Modified Chebyshev algorithm for Szegő polynomials: A stability analysis*

E.X.L. de Andrade and **A. Sri Ranga**†
DCCE, IBILCE, UNESP - Universidade Estadual Paulista
Rua Cristóvão Colombo, 2265
15054-000 São José do Rio Preto, SP, Brazil.

1 Introduction

Let $-\infty \leq d_1 < d_2 \leq \infty$ and let ϕ be a non-decreasing function defined on $[d_1, d_2]$, with at least $N+1$ points of increase ($N \geq 1$) and such that the $2N$ associated moments

$$\mu_r^\phi = \int_{d_1}^{d_2} t^r \mathrm{d}\phi(t), \quad r = 0, 1, \ldots, 2N-1,$$

exist. Then it is known that there exists a unique set of polynomials $\{Q_r^\phi\}_{r=0}^N$, defined by

$$Q_r^\phi \text{ is monic and of exact degree r,}$$
$$\int_{d_1}^{d_2} Q_s^\phi(t) Q_r^\phi(t) \mathrm{d}\phi(t) = 0 \quad \text{if} \quad s \neq r.$$

The polynomials Q_r^ϕ (see [3, 8]) are said to be the monic orthogonal polynomials associated with the non-negative measure ϕ. It is also known that these polynomials satisfy the three term recurrence relation

$$Q_r^\phi(x) = (x - \beta_r^\phi) Q_{r-1}^\phi(x) - \alpha_r^\phi Q_{r-2}^\phi(x), \quad (1.1)$$

for $r = 2, \ldots, N$, with $Q_0^\phi(x) = 1$ and $Q_1^\phi(x) = x - \beta_1^\phi$.

In [5, 6], Gautschi provide some insight on the numerical techniques, namely, the Stieltjes method, the Chebyshev algorithm and the modified Chebyshev algorithm, that are used in practice for generating the orthogonal polynomials. These techniques are designed to produce only the coefficients in the associated recurrence relation (1.1). Gautschi, in [5], also points out the various advantages in generating these coefficients as a first step to obtain further results regarding the corresponding orthogonal polynomials. For questions regarding stability of the modified Chebyshev algorithm, we also refer to the more recent paper [1] of Beckermann and Bourreau.

*Title in Portuguese: **Algoritmo modificado de Chebyshev para polinômios de Szegő: Um estudo de estabilidade**.

†These authors research are supported by grants from CNPq and FAPESP.

2 Modified Chebyshev algorithm

The modified Chebyshev algorithm to generate the coefficients α_r^ϕ, $r = 2, \ldots, n$ and β_r^ϕ, $r = 1, 2, \ldots, n$, where $2 \leq n \leq N$, can be given as follows.

Algorithm I

Initialization: Choose a_l, $l = 2, \ldots, 2n-1$ and b_l, $l = 1, 2, \ldots, 2n-1$.

Let
$$P_{-1}(t) = 0, \; P_0(t) = 1,$$
$$P_l(t) = (t - b_l) P_{l-1}(t) - a_l P_{l-2}(t), \quad l = 1, 2, \ldots, 2n-1,$$
$$\sigma_{-1,l} = 0, \quad l = 0, 1, \ldots, 2n-1,$$
$$\sigma_{0,l} = \tilde{\mu}_l^\phi = \int_{d_1}^{d_2} P_l(t) \mathrm{d}\phi(t), \quad l = 0, 1, \ldots, 2n-1$$

and

$$\beta_1^\phi = b_1 + \frac{\sigma_{0,1}}{\sigma_{0,0}}, \qquad \alpha_1^\phi = 0.$$

Continuation: For $k = 1, 2, \ldots, n-1$,

$$\sigma_{k,l} = \sigma_{k-1,l+2} + (b_{k+l+1} - \beta_k^\phi)\sigma_{k-1,l+1} - \alpha_k^\phi \sigma_{k-2,l+2} + a_{k+l+1}\sigma_{k-1,l}\ , \quad l = 0, 1, \ldots, 2n-2k-1$$

and

$$\beta_{k+1}^\phi = b_{k+1} + \frac{\sigma_{k,1}}{\sigma_{k,0}} - \frac{\sigma_{k-1,1}}{\sigma_{k-1,0}},$$
$$\alpha_{k+1}^\phi = \frac{\sigma_{k,0}}{\sigma_{k-1,0}}\ .$$

The quantities $\sigma_{k,l}$ in the algorithm represent

$$\sigma_{k,l} = \int_{d_1}^{d_2} Q_k^\phi(t) P_{k+l}(t) \mathrm{d}\phi(t),$$

for $l = 0, 1, \ldots, 2n-2k-1$ and $k = 0, 1, \ldots, n-1$.

