Linear Algebra

Sartaj Ul Hasan

Department of Mathematics
Indian Institute of Technology Jammu
Jammu, India - 181221

Email: sartaj.hasan@iitjammu.ac.in

Lecture 37

(Nov 06, 2019)

Inner Products

Inner Products

- **Definition:** An inner product on a (real) vector space V is a function, that to each pair of vectors u and v in V associates a scalar (real number) $< u, v >$ and satisfies the following axioms for all vectors u, v, w in V and all scalars c:
  1. $< u, v >=< v, u >$
  2. $< u + v, w >=< u, w > + < v, w >$
  3. $< cu, v >= c < u, v >$
  4. $< u, u > \geq 0$ and $< u, u >= 0$ if and only if $u = 0$
- **Definition**: A vector space with an inner product is called an **inner product space**.
- **Note:** The above definition holds for real inner products. For complex inner products, the first axiom above becomes: $< u, v >= \overline{< v, u >}$ (in other words, the complex conjugate)

Standard Example of an Inner Product

- **Definition:** If we regard vectors u, v in $\mathbb{R}^n$ as $n \times 1$ matrices, then the transpose u^T is a $1 \times n$ matrix. Thus the matrix product $u^T v$ is a 1×1 matrix which we regard as a real number. This real number is called the **inner product** or **dot product**, written as $u.v$.
- If $u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, then
  $u.v = u^T v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$.

Inner Products in $\mathbb{R}^2$ and $\mathbb{R}^3$

In $\mathbb{R}^2$ and $\mathbb{R}^3$, we can adopt another definition for the inner product. Given $u = u_1 e_1 + u_2 e_2 + u_3 e_3$ and $v = v_1 e_1 + v_2 e_2 + v_3 e_3$, we can define inner product in two ways as follows:

$$< u, v >= u^T v = u_1 v_1 + u_2 v_2 + u_3 v_3 \tag{1}$$

$$< u, v >= |u| \ \ |v| \cos\theta \tag{2}$$

where θ is the angle between the vectors.

It is left as an exercise to visualize that the two definitions are indeed equivalent [Hint: Use cosine rule].

Another Example of an Inner Product Space

- The space $\mathbb{R}_n[t]$ of all polynomials of degree less than or equal to n can be made into an inner product space in the following way. Let $t_0, t_1, t_2, \ldots, t_n$ be distinct real numbers (Note: there are $n+1$ numbers). For any two polynomials p and q in $\mathbb{R}_n[t]$, we define:

$$< p, q >= p(t_0)q(t_0) + p(t_1)q(t_1) + \cdots + p(t_n)q(t_n)$$

- **Remark:** It can be verified that the four axioms for an inner product hold with the above definition (exercise !).
- The above inner product for polynomials is used when the values at specific points are important (interpolation problems).

Yet Another Example of an Inner Product Space

- The space $C[a, b]$ of all continuous functions on the closed interval [a,b] can be made into an inner product space with the following definition:
$$< f, g >= \int_a^b f(t)g(t)dt$$
- **Remark:** Again, it can be verified that the four axioms for an inner product hold with the above definition (exercise !).
- The above inner product plays a very important role in the study of continuous functions and their applications in signals and systems.

Length and Distance in Inner Product Spaces

- **Definition:** The length or norm of any vector u in V is the nonnegative number $||u|| = \sqrt{< u, u >}$.
- **Remark:** In the special case of $\mathbb{R}^n$, we get that the length or norm is the nonnegative number

$$||u|| = \sqrt{< u, u >} = \sqrt{u.u} = \sqrt{(u_1{}^2 + u_2{}^2 + \cdots + u_n{}^2)}$$

- This coincides with the usual notion of length from the origin to the point (u_1, u_2) or (u_1, u_2, u_3) in $\mathbb{R}^2$ or $\mathbb{R}^3$. We can easily see that for any scalar c, $||cu|| = |c|||u||$.
- A vector whose length is one is called a unit vector. Given any non-zero vector u, the vector $\frac{u}{||u||}$ has norm one – this step is called **normalizing**.
- The distance between any two vectors u and v in V is defined as dist $(u, v) = ||u - v||$.

Orthogonality

- **Definition:** Two vectors u, v in V are said to be **orthogonal** to each other if $< u, v >= 0$. (Note: the zero vector is orthogonal to every vector – exercise !); the notation for orthogonality is: $u \perp v$
- **Definition:** A set of vectors $\{u_1, u_2, \dots, u_p\}$ is said to be an orthogonal set if any two distinct vectors in the set are orthogonal to each other, in other words if $< u_i, u_j >= 0$ whenever $i \neq j$.
- **Proposition 46:** An orthogonal set of nonzero vectors in V is linearly independent.

Proof: Suppose $S = \{u_1, u_2, \dots, u_p\}$ is an orthogonal set of nonzero vectors and suppose $c_1 u_1 + c_2 u_2 + \cdots + c_p u_p = 0$. Taking the inner product with u_1, we get:

$$c_1 < u_1, u_1 > + c_2 < u_1, u_2 > + \cdots + c_p < u_1, u_p >= 0$$

Since $< u_1, u_i >= 0$ for $i = 2, \dots .p$, this forces $c_1 < u_1, u_1 >= 0$, and since $< u_1, u_1 > > 0$, we get $c_1 = 0$. Similarly, $c_2 = c_3 = \cdots = c_p = 0$. Hence the set S is linearly independent.

