TRANSFORMAREA RICHARDSON - SHANKS

THE RICHARDSON – SHANKS TRANSFORMATION

Preparator univ. drd, Cristina Popîrlan
Preparator univ. drd, Claudiu-Ionuţ Popîrlan
University of Craiova
Matematics-Informatics Faculty,
Computer Science Departament

Abstract: The general structure of the paper is the following:

- Definition of extrapolation methods;
- The Richardson transformation, definition and algorithm;
- The Shanks transformation, representation;
- The Richardson – Shanks transformation, definition and algorithm.

Keywords: extrapolation, method, convergence, transformation.

1. Introduction

An important problem is to find or to approximate the limit of an infinite sequence $\{x_n\}$. The elements x_n of this sequence can show up in different forms. To solve this problem we can use an extrapolation method or a convergence acceleration method.

The Aitken Δ^2 – process and the Richardson transformation are the most popular representatives of extrapolation methods.

Let the sequence $\{x_n\}$ be such that :

$x_n=x+a\lambda^n+r_n$, where x, a, λ, r_n are in general complex scalar.

The Aitken Δ^2 – process applied to $\{x_n\}$ produces the sequence $\{X_n\}$:

$$X_n=\frac{x_n x_{n+2}-x^2{}_{n+1}}{x_n-2x_{n+1}+x_{n+2}}$$

Definition: Let the sequence $\{x_n\}$and the sequence generated by applying the extrapolation method $\{X_n\}$. we assume that $\lim\limits_{n\to\infty}\frac{|X_n-x|}{|x_{\zeta_n}-x|}=0$.

We say that the extrapolation method accelerates the convergence of $\{x_n\}$. The number $\frac{|X_n-x|}{|x_{\zeta_n}-x|}$ *is called the acceleration factor of $\{X_n\}$.*

Remark: A good extrapolation method is one with a factor of acceleration that goes to 0.

An important subject regarding extrapolation methods is to analyze the convergence and the stability of the method. The first step to solve the problem of convergence is to find the conditions that need to be impose to the sequence. Another step is to analyze the errors (X_n - x) that arise from the algorithm considering those preliminaries conditions.

When we compute the sequence $\{X_n\}$ in finite precision arithmetic we obtain a sequence $\{X'_n\}$ that is different from $\{X_n\}$, the exact transformed sequence. This is caused by errors in x_n. The stability problem is to determinate how much $\{X'_n\}$ differs from $\{X_n\}$, so we want to estimate $|X'_n - X_n|$.

It is proved that the cumulative error $|X'_n - x|$ is at least of the order of the corresponding theoretical error $|X_n - x|$.

When we apply an extrapolation method to a convergent sequence, numerically, we must be able to compute the sequence $\{X_n\}$ without $|X'_n - X_n|$ becoming unbounded for increasing n.

It is very important the development of an efficient algorithm for implementing extrapolation methods. We must try to use a small number of arithmetic operations.

2. The Richardson transformation

In some problems, the sequence $\{x_n\}$ is related to a function f(y) that we know. So we have the next relation:

$$x_n = f(y_n), \text{ for } n=0,1,\dots$$

The sequence assumes that if $\lim_{n\to\infty} x_n = x$ then:

$$f(y) = x + \sum_{k=1}^{s} \alpha_k y^{\sigma_k} + O(y^{\sigma_{k+1}})$$

where $\sigma_k \neq 0, k = 1,2,\dots,s+1.$

The idea of Richardson extrapolation method is to eliminate y^{σ_k} from the relation above and to obtain a new, a better approximation of x.

We have the following equivalent relation:

$$f(y) \approx x + \sum_{k=1}^{s} \alpha_k y^{\sigma_k}$$

An algorithm for this problem can be:

- Set $f_0^{(j)} = f(y_j), j = 0,1,2,\dots$
- Set $c = \omega^{\sigma_n}$ and then compute $f_n^{(j)} = \dfrac{f_{n-1}^{(j+1)} - c_n f_{n-1}^{(j)}}{1-c_n}$, j=0, 1, …, n=1, 2, …

The theorems from below prove the convergence of the Richardson transformation.

