ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 21, Number 1, Winter 1991

A CLASS OF ADAPTIVE MULTIVARIATE NONLINEAR ITERATIVE METHODS

ANNIE CUYT

ABSTRACT. Multivariate iterative procedures, based on the use of multivariate rational Hermite interpolants, are introduced for the solution of systems of nonlinear equations. To this end, the most recent iteration points are used as interpolation points. Coalescent interpolation points are obtained by repeating the same iteration point a number of times.

This new type of iterative procedure can be completely adapted to the cost of evaluating and/or differentiating the nonlinear equations in the system. Section 1 deals with the univariate case and repeats both one-point and multipoint iterative procedures resulting from the use of approximating rational functions. Section 2 generalizes the multipoint iterations to the multivariate case while Section 3 generalizes the one-point iterations.

1. Nonlinear methods for the solution of systems of nonlinear equations. Suppose we want to find a root x^* of the nonlinear equation

$$f(x) = 0,$$

where the univariate function f may be real- or complex-valued. If f is replaced by a local approximation, then a zero of that local approximation can be considered as an approximation for x^*. Methods based on this reasoning are called direct methods. One could also consider the inverse function g of f in a neighborhood of 0, if it exists, and replace g by a local approximation. Then an evaluation of this local approximation at 0 can be considered as an approximation for x^* since

$$g(0) = x^*.$$

Methods using this technique are called inverse methods. We now look at some univariate nonlinear direct and inverse methods that will

Research Associate NFWO
Received by the editors on September 25, 1988, and, in revised form, on June 8, 1989.
AMS *Mathematics Subject Classification.* 65H10, 65D05.

inspire us. Let $x^{(i)}$ be an approximation for the root x^* of f, and let

$$r_i(x) = \frac{p_i}{q_i}(x) \tag{1a}$$

be the Padé approximant of order (n, m) for f in $x^{(i)}$. Then the next approximation $x^{(i+1)}$ is calculated such that

$$p_i(x^{(i+1)}) = 0. \tag{1b}$$

In case $p_i(x)$ is linear ($n = 1$), the value $x^{(i+1)}$ is uniquely determined. It is clear that this is to be preferred for the sake of simplicity. A well-known method obtained in this way is Newton's method ($n = 1, m = 0$). Another famous method is Halley's method, based on the use of the Padé approximant of order (1,1) for f at $x^{(i)}$, which is given by

$$x^{(i+1)} = x^{(i)} - \frac{f(x^{(i)})/f'(x^{(i)})}{1 - \frac{1}{2}f''(x^{(i)})\frac{f(x^{(i)})}{f'(x^{(i)})^2}}. \tag{2}$$

Since these iterative procedures only use information in the point $x^{(i)}$ to calculate the next iteration point $x^{(i+1)}$, they are called one-point methods. Examples exist where methods based on the use of (n, m) Padé approximants with $m > 0$ give better results than linear methods if the function f has singularities. The nonlinear methods may even converge while the linear methods diverge. The reasoning in (1) can be generalized by using rational Hermite interpolants instead of Padé approximants as local approximations for $f(x)$. Let

$$r_i(x) = \frac{p_i}{q_i}(x),$$

with p_i and q_i of degrees n and m, be such that, in approximations $x^{(i)}, \ldots, x^{(i-s)}$ for the root x^* of f,

$$\begin{aligned}
&r_i^{(l)}(x^{(i)}) = f^{(l)}(x^{(i)}), \quad l = 0, \ldots, l_0 - 1 \\
&r_i^{(l)}(x^{(i-1)}) = f^{(l)}(x^{(i-1)}), \quad l = 0, \ldots, l_1 - 1, \\
&\vdots \\
&r_i^{(l)}(x^{(i-s)}) = f^{(l)}(x^{(i-s)}), \quad l = 0, \ldots, l_s - 1,
\end{aligned} \tag{3a}$$

with $n+m+1 = l_0 + \cdots + l_s$. Then the next iteration point $x^{(i+1)}$ is computed such that

$$p_i(x^{(i+1)}) = 0. \tag{3b}$$

For the calculation of $x^{(i+1)}$, we now use information in more than one previous point. Hence, such methods are called multipoint. If we take $n=1$, $m=1$, $s=2$ and $l_0 = l_1 = l_2 = 1$, then $x^{(i+1)}$ is given by

$$x^{(i+1)} = x^{(i)} - \frac{f(x^{(i)})[f(x^{(i-1)}) - f(x^{(i-2)})]}{f(x^{(i-1)})\frac{f(x^{(i-2)})-f(x^{(i)})}{x^{(i-2)}-x^{(i)}} - f(x^{(i-2)})\frac{f(x^{(i-1)})-f(x^{(i)})}{x^{(i-1)}-x^{(i)}}}. \tag{4}$$

