The MPE (Minimal Polynomial Extrapolation) and Vector-ϵ Methods: Numerical Demonstration of their Equivalence in Certain Cases

Yoshio Takane

McGill University

1 Introduction

We exemplify the equivalence between the MPE acceleration method and the vector ϵ (v-ϵ) method when the iterate is linear and the exact k is chosen. General results have been given in McLeod (1971), and Graves-Morris (1983) among others, as discussed by Smith et al. (1987). Let

$$\mathbf{x}^{(q+1)} = \mathbf{H}\mathbf{x}^{(q)} + \mathbf{b} \tag{1}$$

represent the basic iterate, where it is assumed that the largest absolute eigenvalue of $\mathbf{H}$ is strictly smaller than unity. The closed-form solution to the above system is given by

$$\mathbf{s} = \mathbf{x}^{(\infty)} = (\mathbf{I} - \mathbf{H})^{-1}\mathbf{b}, \tag{2}$$

where $\mathbf{I}-\mathbf{H}$ is nonsingular.

Let

$$\mathbf{u}^{(q)} = \mathbf{x}^{(q+1)} - \mathbf{x}^{(q)}. \tag{3}$$

In the MPE method, k represents the order of minimal polynomials that annihilates $\mathbf{u}^{(0)}$. Let

$$\mathbf{U} = [\mathbf{u}^{(0)}, \cdots, \mathbf{u}^{(k-1)}]. \tag{4}$$

We define $\mathbf{c} = (c_0, \cdots, c_{k-1}, 1)'$ by

$$\mathbf{c} = \begin{pmatrix} -\mathbf{U}^{+}\mathbf{u}^{(k)} \\ 1 \end{pmatrix}. \tag{5}$$

In the MPE method, $\mathbf{s}$ is obtained by

$$\mathbf{s} = \mathbf{X}\mathbf{c}/d, \tag{6}$$

where $\mathbf{X} = [\mathbf{x}^{(0)}, \cdots, \mathbf{x}^{(k)}]$, and $d = \mathbf{1}'_{k+1}\mathbf{c}$.

In the v-ϵ method, we define $\epsilon_{-1}^{(q)} = \mathbf{0}$, $\epsilon_0^{(q)} = \mathbf{x}^{(q)}$ for $q = 0, 1, \cdots$, and

$$\epsilon_{k+1}^{(q)} = \epsilon_{k-1}^{(q+1)} + (\epsilon_k^{(q+1)} - \epsilon_k^{(q)})^{-1}, \tag{7}$$

where the inverse of a vector $\mathbf{a}$ is defined to be

$$\mathbf{a}^{-1} = \mathbf{a}/\mathbf{a}'\mathbf{a}. \tag{8}$$

(This is called the Samelson inverse of $\mathbf{a}$, and is equal to the transpose of the Moore-Penrose inverse of $\mathbf{a}$ considered as a matrix.) Then,

$$\mathbf{s} = \epsilon_{2k}^{(0)}. \tag{9}$$

2 Demonstrations of the equivalences

2.1 A case of scalar variable

We assume there is a single variable in $\mathbf{x}$, which will be denoted as x. Suppose we have the following updating formula:

$$x^{(q+1)} = (1/2)x^{(q)} + 1. \tag{10}$$

the closed form solution to this system is $s = 1/(1/2) = 2$. Suppose that the iteration starts at $x^{(0)} = 0$. Then we have

Table 1: Successive updates of $x^{(q)}$ and resultant $u^{(q)}$.

$x^{(0)}$	$x^{(1)}$	$x^{(2)}$
0	1	1.5

$u^{(0)}$	$u^{(1)}$
1	.5

In the MPE method, we obtain $\mathbf{c}_0 = -u^{(1)}/u^{(0)} = -.5$, and $c_1 = 1$ (by definition). so that

$$s = \frac{c_0 x^{(0)} + c_1 x^{(1)}}{c_0 + c_1} = \frac{-.5(0) + 1(1)}{-.5 + 1} = 2. \tag{11}$$

The general formula for s is given by

$$s = \frac{x^{(0)}u^{(1)} - x^{(1)}u^{(0)}}{u^{(1)} - u^{(0)}}. \tag{12}$$

In the v-ϵ method, we would like to get

$$\epsilon_2^{(0)} = \epsilon_0^{(1)} + (\epsilon_1^{(1)} - \epsilon_1^{(0)})^{-1},$$

where

$$\epsilon_0^{(1)} = x^{(1)},$$

$$\epsilon_1^{(1)} = \epsilon_{-1}^{(2)} + (\epsilon_0^{(2)} - \epsilon_0^{(1)})^{-1},$$

and

$$\epsilon_1^{(0)} = \epsilon_{-1}^{(1)} + (\epsilon_0^{(1)} - \epsilon_0^{(0)})^{-1}.$$