Theorem: Let the function $f(y) = x + \sum_{k=1}^{s} \alpha_k y^{\sigma_k} + O(y^{\sigma_{k+1}})$.

In case the integer s is finite and largest possible, for $n \geq s+1$ *we have the relation:*

$$f_n^{(j)} - x = O(y_j^{\sigma_{s+1}}) \text{ when } j \to \infty.$$

In case we have s=1,2,3,…, we have the relation:

$$f_n^{(j)} - x \approx \sum_{k=n+1}^{\infty} U_n(c_k)\alpha_k y_j^{\sigma_k} \text{ when } j \to \infty.$$

Theorem: Let the function $f(y) = x + \sum_{k=1}^{s} \alpha_k y^{\sigma_k} + O(y^{\sigma_{k+1}})$.

In case s is finite and largest possible then we have:

$$f_n^{(j)} - x = O(\omega^{\sigma_{s+1} n}) \text{ when } n \to \infty.$$

In case we have s=1,2,3,…, j fixed we have the relation:

$$f_n^{(j)} - x = O(\omega^{\mu n}) \text{ when } n \to \infty.$$

A particular case of the Richardson transformation is the polynomial Richardson extrapolation. We have the polynomial function f, and the constants t_k.

Using Aitken formula we obtain a new polynomial function:

$$P_{k+1}^{(n)}(y) = \frac{P_k^{(n+1)}(y)(y-t_n) - P_k^{(n)}(y)(y-t_{n+k+1})}{t_{n+k+1} - t_n}$$

Let $R_k^{(n)} = P_k^{(n)}(0).$ We have the relation:

$$R_{k+1}^{(n)} = \frac{t_n R_k^{(n+1)} - t_{n+k+1} R_k^{(n)}}{t_n - t_{+k+1n}}$$

$$\text{and } R_0^{(n)} = f(t_n)$$

We have the following theorem:

Theorem: If we have a polynomial function f and the constants t_k then:

$$\lim_{n\to\infty} R_k^{(n)} = \lim_{n\to\infty} f(t_n),\ \forall k$$

$$\lim_{k\to\infty} R_k^{(n)} = \lim_{p\to\infty} f(t_p),\ \forall n$$

Theorem: Let a polynomial function f, a sequence of constants $\{t_n\}$, $\lim_{n\to\infty} t_n = 0$, $\frac{t_p}{t_{p=1}} \geq a, \forall p, a \geq 1$.

The sequence $\left(R_{k+1}^{(n)}\right)_n$ *converges to* $x = \lim_{n\to\infty} f(t_n)$ *faster then the sequence* $\left(R_k^{(n)}\right)_n$ *if and only if*

$$\lim_{n\to\infty} \frac{R_k^{(n+1)} - x}{R_k^{(n)} - x} = \lim_{n\to\infty} \frac{t_{n+k+1}}{t_n}$$

3. The Shanks transformation

The Shanks transformation is a very effective extrapolation method. It is an extrapolation method that we apply to the sequence $\{x_n\}$to obtain the sequence $\{X_n=e_k(x_n)\}$ where $e_k(x_n)$ has the form of a ratio of two determinants:

$$e_k(x_n) = \frac{\begin{vmatrix} x_n & x_{n+1} & \cdots & x_{n+k} \\ \Delta x_n & \Delta x_{n+1} & \cdots & \Delta x_{n+k} \\ \vdots & \vdots & \vdots & \vdots \\ \Delta x_{n+k-1} & \Delta x_{n+k} & \cdots & \Delta x_{n+2k-1} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \Delta x_n & \Delta x_{n+1} & \cdots & \Delta x_{n+k} \\ \vdots & \vdots & \vdots & \vdots \\ \Delta x_{n+k-1} & \Delta x_{n+k} & \cdots & \Delta x_{n+2k-1} \end{vmatrix}}$$

where $\Delta x_n = x_{n+1} - x_n$.

Remark: k is the order of the Shanks transformation.