Take $n=1$, $m=1$, $s=1$ and $l_0 = 2$, $l_1 = 1$. Then $x^{(i+1)}$ is given by

$$x^{(i+1)} = x^{(i)} + \frac{f(x^{(i)})(x^{(i)} - x^{(i-1)})}{f(x^{(i-1)})f'(x^{(i)})\frac{x^{(i)}-x^{(i-1)}}{f(x^{(i)})-f(x^{(i-1)})} - f(x^{(i)})}. \tag{5}$$

The case $n=1$, $m=0$, $s=1$ and $l_0 = l_1 = 1$ reduces to the secant method. The case $n=1$, $m=1$, $s=0$, $l_0 = 3$ reduces to Halley's method because the rational Hermite interpolant is then a Padé approximant. The rational function satisfying (3) can also be obtained by reformulating the rational Hermite interpolation problem as a Newton-Padé approximation problem. Let us introduce the interpolation points z_l by repeating the approximations $x^{(i)}, \ldots, x^{(i-s)}$ for the root x^* of f as many times as they are used:

$$\begin{aligned} z_0 &= \cdots = z_{l_0-1} = x^{(i)} \\ z_{l_0} &= \cdots = z_{l_0+l_1-1} = x^{(i-1)} \\ &\vdots \\ z_{l_0+\cdots+l_{s-1}} &= \cdots = z_{l_0+\cdots+l_s-1} = x^{(i-s)}. \end{aligned}$$

Then

$$r_i(x) = \frac{p_i}{q_i}(x),$$

with p_i and q_i of degrees n and m, is computed such that the first $n+m+1$ divided differences,

$$\begin{aligned} &(fq-p)[z_0, \ldots, z_l] \\ =&(fq-p)[\underbrace{x^{(i)}, \ldots, x^{(i)}}_{l_0 \text{ times}}, \underbrace{x^{(i-1)}, \ldots, x^{(i-1)}}_{l_1 \text{ times}}, \ldots] = 0, \quad l = 0, \ldots, n+m, \end{aligned} \tag{6a}$$

in the formal Newton series representation of $(fq-p)$ disappear, in other words,

$$(fq-p)(x) = \sum_{l=n+m+1}^{\infty} (fq-p)[z_0,\ldots,z_l]B_l(x)$$

where

$$B_l(x) = (x-z_0)\ldots(x-z_{l-1}), \quad B_0(x) = 1.$$

It is explained in [**2**] and [**6**] how one computes these divided differences if some of the interpolation points coincide. The order of convergence of methods based on the use of (2) and (3) can be calculated as follows [**10**].

Theorem . *If the sequence* $(x^{(i)})_{i\in\mathbf{N}}$ *converges to a simple root* x^* *of* f *and* $f^{(n+m+1)}(x)$ *with* $m>0$ *continuous in a neighborhood of* x^* *with*

$$\begin{vmatrix} f^{(n)}(x^*) & f^{(n-1)}(x^*) & \ldots & f^{(n-m+1)}(x^*) \\ f^{(n+1)}(x^*) & & & \vdots \\ \vdots & & & f^{(n-1)}(x^*) \\ f^{(n+m-1)}(x^*) & \ldots & f^{(n+1)}(x^*) & f^{(n)}(x^*) \end{vmatrix} \neq 0,$$

where $f^{(l)}(x^*)=0$ *if* $l<0$*, then the order of the iterative method based on the use of* (3) *is the unique positive root of the polynomials*

$$x^{s+1} - l_0x^s - l_1x^{s-1} - \cdots - l_s = 0.$$

By means of the multivariate rational Hermite interpolants introduced in [**6**], the formulas (2), (4) and (5) will now be generalized for the solution of systems of nonlinear equations

$$\begin{cases} f_1(x_1,\ldots,x_k) = 0 \\ \quad\vdots \\ f_k(x_1,\ldots,x_k) = 0. \end{cases}$$

We can now refer the reader to [**4**] for a third order multivariate generalization of Halley's iteration. This generalization is not yet adaptive to the informational usage of the functions f_j, $j = 1, \ldots, k$, but it can be altered to be so by following the reasoning presented in the next sections. Let us first discuss "why" it is necessary that such adaptive techniques exist. First of all, available information about the system of nonlinear equations is not always the same. One may know some partial derivatives or function values, at one or more points. Secondly, some function evaluations may be much more difficult or time consuming than others. Some nonlinear functions depend on some of the variables in a more complex way than on the other variables. We shall introduce here a very flexible multivariate generalization that allows the construction of tailor-made iterative procedures.

