1212: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 6

Euclidean spaces

Informally, a Euclidean space is a vector space with a scalar product. Let us formulate a precise definition. In this lecture, we shall assume that our scalars are real numbers.

Definition 1. A vector space V is said to be a Euclidean space if it is equipped with a bilinear function (scalar product) $V \times V \to \mathbb{R}$, $v_1, v_2 \mapsto (v_1, v_2)$ satisfying the following conditions:

- bilinearity: $(c_1v_1 + c_2v_2, v) = c_1(v_1, v) + c_2(v_2, v)$ and $(v, c_1v_1 + c_2v_2) = c_1(v, v_1) + c_2(v, v_2)$,
- symmetry: $(v_1, v_2) = (v_2, v_1)$ for all v_1, v_2,
- positivity: $(v, v) \geqslant 0$ for all v, and $(v, v) = 0$ only for $v = 0$.

Example 1. Let $V = \mathbb{R}^n$ with the *standard scalar product*

$$\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}\right) = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

All the three properties are trivially true.

Example 2. Let V be the vector space of continuous functions on $[0, 1]$, and

$$(f(t), g(t)) = \int_0^1 f(t)g(t)\,dt.$$

The symmetry is obvious, the bilinearity follows from linearity of the integral, and the positivity follows from the fact that if $\int_0^1 h(t)\,dt = 0$ for a nonnegative continuous function $h(t)$, then $h(t) = 0$.

Lemma 1. *For every scalar product and every basis* $e_1, \ldots, e_n$ *of* V*, we have*

$$(x_1e_1 + \ldots + x_ne_n, y_1e_1 + \ldots + y_ne_n) = \sum_{i,j=1}^{n} a_{ij}x_iy_j,$$

where $a_{ij} = (e_i, e_j)$.

This follows immediately from the bilinearity property of scalar products.

Orthonormal bases

A system of vectors $e_1, \ldots, e_k$ of a Euclidean space V is said to be orthogonal, if it consists of nonzero vectors, which are pairwise orthogonal: $(e_i, e_j) = 0$ for $i \neq j$. An orthogonal system is said to be orthonormal, if all its vectors are of length 1: $(e_i, e_i) = 1$. Note that a basis $e_1, \ldots, e_n$ of V is orthonormal if and only if

$$(x_1e_1 + \ldots + x_ne_n, y_1e_1 + \ldots + y_ne_n) = x_1y_1 + \ldots + x_ny_n.$$

In other words, an orthonormal basis provides us with a system of coordinates that identifies V with $\mathbb{R}^n$ with the standard scalar product.

Lemma 2. *An orthonormal system is linearly independent.*

Proof. Indeed, assuming $c_1 e_1 + \ldots + c_k e_k = 0$, we have

$$0 = (0, e_p) = (c_1 e_1 + \ldots + c_k e_k, e_p) = c_1(e_1, e_p) + \ldots + c_k(e_k, e_p) = c_p(e_p, e_p) = c_p.$$

□

Lemma 3. *Every finite-dimensional Euclidean space has an orthonormal basis.*

Proof. We shall start from some basis $f_1, \ldots, f_n$, and transform it into an orthogonal basis which we then make orthonormal. Namely, we shall prove by induction that there exists a basis $e_1, \ldots, e_{k-1}, f_k, \ldots, f_n$, where the first $(k-1)$ vectors form an orthogonal system and are equal to linear combinations of the first $(k-1)$ vectors of the original basis. For $k = 1$ the statement is empty, so there is nothing to prove. Assume that our statement is proved for some k, and let us show how to deduce it for $k+1$. Let us search for e_k of the form $f_k - a_1 e_1 - \ldots - a_{k-1} e_{k-1}$; this way the condition on linear combinations on the first k vectors of the original basis is automatically satisfied. Conditions $(e_k, e_j) = 0$ for $j = 1, \ldots, k-1$ mean that

$$0 = (f_k - a_1 e_1 - \ldots - a_{k-1} e_{k-1}, e_j) = (f_k, e_j) - a_1(e_1, e_j) - \ldots - a_{k-1}(e_{k-1}, e_j),$$

and the induction hypothesis guarantees that the latter is equal to

$$(f_k, e_j) - a_j(e_j, e_j),$$

so we can put $a_j = \frac{(f_k, e_j)}{(e_j, e_j)}$ for all $j = 1, \ldots, k-1$. Clearly, the linear span of the vectors $e_1, \ldots, e_{k-1}, f_k, \ldots, f_n$ is the same as the linear span of the vectors $e_1, \ldots, e_{k-1}, e_k, f_{k+1}, \ldots, f_n$ (because we can recover the original set back: $f_k = e_k + a_1 e_1 + \ldots + a_{k-1} e_{k-1}$). Therefore, $e_1, \ldots, e_{k-1}, e_k, f_{k+1}, \ldots, f_n$ are n vectors in an n-dimensional vector space that form a spanning set; they also must form a basis.

To complete the proof, we normalise all vectors, replacing each e_k by $\frac{1}{\sqrt{(e_k, e_k)}} e_k$. □

The process described in the proof is called *Gram–Schmidt orthogonalisation procedure.*

Example 3. Consider $V = \mathbb{R}^2$ with the usual scalar product, and the vectors $f_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ and $f_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $f_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. Then the Gram–Schmidt orthogonalisation works as follows:

- at the first step, there are no previous vectors to take care of, so we put $e_1 = f_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$,
- at the second step we alter the vector f_2, replacing it by $e_2 = f_2 - \frac{(e_1, f_2)}{(e_1, e_1)} e_1 = f_2 - \frac{1}{2} e_1 = \begin{pmatrix} 1/2 \\ -1/2 \\ 1 \end{pmatrix}$,
- at the third step we alter the vector f_3, replacing it by $e_3 = f_3 - \frac{(e_1, f_3)}{(e_1, e_1)} e_1 - \frac{(e_2, f_3)}{(e_2, e_2)} e_2 = \begin{pmatrix} -2/3 \\ 2/3 \\ 2/3 \end{pmatrix}$,
- finally, we normalise all the vectors, obtaining

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \frac{\sqrt{2}}{\sqrt{3}} \begin{pmatrix} 1/2 \\ -1/2 \\ 1 \end{pmatrix}, \quad \frac{\sqrt{3}}{2} \begin{pmatrix} -2/3 \\ 2/3 \\ 2/3 \end{pmatrix}.$$