Convergence of Multipower Defect Correction for Spectral Computations of Integral Operators *

Filomena D. d'Almeida

Faculty of Engineering of the University of Porto and Center of Mathematics at the University of Porto, Portugal

Paulo B. Vasconcelos

Faculty of Economics of the University of Porto and Center of Mathematics at the University of Porto, Portugal

Abstract

We propose an extension to the double iteration method, the Multipower Defect Correction method, which is used to refine approximate eigenelements of integral operators. It consists of an inner/outer iteration where, inside a defect correction iteration, p steps of a power iteration are used. The approximate inverse used in the defect correction is built with an approximation to the reduced resolvent operator of a small size discretization of the integral operator. The kernel of the integral operator may be weakly singular. The proof of the convergence of this Multipower Defect Correction method is presented and a numerical example illustrates the behavior of the method.

AMS: 45P05, 33F05, 32A55, 45C05, 34L16, 34L05.

Keywords: Integral operator, weakly singular kernel, projection approximation, defect correction, eigenvalue approximation.

1 Introduction

Let us consider the computation of a cluster of nonzero eigenvalues and associated invariant subspace basis, for an integral operator T defined on a Banach space X. Let $\mathcal{L}(X)$ be the Banach algebra of all bounded linear operators from X into itself, $T \in \mathcal{L}(X)$ and all the norms be denoted by $\|\cdot\|$.

The Multipower Defect Correction (MDC) method, which is a generalization of the double iteration described in [2], starts with an initial approximation obtained, for instance, with a Kantorovitch method, where the operator T is

^{*}This work was supported, in part, by Fundação para a Ciência e a Tecnologia (FCT) through Centro de Matemática da Universidade do Porto (CMUP).

projected onto X_n , a finite dimensional subspace of X. The use of this refinement method requires only the solution of an initial matrix eigenproblem of moderate size (to be solved only once). These initial spectral elements are then refined by a special defect correction formula to yield an approximation to the spectral elements of T. Let $T: X \to X$ be defined by,

$$(Tx)(\tau) = \int_{\Omega} g(|\tau - \tau'|) x(\tau') d\tau', \tau \in \Omega.$$

where the kernel g is possibly weakly singular (in the sense given in [3]) and Ω is a real interval. The spectral problem to be solved is

$$\underline{T}\Phi = \Phi\Theta,\tag{1}$$

where $\Phi \in X^{\mu}$ is the invariant subspace basis and Θ (a complex $\mu \times \mu$ matrix) has as eigenvalues the cluster we seek. X^{μ} is the product space having μ factors equal to X, the norm in this space will still be denoted by $\|\cdot\|$ for the sake of simplicity. <u>T</u> denotes the application of T to each element of an ordered family of μ elements of X. The meaning of $\Phi\Theta$ is $\Phi\Theta$ = $\left| \sum_{j=1}^{\mu} \Theta(j,1)\Phi(j), \dots, \sum_{j=1}^{\mu} \Theta(j,\mu)\Phi(j) \right| \text{ for } \Phi = [\Phi(1), \dots, \Phi(\mu)].$

$$\begin{bmatrix} j=1 & j=1 & j=1 \\ Given a finite dimensional subspace X_n spanned by $e_n = \{(e_{n,j}), j = 1, ..., n\}$ let us denote by $e_n^* = \{(e_{n,j}^*), j = 1, ..., n\}$ the adjoint basis of e_n .
The Kantorovitch method, consists in approximating (1) by$$

$$T_n \Phi_n = \Phi_n \Theta_n \tag{2}$$

 e_n .

where $\underline{T_n}$ denotes the application of T_n to each element of an ordered family of μ elements of X_n , and T_n is such that $T_n x = \pi_n T x$ with the projections π_n defined by $\pi_n x = \sum_{j=1}^n \langle x, e_{n,j}^* \rangle e_{n,j}, \langle \cdot, e_{n,j}^* \rangle$ being the duality product. The sequence of these projections converges to the identity operator on X pointwise, when n goes to infinity.

