Symbolic Summation

Victor Adamchik

Carnegie Mellon University

Sister Celine's Algorithm

"The interesting problem of the pure recurrence relation for hypergeometric polynomials received probably its first systematic attack at the hands of Sister Mary Celine Fasenmyer" - E.D.Rainville [1]

Mary Celine Fasenmyer

Sister Celine grew up in Pennsylvania's oil country and displayed mathematical talent in high school. For ten years after her graduation she taught and studied at Mercyhurst College in Erie, where she joined the Sisters of Mercy. She pursued her mathematical studies in Pittsburgh and the University of Michigan, obtaining her doctorate in 1946.

After getting her Ph.D., Sister Celine published two papers which expanded on her doctorate work. These papers would be further elaborated by Doron Zeilberger and Herbert Wilf into "WZ theory", which allowed computerized proof of many combinatorial identities. After this, she returned to Mercyhurst to teach and did not engage in further research.

http://en.wikipedia.org/wiki/Mary_Celine_Fasenmyer

Introduction

The idea of Celine's algorithm is to search for a recurrence equation *for the summand* $F(n, k)$

$$S(n) = \sum_{k=0}^{n} F(n, k)$$

We will outline major steps of the algorithm on the following example.

$$S(n) = \sum_{k=0}^{n} 2^k \binom{n}{k} \tag{1}$$

Let us assume (!) that the summand $F(n, k) = 2^k \binom{n}{k}$ satisfies the following recurrence equation

$$a(n)\, F(n, k) + b(n)\, F(n+1, k) + c(n)\, F(n, k+1) + d(n)\, F(n+1, k+1) = 0 \tag{2}$$

where a, b, c, d are unknown polynomials that depend only on n. We convert (2) to a polynomial form

$$a(n)\, 2^k \binom{n}{k} + b(n)\, 2^k \binom{n+1}{k} + c(n)\, 2^{k+1} \binom{n}{k+1} + d(n)\, 2^{k+1} \binom{n+1}{k+1} = 0$$

$$a(n) + b(n) \frac{\binom{n+1}{k}}{\binom{n}{k}} + 2\, c(n) \frac{\binom{n}{k+1}}{\binom{n}{k}} + 2\, d(n) \frac{\binom{n+1}{k+1}}{\binom{n}{k}} = 0$$

by converting binomials into factorials and them simplifying the ratio of factorials. For example,

$$\frac{\binom{n+1}{k}}{\binom{n}{k}} = \frac{(n+1)!}{(n+1-k)!\, k!} \frac{(n-k)!\, k!}{n!} = \frac{(n+1)!}{n!} \frac{(n-k)!}{(n+1-k)!} = \frac{n+1}{n+1-k}$$

Thus, we arrive at

$$a(n) + b(n) \frac{n+1}{n+1-k} + 2\, c(n) \frac{n-k}{k+1} + 2\, d(n) \frac{n+1}{k+1} = 0$$

Multiplying it by $(n+1-k)\,(k+1)$, we obtain the following polynomial in k

$$a\,(n+1-k)\,(k+1) + b\,(n+1)\,(k+1) + 2\,c\,(n-k)\,(n+1-k) + 2\,d\,(n+1)\,(n+1-k) = 0$$

Since the equation must be valid for all k, we equal coefficients by k to zero. This will give us a system of algebraic equation wrt unknown a, b, c and d:

$$\begin{cases} n^2\,(2\,c + 2\,d) + n\,(a + b + 2\,c + 4\,d) + a + b + 2\,d = 0 \\ n\,(a + b - 4\,c - 2\,d) + b - 2\,c - 2\,d = 0 \\ a - 2\,c = 0 \end{cases}$$

This system always has a solution. (why-?). Solving it, yields

$$b = 0, \;\; a = 2\,c, \; d = -c$$

Substituting this back to (2), we get an equation for the summand

$$2\,F(n,\ k)\ +\ F(n,\ k+1)\ -F(n+1,\ k+1) = 0 \tag{3}$$

This is a two-variables recurrence equation. It does not look simple for solving...but won't solve it. We formally sum it wrt k to get a recurrence for the original sum $S(n)$. Due to the following formal manipulations

$$\sum_{k=0}^{n} F(n,\ k) = S(n)$$

$$\sum_{k=0}^{n} F(n,\ k+1) = \sum_{k=1}^{n+1} F(n,\ k) =$$

$$-F(n,\ 0) + F(n,\ n+1) + \sum_{k=0}^{n} F(n,\ k) = S(n) - F(n,\ 0) + F(n,\ n+1)$$

$$\sum_{k=0}^{n} F(n+1,\ k+1) = -F(n+1,\ 0) + \sum_{k=0}^{n+1} F(n+1,\ k) = S(n+1) - F(n+1,\ 0)$$

