1212: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 7

Lengths and angles, Cauchy–Schwartz inequality

Definition 1. Let V be an Euclidean space. We define the length of a vector v as $|v| = \sqrt{(v,v)}$, and the angle between two nonzero vectors v and w as the only angle α such that $0 \leqslant \alpha \leqslant 180^\circ$ and

$$\cos\alpha = \frac{(v,w)}{|v||w|}.$$

Remark 1. In the case of usual 3D vectors we could *prove* that $(v,w) = |v||w|\cos\alpha$, because we worked with a particular scalar product that was *defined* on $V = \mathbb{R}^3$. Now, the scalar product is a part of the structure, and can be somewhat arbitrary, so we use our intuition from 3D to *define* the angle between two vectors.

Why are angles well defined?

Theorem 1 (Cauchy–Schwartz Inequality)**.** *For any two vectors v, w of a Euclidean space V we have*

$$(v,w)^2 \leqslant (v,v)(w,w),$$

with equality attained if and only if v and w are proportional.

In particular, for nonzero vectors v and w this implies that

$$-1 \leqslant \frac{(v,w)}{|v||w|} \leqslant 1,$$

so the angle α between v and w is well defined.

Proof. If $v = 0$, the inequality states $0 \leqslant 0$, so there is nothing to prove. Otherwise, let us consider the function $f(t) = (tv - w, tv - w)$ defined for a real argument t. Expanding the brackets using the bilinearity and symmetry of scalar products, we obtain

$$f(t) = t^2(v,v) - 2t(v,w) + (w,w),$$

so $f(t)$ is, for fixed v and w, a quadratic polynomial in t whose leading coefficient (v,v) is positive. Also, $f(t)$ assumes non-negative values for all t. This can only happen if the discriminant of $f(t)$ is non-positive, for if it is positive, then $f(t)$ has two distinct roots t_1 and t_2, and we have $f(t) < 0$ for $t_1 < t < t_2$. The discriminant of $f(t)$ is $(2(v,w))^2 - 4(v,v)(w,w) = 4((v,w)^2 - (v,v)(w,w))$, so we conclude that

$$(v,w)^2 \leqslant (v,v)(w,w),$$

as required. The discriminant is zero if and only if $f(t)$ assumes the value 0, and if t_0 is the corresponding value of t, then $t_0 v = w$, so v and w are proportional. □

Orthogonal complements, and orthogonal direct sums

Now that we defined angles, we can in particular make better sense of orthogonality: $(v, w) = 0$ implies that the angle between v and w is equal to $90°$, so v and w are orthogonal in the usual sense.

Definition 2. Let U be a subspace of a Euclidean space V. The set of all vectors v such that $(v, u) = 0$ for all $u \in U$ is called the orthogonal complement of U, and is denoted by $U^\perp$.

Lemma 1. *For every subspace U, $U^\perp$ is also a subspace.*

Proof. This follows immediately from the bilinearity property of scalar products: for example, if $v_1, v_2 \in U^\perp$, then for each $u \in U$ we have $(u, v_1 + v_2) = (u, v_1) + (u, v_2) = 0$. □

Lemma 2. *For every subspace U, we have $U \cap U^\perp = \{0\}$.*

Proof. Indeed, if $u \in U \cap U^\perp$, we have $(u, u) = 0$, so $u = 0$. □

Lemma 3. *For every finite-dimensional subspace $U \subset V$, we have $V = U \oplus U^\perp$. (This justifies the name "orthogonal complement" for $U^\perp$.)*

Proof. Let $e_1, \ldots, e_k$ be an orthonormal basis of U. To prove that the direct sum coincides with V, it is enough to prove $V = U + U^\perp$, or in other words that every vector $v \in V$ can be represented in the form $u + u^\perp$, where $u \in U$, $u^\perp \in U^\perp$. Equivalently, we need to represent v in the form $c_1 e_1 + \ldots + c_k e_k + u^\perp$, where $c_1, \ldots, c_k$ are unknown coefficients. Computing scalar products with e_j for $j = 1, \ldots, k$, we get a system of equations to determine c_i:

$$(c_1 e_1 + \ldots + c_k e_k + u^\perp, e_j) = (v, e_j).$$

Due to orthonormality of our basis and the definition of the orthogonal complement, the left hand side of this equation is c_j. On the other hand, it is easy to see that for every v, the vector

$$v - (v, e_1)e_1 - \ldots, (v, e_k)e_k$$

is orthogonal to all e_j, and so to all vectors from U, and so belongs to $U^\perp$. □

Corollary 1 (Bessel's inequality)**.** *For any vector $v \in V$ and any orthonormal system $e_1, \ldots, e_k$ (not necessarily a basis) we have*

$$(v, v) \geqslant (v, e_1)^2 + \ldots + (v, e_k)^2.$$

Proof. Indeed, we can take $U = \mathrm{span}(e_1, \ldots, e_k)$ and represent $v = u + u^\perp$. Then

$$(v, v) = (u + u^\perp, u + u^\perp) = (u, u) + (u^\perp, u^\perp)$$

because $(u, u^\perp) = 0$, so

$$|v|^2 = |u|^2 + |u^\perp|^2 \geqslant |u|^2 = (u, e_1)^2 + \ldots + (u, e_k)^2 = (v, e_1)^2 + \ldots + (v, e_k)^2.$$

□