For this method the convergence of $(T_n)_{n>1}$ to T is uniform which is a special case of the ν -convergence defined in [2]: $||T_n||$ is bounded, $||(T_n - T)T|| \to 0$ and $||(T_n - T)T_n|| \to 0.$

Let Φ be normalized by $\langle \Phi, \Phi_n^* \rangle = I_\mu$ where $\langle \cdot, \cdot \rangle$ is a Gram matrix of the duality products of each function of Φ by each function of Φ_n^* (so Φ may depend on n). The normalizing functions are such that $\underline{T}_n \Phi_n^* = \Phi_n^* \overline{\Theta}_n$.

We make the following assumptions: $\Theta \neq 0$ has eigenvalues isolated from 0, $\|\Phi_n\| = 1, \ \langle \Phi_n \,, \, \Phi_n^* \rangle = I_\mu, \ \langle \Phi \,, \, \Phi_n^* \rangle = I_\mu.$

For sufficiently large n ([2],[6]), there exists μ clustered nonzero eigenvalues of Θ_n and T_n , and the corresponding invariant subspace basis Φ_n , such that

$$\max\{\left|tr(\Theta)/\mu - tr(\Theta_n)/\mu\right|, \left\|\Phi - \Phi_n\right\|\} \le c\left\|(\underline{T}_n - \underline{T})\Phi\right\|.$$
(3)

Proposition 1. For each $k \in \mathbb{N}$ there exists a constant c_k such that, for all n large enough,

$$\max\{|tr(\Theta)/\mu - tr(\Theta_n)/\mu|, \|\Phi - \Phi_n\|\} \le c_k \|(T_n - T)T^k\|.$$
(4)

Proof. It follows from (3), since $\Phi = (\underline{T}^k \Phi)(\Theta^k)^{-1}$ and $\|\Phi\| \leq 2$ ([2], p 98, where the corresponding to our Φ is denoted by $\varphi(n)$).

The matrix eigenproblem $A_n u_n = u_n \Theta_n$, where $A_n(i, j) = \langle Te_{n,j}, e_{n,i}^* \rangle$, for all i = 1, ..., n and j = 1, ..., n, a representation of T_n restricted to X_n , yields the initial spectral elements of $T_n : \Theta_n$ and $\Phi_n = e_n u_n$. Remark that u_n is a $n \times \mu$ matrix and Θ_n is a $n \times n$ complex matrix.

2 Multipower Defect Correction Algorithm

The accuracy of the initial approximation may be poor for moderate n and can be improved refining it by an inner/outer iteration that has p steps of a power iteration inside a defect correction applied to F(x) = 0, where $F: X^{\mu} \to X^{\mu}$ is defined by

$$F(x) = \underline{T}x - x\langle \underline{T}x, \Phi_n^* \rangle.$$
(5)

Let $\Sigma_n : X^{\mu} \to X^{\mu}$ be the block reduced resolvent operator of T_n corresponding to the spectrum of Θ_n , i.e., $\Sigma_n := G_n^{-1}(I - P_n)$, where $G_n x := (I - P_n)\underline{T_n}x - x\Theta_n$, for all $x \in X^{\mu}$, and P_n is the spectral projection of T_n corresponding to the spectrum of Θ_n : $P_n x := \Phi_n \langle x, \Phi_n^* \rangle$. The computation of $x = \Sigma_n y$ for a given $y \in X^{\mu}$, amounts to solve the Sylvester equation

$$T_n(I - P_n)x - x\Theta_n = (I - P_n)y.$$
(6)

The algorithm starts with an approximation Φ_n to the invariant subspace basis of T_n and produces an approximation $\xi^{(k+1)}$ to the invariant subspace basis of the operator T.

3 Convergence

The proof of convergence of this method follows the proof of the convergence of double iteration ([2], p. 145) but it is essentially different, as we can see in the following.

In order to simplify the notation we introduce the following abbreviations: $e^{(k)} := \xi^{(k)} - \Phi$ and $e^{(k,j)} := \varphi^{(k,j)} - \Phi$. As it was already defined in the algorithm, we will have $\Theta^{(k,j)} := \langle \underline{T}\varphi^{(k,j)}, \Phi_n^* \rangle$, for j = 0, ..., p - 1.

Proposition 2. For all n, k and j = 1, ..., p, if $\langle \underline{T}\varphi^{(k,j-1)}, \Phi_n^* \rangle \neq 0$ then $P_n\varphi^{(k,j)} = \Phi_n, P_nF(\varphi^{(k,j)}) = 0$ and $P_ne^{(k,j)} = 0$.

