ALGEBRAIC METHODS
FOR MODIFIED ORTHOGONAL POLYNOMIALS

DAVID GALANT

ABSTRACT. Some algebraic methods are given to implement Uvarov's extended Christoffel theorem. The stability of the algorithms is discussed.

1. INTRODUCTION

An earlier paper [1] gave a variant of Christoffel's theorem in orthogonal polynomial theory that is useful for finding numerically Gaussian quadrature rules for some nonclassical weight functions. Subsequently, the theory was simplified (Golub and Kautsky [4], Kautsky and Golub [6]) and extended from polynomial modification of the weight function to rational modification of the weight function (Gautschi [2]) to cover Uvarov's extended Christoffel theorem (Uvarov [8]). This paper describes some algebraic methods for implementing Uvarov's theorem. The methods use only real arithmetic. Strong plausibility arguments for the stability and instability of the algorithms are given.

2. BACKGROUND

A weight function, $w(t)$, is a nonnegative function with infinitely many points of increase for which the monomial moments $\{\mu_0, \mu_1, \ldots\}$ exist and are finite. Associated with each weight function is a sequence of monic orthogonal polynomials $\{\varphi_j\}$, $j = 1, 2, \ldots$, $\deg(\varphi_j) = j$, that satisfy a three-term recurrence relation of the form

$$(1) \qquad \varphi_j = (t - b_j)\varphi_{j-1} - g_j\varphi_{j-2}$$

with $\varphi_0(t) = 1$ and $\varphi_{-1}(t) = 0$. The g_j are positive.

With $\boldsymbol{\Phi}^T(t) = (\varphi_0(t), \varphi_1(t), \varphi_2(t), \ldots)$, we have

$$(2a) \qquad t\boldsymbol{\Phi}(t) = \mathbf{J}\boldsymbol{\Phi}(t),$$

$$(2b) \qquad \mathbf{J} = \begin{bmatrix} b_1 & 1 & 0 & \cdots \\ g_2 & b_2 & 1 & \cdots \\ 0 & g_3 & b_3 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}.$$

$\mathbf{J}$ is the Jacobi matrix associated with $w(t)$. From (2a),

$$(2c) \qquad Q(t)\boldsymbol{\Phi}(t) = Q(\mathbf{J})\boldsymbol{\Phi}(t)$$

for an arbitrary polynomial $Q(t)$ follows by induction. As a result, using the orthogonality of the φ_j, we have

Theorem. *If $E(t)$ has a uniformly convergent power series on $[a, b]$, $-\infty < a < b < \infty$, or $E(t)$ is an entire function, then*

$$\int_a^b E(t)w(t)\,dt = E(\mathbf{J})_{1,1} \int_a^b w(t)\,dt = \mu_0 E(\mathbf{J})_{1,1}.$$

This result seems to be new, and may be useful for the numerical evaluation of this kind of function and for the derivation of Taylor series for them.

The use of $\mathbf{J}$ simplifies the statement of results considerably; formulae using $\mathbf{J}$ are to be interpreted as expression of all the relations given by the rows. Thus, (2c) expresses each of the polynomials on the left as a sum of the original orthogonal polynomials. Furthermore, calculations involving $\mathbf{J}$ are not particularly taxing. For instance, in the above theorem, only the first column of $E(\mathbf{J})$ is needed. Thus only the first column of powers of $\mathbf{J}$ are required, and since $\mathbf{J}$ is tridiagonal, the calculation of these columns is simple because only three contiguous elements are involved.

3. Main problem

Given $w(t)$ and $\mathbf{J}$, and a rational function $F(t)$ that is positive and without poles in the support of $w(t)$, construct the Jacobi matrix for the weight function

$$(3) \qquad\qquad \tilde{w}(t) = F(t)w(t).$$

The recurrence relation parameters can be calculated using the Stieltjes formulas (Gautschi [2])

$$(4a) \qquad\qquad \tilde{b}_j = \frac{\int t\tilde{\varphi}_{j-1}^2(t)\tilde{w}(t)\,dt}{\int \tilde{\varphi}_{j-1}^2(t)\tilde{w}(t)\,dt},$$

$$(4b) \qquad\qquad \tilde{g}_j = \frac{\int \tilde{\varphi}_{j-1}^2(t)\tilde{w}(t)\,dt}{\int \tilde{\varphi}_{j-2}^2(t)\tilde{w}(t)\,dt}.$$

