IS THE RECURRENCE RELATION FOR ORTHOGONAL POLYNOMIALS ALWAYS STABLE?*

WALTER GAUTSCHI

Dept. of Computer Sciences, Purdue University, West Lafayette, IN 47907-1398, USA

Abstract.

Attention is drawn to a phenomenon of "pseudostability" in connection with the three-term recurrence relation for discrete orthogonal polynomials. The computational implications of this phenomenon are illustrated in the case of discrete Legendre and Krawtchouk polynomials. The phenomenon also helps to explain a form of instability in Stieltjes's procedure for generating recursion coefficients of discrete orthogonal polynomials.

AMS(MOS) Subject classification: 33-04, 35C50, 39A11, 65D20.

1. It is our experience, and the experience of many others, that the basic three-term recurrence relation for orthogonal polynomials is generally an excellent means of computing these polynomials, both within the interval of orthogonality and outside of it. The same recurrence relation, on the other hand, is known to become unstable if one attempts to use it for computing other solutions, for example, the minimal solution when the argument is outside the interval of orthogonality (cf. [4]), or the Hilbert transform of Jacobi polynomials when one of the Jacobi parameters is large and the argument close to 1 (cf. [8, §4]). Here we wish to point out instances of "pseudostability" in connection with the computation of discrete orthogonal polynomials.

Our discussion sheds new light on a hitherto unexplained phenomenon of instability that afflicts the Stieltjes procedure for generating the recursion coefficients of discrete orthogonal polynomials (cf. [6, §8]).

2. The (monic) orthogonal polynomials $\{\pi_n(x; d\lambda)\}$ corresponding to a positive measure $d\lambda$ on the real line are known to satisfy a three-term recurrence relation

$$(2.1) \qquad y_{k+1} = (x - \alpha_k)y_k - \beta_k y_{k-1}, \qquad k = 0, 1, 2, \ldots,$$

where $\alpha_k = \alpha_k(d\lambda) \in \mathsf{R}$, $\beta_k = \beta_k(d\lambda) > 0$ are coefficients uniquely determined by the measure $d\lambda$. We are interested in the stability of this recurrence relation with respect to initial values y_0, y_1. That is, letting $\{y_n^*\}$ denote the solution of (2.1) corresponding

to slightly perturbed initial values $y_0^* = y_0(1 + \varepsilon_0)$, $y_1^* = y_1(1 + \varepsilon_1)$, we like to know how much y_n^* differs from y_n for values of n larger than 1. This is an elementary exercise in the theory of linear difference equations. The answers is

$$(2.2) \qquad y_n^* - y_n = \frac{(y_0 y_n z_1 - y_0 y_1 z_n)\varepsilon_0 - (y_1 y_n z_0 - y_0 y_1 z_n)\varepsilon_1}{y_0 z_1 - y_1 z_0},$$

where $\{z_n\}$ is an arbitrary solution of (2.1) linearly independent of $\{y_n\}$. The factors multiplying ε_0 and ε_1 on the right of (2.2), or more precisely, their moduli, determine the extent of error amplification in the absolute error $y_n^* - y_n$. Normally, if $y_n \neq 0$, we prefer to consider relative errors $(y_n^* - y_n)/y_n$. Appropriate amplification factors are then given by

$$(2.3) \qquad \omega_n(x) = \begin{cases} \dfrac{|y_0 z_1 - y_0 y_1(z_n/y_n)| + |y_1 z_0 - y_0 y_1(z_n/y_n)|}{|y_0 z_1 - y_1 z_0|} & \text{if } y_n \neq 0, \\[2mm] 2|y_0 y_1 z_n|/|y_0 z_1 - y_1 z_0| & \text{if } y_n = 0. \end{cases}$$