 $\begin{array}{l} \textit{Proof. From the algorithm } \langle \varphi^{(k,j)} \,, \, \Phi_n^* \rangle = I_\mu, \, P_n \varphi^{(k,j)} = \Phi_n \langle \varphi^{(k,j)} \,, \, \Phi_n^* \rangle = \Phi_n. \\ P_n F(\varphi^{(k,j)}) = \Phi_n \langle \underline{T} \varphi^{(k,j)} \,, \, \Phi_n^* \rangle - \Phi_n \langle \varphi^{(k,j)} \,, \, \Phi_n^* \rangle \langle \underline{T} \varphi^{(k,j)} \,, \, \Phi_n^* \rangle = 0. \end{array} \\ \textbf{Finally,} \\ P_n e^{(k,j)} = P_n(\varphi^{(k,j)} - \Phi) = \Phi_n - \Phi_n = 0, \text{ and the proof is complete.} \qquad \Box \end{array}$

Proposition 3. For all k, and for j = 0, ..., p - 1, if $\langle \underline{T} \varphi^{(k,j)}, \Phi_n^* \rangle \neq 0$,

$$e^{(k,j+1)} = \left[\underline{T} e^{(k,j)} - \Phi \left\langle \underline{T} e^{(k,j)}, \Phi_n^* \right\rangle \right] (\Theta^{(k,j)})^{-1},$$
(7)
$$e^{(k,p)} = \left[\underline{T}^p e^{(k,0)} - \Phi \left\langle \underline{T}^p e^{(k,0)}, \Phi_n^* \right\rangle \right] (\Theta^{(k,0)})^{-1} \cdots (\Theta^{(k,p-1)})^{-1}.$$
(8)

Proof.

$$e^{(k,j+1)} = \varphi^{(k,j+1)} - \Phi$$

= $\underline{T}\varphi^{(k,j)}(\Theta^{(k,j)})^{-1} - \Phi + \underline{T}\Phi(\Theta^{(k,j)})^{-1} - \underline{T}\Phi(\Theta^{(k,j)})^{-1}$
= $\left[\underline{T}e^{(k,j)} - \Phi\langle\underline{T}(\varphi^{(k,j)} - \Phi), \Phi_n^*\rangle\right](\Theta^{(k,j)})^{-1}$
= $\left[\underline{T}e^{(k,j)} - \Phi\langle\underline{T}e^{(k,j)}, \Phi_n^*\rangle\right](\Theta^{(k,j)})^{-1}.$

For p = 1 in formula (8), $e^{(k,1)} = \left[\underline{T}e^{(k,0)} - \Phi\langle \underline{T}e^{(k,0)}, \Phi_n^* \rangle\right] (\Theta^{(k,0)})^{-1}$. If the formula (8) is true for j = 0, ..., p - 1, then

$$\begin{split} e^{(k,p)} &= \left[\underline{T} e^{(k,p-1)} - \Phi \langle \underline{T} e^{(k,p-1)} , \Phi_n^* \rangle \right] (\Theta^{(k,p-1)})^{-1} \\ &= \left[\underline{T} \Big(\underline{T}^{p-1} e^{(k,0)} - \Phi \langle \underline{T}^{p-1} e^{(k,0)} , \Phi_n^* \rangle \Big) (\Theta^{(k,0)})^{-1} \cdots (\Theta^{(k,p-2)})^{-1} - \right. \\ &- \Phi \langle \underline{T} \Big(\underline{T}^{p-1} e^{(k,0)} - \Phi \langle \underline{T}^{p-1} e^{(k,0)} , \Phi_n^* \rangle \Big) (\Theta^{(k,0)})^{-1} \cdots (\Theta^{(k,p-2)})^{-1} , \Phi_n^* \rangle \\ &\quad (\Theta^{(k,p-1)})^{-1} \\ &= \left[\underline{T}^p e^{(k,0)} - \underline{T} \Phi \langle \underline{T}^{p-1} e^{(k,0)} , \Phi_n^* \rangle - \right. \\ &- \Phi \langle \Big(\underline{T}^p e^{(k,0)} - \underline{T} \Phi \langle \underline{T}^{p-1} e^{(k,0)} , \Phi_n^* \rangle \Big) , \Phi_n^* \rangle \Big] \\ &\quad (\Theta^{(k,0)})^{-1} \cdots (\Theta^{(k,p-1)})^{-1} \\ &= \left[\underline{T}^p e^{(k,0)} - \Phi \langle \underline{T}^p e^{(k,0)} , \Phi_n^* \rangle \right] (\Theta^{(k,0)})^{-1} \cdots (\Theta^{(k,p-1)})^{-1} . \quad \Box \end{split}$$

