_1', u_2'}

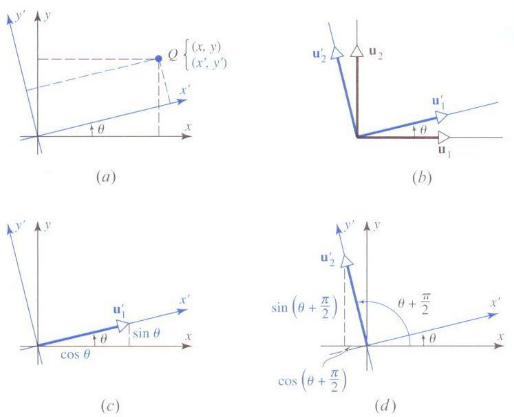


Figure 6.6.1

6-6 Example 5
(Application to Rotation of Axes in 3-Space)

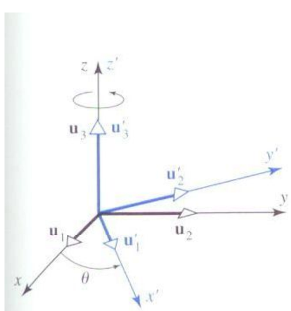


Figure 6.6.2