This algorithm is shown to be numerically stable if the coefficients $a_l, l = 1, \ldots, 2n-1$ and

$b_l, l = 0, 1, \ldots, 2n-1$ are chosen adequately and also that the corresponding modified moments $\tilde{\mu}_l^{\phi}$, $l = 0, 1, \ldots, 2n-1$, can be determined with good precision (see [1] and [6]).

With the choice $a_l = b_l = 0$ for $l = 1, 2, \ldots, 2n-1$, the modified moments $\tilde{\mu}_l^{\phi}$ coincide with the ordinary moments μ_l^{ϕ}. Consequently, the above algorithm corresponds to the ordinary Chebyshev algorithm (known to Chebyshev in 1859), which is in general known to be numerically unstable.

The idea of using modified moments in order to obtain a stable algorithm is due to Sack and Donavan [7]. However, the modified Chebyshev algorithm as described above (with a slight modification to the indices) should be attributed to Wheeler [10].

3 Szegő polynomials

Let ν be a positive measure on the unit circle $\mathcal{C}$. This means $\tilde{\nu}(\theta) = \nu(e^{i\theta})$, defined on $[0, 2\pi]$, is a real, bounded and non-decreasing function with infinitely many points of increase. Then the moments

$$M_m^{\nu} = \int_{\mathcal{C}} z^m d\nu(z) = \int_0^{2\pi} e^{im\theta} d\nu(e^{i\theta}),$$

$m = 0, 1, 2, \ldots$, all exist. We consider the Szegő polynomials $\{S_n^{\nu}\}$ associated with the measure ν defined by

$$\int_{\mathcal{C}} S_n^{\nu}(z)\overline{S_m^{\nu}(z)}\, d\nu(z) = 0, \qquad n \neq m.$$

These polynomials were introduced by Szegő (see for example [8] and [9]).

The Szegő polynomials are known to satisfy the system of recurrence relations

$$\begin{aligned} S_{n+1}^{\nu}(z) = {} & zS_n^{\nu}(z) + a_{n+1}^{\nu} S_n^{\nu *}(z), \\ \left(1 - |a_{n+1}^{\nu}|^2\right) zS_n^{\nu}(z) = {} & S_{n+1}^{\nu}(z) - a_{n+1}^{\nu} S_{n+1}^{\nu *}(z), \end{aligned}$$

for $n \geq 0$. Here $S_n^{\nu}(z)$ is the monic Szegő polynomial of degree n and $S_n^{\nu *}(z) = z^n \overline{S_n^{\nu}(1/\overline{z})}$ is the reciprocal polynomial. The numbers $a_n^{\nu} = S_n^{\nu}(0)$, $n \geq 1$, are known as the reflection coefficients of the Szegő polynomials.

We now see how the modified Chebyshev algorithm for real orthogonal polynomials can be used to obtain an algorithm to determine the real Szegő polynomials. Since $d\nu(e^{i\theta})$ is a symmetric measure,

$$M_m^{\nu} = \int_0^{2\pi} e^{im\theta} d\nu(e^{i\theta}) = \int_0^{2\pi} \cos(m\theta) d\nu(e^{i\theta}).$$

Hence using $d\nu(e^{i\theta}) = -d\phi_1(\cos(\theta/2))$ and $x = \cos(\theta/2)$, we obtain

$$M_m^{\nu} = \int_{-1}^{1} T_{2m}(x) d\phi_1(x), \qquad m \geq 0,$$

where T_m are the m-th degree Chebyshev polynomials of the first kind. Hence using the monic Chebyshev polynomials $\hat{T}_m = 2^{-m+1} T_m$, $m \geq 1$, we can define the modified moments $\hat{\mu}_0^{\phi_1} = M_0^{\nu}$,

$$\hat{\mu}_{2m-1}^{\phi_1} = 0, \;\; \hat{\mu}_{2m}^{\phi_1} = 2^{-2m+1} M_m^{\nu}, \;\; m \geq 1,$$

associated with the measure ϕ_1 and apply the modified Chebyshev algorithm (Algorithm I) to obtain the coefficients $\alpha_n^{\phi_1}$. Since

$$\begin{aligned} &\hat{T}_1(x) = x, \qquad \hat{T}_2(x) = x\hat{T}_m(x) - \tfrac{1}{2}\hat{T}_0(x), \\ &\hat{T}_{m+1}(x) = x\hat{T}_m(x) - \tfrac{1}{4}\hat{T}_{m-1}(x), \end{aligned}$$