Proposition 4. For all n large enough and all k,

$$e^{(k+1)} = \Sigma_n \Big[(\underline{T}_n - \underline{T}) e^{(k,p)} + e^{(k,p)} (\Theta - \Theta_n) + e^{(k,p)} \langle \underline{T} e^{(k,p)} , \Phi_n^* \rangle + (\Phi - \Phi_n) \langle \underline{T} e^{(k,p)} , \Phi_n^* \rangle \Big].$$

Proof. Since $G_n^{-1}(I-P_n) = \Sigma_n$, $(I-P_n)e^{(k,p)} = e^{(k,p)}$, $\Sigma_n \Phi_n = 0$, then $P_n F(\varphi^{(k,p)}) = 0$, $F(\Phi) = 0$, and $(I - P_n)\Phi_n = 0$, and since G_n and P_n commute, we have

$$\begin{aligned} e^{(k+1)} &= \varphi^{(k,p)} - \Phi - \sum_{n} F(\varphi^{(k,p)}) \\ &= G_{n}^{-1} \Big[G_{n} e^{(k,p)} - (I - P_{n}) F(\varphi^{(k,p)}) \Big] \\ &= G_{n}^{-1} (I - P_{n}) \Big[\underline{T_{n}} e^{(k,p)} - e^{(k,p)} \Theta_{n} - F(\varphi^{(k,p)}) + F(\Phi) \Big] \\ &+ (\Phi - \Phi_{n}) \langle \underline{T} e^{(k,p)}, \Phi_{n}^{*} \rangle \Big]. \quad \Box \end{aligned}$$

Proposition 5. For all n large enough and all k, if $\langle \underline{T}\varphi^{(k,j-1)}, \Phi_n^* \rangle \neq 0$ for all j = 1, ..., p,

$$e^{(k+1)} = \sum_{n} \left[(\underline{T_{n}} - \underline{T}) \underline{T}^{p} e^{(k)} + \underline{T}^{p} e^{(k)} (\Theta - \Theta_{n}) + \underline{T}^{p} e^{(k)} \langle \underline{T}^{p} e^{(k,p)} , \Phi_{n}^{*} \rangle + (\Phi - \Phi_{n}) \langle \underline{T}^{p+1} e^{(k)} , \Phi_{n}^{*} \rangle \right] (\Theta^{(k,0)})^{-1} \cdots (\Theta^{(k,p-1)})^{-1} - \sum_{n} \left[\underline{T_{n}} (\Phi - \Phi_{n}) \langle \underline{T}^{p} e^{(k)} , \Phi_{n}^{*} \rangle + (\Phi - \Phi_{n}) \langle \underline{T}^{p} e^{(k)} , \Phi_{n}^{*} \rangle (\Theta - \Theta_{n}) + (\Phi - \Phi_{n}) \langle \underline{T}^{p} e^{(k)} , \Phi_{n}^{*} \rangle \langle \underline{T}^{p} e^{(k,p)} , \Phi_{n}^{*} \rangle \right] (\Theta^{(k,0)})^{-1} \cdots (\Theta^{(k,p-1)})^{-1}.$$

Proof. It follows immediately from Proposition 3 and Proposition 4. \Box

Proposition 6. For all n large enough, all k and j = 0, ..., p, $\|\Theta^{(k,j)}\| \ge \frac{\|\Theta\|}{2}$.