See Table 2 below. We have

$$\epsilon_1^{(0)} = (u^{(0)})^{-1} = (1)^{-1} = 1,$$

and

$$\epsilon_1^{(1)} = (u^{(1)})^{-1} = (.5)^{-1} = 2,$$

so that

$$\epsilon_2^{(0)} = s = x^{(1)} + ((u^{(1)})^{-1} - (u^{(0)})^{-1})^{-1} = 1 + (2-1)^{-1} = 1 + 1 = 2.$$

Table 2: The ϵ table for the v-ϵ method.

$k=-1$	$k=0$	$k=1$	$k=2$
$\epsilon_{-1}^{(0)} = 0$			
	$\epsilon_0^{(0)} = x^{(0)}$		
$\epsilon_{-1}^{(1)} = 0$		$\epsilon_1^{(0)} = (u^{(0)})^{-1}$	
	$\epsilon_0^{(1)} = x^{(1)}$		$\epsilon_2^{(0)}$
$\epsilon_{-1}^{(2)} = 0$		$\epsilon_1^{(1)} = (u^{(1)})^{-1}$	
	$\epsilon_0^{(2)} = x^{(2)}$		

The general formula for s is given by

$$\begin{aligned} s &= x^{(1)} + ((u^{(1)})^{-1} - (u^{(0)})^{-1})^{-1} \\ &= x^{(1)} + \frac{u^{(0)}u^{(1)}}{u^{(0)} - u^{(1)}} \\ &= \frac{u^{(0)}x^{(1)} - u^{(1)}x^{(1)} + u^{(0)}u^{(1)}}{u^{(0)} - u^{(1)}} \\ &= \frac{x^{(0)}u^{(1)} - x^{(1)}u^{(0)}}{u^{(1)} - u^{(0)}}, \end{aligned} \tag{13}$$

which is identical to (12).

The above line of argument remains essentially the same if we start from a different value of $x^{(0)}$, for example, $x^{(0)} = -1$.

2.2 A two-variable case

We consider

$$\mathbf{x}^{(q+1)} = \begin{bmatrix} .7 & 0 \\ 0 & .3 \end{bmatrix} \mathbf{x}^{(q)} + \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \tag{14}$$

(Matrix $\mathbf{H}$ may be assumed diagonal without loss of generality.) The closed-form solution is given by

$$\mathbf{s} = \begin{pmatrix} 3.3333 \\ 2.8571 \end{pmatrix}. \tag{15}$$

We obtain the following table.

For the MPE method with $k = 2$, we have

$$\mathbf{c} = \begin{pmatrix} -[\mathbf{u}^{(1)}, \mathbf{u}^{(2)}]^{-1}\mathbf{u}^{(3)} \\ 1 \end{pmatrix} = \begin{pmatrix} .21 \\ -1 \\ 1 \end{pmatrix}. \tag{16}$$

Table 3: Successive updates of $\mathbf{x}^{(q)}$ and resultant $\mathbf{u}^{(q)}$ for the two-variable case.

$\mathbf{x}^{(0)}$	$\mathbf{x}^{(1)}$	$\mathbf{x}^{(2)}$	$\mathbf{x}^{(3)}$	$\mathbf{x}^{(4)}$
0	1	1.7	2.19	2.533
0	2	2.6	2.78	2.834

$\mathbf{u}^{(0)}$	$\mathbf{u}^{(1)}$	$\mathbf{u}^{(2)}$	$\mathbf{u}^{(3)}$
1	.7	.49	.343
2	.6	.18	.054

Thus,

$$\mathbf{s} = \{.21 \begin{pmatrix} 0 \\ 0 \end{pmatrix} - 1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 1 \begin{pmatrix} 1.7 \\ 2.6 \end{pmatrix} \}/.21 = \begin{pmatrix} .7 \\ .6 \end{pmatrix} /.21 = \begin{pmatrix} 3.3333 \\ 2.8571 \end{pmatrix}. \tag{17}$$

