Controlled accuracy for discrete Chebyshev polynomials

Albertus C. den Brinker
Data Science
Philips Research
Eindhoven, The Netherlands
ORCID 0000-0002-1679-7574

***Abstract*—An algorithm is proposed for stable determination of the normalized discrete Chebyshev polynomials. This is achieved by adaptively restricting the range over which the recurrence relation or difference equation is applied. The adaptation of the range is controlled by a factor indicating how much the normalized basis functions are allowed to deviate from unit norm.**

***Index Terms*—Orthogonal polynomials, discrete Chebyshev polynomials, numerical accuracy**

I. Introduction

Discrete Chebyshev polynomials are being used in various signal processing areas like texture recognition [1], image recognition [2]–[6], image compression [7]–[11] and speech processing and recognition [12], [13]. A problem that has been noted is the numerical evaluation of these functions [14]. Especially at high orders, numerical inaccuracies may accumulate when determining these polynomials from their recurrence or difference equation. We will show where these numerical issues occur, and propose an effective strategy to deal with them. While the practical cases where both large discrete support and high polynomial order are required may be limited, this paper offers an elegant solution that fits nicely with the underlying theory of normed spaces, as the proposed adaptation is norm-driven.

The strategy enables to calculate the discrete Chebyshev polynomials with an upfront specified accuracy. This is achieved by a control mechanism which adds a computational burden. The paper does not address the computational load or speed of the algorithm as this would involve practically relevant operating points in terms of number of samples, order of the approximation and required performance of the considered application. This is beyond the scope of the current work; for aspects like these we refer to [15], [16].

II. Discrete Chebyshev polynomials

A. Orthogonal polynomials

The discrete Chebyshev polynomials fall in the category of classical orthogonal polynomials, a set of functions sharing various properties. Here we only consider the discrete Chebyshev polynomials and especially their normalized variant. We denote the discrete Chebyshev polynomials as t_n. They are defined on a finite discrete interval of N points denoted by the variable x with $x \in [0, N-1]$. There are N polynomials forming an orthogonal basis in $\mathcal{R}^N$ with n denoting the order of the polynomial and running from 0 to $N-1$. The orthogonality implies

$$\sum_{x=0}^{N-1} t_n(x) t_m(x) = \delta_{nm} H_n^2 \tag{1}$$

where δ_{nm} denotes the Kronecker delta and H_n is the norm of the nth polynomial.

It is often desired to use a set of orthonormal basis functions. These can easily be constructed from t_n and H_n by

$$T_n = t_n / H_n. \tag{2}$$

For obvious reasons, we call these the orthonormal discrete Chebyshev functions.

B. Recurrence relation

Like all classical orthogonal polynomials, the discrete Chebyshev polynomials adhere to a 3-terms recurrence relation, in particular

$$\begin{aligned}(n+1)t_{n+1}(x) &= (2n+1)(2x-N+1)t_n(x) \\ &\quad - n(N^2-n^2)t_{n-1}(x)\end{aligned} \tag{3}$$

with $t_0(x) = 1$ and $t_1(x) = 2x - N + 1$. It is clear that the functions $T_n(x)$ adhere to a similar recursion which can be found by substituting (2) in (3). The explicit expression for the norms is given by

$$\begin{aligned}H_n^2 &= \frac{N(N^2-1^2)(N^2-2^2)\cdots(N^2-n^2)}{2n+1} \\ &= \frac{2n-1}{2n+1}\left(N^2-n^2\right) H_{n-1}^2.\end{aligned} \tag{4}$$

C. Difference equation

The discrete Chebyshev polynomials adhere to a second-order difference equation which can also be expressed as a 3-term recurrence relation in the x variable (i.e., for fixed n), in particular,

$$t_n(x) = \alpha t_n(x-1) + \beta t_n(x-2) \tag{5}$$

for $n \in [0, N-1]$ and $x \in [2, N-1]$. From (2) it is obvious that the exact same relation holds for T_n. The coefficients are given by

$$\alpha = \frac{-n(n+1) - (2x-1)(x-N-1) - x}{x(N-x)}, \quad (6)$$

$$\beta = \frac{(x-1)(x-N-1)}{x(N-x)}. \quad (7)$$

D. Numerical and computational issues

Numerical evaluation of the orthonormal functions can be done in various ways. One can start from explicit expressions of $t_n(0)$ and $t_n(1)$ (or $T_n(0)$ and $T_n(1)$) and use (5) to evaluate the polynomials. Numerical problems occur because at larger n the initial values are very small, which means that there is a relatively large quantization error that propagates through the whole calculation. One might think that the initialization of the recurrence could be chosen better (e.g., the middle of the interval); this however does not eliminate the problem.