Proof. Suppose that for a given p (number of power iterations) and all j = 1, ..., p, $\|\Theta^{(k,j-1)}\| \geq \frac{\|\Theta\|}{2}$. Let c_* be such that, for n large enough, $1 \leq \|\Phi_n^*\| \leq c_*$. By Proposition 3, $e^{(k,j)} = \left[\underline{T}e^{(k,j-1)} - \Phi\left\langle \underline{T}e^{(k,j-1)}, \Phi_n^* \right\rangle\right] (\Theta^{(k,j-1)})^{-1}$, so that $\|e^{(k,j)}\| \leq \beta \|e^{(k,j-1)}\|$, where $\beta := \frac{2\|T\|}{\|\Theta\|} \left[1 + 2c_*\right] \geq 6$, since $\|\Phi\| \leq 2$ for n large enough. It follows that for all j = 1, ..., p

$$\begin{aligned} \|\Theta^{(k,j)} - \Theta\| &= |\langle T(\varphi^{(k,j)} - \Phi), \Phi_n^* \rangle| \le c_* \|T\| \|e^{(k,j)}\| \\ &\le c_* \beta \|T\| \|e^{(k,j-1)}\| \le c_* \beta^p \|T\| \|e^{(k)}\|. \end{aligned}$$

Proposition 5 shows that, for *n* sufficiently large, $||e^{(k)}|| \leq ||e^{(0)}||$, and since there is a constant c_p such that, for *n* large enough, $||e^{(0)}|| \leq c_p ||(T_n - T)T^p||$, we conclude that $||\Theta^{(k,j)} - \Theta||$ tends to 0 as *n* tends to infinity uniformly in (k, j). Hence, for all *n* large enough, $||\Theta^{(k,j)} - \Theta|| \leq \frac{||\Theta||}{2}$, and $||\Theta^{(k,j)}|| \geq \frac{||\Theta||}{2}$, independently of *k*, and for all j = 1, ..., p.

Theorem 7. There exist constants γ_0 and γ such that, for n large enough, and all $k \ge 0$,

$$\|e^{(k)}\| \le \gamma_0 \|(T_n - T)T^p\|(\gamma\|(T_n - T)T^p\|)^k.$$
(9)

Proof. Propositions 1 and 5 show that there exists γ such that, for n sufficiently large, $\|e^{(k+1)}\| \leq \gamma \|(T_n - T)T^p\| \|e^{(k)}\|$. But there exists γ_0 such that, for n sufficiently large, $\|e^{(0)}\| \leq \gamma_0 \|(T_n - T)T^p\|$, and the bound (9) follows.

The bound (9) improves over the one for the double iteration when ||T|| is less than one as it is the case of the following example that comes from a real life model with a very interesting weakly singular kernel, the first exponential integral function (see [1]). Although it is a simplified model, it is of real use ([8], [7]). The number of iterations and the dimensions of the discretizations needed are large, however, the total CPU time is quite reasonable.

4 Numerical illustration

To illustrate the performance of the method the following integral operator,

$$(T\varphi)(\tau) = \frac{\varpi}{2} \int_0^{\tau^*} E_1\left(|\tau - \tau'|\right)\varphi\left(\tau'\right) d\tau'$$

issued from a simplified model of the radiative transfer in stellar atmospheres, defined on $X = L^1([0, 1])$ was used, where τ^* is the optical thickness of a stellar atmosphere, $\varpi \in [0, 1]$ is the albedo (assumed to be constant in the present work) (see [3]).

We considered $\tau^* = 4000$, $\varpi = 0.75$, and n = 800. For details on how to derive the corresponding matrix form see [4] or [5]. As the operator T may not be available in closed form, for computational purpose, we used a finer matrix representation of it of dimension m = 4000.

A moderate size eigenproblem, 800×800 , is solved to obtain the initial approximation and another one is solved to obtain the vector used in the normalizations, before starting the iteration MDC.

As the eigenvalues of this operator are simple the computation of $\Sigma_n y$ is here reduced to the solution of a $(n + 1) \times n$ linear system of equations (the extra row in the system corresponds to the restriction of the solution to be in the $(I - P_n)X$ subspace) with the same coefficient matrix and different right hand sides. This solution was done by LU decomposition with partial pivoting once for all.

The computations were done on a personal computer using MATLAB.

Table 1 shows the performance of the MDC method. Increasing values of p imposes a reduction both on the number of outer iterations to convergence and on the CPU time, showing the effectiveness of using cheap inner iterations (although they involve products by the m dimensional matrix A_m representing T in the subspace X_m) to reduce the number of expensive outer iterations (they involve the solution of a moderate size linear system of equations and one prolongation step). Furthermore, the double iteration method, p = 1, could not achieve convergence for the fifth eigenpair. The role of the inner iteration to accelerate the computations is very important as we can see from Table 1.

