A UNIFIED VIEW OF APPROXIMATE RATIONAL ARITHMETIC AND RATIONAL INTERPOLATION

A. Miola

Istituto di Analisi dei Sistemi ed Informatica del CNR
Via Buonarroti 12, 00185 Roma.

Abstract

In this paper we cover the problem of approximation of numbers and of functions by presenting some well known results in a unified view that could help in better understanding the algebraic bases of the problem. In fact the extended Euclid's algorithm happens to be the unique and common tool solving the approximation problems both for numbers and for functions.

1. Introduction

One of the misleading beliefs about computer algebra has been for long time about the unique possibility and the total interest in that field only for exact computation and for the closed form solution of given problems. Even if this aspiration for exact computation mainly characterizes the algebraic manipulation field, nevertheless the interface between numeric and symbolic computations has been deeply studied from the very beginning. Moreover the use of algebraic approaches in many typical numeric problem has been very interesting and fruitful, and it is sufficient to recall here one problem for all: the polynomial zeroes determination.

In this paper we will cover the problem of approximation of numbers and of functions by presenting some well known results in a unified view that we hope would help in better understanding the role of algebraic computation in a direction which happens to be not completely known.

Our unification view starts from two unification results:

(i) the similarity of integers and polynomials (as covered in [1,13])

(ii) the relation between the Newton and Hensel iteration methods (as covered in [22])

and it will be concerned with the problems of the so called error-free computation and of the rational interpolation of functions.

2. Error-free computation

The problem of error-free computation has been investigated by many authors in the recent past.

Two classical solutions of this problem are based on the infinite precision integer and rational arithmetic and on the multiple moduli residue arithmetic[13]. However both these kinds of arithmetic suffer from an unpredictable need of space.

A significant alternative, that lies between these approaches and the floating point arithmetic, is the approximate rational arithmetic that assures low complexity operations together with powerful capabilities. Horn[12] and Matula and Kornerup[17] in fact proposed a method to approximate a rational number, that is based on a classical algorithm for finding successive convergents of a continued fraction expansion of the given number (see [9] sect. 10.7). According to this method since an approximate value of the rational p/q is accurate to within $1/q^2$, it gives similar accurancy as double precision arithmetic.

Another very interesting approximate rational arithmetic is based on the use of the so called p-adic arithmetic, first introduced by Hensel[11] in 1908 and then followed by Hehner and Horspool[10] in order to define an efficient and easy arithmetic.

In the same direction Krishnamurthy, Rao and Subramanian[16] and Gregory[4,5,6] proposed a particular p-adic representation with fixed length numbers that they called Hensel-codes.

CH1892-9/83/0000/0210$01.00 © 1983 IEEE

Let α be a rational number, and let p be an integer, that usually is assumed to be prime, then the sequence of digits

$$a_n a_{n+1} a_{n+2} \cdots \tag{1}$$

with

$$\alpha = \sum_{i=n}^{\infty} a_i p^i \tag{2}$$

and $0 \leq a_i < p$, for $i = n, n+1, \ldots, a_n \neq 0$ and n positive, negative or zero, is the p-adic representation of α.

Since any such α can be uniquely expressed as

$$\alpha = c/d \; p^n \tag{3}$$

with p and n as in (2), and c, d and p pairwise relatively prime, it follows from (2) that:

$$\frac{c}{d} = \sum_{i=n}^{\infty} a_i p^{i-n} \tag{4}$$

This p-adic representation leads to a powerful arithmetic that has very useful and significant properties for error-free computation[10,14].

Let us now suppose we would like to deal only with fixed length p-adic numbers, namely a p-adic sequence as (1) with only, say, r p-adic digits.

Then a rational number α would be expressed by the value of n, that represents its exponent e_α, and by the sequence

$$a_n a_{n+1} \cdots a_{n+r-1} \tag{5}$$

that is a sort of mantissa: m_α. For given p we can then represent α by the so called Hensel-code

$$H(p,r,\alpha) = (m_\alpha , e_\alpha) \tag{6}$$

In order now to operate with this kind of representation we need two algorithms for the direct and inverse mappings of rational numbers into their correspondent Hensel-codes.

Kornerup and Gregory[15] and Miola[19] independently devised algorithms for those two mappings, both based on intensive use of the extended Euclid's algorithm.

However similar results were also reached by Wang[20] and by Wang, Guy and Davenport[21] for the reconstruction of rational numbers from p-adic images.