2. Methods based on the use of multivariate general order rational Hermite interpolants. Let $(x_1^{(0)}, \ldots, x_k^{(0)}), (x_1^{(1)}, \ldots, x_k^{(1)}), (x_1^{(2)}, \ldots, x_k^{(2)}), \ldots$ be given in the k-dimensional space $\mathbf{C}^k$, and let the k-variate function $f(x_1, \ldots, x_k)$ be known at the points $(x_1^{(i_1)}, \ldots, x_k^{(i_k)})$ with $(i_1, \ldots, i_k) \in I \subset \mathbf{N}^k$, where I satisfies the inclusion property, meaning that when $(i_1, \ldots, i_k) \in I$, then $(l_1, \ldots, l_k) \in I$ for $l_j \leq i_j$ with $j = 1, \ldots, k$. We denote $f_{i_1 \ldots i_k} = f(x_1^{(i_1)}, \ldots, x_k^{(i_k)})$. We know from [**6**] how to deal with coalescent interpolation points or coalescent coordinates of interpolation points. In the next section we shall especially focus on coalescent interpolation points. In this section we shall develop the formulas that remain valid both for coalescent and noncoalescent situations. Consider the following set of basis functions for the real-valued polynomials in several variables:

$$B_{l_1 \ldots l_k}(x_1, \ldots, x_k) = \left[(x_1 - x_1^{(0)}) \ldots (x_1 - x_1^{(l_1-1)})\right] \ldots \left[(x_k - x_k^{(0)}) \ldots (x_k - x_k^{(l_k-1)})\right],$$

with $B_{0 \ldots 0}(x_1, \ldots, x_k) = 1$. With our data and these basis functions, we can now formally write a Newton interpolating series for f,

$$f(x_1, \ldots, x_k) = \sum_{(l_1, \ldots, l_k) \in \mathbf{N}^k} c_{0 l_1, \ldots, 0 l_k} B_{l_1 \ldots l_k}(x_1, \ldots, x_k),$$

where $c_{s_1 l_1, \ldots, s_k l_k} = f[x_1^{(s_1)}, \ldots, x_1^{(l_1)}] \ldots [x_k^{(s_k)}, \ldots, x_k^{(l_k)}]$ is a multivariate divided difference with possible coalescence of points [**6**]. The value

$c_{s_1 l_1,\ldots,s_k l_k} = 0$ if, for some $j = 1,\ldots,k$, we have $s_j > l_j$. Let us choose two subsets N and D of $\mathbf{N}^k$, such that

$$N \subset I$$
$$\#I = \#N + \#D - 1,$$

and construct a rational interpolant to $f(x_1,\ldots,x_k)$ as follows:

$$p(x_1,\ldots,x_k) = \sum_{(l_1,\ldots,l_k)\in N} a_{l_1\ldots l_k} B_{l_1\ldots l_k}(x_1,\ldots,x_k)$$
(N from "numerator," $\#N = n+1$)
$$q(x_1,\ldots,x_k) = \sum_{(l_1,\ldots,l_k)\in D} b_{l_1\ldots l_k} B_{l_1\ldots l_k}(x_1,\ldots,x_k)$$
(D from "denominator," $\#D = m+1$)
$$(fq-p)(x_1,\ldots,x_k) = \sum_{(l_1,\ldots,l_k)\in \mathbf{N}^k\setminus I} d_{0l_1,\ldots,0l_k} B_{l_1\ldots l_k}(x_1,\ldots,x_k)$$
(I from "interpolation conditions").

This last condition means that

(6a) $(fq-p)[x_1^{(0)},\ldots,x_1^{(l_1)}]\ldots[x_k^{(0)},\ldots,x_k^{(l_k)}] = 0, \quad (l_1,\ldots,l_k)\in I,$

where

$$(fq)[x_1^{(0)},\ldots,x_1^{(l_1)}]\cdots[x_k^{(0)},\ldots,x_k^{(l_k)}]$$
$$= \sum_{j_1=0}^{l_1}\cdots\sum_{j_k=0}^{l_k} f[x_1^{(0)},\ldots,x_1^{(j_1)}]\cdots[x_k^{(0)},\ldots,x_k^{(j_k)}] \times q[x_1^{(j_1)},\ldots,x_1^{(l_1)}]\cdots[x_k^{(j_k)},\ldots,x_k^{(l_k)}],$$

$$p[x_1^{(0)},\ldots,x_1^{(l_1)}]\cdots[x_k^{(0)},\ldots,x_k^{(l_k)}] = a_{l_1,\ldots,l_k}, \quad (l_1,\ldots,l_k)\in N,$$
$$p[x_1^{(0)},\ldots,x_1^{(l_1)}]\cdots[x_k^{(0)},\ldots,x_k^{(l_k)}] = 0, \quad (l_1,\ldots,l_k)\in I\setminus N,$$
$$q[x_1^{(0)},\ldots,x_1^{(l_1)}]\cdots[x_k^{(0)},\ldots,x_k^{(l_k)}] = b_{l_1,\ldots,l_k}, \quad (l_1,\ldots,l_k)\in D,$$
$$q[x_1^{(0)},\ldots,x_1^{(l_1)}]\cdots[x_k^{(0)},\ldots,x_k^{(l_k)}] = 0, \quad (l_1,\ldots,l_k)\in I\setminus D.$$

