Tutorial—Gram-Schmidt orthonormalization

The Gram-Schmidt procedure builds a set of orthonormal functions, one by one, from a set of input functions which are neither orthogonal nor normalized. Let the input functions be $f_0(x)$, $f_1(x)$, $f_2(x)$, $\ldots$, and let $\phi_0(x)$, $\phi_1(x)$, $\phi_2(x)$, $\ldots$ be the output orthonormal functions. It is convenient to introduce a compact notation for the integrals involved in the orthonormalization. Let

$$\langle f|\ g\rangle \equiv \int_a^b f(x)g(x)w(x)\,dx\,,$$

where a and b are the interval over which orthonormality is to be imposed, and $w(x)$ is the weight factor involved. Then the Gram-Schmidt procedure is as follows:

1. ϕ_0 is simply f_0, adjusted to normalized scale:

$$\phi_0(x) = \frac{f_0(x)}{\langle f_0\ |\ f_0\rangle^{1/2}}\,.$$

2. ϕ_1 is that linear combination of f_1 and ϕ_0 that is orthogonal to ϕ_0, scaled to be normalized. To enforce the orthogonality, we set

$$\langle \phi_0|\ (f_1 + a_{10}\phi_0)\ \rangle = 0\,,$$

from which we solve for a_{10}. The process is simplified by the fact that ϕ_0 is normalized, so that on expansion of the above, the quantity $\langle \phi_0|\ \phi_0\rangle$ reduces to unity. We get $a_{10} = -\langle \phi_0|\ f_1\rangle$. Then we scale $\chi_1 \equiv f_1 + a_{10}\phi_0$ to make it normalized:

$$\phi_1(x) = \frac{\chi_1(x)}{\langle \chi_1|\chi_1\rangle^{1/2}}.$$

3. To make ϕ_2, we take that linear combination of f_2, ϕ_0, and ϕ_1 that is orthogonal to both ϕ_0 and ϕ_1 by setting

$$\begin{aligned}\langle \phi_0|\ (f_2 + a_{20}\phi_0 + a_{21}\phi_1)\ \rangle &= 0\\ \langle \phi_1|\ (f_2 + a_{20}\phi_0 + a_{21}\phi_1)\ \rangle &= 0\end{aligned}$$

Because ϕ_0 and ϕ_1 have already been made orthogonal to each other and normalized, these equations simplify to

$$\begin{aligned}\langle \phi_0|f_2\rangle + a_{20} &= 0\,,\\ \langle \phi_1|f_2\rangle + a_{21} &= 0\,.\end{aligned}$$

Each of the above equations contains only one unknown, so they are easily solved: $a_{20} = -\langle\phi_0|f_2\rangle$, $a_{21} = -\langle\phi_1|f_2\rangle$. This means that $\chi_2 = f_2 + a_{20}\phi_0 + a_{21}\phi_1$ is orthogonal to both ϕ_0 and ϕ_1, and we simply need to normalize it:

$$\phi_2(x) = \frac{\chi_2(x)}{\langle\chi_2|\chi_2\rangle^{1/2}}\,.$$

The process can be continued for arbitrary numbers of additional functions f_i. Notice that the set of orthonormal functions depends not only on the set of f_i but on the order in which they are used. If we had, for example, taken f_2 first, then f_2 would be one of the orthonormal functions. But because we took it after f_0 and f_1, all the orthonormal functions containing f_2 would also contain admixtures of f_0 and f_1, and there would be no function containing only f_2.

Maple code implementing the Gram-Schmidt process is in a separate link written in plain text so that it can be more easily copied.