EE951: Linear Algebra **eMasters**

Chapter 7: Inner Product

Author: Ketan Rajawat

Disclaimer: *These notes are taken from various sources, sometimes without proper citation. Do not distribute outside the class or upload on the Internet.*

Contents

7.1 Introduction

We have already seen that the inner product for vectors $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ is given by

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + u_2v_2 + \ldots + u_nv_n$$

In other words, it is the sum of the products of the corresponding components of the two vectors. For instance, suppose we have two vectors:

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$$

To calculate the inner product $\langle \mathbf{u}, \mathbf{v} \rangle$, we multiply the corresponding components and sum them up:

$$\langle \mathbf{u}, \mathbf{v} \rangle = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 4 + 10 + 18 = 32$$

In the case of one-dimensional vectors, which can be seen as scalar values, the inner product reduces to a simple multiplication.

7.1.1 Properties of Inner Product

For the real field $\mathbb{R}$, the basic properties of the inner product are as follows.

1. *Commutativity:* The inner product of two vectors is commutative, meaning that the order of the vectors does not affect the result. For any vectors $\mathbf{u}$ and $\mathbf{v}$, the commutativity property states:
$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle \tag{7.1}$$
In other words, swapping the order of the vectors in the inner product does not change the value of the inner product.

2. *Associativity ove Scalar Multiplication:* The inner product is associative over scalar multiplication. This means that scalar multiplication can be distributed across the inner product. For any vector $\mathbf{u}$ and scalars a, b, the associativity property states:
$$\langle a\mathbf{u}, b\mathbf{v} \rangle = ab\langle \mathbf{u}, \mathbf{v} \rangle \tag{7.2}$$
In other words, we can pull out the scalars and multiply them together outside the inner product.

3. *Distributivity over Vector Addition:* The inner product is distributive over vector addition. This property states that the inner product of the sum of two vectors is equal to the sum of their individual inner products. For any vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, the distributivity property states:
$$\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle \tag{7.3}$$
In other words, we can distribute the inner product across vector addition.

The associativity and distributitivity properties together imply that for $\mathbb{F} = \mathbb{R}$, the inner product is *bilinear*, i.e., it holds that
$$\langle \alpha\mathbf{u} + \beta\mathbf{y}, \mathbf{v} \rangle = \alpha\langle \mathbf{u}, \mathbf{v} \rangle + \beta\langle \mathbf{y}, \mathbf{v} \rangle \tag{7.4}$$
$$\langle \mathbf{u}, \alpha\mathbf{v} + \beta\mathbf{y} \rangle = \alpha\langle \mathbf{u}, \mathbf{v} \rangle + \beta\langle \mathbf{u}, \mathbf{y} \rangle \tag{7.5}$$
The bilinearity property does not hold for complex fields.

Example 7.1. Let us consider an example to illustrate how the properties of commutativity and distributivity can be combined to calculate the inner product of $(\mathbf{u} + \mathbf{v})$ with $(\mathbf{x} + \mathbf{w})$. Using the distributivity property, we have:

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

Now, applying the distributivity property again to each term, we get:

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

Finally, substituting these results back into the initial equation, we have:

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

In this expression, we have broken down the inner product of $(\mathbf{u} + \mathbf{v})$ with $(\mathbf{x} + \mathbf{w})$ into four separate inner products: $\langle \mathbf{u}, \mathbf{x} \rangle$, $\langle \mathbf{u}, \mathbf{w} \rangle$, $\langle \mathbf{v}, \mathbf{x} \rangle$, and $\langle \mathbf{v}, \mathbf{w} \rangle$. Each of these inner products can be calculated independently using the properties of the inner product.

Consider now the special case where we need to calculate the inner product of $\mathbf{u} + \mathbf{v}$ with itself. Using the earlier formula, we obtain

$$\langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle \tag{7.6}$$

By applying the commutativity property, we can further rearrange the terms within each inner product, so as to yield:

$$\langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle. \tag{7.7}$$

Example 7.2. Consider some simple examples:

1. Multiplication with a standard unit vector can be used to extract a specific element of the vector, i.e., $\langle \mathbf{e}_i, \mathbf{v} \rangle = v_i$.
2. The sum of the elements of a vector $\mathbf{v}$ can be obtained by taking inner product with all-one vector, i.e., $\langle \mathbf{1}, \mathbf{v} \rangle = \sum_i v_i$.
3. The average of the elements of a vector $\mathbf{v}$ can be obtained by taking inner product with all-one vector and dividing by n, i.e., $\langle \mathbf{1}, \mathbf{v} \rangle / n = \langle \frac{1}{n}\mathbf{1}, \mathbf{v} \rangle = \frac{1}{n} \sum_i v_i$.
4. Sum of squares of the elements of a vector $\mathbf{v}$ can be obtained by taking its inner product with itself, i.e., $\langle \mathbf{v}, \mathbf{v} \rangle = \sum_i v_i^2$.

