Petkovšek's algorithm Hyper

Student seminar on automatic proofs of binomial identities

Stefan Herytash

May 28, 2020

Contents

This report presents the algorithm Hyper introduced by Petkovšek in his PhD thesis [2]. The following text is based mainly on chapters 8.1-8.4 from the book [4].

1 Introduction

We are given a sum

$$f(n) = \sum_k F(n,k), \tag{1}$$

where $F(n,k)$ is a hypergeometric funciton in n and k, i.e. $F(n+1,k)/F(n,k)$ and $F(n,k+1)/F(n,k)$ are rational functions in n and k. We are looking for a closed form of $f(n)$.

Definition 1.1 (Closed form). *A function $f(n)$ is said to have a hypergeometric closed form, if there exists a representation of $f(n)$ as a linear combination of r hypergeometric terms, where $r \in \mathbb{N}$ is fixed and does not depend on n or any other variables.*

A closed form does not neccessarily exist and even if it exists it can be rather hard to find it. Remark that a closed form as we defined it does not neccesarrily mean a pretty formula. There are sums of the form (1) for which no closed form exists but indeed a very short formulas does exist. See for example the derangement function $d(n)$ which counts the number of derangements of n objects. A derangement is a permutation without any fixed points. Indeed the derangement function does not have a closed form but still $d(n) = \|n!/\mathrm{e}\|$ for $n > 0$, where $\|\cdot\|$ denotes the nearest integer (cf. [1, Section 1]). How might one go about finding such a closed form. First, for a given hypergeometric F we might try to find $F(k) = G(k+1) - G(k)$, where G is hypergeometric aswell. Gosper's algorithm finds this if it exists. But even for hypergeometric sums where a nice closed form exists, this does not always find it. Next we try to find a recurrence relation of the form

$$F(n+1,k) - F(n,k) = G(n,k+1) - G(n,k)$$

or more generally of the form

$$\sum_{j=0}^{J} a_j(n)F(n+j,k) = G(n,k+1) - G(n,k), \tag{2}$$

where $(a_j(n))_{j=0}^{J}$ are polynomials in n. Zeilberger's algorithm (cf. [4, Chapter 6]) also known as creative telescoping provides such a recurrence under very general assumptions. Assuming we have found a recurrence of the form (2) where the function $G(n,k)$ has compact support we can sum over all k to obtain

$$\sum_{j=0}^{J} a_j(n)f(n+j) = 0. \tag{3}$$

Now we have a recurrence for the function $f(n)$ itself. If the recurrence relation is of order 1 or of order ≥ 2 but the polynomials a_j are constant for all j then solving the charactersitic polynomial gives the solution. If the relation is of order ≥ 2 with polynomial coefficients then this method does not work. In the following we will see how to find a closed form for the function $f(n)$ if one exists or otherwise prove that there does not exist such a form. Our approach is going to be as follows. First, first a recurrence of the form (3). Next, find all hypergeometric solutions of this recurrence. Lastly, find a linear combination of those solutions that matches $f(n)$ for enough consecutive values of n. We will show how to do the second and third steps.

2 Prerequisites

First let us collect some notions that will be needed.

2.1 Ring of sequences

Let us briefly recall the ring of sequences and the shift operator. Let K be a field of characteristic zero. You might think of K as the field of rational numbers $\mathbb{Q}$ or the field of real numbers $\mathbb{R}$. Denote by $K^{\mathbb{N}}$ the ring of sequences with elements in K, that is $(a(n))_{n=0}^{\infty} = (a(0), a(1), \ldots) \in K^{\mathbb{N}}$, where $a(i) \in K$ for all i. Indeed, $K^{\mathbb{N}}$ is a ring with addition and multiplication defined termwise. The field K is naturally embedded in $K^{\mathbb{N}}$ as a subring by identifying $u \in K$ with the constant sequence $(u, u, \ldots) \in K^{\mathbb{N}}$. $K^{\mathbb{N}}$ is a K-vector space. We are interested in recurrence relations. For this end, define the shift operator

