On Economization of Rational Functions

ANTHONY RALSTON

Stevens Institute of Technology, Hoboken, New Jersey

1. Maehly [4] has extended the method of Lanczos [2, 3] for the economization of power series to continued fractions and, therefore, to a certain class of rational functions. It is our purpose here to derive an economization technique which is similar to Maehly's but which applies to all Padé approximations [6] and whose derivation and (in some cases) application is simpler than those described in [4].

2. We consider rational approximations to an analytic function $f(x)$ on an interval $[-\epsilon, \epsilon]$ (where generally $\epsilon < 1$), so that by a simple change of variable the results can be extended to any finite interval. The approximations we consider have the form

$$R_{mk}(x) = \frac{P_m(x)}{Q_k(x)}, \tag{2.1}$$

where

$$P_m(x) = \sum_{j=0}^{m} a_j x^j \tag{2.2}$$

and

$$Q_k(x) = \sum_{j=0}^{k} b_j x^j \quad (b_0 = 1). \tag{2.3}$$

We will say the rational function (2.1) has index N, where $m+k = N$. Then a Padé approximation of index N to the function $f(x)$ has the property that [1]

$$\lim_{x \to 0} \frac{f(x) - R_{mk}(x)}{x^{N+1}} = d_{N+1}^{(m,k)}, \tag{2.4}$$

where

$$d_{N+1}^{(m,k)} = \sum_{j=0}^{k} c_{N+1-j}\, b_j, \tag{2.5}$$

with the c_j's the coefficients of the Maclaurin series for $f(x)$,

$$f(x) = \sum_{j=0}^{\infty} c_j x^j. \tag{2.6}$$

We are interested in using $R_{mk}(x)$ to derive another rational approximation whose maximum error on $[-\epsilon, \epsilon]$ will be smaller than that of $R_{mk}(x)$ and which will, therefore, approach more closely the Chebyshev approximation on the interval.

Following Maehly, we make the change of variable $x = \epsilon u$ in order to convert the interval of interest to $[-1, 1]$, and rewrite (2.4) as

$$\lim_{\epsilon\to 0} \frac{f(\epsilon u) - R_{mk}(\epsilon u)}{\epsilon^{N+1}} = d_{N+1}^{(m,k)} u^{N+1}. \tag{2.7}$$

We derive a method for obtaining a new rational approximation $C_{mk}(x)$ with the property that

$$\lim_{\epsilon\to 0} \frac{f(\epsilon u) - C_{mk}(\epsilon u)}{\epsilon^{N+1}} = d_{N+1}^{(m,k)} (T_{N+1}(u))/2^N, \tag{2.8}$$

where $T_{N+1}(u)$ is the Chebyshev polynomial (of the first kind) of degree $N+1$. For if $C_{mk}(x)$ is such that (2.8) is satisfied, then for sufficiently small ϵ the maximum error on the interval $[-\epsilon, \epsilon]$ of the approximation $C_{mk}(x)$ will be less than that of $R_{mk}(x)$. This follows from (2.7), (2.8) and the well-known property of the Chebyshev polynomials, that of all polynomials of degree $N+1$ with coefficients of u^{N+1} equal to one, $(T_{N+1}(u))/2^N$ has the smallest maximum value on $[-1, 1]$.

3. Let

$$R_{i,j-i}^{(j)}(x) = \frac{P_i^{(j)}(x)}{Q_{j-i}^{(j)}(x)} \qquad j = 0, \cdots, N-1 \tag{3.1}$$

be a sequence of Padé approximations to $f(x)$, where i is selected arbitrarily so that $0 \leqq i \leqq m$ and $0 \leqq j-i \leqq k$. If $R_{i,j-i}^{(j)}(x) = 0$, set $P_i^{(j)}(x) = 0$, $Q_{j-i}^{(j)}(x) = 1$. When $j = 0$, we must have $i = 0$ independent of m and k. When $m = 0$ we must also have $i = 0$, and when $k = 0$ we must have $i = j$, but in all other cases there is more than one choice for i. Analogous to (2.7),

$$\lim_{\epsilon\to 0} \frac{f(\epsilon u) - R_{i,j-i}^{(j)}(\epsilon u)}{\epsilon^{j+1}} = d_{j+1}^{(i,j-i)} u^{j+1}. \tag{3.2}$$