Alternatively, starting from the first two polynomials one can use the recurrence relation (3) to generate t_n and use (4) to establish T_n. This is also numerically problematic in the sense that the calculated values of t_n grow rather quickly (as can also be seen from the norm H_n). A solution to this is to establish the recurrence relation for T_n directly and evaluate this. In particular we have

$$T_{n+1}(x) = a_n(2x - N + 1)T_n(x) - b_n T_{n-1}(x) \quad (8)$$

with

$$a_n = \frac{\sqrt{(2n+1)(2n+3)}}{n+1} \frac{1}{\sqrt{N^2 - (n+1)^2}}, \quad (9)$$

$$b_n = \frac{n}{n+1}\sqrt{\frac{2n+3}{2n-1}}\sqrt{\frac{(N^2 - n^2)}{(N^2 - (n+1)^2)}} \quad (10)$$

where the recursion can be started with $T_0(x) = 1/\sqrt{N}$ and $T_1(x) = \sqrt{3}(2x - N + 1)/\sqrt{N(N^2 - 1)}$. This reduces the propagation of errors and results in an extended but still limited range of polynomial orders for a given error. We programmed these approaches in Matlab and tried to generate the polynomials for all orders. The deviation of the norm of the constructed sequences from the expected unit norm were determined. The maximum deviation over all polynomials at a given interval size N was taken and this is shown in Fig. 1.

We see that for the interval size up to 30, the maximum deviation is at the level of the machine accuracy. From 30 onwards, there is a steady increase in deviation, with the maximum deviation always occurring at the highest polynomial order. At $N = 60$, the deviation is larger than the expected norm itself.

The problems arise at particular positions of x as is exemplified in Fig. 2, where the discrete Chebyshev function is plotted for $n = N - 1 = 55$. For this setting, the deviation from unit norm is about 18%, and thus should be visually observable when displayed graphically. The circles are the approximation to the true values stemming from the algorithm

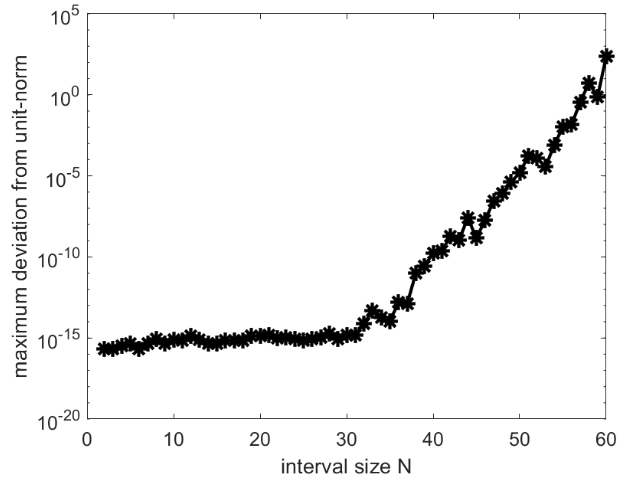

Fig. 1. Maximum deviation from unit norm over all polynomial orders as a function of interval size N when calculating the discrete Chebyshev using the recurrence relation.

proposed later, the dashed line is the data obtained by using the recurrence relation for T_n. In this case, the error by using the standard recursion is still relatively small; as mentioned the squared norm is about 18% more than it is supposed to be. At the endpoints ($x = 0$ and $X = N$) the function becomes small. The coefficients on the other hand become large at high n, especially near the endpoints. These are the areas where the propagation errors need to be controlled [15]. For all practical applications, the $T_n(x)$ values at which the propagation errors occur can be set to zero as the effective support of the Chebyshev function narrows with polynomial order. This is shown in the following.

