A refinement of Egghe's increment studies: an alternative version

Ronald Rousseau [1,2]

[1]Institute for Education and Information Sciences, IBW, University of Antwerp (UA), Venusstraat 35, Antwerp, B-2000, Belgium

[2]KU Leuven, Leuven B-3000, Belgium

Abstract

In this contribution we show how results obtained in a series of papers by Egghe can be refined in the sense that we need fewer conditions. In these articles Egghe considered a general *h*-type index which has a value *n* if *n* is the largest natural number such that the first *n* publications (ranked according to the number of received citations) have received at least *f(n)* citations, with *f(n)* any increasing function defined on the strictly positive numbers. His results deal with increments I_2 and I_1 defined by: $I_2 n = I_1(n+1) - I_1(n)$ where $I_1(n) = (n+1)f(n+1) - nf(n)$. Our results differ from Egghe's because we also consider $I_0(n) = nf(n)$. This version differs from the original one (Rousseau, 2014) by the fact that we (try to) use standard methods for solving difference equations. These methods are recalled in an appendix.

Keywords: indicator characterizations; increments; forward differences; difference equations

1. Introduction

In a series of papers (Egghe, 2013a, 2013b, 2014) Egghe characterized the *h*-index, the threshold index (also known as the highly cited publications indicator), the Wu-index and several far-reaching generalizations by increments (defined further on) of increasing functions *f*. His investigations are motivated by the following observations. It is well-known that a set of publications *S* has an *h*-index equal to *h*, in the sense of Hirsch (2005), if *h* is the largest natural number such that this set contains *h* publications having at least *h* citations. Among all sets with an *h*-index equal to *h* some can be characterized as having the least total number of citations, namely sets such that *h* publications have exactly *h* citations each and the other publications have no citations, or there do not exist other publications in this set. In this case the publications in the set *S* have received h^2 citations in total. Similarly, sets of publications S_1 among which exactly $h+1$ have exactly $h+1$ citations and the other ones (again, if they exist) have none, form a minimum set of publications with *h*-index equal to $h+1$. Such sets S_1 have received $(h+1)^2$ citations in total. Egghe (2013a) refers to the difference between $(h+1)^2$ and h^2, denoted as $I_1(h)$, as the increment of order *1*:

$$I_1(h) = (h+1)^2 - h^2 = 2h + 1 \tag{1}$$

The increment of order *2* is then defined as:

$$I_2(h) = I_1(h+1) - I_1(h) = 2 \tag{2}$$

This special case is then generalized in (Egghe, 2013a) by considering a general *h*-type index which has a value *n* if *n* is the largest natural number such that the first *n* publications (ranked according to the number of received citations) have received at least *f(n)* citations. One obtains the *h*-index if f(n) = n, the threshold index (with threshold *C*) if *f(n) = C* and the Wu-index if *f(n) = 10n* (Wu, 2010).

In this case the increments of order *1* and order *2* are defined as:

$$I_1(n) = (n+1)f(n+1) - nf(n) \tag{3}$$

and

$$I_2 n = I_1(n+1) - I_1(n) = (n+2)f(n+2) - 2(n+1)f(n+1) + nf(n) \tag{4}$$

We note that these increments are actually special kinds of forward differences (Hosking et al., 1996). This becomes clear if we put *g(n) = nf(n)*. Then $I_1(n) = g(n+1) - g(n)$, which is generally known as the first forward difference of the function *g*, with step *1* (Hosking et al., 1996; Rousseau, 1997). We note that one always has the initial condition *g(0) = 0.f(0) = 0*. Among other results Egghe (2013a) finds that the following two assertions are equivalent:

(1) $I_2(n) = 2$ for all natural numbers *n = 1,2,…*

(2) $f(n) = \frac{2(n-1)f(2)-(n-2)f(1)+(n-1)(n-2)}{n}$ for all natural numbers *n = 1,2, …*

This result leads to a corollary characterizing the case of the *h*-index: the following two assertions are equivalent:

(1) $I_2(n) = 2$ for all strictly positive natural numbers and *f(1) = 1* and *f(2) = 2*

(2) *f(n) = n* for all strictly positive natural numbers

All other results obtained in (Egghe 2013a,b,2014) have a similar structure. They all can be expressed in the form:

The following two assertions

(1) $I_2(n) = A(n)$, a certain function of *n*, for all strictly positive natural numbers and $f(1) = C_1$ and $f(2) = C_2$, where C_1 and C_2 are given numbers;

(2) $f(n) = B(n)$, another function of n, for all strictly positive natural numbers

are equivalent. Egghe (2014) also proves similar results about I_1. In that case results have the form:

The following two assertions

(1) $I_1(n) = A_1(n)$, a certain function of n, for all strictly positive natural numbers and $f(1) = C_3$, where C_3 is a given number;

(2) $f(n) = B_1(n)$, another function of n, for all strictly positive natural numbers

are equivalent.

In this contribution we show how all these results can be refined in the sense that instead of two extra conditions (case I_2) we only need one, and in case of I_1 we do not need another condition. The point is that we also consider $I_0(n) = n\ f(n)$ for all natural numbers n, and hence consider $f(n)$ to be defined in all natural numbers, not just the strictly positive ones. However, the value of f in the point zero plays no role and one is completely free to choose it. From now on we do not repeat this point.

2. Initial results

In this section we introduce our method and illustrate it by proving the characterization of the h-index. Let $f(n)$ be a function defined on the set of all natural numbers.

Definition: Increment functions for f

The increment functions for f, denoted as $I_k^{(f)}(n)$, $k = 0,1,\ldots$ form a sequence of functions, defined on the natural numbers. We denote them as $I_k(n)$ when it is clear which function f is meant or when the exact function does not matter. These increment functions are defined recursively for all natural numbers as:

$$I_0(n) = n.f(n)$$

$$I_k(n) = I_{k-1}(n+1) - I_{k-1}(n),\ k \text{ a strictly positive natural number.}$$

In particular: $I_1(n) = I_0(n+1) - I_0(n) = (n+1).f(n+1) - n.f(n)$, which is equation (3) above. If $f(n) = n$, then $I_0(n) = n^2$, $I_1(n) = 2n+1$, $I_2(n) = 2$ and $I_k(n) = 0$ for $k > 2$. We note that for every f $I_0^{(f)}(0) = 0$.

First, we prove our characterization of the h-index.

Theorem 1.

The following two assertions are equivalent

(1) $I_2(n) = 2$ (the constant function *2*), for all natural numbers *n*, and $f(1) = C$

(2) $f(n) = C + (n-1),\ n > 0$.

Proof. (1) implies (2)

The equation $I_2(n) = 2$ can be rewritten as $(n+2)f(n+2) - 2(n+1)f(n+1) + nf(n) = 2$. This clearly is a linear difference equation. Using the substitution $g(n) = nf(n)$ reduces it to a linear difference equation with constant coefficients:

$$g(n+2) - 2g(n+1) + g(n) = 2$$

Because of the substitution and the requirement $f(1) = C$ we have the following initial values: $g(0) = 0$ and $g(1) = C$. Using the method of the annihilator (explained in the appendix) with $D = (S-1)$ leads, in a first step to $g(n) = C_1n^2+C_2n + C_3$, which then becomes: $g(n) = n^2+C_2n + C_3$. The two initial values require that $C_3 = 0$ and $C_2 = C-1$. Hence the final solution is:

$$g(n) = n^2 + n(C-1) \text{ or } f(n) = C + (n-1)\ (n > 0).$$

Conversely (2) implies (1). If for $n > 0$, $f(n)= C + (n-1)$ (hence $f(1) = C$), then, for $n \neq 0$: $I_2(n) = (n+2)f(n+2) – 2 (n+1)f(n+1) + nf(n) = (n+2)[(n+1) + C] – 2.(n+1).[n + C]+ n.[n-1 + C] = n^2+3n+2-2n^2-2n+n^2-n+ [(n+2)-2(n+1) +n].C = 2$. If $n = 0$, then $I_2(0) = 2.f(2) – 2 f(1) + 0 = 2.(1+C) – 2C = 2.$

This proves Theorem 1.