We have proved in [**6**] that, when the rank of the homogeneous system (6b) involving the denominator coefficients $b_{l_1\ldots l_k}$ is maximal—in other

words, if, from $I\backslash N$, no homogeneous equations have to be deleted before applying Cramer's rule to construct a solution—then an explicit determinant formula for this rational interpolant is given by

$$(7)\quad \frac{p(x_1,\ldots,x_k)}{q(x_1,\ldots,x_k)} = \frac{\begin{vmatrix} \sum_{(l_1,\ldots,l_k)\in N} c_{d_1^{(0)}l_1,\ldots,d_k^{(0)}l_k} B_{l_1\ldots l_k} & \cdots & \sum_{(l_1,\ldots,l_k)\in N} c_{d_1^{(m)}l_1,\ldots,d_k^{(m)}l_k} B_{l_1\ldots l_k} \\ c_{d_1^{(0)}h_1^{(1)},\ldots,d_k^{(0)}h_k^{(1)}} & \cdots & c_{d_1^{(m)}h_1^{(1)},\ldots,d_k^{(m)}h_k^{(1)}} \\ \vdots & & \vdots \\ c_{d_1^{(0)}h_1^{(m)},\ldots,d_k^{(0)}h_k^{(m)}} & \cdots & c_{d_1^{(m)}h_1^{(m)},\ldots,d_k^{(m)}h_k^{(m)}} \end{vmatrix}}{\begin{vmatrix} B_{d_1^{(0)}\ldots d_k^{(0)}} & \cdots & B_{d_1^{(m)}\ldots d_k^{(m)}} \\ c_{d_1^{(0)}h_1^{(1)},\ldots,d_k^{(0)}h_k^{(1)}} & \cdots & c_{d_1^{(m)}h_1^{(1)},\ldots,d_k^{(m)}h_k^{(1)}} \\ \vdots & & \vdots \\ c_{d_1^{(0)}h_1^{(m)},\ldots,d_k^{(0)}h_k^{(m)}} & \cdots & c_{d_1^{(m)}h_1^{(m)},\ldots,d_k^{(m)}h_k^{(m)}} \end{vmatrix}}$$

where the points in D are denoted by

$$(d_1^{(0)},\ldots,d_k^{(0)}),\ldots,(d_1^{(m)},\ldots,d_k^{(m)})$$

and those in $I\backslash N$ by

$$(h_1^{(1)},\ldots,h_k^{(1)}),\ldots,(h_1^{(m)},\ldots,h_k^{(m)}).$$

In order to generalize (4) we shall, for each of the multivariate functions $f_j(x_1,\ldots,x_k)$ with $j = 1,\ldots,k$, choose N and D such that $p(x_1,\ldots,x_k)$ and $q(x_1,\ldots,x_k)$ are linear expressions and use information in the last three iteration points

$$(x_1^{(i)},\ldots,x_k^{(i)}),(x_1^{(i-1)},\ldots,x_k^{(i-1)}),(x_1^{(i-2)},\ldots,x_k^{(i-2)}).$$

Take

$$D = N = \{(0,\ldots,0),(1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)\}\subset \mathbf{N}^k,$$
$$I = N\cup\{(2,0,\ldots,0),(0,2,0,\ldots,0),\ldots,(0,\ldots,0,2)\}\subset \mathbf{N}^k.$$

This interpolation set I expresses interpolation conditions in the points

$$\left(x_1^{(i)},\ldots,x_k^{(i)}\right),\left(x_1^{(i-1)},x_2^{(i)},\ldots,x_k^{(i)}\right),\ldots,\left(x_1^{(i)},\ldots,x_{k-1}^{(i)},x_k^{(i-1)}\right),$$
$$\left(x_1^{(i-2)},x_2^{(i)},\ldots,x_k^{(i)}\right),\ldots,\left(x_1^{(i)},\ldots,x_{k-1}^{(i)},x_k^{(i-2)}\right),$$

as will become clearer further on. We remark that this set of $(2k+1)$ interpolation points is constructed from only three successive iteration points. The numerator of