Algebraic expression. When $F(t)$ is rational, it is unnecessary to use these integral formulas. Observe that if $w(t) = \tilde{w}(t)/Q(t)$, $Q(t)$ a polynomial, then for any integrable function $f(t)$

$$\int f(t)\tilde{w}(t)\,dt = \int Q(t)f(t)w(t)\,dt.$$

From equation (2c) the Stieltjes formulas can be expressed in terms of the matrices $\mathbf{J}$ and $Q(\mathbf{J})$ as:

$$(5a) \qquad \tilde{b}_1 = \frac{[\mathbf{J}Q(\mathbf{J})]_{11}}{[Q(\mathbf{J})]_{11}}, \qquad \tilde{\varphi}_1(t) = (t - \tilde{b}_1),$$

$$(5b) \qquad \tilde{g}_2 = \frac{[\tilde{\varphi}_1^2(\mathbf{J})Q(\mathbf{J})]_{11}}{[Q(\mathbf{J})]_{11}},$$

$$(5c) \qquad \tilde{b}_2 = \frac{[\tilde{\varphi}_1^2(\mathbf{J})\mathbf{J}Q(\mathbf{J})]_{11}}{[\tilde{\varphi}_1^2(\mathbf{J})Q(\mathbf{J})]_{11}}, \qquad \tilde{\varphi}_2(t) = (t - \tilde{b}_2)\tilde{\varphi}_1(t) - \tilde{g}_2\varphi_0(t),$$

$$(5d) \qquad \tilde{g}_3 = \frac{[\varphi_2^2(\mathbf{J})Q(\mathbf{J})]_{11}}{\varphi_1^2(\mathbf{J})Q(\mathbf{J})]_{11}},$$

and so forth.

A more direct approach is:

Theorem 1. *Let $P(t)$ be a polynomial that is positive on the support of $w(t)$. Write $P(\mathbf{J}) = \mathbf{L_e R}$, where $\mathbf{L_e}$ is a lower triangular matrix with unit diagonal and $\mathbf{R}$ is an upper triangular matrix. Then*

(6)
$$P(\widetilde{\mathbf{J}}) = \mathbf{R L_e},$$

where

$$\widetilde{\mathbf{J}} = \begin{bmatrix} \tilde{b}_1 & 1 & 0 & \cdots \\ \tilde{g}_2 & \tilde{b}_2 & 1 & \cdots \\ 0 & \tilde{g}_3 & \tilde{b}_3 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix},$$

is the Jacobi matrix associated with $\tilde{w}(t)$.

This result is in Kautsky and Golub [6]. They use symmetric matrices and orthonormal polynomials, but the difference is minor. With $P(t) = \sigma(t - v)$, and σ chosen to make $P(t)$ positive in the appropriate interval, this result was given by Galant [1], who also gave an algorithm based upon the QD algorithm (Henrici [5]). This latter algorithm was generalized by Gautschi [2], who showed, essentially, that it could be run backwards to handle the case when $Q(t) = \sigma(t - v)$. In our terms, the result is stated as follows:

Theorem (Linear backward method). *Let $Q(t) = \sigma(t - v)$, and assume $\tilde{b}_1$ and $\tilde{g}_2$ are known. Write*

(7)
$$Q(\mathbf{J}) = \mathbf{R L_e},$$

where $\mathbf{R}_{11} = \sigma(\tilde{b}_1 - v)$, $\mathbf{L}_{21} = \tilde{g}_2/\mathbf{R}_{11}$, $\mathbf{R}_{12} = 1$, and $\mathbf{L}_{11} = 1$. Then

(8)
$$Q(\widetilde{\mathbf{J}}) = \mathbf{L_e R}.$$

The proof follows from examining the matrices generated and comparing them with the matrices in Theorem 1, using $P(t) = \sigma(t - v)$. They are identical.