We say that the recurrence relation (2.1) is *unstable* for the solution $\{y_n\}$ if $\omega_n \to \infty$ as $n \to \infty$. In view of (2.3), if $y_n \neq 0$ for n sufficiently large, this is equivalent to $\lim_{n\to\infty} |z_n/y_n| = \infty$, i.e., to $\{y_n\}$ being a minimal solution of (2.1). There are various "backward recurrence" algorithms (see, e.g., [3], [10]) that can be used to compute minimal solutions. A more perfidious predicament (since there are no easy counter-measures) is *pseudostability*; by this we mean that ω_n is uniformly bounded as $n \to \infty$, but the bound is extremely large. We refer to pseudostability also in the case (of particular interest here) where n can assume only a finite number of values, and some of the ω_n are extremely large. (Isolated large values of ω_n may be due to "near zeros", $y_n \approx 0$, and may well be harmless in practice.)

In the case of orthogonal polynomials $y_n = \pi_n(x; d\lambda)$, we have $y_{-1} = 0$, $y_0 = 1$, and we may choose for z_n the solution of (2.1) satisfying $z_{-1} = 1$, $z_0 = 0$. The amplification factor ω_n in (2.3) then simplifies to

$$(2.4) \qquad \omega_n(x) = \begin{cases} \left|1 - \dfrac{y_1}{z_1}\dfrac{z_n}{y_n}\right| + \left|\dfrac{y_1}{z_1}\dfrac{z_n}{y_n}\right| & \text{if } y_n \neq 0, \\[3mm] 2\left|\dfrac{y_1}{z_1} z_n\right| & \text{if } y_n = 0, \end{cases} \qquad y_n = \pi_n(x; d\lambda).$$

The quantities ω_n in (2.3) and (2.4) characterize stability with respect to *initial values* y_0, y_1. A more complete picture of stability is provided by the following stability measure relative to arbitrary *starting values* y_m, y_{m+1}:

$$(2.5) \qquad \omega_{m\to n}(x) = \begin{cases} \dfrac{|y_m z_{m+1} - y_m y_{m+1}(z_n/y_n)| + |y_{m+1} z_m - y_m y_{m+1}(z_n/y_n)|}{|y_m z_{m+1} - y_{m+1} z_m|} & \text{if } y_n \neq 0, \\[3mm] \dfrac{2|y_m y_{m+1} z_n|}{|y_m z_{m+1} - y_{m+1} z_m|} & \text{if } y_n = 0. \end{cases}$$

This number indicates to what extent errors committed at $k = m$ and $k = m + 1$ are amplified at $k = n$. We may have $n \geq m$ or $n < m$; clearly, $\omega_{m \to m} = \omega_{m \to m+1} = 1$ if $y_m y_{m+1} \neq 0$, and $\omega_n = \omega_{0 \to n}$.

3. We now apply the tools of §2 to discrete orthogonal polynomials. Here, $d\lambda = d\lambda_N$ is a discrete Dirac measure

$$(3.1) \qquad d\lambda_N(x) = \sum_{v=1}^{N} \omega_v \delta(x - x_v) \, dx,$$

where

$$(3.2) \qquad x_1 < x_2 < x_3 < \cdots < x_N, \quad \omega_v > 0, \quad v = 1, 2, \ldots, N.$$

In this case there are exactly N orthogonal polynomials, $\pi_k(\cdot, d\lambda_N)$, $k = 0, 1, \ldots, N - 1$, and the same number of associated recursion coefficients $\alpha_k(d\lambda_N)$ and $\beta_k(d\lambda_N)$, $k = 0, 1, \ldots, N - 1$. We present two examples, believed to be representative for a wide class of discrete orthogonal polynomials, exhibiting phenomena of pseudostability. A third example illustrates a case of almost perfect stability. All our computations were done on the Cyber 205, which has machine precisions of 7.11×10^{-15} and 5.05×10^{-29} in single, resp. double precision.

EXAMPLE 3.1. Equally spaced and equally weighted measure $d\lambda_N$: $x_v = -1 + 2(v - 1)/(N - 1)$, $\omega_v = 2/N$, $v = 1, 2, \ldots, N$.