A first significant observation can be made at this point by recognizing that all the approximate rational arithmetics we have mentioned here, namely the proposals and the results of [4,5,6,12,15,16,17,19] lie on the basic algebraic and number theoretic properties of the extended Euclid's algorithm and its relation with diophantine analysis and continued fraction expansion of rational numbers. Those properties are very well known (see for instance [1,9,13], however we would like to give a short overview of that, so we will also be able to immediately refer to that in the next sections.

3. The extended Euclid's algorithm and its properties

The extended Euclid's algorithm (EEA) [1,13] for given a_o , a_1 non negative integers, computes x and y, as well as the greatest common divisor (gcd) of a_o and a_1 , such that:

$$a_o x + a_1 y = d = \gcd(a_o , a_1) \tag{7}$$

The computation is carried out by using three vectors $\langle a_o , a_1, \ldots, a_n \rangle$, $\langle x_o, x_1, \ldots, x_n \rangle$ and $\langle y_o, y_1, \ldots, y_n \rangle$ such that at each step the following relations hold:

$$\begin{aligned} a_{i+1} &= a_{i-1} + q_i a_i \quad \text{for given } a_o , a_1 \\ X_{i+1} &= x_{i-1} + q_i x_i \quad \text{for } x_o = 1, x_1 = 0 \\ y_{i+1} &= y_{i-1} + q_i y_i \quad \text{for } y_o = 0, y_1 = 1 \end{aligned} \tag{8}$$

where $q_i = \lfloor a_{i-1}/a_i \rfloor$, $a_n = 0$ with $a_{n-1} = \gcd(a_o, a_1)$, and also such that the property

$$a_o x_i + a_1 y_i = a_i \tag{9}$$

holds for all $i = 0,1,\ldots,n-1$.
The values x_{n-1} , y_{n-1} are then the solution of (7).

This equation (7) has the general form

$$a_o x + a_1 y = c \tag{10}$$

that is known as a diophantine equation in the integer unknown x and y, for given a_o, a_1, c integers.

If $\gcd(a_o,a_1)$ divides c then the solutions $\bar{x},\bar{y}$ of (10) are related to the solution x_{n-1}, y_{n-1} of the equation (7) (see for instance [13]), in the following way:

$$\bar{x} = mx_{n-1} + ka_1/\gcd(a_o,a_1)$$
$$\bar{y} = my_{n-1} + ka_o/\gcd(a_o,a_1) \qquad (11)$$

for any integer k and m such that $c = m \cdot \gcd(a_o,a_1)$.

On the other hand the EEA and the solution (11) of the equation (10) have also interesting relation with the continued fraction expansion of numbers.

In fact the problem of solving the equations (7), (10) can be also formulated as the calculation of an approximation to a_o/a_1 since in (7)

$$\frac{a_o}{a_1} + \frac{y}{x} = \frac{d}{xa_1} < \frac{1}{x^2} \qquad (12)$$

if we assume $0 \leq x < a_1$ [9].

Therefore $-y/x$ is an approximation to a_o/a_1 that is quite close to a_o/a_1 and has a very small denominator.

From this observation wa can also verify that the successive values of q in the EEA are the successive terms of the continued fraction expansion of a_o/a_1 :

$$a_o/a_1 = q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \ddots}} \qquad (13)$$

The successive convergents of a_o/a_1 are then the values $-y_i/x_i$ for $i \geq 0$, with x_i,y_i computed throughout the EEA as in (8) [9,13].

4. The use of EEA for error-free computation

The proposals of [12] and of [17] directly refer to the representation of a number by a convergent of its continued fraction expansions. While the result of [19], as far as the mapping of a rational number $\alpha = a/b$ onto its Hensel-code $H(p,r,a/b) = (m_\alpha , e_\alpha)$ is concerned, is based on the simple computation of $e_\alpha = n$ such that

$$\frac{a}{b} = \frac{c}{d} \cdot p^n \qquad (14)$$

for a given p, and on the determination of the unique solution of the following diophantine equation

$$m_\alpha d + c = kp^r \qquad (15)$$

for the unknown m_α , given c,d,p^r pairwise relatively prime.

The inverse mapping, namely the reconstruction of a rational number c/d, and then of a/b as in (14), from its image mod p^r, is then defined in [19] and it is based again on the use of EEA to generate continued fraction convergents of m_α/p^r.

However a basic observation must be made here.

