JUSTIFICATIONS FOR THE GRAM-SCHMIDT PROCESS

The purpose of this file is to verify the assertions needed to show that the Gram-Schmidt Process described in `gram-schmidt.pdf` has all the required properties. Specifically, given a basis $v_1, \cdots, v_n$ for a finite-dimensional inner product space V, there is a basis $w_1, \cdots, w_n$ such that the following hold:

(1) For all j between 1 and n, the vectors $v_1, \cdots, v_j$ and $w_1, \cdots, w_j$ span the same subspace of V.

(2) For all j between 1 and n, the vectors $w_1, \cdots, w_j$ form an orthogonal set.

These properties are formulated so that one can construct the vectors w_j by induction, starting with $w_1 = v_1$. If $j - 1 \leq 1$ and $w_1, \cdots, w_{j-1}$ has the required properties, we define a candidate for w_j by the following formula:

$$w_j \quad = \quad v_j \ - \sum_{k<j} \frac{\langle v_j, w_k \rangle}{\langle w_k, w_k \rangle} \, w_k$$

Geometrically, the sum is supposed to represent the perpendicular projection of v_j onto the subspace W_{j-1} spanned by $v_1, \cdots, v_{j-1}$ and $w_1, \cdots, w_{j-1}$, and this will be considered further in `math132notes6C.pdf`, but for now we shall concentrate on the algebraic formula. Completion of the inductive step requires us to verify that the two properties stated above are true for the vectors $w_1, \cdots, w_j$.

One place to start is by checking that w_j is nonzero. But if it is zero, then v_j will be a linear combination of $w_1, \cdots, w_{j-1}$ and hence by the inductive hypothesis it will also be a linear combination of $v_1, \cdots, v_{j-1}$. This does not happen because the vectors $v_1, \cdots, v_j$ are a (subset of a) linearly independent set. Note further that by the formula and the inductive hypothesis, the vector w_j is a linear combination of $v_1, \cdots, v_j$, so that $\mathbf{SPAN}(w_1, \cdots, w_j) \subset \mathbf{SPAN}(v_1, \cdots, v_j)$.

To prove the reverse inclusion, use the induction hypothesis that $\mathbf{SPAN}(w_1, \cdots, w_{j-1}) = \mathbf{SPAN}(v_1, \cdots, v_{j-1})$ and the identity

$$v_j \quad = \quad w_j \ + \sum_{k<j} \frac{\langle v_j, w_k \rangle}{\langle w_k, w_k \rangle} \, w_k$$

to see that $\mathbf{SPAN}(v_1, \cdots, v_j) \subset \mathbf{SPAN}(w_1, \cdots, w_j)$. This completes the verification of property (1).■

We now turn to the second property. By induction we know that the vectors $w_1, \cdots, w_{j-1}$ are pairwise orthogonal, so to complete the inductive step it is only necessary to show that if $m < j$ then $\langle w_j, w_m \rangle = 0$. By the definition of w_j, the inner product $\langle w_j, w_m \rangle$ is equal to

$$\langle v_j \ - \sum_{k<j} \frac{\langle v_j, w_k \rangle}{\langle w_k, w_k \rangle} \, w_k, \ w_m \, \rangle$$

which we can rewrite as follows:

$$\langle v_j, w_m \rangle \ - \sum_{k<j} \frac{\langle v_j, w_k \rangle}{\langle w_k, w_k \rangle} \, \langle w_k, w_m \rangle$$

Observe that most of the terms in the summation will vanish because $\langle w_k, w_m \rangle = 0$ when $k \neq m$. This means that at most one term in the summation is nonzero and we can simplify the displayed expression to another one with only two terms:

$$\langle w_j, w_m \rangle \quad = \quad \langle v_j, w_m \rangle \; - \; \frac{\langle v_j, w_m \rangle}{\langle w_m, w_m \rangle} \, \langle w_m, w_m \rangle$$

The terms $\langle w_m, w_m \rangle$ in the numerator and denominator cancel each other out, and this leaves us with $\langle v_j, w_m \rangle - \langle v_j, w_m \rangle = 0$, proving that w_j is orthogonal to $w_1, \cdots, w_{j-1}$.■

Postscript. Very often one sees a stronger conclusion that there is an *orthonormal* basis of V with the given properties (*i.e.*, the basis elements also have unit length). The existence of such a basis is an immediate consequence of our construction; it is only necessary to take each of the vectors w_j and divide by their respective lengths. One advantage of doing things as in `gram-schmidt.pdf` is that if we start with a subspace W of $\mathbb{R}^n$ and a basis for W such that the coordinates of each basis element are rational, then the orthogonal basis we obtain will also have rational coordinates. This makes computations much easier and less susceptible to errors.■