$$(Na)(n) = a(n+1).$$

The shift operator N is a linear operator on $K^{\mathbb{N}}$, i.e. $N(\lambda a + b) = \lambda Na + Nb$ for $a, b \in K^{\mathbb{N}}$ and $\lambda \in K$. Consider the linear recurrence operator

$$L = \sum_{k=0}^{r} a_k N^k,$$

where $a_k \in K^{\mathbb{N}}$, $a_0, a_r \neq 0$. The order of L is $\mathrm{ord} L = r$. For $f \in K^{\mathbb{N}}$

$$Ly = f$$

is a linear recurrence equation. It is homogeneous if $f = 0$ and inhomogeneous otherwise. We desire $\dim \ker L = \mathrm{ord} L$. But this fails on $K^{\mathbb{N}}$. See for example [4, Example 8.2.1] or [4, Example 8.2.2]. For $a, b \in K^{\mathbb{N}}$, say

$$a(n) = b(n) \text{ a.e.},$$

if they agree for all but finitely many n. Identify a with b if $a = b$ a.e. Denote the ring of sequences with regard to this identification as $\mathcal{S}(K)$. The space $\mathcal{S}(K)$ is a K-vector space. Then any element with finitely many nonzero terms is identified with zero. In particular any sequence that is eventually zero is identified with zero. In this domain there are still zero divisors but those have infinitely many nonzero and infinitely many zero terms. A nonzero $a \in \mathcal{S}(K)$ is a unit if and only if it is eventually nonzero. We find $\dim \ker L = \mathrm{ord} L$ on $\mathcal{S}(K)$.

Theorem 2.1 ([4, Theorem 8.2.1])**.** *Let $L = \sum_{k=0}^{r} a_k N^k$ be a linear recurrence operator on $\mathcal{S}(K)$ with $\mathrm{ord} L = r$. If a_0, a_r are units then*

$$\dim \ker L = r.$$

The proof is ommitted.

Definition 2.2. *Let $a \in \mathcal{S}(K)$. The sequence $a(n)$ is called polynomial if there exists a polynomial $p(x) \in K[x]$ such that $a(n) = p(n)$ a.e. Similiarly it is called rational if there exists a rational function $r(x) \in K(x)$ such that $a(n) = r(n)$ a.e. The sequence $a(n)$ is called hypergeometric if there exists a rational function $r(x) \in K(x)$ such that $a(n+1) = r(n)a(n)$ a.e. We denote the set of polynomial, rational and hypergeometric sequences over K as $\mathcal{P}(K), \mathcal{R}(K)$ and $\mathcal{H}(K)$ respectively.*

For further reading see [4] or [5].

2.2 Normal form for rational functions

Let us recall the normal form for rational functions.

Theorem 2.3 (cf. [4, Theorem 5.3.1] or [3, Lemma 3.1])**.** *Let K be a field of characteristic zero and $r \in K[x]$ be a nonzero rational function. Then there exist monic polynomials $a, b, c \in K[n]$ and $z \in K \setminus \{0\}$ such that*

$$r(n) = z\frac{a(n)}{b(n)}\frac{c(n+1)}{c(n)},$$

where

(i) $\gcd(a(n), b(n+h)) = 1$ *for every nonnegative integer h,*

(ii) $\gcd(a(n), c(n)) = 1$,

(iii) $\gcd(b(n), c(n+1)) = 1$.

The polynomials a, b, c are constructed in step 2 of Gosper's algorithm (cf [4, page 80]).

2.3 Method of undetermined coefficients

We shall briefly demonstrate the method of undetermined coefficients. Let us consider

$$3y(n+2) - ny(n+1) + (n-1)y(n) = 0. \tag{4}$$

We wish to find all polynomial solutions of (4). Let $y(n) = \sum_{k=0}^{d} a_k n^k$ be a general solution of (4). Assume we know $\deg(y) = d \leq 2$. Then we obtain