This form is obtained by solving with Cramer method the next nonlinear system:

$$x_r = e_k(x_n) + \sum_{i=1}^{k} \beta_i \Delta_{r+i-1},\ n < r \leq n+k.$$

To use in an algorithm that calculate $e_k(x_n)$ the form above is very difficult. The ε - algorithm of Wynn is very efficient as it produces all of $e_k(x_n)$:

$$\varepsilon_{-1}^{(n)} = 0, n = 0,1,\cdots$$

$$\varepsilon_0^{(n)} = x_n, n = 0,1,\cdots$$

$$\varepsilon_{k+1}^{(n)} = \varepsilon_{k-1}^{(n+1)} + \frac{1}{\varepsilon_k^{(n+1)} - \varepsilon_k^{(n)}}, n, k = 0,1,\cdots$$

We have the following relations:

$$e_k(x_n) = \varepsilon_{2k}^{(n)}, n, k = 0,1,\cdots$$
$$\varepsilon_{2k+1}^{(n)} = \frac{1}{e_k(\Delta x_n)}$$

4. The Richardson – Shanks transformation

Let say we have the following sequence $\{x_n\}$ to witch we associate the relation:

$$x_n \approx x + \sum_{k=1}^{\infty} \alpha_k \lambda_k^n \text{ , where } n \to \infty .$$

We assume that λ_k are known, λ_k are distinct for all k, $\lim_{k\to\infty} \lambda_k = 0$.

The first step in this extrapolation method is to eliminate λ_k from the relation from above., to obtain a new sequence $\{X_n\}$. We do this using the Richardson extrapolation method.

$$a_{-1}^{(n)} = 0, n = 0,1,\cdots$$
$$a_0^{(n)} = x_n, n = 0,1,\cdots$$
$$a_p^{(n)} = \frac{a_{p-1}^{(n+1)} - \lambda_p a_{p-1}^{(n)}}{1-\lambda_p}, n = 0,1,\cdots, p = 1,2,\cdots,s$$
$$X_n = a_s^{(n)}, n = 0,1,\cdots$$

We obtain the following relations:

$$X_n \approx x + \sum_{k=s+1}^{\infty} \tilde{\alpha}_k \lambda_k^n, \; n \to \infty$$

where $\tilde{\alpha}_k = \alpha_k \prod_{i=1}^{s} \frac{\lambda_k - \lambda_i}{1-\lambda_i}$.

The second step of this extrapolation method is to apply the Shanks transformation to $\{X_n\}$. We denote the resulting approximations $\varepsilon_{2n}^{(j)}$ by $\tilde{\varepsilon}_{2n}^{(j)}(\{X_k\})$.

An interesting application for the Richardson – Shanks transformation is to problems in which a sequence $\{x_n\}$ satisfies:

$$x_n \approx x + \sum_{k=1}^{s} \beta_k \mu_k^n + \sum_{k=1}^{\infty} \delta_k \nu_k^n \text{ , where } n \to \infty$$

We want to calculate:

$$X_n = x + \sum_{k=1}^{s} \beta_k \mu_k^n$$

First, we apply the Richardson – Shanks transformation to this sequence with the sequence $\{x_n\}$ and $\lambda_k = \mu_k$, k=1,2,. . ., s. The approximation of x is $\tilde{\varepsilon}_{2n}^{(j)}(\{X_k\})$.

So, we have the sequence:

$$X_n \approx X + x\left(\frac{1}{\mu_r}\right)^n + \sum_{\substack{k=1\\k\neq r}}^{s} \beta_k \left(\frac{\mu_k}{\mu_r}\right)^n + \sum_{k=1}^{\infty} \delta_k \left(\frac{\nu_k}{\mu_r}\right)^n, \; n \to \infty$$

Next, we apply the Richardson – Shanks transformation to the sequence $\{X_n\}$ with $\lambda_k = \frac{\mu_k}{\mu_r}$ for $1 \le k \le s, k \ne r$, and $\lambda_r = \frac{1}{\mu_r}$ and we obtain the approximations $\overset{\sim}{\varepsilon}{}_{2n}^{(j)}(\{x_k \mu_r^{-k}\})$.

