ARTICLE IN PRESS

APNUM:3066

Applied Numerical Mathematics ••• (••••) •••–•••

Contents lists available at ScienceDirect

Applied Numerical Mathematics

www.elsevier.com/locate/apnum

Efficient cyclic reduction for Quasi-Birth–Death problems with rank structured blocks [☆]

Dario A. Bini [a], Stefano Massei [b,*], Leonardo Robol [c]

[a] *Dipartimento di Matematica, University of Pisa, Italy*
[b] *Scuola Normale Superiore, Pisa, Italy*
[c] *Department of Computer Science, KU Leuven, Belgium*

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Cyclic reduction
Matrix equations
Double QBD processes
Markov chains
H matrices
Quasiseparable matrices

ABSTRACT

We provide effective algorithms for solving block tridiagonal block Toeplitz systems with $m \times m$ quasiseparable blocks, as well as quadratic matrix equations with $m \times m$ quasiseparable coefficients, based on cyclic reduction and on the technology of rank-structured matrices. The algorithms rely on the exponential decay of the singular values of the off-diagonal submatrices generated by cyclic reduction. We provide a formal proof of this decay in the Markovian framework. The results of the numerical experiments that we report confirm a significant speed up over the general algorithms, already starting with the moderately small size $m \approx 10^2$.

1. Introduction

Cyclic reduction (CR) is an effective tool that can be used for solving several problems in linear algebra and in polynomial computations [4]. It has been originally introduced by G.H. Golub and R.W. Hockney in the mid 1960s [14,8], for the numerical solution of block tridiagonal linear systems stemming from the finite differences solution of elliptic problems, and has been generalized to solve nonlinear matrix equations associated with matrix power series with applications to queuing problems, Markov chains and spectral decomposition of polynomials. In fact, an important application of CR concerns the computation of the minimal nonnegative solution of the matrix equation $X = A_{-1} + A_0X + A_1X^2$, encountered in Quasi Birth–Death (QBD) Markov chains, where A_{-1}, A_0, and A_1 are given $m \times m$ nonnegative matrices such that $A_{-1} + A_0 + A_1$ is irreducible and stochastic and where X is the matrix unknown [2,4]. The computation of the solution X allows to recover the steady state vector π of the Markov chain.

Rewriting the matrix equation as $A_{-1} + (A_0 - I)X + A_1X^2 = 0$, CR computes four sequences of matrices, $A_i^{(h)}$, $i = -1, 0, 1$ and $\widehat{A}_0^{(h)}$, according to the following equations

$$
\begin{aligned}
&A_1^{(h+1)} = -A_1^{(h)}\, S^{(h)}\, A_1^{(h)}, \qquad S^{(h)} = (A_0^{(h)} - I)^{-1}\\
&A_0^{(h+1)} = A_0^{(h)} - A_1^{(h)}\, S^{(h)}\, A_{-1}^{(h)} \; - \; A_{-1}^{(h)}\, S^{(h)}\, A_1^{(h)},\\
&A_{-1}^{(h+1)} = -A_{-1}^{(h)}\, S^{(h)}\, A_{-1}^{(h)}, \quad \widehat{A}_0^{(h+1)} = \widehat{A}_0^{(h)} - A_1^{(h)}\, S^{(h)}\, A_{-1}^{(h)},
\end{aligned}
\tag{1}
$$

☆ Work supported by Gruppo Nazionale di Calcolo Scientifico (GNCS) of INdAM, and by the PRA project "Mathematical models and computational methods for complex networks" funded by the University of Pisa.

* Corresponding author.
E-mail addresses: bini@dm.unipi.it (D.A. Bini), stefano.massei@sns.it (S. Massei), leonardo.robol@sns.it (L. Robol).

