An efficient procedure deciding positivity for a class of holonomic functions

Veronika Pillwein, Miriam Schussler *
RISC
Johannes Kepler University
Linz, Austria, A-4040
vpillwei@risc.jku.at mschussl@risc.jku.at

Abstract

We present an efficient decision procedure for positivity on a class of holonomic sequences satisfying recurrences of arbitrary order.

1 Introduction

In 2005, Gerhold and Kauers [2] proposed a method that is applicable to proving inequalities concerning sequences that satisfy recurrence equations of a very general type. The basic idea is to prove the inequalities by induction and their method consists of constructing a sequence of polynomial sufficient conditions that would imply the non-polynomial inequality under consideration. The truth of these conditions is tested using Cylindrical Algebraic Decomposition (CAD) [1]. If the inequality does not hold, then the method terminates after a finite number of steps and returns a counterexample. If the inequality holds, then either the program terminates and returns True or it may fail to detect this and run forever. Besides termination not being guaranteed another drawback of using a method based on CAD is that it is computationally expensive. In [4] and [5] a main goal was to find termination conditions. Fortunately the proof produced in [5] to extend the domain where termination can be proven indicates a more efficient procedure for determining positivity on a *restricted set* of holonomic sequences. The work presented here freely uses proofs and follows notation found in [4, 5].

2 Preliminaries

A sequence $f : \mathbb{N} \to K$ where K is a computable subfield of $\mathbb{C}$ is P-finite (or *holonomic*) of order d if there exist polynomials $p_0, \ldots, p_d \in K[x]$, not all zero, such that

$$p_0(n)f(n) + p_1(n)f(n+1) + ... + p_d f(n+d) = 0.$$

*The authors are funded by the Austrian Science Fund (FWF) under grant F50-07.

We also refer to the recurrence as P-finite. If all the coefficients in the recurrence are constant, then we call the sequence C-finite. A P-finite recurrence is called *balanced* if $\deg p_0 = \deg p_d$ and $\deg p_i \leq \deg p_0$ $(i = 1, \ldots, d)$. We will often find it more useful here to write the recurrence with rational function coefficients in a form equal to $f(n+d)$:

$$f(n+d) = r_{d-1}(n)f(n+d-1) + \cdots + r_0(n)f(n). \tag{1}$$

The characteristic polynomial of a balanced recurrence is defined as:

$$\chi(x) = lc_y(p_0(y) + p_1(y)x + p_2(y)x^2 + \cdots + p_d(y)x^d). \tag{2}$$

Its roots $\alpha_0, ..., \alpha_s$ are called the *eigenvalues* of the recurrence. Here the α_i are distinct and the sum of their multiplicities is equal to d. An eigenvalue α_i is called dominant if $|\alpha_j| \leq |\alpha_i|$ for all $j = 1, \ldots, d$. In what follows we consider P-finite recurrences with one positive dominant eigenvalue. The task is, given a P-finite sequence $f(n)$ from its recurrence coefficients and sufficiently many initial values, decide if $f(n) \geq 0$ for all $n \in \mathbb{N}$. For recurrences where $\alpha_i \neq 1$ we may scale our recurrences and without loss of generality consider only sequences with dominant eigenvalue equal to 1 (because $g(n) = f(n)/\alpha_i^n \geq 0 \Leftrightarrow f(n) \geq 0$).

3 Method

Relevant here is a variant (introduced in [5]) of the original algorithm. In this variant, in order to prove positivity for a particular sequence $f(n)$, we consider the shifted subsequence $f(n+m)$ for some $m > d$. That is, we seek to prove that

$$f(n) \geq 0 \wedge f(n+1) \geq 0 \wedge \cdots \wedge f(n+d-1) \geq 0 \Rightarrow f(n+m) \geq 0, \tag{3}$$

for $n \geq n_0$, for some lower bound n_0. For any $m \geq d$, repeated application of the given recurrence allows us to compute $f(n+m)$ from d consecutive sequence elements.

$$f(n+m) = R_{d-1}(m,n)f(n+d-1) + \cdots + R_0(m,n)f(n), \tag{4}$$

where the $R_k(m,\cdot)$ are rational functions. If for some fixed m_0 all the $R_k(m_0,\cdot)$ are eventually positive, then the implication in (3) is trivially true for n greater than some lower bound n_0.