$m = 2, 3, \ldots, 2n-2$, and since $\sigma_{m,2l+1}$ turn out to be zero for all m, letting $d_{0,0} = \sigma_{0,0}$ and $d_{m,l} = 2^{2(m+l)-1} \sigma_{m,2l}$, we obtain the algorithm

Algorithm II

Initialization: Set

$$d_{0,l} = M_l^{\nu}, \quad l = 0, 1, \ldots, n,$$

$$d_{1,l} = d_{0,l+1} + d_{0,l}, \quad l = 0, 1 \ldots, n-1,$$

$$\Delta_2 = 2\frac{d_{1,0}}{d_{0,0}}.$$

Continuation: For $m = 2, 3, \ldots, n$ do

$$\begin{aligned} d_{m,l} = {} & d_{m-1,l+1} + d_{m-1,l} \\ & -\Delta_m d_{m-2,l+1}, \quad l = 0, 1, \ldots, n-m, \end{aligned}$$

$$\Delta_{m+1} = \frac{d_{m,0}}{d_{m-1,0}}.$$

From this algorithm,

$$\alpha_{m+1}^{\phi_1} = \Delta_{m+1}/4, \quad m = 1, 2, \cdots, n.$$

Having found the values of $\alpha_m^{\phi_1}$, the reflection coefficients a_m^{ν} can be generated by

$$a_1^{\nu} = 1 - 2\alpha_2^{\phi_1}, \;\; a_m^{\nu} = 1 - \frac{4\alpha_{m+1}^{\phi_1}}{1 + a_{m-1}^{\nu}}, \; m = 2, \ldots, n.$$

This algorithm, which requires about n^2 multiplications, has already been tested in [2] with a method to solve the frequency analysis problem. Traditionally, to solve this problem one finds the Szegő polynomials and then determines its roots. It was shown in [2] that one does not need to calculate the Szegő polynomials or even the reflection coefficients a_n^{ν}. The eigenvalue problem associated with the orthogonal polynomials $Q_n^{\phi_1}$, which is much easier to solve than finding all the (complex) roots of Szegő polynomials, leads to a very good solution to the frequency analysis problem. Hence, the simplicity and speed of the method given in [2] could be very important for real time frequency analysis.

4 Stability Analysis

Numerical stability of the modified Chebyshev algorithm has been studied in Gautschi [5, 6] and, more recently, in Beckermann and Bourreau [1]. However, in these papers the analysis is focused on how to choose those polynomials that provide the modified moments.

In the case of our algorithm II, the situation is some what different. The modified moments associated with the measure ϕ_1 within $[-1,1]$ are known, as these are the ordinary moments on the unit circle associated with the measure ν. The polynomials which provide these modified moments are the Chebyshev polynomials. The immediate gain is an algorithm which is simple and easy to implement.

However, can we say some thing about the stability of this algorithm, the principal part of which is an implementation of the map $K : \mathbb{R}^n \to \mathbb{R}^n$ from the vector of modified moments $[d_{0,1}, d_{0,2}, \ldots, d_{0,n}]^T$ to the vector of recurrence coefficients $[\alpha_2^{\phi_1}, \alpha_3^{\phi_1}, \ldots, \alpha_{n+1}^{\phi_1}]^T$? Since the $\alpha_j^{\phi_1}$ are obtained from the ratios $d_{j,0}/d_{j-1,0}$, we will instead look at the map $\tilde{K}$ from $[d_{0,1}, d_{0,2}, \ldots, d_{0,n}]^T$ to $[d_{1,0}, d_{2,0}, \ldots, d_{n,0}]^T$.

