Lecture 23: 6.1 Inner Products

Wei-Ta Chu

2008/12/17

Definition

- An inner product on a real vector space V is 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.

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.

Properties of Euclidean Inner Product

- Theorem 4.1.2
  - If **u**, **v** and **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
  - $(3\mathbf{u} + 2\mathbf{v}) \cdot (4\mathbf{u} + \mathbf{v})$
    $= (3\mathbf{u}) \cdot (4\mathbf{u} + \mathbf{v}) + (2\mathbf{v}) \cdot (4\mathbf{u} + \mathbf{v})$
    $= (3\mathbf{u}) \cdot (4\mathbf{u}) + (3\mathbf{u}) \cdot \mathbf{v} + (2\mathbf{v}) \cdot (4\mathbf{u}) + (2\mathbf{v}) \cdot \mathbf{v}$
    $=12(\mathbf{u} \cdot \mathbf{u}) + 11(\mathbf{u} \cdot \mathbf{v}) + 2(\mathbf{v} \cdot \mathbf{v})$

Euclidean Inner Product vs. Inner Product

- The Euclidean inner product is the most important inner product on R^n. However, there are various applications in which it is desirable to modify the Euclidean inner product by weighting its terms differently.
- More precisely, if $w_1, w_2, \ldots, w_n$ are positive real numbers, and if **u**=$(u_1, u_2, \ldots, u_n)$ and **v**=$(v_1, v_2, \ldots, v_n)$ are vectors in R^n, then it can be shown

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

which is called *weighted Euclidean inner product with weights* $w_1, w_2, \ldots, w_n$ …

Example

- Suppose that some physical experiment can produce any of n possible numerical values $x_1, x_2, \ldots, x_n$.
- We perform m repetitions of the experiment and yield these values with various frequencies; that is, x_1 occurs f_1 times, x_2 occurs f_2 times, and so forth. $f_1+f_2+\ldots+f_n=m$.
- Thus, the **arithmetic average**, or **mean**, is

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + \cdots + f_nx_n}{f_1 + f_2 + \cdots + f_n} = \frac{1}{m}(f_1x_1 + f_2x_2 + \cdots + f_nx_n)$$

Example

$$\bar{x} = \frac{f_1 x_1 + f_2 x_2 + \cdots + f_n x_n}{f_1 + f_2 + \cdots + f_n} = \frac{1}{m}(f_1 x_1 + f_2 x_2 + \cdots + f_n x_n)$$

- If we let **f**=(f_1,f_2,…,f_n), **x**=(x_1,x_2,…,x_n), w_1=w_2=…=w_n=1/m
- Then this equation can be expressed as the weighted inner product

$$\bar{x} = \langle \boldsymbol{f}, \boldsymbol{x} \rangle = w_1 f_1 x_1 + w_2 f_2 x_2 + \cdots + w_n f_n x_n$$

Weighted Euclidean Product

- 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:
  - Note first that if **u** and **v** are interchanged in this equation, the right side remains the same. Therefore, $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.
  - If $\mathbf{w} = (w_1, w_2)$, then

$\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = 3(u_1+v_1)w_1 + 2(u_2+v_2)w_2 =$

$(3u_1w_1 + 2u_2w_2) + (3v_1w_1 + 2v_2w_2) = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$

which establishes the second axiom.

Weighted Euclidean Product

- $\langle k\mathbf{u}, \mathbf{v}\rangle = 3(ku_1)v_1 + 2(ku_2)v_2 = k(3u_1v_1 + 2u_2v_2) = k\langle \mathbf{u}, \mathbf{v}\rangle$ which establishes the third axiom.
- $\langle \mathbf{v}, \mathbf{v}\rangle = 3v_1v_1 + 2v_2v_2 = 3v_1^2 + 2v_2^2$ .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 $\mathbf{v} = (v_1, v_2) = \mathbf{0}$. Thus, the fourth axiom is satisfied.

Definition

- 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

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

- If a vector has norm 1, then we say that it is a ***unit vector***.