(8a)
$$\begin{aligned} r_{i,j}(x_1,\ldots,x_k) &= \frac{p_{i,j}}{q_{i,j}}(x_1,\ldots,x_k) \\ &= \frac{a_{0,\ldots,0}+\sum_{l=1}^{k} a_{0,\ldots,1,\ldots,0}(x_l-x_l^{(i)})}{b_{0,\ldots,0}+\sum_{l=1}^{k} b_{0,\ldots,1,\ldots,0}(x_l-x_l^{(i)})}, \quad j=1,\ldots,k, \end{aligned}$$

satisfying

$$(f_j q_{i,j}-p_{i,j})(x_1,\ldots,x_k)=\sum_{(l_1,\ldots,l_k)\in\mathbf{N}^k\setminus I} d_{0l_1,\ldots,0l_k}B_{l_1\ldots l_k}(x_1,\ldots,x_k),$$

where

$$B_{l_1\ldots l_k}(x_1,\ldots,x_k)=\sum_{l=0}^{l_1-1}(x_1-x_1^{(i-l)})\cdots\prod_{l=0}^{l_k-1}(x_k-x_k^{(i-l)})$$

is then given by

(8b) $p_{i,j}(x_1,\ldots,x_k)=$
$$\begin{vmatrix} N(x_1,\ldots,x_k) & c_{11,00,\ldots,00}(x_1-x_1^{(i)}) & \ldots & c_{00,\ldots,00,11}(x_k-x_k^{(i)}) \\ c_{02,00,\ldots,00} & c_{12,00,\ldots,00} & \ldots & 0 \\ \vdots & & \ddots & \\ c_{00,\ldots,00,02} & 0 & \ldots & c_{00,\ldots,00,12} \end{vmatrix},$$

where $N(x_1,\ldots,x_k)=c_{00,\ldots,00}+\sum_{l=1}^{k} c_{00,\ldots,01,\ldots,00}(x_l-x_l^{(i)})$.

The values $c_{s_1 l_1,\ldots,s_k l_k}$ are multivariate divided differences of $f_j(x_1,\ldots,x_k)$ with possible coalescence of points. Coalescent interpolation points (or coordinates) can be obtained by using the same iteration point (or coordinates) again. We remark that (5b) is only valid if the set $I\backslash N$ provides a system of linearly independent equations. The next iteration point
$(x_1^{(i+1)},\ldots,x_k^{(i+1)})$ is constructed such that

$$\text{(8c)} \qquad \begin{cases} p_{i,1}(x_1^{(i+1)},\ldots,x_k^{(i+1)}) = 0, \\ \vdots \\ p_{i,k}(x_1^{(i+1)},\ldots,x_k^{(i+1)}) = 0. \end{cases}$$

Thanks to the choice of N, (8c) is a linear system of equations. It must be clear, however, that a whole variety of choices for the sets N, D and I is possible, depending on which multivariate divided differences can be computed for use in the determinant representation (7). Some function evaluations may be much more difficult or time consuming than others. Some nonlinear functions depend on some of the variables in a more complex way than on the other variables. Let us illustrate all this by the following. Suppose f_j depends on the variable x_2 in a very complex way, and it is costly to compute divided differences like $c_{00,12,00\ldots,00}$. To avoid this, we can remove the point $(0,2,0,\ldots,0)$ from $I\backslash N$ and, for instance, replace it by $(3,0,\ldots,0)$. Then, in (7), the numerator $p_{i,j}$ becomes

$$\begin{vmatrix} N(x_1,\ldots,x_k) & \cdots & c_{11,00,\ldots,00}(x_1 - x_l^{(i)}) \\ c_{02,00,\ldots,00} & c_{12,00,\ldots,00} \; 0 \quad \cdots & 0 \\ c_{03,00,\ldots,00} & c_{13,00,\ldots,00} \; 0 \quad \cdots & 0 \\ c_{00,00,02,00,\ldots,00} \quad 0 \quad 0 & c_{00,00,12,00,\ldots,00} & 0 \\ \vdots & \ddots & \\ c_{00,\ldots,00,02} & 0 & c_{00,\ldots,00,12} \end{vmatrix},$$

where $N(x_1,\ldots,x_k) = c_{00,\ldots,00} + \sum_{l=1}^{k} c_{00,\ldots,01,\ldots,00}(x_l - x_l^{(i)})$.

