GAUSSIAN QUADRATURE MADE COMPARATIVELY SIMPLE, 
AS THESE THINGS GO

1. Setup. We are doing approximate integration on an interval \([a, b]\) which may be infinite in either or both directions. We have a weight function \(w(x)\) on \([a, b]\) with \(w(x) \geq 0\) everywhere, and we assume that all integrals

\[
\int_a^b [x]^n w(x) \, dx < \infty
\]

(so any function of the form \(p(x) \cdot w(x)\), where \(p(x)\) is a polynomial, is absolutely integrable); moreover, we assume that for every finite open interval \((c, d)\) contained in \([a, b]\) the integral of \(w(x)\) over \((c, d)\) is strictly positive. These are the standard standing assumptions in the theory of orthogonal polynomials, so we know that there is a sequence of polynomials \(\{\varphi_0(x), \ldots, \varphi_n(x), \ldots\}\)—which is uniquely determined up to the choice of leading coefficients—such that \(\deg \varphi_k = k\) \((k = 0, 1, \ldots)\), and \(\langle \varphi_j, \varphi_k \rangle = \delta_{jk}\) where the inner product\(^{(1)}\) is given by

\[
\langle f, g \rangle = \int_a^b f(x) g(x) w(x) \, dx.
\]  

We adopt standard names for the leading coefficients and weighted \(L^2\)-norms:

\[
\varphi_k(x) = A_k x^k + \text{lower-order terms} \quad (1.2)
\]

\[
\gamma_k = \langle \varphi_k, \varphi_k \rangle = \int_a^b \varphi_k(x)^2 w(x) \, dx > 0. \quad (1.3)
\]

2. The Basic Formula. Let \(\varphi_n(x)\) be the \(n\)-th orthogonal polynomial as in the setup above, and let \(\{x_1, \ldots, x_n\}\) be its roots (listed in increasing order). By general orthogonal-polynomial theology (see Atkinson, Thm. 4.4, p. 213 ff.) we know that these are distinct and lie in the interval \([a, b]\). For conceptual purposes—although not for computational purposes—we can regard \(p_{n-1}(x)\) as having been produced by the Newton divided-difference method, and we can then add \(n\) additional terms from a Newton basis to \(p_{n-1}(x)\) to produce the polynomial of degree \((n - 1) + n = 2n - 1\) which interpolates \(f(x)\) to second order, i.e., interpolates \(f'(x)\) also, at each of the points \(x_1, \ldots, x_n\). Adding the error term at \(x \in [a, b]\) as well, we can write

\[
f(x) = p_{n-1}(x) + f[x_1, \ldots, x_n, x_1](x - x_1) \cdots (x - x_n) \\
+ f[x_1, \ldots, x_n, x_1, x_2](x - x_1) \cdots (x - x_n)(x - x_1) + \cdots \\
+ f[x_1, \ldots, x_n, x_1, \ldots, x_n](x - x_1) \cdots (x - x_n)(x - x_1) \cdots (x - x_{n-1}) \\
+ f[x_1, \ldots, x_n, x_1, \ldots, x_n, x](x - x_1)^2 \cdots (x - x_n)^2 \quad \text{this is the error term} \}. \quad (2.1)
\]

Now let us multiply both sides of this relation by \(w(x)\) and integrate over the interval \([a, b]\). On the l. h. s., of course, we get the integral

\[
\int_a^b f(x) w(x) \, dx \quad (2.2)
\]

which is the true value of the integral that we presumably want to approximate by the sum of all or most of the terms on the r. h. s. On the r. h. s., the term

\[
\int_a^b p_{n-1}(x) w(x) \, dx \quad (2.3)
\]

\(^{(1)}\) The inner product is indicated in complex notation, but all the cases we consider will involve only real coefficients.
is the result of the “standard general technique for getting approximate integral formulas by integrating polynomial interpolators”: indeed, since we have

\[ p_{n-1}(x) = \sum_{j=1}^{n} f(x_j) \ell_j(x) \]  