For an arbitrary function f the center of energy M_f and the

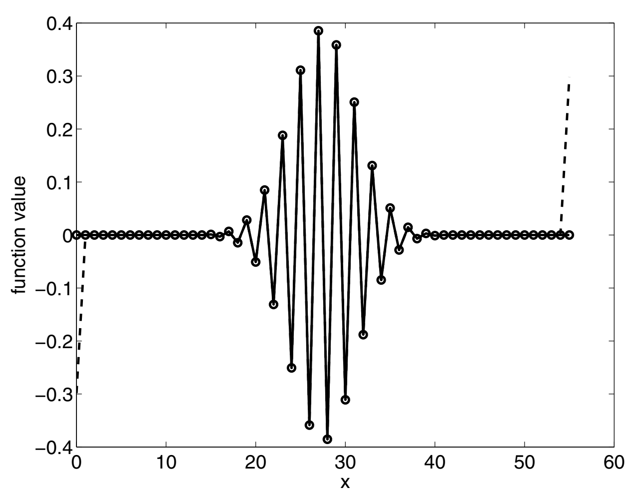

Fig. 2. Discrete Chebyshev polynomial for $n = N - 1 = 55$. The circles give the true values, the dashed line the calculated ones using the recurrence relation.

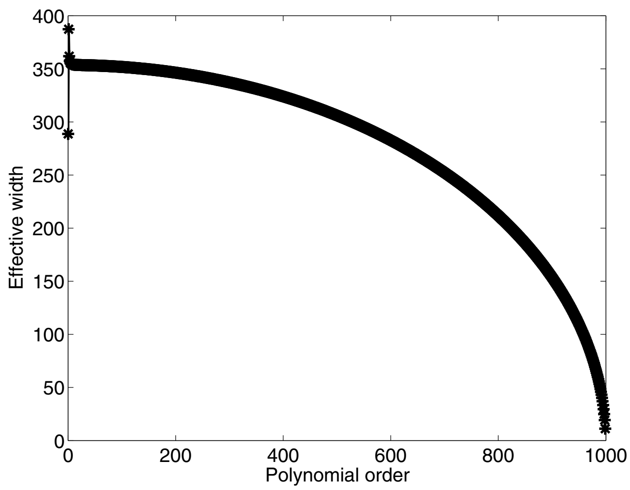

Fig. 3. Effective support as a function of polynomial order.

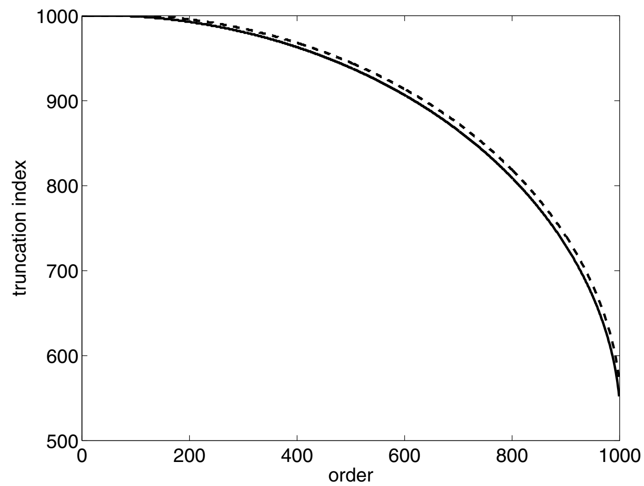

Fig. 4. Truncation index for two tolerances $\epsilon = 10^{-4}$ (solid line) and 10^{-8} (dashed line).