Corollary 1 (characterization of the *h*-index)

The following two assertions are equivalent

(1) $I_2(n) = 2$ (constant function *2*), for all natural numbers *n* and $f(1) = 1$

(2) $f(n) = n,\ n > 0$.

Note that assertion (2) of this corollary is exactly assertion (2) of Egghe's characterization (as $f(0)$ being free is not an extra condition).

3. Further results

We do not intend to repeat all Egghe's results but will illustrate our approach for the case that $I_2(n) = 0$ (the so-called threshold index), for a case involving I_1 and for a very general case involving I_2 (both taken from (Egghe, 2014)).

Theorem 2 (about the threshold index). The following assertions are equivalent:

(1) $I_2(n) = 0$ (null function) and $f(1) = C$;

(2) *f(n) = C* for *n > 0*.

Proof. (1) implies (2)

Using the same substitution *g(n) = nf(n)* as in Theorem 1, we obtain:

$$g(n+2) - 2g(n+1) + g(n) = 0$$

which is a homogeneous linear difference equation with constant coefficients. Its general solution is:

$$g(n) = C_1 n + C_2$$

The initial conditions *g(0) = 0* and *g(1) = C* lead to: $C_2 = 0$ and $C_1 = C$. Hence the solution of this difference equation is:

$$g(n) = Cn \text{ and hence } f(n) = C, n > 1$$

Conversely, (2) implies (1). If *f(0)* is any number and for all *n > 0, f(n) = C* (in particular *f(1) = C*), then, for $n \neq 0$: $I_2(n) = (n+2)f(n+2) - 2(n+1)f(n+1) + nf(n) = (n+2).C - 2.(n+1).C + n.C = 0$; if *n = 0*, then $I_2(0) = 2f(2) - 2f(1) + 0 = 2C-2C+0 = 0$.

This proves Theorem 2, related to the threshold index.

Theorem 3 (Egghe, 2014)

Let $A_1(n)$, *n = 0,1,2,…* be any sequence. Then the following two assertions are equivalent:

(1) For all natural numbers *n*, $I_1(n) = A_1(n)$

(2) For all *n > 0*: $f(n) = \frac{\sum_{k=0}^{n-1} A_1(k)}{n}$.

We were not able to solve the corresponding difference equation using 'standard' methods. However, we provide a solution using an induction method. Another solution is given in (Rousseau, 2014).

Proof. (1) implies (2). $I_1(n) = A_1(n)$ is equivalent with: $(n+1)f(n+1) - n.f(n) = A_1(n)$.

Taking *n = 0* yields: $f(1) = A_1(0)$. Now we prove the induction step: if, for *n > 0*, $f(n) = \frac{\sum_{k=0}^{n-1} A_1(k)}{n}$ then $f(n+1) = \frac{\sum_{k=0}^{n} A_1(k)}{n+1}$. Indeed: we know that $(n+1)f(n+1) - n.f(n) = A_1(n)$. Hence: $(n+1)f(n+1) = \left(\sum_{k=0}^{n-1} A_1(k)\right) + A_1(n)$. From this equality we see that $f(n+1) = \frac{\sum_{k=0}^{n} A_1(k)}{n+1}$. There is no condition on *f(0).*

(2) implies (1). If *n > 0*, then $I_1(n) = (n+1)f(n+1) - n.f(n) = \sum_{k=0}^{n} A_1(n) - \sum_{k=0}^{n-1} A_1(n) = A_1(n)$. Finally, $I_1(0) = 1f(1) - 0 = f(1) = A_1(0)$.

This proves Theorem 3.