For any k such that $k = 1, 2, \ldots, n$, let the k-vectors $\mathbf{x}^{(k)}$ and $\mathbf{d}_0^{(k)}$ be given by

$$\mathbf{x}^{(k)} = \begin{bmatrix} d_{1,n-k} \\ d_{2,n-k} \\ d_{3,n-k} \\ \vdots \\ d_{m,n-k} \end{bmatrix} \quad \text{and} \quad \mathbf{d}_0^{(k)} = \begin{bmatrix} d_{0,n-k} \\ d_{0,n-k} \\ d_{0,n-k} \\ \vdots \\ d_{0,n-k} \end{bmatrix}.$$

Also let the $k+1$-vector $\hat{\mathbf{x}}^{(k)}$ be such that

$$\hat{\mathbf{x}}^{(0)} = d_{0,n} \quad \text{and} \quad \hat{\mathbf{x}}^{(k)} = \begin{bmatrix} d_{0,n-k} \\ \mathbf{x}^{(k)} \end{bmatrix}, \ 1 \le k \le n.$$

We then have the following identities from the equations in algorithm II:

$$\mathbf{x}^{(k)} = \mathbf{H}_k \hat{\mathbf{x}}^{(k-1)} + \mathbf{d}_0^{(k)} \quad \text{and} \quad \hat{\mathbf{x}}^{(k)} = \begin{bmatrix} d_{0,n-k} \\ \mathbf{x}^{(k)} \end{bmatrix},$$

for $k = 1, 2, \ldots, n$, where $\mathbf{H}_1 = 1$ and $\mathbf{H}_k =$

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1-\Delta_2 & 1 & 0 & \cdots & 0 & 0 \\ 1-\Delta_2 & 1-\Delta_3 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1-\Delta_2 & 1-\Delta_3 & 1-\Delta_4 & \cdots & 1 & 0 \\ 1-\Delta_2 & 1-\Delta_3 & 1-\Delta_4 & \cdots & 1-\Delta_k & 1 \end{bmatrix},$$

for $k \ge 2$. The above identities can also be written in the form

$$\mathbf{y}^{(k)} = \mathbf{H}_k \hat{\mathbf{x}}^{(k-1)}, \tag{4.1}$$

$$\hat{\mathbf{x}}^{(k)} = \begin{bmatrix} d_{0,n-k} \\ \mathbf{y}^{(k)} + \mathbf{d}_0^{(k)} \end{bmatrix}, \tag{4.1a}$$

for $k = 1, 2, \ldots, n$. Hence, the map $\tilde{K}$ is achieved by applications of the maps given by (4.1) and (4.1a) for $k = 1, 2, \ldots, n$. Thus, we can study the stability property of the map $\tilde{K}$ by analyzing the maps given by (4.1) and (4.1a).

The instability due to the additions (and/or subtractions) performed in (4.1a) is most critical when $k = n$ since, in order to obtain the Δ_{m+1} in the algorithm, one needs to perform the divisions $d_{m,0}/d_{m-1,0}$, $m = 1, 2, \ldots, n$. However, since $d_{0,0} = M_0^\nu$ and $d_{m,0} = \frac{1}{2} M_0^\nu \Delta_2 \cdots \Delta_{m+1}$, $m = 1, 2, \ldots, n$, one foresees a problem only if the product $\Delta_2 \cdots \Delta_{m+1}$ rapidly goes to zero or infinity as m increases. For example, if $d\phi_1(x) = -d\nu(z)$ is equal to $d\phi^{(T)}(x) = (1-x^2)^{-1/2} dx$ (asociated with the Chebyshev polynomials of the first kind) or equal to $d\phi^{(U)}(x) = (1-x^2)^{1/2} dx$ (associated with the Chebyshev polynomials of the second kind), then we have $\Delta_{m+1} = 1$ for $m = 2, 3, \ldots, n$. Hence the maps given by (4.1a) can be expected to be safe if $d\phi_1$ varies little from $d\phi^{(T)}$ or $d\phi^{(U)}$.

Now the stability behavior of the map in (4.1) can be verified as follows.

Let $\delta \mathbf{y}^{(k)}$ be the perturbation in $\mathbf{y}^{(k)}$ as a result of the perturbations $\delta \mathbf{H}_k$ and $\delta \hat{\mathbf{x}}^{(k-1)}$ in $\mathbf{H}_k$ and $\hat{\mathbf{x}}^{(k-1)}$, respectively. Thus, we easily obtain that

$$\frac{\|\delta \mathbf{y}^{(k)}\|}{\|\mathbf{y}^{(k)}\|} \le c(\mathbf{H}_k) \Big[\frac{\|\delta \mathbf{H}_k\|}{\|\mathbf{H}_k\|} + \big\| I + \delta \mathbf{H}_k \, \mathbf{H}_k^{-1} \big\| \, \frac{\|\delta \hat{\mathbf{x}}^{(k-1)}\|}{\|\hat{\mathbf{x}}^{(k-1)}\|} \Big],$$

where $c(\mathbf{H}_k) = \|\mathbf{H}_k\| \|\mathbf{H}_k^{-1}\|$ is the condition number of $\mathbf{H}_k$. Hence for good stability of (4.1) it is preferable that $c(\mathbf{H}_k)$ does not increase rapidly with k.