Here, the recursion coefficients are explicitly known:

$$(3.3) \qquad \alpha_k = 0, \quad k = 0, 1, \ldots, N - 1;$$

$$\beta_0 = 2, \quad \beta_k = \left(1 + \frac{1}{N-1}\right)^2 \left(1 - \left(\frac{k}{N}\right)^2\right)\left(4 - \frac{1}{k^2}\right)^{-1}, \quad k = 1, 2, \ldots, N - 1.$$

For fixed k, and $N \to \infty$, they converge to the respective recursion coefficients for monic Legendre polynomials.

It turns out that in this example the recurrence relation (2.1) applied with $x = x_v$ is generally pseudostable, particularly so if $v \ll N/2$ and N is large. (There is of course symmetry with respect to the midpoint of $[x_1, x_N]$.) We illustrate this in Figure 3.1, which depicts the amplification factor $\omega_n(x)$ of (2.4) on a logarithmic scale for $1 \leq n \leq N - 1$, $N = 40$, $x = x_v$, $v = 1, 5, 10, 20$. There is clearly a trend of rapidly increasing $\omega_n(x)$ as n approaches $N - 1$ when x is near the ends of the interval $[x_1, x_N]$. Near the center of the interval, the recurrence is quite stable.

The graphs of Figure 3.1 are also indicative of stability with regard to starting values other than y_0, y_1, as is shown in Table 3.1. (Integers in parentheses denote decimal exponents.) Here, the quantity

$$(3.4) \qquad \Omega(x_v) = \max_{0 \leq m < n \leq N-1} \omega_{m \to n}(x_v)$$
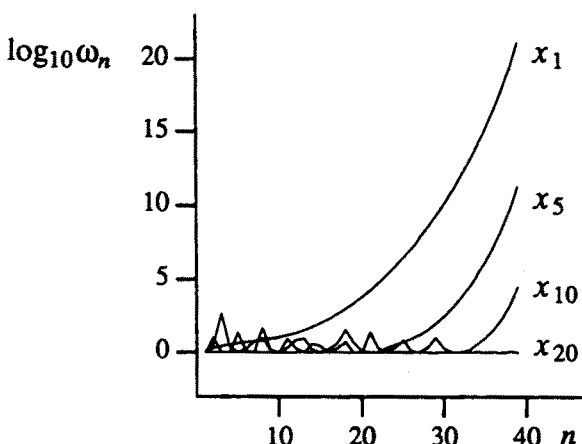
Fig. 3.1. *Amplification factors for Example* 3.1.

is tabulated for selected values of v along with the integers $m = m_v$ and $n = n_v$ for which the maximum in (3.4) is attained, and the maximum relative single-precision error observed in the recurrence.

Table 3.1. *Pseudostability of discrete Legendre polynomials.*

v	$\Omega(x_v)$	m_v	n_v	max err
1	3.771(21)	4	39	1.0310(8)
5	4.148(11)	22	39	3.4959($-$2)
10	6.912(4)	32	39	3.2338($-$8)
20	3.715(0)	25	38	1.1081($-$12)

EXAMPLE 3.2. Krawtchouk polynomials: $x_v = v - 1$, $\omega_v = \binom{N-1}{v-1} p^{v-1} q^{N-v}$,

$v = 1, 2, \ldots, N$, with $p > 0$, $q > 0$, and $p + q = 1$.

Here, too, the recursion coefficients are known explicitly (see, e.g., [1, Eq. (3.5) on p. 161 and Eq. (3.2) on p. 176]),

$$(3.5) \quad \begin{aligned} \alpha_k &= qk + p(N - 1 - k), \quad k = 0, 1, \ldots, N - 1; \\ \beta_0 &= 1, \quad \beta_k = k(N - k)pq, \quad k = 1, 2, \ldots, N - 1. \end{aligned}$$

