MA1112: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 14

Example of Gram–Schmidt orthogonalisation

Example 1. Consider $V = \mathbb{R}^3$ 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

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

We know that every linearly independent system of vectors can be extended to a basis. Let us show that a similar statement holds for orthonormal systems.

Lemma 1. *Every orthonormal system of vectors in an n-dimensional Euclidean space can be included in an orthonormal basis.*

Proof. Indeed, we know that this system is linearly independent. Thus it can be extended to a basis. If we apply the orthogonalisation procedure to this basis, we shall end up with an orthonormal basis containing our system (if the first k vectors in the system are orthonormal, nothing would happen to them during orthogonalisation). $\square$

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.

In principle, two things need to be proved for this "definition" to make sense. We need to establish that angles are well defined, i.e. that $\frac{(v,w)}{|v||w|}$ is the cosine of some angle, and also it would be good to justify why the "length" has various familiar properties. We do the first one here, and the second one you will get to do in your homework.

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. □

Let us outline another proof, illustrating the power of linear algebra. Consider the subspace U of V spanned by v and w. Since it is the span of two vectors, it is of dimension at most two. If its dimension is at most one, then v and w are proportional, and our inequality is manifestly an equality. If the dimension is two, then we can choose an orthonormal basis, thus identifying U with $\mathbb{R}^2$. In $\mathbb{R}^2$, $(v,w) = |v||w|\cos\alpha$, so the inequality is obvious.

Orthogonal complements

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 2. *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 $(u, v_1 + v_2) = (u, v_1) + (u, v_2) = 0$. □