Symbolic Summation

Victor Adamchik

Carnegie Mellon University

Petkovšek's algorithm

"All the wonders of our universe can in effect be captured by simple rules, yet [...] there can be no way to know all the consequences of these rules, except in effect just to watch and see how they unfold."

Stephen Wolfram, "*A New Kind of Science*", p. 846.

Algorithm

The algorithm deals with finding a hypergeometric term solution to

$$\sum_{j=0}^{N} d_j(n)\, S(n+j) = 0$$

where d_j are polynomials in n. For simplicity, let us consider a second order equation

$$d_2(n)\, S(n+2) + d_1(n)\, S(n+1) + d_0(n)\, S(n) = 0 \qquad (1)$$

where $S(n)$ is assumed to be a hypergeometric term

$$\frac{S(n+1)}{S(n)} \in Q(n)$$

If all $d_k(n)$ are constants then we solve the equation by means of the characteristic equation

$$d_2\, \lambda^2 + d_1\, \lambda + d_0 = 0$$

The roots of this equation define the general solution

$$S(n) = c_1\, \lambda_1^n + c_2\, \lambda_2^n$$

If $d_k(n)$ is polynomials in n then proceed in the following way. Divide (1) by $S(n)$

$$d_2(n)\,\frac{S(n+2)}{S(n)} + d_1(n)\,\frac{S(n+1)}{S(n)} + d_0(n) = 0$$

or

$$d_2(n)\,\frac{S(n+2)}{S(n+1)}\,\frac{S(n+1)}{S(n)} + d_1(n)\,\frac{S(n+1)}{S(n)} + d_0(n) = 0 \tag{2}$$

From the Gosper's algorithm, we know that any rational function can be represented as

$$\frac{S(n+1)}{S(n)} = c\,\frac{p(n+1)}{p(n)}\,\frac{q(n+1)}{r(n+1)} \tag{3}$$

where $p(n),\ q(n),\ r(n)$ are *monic* (the leading coefficient is 1) and *coprime*

$$\mathrm{GCD}(q(n),\ r(n+j)) = 1, \quad j \in N_0.$$

$$\mathrm{GCD}(p(n),\ q(n+1)) = 1,$$

$$\mathrm{GCD}(p(n),\ r(n)) = 1,$$

Substituting (3) into (2), yields

$$d_2(n)\,c\,\frac{p(n+2)}{p(n+1)}\,\frac{q(n+2)}{r(n+2)}\,c\,\frac{p(n+1)}{p(n)}\,\frac{q(n+1)}{r(n+1)} + d_1(n)\,c\,\frac{p(n+1)}{p(n)}\,\frac{q(n+1)}{r(n+1)} + d_0(n) = 0$$

Multiply it by $p(n)\,r(n+1)\,r(n+2)$ to get

$$\begin{aligned} &c^2\,d_2(n)\,p(n+2)\,q(n+1)\,q(n+2) + \\ &c\,d_1(n)\,p(n+1)\,q(n+1)\,r(n+2) + \\ &d_0(n)\,p(n)\,r(n+1)\,r(n+2) = 0 \end{aligned} \tag{4}$$

Remember, our goal is to find $p(n),\ q(n),\ r(n)$ and constant c.

The logic in the next paragraph is somewhat similar to the proof of step 2 in Gosper's algorithm.

The first two terms of (4) are divisible by $q(n+1)$, therefore the last term $d_0(n)\,p(n)\,r(n+1)\,r(n+2)$ must be divisible by $q(n+1)$. Since $p,\ q$ and r are relatively prime, then $q(n+1)$ must divide $d_0(n)$. In other words, $q(n+1)$ must be a factor of $d_0(n)$. This leads us to the finite number of choices for $q(n+1)$.

The last two terms of (4) are divisible by $r(n+2)$, therefore the first term must be divisible by $r(n+2)$. Since $p,\ q$ and r are relatively prime, then $r(n+2)$ must divide $d_2(n)$. In other words, $r(n+2)$ must be a factor of $d_2(n)$. This leads us to finite number of choices for $r(n+2)$.

Once we know q and r, we can easily find a rational constant c. Divide (4) by $q(n+1)\,r(n+2)$

$$c^2 \frac{d_2(n)}{r(n+2)} p(n+2)\, q(n+2) +$$
$$c\, d_1(n)\, p(n+1) +$$
$$\frac{d_0(n)}{q(n+1)} p(n)\, r(n+1) = 0$$

This gives us an equation with polynomial coefficients. Equating the leading coefficient to zero, generates a quadratic equation for c. But $p(n)$ is still unknown. ??? hmm... Where is a catch? $p(n)$ is a monic polynomial.

Now the last step.

So far we found the constant c and two polynomials q and r. To find the polynomial p we must solve (5). Since all coefficients in (5) are polynomials, we need to find a polynomial solution.