After we approximate x and β_r, we form the sequence:

$$\tilde{X}{}_n^{(j)}(m) = \overset{\sim}{\varepsilon}{}_{2n}^{(j)}(\{x_k\}) + s\sum_{r=1} \overset{\sim}{\varepsilon}{}_{2n}^{(j)}(\{x_k \mu_r^{-k}\})\mu_r^m$$

which approximate X_n.

Sequence of this type arises when we want to solve numerically time-dependent problems with solutions that are periodic in time. This may be the result of the periodicity built directly into the equations associated to the problem. This can also result from the boundary conditions that are periodic in time.

In some problems, we can have only a part of the sequence $\{x_n\}$, and we want to approximate this sequence. This is possible if we apply the Richardson – Shanks transformation to the sequence $\{X'_k\}_{k=0}^{\infty}$, where $B'_k = B_{j+k}$, k=0,1,2,….

Example: We consider the linear system of differential equations:

$$y'(t) = Cy(t) + f(t),\ t>0;\ y(0)=y_0$$

where C is a constant matrix, the function f is periodic.

Because we have:

$$y(t) = e^{Ct} y_0 + \int_0^1 e^{C(t-s)} f(s)ds$$

we can prove that $y(t) = y^{trans}(t) + y^{steady}(t)$ where $y^{trans}(t)$ is transient and has the limit 0, and $y^{steady}(t)$ is periodic and has the same period as the function f.

Using other formulas we obtain the relation:

$$y_m = \sum_{k=-L}^{L} b_k (\omega^k)^m + \sum_{k=1}^{q} d_k \nu_k^m$$

Now we can apply the Richardson – Shanks transformation to the sequence y_m .

5. Open problems and future works

From this article the following problems arise:

- How to implement an algorithm for same applications. If we consider a class of problems then we can implant a good algorithm, but if we get another problem that isn't in that class we can't use the same program.

This is a very important problem: how to create an algorithm and how to implement it so that we can use it to solve different classes of problems.

For example if we want to solve a system of differential equations, we must implement the same algorithm in different ways considering the form of the system and the initial conditions.

- Another problem is to improve the algorithm for the Richardson – Shanks transformation so that we can use it in the same form for different classes of problems.

An algorithm for the Richardson – Shanks transformation can be this:

1. transform the problem so that we obtain a sequence like this:

$$x_n \approx x + \sum_{k=1}^{\infty} \alpha_k \lambda_k^n,\ \text{where}\ n \to \infty.$$

2. apply the Richardson transformation to this sequence;

3. apply the Shank transformation to the sequence that result from the second step;
4. extract the solution of the problem.

There are a lot of problems when we must transform the problems in order to obtain a sequence to which we can apply first the Richardson transformation. If this problem is solved then the rest is easy.

REFERENCES

Brezinski C.	*Convergence d'une forme confluente de l'ε-algorithm*, C.R. Acad. Sci., Paris, 1971
Brezinski C.	*L'ε-algorithm et les suites totalement monotones et oscillantea*, C.R. Acad. Sci., Paris, 1973
Brezinski C.	*Acceleration de la convergence en analyse numerique*, Springer-Verlag, Berlin, 1977
Ghindeanu Mihaela, Popîrlan Claudiu Ionuţ	"A Natural Language Processing System using Java-Prolog Technology", *Research Notes in Artificial Intelligence and Digital Communications*, Vol.105, Craiova, 2005
Wynn P.	*On the convergence and stability of the epsilon algorithm*, SIAM. 1966
Popîrlan Claudiu Ionuţ	"A Java Implementation of Modeling Results About Stratified Graphs", *Research Notes in Artificial Intelligence and Digital Communications*, Vol.104, Craiova, 2004
Puliafito A.	"A Java-based Distributed Network Management Architecture", *3rd Int'l Conf. on Computer Science and Informatics (CS&I'97)*, 1997.
Sidi Avram	*Practical extrapolation methods*, Cambridge, 2003