Figure 3.2 shows severe cases of pseudostability when $p = 0.1$, $q = 0.9$, $N = 40$, and the recurrence formula (2.1) is applied for $x = x_1, x_5, x_{10}$ and x_{20}. Unlike the previous example, Figure 3.2 does not indicate the full extent of pseudostability, especially not in the case $x = x_{20}$. Indeed, the more general stability measure $\omega_{m \to n}$ in (2.5) reveals considerable additional error amplification. This can be seen from
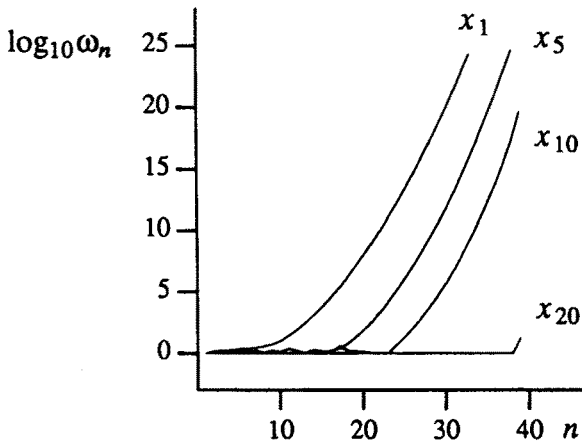
Fig. 3.2. *Amplification factors for Example 3.2.*

Table 3.2, which displays the analogous information as Table 3.1. If p increases, the severity of pseudostability diminishes, the lowest level being attained for $p = q = 1/2$. In this case the quantities in the second and fifth column of Table 3.2 become 7.266(10), 5.797(5), 4.743(2), 5.173(0) and 4.702(-4), 1.217(-8), 1.382(-11), 1.401(-12), respectively.

Table 3.2. *Pseudostability of Krawtchouk polynomials with*
$$p = 0.1, q = 0.9.$$

v	$\Omega(x_v)$	m_v	n_v	max err
1	8.931(25)	4	39	4.859(11)
5	2.053(26)	13	39	4.339(12)
10	5.041(20)	20	39	2.741(7)
20	6.115(8)	30	39	4.995(-5)

The occurrence of pseudostability in Example 3.1 and 3.2 may be due, at least in part, to the equispacing of the abscissae x_v. Choosing as abscissae the Chebyshev points on $[-1, 1]$ indeed may lead to perfectly stable recurrences. This is shown in the next example.

EXAMPLE 3.3. The Fejér measure.
This is the Dirac measure (3.1) underlying the Féjer quadrature rule, i.e.,
$$x_v = \cos\left(\frac{2v - 1}{2N}\right)$$
are the Chebyshev points, and ω_v the Cotes numbers for the corresponding (interpolatory) quadrature rule. The latter are known to be all

positive. This example is of some interest in connection with Stieltjes's procedure (cf. §4).

We computed the recursion coefficients β_k (all $\alpha_k = 0$) in double precision by an orthogonal reduction method using Lanczos's algorithm (cf., e.g., [9], [6, §7]). Applying the recurrence relation (2.1) for each $x = x_\nu$, $\nu = 1, 2, \ldots, N/2$, we then determined (again in double precision) the maximum of all amplification factors in (2.5),

$$(3.6) \qquad \Omega_N = \max_{1 \le \nu \le N/2} \max_{0 \le m < n \le N-1} \omega_{m \to n}(x_\nu).$$

The results are summarized in Table 3.3, where ν_N is the integer ν for which the maximum in (3.6) is attained. In the last column we also show the maximum single-precision error observed. Compared with the previous two examples, the recurrence relation is now remarkably stable.