Corollary 2 (another characterization of the h-index)

Taking $A_1(n) = 2n+1$ yields that the following two assertions are equivalent:

(1) For all natural numbers n, $I_1(n) = 2n+1$

(2) For all $n > 0$: $f(n) = n$.

As we did not prove Theorem 3 (at least not in the context of the methods recalled in the Appendix) we now provide such a proof of this corollary (hence treating this corollary as if it were a theorem).

Proof. (1) implies (2)

Using the same substitution $g(n) = nf(n)$ as in Theorem 1, we obtain:

$$g(n+1) - g(n) = 2n + 1$$

which is a non-homogeneous first order linear difference equation with constant coefficients. The annihilator is $(S-1)^2$. Applying this annihilator leads to $(S-1)^3 g(n) = 0$. Hence, $g(n) = C_1 + C_2n + C_3n^2$. Substituting this preliminary solution in the difference equation gives: $C_1 + C_2(n+1) + C_3(n+1)^2 - C_1 - C_2n - C_3n^2 = 2n+1$. From this we derive that $C_3 = 1$ and $C_2 = 0$. The initial value $g(0) = 0$ leads to $C_1 = 0$. Hence the solution is $g(n) = n^2$ and thus $f(n) = n$ $(n>0)$.

(2) implies (1). If $n > 0$, then $I_1(n) = (n+1)f(n+1) - n.f(n) = n^2 + 2n + 1 - n^2 = 2n+1$. Finally, $I_1(0)= 1f(1)-0 = f(1)= 1$ (the value of the function $2n+1$ in the point $n = 0$).

We also note the following corollary of Theorem 3.

Corollary 3

Let $A_1(n)$ be the constant sequence with value K. Then the following two assertions are equivalent:

(1) For all natural numbers n, $I_1(n) = K$

(2) For all $n > 0$: $f(n) = K$.

Finally, we consider a general case for I_2.

Theorem 4 (Egghe, 2014)

Let *A(n), n = 0,1,2,…* be any sequence. Then the following two assertions are equivalent:

(1) For all natural numbers *n*, $I_2(n) = A(n)$ and *f(1) = C*

(2) For all *n > 0*: $f(n) = C + \frac{\sum_{k=0}^{n-1}(n-1-k)\,A(k)}{n}$.

Also for this equation we were not able to solve the corresponding difference equation using 'standard' methods. Again we provide a proof by induction and refer the reader to (Rousseau, 2014) for a more elegant proof.

Proof

(1) implies (2). First we note that the equation in (2) holds for *n = 1*. For all *n*, $I_2(n) = A(n)$ is equivalent with the expression: for all *n*, *(n+2)f(n+2) – 2 (n+1)f(n+1) + nf(n) = A(n).* For *n = 0* this yields: *2f(2) – 2 f(1) = A(0)* or: *f(2) = f(1)+A(0)/2 = C + A(0)/2*. The cases *f(1)* and *f(2)* form the base of the induction. Just as a verification we also show how one finds *f(3).* For *n = 1* we have: *3f(3)-2.2 f(2) + f(1) = A(1) or 3f(3)-4C-2A(0) + C = A(1)*. Consequently *f(3) = C + (A(1) + 2A(0))/3*.

Next we prove the induction step. If, for *n > 0*, $f(n) = C + \frac{\sum_{k=0}^{n-1}(n-1-k)\,A(k)}{n}$ and $f(n+1) = C + \frac{\sum_{k=0}^{n}(n-k)\,A(k)}{n+1}$ then $f(n+2) = C + \frac{\sum_{k=0}^{n+1}(n+1-k)\,A(k)}{n+2}$.

As *(n+2)f(n+2) – 2 (n+1)f(n+1) + nf(n) = A(n)* we obtain: *(n+2)f(n+2) = 2* $\left((n+1)C + \sum_{k=0}^{n}(n-k)\,A(k)\right) - nC - \sum_{k=0}^{n-1}(n-1-k)\,A(k) + A(n)$.

From this equality we obtain: $f(n+2) = C + \frac{\sum_{k=0}^{n+1}(n+1-k)\,A(k)}{n+2}$. This proves the induction step.