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

$$\|\boldsymbol{u}\| = \langle \boldsymbol{u}, \boldsymbol{u} \rangle^{1/2} = (\boldsymbol{u} \cdot \boldsymbol{u})^{1/2} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

$$d(\mathbf{u},\ \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = < \mathbf{u} - \mathbf{v},\ \mathbf{u} - \mathbf{v} >^{1/2} = \left[(\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v})\right]^{1/2}$$

$$= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \ldots + (u_n - v_n)^2}$$

Weighted Euclidean Inner Product

- The norm and distance depend on the inner product used.
  - If the inner product is changed, then the norms and distances between vectors also change.
  - For example, for the vectors **u** = (1,0) and **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

$$\|\mathbf{u}\| = \langle \mathbf{u},\mathbf{u}\rangle^{1/2} = (3\cdot 1\cdot 1 + 2\cdot 0\cdot 0)^{1/2} = \sqrt{3}$$

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

Unit Circles and Spheres in Inner Product Space

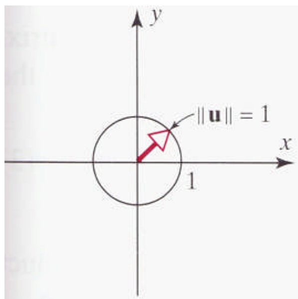
- If V is an inner product space, then the set of points in V that satisfy

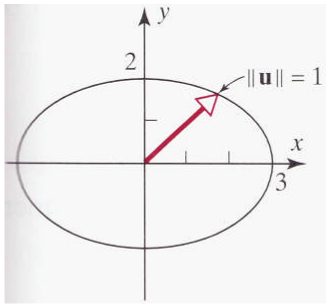
$$||\mathbf{u}|| = 1$$

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.

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 $x^2/9 + y^2/4 = 1$.





Inner Products 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.
- Recalling that 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,

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

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

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

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

Inner Product Generated by the Identity Matrix

- 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

$$A = \begin{bmatrix} \sqrt{w_1} & 0 & 0 \cdots & 0 \\ 0 & \sqrt{w_2} & 0 \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \cdots & \sqrt{w_n} \end{bmatrix}$$

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
  $$\langle U, V\rangle = \text{tr}(U^T V) = \text{tr}(V^T U) = u_1v_1 + u_2v_2 + u_3v_3 + u_4v_4$$
  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 $\langle U, V\rangle = 1(-1)+2(0) + 3(3) + 4(2) = 16$
- 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 $u_1{}^2 + u_2{}^2 + u_3{}^2 + u_4{}^2 = 1$

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, then the following formula defines 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 **p** relative to this inner product is

$$\|\mathbf{p}\| = <\mathbf{p},\ \mathbf{p}>^{1/2} = \sqrt{a_0^2 + a_1^2 + a_2^2}$$

  and the unit sphere in this space consists of all polynomials **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 **u**, **v**, and **w** are vectors in a real inner product space, and *k* is any scalar, then:
  - $\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle = 0$
  - $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
  - $\langle \mathbf{u}, k\mathbf{v} \rangle = k \langle \mathbf{u}, \mathbf{v} \rangle$
  - $\langle \mathbf{u} - \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle - \langle \mathbf{v}, \mathbf{w} \rangle$
  - $\langle \mathbf{u}, \mathbf{v} - \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \langle \mathbf{u}, \mathbf{w} \rangle$

Example

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

Example

- We are guaranteed without any further proof that the five properties given in Theorem 6.1.1 are true for the inner product on R^n generated by any matrix A.
- $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = (\mathbf{v}+\mathbf{w})^T A^{\mathrm{T}} A\mathbf{u}$

$=(\mathbf{v}^T+\mathbf{w}^T)A^T A\mathbf{u}$ [Property of transpose]

$=(\mathbf{v}^T A^T A\mathbf{u}) + (\mathbf{w}^T A^T A\mathbf{u})$ [Property of matrix multiplication]

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

Lecture 23: 6.2
Angle and Orthogonality

Wei-Ta Chu

2008/12/17

Theorem 6.2.1

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

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

  - The inequality can be written in the following two forms

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle^2 \leq \langle \boldsymbol{u}, \boldsymbol{u} \rangle \langle \boldsymbol{v}, \boldsymbol{v} \rangle$$

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle^2 \leq \|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2$$