Let us consider a generic equation of the second order with polynomial coefficients

$$a_2(n)\, Y(n+2) + a_1(n)\, Y(n+1) + a_0(n)\, Y(n) = 0$$

We need to find an upper bound for the degree of a polynomial solution. Assume the following

$$a_2(n) = \alpha_p\, n^P + \alpha_{p-1}\, n^{P-1} + \ldots$$
$$a_1(n) = \beta_p\, n^P + \beta_{p-1}\, n^{P-1} + \ldots$$
$$a_0(n) = \gamma_p\, n^P + \gamma_{p-1}\, n^{P-1} + \ldots$$

where P is the maximal degree of a_2, a_1 and a_0, and all coefficents α_j, β_j, γ_j are known. We are looking for a monic polynomial solution

$$Y(n) = n^M + \delta_{M-1}\, n^{M-1} + \ldots$$
$$Y(n+1) = n^M + (M + \delta_{M-1})\, n^{M-1} + \ldots$$
$$Y(n+2) = n^M + (2\,M + \delta_{M-1})\, n^{M-1} + \ldots$$

where order M and coefficients δ_k are to be determined. Substitute these into the difference equation and take coefficients of the first three dominant terms. We obtain

$\boldsymbol{n^{M+P}}$ **:** $\alpha_p + \beta_p + \gamma_p$

$\boldsymbol{n^{M+P-1}}$ **:** $\alpha_{p-1} + \beta_{p-1} + \gamma_{p-1} + M\,(2\,\alpha_p + \beta_p) + (\alpha_p + \beta_p + \gamma_p)\,\delta_{M-1}$

$\boldsymbol{n^{M+P-2}}$ **:** $\frac{1}{2}(4\,\alpha_p + \beta_p)\, M^2 + \left(2\,\alpha_{p-1} + \beta_{p-1} - 2\,\alpha_p - \frac{\beta_p}{2} + (2\,\alpha_p + \beta_p)\,\delta_{M-1}\right) M +$
$\alpha_{p-2} + \beta_{p-2} + \gamma_{p-2} + (\alpha_p + \beta_p + \gamma_p)\,\delta_{M-2} + (\alpha_{p-1} - 2\,\alpha_p +$
$\beta_{p-1} - \beta_p + \gamma_{p-1})\,\delta_{M-1}$

Each of them must be zero. Start with the first

$$\alpha_p + \beta_p + \gamma_p = 0 \qquad (6)$$

If this condition is not satisfied, then no polynomial solution exists. Suppose (6) is satisfied, then the next coefficient gives

$$\alpha_{p-1} + \beta_{p-1} + \gamma_{p-1} + M\,(2\,\alpha_p + \beta_p) + (\alpha_p + \beta_p + \gamma_p)\,\delta_{M-1} = 0$$

or

$$\alpha_{p-1} + \beta_{p-1} + \gamma_{p-1} + M\,(2\,\alpha_p + \beta_p) = 0 \tag{7}$$

This splits into two subcases

1) case

$$2\,\alpha_p + \beta_p \neq 0$$

then the order M is uniquely defined from (7)

$$M = -\frac{\alpha_{p-1} + \beta_{p-1} + \gamma_{p-1}}{2\,\alpha_p + \beta_p} \tag{8}$$

2) case

$$2\,\alpha_p + \beta_p = 0$$

Then

$$\alpha_{p-1} + \beta_{p-1} + \gamma_{p-1} = 0 \tag{9}$$

and we must look at the next coefficient (a coefficient by n^{M+P-2}) that is

$$\alpha_p\,M^2 + (2\,\alpha_{p-1} + \beta_{p-1} - \alpha_p)\,M + \alpha_{p-2} + \beta_{p-2} + \gamma_{p-2} = 0 \tag{10}$$

Observe, that $\alpha_p \neq 0$, because otherwise $\beta_p = 0$ and then by (6) $\gamma_p = 0$, which will contradict to the assumption that p has a maximal degree of p_2, p_1 and p_0. Therefore, equation (10) has two solutions, one (or both) defines the upper degree of a polynomial solution.

Example

Find a hypergeometric term solution to

$$9\,(n+2)\,S(n+2) - 3\,(n+4)\,S(n+1) - 2\,(n+3)\,S(n) = 0$$

$$S(0) = S(1) = 1$$

We assume that

$$\frac{S(n+1)}{S(n)} = c\,\frac{p(n+1)}{p(n)}\,\frac{q(n+1)}{r(n+1)}$$

where

$q(n+1)$ must divide $d_0(n)$

$r(n+2)$ must divide $d_2(n)$

Here are our choices

$$q(n+1) \text{ is either } 1 \text{ or } n+3$$

$$r(n+2) \text{ is either } 1 \text{ or } n+2$$

■ q(n+1) = 1; r(n+2)=1

We find constant c from the equation (since p, q and r are monic, we replace them by 1 in equation (5))

$$c^2 d_2(n) + c\, d_1(n) + d_0(n) = 0$$

that simplifies to

$$9\,(n+2)\,c^2 - 3\,(n+4)\,c - 2\,(n+3) = 0$$

$$n\left(9\,c^2 - 3\,c - 2\right) + 6\left(3\,c^2 - 2\,c - 1\right) = 0$$

Solving

$$9\,c^2 - 3\,c - 2 = 0$$

we obtain

$$c = -\frac{1}{3} \quad \text{or} \quad c = \frac{2}{3}$$

Note, constant c is defined only by a leading coefficient of equation (5).