Let $\chi(x) = x^d - c_{d-1}x^{d-1} - \cdots - c_1 x - c_0$ be the characteristic polynomial of the given sequence. Then [5] for every fixed m_0, we have that $\lim_{n\to\infty} R_k(m_0, n) = \gamma_k(m_0)$, where each $\gamma_k(m)$ is the C-finite sequence defined by the recurrence

$$\gamma_k(m+d) = c_{d-1}\gamma_k(m+d-1) + ... + c_0\gamma_k(m) \tag{5}$$

with initial values $\gamma_k(j) = \delta_{k,j}$ for $j = 0, 1, ..., d-1$ (where $\delta_{k,j}$ denotes the Kronecker delta). The solution of each recurrence can be explicitly computed in closed form as a linear combination of the sequence $(1)_{m\geq 0}$ and sequences of the form $\alpha^m, m\alpha^m, \ldots, m^{e-1}\alpha^m$ where α is an eigenvalue and e denotes its multiplicity. Let ζ_k be the coefficient of the eigenvalue 1 in this closed form, then $\lim_{m\to\infty} \gamma_k(m) = \zeta_k$. Furthermore, for each fixed m, $\lim_{n\to\infty} R_k(m,n) = \gamma_k(m)$. If all the limits ζ_k are positive, then it remains to determine an m_0 and n_0 such that $R_k(m_0, n) > 0$, for $n > n_0$. Checking positivity of $f(0), \ldots, f(n_0 + m_0)$ concludes the proof of positivity of $f(n)$.

With these notations and considerations at hand we now proceed to use the proof-idea of [5] for a method to prove positivity directly, avoiding the use of CAD.

Given a scaled balanced P-finite recurrence and initial values we will decide if an m_0 (not unique) exists for which the implication in (3) holds with $m = m_0$ and n large. If so we will check a sufficient number of initial values to prove or disprove positivity.

The first task is to determine the characteristic polynomial, eigenvalues and a closed form for each of the C-finite sequences γ_k as defined in (5). These are elementary procedures. From the closed form we check that each ζ_k is positive. If we find any $\zeta_k < 0$ we revert to the CAD based approach. Otherwise we proceed.

We do not require our choice of m_0 to be minimal in any sense, only that for each k it satisfies $|\gamma_k(m)-\zeta_k| \leq \frac{\zeta_k}{2}$ for $m > m_0$. Let $B_k(x)$ be an upper bound for $|\gamma_k(x)-\zeta_k|$. A sufficient requirement for $m_0 \in \mathbb{N}$ is that for each k it satisfy $B_k(x) \leq \frac{\zeta_k}{2}$ for $x > m_0$. We use $B_k(x) = \upsilon_k \alpha^x t_k x^{e-1}$ where υ_k and α are respectively the maximum absolute values of the coefficient of any term in the closed form of γ_k other than 1^m and all eigenvalues other than 1, and t_k and e are the number of terms in the closed form and maximum multiplicity of any eigenvalue.

Note that $B_k(x)$ was constructed in a way to make it easy to determine the maximum used in the following choice for m_0:

$$m_0 = \lceil \max\{0, x \mid \exists k \in \{0, \ldots, d-1\} \colon B_k(x) = \zeta_k/2\} \rceil$$

and to ensure that $B_k(x) < \zeta_k/2$ for all $x > m_0$. Therefore also $|\gamma_k(m) - \zeta_k| < \zeta_k/2$ for all $m > m_0$.

Having set m_0, we now find an $n_0 \geq 0$ such that $|R_k(m_0, n) - \zeta_k| < \zeta_k/2$ for each k and all $n > n_0$. Through iteration of the original recurrence we find $R_0(m_0, \cdot), \ldots, R_{d-1}(m_0, \cdot)$ such that

$$f(n + m_0) = R_{d-1}(m_0, n) f(n + d - 1) + \ldots + R_0(m_0, n) f(n).$$

For each k, $R_k(m_0, x) - \gamma_k(m_0)$ is rational, and $\lim_{x\to\infty} R_k(m_0, x) - \gamma_k(m_0) = 0$. These two facts allow us to set

$$n_0 = \left\lceil \max\{0, x \mid \exists k \in \{0 \ldots d-1\} \colon |R_k(m_0, x) - \gamma_k(m_0)| = \tfrac{\zeta_k}{2}\} \right\rceil .$$

Then for each k we have

$$\begin{aligned} |R_k(m_0, n) - \zeta_k| &= |R_k(m_0, n) - \gamma_k(m_0)| + |\gamma_k(m_0) - \zeta_k| \\ &< \zeta_k/2 + \zeta_k/2 = \zeta_k, \qquad \forall n > n_0. \end{aligned}$$

ensuring that all the rational function coefficients of $f(n + m_0)$ are positive for $n > n_0$. If we check the first $n_0 + m_0$ values and find they are positive we have a proof that $f(n)$ is a positive valued sequence.