(2.4)

where \( \ell_j(x) \) is the \( j \)-th Lagrange “cardinal” polynomial for the nodes \( \{x_1, \ldots, x_n\} \), we know that the value of this integral on the r. h. s. of (2.1) is

\[ \sum_{j=1}^{n} w_j f(x_j), \quad \text{where} \quad w_j = \int_{a}^{b} \ell_j(x) w(x) \, dx. \]  

(2.5)

But now an exceedingly clever thing happens with the integrals of the remaining approximation-terms on the r. h. s. of (2.1). Since every polynomial \( q(x) \) with \( \deg q < n \) can be written as a linear combination of the orthogonal polynomials \( \{\varphi_0(x), \ldots, \varphi_{n-1}(x)\} \), we have \( \langle \varphi_n, q \rangle = 0 \) for every such \( q(x) \). In each of the integrals

\[ f[x_1, \ldots, x_n, x_1, \ldots, x_k] \int_{a}^{b} (x-x_1) \cdots (x-x_n) \cdot \{x-x_1) \cdots (x-x_{k-1}\} w(x) \, dx \]  

(2.6)

the factor \( (x-x_1) \cdots (x-x_n) \)—being the product of all the linear factors corresponding to the roots of \( \varphi_n(x) \)—is exactly \( \varphi_n(x) \) itself up to its leading coefficient: that is, \( (x-x_1) \cdots (x-x_n) = \frac{1}{A_n} \cdot \varphi_n(x) \). On the other hand, for \( k \leq n \) the product in the curly braces \( \{(x-x_1) \cdots (x-x_{k-1} \} = q(x) \) is a polynomial of degree \( < n \). The integral therefore has the form

\[ f[x_1, \ldots, x_n, x_1, \ldots, x_k] \cdot \frac{1}{A_n} \cdot \langle \varphi_n, q \rangle = 0 \]  

(2.7)

because \( \varphi_n(x) \) is orthogonal to every such \( q(x) \). So these terms simply drop out of the integral on the r. h. s.! (This is a just-about-maximal generalization of what we see happen in the midpoint rule and in Simpson’s rule, where we were able to gain an extra single degree of precision by “interpolating \( f(x) \) twice at the midpoint of \([a,b]\) where the integral couldn’t see it” due to parity considerations.) As if this weren’t good enough, the error factor \( (x-x_1)^2 \cdots (x-x_n)^2 \) in the error term is obviously everywhere nonnegative on \([a,b]\), so we can use our usual mean-value theorem argument to write

\[ \int_{a}^{b} f[x_1, \ldots, x_n, x_1, \ldots, x_n, x](x-x_1)^2 \cdots (x-x_n)^2 w(x) \, dx = \]

\[ \frac{f^{(2n)}(\eta)}{(2n)!} \cdot \int_{a}^{b} (x-x_1)^2 \cdots (x-x_n)^2 w(x) \, dx = \frac{f^{(2n)}(\eta)}{(2n)!} \cdot \frac{\gamma_n}{A_n^2} \]  

(2.8)

where we have used the fact that \( (x-x_1)^2 \cdots (x-x_n)^2 = \left( \frac{\varphi_n(x)}{A_n} \right)^2 \) and employed the standard notation \( \gamma_n = \langle \varphi_n, \varphi_n \rangle \). We thus have an approximate integration formula—the Gaussian approximate quadrature formula for \([a,b]\) with weight \( w(x) \)—complete with error term: we have

\[ \int_{a}^{b} f(x) w(x) \, dx = \sum_{j=1}^{n} w_j f(x_j) + \left( \frac{\gamma_n}{A_n} \right)^2 \cdot \frac{f^{(2n)}(\eta)}{(2n)!} \]  

(2.9)

with degree of precision \( 2n \)—even though it uses only \( n \) points—and for which the coefficient in the error is a simple expression in the quantities \( \gamma_n \) and \( A_n \) (which one can look up rather than compute in the most important cases, viz., those in which the \( \{ \varphi_k(x) \}_{k=0}^{\infty} \) are classically understood and tabulated).
3. The Weights \{w_k\}. Of course we have the standard formula

\[ w_k = \int_a^b \ell_k(x)w(x) \, dx \]  

(3.1)

that expresses the weight \(w_k\) that must be assigned to the \(k\)-th root \(x_k\) of \(\varphi_n(x)\); our object now is to find out more about these weights and, if possible, to find out how to calculate them without doing any actual integration.