Of course, one must also take into account that some functions and some variables play a more important role than others, but we only want to illustrate the flexibility and adaptability of the new method. An advantage of iterative procedures that only use function evaluations is that no derivatives must be supplied, although, nowadays automatic differentiation takes away much of the laborious work. Let us now write down an explicit formula for the iterative procedure resulting from

(8b–c). As mentioned, the correction $(x_1^{(i+1)} - x_1^{(i)}, \ldots, x_k^{(i+1)} - x_k^{(i)})$ is computed from a linear system of equations. From (8b) one can easily see that

$$\begin{pmatrix} x_1^{(i+1)} - x_1^{(i)} \\ \vdots \\ x_k^{(i+1)} - x_k^{(i)} \end{pmatrix} = -P^{-1} \begin{pmatrix} f_1(x_1^{(i)}, \ldots, x_k^{(i)}) \\ \vdots \\ f_k(x_1^{(i)}, \ldots, x_k^{(i)}) \end{pmatrix}$$

with $P = (P_{jl})_{1 \le j \le k, 1 \le l \le k}$, where

$$\begin{aligned} P_{jl} = {} & f_j[x_1^{(i)}] \cdots [x_{l-1}^{(i)}][x_l^{(i)}, x_l^{(i-1)}][x_{l+1}^{(i)}] \cdots [x_k^{(i)}] \\ & - f_j(x_1^{(i)}, \ldots, x_{l-1}^{(i)}, x_l^{(i-1)}, x_{l+1}^{(i)}, \ldots, x_k^{(i)}) \\ & \cdot \frac{f_j[x_1^{(i)}] \cdots [x_{l-1}^{(i)}][x_l^{(i)}, x_l^{(i-1)}, x_l^{(i-2)}][x_{l+1}^{(i)}] \cdots [x_k^{(i)}]}{f_j[x_1^{(i)}] \cdots [x_{l-1}^{(i)}][x_l^{(i-1)}, x_l^{(i-2)}][x_{l+1}^{(i)}] \cdots [x_k^{(i)}]}. \end{aligned}$$

If we rewrite the univariate iteration (4) using divided differences, we get

$$x^{(i+1)} - x^{(i)} = f[x^{(i)}, x^{(i-1)}] - f(x^{(i-1)}) \frac{f[x^{(i)}, x^{(i-1)}, x^{(i-2)}]}{f[x^{(i-1)}, x^{(i-2)}]}.$$

Clearly, for $k = 1$ and without coalescence of points, procedure (8) coincides with the univariate iterative method (4). With $k = 2$ and without coalescence of points, we obtain a bivariate generalization of (4). Techniques where coalescent interpolation points are used are treated in the next section.

3. Methods based on the use of multivariate general order Padé approximants. In [**3**] we described that many previously introduced multivariate Padé approximants can, within this framework, be

considered as multivariate general order rational Hermite interpolants by letting all interpolation points coincide. So we continue our construction of new nonlinear methods. For a multivariate generalization of Halley's method as given in [**4**], all first and all second partial derivatives of the functions $f_j(x_1,\ldots,x_k)$, $j=1,\ldots,k$, must be known. This may be a drawback. A completely different situation is encountered in formula (8). If none of the interpolation points and coordinates coincide, this iterative procedure uses only function evaluations. One can also construct iterative methods based on the use of rational Hermite interpolants where only some of the interpolation points (or coordinates) coincide. Let us now focus on a new generalization of Halley's method that needs only some specified partial derivatives as introduced here. Suppose a limited number of partial derivatives of the $f_j(x_1,\ldots,x_k)$ are given or easily computed. Then from this information the sets N, D and I can be chosen so that precisely these pieces of information are used. Consider, for instance, the situation where, for each of the functions $f_j(x_1,\ldots,x_k)$, the following information is given (there are no mixed second partial derivatives):

$$\begin{gathered}
f_j(x_1^{(i)},\ldots,x_k^{(i)}), \qquad j=1,\ldots,k,\\
\frac{\partial f_j}{\partial x_1}(x_1^{(i)},\ldots,x_k^{(i)}),\ldots,\frac{\partial f_j}{\partial x_k}(x_1^{(i)},\ldots,x_k^{(i)}), \qquad j=1,\ldots,k,\\
\frac{\partial^2 f_j}{\partial x_1^2}(x_1^{(i)},\ldots,x_k^{(i)}),\ldots,\frac{\partial^2 f_j}{\partial x_k^2}(x_1^{(i)},\ldots,x_k^{(i)}), \qquad j=1,\ldots,k.
\end{gathered}$$