Example 1 *We use the new method to show positivity of a sequence defined by a P-finite recurrence with eigenvalues outside the previously proven termination bounds for the CAD based approach. Let $f(n)$ be defined by*

$$(8n - 17)f(n + 3) - (4n - 14)f(n + 2) - (8 - 3n)f(n + 1) - (7n + 11)f(n) = 0,$$

with initial values $f(0) = 9, f(1) = 3, f(2) = 7$.

The scaled characteristic polynomial is $\chi(x) = x^3 - \frac{1}{2}x^2 + \frac{3}{8}x - \frac{7}{8}$*, the eigenvalues are* $\alpha_0 = 1$, $\alpha_{1,2} = \frac{1}{4}(-1 \pm i\sqrt{13})$*, and so* $\alpha = \max\{|\alpha_1|, |\alpha_2|\} = \sqrt{7/8}$*.*

The related C-finite recurrence is $\gamma_k(m+3) - \frac{1}{2}\gamma_k(m+2) + \frac{3}{8}\gamma(m+1) - \frac{7}{8}\gamma(m) = 0$*. We compute the closed form for* $\gamma_0(m), \gamma_1(m)$ *and* $\gamma_2(m)$ *with initial values* $\{1,0,0\}, \{0,1,0\}$*, and* $\{0,0,1\}$ *respectively:*

$$\gamma_0(m) = \frac{7}{19} - \frac{6\sqrt{13}-8i}{19\sqrt{13}}\alpha_1^m + \frac{6\sqrt{13}+8i}{19\sqrt{13}}\alpha_2^m$$
$$\gamma_1(m) = \frac{4}{19} - \frac{2\sqrt{13}-28i}{19\sqrt{13}}\alpha_1^m + \frac{2\sqrt{13}-28i}{19\sqrt{13}}\alpha_2^m$$
$$\gamma_2(m) = \frac{8}{19} - \frac{4\sqrt{13}+20i}{19\sqrt{13}}\alpha_1^m + \frac{-4\sqrt{13}+20i}{19\sqrt{13}}\alpha_2^m$$

¿From the closed form we get $\upsilon_0 = \sqrt{\frac{28}{247}}, \upsilon_1 = \sqrt{\frac{44}{247}}, \upsilon_2 = \sqrt{\frac{32}{247}}$*;* $\zeta_0 = \frac{7}{19}, \zeta_1 = \frac{4}{19}, \zeta_2 = \frac{8}{19}$*;* $t_0 = t_1 = t_2 = 3; e = 1$ *With those values we construct* $B_k(x) = \upsilon_k \alpha^x t_k x^{e-1}$ *for each* k*,*

$$B_0(x) = \sqrt{\frac{28}{247}}\left(\frac{7}{8}\right)^{x/2} \qquad B_0(x) = \sqrt{\frac{44}{247}}\left(\frac{7}{8}\right)^{x/2} \qquad B_0(x) = \sqrt{\frac{32}{247}}\left(\frac{7}{8}\right)^{x/2}.$$

Then, as defined above, we set $m_0 = \lceil \max\{0, x \mid \exists k\colon B_k(x) = \zeta_k/2\} \rceil = 32$*, and compute the* $R_k(32, n)$ *with the original recurrence:*

$$(8n-17)R_3(k+3,n) - (4n-14)R_2(k+2,n) - (8-3n)R(k+1,n) - (7n+11)R_0(k,n) = 0$$

to determine $n_0 = \lceil \max\{0, x \mid \exists k\colon |R_k(32,x) - \gamma_k(m_0)| = \zeta_k/2\} \rceil = 117$*. The first* $m_0 + n_0 = 149$ *values are indeed positive which completes our proof that* $f(n) > 0$ *for all* $n \in \mathbb{N}$*.*

This result took approximately 8 seconds with our method vs. almost 96 seconds for the original algorithm using the implementation in SumCracker [3].

References

[1] Bob F. Caviness and Jeremy R. Johnson, editors. Quantifier Elimination and Cylindrical Algebraic Decomposition. *Texts and Monographs in Symbolic Computation*, 1998.

[2] Stefan Gerhold and Manuel Kauers. A procedure for proving special function inequalities involving a discrete parameter. *Proceedings of ISSAC'05*, 156–162, 2005.

[3] Manuel Kauers. SumCracker – A Package for Manipulating Symbolic Sums and Related Objects. Journal of Symbolic Computation 41(9), 1039–1057, 2006.

[4] Manuel Kauers, and Veronika Pillwein. When can we detect that a P-finite sequence is positive? *Proceedings of ISSAC'10*, 195–201, 2010.

[5] Veronika Pillwein. Termination conditions for positivity proving procedures *Proceedings of ISSAC'13*, 315–321, 2013.