effective support (or width) S_f are defined as

$$M_f = \sum_x x f^2(x)/||f||^2, \tag{11}$$

$$S_f^2 = \sum_x (x - M_f)^2 f^2(x)/||f||^2. \tag{12}$$

The two measures are indicative of the position and size of the region where the bulk of the energy of the function f is located. We define M_n as the center of gravity of the nth normalized Chebyshev function T_n. In view of the (anti-)symmetry in the Chebyshev polynomials, we have $M_n = (N-1)/2$. We define S_n as the effective support of the nth normalized Chebyshev function T_n for which it can be easily shown that

$$S_n = \sqrt{1 + b_n^2}/2a_n \tag{13}$$

in view of (8). Substitution of (9)-(10) shows that for large n the effective support is continuously decreasing. We note that the calculation of center of gravity and effective width is extremely simple due to the specific form of the recurrence relation. This carries over to the other classical orthogonal polynomials. In a similar vein, the second-order differential or difference equation of orthogonal polynomials can be used to define compactness in the polynomial domain, see [17].

An example of the effective width is shown in Fig. 3 for a discrete interval of size $N = 1000$. After a short initial increase (the effective width of the first-order polynomial is larger than that of the zeroth order), the effective support is monotonically decreasing. This observation inspired the algorithm discussed below.

III. Proposed algorithm

A. Central idea

In view of the problem sketched above, we propose the following. First, we note symmetry in the polynomials:

$$t_n(x) = (-1)^n t_n(N-1-x). \tag{14}$$

This symmetry reduces the number of computations by half. Therefore, we need only to apply the recursion for $x \in [\lfloor (N-1)/2 \rfloor, N]$.

Suppose that we tolerate a deviation of ϵ from unit norm. After each recursion step, we evaluate

$$L_n(M) = \begin{cases} \displaystyle\sum_{x=N/2}^{M} 2T_n^2(x) \\ \displaystyle T_n^2((N-1)/2) + \sum_{x=(N+1)/2}^{M} 2T_n^2(x) \end{cases} \tag{15}$$

for even and odd N, respectively, with $M = \lceil (N+1)/2 \rceil, \cdots, N-1$. We limit the use of the recursion in step $n+1$ to those centrally located x-values that give a result sufficiently close to unity for the current n. Therefore, we search X_n as the lowest value M for which the following criterion is met:

$$X_n(\epsilon) = \min \arg_M \{L_n(M) > 1 - \epsilon\}, \tag{16}$$

where ϵ is called the tolerance. In Fig. 4, X_n is plotted for two tolerances: $\epsilon = 10^{-4}$ and 10^{-8}. We make two observations. Firstly, except for the first few orders, the curve has a shape that is essentially identical to that of the effective support. Secondly, for two largely different tolerances, the truncation index X_n is very steady.

In the proposed algorithm, we set a certain tolerance and take

$$T_k(X_n) = 0 \text{ for all } k > n \tag{17}$$

or, in other words, we stop executing the recursion algorithm at $x = X_n$ beyond this n. The range of x values over which the recursion is evaluated reduces gradually with the polynomial order n, but the example of Fig. 4 in combination with the general monotonically decreasing behaviour of the effective support suggest that this approach will guarantee a small

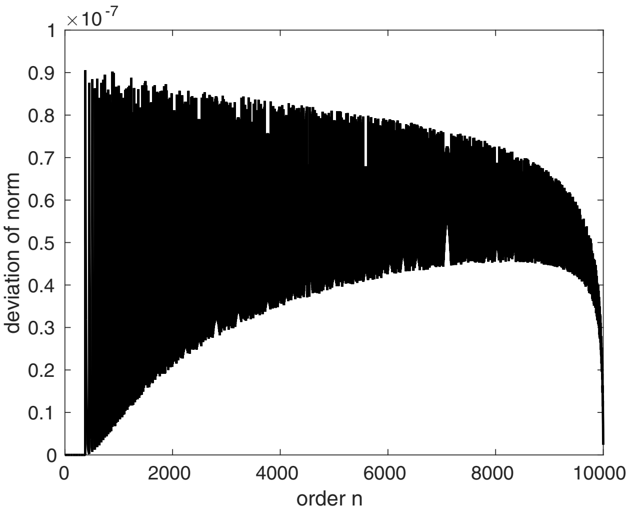

Fig. 5. Difference between unity and the norm of the calculated normalized discrete Chebyshev polynomials for $N = 10000$ and tolerance $\epsilon = 1/(1000N)$.