Example 7.3. The matrix-vector product $\mathbf{u} = \mathbf{A}\mathbf{v}$ can be interpreted in terms of the inner product as follows. Denoting the rows of $\mathbf{A}$ by $\mathbf{a}_{j\cdot}$, we see that

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_{1\cdot} \\ \mathbf{a}_{2\cdot} \\ \vdots \\ \mathbf{a}_{n\cdot} \end{bmatrix} \mathbf{v} \tag{7.8}$$

and therefore,

$$u_j = \mathbf{a}_{j\cdot}\mathbf{v} = \langle \mathbf{a}_{j\cdot}^\top, \mathbf{v} \rangle \tag{7.9}$$

That is, each entry of the resulting vector $\mathbf{u}$ represents the inner product between the vector $\mathbf{v}$ and the corresponding row of the matrix $\mathbf{A}$. The row vector needs to be transposed so that it becomes a column vector which can be used within the inner product operation.

7.1.2 Block vectors

The inner product can likewise be defined for block vectors with conformal sizes. Consider two block vectors:

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_k \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_k \end{bmatrix}$$

where $\mathbf{v}_i$ and $\mathbf{w}_i$ represent the blocks of the vectors. Then, we have that

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}_1, \mathbf{w}_1 \rangle + \langle \mathbf{v}_2, \mathbf{w}_2 \rangle + \ldots + \langle \mathbf{v}_k, \mathbf{w}_k \rangle$$

which can be calculated if the sizes of $\mathbf{v}_i$ and $\mathbf{w}_i$ are the same for each i.

7.2 Inner Product and Matrix-Vector Product

The inner product of $\mathbf{y}$ with $\mathbf{Ax}$ is the same as the inner product of $\mathbf{x}$ with $\mathbf{A}^\mathsf{T}\mathbf{y}$. This can be shown as follows. We first write the expression for inner product as $\mathbf{y}^\mathsf{T}\mathbf{Ax}$, which can then be rearranged using the associative property of matrix multiplication:

$$\mathbf{y}^\mathsf{T}\mathbf{Ax} = (\mathbf{y}^\mathsf{T}\mathbf{A})\mathbf{x}$$

Next, we can apply the transpose of a product property (since both sides are scalars):

$$(\mathbf{y}^\mathsf{T}\mathbf{A})\mathbf{x} = \mathbf{x}^\mathsf{T}(\mathbf{y}^\mathsf{T}\mathbf{A})^\mathsf{T}$$

Now, using the properties of transposition, we have:

$$\mathbf{x}^\mathsf{T}(\mathbf{y}^\mathsf{T}\mathbf{A})^\mathsf{T} = \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{y}$$

Finally, recognizing that $\mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{y}$ represents the inner product of $\mathbf{x}$ and $\mathbf{A}^\mathsf{T}\mathbf{y}$, we can conclude that:

$$\langle \mathbf{y}, \mathbf{Ax} \rangle = \langle \mathbf{x}, \mathbf{A}^\mathsf{T}\mathbf{y} \rangle.$$

7.3 Inner Product Examples

7.3.1 Co-occurrence

Inner product can be used to calculate co-occurrence. Consider two sets $\mathcal{A}$ and $\mathcal{B}$ with elements

$$\mathcal{A} = \{1, 3, 4, 6\}$$

$$\mathcal{B} = \{2, 3, 5, 6\}$$

We can represent sets $\mathcal{A}$ and $\mathcal{B}$ using column vectors, where each element in the vector represents the presence or absence of an object in the corresponding set. We assign a value of 1 if the object is in the set and 0 otherwise. The column vector representations of sets $\mathcal{A}$ and $\mathcal{B}$ would be

$$\mathbf{v}_\mathcal{A} := \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \qquad \mathbf{v}_\mathcal{B} := \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \tag{7.10}$$

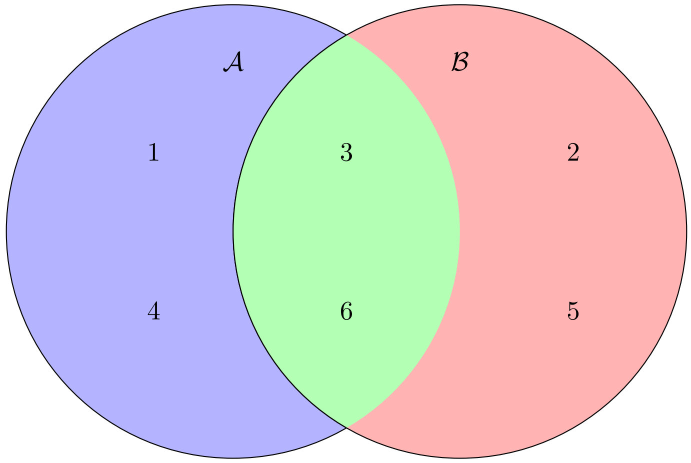

Figure 7.1: Illustration of calculation of cardinality of the intersection

To calculate the cardinality of the intersection of sets $\mathcal{A}$ and $\mathcal{B}$, we can use the inner product of the column vectors $\mathbf{v}_{\mathcal{A}}$ and $\mathbf{v}_{\mathcal{B}}$, so as to yield

$$\langle \mathbf{v}_{\mathcal{A}}, \mathbf{v}_{\mathcal{B}} \rangle = 1(0) + 0(1) + 1(1) + 1(0) + 0(1) + 1(1) = 2 \tag{7.11}$$

Therefore, the cardinality of the intersection of sets A and B is 2. This means that there are 2 objects that are common to both sets A and B. This is illustrated in Fig. 7.1.