$$\begin{aligned}
0 &= 3\left[a_0 + a_1(n+2) + a_2(n+2)^2\right] - n\left[a_0 + a_1(n+1) + a_2(n+1)^2\right] \\
&\quad + (n-1)\left[a_0 + a_1 n + a_2 n^2\right] \\[1ex]
&= 3\left[a_0 + a_1 n + 2a_1 + a_2 n^2 + 4a_2 n + 4a_2)\right] - n\left[a_0 + a_1 n + a_1 + a_2 n^2 + 2a_2 n + a_2\right] \\
&\quad + (n-1)\left[a_0 + a_1 n + a_2 n^2\right] \\[1ex]
&= 3a_0 + 3a_1 n + 6a_1 + 3a_2 n^2 + 12a_2 n + 12a_2 - a_0 n - a_1 n^2 - a_1 n - a_2 n^3 - 2a_2 n^2 \\
&\quad - a_2 n + a_0 n + a_1 n^2 + a_2 n^3 - a_0 - a_1 n - a_2 n^2.
\end{aligned}$$

Hence,

$$\begin{aligned}
0 &= [3a_0 + 6a_1 + 12a_2 - a_0] + n\,[3a_1 + 12a_2 - a_0 - a_1 - a_2 + a_0 - a_1] \\
&\quad + n^2\,[3a_2 - a_1 - 2a_2 + a_1 - a_2] + n^3\,[-a_2 + a_2] \\[1ex]
&= [2a_0 + 6a_1 + 12a_2] + n\,[a_1 + 11a_2]\,.
\end{aligned}$$

We obtain

$$a_1 + 11a_2 = 0, \quad a_0 + 3a_1 + 6a_2 = 0.$$

Let $a_2 = t$, then $a_1 = -11t$ and $a_0 = 27t$. Thus, the general solution of (4) is

$$y(n) = t(27 - 11n + n^2),$$

where t is an arbitrary constant.

2.4 Algorithm Poly

In this part we show how to find all polynomial solutions for an inhomogeneous linear recurrence equation. This will be a neccessary part of algorithm Hyper. Let us consider the problem of finding polynomial solutions for the equation

$$Ly = f, \tag{5}$$

where

$$L = \sum_{i=0}^{r} p_i N^i$$

is a linear recurrence operator with polynomial coefficients $p_i \in \mathcal{P}(K)$, where $p_r, p_0 \neq 0$ and f is a given sequence. In the previous part, we have have seen that assuming we have a bound on the degree of $y(n)$ in (5) it is rather easy to find the general solution by the method of undetermined coefficients. For this matter, our aim will be to find a general bound on the degree of y in (5). First we will rewrite L in terms of the difference operator $\triangle = N - 1$. This gives

$$\begin{aligned} L &= \sum_{i=0}^{r} p_i (\triangle + 1)^i \\ &= \sum_{i=0}^{r} \sum_{j=0}^{i} \binom{i}{j} p_i \triangle^j \\ &= \sum_{j=0}^{r} \sum_{i=j}^{r} \binom{i}{j} p_i \triangle^j \\ &= \sum_{j=0}^{r} q_j \triangle^j, \end{aligned}$$

where $q_j = \sum_{i=j}^{r} \binom{i}{j} p_i$. Let $y(n) = \sum_{k=0}^{d} a_k n^k$ be a general polynomial solution of (5). We now wish to get a bound for the degree d of $y(n)$. For this notice that

$$\begin{aligned}
\triangle^j \, n^k &= (N-1)^j n^k \\
&= (N-1)^{j-1}(N-1)n^k \\
&= (N-1)^{j-1}\left((n-1)^k - n^k\right) \\
&= (N-1)^{j-1}\left(-n^k + \sum_{i=0}^{k} \binom{k}{i} n^{k-i}(-1)^i\right) \\
&= (N-1)^{j-1} \sum_{i=1}^{k} \binom{k}{i} n^{k-i}(-1)^i \\
&= \dots \\
&= (k)_j n^{k-j} + \mathcal{O}(n^{k-j-1}),
\end{aligned}$$

where we write $(a)_n = a(a-1)\cdots(a-n+1)$ for the falling factorial function. Let us write $\mathrm{lc}(P)$ for the leading coefficient of a polynomial P. Then, the above computation yields $\mathrm{lc}(q_j \, \triangle^j \, y(n)) = \mathrm{lc}(q_j \, \triangle^j \, a_d n^d) = \mathrm{lc}(q_j)(d)_j a_d$. Furthermore, since