For the v-ϵ method, we would like to get

$$\epsilon_4^{(0)} = \epsilon_2^{(1)} + (\epsilon_3^{(1)} - \epsilon_3^{(0)})^{-1},$$

where

$$\epsilon_2^{(1)} = \epsilon_0^{(2)} + (\epsilon_1^{(2)} - \epsilon_1^{(1)})^{-1},$$

$$\epsilon_3^{(1)} = \epsilon_1^{(2)} + (\epsilon_2^{(2)} - \epsilon_2^{(1)})^{-1},$$

and

$$\epsilon_3^{(0)} = \epsilon_1^{(1)} - (\epsilon_2^{(1)} - \epsilon_2^{(0)})^{-1}.$$

The $\epsilon_2^{(q)}$ for $q = 0, 1, 2$ we need to calculate $\epsilon_3^{(q)}$ for $q = 0, 1$, on the other hand, are obtained by

$$\epsilon_2^{(0)} = \epsilon_0^{(1)} + (\epsilon_1^{(1)} - \epsilon_1^{(0)})^{-1},$$

$$\epsilon_2^{(1)} = \epsilon_0^{(2)} + (\epsilon_1^{(2)} - \epsilon_1^{(1)})^{-1},$$

and

$$\epsilon_2^{(2)} = \epsilon_0^{(3)} + (\epsilon_1^{(3)} - \epsilon_1^{(2)})^{-1}.$$

The $\epsilon_1^{(q)}$ for $q = 0, 1, 2, 3$, in turn, are obtained by $\epsilon_{-1}^{(q+1)} + (\epsilon_0^{(q+1)} - \epsilon_0^{(q)})^{-1}$. See Table 4 below.

Table 4: The ϵ table for the v-ϵ method.

$k=-1$	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
$\epsilon_{-1}^{(0)} = \mathbf{0}$					
	$\epsilon_0^{(0)} = \mathbf{x}^{(0)}$				
$\epsilon_{-1}^{(1)} = \mathbf{0}$		$\epsilon_1^{(0)} = (\mathbf{u}^{(0)})^{-1}$			
	$\epsilon_0^{(1)} = \mathbf{x}^{(1)}$		$\epsilon_2^{(0)}$		
$\epsilon_{-1}^{(2)} = \mathbf{0}$		$\epsilon_1^{(1)} = (\mathbf{u}^{(1)})^{-1}$		$\epsilon_3^{(0)}$	
	$\epsilon_0^{(2)} = \mathbf{x}^{(2)}$		$\epsilon_2^{(1)}$		$\epsilon_4^{(0)}$
$\epsilon_{-1}^{(3)} = \mathbf{0}$		$\epsilon_1^{(2)} = (\mathbf{u}^{(2)})^{-1}$		$\epsilon_3^{(1)}$	
	$\epsilon_0^{(3)} = \mathbf{x}^{(3)}$		$\epsilon_2^{(2)}$		
$\epsilon_{-1}^{(4)} = \mathbf{0}$		$\epsilon_1^{(3)} = (\mathbf{u}^{(3)})^{-1}$			
	$\epsilon_0^{(4)} = \mathbf{x}^{(4)}$				

We have

$$\epsilon_1^{(q)} = (\mathbf{u}^{(q)})^{-1}$$

for $q = 0, 1, 2, 3$, where

$$(\mathbf{u}^{(0)})^{-1} = \begin{pmatrix} .2 \\ .4 \end{pmatrix}, \quad (\mathbf{u}^{(1)})^{-1} = \begin{pmatrix} .8235 \\ .7059 \end{pmatrix},$$

$$(\mathbf{u}^{(2)})^{-1} = \begin{pmatrix} 1.7982 \\ .6606 \end{pmatrix}, \quad (\mathbf{u}^{(3)})^{-1} = \begin{pmatrix} 2.8449 \\ .4479 \end{pmatrix},$$

$$\begin{aligned} \epsilon_2^{(0)} &= \mathbf{x}^{(1)} + ((\mathbf{u}^{(1)})^{-1} - (\mathbf{u}^{(0)})^{-1})^{-1} \\ &= \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \left(\begin{pmatrix} .8235 \\ .7059 \end{pmatrix} - \begin{pmatrix} .2 \\ .4 \end{pmatrix}\right)^{-1} = \begin{pmatrix} 2.2927 \\ 2.6341 \end{pmatrix}, \end{aligned}$$

