Linear extrapolation in iterative methods

Alexander V. Mitin [1] and Gerhard Hirsch

Bergische Universität – Gesamthochschule Wuppertal, Fachbereich 9, Theoretische Chemie, Gaußstraße 20, D-42097 Wuppertal, Germany

Received 7 December 1992; revised 23 August 1993

The application of extrapolation methods in an iterative process of general type is investigated. It is shown that extrapolation methods are obtained by two successive approximations from the formula defining an iterative process. The first approximation gives the general formula for extrapolation methods, while particular examples are obtained from this formula by introducing different approximations for the operator defining the iterative process.

1. Introduction

Iterative methods are widely used in quantum chemistry to obtain the solution of given equations, for example, in coupled-cluster methods for the calculation of molecular wavefunctions [1], or for the calculation of second derivatives of the energy functional [2]. Often the solution is difficult to obtain due to the slow convergence of the iterative process. One of the ways for accelerating convergence is to use appropriate extrapolation methods in the iterative process, but such methods are used as a rule without proper investigation of their applicability in a given procedure. In this connection the application of extrapolation methods to an iterative process of self-consistent type has been studied in previous work [3], while in the present study their application in an iterative process of general type is investigated. It is shown that extrapolation methods are obtained by two successive approximations from the formula defining the iterative process. At the beginning the problem of the extrapolation of a sequence of vectors, obtained in an iterative process, reduces to one of interpolation of this sequence of vectors. This gives the general formula for the extrapolation. Then the various extrapolation methods are obtained by introducing different approximations for the operator defining the iterative process. This approach to the construction of extrapolation methods follows from the work of Lanczos [4], where the convergence of power series is accelerated by interpolating them by means of a series of Chebyshev polynomials, and it

[1] Alexander von Humboldt fellow. Permanent address: Institute for High Temperatures, Russian Academy of Sciences, Izhorskaya ul. 13/19, Moscow 127412, Russia.

had been used in [3] for the construction of extrapolation methods in iterative processes of self-consistent type.

2. General form of linear extrapolation methods

Consider an iterative process in Hilbert space H [5,6]

$$X_{k+1} = \mathbb{A}X_k + B\,. \tag{1}$$

Let the vectors $X_1, \ldots, X_k, \ldots, X_\infty$ be obtained at the 1st, ..., kth, ..., ∞th steps of this process. Define the vectors

$$D_{k+1} = X_{k+1} - X_k\,. \tag{2}$$

They satisfy the equation

$$D_{k+n} = (\mathbb{A})^n D_k\,. \tag{3}$$

By definition (2) the vector X_{k+mp}, obtained at the $(k+mp)$th iteration, can be written as

$$X_{k+mp} = X_k + D_{k+1} + \ldots + D_{k+mp}\,. \tag{4}$$

Then using (3) one can write (4) as

$$\begin{aligned} X_{k+mp} =& X_k + (D_{k+1} + \ldots + D_{k+p}) + (D_{k+p+1} + \ldots + D_{k+2p}) \\ &+ \ldots + (D_{k+(m-1)p} + \ldots + D_{k+mp}) \\ =& X_k + (D_{k+1} + \ldots + D_{k+p}) + (\mathbb{A})^p (D_{k+1} + \ldots + D_{k+p}) \\ &+ \ldots + (\mathbb{A})^{(m-1)p} (D_{k+1} + \ldots + D_{k+p}) \\ =& X_k + [I + (\mathbb{A})^p + \ldots + (\mathbb{A})^{(m-1)p}](D_{k+1} + \ldots + D_{k+p}) \\ =& X_k + \left[\sum_{l=0}^{m} (\mathbb{A})^{lp}\right] (D_{k+1} + \ldots + D_{k+p})\,, \end{aligned} \tag{5}$$

where I is the identity operator. Passing to the limit with respect to m in (5) one obtains the extrapolating vector

$$X_\infty^* = \lim_{m\to\infty} X_{k+mp} = X_k + \left[\sum_{l=0}^{\infty} (\mathbb{A})^{lp}\right] (D_{k+1} + \ldots + D_{k+p})\,, \tag{6}$$

which is an approximation to the vector X_∞. For a descending sequence we have $\|\mathbb{A}\| < \|I\|$ and hence the series $\sum_{l=0}^{\infty} (\mathbb{A})^{lp}$ is the expansion of the operator $[I - (\mathbb{A})^p]^{-1}$ in powers of $(\mathbb{A})^p$. Noting this, one can write (6) in the form

$$X_\infty^* = X_k + [I - (\mathbb{A})^p]^{-1} (D_{k+1} + \ldots + D_{k+p})\,. \tag{7}$$

Substituting D_{k+1} from (2) in this expression one obtains

$$X_\infty^* = X_k + [I - (\mathbb{A})^p]^{-1}(X_{k+p} - X_k)\,. \tag{8}$$

Equation (8) has the form of the well-known Lagrange interpolation method

$$f(x_0 + \alpha h) = \alpha f(x_0 + h) + (1 - \alpha)f(x_0) + O(h^2)\,. \tag{9}$$

The transformation from the extrapolation of the sequence of vectors $\{X_k\}_{k=0}^{\infty}$ to the problem of interpolation is connected with passing to the limit in (5).

Equation (8) is the basis for the construction of linear extrapolation methods, whereby various types can be obtained from this equation by introducing different approximations for the operator $(\mathbb{A})^p$ or $[I - (\mathbb{A})^p]^{-1}$. In the following we shall consider the most interesting of these methods.