Now define (cf. equation (2.13a) of [4])

$$C_{mk}(x) = \frac{P_m(x) + \sum_{j=0}^{N-1} \beta_{j+1} P_i^{(j)}(x) + \beta_0}{Q_k(x) + \sum_{j=0}^{N-1} \beta_{j+1} Q_{j-i}^{(j)}(x)}, \tag{3.3}$$

where

$$\begin{aligned} \beta_{j+1} &= \frac{d_{N+1}^{(m,k)}}{d_{j+1}^{(i,j-i)}} \frac{\epsilon^{N-j}}{2^N} t_{j+1} \qquad j = 0, \cdots, N-1, \\ \beta_0 &= -d_{N+1}^{(m,k)} \epsilon^{N+1} t_0/2^N, \end{aligned} \tag{3.4}$$

with t_j the coefficient of u^j in $T_{N+1}(u)$. We now show that this $C_{mk}(x)$ satisfies (2.8). We have

$$f(x) - C_{mk}(x)$$
$$= \frac{\left[Q_k(x) + \sum_{j=0}^{N-1} \beta_{j+1} Q_{j-i}^{(j)}(x)\right] f(x) - P_m(x) - \sum_{j=0}^{N-1} \beta_{j+1} P_i^{(j)}(x) - \beta_0}{Q_k(x) + \sum_{j=0}^{N-1} \beta_{j+1} Q_{j-i}^{(j)}(x)}$$
$$= \frac{Q_k(x)f(x) - P_m(x) + \sum_{j=0}^{N-1} \beta_{j+1}[Q_{j-i}^{(j)}(x)f(x) - P_i^{(j)}(x)] - \beta_0}{Q_k(x) + \sum_{j=0}^{N-1} \beta_{j+1} Q_{j-i}^{(j)}(x)}. \tag{3.5}$$

Because of (2.3), $Q_j(0) = 1$ for any j. Thus we may write (3.2) as

$$\lim_{\epsilon\to 0} \frac{Q_{j-i}^{(j)}(\epsilon u)f(\epsilon u) - P_i^{(j)}(\epsilon u)}{\epsilon^{j+1}} = d_{j+1}^{(i,j-i)} u^{j+1}. \tag{3.6}$$

Using this and (3.4),

$$\lim_{\epsilon\to 0} \frac{f(\epsilon u) - C_{mk}(\epsilon u)}{\epsilon^{N+1}} = d_{N+1}^{(m,k)} \left[u^{N+1} + \sum_{j=0}^{N-1} \frac{t_{j+1}}{2^N} u^{j+1} + \frac{t_0}{2^N}\right]$$
$$= d_{N+1}^{(m,k)} T_{N+1}(u)/2^N, \tag{3.7}$$

as we desired.

4. We make the following remarks on this technique:

(1) Since every Chebyshev polynomial contains only even or only odd powers, alternate β_j's are equal to zero.

(2) The cases $d_{j+1}^{(i,j-i)} = 0$ can generally be taken care of by using a different member of the sequence (3.1) or (see below) by noting that the highest power in the numerator or denominator really has a coefficient of zero so that the index may be increased by 1.

(3) When a convergent continued fraction expansion for $f(x)$ is known, a sequence of Padé approximations may be derived directly from the continued fraction (see below) or the method in [4] may be used.

(4) Given (2.6) the technique presented here could be quite practically performed on a digital computer in a general fashion (cf. [5]).

5. An example. As in [4], we consider approximations to tan x, starting from the continued fraction expansion

$$\tan x = \frac{x|}{|1} - \frac{x^2|}{|3} - \frac{x^2|}{|5} - \frac{x^2|}{|7} - \frac{x^2|}{|9} - \cdots \tag{5.1}$$

on the interval $[-.6, .6]$. Converting successive approximants of (5.1) into rational functions, we get a sequence (3.1) of the form

$$\begin{aligned} R_{0,0}^{(0)}(x) &= 0/1 \\ R_{1,0}^{(1)}(x) &= R_{1,1}^{(2)}(x) = x \\ R_{1,2}^{(3)}(x) &= R_{2,2}^{(4)}(x) = \frac{x}{1 - \frac{1}{3}x^2} \\ R_{3,2}^{(5)}(x) &= R_{3,3}^{(6)}(x) = \frac{x - \frac{1}{15}x^3}{1 - \frac{2}{5}x^2}. \end{aligned} \tag{5.2}$$