Table 3.3. *Stability of the recurrence relation for Fejér's measure.*

N	Ω_N	ν_N	max err
20	1.098(2)	2	9.234(-12)
40	1.465(3)	2	2.148(-10)
80	2.958(4)	3	5.554(-9)
160	8.094(4)	21	3.636(-8)

4. Discrete orthogonal polynomials are an important tool in least squares curve fitting. In this context, a common procedure to generate the required recursion coefficients consists in combining the recurrence relation (2.1) with the well-known formulae

$$\alpha_k = \frac{\sum_{\nu=1}^N \omega_\nu x_\nu \pi_k^2(x_\nu)}{\sum_{\nu=1}^N \omega_\nu \pi_k^2(x_\nu)}, \quad k = 0, 1, \ldots, N-1;$$

$$(4.1)$$

$$\beta_0 = \sum_{\nu=1}^N \omega_\nu, \quad \beta_k = \frac{\sum_{\nu=1}^N \omega_\nu \pi_k^2(x_\nu)}{\sum_{\nu=1}^N \omega_\nu \pi_{k-1}^2(x_\nu)}, \quad k = 1, 2, \ldots, N-1.$$

Since $\pi_0 = 1$, one begins by using (4.1) with $k = 0$ to compute α_0, β_0. Then (2.1) is used with $k = 0$ and $x = x_\nu$, $\nu = 1, 2, \ldots, N$, to generate all quantities $\pi_1(x_\nu)$ needed to compute α_1, β_1 from (4.1). Returning to (2.1) with $k = 1$ then yields (for $x = x_\nu$) the quantities $\pi_2(x_\nu)$, which in turn allow us to compute α_2, β_2, etc. In this way, all coefficients $\alpha_k, \beta_k, k = 0, 1, \ldots, N-1$, can be progressively computed, by alternating between (4.1) and (2.1). We have attributed this algorithm to Stieltjes, and called it *Stieltjes's procedure* in [5]. The same procedure has been developed in the 1950's by various authors; see, e.g., Forsythe [2].

Since Stieltjes's procedure relies substantially on the recurrence relation for

discrete orthogonal polynomials, it will necessarily begin to deteriorate, once the recurrence relation starts developing the ill effects of pseudostability. This can be nicely illustrated with the discrete polynomials of Examples 3.1 and 3.2. Using $N = 40, 80, 160$ and 320, we applied Stieltjes's algorithm in single-precision arithmetic and compared the computed coefficients with the known ones in (3.3) and (3.5). The respective relative errors (absolute errors, if $\alpha_k = 0$) are shown in Table 4.1 for Example 3.1. (This is a shortened version of Table 4.1 in [7, §4].) The error growth is not as dramatic as Figure 3.1 would suggest. The reason is that for $x = x_v$ near the endpoints of $[x_1, x_N]$ (where error growth is most severe), the values of the polynomials π_k at $x = x_v$ appearing in (4.1), when k is large, are much smaller than further inside the interval, so that their errors do not contribute as much to the sums in (4.1) as the errors of the more significant terms. Still, there is substantial deterioration of Stieltjes's algorithm after some point (depending on N).[1] The analogous results for Krawtchouk polynomials are shown in Table 4.2 (where err α_k are relative errors).

Table 4.1. *Accuracy of Stieltjes's procedure for Example* 3.1.

N	k	err α_k	err β_k	N	k	err α_k	err β_k
40	≤35	≤1.91(−13)	≤7.78(−13)	160	≤ 76	≤2.98(−13)	≤7.61(−13)
	37	6.93(−11)	3.55(−10)		94	1.25(−4)	1.17(−3)
	39	1.93(−7)	9.58(−7)		112	2.35(−3)	1.16(0)
80	≤53	≤2.04(−13)	≤6.92(−13)	320	≤106	≤8.65(−13)	≤7.39(−13)
	61	3.84(−7)	9.35(−7)		128	2.46(−6)	4.67(−6)
	69	1.87(−1)	6.14(0)		150	1.15(−3)	2.18(−2)

Table 4.2. *Accuracy of Stieltjes's procedure for Example* 3.2.