It’s easy to see that each \(w_k > 0\) (as if Gaussian quadrature weren’t already good enough!). Each \(\ell_k(x)\) is of degree \(n-1\), so \(\ell_k(x)^2\) is of degree \(2n-2\) and therefore Gaussian quadrature is exact when the function \(f(x) = \ell_k(x)^2\). The relation

\[ \ell_k(x_j) = \delta_{jk} \quad \text{or} \quad \ell_k(x_j)^2 = \delta_{jk} \]  

(3.2)

gives—with error term = 0 by the low degree of \(\ell_k(x)^2\)—

\[ \int_a^b \ell_k(x)^2 w(x) \, dx = \sum_{j=1}^n w_j \ell_k(x_j)^2 = \sum_{j=1}^n w_j \delta_{jk} = w_k \]  

(3.3)

and since the l. h. s. of that relation is obviously \(> 0\), we have \(w_k > 0\) for each \(k = 1, \ldots, n\). This keeps computations numerically stable, and also gives us a relation (take \(f(x) = 1\) identically)

\[ \int_a^b w(x) \, dx = \sum_{k=1}^n w_k \]  

(3.4)

that can frequently help in finding the \(w_k\)’s directly, at least when \(n\) is small (= 2, 3, or so). For cases where \(n\) is larger, it is of course desirable to have a closed-form formula for the \(w_k\)’s that can be evaluated without any integration. One way to produce such a formula is to exploit our knowledge of the error term: since a polynomial of degree exactly \(2n\) has a constant \(2n\)-th derivative, the “error term” for Gaussian integration of such a polynomial can be evaluated exactly. Let \(\varphi_{n+1}(x)\) be the next orthogonal polynomial after the polynomial \(\varphi_n(x)\) whose roots are the nodes for the Gaussian rule under consideration. Since the Lagrange-basis polynomials \(\ell_k(x)\) have degree exactly \(n - 1\), they are orthogonal to \(\varphi_{n+1}(x)\):

\[ 0 = \langle \varphi_{n+1}, \ell_k \rangle = \int_a^b \varphi_{n+1}(x) \ell_k(x) w(x) \, dx . \]  

(3.5)

But the integrand “\(f(x)\)” in that integral is the polynomial \(\varphi_{n+1}(x) \cdot \ell_k(x)\), of degree exactly \((n+1)+(n-1) = 2n\). Therefore the integral can be evaluated exactly by applying Gaussian integration, including the error term: with integrand \(f(x) = \varphi_{n+1}(x) \cdot \ell_k(x)\),

\[ 0 = \langle \varphi_{n+1}, \ell_k \rangle = \sum_{j=1}^n \varphi_{n+1}(x_j) \ell_k(x_j) w_j + \frac{\mathcal{f}'(2n)}{(2n)!} \frac{\gamma_n}{A_n^2} \]  

\[ = \varphi_{n+1}(x_k) \cdot w_k + \frac{\mathcal{f}'(2n)}{(2n)!} \cdot \frac{\gamma_n}{A_n^2} , \quad \text{or equivalently} \]

\[ w_k = -\frac{\mathcal{f}'(2n)}{(2n)!} \cdot \frac{\gamma_n}{A_n^2 \cdot \varphi_{n+1}(x_k)} , \]  