Since every real polynomial can be factored into a product of real linear and real quadratic factors, the general case can be treated by factoring the polynomial and iterating the procedure above, along with

The backward quadratic method. Let $Q(t) = t^2 - 2xt + (x^2 + y^2)$, and assume $\tilde{b}_1$, $\tilde{g}_2$, and $\tilde{g}_3$ are known. Perform the decomposition $Q(\mathbf{J}) = \mathbf{R L_e}$, with

$$\begin{aligned}
\mathbf{R}_{11} &= Q(\tilde{b}_1) + \tilde{g}_2, & \mathbf{L}_{11} &= 1, \\
\mathbf{R}_{12} &= \tilde{b}_1 + \tilde{b}_2 - 2x, & \mathbf{L}_{21} &= \tilde{g}_2 \frac{\mathbf{R}_{12}}{\mathbf{R}_{11}}, \\
\mathbf{R}_{13} &= 1, & \mathbf{L}_{31} &= \frac{\tilde{g}_2 \tilde{g}_3}{\mathbf{R}_{11}}.
\end{aligned}$$

Then

$$Q(\widetilde{\mathbf{J}}) = \mathbf{L_e R}.$$

We have chosen the first row of $\mathbf{L_e}$ and the first column of $\mathbf{R}$ to produce the matrices that are formed in Theorem 1 with the weight function $w(t)$.

Recovering $\widetilde{\mathbf{J}}$ from $Q(\widetilde{\mathbf{J}})$ is direct for a linear factor $\sigma(t - v)$, because $\widetilde{\mathbf{J}} = \frac{1}{\sigma} Q(\widetilde{\mathbf{J}}) + v$. For a quadratic factor, the process is almost as direct. Since $\check{b}_1$ is known, we can extract the diagonal, using

$$\widetilde{\mathbf{J}}_{j+1,j+1} = Q(\widetilde{\mathbf{J}})_{j,j+1} + 2x - \widetilde{\mathbf{J}}_{j,j}.$$

Furthermore, since $\check{g}_2$ (and for that matter $\check{g}_3$) is known and $Q(\widetilde{\mathbf{J}})_{j+2,j} = \check{g}_{j+2} \cdot \check{g}_{j+1}$, the subdiagonal of $\widetilde{\mathbf{J}}$ is given by

$$\widetilde{\mathbf{J}}_{j,j+1} = \frac{Q(\widetilde{\mathbf{J}})_{j,j+2}}{\check{g}_{j+1}}.$$

The explicit factorization of the polynomial $Q(t)$ and many intermediate factorization/recombination steps are required. Performing more work initially can eliminate the need for this explicit knowledge. This leads to:

Method 2 (The backward method). Let $Q(t)$ be a polynomial of degree m. Suppose further that $\check{b}_j$ and $\check{g}_j$, $j = 1, 2, \ldots, m$, are known. Calculate the first row of $\mathbf{R}$ and the first column of $\mathbf{L_e}$ from them and complete the $\mathbf{RL_e}$ decomposition. Recover $\mathbf{J}$ from $Q(\widetilde{\mathbf{J}})$, using

$$Q(\widetilde{\mathbf{J}})_{j+m,j} = \prod_{i=2}^{m+1} \check{g}_{i+j-1}, \qquad Q(\widetilde{\mathbf{J}})_{j,j+m-1} = \sum_{i=j}^{j+m-1} \check{b}_i + a,$$

where a is the coefficient of t^{m-1} in $Q(t)$.

Theorem 1 and Method 2 almost solve the problem. The numerical processes are real, but Method 2 requires data which is not usually available. To compute the matrices $\mathbf{L_e}$ and $\mathbf{R}$, knowledge of the first m of the three-term recurrence relation parameters, which are not known, is required. Probably the only practical method to obtain these is given by the algorithm in Gautschi [3]. However, as is shown below, the matrix decomposition is very unstable and results quickly lose all accuracy.

Practically, we truncate $\mathbf{J}$ and work with a principal minor. However, equation (2a) is now incorrect because the last column of the principal minor of $\mathbf{J}$ must be modified. Neglecting this correction makes the last m rows and columns of the corresponding principal minor of $Q(\widetilde{\mathbf{J}})$ wrong, but the other principal minors are correct. So, if the recurrence relation parameters up to index n are desired, work with the $(n + m) \times (n + m)$ principal minor of $\mathbf{J}$. Note that formation of the first column of $\mathbf{L_e}$ and the first row of $\mathbf{R}$ in the backward method requires the $m \times m$ principal minor of $\mathbf{J}$.