http://dx.doi.org/10.1016/j.apnum.2016.06.014
0168-9274/

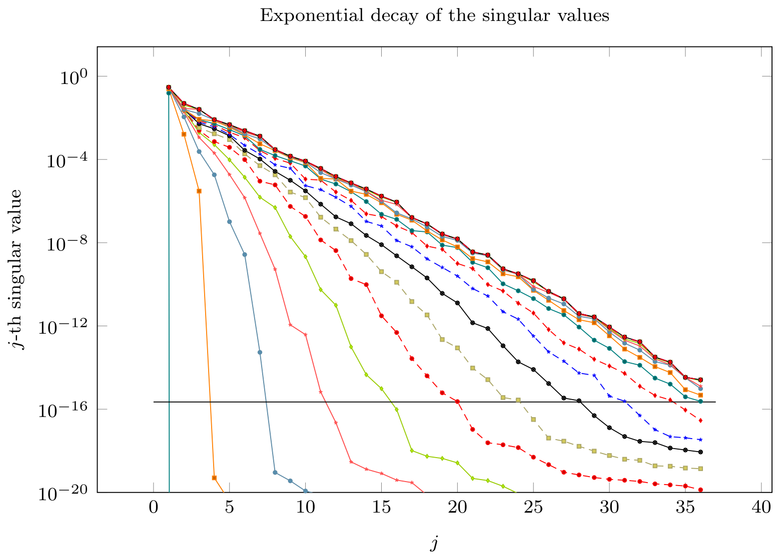

Fig. 1. Log-scale plot of the largest singular values of the largest south-western submatrix of A_0 contained in the lower triangular part, for $m = 1600$. The horizontal line denotes the machine precision threshold. Matrices are randomly generated so that $A_i \geqslant 0$ are tridiagonal matrices and $A_{-1} + A_0 + A_1$ is stochastic. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

for $h = 0, 1, \ldots$, with $A_i^{(0)} = A_i$, $i = -1, 0, 1$ and $\widehat{A}_0^{(0)} = A_0 - I$. It can be proved, in the context of Markov chains, [2] that the sequence $-(\widehat{A}_0^{(h)})^{-1} A_{-1}$ converges to the minimal nonnegative solution G of the matrix equation.

If the QBD process is not null recurrent, applicability and quadratic convergence of the algorithm are guaranteed [2, Theorems 7.5, 7.6]. In the null recurrent case, it has been proved in [12] that convergence is linear with factor $\frac{1}{2}$.

Without any further assumption on the structure of the blocks, each step of CR requires a small number of matrix multiplications and one matrix inversion for the resulting computational cost of $O(m^3)$ arithmetic operations (ops) per step. On the other hand, there are several models from the applications in which the blocks A_i exhibit special structures. In order to decrease the computational complexity of the iterations, variations of CR which exploit these structures have been proposed, see for instance [3,18,1].

Here, we are interested in analyzing the case where the blocks A_i are quasiseparable matrices. That is, the case where the *off-diagonal submatrices* of A_{-1}, A_0 and A_1, strictly contained in the upper or in the lower triangular part, have low rank with respect to m. The maximum of the ranks of the off-diagonal submatrices is called *quasiseparable rank* and a matrix with quasiseparable rank k is called k-quasiseparable. Observe that k-quasiseparable matrices include banded matrices. These structures are encountered in wide and important classes of applications like, bidimensional random walks [17], the Jackson tandem queue model [15] and other QBD processes, or, for instance, in the finite differences discretization of elliptic PDEs.

Our goal is to design a version of CR which exploits the rank structures of the blocks A_i and which can be implemented at a substantially lower cost. This way, we may arrive at designing effective solvers both for block tridiagonal block Toeplitz systems and for the quadratic matrix equations encountered in QBD Markov chains. Indeed, a way to reach this goal is to find out if some structure of the blocks $A_i^{(h)}$ is maintained during the CR steps (1), and then to try to exploit this structure in order to design an effective implementation of CR.

Looking at the iterative scheme (1), one can find out that the quasiseparable rank can grow exponentially at each step. Despite that, plotting the singular values of the off-diagonal blocks of the matrices $A_i^{(h)}$ shows an interesting behavior as reported in Fig. 1. For a randomly generated QBD process with tridiagonal blocks of size 1600, the singular values of an off-diagonal submatrix of size 799×800 in $A_0^{(h)}$ have an exponential decay in the first 20 steps of CR.

It is evident that, even though the number of nonzero singular values grows at each step of CR, the number of singular values above the machine precision – denoted by a horizontal line in Fig. 1 – is bounded by a moderate constant. Moreover, the singular values seem to stay below a straight-line which constitutes an asymptotic bound. That is, they get closer to this line as $h \to \infty$. The logarithm scale suggest that the computed singular values $\sigma_i^{(h)}$ decay exponentially with i and the basis of the exponential grows with h but has a limit less than 1.

We will prove this asymptotic property relying on the technology of rank-structured matrices and relate the basis of the exponential decay to the width of the domain of analyticity of the matrix function $\varphi(z)^{-1}$ for $\varphi(z) = -z^{-1}A_{-1} + I - A_0 - zA_1$.

It is interesting to observe that the three matrix sequences $A_i^{(h)}$, $i = -1, 0, 1$ are related to certain Schur complements of the block tridiagonal block Toeplitz matrix $\mathrm{trid}(A_{-1}, A_0, A_1)$, obtained after permuting block rows and columns with the odd-even permutation [4]. A slightly different sequence is given by the Schur complements obtained by applying Gaussian