(3.6)

where we have indicated no argument in \(\mathcal{f}'(2n)\) because it is a constant. Since \(\frac{\mathcal{f}'(2n)}{(2n)!}\) is exactly the leading coefficient of the \(2n\)-th-degree polynomial \(\varphi_{n+1}(x) \cdot \ell_k(x)\), we shall have evaluated \(w_k\) if we can find that leading coefficient. We have

\[ \varphi_{n+1}(x) = A_{n+1}x^{n+1} + O(x^n) \]

\[ \ell_k(x) = \ell_k'(x) \frac{\varphi_n(x)}{\varphi_n'(x_k)(x - x_k)} \]

\[ = \frac{A_n}{\varphi_n'(x_k)} x^{n-1} + O(x^{n-2}) \]  

(3.7)
and therefore
\[ \varphi_{n+1}(x) \cdot \ell_k(x) = A_{n+1} \cdot \frac{A_n}{\varphi_n(x_k)} x^{2n} + O(x^{2n-1}) . \]  
(3.8)

Finally, therefore, we have the

**Proposition:** The weights for Gaussian quadrature with nodes at the roots \( \{x_1, \ldots, x_n\} \) of \( \varphi_n(x) \) are given by the formulas
\[
w_k = -A_{n+1} \cdot \frac{A_n}{\varphi_n'(x_k)} \cdot \frac{\gamma_n}{A_n^2 \cdot \varphi_{n+1}(x_k)} = -A_{n+1} \cdot \frac{\gamma_n}{\varphi_n'(x_k) \varphi_{n+1}(x_k)} , \quad k = 1, \ldots, n .
\]  
(3.9)

In the case of classical and/or tabulated polynomials, this formula uses “cookbook information” and does not require that any integrals be computed. This is Atkinson’s formula (5.3.11), p. 272; as you see, it isn’t as hard to derive it as he tells you it is on p. 275!

{It is not obvious that these weights are positive, but we had previously established that fact by a different argument. On the other hand, these formulas together with the known positivity of the weights lead to an interesting conclusion about orthogonal polynomials:

**Corollary:** Between any two consecutive roots of \( \varphi_n(x) \) lies a root of \( \varphi_{n+1}(x) \).

**Proof.** The formula for the weights can be rewritten in the form
\[
\varphi_n'(x_k) \cdot \varphi_{n+1}(x_k) = -A_{n+1} \gamma_n \frac{A_n w_k}{A_n^2}
\]  
and clearly the r. h. s. < 0. The roots of orthogonal polynomials are simple, so (assuming \( x_1, \ldots, x_n \) are written in increasing order) the signs alternate in the sequence
\[
\varphi_n'(x_1), \varphi_n'(x_2), \ldots, \varphi_n'(x_n) .
\]
Therefore the signs must also alternate in the sequence
\[
\varphi_{n+1}(x_1), \varphi_{n+1}(x_2), \ldots, \varphi_{n+1}(x_n)
\]
and so \( \varphi_{n+1}(x) \) must take the value zero at some point between each pair of roots \( x_k, x_{k+1} \) of \( \varphi_n(x) \).

Indeed, since each root of \( \varphi_{n+1}(x) \) must also be simple, the number of its roots in each interval between roots of \( \varphi_n(x) \) must be odd. Consequently (count roots!) there can be only one root of \( \varphi_{n+1}(x) \) in each such interval—except possibly there is one such interval containing three roots of \( \varphi_{n+1}(x) \). It is not hard to rule out that possibility—the details are left to the interested reader. This corollary can be useful in situations where one is computing successive \( \varphi_0(x), \varphi_1(x), \ldots \) and their roots: each successive pair of roots of each \( \varphi_n(x) \) furnishes an interval in which a root of the next \( \varphi_{n+1}(x) \) can be found by Newton’s method or some similar technique.

The proposition just proved enables one to compute Gaussian weights from cookbook information when it is available. On rare occasions—for strange choices of the weight function \( w(x) \)—one will have to lift oneself by one’s bootstraps and compute the polynomials \( \varphi_0(x), \ldots, \varphi_n(x) \) by brute force. Then one won’t want to have to generate another polynomial \( \varphi_{n+1}(x) \) just in order to find the Gaussian weights, so the following alternative expression may be useful.