(2) implies (1). Equation (2) implies that *f(1) = C*. For *n > 0*, $I_2(n)$ *= (n+2)f(n+2) – 2 (n+1)f(n+1) + nf(n) = (n+2)C +* $\sum_{k=0}^{n+1}(n+1-k)\,A(k)$ *-2(n+1)C -* $2.\left(\sum_{k=0}^{n}(n-k)\,A(k)\right)$*+ nC +* $\sum_{k=0}^{n-1}(n-1-k)\,A(k)$ *= 0.C + A(n).* Further: $I_2(0)$ *= 2f(2) - 2f(1) = 2C + A(0) – 2C = A(0).*

This proves Theorem 4.

Remark. For *k > 1*: $\sum_{k=0}^{n-1}(n-1-k)\,A(k) = \sum_{k=0}^{n-2}(n-1-k)\,A(k)$, as for *k=n-1* the factor *(n-1-k)* is zero. We wrote the first form as it is also meaningful for *n=1*, while the second is not.

As a corollary we consider the special case that *A(n)* is a constant sequence: *A(n) = K*.

Corollary.

The following two assertions are equivalent:

(1) For all natural numbers n, $I_2(n) = K$ and $f(0) = C$

(2) For all $n > 0$: $f(n) = C + \frac{n-1}{2} K$.

As we did not prove Theorem 4 (at least not in the context of the methods recalled in the Appendix) we provide such a proof here.

Proof.

(1) implies (2). The substitution $g(n) = nf(n)$ leads to:

$$g(n+2) - 2g(n+1) + g(n) = K$$

Using the method of the annihilator (explained in the Appendix) with $D = (S-1)$ leads to $g(n) = Kn^2/2 + C_2 n + C_3$. The two initial values $g(0) = 0$ and $g(1) = C$ require that $C_3 = 0$ and $C_2 = C-K/2$. Hence the final solution is:

$$g(n) = (K/2)n^2 + n(C-K/2) \text{ or } f(n) = C + K(n-1)/2 \ (n > 0).$$

(2) implies (1). Equation (2) implies that $f(1) = C$. For $n > 0$, $I_2(n) = (n+2)f(n+2) - 2(n+1)f(n+1) + nf(n) = (n+2)C+(n+2)(n+1)K/2 -2(n+1)C - (n+1)nK + nC + n(n-1)K/2 = 0C+K$. Further: $I_2(0) = 2f(2) - 2f(1) = 2C + K - 2C = K$.

In particular we note that if $K = 0$, $f(n)$ in equation (2) of the previous corollary is equal to C (the threshold index); if $K = 1$, $f(n) = C + (n-1)/2$ (a case not explicitly considered by Egghe), and if $K=2$, $f(n) = C + (n-1)$ (the *h*-index).

Remarks

1. If f is strictly positive and increasing then I_0 is strictly increasing.

Indeed: if f is positive and increasing then $f(n+1) \geq f(n) > n/(n+1).f(n)$, and hence: $(n+1)f(n+1) - n\, f(n) > 0$, or $I_0(n) > 0$.

2. In the previous sections we considered I_1 and I_2. Now we briefly mention the (trivial) case I_0. The relation $I_0(n) = B_0(n)$ for all n is equivalent with $n.f(n) = B_0(n)$ for all n. This implies that $B_0(0)= 0$ and $f(n) = B_0(n)/n$ for all $n > 0$.

3. As Egghe's increments are actually forward differences they satisfy the same relations such as: for all natural numbers k and n:

$$I_k(n) = \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} (n+j) f(n+j) \qquad (5)$$

This formula actually provides a non-recursive definition of the sequence I_k.

4. Conclusion

We considered some variations on the study of so-called increments by Egghe (2013a,b, 2014). In this way we provided slightly new results. In this alternative version of (Rousseau, 2014) we tried to apply the standard theory of solving linear difference equations with constant coefficients. Yet, we were not always successful. We note that, because these difference equations have initial conditions we could have applied the theory of z-transforms, at least for those equations we were able to solve (Oppenheim et al., 1983). Yet, we did not go that far.

