Lec 30: Dot product and its properties.

Recall that the dot product of n-vectors

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

is (the real number) $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$. From this definition one can see that

(1) $\mathbf{u} \cdot \mathbf{u} \geq 0$, and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$;

(2) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$;

(3) $\mathbf{u} \cdot (a\mathbf{v} + b\mathbf{w}) = a(\mathbf{u} \cdot \mathbf{v}) + b(\mathbf{u} \cdot \mathbf{w})$.

These three properties will serve for the definition of inner product of vectors in arbitrary vector space. For this reason it is convenient to write $\mathbf{u} \cdot \mathbf{v} = (\mathbf{u}, \mathbf{v})$ (simply another notation). Let's see some consequences of $(1) - -(3)$. Recall that the length of vector $\mathbf{u}$ in $\mathbb{R}^n$ is $\|\mathbf{u}\| = \sqrt{(\mathbf{u}, \mathbf{u})}$. For instance, the length of the vector $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ in $\mathbb{R}^2$ is $\sqrt{3^2 + 4^2} = \sqrt{25} = 5$.

(4) Cauchy-Bunyakovsky-Schwarz inequality (CBS inequality):

$$|(\mathbf{u}, \mathbf{v})| \leq \|\mathbf{u}\| \|\mathbf{v}\|.$$

Let's prove this. We have for any number r:

$$0 \leq (r\mathbf{u} + \mathbf{v}, r\mathbf{u} + \mathbf{v}) = (\mathbf{u}, \mathbf{u}) r^2 + 2(\mathbf{u}, \mathbf{v}) r + (\mathbf{v}, \mathbf{v}) = q(r).$$

The case $(\mathbf{u}, \mathbf{u}) = 0$ is obvious, because then $\mathbf{u} = 0$ and CBS holds: $0 \leq 0$. So, suppose $\mathbf{u} \neq 0$. Then $q(r)$ is a quadratic polynomial with nonpositive discriminant (otherwise it would have two real roots a and b, and for all r between a and b it would be $q(r) < 0$, which is contradictory). The discriminant of $q(r)$ is $4(\mathbf{u}, \mathbf{v})^2 - 4(\mathbf{u}, \mathbf{u})(\mathbf{v}, \mathbf{v})$, and it is ≤ 0 if and only if $(\mathbf{u}, \mathbf{v})^2 \leq (\mathbf{u}, \mathbf{u})(\mathbf{v}, \mathbf{v})$. Taking the square root, we obtain CBS inequality.

Take, for example, $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} \sqrt{3} \\ 1 \end{bmatrix}$. Then $(\mathbf{u}, \mathbf{v}) = 1$, $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 2$. Clearly, $1 = |(\mathbf{u}, \mathbf{v})| \leq \|\mathbf{u}\| \|\mathbf{v}\| = 2$.

By CBS inequality,

$$-1 \leq \frac{(\mathbf{u}, \mathbf{v})}{\|\mathbf{u}\| \|\mathbf{v}\|} \leq 1.$$

Then there is a unique real number $0 \leq \varphi \leq \pi$ such that $\cos \varphi = \frac{(\mathbf{u}, \mathbf{v})}{\|\mathbf{u}\| \|\mathbf{v}\|}$. This number φ is called the *angle* between $\mathbf{u}$ and $\mathbf{v}$. In the above example $\cos \varphi = \frac{1}{2}$, so $\varphi = 60°$.

Another example: take $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$. Then $(\mathbf{u}, \mathbf{v}) = 0$ and $\cos \varphi = 0$. Hence $\varphi = 90°$. This suggests the definition: vectors $\mathbf{u}$ and $\mathbf{v}$ are called *orthogonal*, if $(\mathbf{u}, \mathbf{v}) = 0$.