**Proposition:** The weights for Gaussian quadrature with nodes at the zeros of \( \varphi_n(x) \) are given by the formulas
\[
\frac{1}{w_k} = \sum_{j=0}^{n-1} \frac{\varphi_j(x_k)^2}{\gamma_k} , \quad k = 1, \ldots, n .
\]  
(3.11)
Proof. This is most easily formalized as a matrix computation. Define $n$ vectors in $\mathbb{R}^n$ (coordinates indexed by $i = 0, \ldots, n-1$) for $j = 1, \ldots, n$ by

$$
\Phi_j = \begin{bmatrix}
\varphi_0(x_j) \\
\varphi_1(x_j) \\
\vdots \\
\varphi_{n-1}(x_j)
\end{bmatrix}.
$$

(3.12)

Then let $\Phi$ denote the $n \times n$ matrix whose $j$-th column is $\Phi_j$, $j = 1, \ldots, n$. Define two diagonal matrices $W$ and $\Gamma$ by

$$
W = \text{diag}[w_1, \ldots, w_n] \quad \text{and} \quad \Gamma = \text{diag}[\gamma_0, \ldots, \gamma_{n-1}].
$$

(3.13)

Now consider the inner product $\langle \varphi_i, \varphi_j \rangle$ for indices $i, j \leq n - 1$. Because the degree of the polynomial $\varphi_i(x) \cdot \varphi_j(x)$ in the integral defining the inner product

$$
\langle \varphi_i, \varphi_j \rangle = \int_a^b \varphi_i(x) \varphi_j(x) w(x) \, dx
$$

is $\leq (n - 1) + (n - 1) = 2n - 1$, the integral can be evaluated exactly by Gaussian quadrature. Therefore, since we know the values of these inner products—0 when $i \neq j$ and $\gamma_i$ otherwise—

$$
\gamma_i \delta_{ij} = \langle \varphi_i, \varphi_j \rangle = \sum_{k=1}^n \varphi_i(x_k) w_k \varphi_j(x_k).
$$

(3.14)

These equations are valid for all choices of $0 \leq i, j \leq n - 1$. The sums look very much like those encountered in matrix multiplication, and a moment’s thought will show the reader that these $n^2$ scalar equations say the same thing as the single matrix equation $\Gamma = \Phi^T W \Phi$, or

$$
\Phi^T W \Phi = \Gamma.
$$

(3.15)

Since every $\gamma_i > 0$, the r. h. s. of this equation is an invertible matrix. Taking determinants and using the multiplicative property of the determinant says

$$
(det \Phi)(det W)(det \Phi) = det \Gamma > 0
$$

so $\Phi$ and $W$ must also be invertible. It is therefore legitimate to make the calculation

$$
\Phi^{-1} W^{-1} (\Phi^T)^{-1} = \Gamma^{-1}
$$

$$
W^{-1} = \Phi \Gamma^{-1} \Phi^T.
$$

(3.16)

Since the inverses of the diagonal matrices $W$ and $\Gamma$ are just the diagonal matrices with the reciprocals of the entries of the respective matrices on their diagonals, this last matrix equation is equivalent to the $n$ scalar equations

$$
\frac{1}{w_i} = \sum_{k=0}^{n-1} \varphi_k(x_i) \frac{1}{\gamma_k} \varphi_k(x_i)
$$

$$
= \sum_{k=0}^{n-1} \frac{\varphi_k(x_i)^2}{\gamma_k}.
$$

(3.17)

Knowing the l. h. s. is obviously equivalent to knowing $w_i$; the r. h. s. can be computed from information involving only the polynomials $\varphi_0(x), \ldots, \varphi_{n-1}(x)$. {Incidentally, the r. h. s. is obviously positive, so this is another way to see that all the Gaussian weights are positive.}