Acknowledgements

The author is supported by NSFC Grant No 71173185. He further thanks Leo Egghe and Raf Guns for helpful suggestions and for giving him access to (Egghe, 2014) before publication.

References

Egghe, L. (2013a). A mathematical characterization of the Hirsch-index by means of minimal increments. *Journal of Informetrics*, 5(3), 439-445.

Egghe, L. (2013b). Mathematical characterizations of the Wu- and Hirsch- indices using two types of minimal increments. *Proceedings of ISSI 2013* (J. Gorraiz, E. Schiebel, C. Gumpenberger, M. Hörlesberger, H. Moed, Eds.), pp. 1159-1169. Vienna, Austrian Institute of Technology.

Egghe, L. (2014). A general theory of minimal increments for Hirsch-type indices and applications to the mathematical characterization of Kosmulski-indices (to appear).

Hirsch, J.E. (2005). An index to quantify an individual’s scientific research output. *Proceedings of the National Academy of Sciences of the USA*, 102(46), 16569-16572.

Hosking, R.J., Joe, S., Joyce, D.C., Turner, J.C. (1996). *First steps in numerical analysis (2nd edition)*. London: Arnold.

Oppenheim, A.V., Willsky, A.S., Young, I.T. (1983). *Signals and systems*. London: Prentice-Hall.

Rousseau, R. (1997). *Numerical Mathematics*. Course notes KHBO (in Dutch).

Rousseau, R. (2014). A refinement of Egghe’s increment studies. *Journal of Informetrics*, 8(1), to appear.

Wu, Q. (2010). The w-index: A measure to assess scientific impact by focusing on widely cited papers. *Journal of the American Society for Information Science and Technology,* 61(3), 609-614.

Appendix

Basic methods to solve linear difference equations with constant coefficients.

A1. Definitions

A real difference equation is a recursion relation for which one wants to find an unknown real sequence $(a_n)_n$. If the equation contains only powers of a_n then the highest power occurring is called the degree of the difference equation. A difference equation of degree one is called a linear difference equation. A difference equation expressing a relation between $(a_n, a_{n+1}, \ldots, a_{n+k})$ with k a fixed natural number is a difference equation of order k. Consequently, the general form of a linear difference equation of order k is:

$$a_{n+k} + b_1 a_{n+k-1} + \ldots + b_k a_n = f_n \qquad (6)$$

The coefficients b_i in equation (6) may in general depend on n. If this is not the case (the b_i are constants) then equation (6) is a linear difference equation with constant coefficients. If $f_n = 0$ for each n (the sequence $(f_n)_n$ is the null sequence) then the difference equation is said to be homogeneous; otherwise it is non-homogeneous. A homogeneous difference equation of order k has the following form:

$$a_{n+k} + b_1 a_{n+k-1} + \ldots + b_k a_n = 0 \qquad (7)$$

A2. Solving a homogeneous linear difference equation with constant coefficients

A2.1 The characteristic equation

Introducing the operator S, defined as $S(a_n) = a_{n+1}$, equation (7) can be rewritten as:

$$(S^k + b_1 S^{k-1} + b_2 S^{k-2} + \cdots + b_{k-1} S + b_k) a_n = 0 \qquad (8)$$

The characteristic equation of equation (8) is then defined as:

$$r^k + b_1 r^{k-1} + b_2 r^{k-2} + \cdots + b_{k-1} r + b_k = 0 \qquad (9)$$

A2.2. The second order linear difference equation with constant coefficients

The characteristic equation for this case is: $r^2 + b_1 r + b_2 = 0$.

- If this equation has two different real roots s_1 and s_2 then it can be shown that the solution of the corresponding difference equation is (we give the *n-th* term of the solution sequence $(a_n)_n$):

$$a_n = C_1 s_1^n + C_2 s_2^n$$

with two arbitrary real constants C_1 and C_2. If initial values for the sequence $(a_n)_n$ are given then these constants can be determined.