3. Extrapolation methods

At first it can be noted that expression (8) may be written with two iteration steps.

$$X_\infty^* = X_k + [I - (\mathbb{A})^p]^{-1}(X_{k+p} - X_k)\,, \tag{10a}$$

$$X_\infty^* = X_{k+d} + [I - (\mathbb{A})^p]^{-1}(X_{k+d+p} - X_{k+d})\,. \tag{10b}$$

This permits one to eliminate the operator $(\mathbb{A})^p$ and to obtain

$$X_\infty^* = \frac{X_{k+d+p}X_k - X_{k+p}X_{k+d}}{X_{k+d+p} - X_{k+p} - X_{k+d} + X_k}\,, \quad p = 1, 2, \ldots; d = 1, 2, \ldots\,. \tag{11}$$

This expression defines the linear four-point Δ^2-method of extrapolation. In the case of $p = 1, d = 1$ this procedure reduces to Aitken's three-point δ^2-method of extrapolation [7]. Comparing (11) and (8) one can see that the method (11) is obtained from (8) when the operator $(\mathbb{A})^p$ is approximated by

$$(\mathbb{A})_{ij}^p = \begin{cases} \dfrac{X_{k+d+p,j} - X_{k+p,j}}{X_{k+d,j} - X_{k,j}}\,, & i = j\,, \\ 0, & i \neq j\,, \end{cases} \tag{12}$$

where $(\mathbb{A})_{ij}^p$ is the i,j component of the operator $(\mathbb{A})^p$ and $X_{n,j}$ is the j component of vector X_n.

As known [5], the linear operator defining an iterative process can be restored from the vectors obtained during this process. The approximate form of this operator can be determined by using a limited number of vectors, and in accordance with the approach described above this approximate form can be used to construct the extrapolation methods based on eq. (8).

Consider the approximation of the operator $R = [I - (\mathbb{A})^p]^{-1}$ by an operator R_n in the subspace H_n of the complete space H. Let Q_n be the projector onto subspace H_n. Then, using relationship $Q_n = (Q_n)^2$, operator R_n approximating R can be represented in the form

$$R_n = Q_n R Q_n = Q_n[Q_n I Q_n - Q_n(\mathbb{A})^p Q_n]^{-1} Q_n \,. \tag{13}$$

This approximation of the operator leads to the general form of linear projection extrapolation methods

$$X^*_\infty = X_k + Q_n[Q_n I Q_n - Q_n(\mathbb{A})^p Q_n]^{-1} Q_n(X_{k+p} - X_k) \,. \tag{14}$$

This general form can be simplified if projectors P^n_i onto the subspace corresponding to the eigenvalues λ^n_i of the operator $Q_n \mathbb{A} Q_n$ are known. In this case the linear projection extrapolation method has the form

$$X^*_\infty = X_k + \left[\sum \frac{P^n_i}{1 - (\lambda^n_i)^p}\right](X_{k+p} - X_k) \,. \tag{15}$$

Let us consider the expansion of the operator $[I - (\mathbb{A})^p]^{-1}$ in a power series. Then expression (8) is also a power series, and different extrapolation methods can be constructed by using the technique of accelerating convergence of a power series.

As is well known [8], the Padé approximants of a power series give a better estimate of the sum of the series then the direct summation of the terms of the series. Consider the expansion of the operator $[I - (\mathbb{A})^p]^{-1}$ in a power series at $p = 1$:

$$X^*_\infty - X_k = D_k + \mathbb{A} D_{k+1} + \mathbb{A}^2 D_{k+2} + \ldots + \,. \tag{16}$$

Then extrapolation methods of Padé type are obtained by replacing the finite terms of the series (16) by their Padé approximants

$$X^*_\infty = X_k + P_{m,n}(D) \,. \tag{17}$$

We can write a number of Padé approximants for series (16):

$$P_{0,1}(D) = D_{k+1}(1 - D_{k+2}/D_{k+1})^{-1} \,, \tag{18a}$$

$$P_{0,2}(D) = (D_{k+1})^3[(D_{k+1})^2 - D_{k+2}D_{k+1} + (D_{k+2})^2 - D_{k+3}D_{k+1}]^{-1} \,, \tag{18b}$$

$$P_{1,1}(D) = D_{k+1}[D_{k+2}D_{k+1} + (D_{k+2})^2 - D_{k+3}D_{k+1}][D_{k+2}(D_{k+1} - D_{k+2})]^{-1} \,, \tag{18c}$$

$$P_{2,1}(D) = D_{k+1} + D_{k+2} + D_{k+3}(1 - D_{k+4}/D_{k+3})^{-1} \,. \tag{18d}$$

It is interesting to note that the extrapolation method with Padé approximant $P_{0,1}(D)$ reduces to the Aitken δ^2-method of extrapolation [7].

Acknowledgements

One of us (AVM) is grateful to the Alexander von Humboldt foundation for supporting this work. We are grateful to Professor R.J. Buenker for many useful discussions and his support of this work.

References

[1] P. Pulay, Adv. Chem. Phys. 69 (1987) 241.
[2] R.J. Bartlett, Ann. Rev. Phys. Chem. 32 (1981) 359.
[3] A. Mitin, Zh. Vitchislit. Matem. i Matem. Fiz. 25 (1985) 325.
[4] C. Lanczos, J. Math. Phys. 17 (1938/1939) 123.
[5] D.K. Faddeev and V.N. Faddeeva, *Computational Methods of Linear Algebra* (Freeman, San Francisco/London, 1963).
[6] J.H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon, Oxford, 1988).
[7] A. Aitken, Proc. Roy. Soc. Edinb. A57 (1937) 269.
[8] G.A. Baker, Jr., *Essentials of Padé Approximants* (Academic Press, New York, 1975).