Case 1. $c = -\frac{1}{3}$

We need to find a polynomial solution to

$$(n+2)\; p(n+2) + (n+4)\; p(n+1) - 2\,(n+3)\; p(n) = 0$$

Compute

$$\alpha_p + \beta_p + \gamma_p = 1 + 1 - 2 = 0$$

$$\alpha_{p-1} + \beta_{p-1} + \gamma_{p-1} + M\,(2\,\alpha_p + \beta_p) = 2 + 4 - 6 + M(2+1) = 0$$

The solution $p(n)$ is a constant, since $M = 0$. Therefore, we have

$$\frac{S(n+1)}{S(n)} = c\,\frac{p(n+1)}{p(n)}\,\frac{q(n+1)}{r(n+1)} = \frac{-1}{3}\,\frac{p(n+1)}{p(n)}\,\frac{1}{1} = -\frac{1}{3}$$

or

$$Y(n) = \left(-\frac{1}{3}\right)^n$$

Case 2. $c = \frac{2}{3}$

We find an upper bound on the polynomial solution

$$(2n+4)\,p(n+2) - (n+4)\,p(n+1) - (n+3)\,p(n) == 0$$

The first condition (the sum of leading coefficients should be zero)

$$\alpha_p + \beta_p + \gamma_p = 0$$

The next condition

$$\alpha_{p-1} + \beta_{p-1} + \gamma_{p-1} + M\,(2\,\alpha_p + \beta_p) = 0$$

which is

$$3\,M - 3 = 0$$

So, the solution is linear

$$p(n) = n + x$$

where x is unknown. We find x by substituting $p(n)$ into the original equation

$$(2n+4)\,(n+2+x) - (n+4)\,(n+1+x) - (n+3)\,(n+x) == 0$$

This simplifies to

$$4 - 3\,x = 0$$

Thus,

$$p(n) = n + \frac{4}{3}$$

And

$$\frac{S(n+1)}{S(n)} = c\,\frac{p(n+1)}{p(n)}\,\frac{q(n+1)}{r(n+1)} = \frac{2}{3}\,\frac{n+\frac{7}{3}}{n+\frac{4}{3}}\,\frac{1}{1} = \frac{2}{3}\,\frac{3\,n+7}{3\,n+4}$$

or

$$S(n) = \frac{3\,n+4}{4}\left(\frac{2}{3}\right)^n$$

Finally, we combine the above two cases to get a general solution

$$S(n) = c_1\left(-\frac{1}{3}\right)^n + c_2\,\frac{3\,n+4}{4}\left(\frac{2}{3}\right)^n$$

where unknown c_1 and c_2 can be easily found from the initial conditions

$$S(0) = S(1) = 1$$

We get a system of two linear equations

$$c_1 + c_2 = 1$$

$$c_1\left(-\frac{1}{3}\right) + c_2\,\frac{7}{4}\left(\frac{2}{3}\right) = 1$$

solving each, yields

$$c_1 = \frac{1}{9} \text{ and } c_2 = \frac{8}{9}$$

Hence,

$$S(n) = \frac{1}{9}\left(-\frac{1}{3}\right)^n + \frac{8}{9}\,\frac{3n+4}{4}\left(\frac{2}{3}\right)^n$$

- **q(n+1) = 1; r(n+2)=n+2**

Equation for constant c

$$-2n^2 + (-3c-8)\,n + 3\left(3c^2 - 4c - 2\right) = 0$$

Such constant c does not exist

- **q(n+1) = n+3; r(n+2)=1**

Equation for constant c

$$9c^2 n^2 + 3c\,(18c-1)\,n + 2\left(36c^2 - 6c - 1\right) = 0$$

It has only a trivial solution $c = 0$.

- **q(n+1) = n+3; r(n+2)=n+2**

This case cannot be chosen, since they have a common polynomial GCD:

```
q[n_] := n + 2;
r[n_] := n
Table[PolynomialGCD[q[n], r[n + j]], {j, 0, 5}]
```

```
{1, 1, 2 + n, 1, 1, 1}
```

References

[1] M. Petkovsek, H. Wilf, D. Zeilberger, *A = B*, Algorithms and Computations in Mathematics, AK Peters, 1996.

[2] M. Petkovšek, Hypergeometric solutions of linear recurrences with polynomial coefficients, *J. Symbolic Computation*, (**14**)1992, 243-264.