Finding an orthonormal spanning set

If we are given a finite set of vectors, we can, in principle, derive from it, in an algorithmic way, an orthonormal set of vectors that has the same span [1] as the original set of vectors. The same process works when the set of vectors is an infinite sequence of vectors. This is so because we generate the orthonormal vectors one at a time, making sure that each new one is orthogonal to each of the *preceding* ones. It might happen that one of our vectors is already in the span of the ones that came before it. The process will actually give the vector zero as the "next" one, but we will just discard vectors like that, and go on to the next one.

The Gram-Schmidt process

Suppose we are given vectors $\{x_1, x_2, \ldots, x_n, \ldots\}$, arranged in sequential order. The sequence can be finite or infinite, and there is no requirement that the vectors comprise a linearly independent set. Indeed, there can be repetitions in the list, and the zero vector can be there too. We look thru the list in sequential order, and we discard any zero vectors, until we find the first one that is not zero. Let's suppose that $x_1 \neq 0$. We set v_1, our first "orthogonal vector," equal to x_1: $v_1 := x_1$. We calculate $\lambda_1 := \|v_1\|$. Actually, we should calculate $\lambda_1^2 = \|v_1\|^2$. The first "real" step is the following one, where we work with the next non-zero vector in the list; let's assume that $x_2 \neq 0$:

$$v_2 = x_2 - \langle x_2, v_1\rangle v_1/\lambda_1^2.$$

You might recognize $\langle x_2, v_1\rangle v_1/\lambda_1^2$: it's the projection of x_2 onto the "line" [2] determined by v_1. It turns out that v_2 is perpendicular to v_1. Here are the details:

$$\begin{aligned}
\langle v_2, v_1\rangle &= \langle x_2 - \langle x_2, v_1\rangle v_1/\lambda_1^2, v_1\rangle \\
&= \langle x_2, v_1\rangle - \langle \langle x_2, v_1\rangle v_1/\lambda_1^2, v_1\rangle \\
&= \langle x_2, v_1\rangle - \langle x_2, v_1\rangle \langle v_1/\lambda_1^2, v_1\rangle \\
&= \langle x_2, v_1\rangle - \langle x_2, v_1\rangle \langle v_1, v_1\rangle/\lambda_1^2 \\
&= \langle x_2, v_1\rangle - \langle x_2, v_1\rangle \lambda_1^2/\lambda_1^2 = 0,
\end{aligned}$$

which is what we mean by "perpendicular." It is possible that $v_2 = 0$. In that case, we would not do this last calculation; we would just discard *this* v_2, and construct a new one from the next non-zero vector in our list. To avoid this bookkeeping, let's assume in what follows that all the x_n's are non-zero, and that all the v_n's we construct are non-zero as well. [3]

We can now proceed by induction, using the same calculations as above. Suppose we are given a linearly independent set $\{x_1, x_2, \ldots, x_n, \ldots\}$ of vectors and that we have constructed vectors $\{v_1, v_2, \ldots, v_n\}$ such that:

1) For $1 \le k \le n$, $\quad \operatorname{span}(\{v_1, \ldots, v_k\}) = \operatorname{span}(\{x_1, \ldots, x_k\})$;
2) for $1 \le k \le n$, and $1 \le l \le n$, $\quad \langle v_k, v_l\rangle = 0$ if $k \neq l$.

Then, with $\lambda_k := \langle v_k, v_k\rangle^{1/2}$, the vector

$$v_{n+1} := x_{n+1} - \sum_{k=1}^{n} \langle x_{n+1}, v_k\rangle v_k/\lambda_k^2$$

gives us a set $\{v_1, v_2, \ldots, v_n, v_{n+1}\}$ that satisfies the conditions 1) and 2). Whenever we wish, we can divide each v_k by its norm, λ_k, to obtain vectors $\mathcal{V}_k := v_k/\lambda_k$ that satisfy 1) and 2) *and* 3) : for $1 \le k \le n$, $\quad \|\mathcal{V}_k\| = 1$.

---

[1] The *span* of a set S of vectors is the set of *all* linear combinations that can be formed using vectors in the set S. The span of S, $\operatorname{span}(S)$, is a subspace; we say the set *spans* the subspace. *Important:* a linear combination always involves a *finite* number of vectors! What we intuitively think of as "infinite linear combinations" are actually series, and involve the use of limits. We *do* talk about the span of an infinite set; each vector in the span involves only a finite number of the vectors in the set.

[2] One-dimensional subspace.

[3] This amounts to assuming that our sequence of vectors comprises a linearly independent set.