OBSERVATION 1. The solution of the inverse mapping problem is unique only when a restriction on the rational number c/d we are looking for is posed; namely if this ratio is limited as:

$$0 \leq |c| < N \text{ and } 0 < d \leq N \qquad (16)$$

A discussion of this observation, and the proof of a related theorem can be found in [15,19].

5. Rational approximation of functions

Let $x_o,x_1,x_2,\ldots$ be a sequence of points some of which may be repeated. The problem of rational Hermite interpolation of degree-type (μ,ν), with $\mu+\nu = N$, is to determine a rational function $A(x)/B(x)$ with degree $(A) \leq \mu$ and degree $(B) \leq \nu$, which interpolates an analytic function $\alpha(x)$ at the first N+1 points of the sequence: $x_o,x_1,\ldots,x_{\mu+\nu}$.

For $\mu = N$ and $\nu = 0$ we have the polynomial Hermite interpolation problem solved by a polynomial P(x) of degree N such that

$$\alpha(x) - P(x) = \beta(x) \prod_{i=0}^{N} (x-x_i) \qquad (17)$$

where $\beta(x)$ is a analytic function.

In general we have that the rational function $A(x)/B(x)$ is such that

$$\alpha(x) - \frac{A(x)}{B(x)} = \beta(x) \prod_{i=0}^{N}(x-x_i) \qquad (18)$$

When the points x_i are distinct the rational interpolation problem is known as Cauchy interpolation, and in the other side when all the points x_i are the same this interpolation problem is called Padè interpolation. Then every problem and every result concerning Cauchy interpolation has its Padè equivalent, so we will refer to these two situations interchangeably.

An algebraic formulation of (18) can be immediately found. Let in fact

$$\alpha(x) = a_o + a_1 x + a_2 x + \ldots \qquad (19)$$

be a power series expansion (in x, over a field) of the function α. A Padè approximant (or interpolant) to $\alpha(x)$ is a rational function $A(x)/B(x)$ such that:

$$\begin{aligned} &a) \quad \deg(A) \leq \mu \\ &b) \quad \deg(B) \leq \nu \qquad (20) \\ &c) \quad B(x)\alpha(x) \equiv A(x) \bmod x^{N+1} \end{aligned}$$

for $N = \mu+\nu$.

If we consider the truncation of $\alpha(x)$ at the N-power term we have

$$\alpha_N(x) = a_o + a_1 x + \ldots + a_N x^N$$

and

$$B(x)\alpha_N(x) \equiv A(x) \bmod x^{N+1} \qquad (21)$$

The formulations (20) and (21) show how, for a given function $\alpha(x)$, we will have a table of all the interpolants of $\alpha(x)$ for different values of μ and ν. We would like now to discuss how this table can be constructed using the EEA, so we do not go into more details of rational interpolation.

However for a good survey on rational interpolation see [3].

6. The use of EEA for rational interpolation

The problem of rational interpolation as formulated in (20) and (21) has a clear relation with the problem of reconstructing rational numbers from their images modulo a prime, as formulated in section 4 and solved in [19]. The formulation (21) in fact immediately suggests the use of the EEA to determine the rational approximant of a given function $\alpha(x)$. Brent, Gustavson and Yun [2,7,8] and McEliece and Shearer [18] independently discovered the relation of the EEA with the problem of Hermite rational interpolation.

According to these results if we apply the EEA to $A_o(x) = x^{N+1}$ and $A_1(x) = \alpha_N(x)$ (or to $A_o(x) = \prod_{i=0}^{N}(x-x_i)$ and $A_1(x) = \alpha_N(x)$ in the case of Cauchy interpolation) we obtain the following properties:

a) each step of the EEA furnishes a unique reference to the rational interpolation table.
b) the rational function $A_i(x)/Y_i(x)$ obtained by the EEA furnishes as many as $\deg(Q_i)$ equal entries to the interpolation table, along the $(\mu+\nu)$-th antidiagonal.
c) all entries along the $(\mu+\nu)$-th antidiagonal of the interpolation table are computed uniquely by the EEA.

In formulating these properties we have made use of a symbolism for the EEA elements which is a natural extension of that used in section 3. In fact the EEA applies also for polynomials over a field and here we have indicated polynomials by capitalizing the same letters we used in section 3.

A basic observation can then be made here on these properties.

OBSERVATION 2. The solution of the rational interpolation problem is uniquely determinated by the EEA, because a natural restriction on the desired solution has been assumed by the limitation on the degree of the polynomials A(x) and B(x) as in (20).