7.3.2 Polynomial Evaluation

The inner product can be used to evaluate a general polynomial of degree $n-1$. Let us consider a polynomial given by:

$$p(x) = c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \ldots + c_1 x + c_0$$

We can represent this polynomial in vector form as:

$$\mathbf{p} = \begin{bmatrix} c_{n-1} \\ c_{n-2} \\ \vdots \\ c_1 \\ c_0 \end{bmatrix}$$

Now, let us define another vector:

$$\mathbf{x} = \begin{bmatrix} x^{n-1} \\ x^{n-2} \\ \vdots \\ x \\ 1 \end{bmatrix}$$

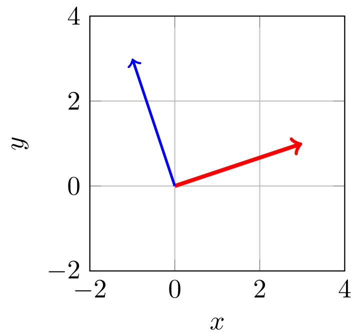

Figure 7.2: Orthogonal displacement vectors.

The entries of $\mathbf{x}$ represent the powers of the variable x in decreasing order, starting from x^{n-1} down to x^1 and finally x^0.

To evaluate the polynomial $p(x)$ for a specific value of x, we can take the inner product of $\mathbf{p}$ and $\mathbf{x}$. Suppose we want to evaluate $p(a)$, then x will take the value of a. We have:

$$\langle \mathbf{p}, \mathbf{x} \rangle = \langle \begin{bmatrix} c_{n-1} \\ c_{n-2} \\ \vdots \\ c_1 \\ c_0 \end{bmatrix}, \begin{bmatrix} a^{n-1} \\ a^{n-2} \\ \vdots \\ a \\ 1 \end{bmatrix} \rangle = c_{n-1}a^{n-1} + c_{n-2}a^{n-2} + \ldots + c_1 a + c_0$$

.

So, the inner product of $\mathbf{p}$ and $\mathbf{x}$ evaluates the polynomial $p(x)$ at the value $x = a$.

7.4 Orthogonal Vectors

Two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{V}$ are said to be orthogonal, if

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^\mathsf{T}\mathbf{v} = 0 \tag{7.12}$$

If the vectors represent displacements in $\mathbb{R}^2$, then it can be seen that orthogonal vectors represent displacements that are perpendicular or at 90° to each other. As an example, consider

$$\mathbf{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

To check if they are orthogonal, we calculate their inner product:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^\mathsf{T}\mathbf{v} = \begin{bmatrix} 3 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = (3 \cdot -1) + (1 \cdot 3) = 0$$

Since the inner product is zero, the vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal to each other. Visually, this means their displacements are perpendicular or at a 90-degree angle to each other, as depicted in Fig. 7.2.

7.5 Matrix Inner Product

The matrix inner product extends the idea of inner product for vectors. Given two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same size, the inner product is denoted as $\langle \mathbf{A}, \mathbf{B} \rangle$. The inner product is calculated by performing element-wise multiplication between corresponding elements of the matrices and summing up the results. Mathematically, it can be expressed as:

$$\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} B_{ij}$$

For example, consider the following matrices:

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} \tag{7.13}$$

The inner product of $\mathbf{A}$ and $\mathbf{B}$ is calculated as:

$$\langle \mathbf{A}, \mathbf{B} \rangle = (2 \cdot 1) + (4 \cdot (-1)) + (1 \cdot 0) + (3 \cdot 2) = 2 - 4 + 0 + 6 = 4$$

As with vectors, the matrix inner product is also commutative, i.e., $\langle \mathbf{A}, \mathbf{B} \rangle = \langle \mathbf{B}, \mathbf{A} \rangle$ and linear, i.e., $\langle \alpha \mathbf{A}, \mathbf{B} \rangle = \alpha \langle \mathbf{A}, \mathbf{B} \rangle$ and $\langle \mathbf{A} + \mathbf{C}, \mathbf{B} \rangle = \langle \mathbf{A}, \mathbf{B} \rangle + \langle \mathbf{C}, \mathbf{B} \rangle$.

The inner product can also be compactly represented using matrix-matrix multiplication and trace as follows:

$$\langle \mathbf{A}, \mathbf{B} \rangle = \operatorname{tr}\left(\mathbf{A}^{\mathsf{T}} \mathbf{B}\right) \tag{7.14}$$

which may be verified by expanding the two sides. Recall that the $\operatorname{vec}()$ operator reshapes the matrix into a column vector. Then we have that

$$\langle \mathbf{A}, \mathbf{B} \rangle = \langle \operatorname{vec}(\mathbf{A}), \operatorname{vec}(\mathbf{B}) \rangle. \tag{7.15}$$