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

Theorem 6.2.2

- 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 :
    - $\|\mathbf{u}\| \geq 0$
    - $\|\mathbf{u}\| = 0$ if and only if $\mathbf{u} = \mathbf{0}$
    - $\|k\mathbf{u}\| = |k|\,\|\mathbf{u}\|$
    - $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ (Triangle inequality)
- Proof of (d)

$$
\begin{aligned}
&\|\boldsymbol{u}+\boldsymbol{v}\|^2 = \langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}\rangle \\
&= \langle \boldsymbol{u}, \boldsymbol{u}\rangle + 2\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \langle \boldsymbol{v}, \boldsymbol{v}\rangle \\
&\leq \langle \boldsymbol{u}, \boldsymbol{u}\rangle + 2|\boldsymbol{u}, \boldsymbol{v}| + \langle \boldsymbol{v}, \boldsymbol{v}\rangle && \text{(Property of absolute value)} \\
&\leq \langle \boldsymbol{u}, \boldsymbol{u}\rangle + 2\|\boldsymbol{u}\|\|\boldsymbol{v}\| + \langle \boldsymbol{v}, \boldsymbol{v}\rangle && \text{(Theorem 6.2.1)} \\
&= \|\boldsymbol{u}\|^2 + 2\|\boldsymbol{u}\|\|\boldsymbol{v}\| + \|\boldsymbol{v}\|^2 \\
&= (\|\boldsymbol{u}\| + \|\boldsymbol{v}\|)^2
\end{aligned}
$$

Theorem 6.2.3

- 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)

Angle Between Vectors

- Cauchy-Schwarz inequality can be used to define angles in general inner product cases.

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle^2 \leq \|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2$$

$$\Rightarrow \left[\frac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}{\|\boldsymbol{u}\| \|\boldsymbol{v}\|} \right]^2 \leq 1 \qquad \Rightarrow -1 \leq \frac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}{\|\boldsymbol{u}\| \|\boldsymbol{v}\|} \leq 1$$

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

- We define θ to be the angle between **u** and **v**.

Example

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

$$\|\boldsymbol{u}\| = \sqrt{30} \qquad \|\boldsymbol{v}\| = \sqrt{18} \qquad \langle \boldsymbol{u}, \boldsymbol{v} \rangle = -9$$

$$\cos\theta = \frac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|} = -\frac{9}{\sqrt{30}\sqrt{18}} = -\frac{3}{2\sqrt{15}}$$

Orthogonality

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

- Example ($\langle U, V \rangle = \text{tr}(U^T V) = \text{tr}(V^T U) = u_1 v_1 + u_2 v_2 + u_3 v_3 + u_4 v_4$)
  - 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 $\langle U, V \rangle = 1(0) + 0(2) + 1(0) + 1(0) = 0$.

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

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

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

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

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

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

- Proof

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

Example

- 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:

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

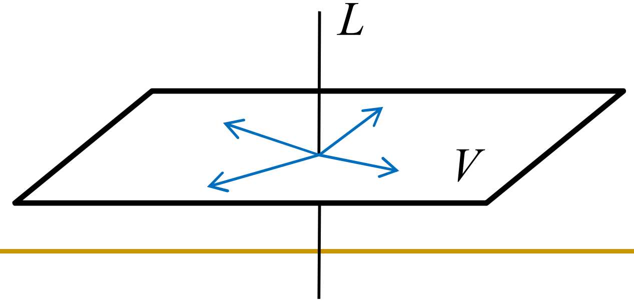
- We can check this result by direct integration:

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

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

Orthogonality

- Definition
  - 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 (正交補餘) of W.
  - If V is a plane through the origin of R^3 with Euclidean inner product, then the set of all vectors that are orthogonal to every vector in V forms the line L through the origin that is perpendicular to V.



Theorem 6.2.5

- 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. (read "W perp")
    - 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 , $(W^\perp)^\perp = W$.

Proof of Theorem 6.2.5(a)

- Note first that $\langle \mathbf{0}, \mathbf{w} \rangle = \mathbf{0}$ for every vector **w** in *W*, so $W^{\perp}$ contains at least the zero vector.
- We want to show that the sum of two vectors in $W^{\perp}$ is orthogonal to every vector in *W* (closed under addition) and that any scalar multiple of a vector in $W^{\perp}$ is orthogonal to every vector in *W* (closed under scalar multiplication).

Proof of Theorem 6.2.5(a)

- Let **u** and **v** be any vector in $W^{\perp}$, let k be any scalar, and let **w** be any vector in W. Then from the definition of $W^{\perp}$, we have $\langle \mathbf{u}, \mathbf{w}\rangle=0$ and $\langle \mathbf{v}, \mathbf{w}\rangle=0$.
- Using the basic properties of the inner product, we have

$$\langle \mathbf{u}+\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle = 0 + 0 = 0$$

$$\langle k\mathbf{u}, \mathbf{w}\rangle = k\langle \mathbf{u}, \mathbf{w}\rangle = k(0) = 0$$

- Which proves that **u**+**v** and k**u** are in $W^{\perp}$.