The inverse of the matrix $\mathbf{H}_k$ is easily found to be

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ h_{2,1} & 1 & 0 & \cdots & 0 & 0 \\ h_{3,1} & h_{3,2} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ h_{k-1,1} & h_{k-1,2} & h_{k-1,3} & \cdots & 1 & 0 \\ h_{k,1} & h_{k,2} & h_{k,3} & \cdots & h_{k,k-1} & 1 \end{bmatrix},$$

where, for $j = 1, 2, \ldots, k-1$,

$$h_{j+1,j} = -(1 - \Delta_{j+1}),$$
$$h_{i,j} = -(1 - \Delta_{j+1}) \prod_{s=j+2}^{i} \Delta_s, \quad i = j+2, \ldots, k.$$

These results indicate that the condition number $c(\mathbf{H}_k)$ depends on the values of Δ_j, $j = 2, \ldots, k$. Normally, what one can say is $\Delta_{j+1} > 0$ for $j \ge 1$ and $\{\Delta_{j+1}/4\}$ is a chain sequence. This means,

$$0 \le \Delta_{j+1} \le 4, \quad j \ge 1.$$

Clearly if all the Δ_j are closer to one then $c(\mathbf{H}_k)$ also takes a value closer to one. This is the case if $\mathrm{d}\phi_1(x) = -\mathrm{d}\nu(z) = \mathrm{d}\phi^{(U)}(x)$, when we have $\Delta_{j+1} = 1,\ j = 1, 2, \ldots$. Consequently, we can expect (4.1) to be stable if $\mathrm{d}\phi_1$ varies little from $\mathrm{d}\phi^{(U)}(x)$. That is, all the Δ_{j+1} stay closer to 1.

Stability of (4.1) when $\mathrm{d}\phi_1$ varies little from $\mathrm{d}\phi^{(T)}(x)$ can also be established since we have $d_{0,0} \simeq \pi$ and $|d_{0,j}| \leq \epsilon$, $j = 1, 2, \ldots, n$.

References

[1] B. Beckermann and E. Bourreau, How to choose modified moments, *J. Comput. Appl. Math.*, **98** (1998), 81-98.

[2] C.F. Bracciali, X. Li and A. Sri Ranga, Real orthogonal polynomials in frequency analysis, *Math. Comp.*, **74**, 341-362.

[3] T.S. Chihara, *An Introduction to Orthogonal Polynomials*, Mathematics and its Applications Series, New York: Gordon and Breach, 1978.

[4] P. Delsarte and Y. Genin, An introduction to the class of split Levinson algorithms, in Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms (G.H. Golub and P. Van Dooren, eds.), NATO ASI Series F, Vol. 70, pp. 111-130, Springer-Verlag, 1991.

[5] W. Gautschi, On generating orthogonal polynomials, *SIAM J. Sci. Statist. Comput.*, **3** (1982), 289-317.

[6] W. Gautschi, Orthogonal polynomials - Constructive theory and applications *J. Comput. Appl. Math.*, **12** & **13** (1985), 61-76.

[7] R.A. Sack and A.F. Donavan, An algorithm for Gaussian quadrature given modified moments *Numer. Math.* , **18** (1971/72), 465-478.

[8] G. Szegö, *Orthogonal Polynomials*, 4th ed., Amer. Math. Soc. Colloq. Publ., vol. **23**, (Amer. Math. Soc., Providence, RI, 1975).

[9] W. Van Assche, Orthogonal polynomials in the complex plane and on the real line, in Field Institute Communications 14: Special functions, q-series and related topics (M.E.H. Ismail et al., eds.), Amer. Math. Soc. (1997), 211-245.

[10] J.C. Wheeler, Modified moments and Gaussian quadrature, *Rocky Mountain Journal of Maths.*, **4** (1974), 287-296.