		(-		~ ~
	p = 1 (double it.)		p = 5		p = 20	
eig	n. it.	cpu time	outer iter.	total cpu time	outer iter.	total cpu time
1	349	346.2	138	146.9	72	92.3
2	411	406.6	161	170.2	84	106.9
3	495	489.0	192	201.4	99	125.1
4	630	620.5	241	251.8	122	152.5
5	*	*	265	278.2	125	156.3

Table 1: Number of outer iterations and total CPU time (sec) for several eigenvalues (and corresponding eigenvectors), for a tolerance less than 1e-12.

For the five largest eigenvalues in magnitude, the initial approximation, the refined one computed with the MDC method for a tolerance of 1e - 12, and the respective eigenvalues of the A_m matrix, are presented in Table 2. We can check that the refined values are as good as if we had proceeded with the Kantorovitch method to the dimension m = 4000.

Table 2: Initial eigenvalues approximation to be refined, corrected approximate eigenvalues for tol = 1e - 12.

eig	initial eigenvalue approximation	refined eigenvalue $(tol = 1e - 12)$	eigenvalue of A_{4000}
1	0.7499992074916	0.7499997185840	0.74999971858403
2	0.7499970999445	0.7499990781097	0.74999907810973
3	0.7499942701664	0.7499978858383	0.74999788583835
4	0.7499894135218	0.7499959814241	0.74999598142415
5	0.7499814539823	0.7499937902837	0.74999379028378

In Table 3 we show number of iterations and CPU time (sec) for the approximation of the largest eigenvalues and associated eigenvectors for m = 4000 and p = 20 for coarse and fine tolerances. The method shows good behavior for both tolerances with a slight degradation in performance when looking for eigenvalues of smaller absolute value, as expected.

The MDC method is very performant even for coarse tolerances on the norm of the residual, especially when the number of eigenvalues to compute is not large. Moreover it has the advantage of being simple to understand and implement.

Acknowledgments. The authors are deeply indebted to Mario Ahues for his help with part of the proof and thank also Alain Largillier for helpful suggestions.

References

 M. ABRAMOWITZ AND I. A. STEGUN, Handbook of Mathematical Functions, Dover, New York, 1960.

	tol =	= 1e - 6	tol = 1e - 12		
eig	n. it.	cpu time	n. it.	cpu time	
1	13	21.4	72	92.3	
2	11	19.1	84	106.9	
3	18	27.4	99	125.1	
4	26	37.0	122	152.5	
5	25	35.8	125	156.3	

Table 3: Number of iterations and CPU time (sec) for several eigenvalues and p = 20.

- [2] M. AHUES, A. LARGILLIER, B. V. LIMAYE, Spectral Computations with Bounded Operators, CRC, Boca Raton, 2001.
- [3] M. AHUES, F. D. D'ALMEIDA, A. LARGILLIER, O. TITAUD, P. VASCON-CELOS, An L¹ Refined Projection Approximate Solution of the Radiation Transfer Equation in Stellar Atmospheres, J. Comput. Appl. Math., 140 (2002), pp. 13–26.
- [4] M. AHUES, F. D. D' ALMEIDA, A. LARGILLIER, P. B. VASCONCELOS, Defect correction for spectral computations for a singular integer operator, Comm. Pure Appl. Math., 5-2 (2006), pp. 241–250.
- [5] F. D. D' ALMEIDA, P. B. VASCONCELOS, R. FERNANDES, Discretization of Iterative Refinement Methods for a Transfer Integral Equation, J. Anal., 14 (2006), pp. 33–53.
- [6] F. CHATELIN, Spectral Approximations of Linear Operators, Academic Press, New York, 1983.
- [7] L. CHEVALLIER, F. PALETOU, B. RUTILY, On the accuracy of the ALI method for solving the radiative transfer equation, Astronomy & Astrophysics, 411 (2003), pp. 221-227.
- [8] B. RUTILY, Multiple Scattering Theoretical and Integral and Numerical Technique, C. Constanda and M. Ahues and A. Largillier, eds., Birkhauser, 2004, pp. 211–231.
- [9] H. STETTER, The defect correction principle and discretization methods, Numer. Math., 29 (1978), pp. 425–443.