Proof of Theorem 6.2.5(b)

The only vector common to W and $W^{\perp}$ is $\mathbf{0}$; that is ,$W \cap W^{\perp} = \mathbf{0}$.

- If $\mathbf{v}$ is common to W and $W^{\perp}$, then $\langle \mathbf{v}, \mathbf{v} \rangle=0$, which implies that $\mathbf{v}=\mathbf{0}$ by Axiom 4 for inner products.

Theorem 6.2.6

- 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.

Proof of Theorem 6.2.6(a)

- We must show that if a vector **v** is orthogonal to every vector in the row space, then $A\mathbf{v}=\mathbf{0}$, and conversely, that if $A\mathbf{v}=\mathbf{0}$, then **v** is orthogonal to every vector in the row space.
- Assume that **v** is orthogonal to every vector in the row space of A. Then in particular, v is orthogonal to the row vectors $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m$ of A; that is

$$\boldsymbol{r}_1 \cdot \boldsymbol{v} = \boldsymbol{r}_2 \cdot \boldsymbol{v} = \cdots = \boldsymbol{r}_m \cdot \boldsymbol{v} = 0$$

Proof of Theorem 6.2.6(a)

$$\boldsymbol{r}_1 \cdot \boldsymbol{v} = \boldsymbol{r}_2 \cdot \boldsymbol{v} = \cdots = \boldsymbol{r}_m \cdot \boldsymbol{v} = 0$$

- The linear system $A\mathbf{x}=\mathbf{0}$ can be expressed in dot product notation as

$$\begin{bmatrix} \boldsymbol{r}_1 \cdot \boldsymbol{x} \\ \boldsymbol{r}_2 \cdot \boldsymbol{x} \\ \vdots \\ \boldsymbol{r}_m \cdot \boldsymbol{x} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- And it follows that **v** is a solution of this system and hence lies in the nullspace of A.

Proof of Theorem 6.2.6(a)

- Conversely, assume that **v** is a vector in the nullspace of A, so A**v**=**0**. It follows that $\boldsymbol{r}_1 \cdot \boldsymbol{v} = \boldsymbol{r}_2 \cdot \boldsymbol{v} = \cdots = \boldsymbol{r}_m \cdot \boldsymbol{v} = 0$
- But if **r** is any vector in the row space of A, then **r** is expressible as a linear combination of the row vectors of A, say $\boldsymbol{r} = c_1\boldsymbol{r}_1 + c_2\boldsymbol{r}_2 + \cdots + c_m\boldsymbol{r}_m$
- Thus $$\begin{aligned} \boldsymbol{r} \cdot \boldsymbol{v} &= (c_1\boldsymbol{r}_1 + c_2\boldsymbol{r}_2 + \cdots + c_m\boldsymbol{r}_m) \cdot \boldsymbol{v} \\ &= c_1(\boldsymbol{r}_1 \cdot \boldsymbol{v}) + c_2(\boldsymbol{r}_2 \cdot \boldsymbol{v}) + \cdots + c_m(\boldsymbol{r}_m \cdot \boldsymbol{v}) \\ &= 0 + 0 + \cdots + 0 = 0 \end{aligned}$$
- Which proves that **v** is orthogonal to every vector in the row space of A.

Example (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

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

  - By Theorem 6.2.6, the nullspace of A is the orthogonal complement of W.
  - In Example 4 of Section 5.5 we showed that

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

  form a basis for this nullspace.
  - Thus, vectors $\mathbf{v}_1$ = (-1, 1, 0, 0, 0) and $\mathbf{v}_2$ = (-1, 0, -1, 0, 1) form a basis for the orthogonal complement of W.

Remarks

- In any inner product space V, the zero space $\{\mathbf{0}\}$ and the entire space V are orthogonal complements.
- If A is an $n \times n$ matrix, to say that $A\mathbf{x}=\mathbf{0}$ has only the trivial solution is equivalent to saying that the orthogonal complement of the nullspace of A is all of R^n, or equivalently, that the row space of A is all of R^n.

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 $\{\mathbf{0}\}$.