Orthogonality (Cont'd)

- **Definition:** If W is a subspace of V, then a vector v is said to be orthogonal to W if v is orthogonal to every vector in W. The set of all vectors orthogonal to W is called the orthogonal complement of W, written $W^{\perp}$.
- **Proposition 47: (a)** A vector v belongs to $W^{\perp}$ if and only if v is orthogonal to every vector in a spanning set for W.
  (b) $W^{\perp}$ is a subspace of V and $W \cap W^{\perp} = \{0\}$.

  Proof of Part (a): $[\implies]$ This implication is ovbious since if $v \in W^{\perp}$, then $v \perp w$ for all $w \in W$.

  $[\Longleftarrow]$ Suppose v is orthogonal to every vector in a spanning set S for W $[W = \text{Span}(S)]$. Let $w \in W$. Then,
  $w = c_1 u_1 + \cdots + c_p u_p, u_i \in S$.

$$
\begin{aligned}
\text{Now,} < v, w > &= < v, c_1 u_1 + \cdots + c_p u_p > \\
&= c_1 < v, u_1 > + \cdots + c_p < v, u_p > \\
&= 0, \quad \text{(by hypothesis)} \implies v \in W^{\perp} \text{ as required.}
\end{aligned}
$$

Proof of Proposition 47 (Cont'd)

- **Proof of Part (b):** Suppose $v_1, v_2 \in W^{\perp}$ and $w \in W$. Then:
  1. Clearly $0 \in W^{\perp}$
  2. $< v_1 + v_2, w > = < v_1, w > + < v_2, w > = 0 + 0 = 0 \implies v_1 + v_2 \in W^{\perp}$.
  3. If $c \in \mathbb{F}$, then $< cv_1, w > = c < v_1, w > = c0 = 0 \implies cv_1 \in W^{\perp}$

  By (i), (ii) and (ii), $W^{\perp}$ is a subspace of V. Finally, suppose $w \in W \cap W^{\perp}$. Then $< w, w > = 0$ and hence by axiom 4 for inner products, $w = 0$.

Orthogonal Bases

- **Definition:** An orthogonal basis for a subspace W is a basis which is also an orthogonal set.
- **Proposition 48:** Let $\{u_1, u_2, \ldots, u_p\}$ be an orthogonal basis for a subspace W. Then if $y = c_1u_1 + c_2u_2 + \cdots + c_pu_p$ is any vector in W, we have: $c_j =< y, u_j > / < u_j, u_j >$ for $j = 1, \ldots, p$.
- **Remark 1:** The above result shows that it is easy to find the coordinates relative to an orthogonal basis, i.e. it is only needed to take an inner product and divide by the inner product of the basis vector with itself.
  Recall that the norm or length of a vector u, i.e. $||u|| = \sqrt{< u, u >}$.
  If it is an orthonormal basis, then the length of each basis vector is 1, and even the step of division is avoided.

Proof of Proposition 48

Suppose $y = c_1 u_1 + \cdots + c_p u_p$. We need to determine the coefficients $c_1, c_2, \ldots, c_p$. Let us take the inner product with u_j, where $1 \leq j \leq p$. Then:

$$\begin{aligned} < y, u_j > &= < c_1 u_1 + \cdots + c_p u_p, u_j > \\ &= c_1 < u_1, u_j > + \cdots + c_p < u_p, u_j > \\ &= c_j < u_j, u_j >, \quad \text{since other terms are 0.} \end{aligned}$$

Therefore, $c_j = \dfrac{< y, u_j >}{< u_j, u_j >}$. Hence, y is given by $y = \displaystyle\sum_{j=1}^{p} \frac{< y, u_j >}{< u_j, u_j >} u_j$.

An example for Proposition 48

Given an orthogonal basis $B = \{v_1, v_2, v_3\}$ of $\mathbb{R}^3$, where

$$v_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, v_3 = \begin{bmatrix} -1 \\ 4 \\ -1 \end{bmatrix}.$$

Find the cooerdinates of $v = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ w.r.t this basis.

First check:

$< v_1, v_2 >= 2-2 = 0, < v_1, v_3 >= -2+4-2 = 0, < v_2, v_3 >= -1+1 = 0.$

Now:

$$c_1 = \frac{< v, v_1 >}{< v_1, v_1 >} = \frac{2+2+6}{9} = \frac{10}{9}$$

$$c_2 = \frac{< v, v_2 >}{< v_2, v_2 >} = \frac{1-3}{2} = \frac{-2}{2} = -1$$

Example (Cont'd)

$$c_3 = \frac{< v, v_3 >}{< v_3, v_3 >} = \frac{-1 + 8 - 3}{18} = \frac{2}{9}$$

Check that:

$$\frac{10}{9}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + (-1)\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + \frac{2}{9}\begin{bmatrix} -1 \\ 4 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{20}{9} - 1 - \frac{2}{9} \\ \frac{10}{9} + \frac{8}{9} \\ \frac{20}{9} + 1 - \frac{2}{9} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

as required.