7. Conclusion

This paper has presented some well known results on the use of the extended Euclid's algorithm for approximation processes in a hopefully meaningful unified view, underlining the generalities and the properties of this algorithm.

The unification remarks coming out from this presentation can be the following:

(i) The integers and the polynomials have a well known similarity which is very useful in algebraic algorithm designing [1,13].

(ii) The iterative methods as Newton and Hensel constructions have been shown to be very related [22].

(iii) Approximate rational arithmetic and rational interpolation of functions are also related, and both together have strong dependency from the use of the extended Euclid's algorithm [15,19,18,2,7,8].

(iv) The solutions of recovering a rational expression for numbers and for functions are unique when the desired answers are assumed to be limited in their sizes, (see observations above).

These remarks once again make evident the unification strategy in algebraic computation. This strategy is more and more powerful in these days and it is the significant base for developing new systems and it can help for incorporating also the p-adic arithmetic for error free computation and in general more numeric-like methods.

References

1 Aho, A.V., Hopcroft, J.E., and Ullman, J.D., The Design and Analysis of Computer Algorithms, Addison Wesley, Reading, Mass. (1974).

2 Brent, R.P., Gustavson, F.G., Yun, D.Y.Y., Fast solution of Toeplitz systems of equations and computation of Padè approximants, J. of Algorithms, 1, (1980).

3 Gragg, W.B., The Padè table and its relation to certain algorithms of numerical analysis, SIAM Rev. 14, n. 1, (1972).

4 Gregory, R.T., The use of finite-segment p-adic arithmetic for exact computations, BIT, 18, 282-300 (1978).

5 Gregory, R.T., Error-Free Computation, Robert E. Krieger Pub. Co. Huntington, N.Y. (1980).

6 Gregory, R.T., Error-free computation with rational numbers, BIT, 21, 194-202 (1981).

7 Gustavson, F.G., Yun, D.Y.Y., Fast algorithms for rational Hermite approximation and solution of Toeplitz systems, IEEE Trans.Circuits and Systems, CAS 26, n. 9 (1979).

8 Gustavson, F.G., Yun, D.Y.Y., Fast Computation of Padè Approximants and Toeplitz Systems of Equations via the Extended Euclidean Algorithm, IBM Research Report RC7551 (1979).

9 Hardy, G.H., Wright, E.M., An Introduction to the theory of numbers, Fourth Ed., Clarendon Press, Oxford (1960).

10 Hehner, E.C.R., Horspool, R.N.S., A new representation of the rational numbers for fast easy arithemtic, SIAM J. Comput., 8, 124-134 (1979).

11 Hensel, K., Theorie der Algebraischen Zahlen, Teubner, Leipzing-Stuttgart (1908).

12 Horn, B.K.P., Rational Arithmetic for Minicomputers, Software-Prac. and Exper. 8, 171-176 (1978).

13 Knuth, D.E., The Art of Computer Programming: Volume II Seminumerical Algorithms, Addison Wesley, Reading, Mass. (1980).

14 Koblitz, N., P-Adic Numbers, P-Adic Analysis, and Zeta Functions, Springer-Verlag, New York (1979).

15 Kornerup, P., Gregory, R.T., Mapping integers and Hensel-codes onto Farey fractions, DAIMI PB 149, Comp. Sc. Dept. of Aarhus University, Denmark (1982).

16 Krishnamurthy, E.V., Rao, T.M., Subramanian,K., Finite segment p-adic number systems with applications to exact computation, Proc. Indian Acad. Sci., 81A, 58-79 (1975).

17 Matula, D.W., Kornerup, P., Approximate rational arithmetic systems: analysis of recovery of simple fractions during expression evaluation, Lecture Notes in Computer Science, n. 72 (1979).

18 McEliece, R.J., Skearer, J.B., A property of Euclid's algorithm and an application to Padè approximation, SIAM J. Appl. Math. 34, n. 4 (1978).

19 Miola, A., The conversion of Hensel-codes to their rational equivalents, SIGSAM Bulletin n. 16 November (1982).

20 Wang, P.S., A p-adic Algorithm for Univariate Partial Fractions. Proceedings of the 1981 ACM Symposium on Symbolic and Algebraic Computation, ACM Inc., New York, (1981).

21 Wang, P., Guy, M., Davenport, J., P-adic reconstruction of rational numbers, SIGSAM Bulletin 62, vol. 16, n. 2, (1982).

22 Yun, D.Y.Y., Algebraic algorithms using p-adic constructions, Proc. of ACM SYMSAC (1976).