$$\begin{aligned} \epsilon_2^{(1)} &= \mathbf{x}^{(2)} + ((\mathbf{u}^{(2)})^{-1} - (\mathbf{u}^{(1)})^{-1})^{-1} \\ &= \begin{pmatrix} 1.7 \\ 2.6 \end{pmatrix} + \left(\begin{pmatrix} 1.7982 \\ .6606 \end{pmatrix} - \begin{pmatrix} .8235 \\ .7059 \end{pmatrix}\right)^{-1} = \begin{pmatrix} 2.7238 \\ 2.5524 \end{pmatrix}, \end{aligned}$$

and

$$\begin{aligned} \epsilon_2^{(2)} &= \mathbf{x}^{(3)} + ((\mathbf{u}^{(3)})^{-1} - (\mathbf{u}^{(2)})^{-1})^{-1} \\ &= \begin{pmatrix} 2.19 \\ 2.78 \end{pmatrix} + \left(\begin{pmatrix} 2.8449 \\ .4479 \end{pmatrix} - \begin{pmatrix} 1.7982 \\ .606 \end{pmatrix}\right)^{-1} = \begin{pmatrix} 3.1075 \\ 2.5936 \end{pmatrix}. \end{aligned}$$

Finally,

$$\begin{aligned} \epsilon_3^{(0)} &= (\mathbf{u}^{(1)})^{-1} + (\mathbf{u}^{(1)} + ((\mathbf{u}^{(2)})^{-1} - (\mathbf{u}^{(1)})^{-1})^{-1} - ((\mathbf{u}^{(1)})^{-1} - (\mathbf{u}^{(0)})^{-1})^{-1})^{-1} \\ &= \begin{pmatrix} 3.0625 \\ .2813 \end{pmatrix}, \end{aligned}$$

$$\begin{aligned} \epsilon_3^{(1)} &= (\mathbf{u}^{(2)})^{-1} + (\mathbf{u}^{(2)} + ((\mathbf{u}^{(3)})^{-1} - (\mathbf{u}^{(2)})^{-1})^{-1} - ((\mathbf{u}^{(2)})^{-1} - (\mathbf{u}^{(1)})^{-1})^{-1})^{-1} \\ &= \begin{pmatrix} 4.3750 \\ .9375 \end{pmatrix}, \end{aligned}$$

and

$$\epsilon_4^{(0)} = \begin{pmatrix} 2.7238 \\ 2.5524 \end{pmatrix} + \left(\begin{pmatrix} 4.3750 \\ .9375 \end{pmatrix} - \begin{pmatrix} 3.0625 \\ .2813 \end{pmatrix}\right)^{-1} = \begin{pmatrix} 3.3333 \\ 2.8571 \end{pmatrix}.$$

Again, we get an identical result from the two methods.

A general formula for $\epsilon_4^{(0)}$ is given by

$$\epsilon_4^{(0)} = \mathbf{x}^{(2)} + ((\mathbf{u}^{(2)})^{-1} - (\mathbf{u}^{(1)})^{-1})^{-1} + \\ \{(\mathbf{u}^{(2)})^{-1} + (\mathbf{u}^{(2)} + ((\mathbf{u}^{(3)})^{-1} - (\mathbf{u}^{(2)})^{-1})^{-1} - ((\mathbf{u}^{(2)})^{-1} - (\mathbf{u}^{(1)})^{-1})^{-1})^{-1} \\ -(\mathbf{u}^{(1)})^{-1} - (\mathbf{u}^{(1)} + ((\mathbf{u}^{(2)})^{-1} - (\mathbf{u}^{(1)})^{-1})^{-1} - ((\mathbf{u}^{(1)})^{-1} - (\mathbf{u}^{(0)})^{-1})^{-1})^{-1}\}^{-1}. \tag{18}$$

This should theoretically be identical to (Smith, et al., 1987, Sect. 5)

$$\mathbf{s} = \sum_{q=0}^{k} \gamma_q \mathbf{x}^{(q)},$$

where

$$\gamma_q = c_q / \sum_{j=1}^{k} c_j,$$

although it is by no means obvious from these formulas.

References

Graves-Morris, P. R. (1983). Vector valued rational interpolants I. *Numerische Mathematik*, **42**, 331-348.

McLeod, J. B. (1971). A note on the ϵ-algorithm. *Computing*, **7**, 17-24.

Smith, D. A., Ford, W. F., and Sidi, A. (1987). Extrapolation methods for vector sequences. *SIAM Review*, **29**, 199-233.