(5) Triangle inequality:

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

[Try to find the geometrical meaning of this for $\mathbb{R}^2$. Why this inequality is called triangle?] This is a simple consequence of CBS. Indeed, taking squares of both parts of the inequality above, one has

$$(\mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}) \le (\mathbf{u},\mathbf{u}) + 2\|\mathbf{u}\|\|\mathbf{v}\| + (\mathbf{v},\mathbf{v}).$$

The left-hand side is $(\mathbf{u},\mathbf{u}) + 2(\mathbf{u},\mathbf{v}) + (\mathbf{v},\mathbf{v})$, so the inequality becomes

$$(\mathbf{u},\mathbf{v}) \le |\mathbf{u}\|\|\mathbf{v}\|,$$

and the latter follows directly from CBS inequality.

For example, if $\mathbf{u} = \begin{bmatrix}3\\4\end{bmatrix}$, $\mathbf{v} = \begin{bmatrix}4\\3\end{bmatrix}$, then $\mathbf{u}+\mathbf{v} = \begin{bmatrix}7\\7\end{bmatrix}$ and $\|\mathbf{u}+\mathbf{v}\| = 7\sqrt{2}$, $\|\mathbf{u}\| + \|\mathbf{v}\| = 5+5 = 10$. So, according to the triangle inequality $7\sqrt{2} \le 10$, or after squaring $98 \le 100$.

We call a set of vectors *orthogonal*, if each pair of them is orthogonal. There is an important property of such sets:

(6) Orthogonal set of nonzero vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ is linearly independent.

To show this, suppose

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k = \mathbf{0}.$$

Then, for any i:

$$(\mathbf{u}_i, a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_k\mathbf{u}_k) = (\mathbf{u}_i, \mathbf{0}) = 0.$$

The left-hand side is $a_i(\mathbf{u}_i,\mathbf{u}_i)$, by orthogonality. Since $\mathbf{u}_i \ne 0$ by assumption, $(\mathbf{u}_i,\mathbf{u}_i) > 0$. Then $a_i = 0$ (for all i). This means that (6) is true.

In particular, if $k = n$ (i.e. the number of vectors in the orthogonal set equal the dimension), then the orthogonal set form a basis (why?).

Especially useful are orthogonal sets $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ in which $\|\mathbf{u}_i\| = 1$. Such sets are called *orthonormal*. For example, the standard basis in $\mathbb{R}^n$ is orthonormal. There are many orthonormal bases in $\mathbb{R}^n$.

Example 1. Verify that

$$\mathbf{u}_1 = \begin{bmatrix}\frac{2}{3}\\-\frac{2}{3}\\\frac{1}{3}\end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix}\frac{2}{3}\\\frac{1}{3}\\-\frac{2}{3}\end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix}\frac{1}{3}\\\frac{2}{3}\\\frac{2}{3}\end{bmatrix},$$

is an orthonormal basis in $\mathbb{R}^3$.

In general, given a basis $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ of $\mathbb{R}^n$, in order to find the coordinates of a $\mathbf{u}$ in this basis, we need to solve an $n \times n$ linear system. But when S is orthonormal, the coefficients can be found rather easily.

(7) If a basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ of $\mathbb{R}^n$ is orthonormal, then for any vector $\mathbf{u}$:

$$\mathbf{u} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_n\mathbf{u}_n,$$

where $a_i = (\mathbf{u},\mathbf{u}_i)$.

Indeed, $(\mathbf{u},\mathbf{u}_i) = (a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_n\mathbf{u}_n, \mathbf{u}_i) = a_1(\mathbf{u}_1,\mathbf{u}_i) + \cdots + a_n(\mathbf{u}_n,\mathbf{u}_i) = a_i(\mathbf{u}_i,\mathbf{u}_i) = a_i$.

Let's find the coordinates of the vector $\mathbf{u} = \begin{bmatrix}3\\4\\5\end{bmatrix}$ in the basis from example 1. We have

$$a_1 = (\mathbf{u},\mathbf{u}_1) = 3\cdot\frac{2}{3} + 4\cdot(-\frac{2}{3}) + 5\cdot\frac{1}{3} = 1, \;\; a_2 = (\mathbf{u},\mathbf{u}_2) = 0, \;\; a_3 = (\mathbf{u},\mathbf{u}_3) = 7.$$

In other words,

$$\mathbf{u} = \mathbf{u}_1 + 7\mathbf{u}_3.$$