we get

$$\sum_{k=0}^{n} [2\,F(n,\ k)\ +\ F(n,\ k+1)\ -F(n+1,\ k+1)]\ =$$

$$2\,S(n)\ + S(n) - F(n,\ 0) + F(n,\ n+1) - S(n+1) + F(n+1,\ 0)$$

where the boundary cases $F(n,\ 0)$, $F(n,\ n+1)$ and $F(n+1,\ 0)$ are computed by

$$F(n,\ k) = 2^k \binom{n}{k}$$

Thus,

$$\begin{aligned} F(n,\ 0)\ &=\ 1 \\ F(n,\ n+1)\ &=\ 0 \\ F(n+1,\ 0)\ &=\ 1 \end{aligned}$$

It follows, that formally summing up the recurrence equation (3), we derive a recurrence equation for the sum $S(n)$:

$$\begin{aligned} 3\,S(n) - S(n+1)\ &=\ 0 \\ S(0) &= 1 \end{aligned}$$

This can be easily solved by iteration,

$$S(n) = 3^n = \sum_{k=0}^{n} 2^k \binom{n}{k}$$

Algorithm

The algorithm was developed by Sister Mary Celine Fasenmyer in her Ph.D. thesis in 1945 (see [1, p. 233] or [2, 3]). Given

$$S(n) = \sum_{k=0}^{n} F(n,\ k)$$

We search for the recurrence equation for the summand $F(n,\ k)$ in the form

$$\sum_{i=0}^{I} \sum_{j=0}^{J} a_{i,j}(n) \frac{F(n+j,\ k+i)}{F(n,\ k)} = 0 \tag{4}$$

The important assumptions of this method are

a) the summand is *hypergeometric term* in n and k is rational

$$\frac{F(n+1,\ k)}{F(n,\ k)} \in Q(n) \text{ and } \frac{F(n,\ k+1)}{F(n,\ k)} \in Q(k)$$

A *geometric series* is a series $a_0 + a_1 + ...$ in which the ratio $\frac{a_{k+1}}{a_k}$ of two consecutive terms is constant for all k. In contrast, a *hypergeometric series* is a series in which the ratio of consective terms is not constant, but rather a rational function. It says that the summand is *double hypergeometric* if the above two conditions are met wrt to two parameters.

b) coefficients a, b, c, and d are free of the summation index n

The crucial step of this method is that system for coefficients a, b, c, and d has a nontrivial solution. This will be always a case if the number of variables exceeds the number of equations. This seems is not hard to prove. Recall that we require the summand to be double hypergeometric, in other words, each ratio in the above system is a rational function with respect to n and k. Then, when we equating to zero coefficients by k to zero and see that the system has too few equations, we increase the order of a difference equation.

Increasing I and J in (3) will increase the number of equations as well as the number of unknowns $a_{i,j}(n)$. The number of unknowns grows as $O(I * J)$, however the number of equations can grow at the same (or faster) rate. So we must restrict the class of summands.

Definition.

We say that $F(n,\ k)$ is a *proper hypergeometric term* if it can be written as

$$F(n,\ k) = P(n,\ k) \frac{\prod_{j=1}^{G} \left(a_j\, n + b_j\, k + c_j\right)!}{\prod_{j=1}^{H} \left(u_j\, n + v_j\, k + w_j\right)!}\, x^k$$

where P is a polynomial, and all coefficients $a_j,\ b_j,\ u_j, v_j$ are integers.

Observe,

$$\frac{(n+1)!}{n!} = O(n)$$
$$\frac{(n+2)!}{n!} = O(n^2)$$
$$\frac{(n+p)!}{n!} = O(n^p)$$

Thus, if the summand is a proper hypergeometric term then the number of equations defined by a ratio

$$\frac{F(n+j,\ k+i)}{F(n,\ k)}$$

will grow as $O(I\ +\ J)$. This guarantees that the number of unknowns will eventually exceed the number of equations.

Drawbacks.

The question remains, what is maximal order of the recurrence (3)? There are some estimates, but all of them are impractical. The algorithm that it does not necessarily produces a difference equation of the lowest order.