$$Ly(n) = \left(\sum_{j=0}^{r} q_j \, \triangle^j\right)\left(\sum_{k=0}^{d} a_k n^k\right),$$

we have $\deg(Ly) \le d+b$, where $b = \max_{0\le j\le r}(\deg q_j - j)$. Now we distinguish two cases. First, if $d+b<0$, then $d \le -b-1$ is the desired bound. Notice that b may be negative. On the other hand, if $d+b \ge 0$

$$\begin{aligned}
Ly(n) &= \left(\sum_{j=0}^{r} q_j \, \triangle^j\right)\left(a_d n^d + \sum_{k=0}^{d-1} a_k n^k\right) \\
&= a_d \left(\sum_{\substack{0\le j\le r \\ \deg q_j - j = b}} q_j \, \triangle^j\right) n^d + \mathcal{O}(n^{d+b-1}).
\end{aligned}$$

The coefficient in front of n^{d+b} is

$$a_d \sum_{\substack{0\le j\le r \\ \deg q_j - j = b}} \mathrm{lc}(q_j)(d)_j. \tag{6}$$

If $\deg(Ly) = d + b$, then because of (5) we have $\deg f = d + b$. Thus $d = \deg f - b$ is the desired bound. Otherwise, if $\deg(Ly) < d + b$, then the coefficient in front of n^{d+b} vanishes, i.e. (6) is zero. In other words d is a root of the polynomial

$$\alpha(x) = \sum_{\substack{0 \le j \le r \\ \deg q_j - j = b}} \mathrm{lc}(q_j)(x)_j.$$

Which leaves us with a finite set of candidates for d to check. In any case, we have found a general bound on the degree of y. Next we plug a general polynomial y of degree d into the recurrence equation. By the method of undetermined coefficients we equate the coefficients of like power of n and solve the linear system of $d+1$ unknowns. The algorithm can be easily generalized for higher dimensions. Furthermore, we can even find solutions over an extension field of the original field. You might think of the original field as the field of rational numbers $\mathbb{Q}$ and take a quadratic extension of $\mathbb{Q}$ as the extension field. The generalized algorithm looks as follows.

Algorithm 2.4 (Poly)**.** *Let F be a field of characteristic zero and K an extension field of F. Let $p_i(n)$ be polynomials over F for $i = 0, 1, \ldots, d$. The algorithm will yield the general polynomial solution of $Ly = f$ over K.*

1. *For $0 \le j \le r$ compute*
$$q_j = \sum_{i=j}^{r} \binom{i}{j} p_i.$$

2. *Let*
$$b = \max_{0 \le j \le r} (\deg q_j - j),$$
$$\alpha(x) = \sum_{\substack{0 \le j \le r \\ \deg q_j - j = b}} lc(q_j)(x)_j,$$
$$d_1 = \max\{m \in \mathbb{N} : \alpha(m) = 0\}.$$
Then
$$d = \max\{\deg f - b, -b - 1, d_1\}.$$

3. *Find all $y(n) = \sum_{k=0}^{d} c_k n^k$ that satisfy $Ly = f$ by the method of undetermined coefficients.*

3 Algorithm Hyper

Our procedure of finding a closed form solution for the sum (1) or proving that none exists is going to be as follows. First we find a recurrence of the form (3). This can be done by using Zeilberger's algorithm ([4, Chapter 6]). Next we find all hypergeometric solutions $y(n)$ satisfying

$$Ly = 0 \tag{7}$$

Lastly, we check if any linear combination of $y(n)$ coincides with $f(n)$ for enough consecutive values of n. Let us consider a recurrence relations of the form (7) of order 2, i.e. a reccurence of the form

$$p(n)y(n+2) + q(n)y(n+1) + r(n)y(n) = 0. \tag{8}$$

Assume that $y(n)$ is a hypergeometric solution of (8). Then there exists a rational sequence $S(n)$ such that $y(n+1) = S(n)y(n)$. Substituiting into (8) we obtain