4. STABILITY

Galant [1], Gautschi [2], and Golub and Kautsky [4] have all noted that the forward algorithm is stable and Gautschi has noted that the backward algorithm appears unstable. While the stability of the forward algorithm follows from the stability of the Cholesky decomposition of a positive definite matrix, the reason for the instability of the backward algorithm is less direct. It is sufficient to consider the linear backward method and the quadratic backward method,

which are given in §3. A strong plausibility argument for the inherent instability of the algorithm in many cases can be based on

Theorem (Szegő [7, Chapter 12]). *Let* $w(t)$ *be a weight function on the interval* $[-1, 1]$ *with the property that the function* $f(\theta) = w(\cos\theta)|\sin\theta|$ *is measurable and positive and the integrals* $\int_{-\pi}^{\pi} f(\theta)\, d\theta$ *and* $\int_{-\pi}^{\pi} |\log f(\theta)|\, d\theta$ *are positive and finite. Then the three-term recurrence relation parameters satisfy*

$$\lim_{n\to\infty} b_n = 0, \qquad \lim_{n\to\infty} g_n = \tfrac{1}{4}.$$

By applying the asymptotic behavior which is the result of this theorem to the forward algorithm with a quadratic polynomial, we have

$$\mathbf{R}_{j,j} \approx \frac{1/16}{3/2 - \mathbf{R}_{j-2,j-2}}.$$

Asymptotically, then,

$$\mathbf{R}_{j,j} \approx \frac{(3 + \sqrt{8})}{4}$$

and, consequently,

$$\mathbf{L}_{j+2,j} \approx \frac{(3 - \sqrt{8})}{4}.$$

Forming the elements of $Q(\mathbf{J})$, decomposing the matrix, forming $Q(\mathbf{J})$, and extracting $\mathbf{J}$ are all provably stable; the only errors are roundoff errors on positive numbers and perhaps an additional small error in calculating $\check{g}_n$. However, the backward algorithm is considerably different. With the same choice of Q and $\mathbf{J}$, we have the same asymptotic behavior of the elements of $\mathbf{R}$ and $\mathbf{L_e}$, but the rules are

$$\mathbf{R}_{j,j} = \frac{g_j g_{j-1}}{\mathbf{L}_{j,j-2}}, \qquad \mathbf{L}_{j,j-2} = 1 + g_{j-2} + g_{j-1} - \mathbf{R}_{j-2,j-2}.$$

Here, $\mathbf{L}_{j,j-2}$ is the difference of two nearly equal numbers, and errors are magnified by about $(3 + \sqrt{8})^2$ per double step, or about 0.766 digits per step in the calculation of the successive principal minors of $\widetilde{\mathbf{J}}$. This severe error magnification is the same as when starting from the moments of the weight function, so the method has dubious practical value.

BIBLIOGRAPHY

1. David Galant, *An implementation of Christoffel's theorem in the theory of orthogonal polynomials*, Math. Comp. **25** (1971), 111–113.

2. Walter Gautschi, *An algorithmic implementation of the generalized Christoffel theorem*, Numerical Integration (G. Hämmerlin, ed.), Internat. Ser. Numer. Math., vol. 57, Birkhäuser, Basel, 1982, pp. 89–106.

3. ____, *Minimal solutions of three-term recurrence relations and orthogonal polynomials*, Math. Comp. **36** (1981), 547–554.

4. G. H. Golub and J. Kautsky, *Calculation of Gauss quadratures with multiple free and fixed knots*, Numer. Math. **41** (1983), 147–163.

5. Peter Henrici, *Elements of numerical analysis*, Wiley, 1964.

6. J. Kautsky and G. H. Golub, *On the calculation of Jacobi matrices*, Linear Algebra Appl. **52/53** (1983), 439–455.

7. Gabor Szegő, *Orthogonal polynomials*, Amer. Math. Soc. Colloq. Publ., vol. 23, Amer. Math. Soc., Providence, R. I., 1959.

8. V. B. Uvarov, *On the connection between systems of polynomials orthogonal with respect to different distribution functions*, U.S.S.R. Comput. Math. and Math. Phys. **9** (1969), no. 6, 25–36.

NASA, Ames Research Center, Intelligent Systems Technology Branch, Moffett Field, California 94035-1000