• If this equation has a double (real) root $s = -b_1/2$ then the solution of the corresponding difference equation is:

$$a_n = C_1 s^n + C_2\, n s^n = C_1\left(-\frac{b}{2}\right)^n + C_2\, n\left(-\frac{b}{2}\right)^n$$

• Finally, if the characteristic equation has two complex conjugate roots $s_1 = |s|e^{it}$ and $s_2 = |s|e^{-it}$, then the solution is:

$$a_n = C_1 |s|^n \cos(nt) + C_2 |s|^n \sin(nt)$$

A2.3 An example: Fibonacci's sequence

This famous sequence is given by the recursion relation $f_n - f_{n-1} = f_{n-2}$ and initial values $f_1=1$ and $f_2 = 1$. This recursion can be rewritten as a homogeneous difference equation $f_{n+2} - f_{n+1} - f_n = 0$ with characteristic equation $r^2 - r - 1 = 0$. Its solutions are:

$$r_{1,2} = \frac{1 \pm \sqrt{1+4}}{2}$$

Hence $f_n = C_1\left(\frac{1+\sqrt{5}}{2}\right)^n + C_2\left(\frac{1-\sqrt{5}}{2}\right)^n$. As $f_1 = f_2 = 1$, it can be shown that:

$$a_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$$

A2.4 Solving the general linear homogeneous difference equation with constant coefficients

Also here one starts from the characteristic equation and finds its roots. Real roots that occur just once are treated as s_1 and s_2 before. Concretely, if $s_1, \ldots, s_j$ are different unique real roots, then $(a_n)_n = (C_l s_l^n)_n, l = 1, \ldots, j$ are j different independent solutions. If s_t is a real root with multiplicity m, then $C_1 s_t^n, C_2 n s_t^n, \ldots, C_m n^{m-1} s_t^n$ are m independent solutions. Complex root always occur in conjugate pairs and are treated similarly (and are not be needed in this article).

A3. Solution of a non-homogeneous difference equation with constant coefficients

We will briefly discuss the annihilator method.

A3.1 Consider the non-homogeneous difference equation

$$(S^k + b_1 S^{k-1} + b_2 S^{k-2} + \cdots + b_{k-1} S + b_k) a_n = f_n \tag{10}$$

This method consists in finding an operator D (the annihilator) such that $Df_n = 0$. Applying this operator to equation (10) reduces it to a homogeneous difference equation. Besides the problem of finding this annihilator D this method also has the disadvantage of introducing extra variables C which have to be eliminated by substituting the preliminary solution into the equation. Luckily, many difference equations have a sequence $(f_n)_n$ for which an annihilator is known. In particular if $(f_n)_n$ is a polynomial of degree p in the variable n, then the annihilator is $(S\text{-}1)^{p+1}$.

A3.2 An example

Solve: $a_{n+1} = a_n + 2n$ (n ≥ 1), or: $(S-1)a_n = 2n$.

The right-hand side is a polynomial of degree one, hence its annihilator is $(S-1)^2$ (this can easily be verified).

Applying the annihilator leads to: $(S-1)^3\, a_n = 0$ and a (preliminary) solution:

$$a_n = C_1 + C_2 n + C_3 n^2.$$

Substituting this preliminary solution into the original difference equation leads to:

$$(S-1)(\, C_1 + C_2 n + C_3 n^2) = 2n$$

$$\Rightarrow C_1 + C_2(n+1) + C_3(n+1)^2 - C_1 - C_2 n - C_3\, n^2 = 2n$$

$$\Rightarrow C_2 + 2nC_3 + C_3 = 2n$$

Hence, we only have a solution if $C_3 = 1$ and $C_2 = -1$

We conclude that the general solution of this difference equation is: $a_n = C_1 - n + n^2$. If an initial value is given one can determine the arbitrary constant C_1.