Mathematica session

In this section we provide all steps of Celine's algorithm the way they can be done in *Mathematica*. As an example, we evaluate

$$S(n) = \sum_{k=0}^{n} 2^k \binom{n}{k}$$

First we define the summand as a *Mathematica* function

```
F[n_, k_] := 2^k Binomial[n, k]
```

Then we assume that the summand satisfies the recurrence equation of the first order

$$a(n)\,F(n,\ k)\ +\ b(n)\,F(n+1,\ k)\ +\ c(n)\,F(n,\ k+1)\ +\ d(n)\,F(n+1,\ k+1) = 0$$

```
a + b F[n+1, k]/F[n, k] + c F[n, k+1]/F[n, k] + d F[n+1, k+1]/F[n, k]
```

We call ***FunctionExpand*** to cancel binomial coefficients

```
FunctionExpand[%]
```

In this input, we convert the rational expression to a polynomial form

```
Numerator[Together[%]]
```

Next, we take coefficients by k and set them to 0:

```
CoefficientList[%, k]
```

```
eq = Map[Equal[#, 0] &, %]
```

As in a style of functional programming, we map a function over the list

```
Map[f, {a, b, c}]
```

Same, for a multivariate function

```
Map[f[#, y] &, {a, b, c}]
```

Here, Function[body] or body & is a pure function with a formal parameter #.

The system of linear equations is efficiently solved by

```
Solve[eq, {a, b, c, d}]
```

Thus, we found the recurrence for the summand

```
Clear[F]; a F[n, k] + b F[n+1, k] + c F[n, k+1] + d F[n+1, k+1] /.
  {b → 0, c → a/2, d → -a/2}
```

which is, after canceling a parameter a,

$$F(n+1,\ k+1) - F(n,\ k+1) - 2\,F(n,\ k) = 0$$

Summing it up wrt k, we arrive at

$$S(n+1) = 3\,S(n)$$

which is solved by

```
RSolve[{S[n+1] == 3 S[n], S[0] == 1}, S[n], n]
```

The Euler Gamma Function

Almost always when you deal with binomial coefficients and/or factorials, *Mathematica* may replace then by the Gamma function

$$n! = \Gamma(n+1) = \int_0^\infty x^n e^{-x} \, dx$$

This functiion is an analytic continuation of $n!$ into a complex plane. By means of Γ-function we can define a fractional factorial

$$\left(\frac{1}{2}\right)! = \frac{\sqrt{\pi}}{2}$$

Here are two functional properties of the function:

$$\Gamma(z+1) = z\,\Gamma(z)$$

$$\Gamma(z)\,\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}, \quad 0 < z < 1$$

Read more on the Gamma function at http://mathworld.wolfram.com/GammaFunction.html

Another Example

$$S(n) = \sum_{k=0}^{n} k \binom{n}{k}$$

```
F[n_, k_] := k Binomial[n, k]
```

Then we assume that the summand satisfies the recurrence equation of the first order

$$a(n)\,F(n, k) + b(n)\,F(n+1, k) + c(n)\,F(n, k+1) + d(n)\,F(n+1, k+1) = 0$$

```
a + b F[n+1, k]/F[n, k] + c F[n, k+1]/F[n, k] + d F[n+1, k+1]/F[n, k]
```

```
Numerator[Together[FunctionExpand[%]]]
```

```
CoefficientList[%, k]
```

```
eq = Map[Equal[#, 0] &, %]
```

```
Solve[eq, {a, b, c, d}]
```

```
Clear[F]; a F[n, k] + b F[n + 1, k] + c F[n, k + 1] + d F[n + 1, k + 1] /.
 %[[1]]
```

Thus, we found the following recurrence for the summand

$$n\,F(n+1,\ k+1) - \ (n+1)\,F(n,\ k+1) - (n+1)\,F(n,\ k) = 0$$

Summing it up wrt k and using these formal manipulations

$$\sum_{k=0}^{n} F(n,\ k) = S(n)$$

$$\sum_{k=0}^{n} F(n,\ k+1) = \sum_{k=1}^{n+1} F(n,\ k) = -F(n,\ 0) + F(n,\ n+1) + \sum_{k=0}^{n} F(n,\ k) = S(n) - F(n,\ 0)$$

$$\sum_{k=0}^{n} F(n+1,\ k+1) = -F(n+1,\ 0) + \sum_{k=0}^{n+1} F(n+1,\ k) = S(n+1) - F(n+1,\ 0)$$

we arrive at

$$n\,(S(n+1) - F(n+1,\ 0)) - \ (n+1)\,(S(n) - F(n,\ 0)) - (n+1)\,S(n) = 0$$

or

$$n\,S(n+1) - \ (n+1)\,S(n) - (n+1)\,S(n) = 0$$

which is solved by

```
RSolve[{n S[n + 1] == 2 (n + 1) S[n], S[0] == 0}, S[n], n]
```

References

[1] E.D.Rainville, *Special Functions*, McMillan Company, NY, 1965.

[2] Sister M. Celine Fasenmyer, A note on pure recurrence relation, *Amer. Math. Monthly*, **56**(1949), 14-17.

[3] Sister M. Celine Fasenmyer, Some generalized hypergeometric polynomials. *Bull. Amer. Math. Soc.*, **53**(1947.), 806-812.