With the same sets N, D and I as in the previous section,

$$\begin{gathered}
D = N = \{(0,\ldots,0),(1,0,\ldots,0),\ldots,(0,\ldots,0,1)\},\\
I\backslash N = \{(2,0,\ldots,0),\ldots,(0,\ldots,0,2)\},
\end{gathered}$$

but now, with all the interpolation points coinciding, this is precisely the informational usage of the iterative procedure (9), as shown below. The numerator of

$$\begin{aligned}
r_{i,j}(x_1,\ldots,x_k) &= \frac{p_{i,j}}{q_{i,j}}(x_1,\ldots,x_k)\\
&= \frac{\displaystyle\sum_{(l_1,\ldots,l_k)\in N} a_{l_1\ldots l_k}(x_1-x_1^{(i)})^{l_1}\ldots(x_k-x_k^{(i)})^{l_k}}{\displaystyle\sum_{(l_1,\ldots,l_k)\in D} b_{l_1\ldots l_k}(x_1-x_1^{(i)})^{l_1}\ldots(x_k-x_k^{(i)})^{l_k}},
\end{aligned}$$

satisfying

$$(9a)\quad (f_j q_{i,j} - p_{i,j})(x_1,\ldots,x_k) = \sum_{(l_1,\ldots,l_k)\in\mathbf{N}^k\setminus I} d_{l_1\ldots l_k}(x_1 - x_1^{(i)})^{l_1}\ldots(x_k - x_k^{(i)})^{l_k},$$

is given by

$$(9b)\quad p_{i,j}(x_1,\ldots,x_k) = \begin{vmatrix} f_j(x^{(i)}) + \sum_{l=1}^{k} \frac{\partial f_j}{\partial x_l}(x_l - x_l^{(i)}) & f_j(x^{(i)})(x_1 - x_1^{(i)}) & \ldots & f_j(x^{(i)})(x_k - x_k^{(i)}) \\ \frac{1}{2}\frac{\partial^2 f_j}{\partial x_1^2}(x_1^{(i)},\ldots,x_k^{(i)}) & \frac{\partial f_j}{\partial x_1}(x_1^{(i)},\ldots,x_k^{(i)}) & & 0 \\ \vdots & & \ddots & \\ \frac{1}{2}\frac{\partial^2 f_j}{\partial x_k^2}(x_1^{(i)},\ldots,x_k^{(i)}) & 0 & & \frac{\partial f_j}{\partial x_k}(x_1^{(i)},\ldots,x_k^{(i)}) \end{vmatrix}.$$

We have, of course, assumed that the set $I\setminus N$ provided a system of linearly independent equations. As in the previous section, the next iteration point is constructed such that

$$(9c)\quad \begin{cases} p_{i,1}(x_1^{(i+1)},\ldots,x_k^{(i+1)}) = 0, \\ \quad\vdots \\ p_{i,k}(x_1^{(i+1)},\ldots,x_k^{(i+1)}) = 0. \end{cases}$$

If we give an explicit formula for this iterative procedure, we find that

$$\begin{pmatrix} x_1^{(i+1)} - x_1^{(i)} \\ \vdots \\ x_k^{(i+1)} - x_k^{(i)} \end{pmatrix} = -P^{-1} \begin{pmatrix} f_1(x_1^{(i)},\ldots,x_k^{(i)}) \\ \vdots \\ f_k(x_1^{(i)},\ldots,x_k^{(i)}) \end{pmatrix}$$

with $P = (P_{jl})_{1\le j\le k, 1\le l\le k}$, where

$$P_{jl} = \left(\frac{\partial f_j}{\partial x_l} - \frac{1}{2} f_j \frac{\partial^2 f_j}{\partial x_l^2} \Big/ \frac{\partial f_j}{\partial x_l}\right)_{|(x_1,\ldots,x_k)=(x_1^{(i)},\ldots,x_k^{(i)})}.$$

Clearly, (9b–c) is a multivariate generalization of Halley's method (2). Finally, we emphasize once more that this section covers iterative

procedures adapted to the cost of evaluating and differentiating the nonlinear equations in the system. By varying the set I, more multivariate "Halley-like" procedures can be derived, as must be clear from the foregoing. In the same way as we obtained (9), we can also set up nonlinear methods based on the use of Canterbury approximants [**1**], Karlsson and Wallin approximants [**7**], Lutterodt approximants [**8**] and all types of multivariate general order Padé approximants that can be described in this framework [**3**].

4. Numerical considerations. Just as for the univariate techniques, the newly introduced nonlinear methods behave much better than analogous linear methods when the considered equations have singularities in the neighborhood of the roots [**4**]. Since the numerical behavior in such situations is similar using either iteration (8) or iteration (9), except for a difference in convergence order, we shall only illustrate (8). It is obvious that the exact order of convergence of this type of method is a difficult problem and should be the subject of further investigation.

Another advantage of these rational procedures is that one can better handle nonlinear systems which have a singular or nearly-singular Jacobian in the neighborhood of the root (multiple zeros, bifurcation points, ...) because the matrix P is not an approximation to the Jacobian but is "rational" in nature. This is good because Newton-like methods will generate ill-conditioned approximations to the (nearly)-singular Jacobian.