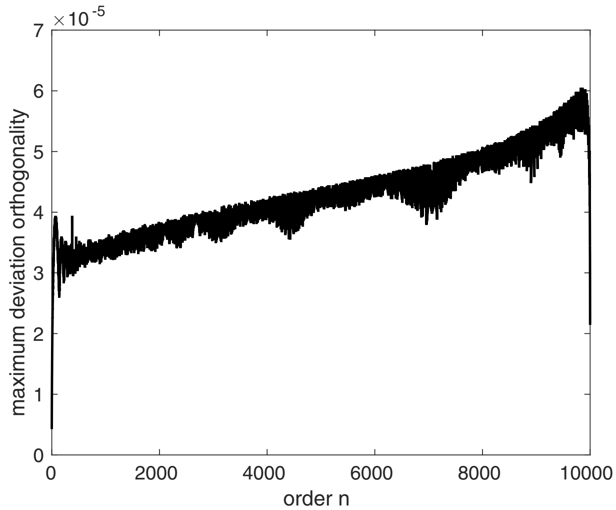

Fig. 6. Maximum of the absolute value of the inner product of function T_n with any other function T_k. Settings are $N = 10000$ and tolerance $\epsilon = 1/(1000N)$.

deviation from the unit norm for each constructed function T_n where the maximum deviation in norm is completely controlled by setting ϵ.

In practice, it is advised to choose the tolerance ϵ based on N. An evenly distributed signal strength (first-order polynomial) would be $1/N$ power addition per sample. Setting ϵ a fraction of that (e.g., $\epsilon = N/1000$) gives visually nice results. A lower ϵ gives more accuracy, a higher ϵ gives more speed in the calculation. In view of Fig. 4 the speed gain is rather limited. It is obvious that the tolerance must be chosen several orders of magnitude above machine precision.

We note that the procedure only makes sense for general purpose applications. For many dedicated applications, precomputation of the discrete Chebyshev functions (preventing repeated computation, including the additional computations required to control the accuracy) is likely a more logical choice.

B. Results

In Fig. 5 we have plotted the difference between unity and the norm of the orthonormal functions as calculated by the algorithm for $N = 10000$ and $\epsilon = 1/(1000N)$. Though it is very doubtful if anybody is ever going to use these functions for this large number of data and such high orders, it demonstrates the performance very well.

We observe that the algorithm guarantees a difference that is less than the tolerance for each order n, even though the size over which the recurrence operates is only allowed to shrink and never to increase. Secondly, we observe that at about $n = 400$ the support shrinkage kicks in. Having shrunk the size, the norm is low (close to tolerance) but then starts to recover (the deviation reduces). This is due to the fact that the functions become increasingly compact with increased order. This continues until the error propagation at the boundaries starts to play a role again, forcing a new shrinkage of the range where the recurrence relation is operated.

In Fig. 6 we have plotted the maximum of the absolute value of the inner product of function T_n with any other T_k ($k \neq n$). We make a number of observations from inspection of this graph. First of all, the deviation (the desired inner product being 0) is larger than the tolerance. We also note that at low orders n there is an immediate deviation, in contrast to the norm of the functions where, for low orders n, the deviation is negligible (see Fig 5). Throughout the figure there is a linear trend with the deviation doubling over the full range. We inspected the results for lower and higher tolerances. The curves remain remarkably the same except for scaling. The results in Fig. 5 obviously scale with ϵ, the results in Fig. 6 scale with $\sqrt{\epsilon}$.

IV. Conclusion

We have proposed an algorithm for calculating the discrete orthonormal Chebyshev functions. The algorithm is based on the recurrence relation and mitigates its numerical problems by shrinking the range over which the recurrence is executed. A similar method can be constructed for the case where the functions are calculated based on the difference equation, where, preferably, the difference equation is initialized at the mid of the interval.

Acknowledgements

The author thanks L. Tolhuizen and R. van Dinther for suggestions on a early version of the paper.