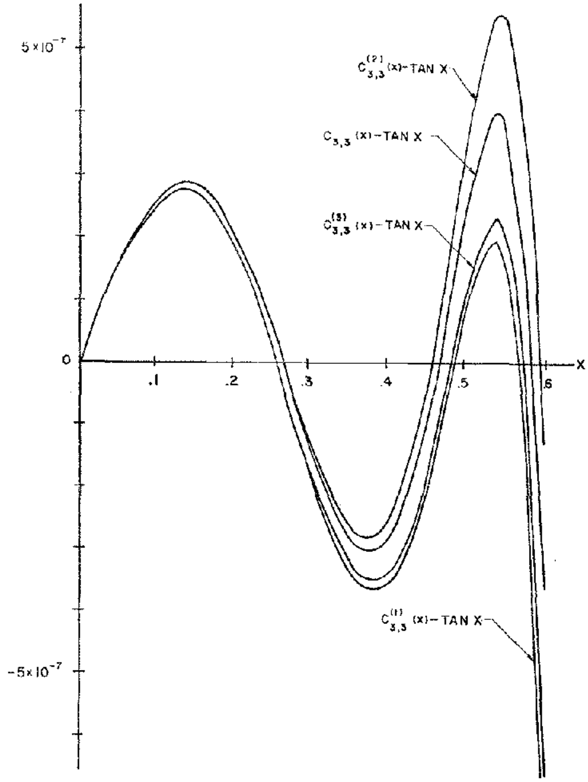

FIG. 1

The equality between successive members of the sequence follows because the numerators contain only odd, and the denominators only even powers of x. The Maclaurin series for $\tan x$ is

$$\tan x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \cdots. \tag{5.3}$$

Then using (5.2), (5.3) and the equations of section 3 with $m = 3$, $k = 3$ and $N = 6$, compute

$$C_{3,3}(x) = x\,\frac{15.0000495 - 1.0181094x^2}{15.0000000 - 6.0170263x^2}. \tag{5.4}$$

In [4] Maehly presented three approximations to $\tan x$ of the type (5.4), each derived using a slightly different technique. These are

$$\begin{aligned} C^{(1)}_{3,3} &= x\,\frac{15.0000486 - 1.0180033x^2}{15.0000000 - 6.0169200x^2} \\ C^{(2)}_{3,3} &= x\,\frac{15.0000486 - 1.0181133x^2}{15.0000000 - 6.0170465x^2} \\ C^{(3)}_{3,3} &= x\,\frac{15.0000486 - 1.0180000x^2}{15.0000000 - 6.0169200x^2}. \end{aligned} \tag{5.5}$$

In Figure 1 the errors in these four approximations are plotted[1] for $0 \leqq x \leqq .6$. The magnitudes of the maximum errors for $C_{3,3}(x)$, $C_{3,3}^{(1)}(x)$, $C_{3,3}^{(2)}(x)$ and $C_{3,3}^{(3)}(x)$ are approximately 4.0×10^{-7}, 7.0×10^{-7}, 5.6×10^{-7} and 6.4×10^{-7}, respectively. The fact that $C_{3,3}(x)$ gives a smaller maximum error than the others (and is, in fact, nearly a Chebyshev (i.e. minimum maximum error) approximation) is probably not significant, for there is no obvious reason why the method of this paper should be any more efficient than that of [4] in comparable situations.

REFERENCES

1. KOGBETLIANTZ, E. G. Generation of elementary functions. In A. Ralston and H. S. Wilf, Eds., *Mathematical Methods for Digital Computers*, pp. 13–14 (John Wiley and Sons, 1960).
2. LANCZOS, C. Trigonometric interpolation of empirical and analytical functions. *J. Math. Phys. 17* (1938), 123–199.
3. ——. *Applied Analysis*, pp. 457–463. Prentice Hall, 1956.
4. MAEHLY, H. J. Methods for fitting rational approximations. *J. ACM 7* (1960), 150–162.
5. SPIELBERG, K. Representation of power series in terms of polynomials, rational approximations and continued fractions. *J. ACM 5* (1961), 613–627.
6. WALL, H. S. *Continued Fractions*, Ch. 20. Van Nostrand, 1948.

---

[1] The calculations for Figure 1 were performed on the IBM 1620 computer at the Stevens Institute of Technology, which is partly supported by the National Science Foundation.