$$p(n)S(n+1)S(n)y(n) + q(n)S(n)y(n) + r(n)y(n) = 0$$

and thus

$$p(n)S(n+1)S(n) + q(n)S(n) + r(n) = 0.$$

By Theorem 2.3 we can write

$$S(n) = z\frac{a(n)}{b(n)}\frac{c(n+1)}{c(n)},$$

where $z \in K \setminus \{0\}$ and $a, b, c \in K[n]$ are monic polynomials that satisfy conditions (i), (ii) and (iii) of the theorem. Using this representation for $S(n)$ we obtain

$$z^2 p(n)\frac{a(n+1)}{b(n+1)}\frac{a(n)}{b(n)}\frac{c(n+2)}{c(n)} + zq(n)\frac{a(n)}{b(n)}\frac{c(n+1)}{c(n)} + r(n) = 0.$$

Clearing denominators we get

$$z^2 p(n)a(n+1)a(n)c(n+2) + zq(n)b(n+1)a(n)c(n+1) + r(n)b(n+1)b(n)c(n) = 0.$$

Notice that $a(n)$ divides the first and the second term of the left hand side of the equation, thus it follows that $a(n) \mid r(n)b(n+1)b(n)c(n)$. By (i) and (ii) we get that $a(n) \mid r(n)$. Similiarly $b(n+1) \mid p(n)a(n+1)a(n)c(n+2)$ and thus $b(n+1) \mid p(n)$. In this matter we found that there are only finitely many possibilities for $a(n)$ and $b(n)$, i.e. the monic factors of $r(n)$ and $p(n-1)$. Next, we collect the equation

$$z^2 \frac{p(n)}{b(n+1)}a(n+1)c(n+2) + zq(n)c(n+1) + \frac{r(n)}{a(n)}b(n)c(n) = 0. \tag{9}$$

Since this polynomial has to vanish for all n, all coefficients have to equal zero. If we simply consider the leading coefficient of this polynomial as a polynomial in n we find a quadratic equation in z. Since p, q, r and a, b are known, all the coefficients but the polynomial c of this quadratic equation are known. However, since the quadratic equation is independent of $c(n)$. Indeed, $c(n)$ does not change the value of the leading coefficient in (9). Thus, we get at most two solutions $z \neq 0$. For fixed $a(n), b(n)$ and z

we can use algorithm Poly to find a nonzero monic polynomial $c(n)$ that satisfies (9). In this case we have found a hypergeometric solution of (7). Otherwise, i.e. if the algorithm does not find such a polynomial $c(n)$ then there does not exist a hypergeometric solution with this choice of $a(n), b(n)$ and z. In either case we continue in this fashion for all possible values of $a(n), b(n)$ and z. If the algorithm finds no solution then this proves that there does not exist a hypergeometric solution of (7). Reccurences of higher order are solved similiarly. Similiarly to algorithm Poly we can find solutions in an extension field. Again, you might think of the original field as the field of rational numbers $\mathbb{Q}$ and take a quadratic extension of $\mathbb{Q}$ as the extension field. Algorithm Hyper as is programmed in Mathematica as the function Hyper[eqn, y[n]] will look for solutions over $\mathbb{Q}$ but we can force it to search for solutions over a quadratic extension of $\mathbb{Q}$. The generalized algorithm looks as follows.

Algorithm 3.1 (Hyper)**.** *Let F be a field of characteristic zero and K an extension field of F. Let $p_i(n)$ be polynomials over F for $i = 0, 1, \dots, d$. The algorithm will yield a hypergeometric solution of $Ly = 0$ over K if one exists and 0 otherwise.*

1. *For all monic factors $a(n)$ of $p_0(n)$ and $b(n)$ of $p_d(n - d + 1)$ over K do for $i = 0, 1, \dots, d$*

$$P_i(n) = p_i(n) \prod_{j=0}^{i-1} a(n+j) \prod_{j=i}^{d-1} b(n+j)$$

   Let

$$m = \max_{0 \le i \le d} \deg P_i(n)$$

   and α_i be the coefficient of n^m in $P_i(n)$ for $i = 0, 1, \dots, d$.