References

[1] J. V. Marcos and G. Cristóbal, "Texture classification using discrete Tchebichef moments," *J. Opt. Soc. Am. A*, vol. 30, no. 8, pp. 1580-1591, Aug. 2014.

[2] R. Mukundan, S. H. Ong and P.A. Lee, "Image analysis by Tchebichef moments," *IEEE Trans. Image Process.*, vol. 10, no. 9, pp. 1357-1364, 2001.

[3] J.-L. Rose, C. Revol-Muller, D. Charpigny, and C. Odet, "Shape prior criterion based on Tchebichef moments in variational region growing," in *2009 16th IEEE Int. Conf. Image Process. (ICIP)*, Cairo, Egypt, 7-10 Nov. 2009, pp. 1081-1084.

[4] H. Zhang, X. Dai, P. Sun, H. Zhu, and H. Shu, "Symmetric image recognition by Tchebichef moment invariants," in *2010 17th IEEE Int. Conf. Image Process. (ICIP)*, Hong Kong, 26-29 Sep. 2010, pp. 2273-2276.

[5] Q. Li, H. Zhu, and Q. Liu, "Image recognition by combined affine and blur Tchebichef moment invariants," in *2011 4th Int. Congress Image Signal Process. (CISP)*, Shanghai, China, 15-17 Oct. 2011, pp. 1517-1521.

[6] G. Wu and L. Xu, "Shape description and recognition by implicit Chebyshev moments," *Pattern Recognition Letters*, vol. 128, pp. 137-145, 2019.

[7] N. A. Abu, S. L. Wong, N. Herman, and R. Mukundan, "An efficient compact Tchebichef moment for image compression," in *2010 10th Int. Conf. Inform. Sciences Signal Process. Appl. (ISSPA)*, Kuala Lumpur, Malaysia, 10-13 May 2010, pp. 448-451.

[8] Q. Li and H. Zhu, "Block-based compressed sensing of image using directional Tchebichef transforms," in *2012 IEEE Int. Conf. Systems Man Cybern. (SMC)*, Seoul, Korea, 14-17 Oct. 2012, pp. 2207-2212.

[9] R. K. Senapati, U. C. Pati, and K. K. Mahapatra, "Reduced memory, low complexity embedded image compression algorithm using hierarchical listless discrete Tchebichef transform," *IET Image Process.*, vol. 8, pp. 213-238, Apr. 2014.

[10] F. M. H. AL-Shareefi, "Tchebichef moment based Hilbert scan for image compression," *Comput. Eng. Intelligent Syst.*, vol. 5, no. 8, 2014, pp. 18-25.

[11] B. Xiao, G. Lu, Y. Zhang, W. Li, and G. Wang, "Lossless image compression based on integer Discrete Tchebichef Transform," *Neurocomputing*, vol. 214, 19 Nov. 2016, pp. 587-593.

[12] W. A. Jassim, P. Raveendran and R. Mukundan, "New orthogonal polynomials for speech signal and image processing," *IET Signal Process.*, vol. 6, no. 8, pp. 713-7232, Oct. 2012.

[13] F. Ernawan, E. Noersasonko, N. A. Abu, "Using discrete Tchebichef transform on speech recognition," in *Proc. SPIE*, vol. 8350, pp. 83501J1-83501J5, 2012.

[14] R. Mukundan, "Some computational aspects of discrete orthonormal moments," *IEEE Trans. Image Process.*, vol. 13, no. 8, pp. 1055-1059, 2004.

[15] S. H. Abdulhussain, A. R. Ramli, S. A. R. Al-Haddad, B. M. Mahmmod, and W. A. Jassim, "On computational aspects of Tchebichef polynomials for higher polynomial order," *IEEE Access*, vol. 5, pp. 2470-2478, 2017.

[16] C. Camacho-Bello and J. S. Rivera-Lopez, "Some computational aspects of Tchebichef moments for higher orders," *Pattern Recognition Letters*, vol. 112, pp. 332-339, 2018.

[17] A. C. den Brinker and H. J. W. Belt, "Optimal free parameters in orthonormal approximations," *IEEE Trans. Signal Process.*, vol. 46, no. 8, pp. 2081-2087, 1998.