Let us illustrate (8) by solving the system [**5**]

$$\begin{cases} f_1(x,y) = e^{-x+y} - 0.1 = 0, \\ f_2(x,y) = e^{-x-y} - 0.1 = 0, \end{cases}$$

with the points (3.2,−0.95), (3.4,−1.15) and (3.3,−1.00) as initial points for the iterative procedure (8). The numerical results are displayed in Table 1. The consecutive iteration steps in a discretized Newton method with the same but fewer initial points as above can be found in Table 2. Here the partial derivatives of the Jacobian are approximated

by the difference quotients

$$\left.\frac{\partial f_j}{\partial x_l}\right|_{(x_1^{(i)},\ldots,x_k^{(i)})} \approx \frac{f_j(x_1^{(i)},\ldots,x_{l-1}^{(i)},x_l^{(i-1)},x_{l+1}^{(i)},\ldots,x_k^{(i)}) - f_j(x_1^{(i)},\ldots,x_k^{(i)})}{x_l^{(i-1)} - x_l^{(i)}},$$

which result from choosing $N = I = \{(0,\ldots,0),(1,0,\ldots,0),\ldots,(0,\ldots,0,1)\}$ and $D = \{(0,\ldots,0)\}$. The simple root is (2.302585092994046,0). All computations are performed in double precision. The rational method is rapidly converging while the linear method is diverging because during the iteration $f_1(x,y)$ comes close to -0.1, which is precisely a singularity of the inverse operator

$$\begin{cases} g_1(x,y) = \frac{-\ln(x+0.1)-\ln(y+0.1)}{2}, \\ g_2(x,y) = \frac{\ln(x+0.1)-\ln(y+0.1)}{2} \end{cases}$$

for the considered system of nonlinear equations.

i—	$x^{(i)}$—	$y^{(i)}$
—	$0.32000000D+01$	$-0.95000000D+00$
—	$0.34000000D+01$	$-0.11500000D+01$
0—	$0.33000000D+01$	$-0.10000000D+01$
1—	$0.25249070D+01$	$-0.22072875D+00$
2—	$0.22618832D+01$	$0.41971944D-01$
3—	$0.23127609D+01$	$-0.10164490D-01$
4—	$0.23030978D+01$	$-0.51269373D-03$
5—	$0.23025801D+01$	$0.49675854D-05$
6—	$0.23025851D+01$	$-0.25696929D-08$
7—	$0.23025851D+01$	$-0.12778916D-13$
8—	$0.23025851D+01$	$-0.11350932D-16$

TABLE 1.

i	$x^{(i)}$	$y^{(i)}$
—	$0.34000000D+01$	$-0.11500000D+01$
0	$0.33000000D+01$	$-0.10000000D+01$
1	$-0.29618530D+00$	$0.21743633D+01$
2	$0.32743183D+01$	$0.20884933D+01$
3	$0.22114211D+01$	$-0.84011352D+01$
4	$0.36513339D+01$	$-0.72149651D+01$
5	$-0.17900983D+04$	$0.20854111D+04$

TABLE 2.

REFERENCES

1. J.S.R. Chisholm, *N-variable rational approximants*, in [**9**], 23–42.

2. G. Claessens, *On the Newton-Padé approximation problem*, J. Approx. Theory **22** (1978), 150–160.

3. Annie A.M. Cuyt, *Multivariate Padé approximants revisited*, BIT **26** (1986), 71–79.

4. ———, *Padé approximants for operators: theory and applications*, LNM **1065**, Springer Verlag, Berlin, 1984.

5. ——— and P. Van der Cruyssen, *Abstract Padé approximants for the solution of a system of nonlinear equations*, Comput. Math. Appl. **9** (1983), 617–624.

6. A. Cuyt and B. Verdonk, *General order Newton-Padé approximants for multivariate functions*, Numer. Math. **43** (1984), 293–307.

7. J. Karlsson and H. Wallin, *Rational approximation by an interpolation procedure in several variables*, in [**9**], 83–100.

8. C. Lutterodt, *Rational approximants to holomorphic functions in n dimensions*, J. Math. Anal. Appl. **53** (1976), 89–98.

9. E. Saff and R. Varga, *Padé and rational approximation: theory and applications*, Academic Press, New York, 1970.

10. L. Tornheim, *Convergence of multipoint iterative methods*, J. Assoc. Comp. Mach. **11** (1964), 210–220.

Department of Mathematics and Computer Science, University of Antwerp, Universiteitsplein 1, B-2610 Wilrijk-Antwerp, Belgium