2. *For all nonzero $z \in K$ with*

$$\sum_{i=0}^{d} \alpha_i z^i = 0,$$

   if the recurrence

$$\sum_{i=0}^{d} z^i P_i(n) c(n+i) = 0 \tag{10}$$

   has a nonzero polynomial solution $c(n)$ over K then let

$$S(n) = z \frac{a(n)}{b(n)} \frac{c(n+1)}{c(n)}.$$

3. *Return a nonzero solution $y(n)$ of $y(n+1) = S(n)y(n)$ and stop.*

4. *Return 0 and stop.*

Let us show that this algorithm finds all existing solutions.

Theorem 3.2 ([4, Theorem 8.4.1])**.** *Let $y(n)$ be a nonzero solution of* (7) *such that $y(n+1) = S(n)y(n)$ where $S(n)$ is a rational sequence. Let*

$$S(n) = z\frac{a(n)}{b(n)}\frac{c(n+1)}{c(n)},$$

where a, b, c are monic polynomials satisfying conditions (i), (ii) and (iii) of Theorem 2.3. Let $P_i(n)$ and α_i for $i = 0, 1, \dots, d$ be defined as in algorithm Hyper. Then

1. $\sum_{i=0}^{d} \alpha_i z^i = 0$,
2. *$a(n)$ divides $p_0(n)$,*
3. *$b(n)$ divides $p_d(n-d+1)$ and*
4. *$c(n)$ satisfies* (10).

Proof. Using $y(n+1) = S(n)y(n)$ in (7) gives

$$\sum_{i=0}^{d} p_i(n)\left(\prod_{j=0}^{i-1} S(n+j)\right) y(n) = 0.$$

Dividing by $y(n)$ and using the canonical form of $S(n)$ we obtain

$$\sum_{i=0}^{d} p_i(n)z^i\left(\prod_{j=0}^{i-1} \frac{a(n+j)}{b(n+j)}\right)\frac{c(n+i)}{c(n)} = 0.$$

Clearing denominator and multiplication by $\prod_{j=i}^{d-1} b(n+j)$ gives

$$\sum_{i=0}^{d} z^i p_i(n)c(n+i)\prod_{j=0}^{i-1} a(n+j)\prod_{j=i}^{d-1} b(n+j) = 0. \tag{11}$$

All summands for $i \le i \le d$ contain the factor $a(n)$. Hence $a(n)$ divides the summand for $i = 0$, i.e. $a(n)$ divides $p_0(n)c(n)\prod_{j=0}^{d-1} b(n+j)$. By the conditions on the gcd, it follows that $a(n)$ divides $p_0(n)$. Similiarly $b(n+d-1)$ must divide $z^d p_d(n)c(n+d)\prod_{j=0}^{d-1} a(n+j)$. Thus $b(n+d-1)$ divides $p_d(n)$ in other words $b(n)$ divides $p_d(n-d+1)$. Lastly, considering the leading coefficient in (11) gives $\sum_{i=0}^{d} \alpha_i z^i = 0$. □

References

[1] Mehdi Hassani. "Derangements and applications". In: *J. Integer Seq.* 6.1 (2003), Article 03.1.2, 8.

[2] Marko Petkovšek. "Finding closed-form solutions of difference equations by symbolic methods". PhD thesis. Pittsburgh PA: Carnegie Mellon University, 1990.

[3] Marko Petkovšek. "Hypergeometric solutions of linear recurrences with polynomial coefficients". In: *J. Symbolic Comput.* 14.2-3 (1992), pp. 243–264. ISSN: 0747-7171. DOI: `10.1016/0747-7171(92)90038-6`. URL: `https://doi.org/10.1016/0747-7171(92)90038-6`.

[4] Marko Petkovšek, Herbert S. Wilf, and Doron Zeilberger. $A = B$. A K Peters, Ltd., Wellesley, MA, 1996, pp. xii+212. ISBN: 1-56881-063-6.

[5] Gian-Carlo Rota, David K. Kahaner, and Andrew Odlyzko. "On the foundations of combinatorial theory. VIII. Finite operator calculus". In: *J. Math. Anal. Appl.* 42 (1973), pp. 684–760. ISSN: 0022-247X. DOI: `10.1016/0022-247X(73)90172-8`. URL: `https://doi.org/10.1016/0022-247X(73)90172-8`.