N	k	err α_k	err β_k	N	k	err α_k	err β_k
40	≤26	≤5.71(−13)	≤5.83(−13)	160	≤ 54	≤8.00(−13)	≤1.29(−12)
	31	3.27(−6)	3.38(−6)		63	4.96(−7)	5.81(−7)
	36	9.63(−2)	5.07(0)		72	2.06(−1)	1.16(0)
80	≤37	≤2.75(−13)	≤7.11(−13)	320	≤ 84	≤9.25(−13)	≤2.52(−12)
	43	1.23(−7)	1.35(−7)		95	4.17(−7)	5.26(−7)
	49	2.41(−1)	3.61(−1)		106	2.00(−1)	6.42(−1)

For the Fejér measure, we compared single-precision results furnished by the Stieltjes procedure with double-precision results produced by the Lanczos algorithm. The maximum (absolute) error in the α's and the maximum (relative) error in the β's are shown in Table 4.3. The results confirm the remarkable stability of Stieltjes's algorithm in this case.

---

[1] This has already been observed in [5, Example 4.1], but was incorrectly attributed to the ill-conditioning of an underlying map, the map H_n of Eq. (3.4) in [5]. (The discussion of the condition of H_n in [5, §3.1] is incomplete inasmuch it does not take into account the dependence of the polynomials π_k on the abscissae τ_v and weights λ_v.)

Table 4.3. *Accuracy of Stieltjes's procedure for Example 3.3.*

N	max err α	max err β
40	1.35(-13)	5.19(-13)
80	2.34(-13)	1.80(-12)
160	5.21(-13)	3.14(-12)
320	5.37(-13)	6.05(-12)

Stieltjes's procedure becomes relevant also in connection with absolutely continuous measures $d\lambda$ if one adopts the following idea (cf. [5, §2.2]). Approximate $d\lambda$ by a discrete measure $d\lambda_N$ such that $\alpha_k(d\lambda_N) \to \alpha_k(d\lambda)$ and $\beta_k(d\lambda_N) \to \beta_k(d\lambda)$ as $N \to \infty$, for fixed k. The discretization $d\lambda \approx d\lambda_N$ can often be accomplished by applying a suitable N-point quadrature rule to the inner product associated with $d\lambda$. (In this connection, Example 3.3 suggests the use of Fejér's quadrature rule as especially appropriate.) Possible occurrences of pseudostability, in such applications, are usually of no concern, since convergence is realized for a value of N that is considerably larger than the maximum value of k for which the α_k, β_k are desired. The onset of pseudostability is thereby avoided; see [6, §8] for a numerical illustration. The same is true in the curve fitting context, where the number of data points, N, is usually much larger than the degree k of the least squares approximant.

REFERENCES

1. T. S. Chihara, *An Introduction to Orthogonal Polynomials*, Gordon and Breach, New York, 1978.
2. G. E. Forsythe, *Generation and use of orthogonal polynomials for data-fitting with a digital computer*, J. Soc. Indust. Appl. Math. 5 (1957), 74–88.
3. W. Gautschi, *Computational aspects of three-term recurrence relations*, SIAM Rev. 9 (1967), 24–82.
4. W. Gautschi, *Minimal solutions of three-term recurrence relations and orthogonal polynomials*, Math. Comp. 36 (1981), 547–554.
5. W. Gautschi, *On generating orthogonal polynomials*, SIAM J. Sci. Statist. Comput. 3 (1982), 289–317.
6. W. Gautschi, *Computational problems and applications of orthogonal polynomials*, in *Orthogonal Polynomials and Their Applications* (C. Brezinski et al., eds.), pp. 61–71, IMACS Annals on Computing and Applied Mathematics, vol. 9, J. C. Baltzer AG, Basel, 1991.
7. W. Gautschi, *Algorithm xxx – ORTHPOL: A package of routines for generating orthogonal polynomials and Gauss-type quadrature rules*, submitted for publication.
8. W. Gautschi and J. Wimp, *Computing the Hilbert transform of a Jacobi weight function*, BIT 27 (1987), 203–215.
9. W. B. Gragg and W. J. Harrod, *The numerically stable reconstruction of Jacobi matrices from spectral data*, Numer. Math. 44 (1984), 317–335.
10. F. W. J. Olver, *Numerical solution of second-order linear difference equations*, J. Res. Nat. Bur. Standards 71B (1967), 111–129.